

Jiadong Sun · Changfeng Yang
Shuren Guo
Editors

China Satellite Navigation Conference (CSNC) 2018 Proceedings

Volume I



Lecture Notes in Electrical Engineering

Volume 497

Board of Series editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Bijaya Ketan Panigrahi, New Delhi, India
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Shanben Chen, Shanghai, China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Haibin Duan, Beijing, China
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Limin Jia, Beijing, China
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Qilian Liang, Arlington, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Mukhopadhyay, Palmerston North, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Kyoto, Japan
Federica Pascucci, Roma, Italy
Yong Qin, Beijing, China
Gan Woon Seng, Singapore, Singapore
Germano Veiga, Porto, Portugal
Haitao Wu, Beijing, China
Junjie James Zhang, Charlotte, USA

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Springerlink ****

Lecture Notes in Electrical Engineering (LNEE) is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering
- Engineering

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNEE will be distributed through Springer's print and electronic publishing channels.

For general information about this series, comments or suggestions, please use the contact address under "service for this series".

To submit a proposal or request further information, please contact the appropriate Springer Publishing Editors:

Asia:

China, *Jessie Guo, Assistant Editor* (jessie.guo@springer.com) (Engineering)

India, *Swati Meherishi, Senior Editor* (swati.meherishi@springer.com) (Engineering)

Japan, *Takeyuki Yonezawa, Editorial Director* (takeyuki.yonezawa@springer.com)
(Physical Sciences & Engineering)

South Korea, *Smith (Ahram) Chae, Associate Editor* (smith.chae@springer.com)
(Physical Sciences & Engineering)

Southeast Asia, *Ramesh Premnath, Editor* (ramesh.premnath@springer.com)
(Electrical Engineering)

South Asia, *Aninda Bose, Editor* (aninda.bose@springer.com) (Electrical Engineering)

Europe:

Leontina Di Cecco, Editor (Leontina.dicecco@springer.com)

(Applied Sciences and Engineering; Bio-Inspired Robotics, Medical Robotics, Bioengineering; Computational Methods & Models in Science, Medicine and Technology; Soft Computing; Philosophy of Modern Science and Technologies; Mechanical Engineering; Ocean and Naval Engineering; Water Management & Technology)
(christoph.baumann@springer.com)

(Heat and Mass Transfer, Signal Processing and Telecommunications, and Solid and Fluid Mechanics, and Engineering Materials)

North America:

Michael Luby, Editor (michael.luby@springer.com) (Mechanics; Materials)

More information about this series at <http://www.springer.com/series/7818>

Jiadong Sun · Changfeng Yang
Shuren Guo
Editors

China Satellite Navigation Conference (CSNC) 2018 Proceedings

Volume I

 Springer

Editors

Jiadong Sun
China Aerospace Science
and Technology Corporation
Beijing
China

Shuren Guo
China Satellite Navigation
Engineering Center
Beijing
China

Changfeng Yang
China Satellite Navigation
Engineering Center
Beijing
China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-0004-2 ISBN 978-981-13-0005-9 (eBook)
<https://doi.org/10.1007/978-981-13-0005-9>

Library of Congress Control Number: 2018938375

© Springer Nature Singapore Pte Ltd. 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. part of Springer Nature
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Editorial Board

Topic: S1: Satellite Navigation Technology

Chairman

Qin Zhang, Chang'an University, Shaanxi, China

Vice-Chairman

Feixue Wang, National University of Defense Technology, Changsha, China

Shuanggen Jin, Shanghai Astronomical Observatory Chinese Academy of Sciences, Shanghai, China

Xiaolin Meng, University of Nottingham, UK

Topic: S2: Navigation and Location Service

Chairman

Yamin Dang, Chinese Academy of Surveying & Mapping, Beijing, China

Vice-Chairman

Baoguo Yu, The 54th Research Institute of China Electronics Technology Group Corporation, Hebei, China

Qun Ding, The 20th Research Institute of China Electronics Technology Group Corporation, Shaanxi, China

Kefei Zhang, RMIT University, Melbourne, Australia

Topic: S3: Satellite Navigation Signal and Anti-Jamming Technologies

Chairman

Xiaochun Lu, National Time Service Center, Chinese Academy of Sciences, Shaanxi, China

Vice-Chairman

Yanhong Kou, Beihang University, Beijing, China

Zheng Yao, Tsinghua University, Beijing, China

Topic: S4: Satellite Orbit and Satellite Clock Error**Chairman**

Xiaogong Hu, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China

Vice-Chairman

Rongzhi Zhang, Xi'an Satellite Control Center, Shaanxi, China

Geshi Tang, Beijing Aerospace Control Center, Beijing, China

Maorong Ge, Geo Forschungszentrum (GFZ), Potsdam, Germany

Topic: S5: Precise Positioning Technology**Chairman**

Qile Zhao, Wuhan University, Hubei, China

Vice-Chairman

Jianwen Li, The PLA Information Engineering University, Henan, China

Shuli Song, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China

Yanming Feng, Queensland University of Technology, Brisbane, Australia

Topic: S6: Time–Space Datum and Temporal Frequency Technologies**Chairman**

Lianshan Gao, The 203th Research Institute of China Aerospace Science & Industry Corporation, Beijing, China

Vice-Chairman

Chunhao Han, Beijing Satellite Navigation Center, Beijing, China

Xiaohui Li, National Time Service Center, Chinese Academy of Sciences, Shaanxi, China

Nikolay Demidov, Vremya CH Corporation, Novgorod, Russia

Topic: S7: Satellite Navigation Augmentation Technology**Chairman**

Junlin Zhang, OLinkStar Co., Ltd., Beijing, China

Vice-Chairman

Jinping Chen, Beijing Satellite Navigation Center, Beijing, China

Rui Li, Beihang University, Beijing, China

Shaojun Feng, Imperial College London Qianxun Positioning Network, Co., Ltd., Shanghai, China

Dun Wang, Space Star Aerospace Technology Applications Co., Ltd., Heilongjiang, China

Topic: S8: Test and Assessment Technology

Chairman

Xiaolin Jia, Xi'an Institute of Surveying and Mapping, Shaanxi, China

Vice-Chairman

Jun Yang, National University of Defense Technology, Hunan, China

Jianguang Qu, Heilongjiang Institute of Technology, Heilongjiang, China

Yang Gao, University of Calgary, Alberta, Canada

Topic: S9: User Terminal Technology

Chairman

Haibo He, Beijing Satellite Navigation Center, Beijing, China

Vice-Chairman

Baowang Lian, Northwestern Polytechnic University, Shaanxi, China

Hong Li, Tsinghua University, Beijing, China

Topic: S10: Multi-Source Fusion Navigation Technology

Chairman

Zhongliang Deng, Beijing University of Posts and Telecommunications, Beijing, China

Vice-Chairman

Hong Yuan, Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China

Yongbin Zhou, National University of Defense Technology

Jinling Wang, University of New South Wales, Australia

Topic: S11: PNT New Concept, New Methods and New Technology**Chairman**

Mingquan Lu, Tsinghua University, Beijing, China

Vice-Chairman

Wei Wang, The 20th Research Institute of China Electronics Technology Group Corporation, Shaanxi, China

Ying Xu, Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China

Sang Jeong Lee, Chungnam National University, Daejeon, South Korea

Topic: S12: Policies and Regulations, Standards and Intellectual Properties**Chairman**

Junlin Yang, Beihang University, Beijing, China

Vice-Chairman

Daiping Zhang, China Defense Science and Technology Information Center, Beijing, China

Yonggang Wei, China Academy of Aerospace Standardization and Product Assurance, Beijing, China

Huiying Li, Electronic Intellectual Property Center, Ministry of Industry and Information Technology, PRC Beijing, China

Preface

BeiDou Navigation Satellite System (BDS) is China's global navigation satellite system which has been developed independently. BDS is similar in principle to global positioning system (GPS) and compatible with other global satellite navigation systems (GNSSs) worldwide. BDS will provide highly reliable and precise positioning, navigation and timing (PNT) services and short-message communication for all users under all-weather, all-time and worldwide conditions.

China Satellite Navigation Conference (CSNC) is an open platform for academic exchanges in the field of satellite navigation. It aims to encourage technological innovation, accelerate GNSS engineering, and boost the development of the satellite navigation industry in China and in the world.

The 9th China Satellite Navigation Conference (CSNC 2018) is held during May 23–25, 2018, Harbin, China. The theme of CSNC2018 is Location, Time of Augmentation, including technical seminars, academic exchanges, forums, exhibitions, and lectures. The main topics are as followed:

Conference Topics

- S1 Satellite Navigation Technology
- S2 Navigation and Location Service
- S3 Satellite Navigation Signal and Anti-Jamming Technologies
- S4 Satellite Orbit and Satellite Clock Error
- S5 Precise Positioning Technology
- S6 Time–Space Datum and Temporal Frequency Technologies
- S7 Satellite Navigation Augmentation Technology
- S8 Test and Assessment Technology
- S9 User Terminal Technology
- S10 Multi-Source Fusion Navigation Technology
- S11 PNT New Concept, New Methods and New Technology
- S12 Policies and Regulations, Standards and Intellectual Properties

The proceedings have 208 papers in twelve topics of the conference, which were selected through a strict peer-review process from 588 papers presented at CSNC2018. In addition, another 274 papers were selected as the electronic proceedings of CSNC2018, which are also indexed by “China Proceedings of Conferences Full-text Database (CPCD)” of CNKI and Wan Fang Data.

We thank the contribution of each author and extend our gratitude to 279 referees and 55 session chairmen who are listed as members of editorial board. The assistance of CSNC2018’s organizing committees and the Springer editorial office is highly appreciated.

Beijing, China

Jiadong Sun
Changfeng Yang
Shuren Guo

Scientific Committee

Chairman

Jiadong Sun, China Aerospace Science and Technology Corporation, Beijing, China

Vice-Chairman

Rongjun Shen, China Satellite Navigation System Committee, Beijing, China
Jisheng Li, China Satellite Navigation System Committee, Beijing, China
Qisheng Sui, China Satellite Navigation System Committee, Beijing, China
Changfeng Yang, China Satellite Navigation System Committee, Beijing, China
Zuhong Li, China Academy of Space Technology, Beijing, China
Shusen Tan, Beijing Satellite Navigation Center, Beijing, China

Executive Chairman

Jingnan Liu, Wuhan University, Hubei, China
Yuanxi Yang, China National Administration of GNSS and Applications, Beijing, China
Shiwei Fan, China Satellite Navigation Engineering Center, Beijing, China
Jun Xie, China Academy of Space Technology, Beijing, China

Committee Members: (By Surnames Stroke Order)

Xiancheng Ding, China Electronics Technology Group Corporation, Beijing, China
Qingjun Bu, China National Administration of GNSS and Applications, Beijing, China
Weixing Wan, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China
Liheng Wang, China Aerospace Science and Technology Corporation, Beijing, China
Yuzhu Wang, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, China

Guoxiang Ai, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China

Shuhua Ye, Shanghai Astronomical Observatories, Chinese Academy of Sciences, Shanghai, China

Daren Lv, The Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Yongcai Liu, China Aerospace Science & Industry Corporation, Beijing, China

Zhaowen Zhuang, National University of Defense Technology, Hunan, China

Qifeng Xu, PLA Information Engineering University, Henan, China

Houze Xu, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Hubei, China

Tianchu Li, National Institute of Metrology, Beijing, China

Jiancheng Li, Wuhan University, Hubei, China

Guirong Min, China Academy of Space Technology, Beijing, China

Xixiang Zhang, The 29th Research Institute of China Electronics Technology Group Corporation, Sichuan, China

Lvqian Zhang, China Aerospace Science and Technology Corporation, Beijing, China

Junyong Chen, National Administration of Surveying, Mapping and Geo-information, Beijing, China

Benyao Fan, China Academy of Space Technology, Beijing, China

Dongjin Luo, China People's Liberation Army, Beijing, China

Guohong Xia, China Aerospace Science & Industry Corporation, Beijing, China

Shuren Guo, China Satellite Navigation Engineering Center, Beijing, China

Peikang Huang, China Aerospace Science & Industry Corporation, Beijing, China

Chong Cao, China Research Institute of Radio Wave Propagation (CETC 22), Beijing, China

Faren Qi, China Academy of Space Technology, Beijing, China

Rongsheng Su, China People's Liberation Army, Beijing, China

Ziqing Wei, Xi'an Institute of Surveying and Mapping, Shaanxi, China

Executive Members: (By Surnames Stroke Order)

Qin Zhang, Chang'an University, Shaanxi, China

Yamin Dang, Chinese Academy of Surveying & Mapping, Beijing, China

Xiaochun Lu, National Time Service Center, Chinese Academy of Sciences, Shaanxi, China

Xiaogong Hu, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China

Qile Zhao, Wuhan University, Hubei, China

Lianshan Gao, The 203th Research Institute of China Aerospace Science & Industry Corporation, Beijing, China

Junlin Zhang, OLinkStar Co., Ltd., Beijing, China

Xiaolin Jia, Xi'an Institute of Surveying and Mapping, Shaanxi, China

Haibo He, Beijing Satellite Navigation Center, Beijing, China

Zhongliang Deng, Beijing University of Posts and Telecommunications, Beijing, China

Mingquan Lu, Tsinghua University, Beijing, China

Junlin Yang, Beihang University, Beijing, China

Organizing Committee

Secretary-General

Haitao Wu, Satellite Navigation Headquarters, Chinese Academy of Sciences, Beijing, China

Executive Deputy Secretary-General

Weina Hao, Navigation Headquarters, Chinese Academy of Sciences, Beijing, China

Deputy Secretary-General

Wenhai Jiao, China Satellite Navigation Engineering Center, Beijing, China

Wenjun Zhao, Beijing Satellite Navigation Center, Beijing, China

Tonghuang Wang, Space Star Aerospace Technology Applications Co., Ltd., Heilongjiang, China

Bo Wang, Academic Exchange Center, China Satellite Navigation Office, Beijing, China

Committee Members: (By Surnames Stroke Order)

Qun Ding, The 20th Research Institute of China Electronics Technology Group Corporation, Shaanxi, China

Guang Yu, Harbin Institute of Technology, Heilongjiang, China

Li Wang, International Cooperation Research Center, China Satellite Navigation Office, Beijing, China

Chunsheng Wang, Haidian Investment Promotion Bureau, Beijing, China

Ying Liu, China Satellite Navigation Engineering Center, Beijing, China

Wanming Yang, Administrative Committee of Nanjing New & High Technology Industry Development Zone, Jiangsu, China

Jun Shen, Beijing UniStrong Science & Technology Co., Ltd., Beijing, China

Liquan Song, Heilongjiang Institute of Technology, Heilongjiang, China

Mingquan Lu, Tsinghua University, Beijing, China

Lu Chen, Beijing Institute of Space Science and Technology Information,
Beijing, China

Xiuwan Chen, Peking University, Beijing, China

Gang Hu, Beijing Unicore Communications, Inc., Beijing, China

Qile Zhao, Wuhan University, Hubei, China

Min Shui, The National Remote Sensing Center of China, Beijing, China

Contents

Part I Satellite Navigation Technology

A Troposphere Tomography Method by Combining the Truncation Coefficient and Variance Component Analysis	3
Qingzhi Zhao, Yibin Yao and Linyang Xin	
The Design and Implementation of a SDS-TWR Based Wireless Location System	17
Yao Wang, He Yang and Moyu Sha	
Research on Smoothing Filtering Algorithm of BDS/GPS Slow Deformation Monitoring Sequence	33
Lei Ren and Tian-He Xu	
Research on Space-Based Measurement and Control Scheme of Launch Vehicle Based on BeiDou Navigation Satellite System	45
Chong Li, Xianqing Yi, Yue Zhao and Zhenwei Hou	
Online GPR-KF for GNSS Navigation with Unmodelled Measurement Error	59
Panpan Huang, Chris Rizos and Craig Roberts	
The Application of a New Multi-gross Errors Detection Method with the Epoch Variation Constraint in UAV	71
Hongyu Shan, Chunhua Li, Xiaoyu Shi and Chenggang Li	
Research on Multi-satellites Fusion Inversion Model of Soil Moisture Based on Sliding Window	85
Yueji Liang, Chao Ren and Yibang Huang	
A Study of TEC Storm on 13 October 2016.	97
Qingtao Wan, Jinghua Li, Xiaolan Wang, Jiangtao Fan, Jie Zhang and Guanyi Ma	

An Alternative Approach for Estimating SNR Metrics in GPS-IR 105
 Jizhong Wu and Wei Wu

Research on Refined 3D Attitude Model of Smart Construction Machinery Based on Multi-sensor Fusion 117
 Yanning Zheng, Shengli Wang, Yang Liu, Ying Xu, Xu Li and Guiping Chen

Feasibility Study of Low Cost Receiver for Deformation Monitoring . . . 129
 Hongke Hou, Cuilin Kuang, Yaozong Zhou and Yonglin Zhang

Algorithm and Performance of Precipitable Water Vapor Retrieval Using Multiple GNSS Precise Point Positioning Technology 139
 Peng Hu, Guanwen Huang, Qin Zhang, Xiaolei Wang and Min Mao

Retrieval of PWV Based on GPS and Multi-level Isobaric Surface Data 153
 Hongkai Shi, Xiufeng He and Xinyuan Wang

Multipath Mitigation Method in GPS/BDS Deformation Monitoring Based on Observation Domain 163
 Qinggen Yi, Guoli Lin, Wanke Liu and Jiuhui Pan

Analysis of Sea Surface Multipath and Impact on GNSS Precision Positioning 173
 Zhengxun He, Shuangcheng Zhang, Yang Nan, Kai Liu, Wei Qu, Huilin Wu and Qi Liu

HECTOR for Analysis of GPS Time Series 187
 Yuefan He, Shuangcheng Zhang, Qianyi Wang, Qi Liu, Wei Qu and Xiaowei Hou

Inversion of Water Vapor Variation During Typhoon by Quad-Constellation GNSS Tomography 197
 Yonglin Zhang, Changsheng Cai, Xiaotao Bai, Hongke Hou and Xianqiang Cui

GPS-MR for Altai Snow Depth Monitoring 211
 Chenglong Zhang, Shuangcheng Zhang, Tao Che, Yajie Wang, Ning Zhang, Wei Qu and Tianhe Wan

Snow Depth Detection Based on L2 SNR of GLONASS Satellites and Multipath Reflectometry 221
 Wei Zhou, Lilong Liu, Liangke Huang, Jun Chen and Songqing Li

Analysis of Ionospheric Anomalous Disturbance During a Heavy Rainfall 233
 Lulu Shan and Yibin Yao

Water Level Changes Measurement of Reservoir Using Multi-GNSS Reflectometry 243
 Kai Liu, Cuilin Kuang, Wei Zhang, YaXin Wen and Chunlong Gong

The Impact of Atmospheric Mass Loading on Precise Positioning of National GNSS Stations 255
 Yangyang Liu, Yamin Dang and Changhui Xu

Part II Navigation and Location Service

A Unidirectional Communication Architecture for Extended Location-Based Services 267
 Yanming Feng and Charles Wang

Virtual Reference Station (VRS) Coordinate’s Pattern of QianXun Ground-Based Augmentation System 285
 Xinyu Chen, Chenhao Ouyang and Junbo Shi

A Novel Fingerprinting Method of WiFi Indoor Positioning Based on Weibull Signal Model 297
 Zheng Li, Jingbin Liu, Zemin Wang and Ruizhi Chen

Development and Assessment of GNSS Online Differential Data Processing System Based on HNCORS 311
 Yaozong Zhou, Cuilin Kuang, Shaohua Dou and Ziping Liu

A Robust Turn Detection Algorithm Based on Periodic Signal Identification 325
 Yu Chen, Haiyong Luo, Fang Zhao, Wenhua Shao and Qu Wang

The Indoor Localization Algorithm for Combination of Signal Strength and Anti-disturbance 341
 Wanqing Liu, Juqing Zhang, Guanwen Huang, Ge Wang and Zhen Zhang

A Bluetooth Localization Algorithm Based on Map Path Calibration and Time Series Filtering 355
 Shaomeng Chen, Haiyong Luo, Qu Wang, Wenhua Shao and Fang Zhao

Synthetization of Fingerprint Recognition and Trilateration for Wi-Fi Indoor Localization Through Linear Kalman Filtering 373
 Junxi Tian, Zebo Zhou, Jin Wu, Shuang Du, Changgan Xiang and Changhong Kuang

Positioning Performance Analysis of Combined BDS/GPS-RTK Based on iGMAS Products 387
 Subo Wan, Qianxin Wang, Wenwen She and Mingbin Zhang

Environment Recognition Based on Temporal Filtering SVM	403
Yuze Wang, Peilin Liu, Xiaoguang Zhu, Xiaoxi Jin, Qiang Liu and Jiuchao Qian	
Pseudorange Double Difference Algorithm Based on Duty-Cycled Carrier Phase Smoothing on Low-Power Smart Devices	415
Qiang Liu, Rendong Ying, Yuze Wang, Jiuchao Qian and Peilin Liu	
Communication Tower Based Experiment and Analysis of Differential Augmentation for Auto-Steering Guidance of Agricultural Machinery	431
Bingbing Hu, Caicong Wu, Jinjiang Li, Xiaolong Li and Xingtao Liu	
Low-Cost Receiver Robust-RTD Algorithm Based on the Constraints of Velocity	443
Jian Zhang, Guorong Yu, Shuguo Pan, Guoliang Liu and Fei Shen	
The Research on Single Satellite Positioning System of Aircraft Based on Communication Signals of Beijing Inmarsat Ground Earth Station	455
Yuanliang Liu and Hongxia Kang	
Assessment of the Contribution of QZSS Combined GPS/BeiDou Positioning in Asia-Pacific Areas	467
Yize Zhang, Nobuaki Kubo, Junping Chen, Hu Wang and Jiexian Wang	
The Service Improvement of BDS Positioning Based on Advanced Equivalent Satellite Clock Calculation	479
Yangfei Hou, Junping Chen, Bin Wang and Jiexian Wang	
Rapid Ambiguity Resolution Algorithm for Multi-constellation Between Reference Stations Based on Ambiguity Tight Constraint	491
Ruicheng Zhang, Chengfa Gao, Shuguo Pan, Zhiyue Yan and Qing Zhao	
Part III Time-Space Datum and Temporal Frequency Technologies	
Using IGS RTS Products for Real-Time Subnanosecond Level Time Transfer	505
Guojun Li, Yuting Lin, Fengfeng Shi, Jialin Liu, Yuting Yang and Junwen Shi	
Time Transfer Analysis of GPS- and BDS-Precise Point Positioning Based on iGMAS Products	519
Yulong Ge, Xuhai Yang, Weijin Qin, Hang Su, Meifang Wu, Yuanxi Wang and Shengli Wang	
Simulation and Experiment of Multi-section Coil for C Field Used in Passive Hydrogen Maser	531
Zhibing Pan, Yonghui Xie, Tao Shuai, Pengfei Chen, Yuxian Pei, Chao Shen, Xiaoyan Pan, Yang Zhao and Chuanfu Lin	

Periodicity Systematic Error Analysis of GPS in Nonlinear Variations of Station Coordinates 539
 Yanbo Fu, Fuping Sun, Xinhui Zhu, Kai Xiao, Hailiang Dai and Longlong Zhang

An Accurate and Efficient Detecting and Correcting Algorithm for Abnormal Clock Data 549
 Fan Yang and Ran Zhang

Processing Method of Clock Break in VLBI Data Analysis 561
 Han Lei, Lu Cao, Ke Xu and Chong Wang

An Improved Abnormity Detection Method for Atomic Clock Frequency Based on Chi-Squared Statistics 573
 Huijie Song, Shaowu Dong, Meng Jiang, Yongnan Rao, Haibo Yuan and Hong Zhang

Research on Robust Timing Receiver with Joint Verification of Location and Clock Offset 581
 Dan Shen, Hang Gong, Zengjun Liu, Xiangwei Zhu and Feixue Wang

The Long-Term Performance Analysis and Comparison of BDS on-Board Atomic Clock Based on Precise Clock Products of IGS Analysis Centers 593
 Huijie Xue, Tianhe Xu, Qingsong Ai and Shengchao Wang

A Method on Laser Power Stabilization in Optical Detection Cesium Atomic Clock 607
 Ziyu Chen, Chang Liu, Shuhui Wang and Yanhui Wang

Analysis of BDS Satellite Clock in Orbit with ODTs and TWTT Satellite Clock Data 615
 Bin Wang and Junping Chen

Analytical Analysis Method of Whispering Gallery Mode of Sapphire Dielectric Resonators and the Optimization of the Q Value in Liquid Helium Temperature 623
 Xi Zhu, Haibo Chen, Kai Huang and Lianshan Gao

Part IV Satellite Navigation Augmentation Technology

Application of Gaussian Overbounding on Ionospheric Error Analysis for China SBAS 635
 Dun Liu, Liang Chen, Xiao Yu and Weimin Zhen

Analysis on Characteristics of Delay Errors Under Ionospheric Anomaly in China Area 649
 Dun Liu, Liang Chen, Xiao Yu and Weimin Zhen

Design of Mega-Constellations of LEO Satellites for Positioning 663
Xingchi He and Urs Hugentobler

A Required Navigation Performance Based Approach to Monitor the Accuracy and Integrity Performance of UAVs for Delivery Applications 675
Rui Sun, Yucheng Zhang, Bojia Ye and Washington Yotto Ochieng

SBAS Integrity Verification Based on the Extreme Value Theory 689
Yan Zhang, Zengjun Liu, Caihua Li, Xiaomei Tang and Gang Ou

Reduced ARAIM Subsets Method Determined by Threshold for Integrity Risk 701
Qian Meng, Jianye Liu, Qinghua Zeng, Shaojun Feng and Rui Xu

Comparison and Analysis of the Optimized AGC Setting and the Improved CNMC Method for the Ground Pseudolite Enhanced Beidou System 713
Jingyang Fu, Li Wang, Guangyun Li, Yanglin Zhou and Shuaixin Li

Research on BDS/GPS Integrated Navigation Satellite Selection Algorithm Based on Particle Swarm Optimization 727
Ershen Wang, Chaoying Jia, Tao Pang, Pingping Qu and Zhixian Zhang

A Minor Fault Detection Algorithm Based on Sliding-Window Accumulated Parity Vector 739
Rong Wang, Zhi Xiong, Jianye Liu, Chuanyi Li and Hangshuai Ma

The Construction Method of Ionospheric Delay Model with Combined BDS Phase and Range Observation 749
Zongpeng Pan, Hongzhou Chai, Yulong Kong, Rui Wang and Chunhe Liu

Improved RAIM Algorithm Based on Kalman Innovation Monitoring Method 759
Zhengnan Yang, Huaijian Li and Xiaojing Du

LEO Constellation Optimization Model with Non-uniformly Distributed RAAN for Global Navigation Enhancement 769
Deng Pan, Dan Sun, Jing Ren and Mingtao Li

Design of Navigation Message Authentication Assisted by Ground Based Augmentation Systems 779
Muzi Yuan, Zhe Liu, Xiaomei Tang, Shengqiang Lou and Gang Ou

Configuration Method of the Dynamic Navigation Network Based on Near Space 789
Lizhi Deng, Yingkui Gong, Hongxia Wan and Bingcheng Liu

Research on Ionospheric Grid Correction Strategy Based on Chinese Region 807
Kewu Chang, Rui Li and Haipeng Li

Satellite Integrity Autonomous Monitoring (SAIM) of BDS and Onboard Performance Evaluation 819
Lang Bian, Wenshan Liu, Tao Yan, Yueling Cao, Rui Li, Wei Wang, Xiao Liu, Wenyin Lei, Yansong Meng and Lixin Zhang

Part I
Satellite Navigation Technology

A Troposphere Tomography Method by Combining the Truncation Coefficient and Variance Component Analysis



Qingzhi Zhao, Yibin Yao and Linyang Xin

Abstract Traditional troposphere tomography method cannot use the GNSS signals penetrating from the side face of research area, which not only decreases the utilization rate of GNSS observation but also leads to a low percentage of voxels crossed by rays. In order to overcome this issue, the GNSS signals penetrating from the model's side face are also used to build the observation equation by introducing the truncation coefficient in this paper. Due to the fact that the tomography modeling is consists of various equations, including observation equation (using signals from the side and top faces of research area to build equations), horizontal and vertical equations, how to determine the weightings of different equations is a key to obtain the reliable tomographic result. Therefore, a method is proposed to determine the weightings of various equations based on the variance component analysis (VCA). The data from Satellite Positioning Reference Station Network (SatRef) of Hong Kong over the period of 27 days is selected for the tomography experiment. The tomographic result shows that the proposed method is of ability to obtain a good quality. Comparing to the traditional method, the utilization rate and number of voxels crossed by rays have been improved by 32.21 and 12.23%, respectively. When compared to the radiosonde data, the RMS error of the reconstructed integral water vapor (IWV) derived from the proposed method (4.2 mm) superior to that from the traditional method (5.2 mm). The comparison of water vapor profiles also shows that the proposed method with a RMS value of 1.30 g/m^3 , is smaller than that of traditional method with a value of 1.58 g/m^3 , and the accuracy of tomographic result based on the proposed method is increased by 17.7%.

Keywords Troposphere tomography · Truncation coefficient · VCA Radiosonde

Q. Zhao (✉)

College of Geomatics, Xi'an University of Science and Technology, Xi'an, China
e-mail: zhaqingzhia@163.com

Y. Yao · L. Xin

School of Geodesy and Geomatics, Wuhan University, Wuhan, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_1

1 Introduction

The earth observing technique has been developed significantly and applied into various areas since the emergence of global positioning system (GPS) in 1990s. The conception of GNSS meteorology was first proposed by Bevis et al. [1], and obtaining a rapid development in the related areas, like enhancing the accuracy of the numerical forecasting model and the predictability of disaster weathers [2–7, 8, 9]. However, the integrated water vapour (IWV) is the average value of some satellite rays mapped into the zenith direction, which cannot reflect the three-dimensional (3D) spatial-temporal variation of atmosphere water vapour. Therefore, the development of two-dimensional (2D) IWV is limited due to its disadvantages [10, 11].

Braun et al. [12] was first proposed the troposphere tomography conception using observed data of twenty stations in a local network based on the tomography idea introduced by Radon et al. [13]. In the next year, Flores et al. [14] first realized the reconstruction of wet refractivity based on the troposphere tomography technique in a local GPS network, and manifested the feasibility of the tomography technique to monitor the water vapour changed by comparing the tomographic result with that derived from the European Center for Medium Range Weather Forecast (ECMWF). Hereafter, GNSS tomography technique has obtained the rapid development in acquiring the 3D water vapour information, including the obtaining of slant water vapour (SWV) based on the un-differenced and doubled-differenced method [14, 11], as well as the establishment of troposphere tomography modelling by combing the Interferometric Synthetic Aperture Radar (InSAR) data [15, 16].

However, the most of studies mentioned above only used the satellite rays penetrating from the top boundary of tomography area to establish the observation equation, while the signals crossing from the model's side face were abandoned as the "ineffective information". Such behaviour not only decreases the utilization rate of satellite signals, but also reduces the coverage rate of tomography voxels. Rohm and Bosy [17] proposes a method to obtain the values of the outter part of the ray based on the ray-tracing water vapor model with the help of UNB3 model. Van baelen et al. [18] and Benevides et al. [19] use an empirically exponential negative function to calculate the value of the water vapor content inside the tomography area. However, some issues are remained to be discussed in those previous studies, including the method to use the signals penetrating from the side face of tomography area, the accuracy and reliability evaluation of the estimated values of the satellite rays and the reasonability of the selected data, etc.

Therefore, how to use the signals crossing out from the model's side face effectively is a key of troposphere tomography technique. In this paper, a troposphere tomography method is proposed, which use the signals mentioned above to build the observation equation by introducing the truncation coefficient. In addition, the variance component analysis (VCA) is also applied in the process of tomography modeling resolving so as to obtain the realistic 3D water vapour distribution information.

2 Tomography Theory

GNSS troposphere tomography refers to use amount of data from the ground-based receivers to reconstruct the 3D water vapor distribution of local area. At present, two methods are usually applied to build the observation equation of tomography modeling. The one is the parameterization method proposed by Flores et al. [14], which considers the water vapor value of the divided grid as a constant during the given period. This method is very convenient for ray tracing and travel time-computation, but introduces the boundary of the divided voxel and leads to the discontinuity of modeling parameter, in addition, the area of the numerical anomaly is only presented as a block. The other parameterized method is based on the node of grids, which regards the node as the unknown parameter and the value of other areas is interpolated using the nearby eight nodes. Comparing to the former method, this method is of ability to reflect the continuity of the water vapor information and eliminate the influence of artificial boundaries. Regardless of which method is used, the integrated form of SWV can be expressed as follows:

$$SWV = 10^{-6} \cdot \int_l \rho_v dl \quad (1)$$

where ρ_v is water vapor density and l is the distance between the satellite and receiver.

According to the tomography theory, the research area is divided into many 3D voxels and the value of each voxel center is regarded as an unknown parameter. Assuming that the water vapor density is a constant during the period (0.5 h), the water vapor content of satellite signal path can be discretized into the following expression:

$$SWV = \sum_{ijk} (a_{ijk} \cdot x_{ijk}) \quad (2)$$

where a_{ijk} represents the distance of satellite signals in the voxel (i, j, k) , while x_{ijk} refers to the water vapor density value in the voxel (i, j, k) .

Due to the lacking of sufficient data and the specificity of the tomography area, the distribution of satellite signal path is not the spatial optimum for the established observation equation. Therefore, the design matrix of the observation equation is ill-posed without the observed data covered the whole tomography area, which leads to a poor tomographic result. Generally, some physical or priori information are required to express the relationship between the voxels in horizontal and vertical directions [14, 20–24]. The distribution of water vapor density is continuous in the horizontal direction, and the correlation is large with respect to the close voxels while the correlation is decreased with the increasing of distance between voxels.

Therefore, the horizontal constraint equation can be established based on the mean filter method to reflect the relationship between horizontal voxels [25] and expressed as:

$$\mathbf{H} \cdot \mathbf{x} = \mathbf{0} \quad (3)$$

where \mathbf{H} is the coefficient matrix of horizontal constraint equation.

Due to the summation of the water vapor content for voxels vertically is irrelevant with the order of the voxels in different layers, therefore, the tomographic result only using the observation equation and horizontal constraint sometimes is unsatisfactory under the condition that height difference between GNSS stations is small in the flat area. In order to overcome this issue, the relationship between voxels vertically also needs to be established. In this paper, the vertical relationship between voxels is obtained according to the exponential negative relationship of water vapor with heights and expressed as:

$$x_h = x_0 \cdot e^{-z/H} \quad (4)$$

where x_h represents the water vapor density value at height of h . x_0 refers to the water vapor density at the ground face. H is the water vapor elevation with a value of 1–2 km, which can be fitted using observed data.

Consequently, the conventional troposphere tomography modeling can be obtained as:

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{H} \\ \mathbf{V} \end{pmatrix} \cdot \mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \quad (5)$$

where \mathbf{V} represents the coefficient matrix of the vertical constraint equation.

3 Establishing the Observation Equation of Signals Penetrating from the Model's Side Face Using the Truncation Coefficient

For the traditional troposphere tomography method, only the signals crossing the whole area can be used, as presented by yellow line in Fig. 1a, while the satellite rays penetrating from the model's side face are abandoned, as described by the green line in Fig. 1b. The above behavior reduces the utilization rate of observation data and the coverage rate of voxels crossed by rays, as described by the red blocks in Fig. 1b. Therefore, how to use those rays effectively to build the observation equation is a key to improve the stability of tomography modeling as well as the accuracy of tomographic result.

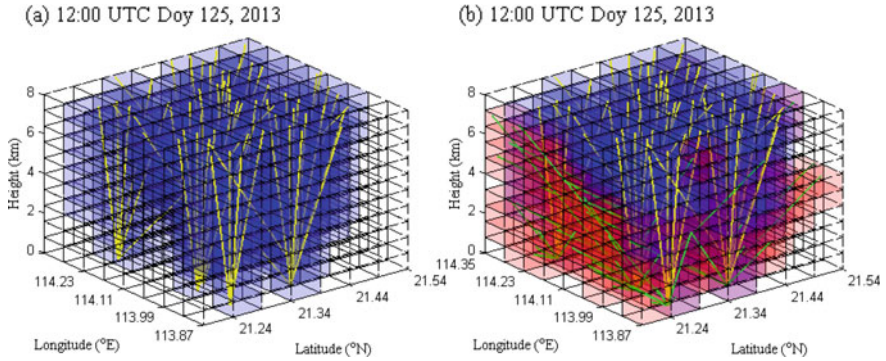


Fig. 1 Distribution of GPS signals, where **a** is the signals crossing out from the top face of modeling area whereas **b** includes both signals penetrating from the top and side face of modeling area. The shades of the blue and red represent the times of voxels crossed by rays

In this paper, the water vapor content of satellite rays in the tomography area for those signals penetrating from the model’s side face are estimated by introducing the truncation coefficient, and used for building the observation equation. The key steps of the proposed method are as follows:

- (1) Calculating the initial water vapor density of each voxels using the satellite rays cross the whole tomography area.
- (2) Narrowing the research area (as presented by the pink and blue rectangles in Fig. 2a) and makes the signals penetrating from the model’s top face cross from the side face of tomography area (see yellow line OQ in Fig. 2). Then, the truncation coefficient is defined, which refers to the proportion between the water vapor content of signal in the tomography area and the total water vapor value of this signal. As described in Fig. 2b, the truncation coefficient can be expressed as:

$$\alpha_{op} = \frac{SWV^{op}}{SWV^{Oq}} \tag{6}$$

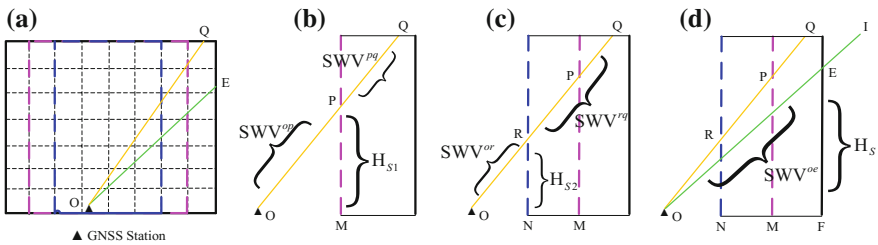


Fig. 2 Schematic of building the observation equation using the signals from the side face based on the truncation coefficient

where α_{op} is a truncation coefficient of signal OP for the whole signal OQ and $SWV^{oq} = SWV^{op} + SWV^{pq}$. Similarly, another truncation coefficient α_{or} of signal OQ can be calculated, as described in Fig. 2c. By this way, all the truncation coefficients of satellite signals penetrating from the model's top face can be calculated.

- (3) Establishing the truncation coefficient model (expressed as $\alpha = a + b \cdot \exp(1/H_s)$) based on the calculated truncation coefficients in step (2), which is used to estimate the water vapor content in the tomography area for those rays penetrating from the model's side face. As presented by the line OI in Fig. 2d, the water vapor value of OI can be expressed as

$$SWV_{oe} = \alpha_{oi} \times SWV_{oi} \quad (7)$$

where SWV_{oe} represents the water vapor content of signal OI in the tomography area while SWV_{oi} refers to the total water vapor value of signal OI. α_{oi} is the truncation coefficient of signal OI with respect to the height of H_s , the values of a and b in truncation coefficient model can be estimated by the least square method.

- (4) The tomography modeling proposed in this paper thus can be obtained by combining the traditional method and the added observation equation using the signals penetrating from the model's side face as follows:

$$\begin{pmatrix} \mathbf{A}_{m \times n} \\ \mathbf{H}_{m \times n} \\ \mathbf{V}_{m \times n} \\ \mathbf{As}_{m \times n} \end{pmatrix} \cdot \mathbf{x}_{n \times 1} = \begin{pmatrix} \mathbf{y}_{m \times 1} \\ \mathbf{0}_{m \times 1} \\ \mathbf{0}_{m \times 1} \\ \mathbf{ys}_{l \times 1} \end{pmatrix} \quad (8)$$

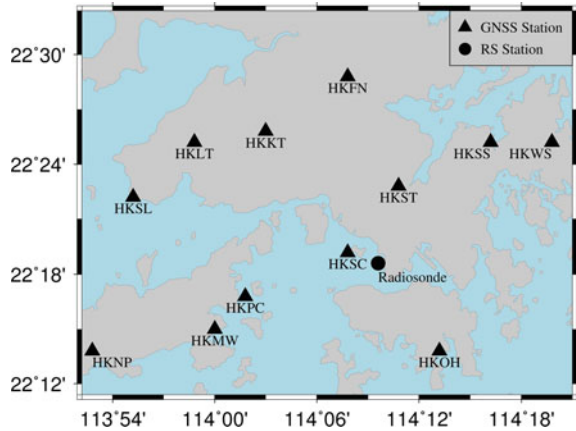
where \mathbf{As} is the coefficient matrix of observation equation using satellite rays penetrating from the model's side face, while \mathbf{ys} is the column vector of water vapor values for the above signals.

4 Experiment Description and Tomography Strategy

4.1 Experiment Description

To validate the proposed method in this paper, the data from twelve Satellite Positioning Reference Station Network (SatRef) of Hong Kong is selected for the period of Doy 124–150, 2013 to perform the tomography experiment. Tomography area ranges from N21.19° to N21.54° and E113.87° to E114.35° in latitudinal and longitudinal directions, respectively, while the tomography height is selected as 8 km. The steps in three directions are 0.06°, 0.05° and 0.8 km, respectively. The GPS observation is processed based on the GAMIT/GLOBK (v10.5) software,

Fig. 3 Geographic distribution of GNSS receivers and radiosonde



while the ZTD and the gradient parameters in east-west and south-north directions are estimated with the intervals of 0.5 and 2 h, respectively. In addition, there is a radiosonde station in the experiment area, which is used to validate the accuracy of the tomographic result.

For any tomography modeling, the accuracy of the tomographic result is a key to evaluate the quality of established modeling. Therefore, two strategies are selected to build the tomography modeling and evaluate the quality of the established modeling. Two strategies are as follows, (1) only considering the satellite signals crossing the whole area to build the observation equation and the final tomography modeling is described as Eq. (5); (2) both rays penetrating from the top and side faces of tomography area are considered to establish the observation equation and the final modeling is described as Eq. (8) (Fig. 3).

4.2 Troposphere Tomography Resolving Strategy Based on the VCA

For troposphere tomography, posterior variance component analysis (VCA) is a method to determine the weightings of different kinds of input information. In this paper, the idea of co-integration test in the macro-economics is introduced to analyze the posterior unit weight variances of different input information, and then to judge whether the linear combination of a set of posterior unit weight variances is of the stable equilibrium relations. Based on the above method, the stability of unit weight variances of different equations can be determined, which means the ratio between arbitrary two unit weight variances is equal to one. The specific steps can be expressed as follows:

- (1) Initializing the unit weight variances of different equations as one, which means $\hat{\sigma}_{0_A}^2 = \hat{\sigma}_{0_{A_s}}^2 = \hat{\sigma}_{0_H}^2 = \hat{\sigma}_{0_V}^2 = 1$, where $\hat{\sigma}_{0_i}^2$, ($i = A, A_s, H, V$) represent the initial unit weight variances of observation equations using the signals penetrating from the top, and side face of tomography area, the horizontal and vertical constraints, respectively.
- (2) Calculating the posterior residuals of input information $\mathbf{v}_A, \mathbf{v}_H, \mathbf{v}_V$ and \mathbf{v}_{A_s} based on the following formula:

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_A \\ \mathbf{v}_H \\ \mathbf{v}_V \\ \mathbf{v}_{A_s} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{m1 \times n1}^A \\ \mathbf{A}_{m1 \times n1}^H \\ \mathbf{A}_{m1 \times n1}^V \\ \mathbf{A}_{m1 \times n1}^{A_s} \end{bmatrix} \cdot \hat{\mathbf{X}} - \begin{bmatrix} \mathbf{y}_{m1 \times 1} \\ \mathbf{0}_{m2 \times 1} \\ \mathbf{0}_{m3 \times 1} \\ \mathbf{y}_{sm4 \times 1} \end{bmatrix} \quad (9)$$

- (3) Updating the unit weight variances of input information $\hat{\sigma}_{0_A}^2, \hat{\sigma}_{0_H}^2, \hat{\sigma}_{0_V}^2$ and $\hat{\sigma}_{0_{A_s}}^2$. Here, the simplified variance component estimation [26] is used to calculate the unit weight variances of various equations:

$$\hat{\sigma}_{0_{A_s}}^2 = \frac{\mathbf{v}_q^T \mathbf{P}_{0q} \mathbf{v}_q}{n_q - \text{tr}(\mathbf{N}^{-1} \mathbf{N}_q)} \quad (10)$$

where $\mathbf{N} = \mathbf{A}^T \mathbf{P} \mathbf{A}$, $\mathbf{N}_q = \mathbf{A}_q^T \mathbf{P}_q \mathbf{A}_q$, $\mathbf{A} = [\mathbf{A}_A \ \mathbf{A}_H \ \mathbf{A}_V \ \mathbf{A}_{A_s}]^T$. n_q ($q = A, H, V, A_s$) refers to the number of different input information. tr represents the rank of the matrix. In addition, the updated unit weight variance and the formula $v_i^A > \lambda \hat{\sigma}_{0_A}$ can be used to remove the outliers in the observation equation, λ is an empirical value and $\lambda = 3$ is selected in our experiment.

- (4) Judging the reasonability of the posterior unit weight variances of various input information based on the co-integration test [26]. The main idea is to determine whether the values is of the stable equilibrium relations. The specific testing steps are as follows:

(a) Establishing the relationship of different unit weight variances, here, the first-order self-regression variables sequence is selected as:

$$\hat{\sigma}_{0_q}^2 = \varphi \cdot \hat{\sigma}_{0_{q-1}}^2 + \varepsilon_{0_q} \quad (11)$$

where $\varphi = 1$.

- (b) Calculating the improved Dickey-Fuller statistic $t(\varphi)$ [27, 28]:

$$t(\varphi) = \frac{\text{abs}(\hat{\varphi} - \varphi)}{S(\varphi)} \quad (12)$$

where $\hat{\varphi}$ is the estimated value of φ using the least square method, and

$$S(\varphi) = \sqrt{\frac{S_N^2}{\sum_{q=2}^N \hat{\sigma}_{0_{q-1}}^2}}, \quad S_N^2 = \frac{\sum_{q=2}^N (\hat{\sigma}_{0_q}^2 - \varphi \cdot \hat{\sigma}_{0_{q-1}}^2)}{N-1} \quad (13)$$

where N is the number of different kinds of equations.

(c) Giving the assumption conditions. Due to the accurate weightings cannot be given for the first time, the assumption is presented as follows: H_0 refers to the instable variables sequence, which means $\hat{\sigma}_{0_A}^2 \neq \hat{\sigma}_{0_H}^2 \parallel \hat{\sigma}_{0_A}^2 \neq \hat{\sigma}_{0_V}^2 \parallel \hat{\sigma}_{0_A}^2 \neq \hat{\sigma}_{0_P}^2 \parallel \hat{\sigma}_{0_H}^2 \neq \hat{\sigma}_{0_V}^2 \parallel \hat{\sigma}_{0_H}^2 \neq \hat{\sigma}_{0_{As}}^2 \parallel \hat{\sigma}_{0_V}^2 \neq \hat{\sigma}_{0_{As}}^2$; H_1 refers to the stable variable sequence, which means $\hat{\sigma}_{0_A}^2 = \hat{\sigma}_{0_H}^2 = \hat{\sigma}_{0_V}^2 = \hat{\sigma}_{0_{As}}^2$.

(d) Determining the reasonable threshold of $t_x(\varphi)$. For different conditions, this value is various and should be determined based on the experiment. In our experiment, $t_x(\varphi) = 0.1$ is selected.

(e) Accepting the assumption based on the calculated $t(\varphi)$.

$$\begin{cases} \text{if } t(\varphi) > t_x(\varphi), & \text{accept } H_0, \text{ reject } H_1 \\ \text{if } t(\varphi) \leq t_x(\varphi), & \text{accept } H_1, \text{ reject } H_0 \end{cases} \quad (14)$$

(5) If the condition H_0 is accepted, then updating the weight matrix of various equations and back to the step (2).

$$\mathbf{P}_q^{ite} = \frac{c}{\hat{\sigma}_{0_q}^2} \cdot \mathbf{P}_q^{(ite-1)} \quad (15)$$

where c is an arbitrary value and ite is the iteration times.

(6) If H_1 is accepted, then the final unit weight variances of different input information are determined.

5 Result Analysis

5.1 IWV Comparison

The IWV time series derived from the different tomography strategies are first compared with that from the radiosonde data over the period of Doy 124–150, 2013. Figure 4 gives the comparison result of different tomographic strategies and the radiosonde data, from which it can be concluded that the change trends of IWV

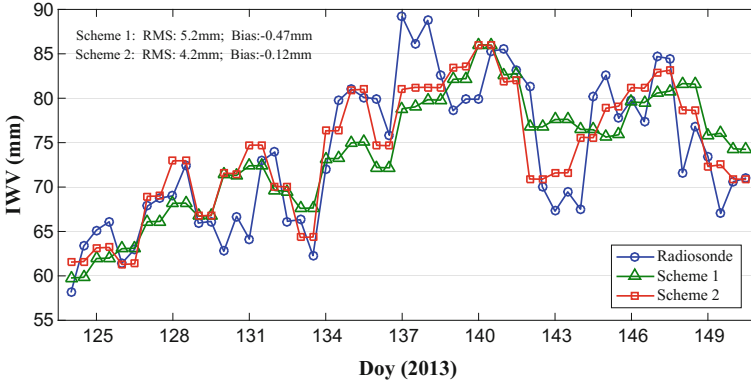


Fig. 4 Comparison of IWV derived from different methods and radiosonde data

time series are similar for different tomography strategies. The statistical result over the 27-days reveals that the root mean square (RMS) error and bias of strategy 2 with values of 4.2 and -0.14 mm, respectively, are less than that of strategy 1 with values of 5.2 and -0.47 mm, respectively.

It is worth noting that the IWV is the integrated value of water vapor density vertically, which cannot reflect the 3D distribution of water vapor information. Therefore, the tomographic result of strategy 2 not always superior to that of strategy 1. Because the value of IWV is unchanged if the position of arbitrary two layers changed, however, the spatial distribution of water vapor has changed significantly. Therefore, it is necessary to further compare the water vapor profiles of different strategies.

5.2 Water Vapour Profile Comparison

Water vapor profiles at two specific epochs (00:00 UTC Doy 124, 2013 and 12:00 UTC, Doy 137) are first compared and Fig. 5 presents the comparison of water vapor profiles derived from different strategies with that from radiosonde data. It can be concluded from Fig. 5 that, comparing to strategy 1, the water vapor profile of strategy 2 has a better agreement with that from radiosonde data. Statistical result reveals that the RMS errors of strategy 2 at two epochs are 1.31 and 1.83 g/m^3 , respectively, while the values of strategy 1 are 1.52 and 2.52 g/m^3 , respectively. Figure 6 present the daily RMS error of differences between different strategies and radiosonde data at UTC 00:00 and 12:00, respectively, during the experimental period. It can be seen from Fig. 6 that the tomographic result of strategy 2 is superior to that of strategy 1 and the statistical result reveals that the RMS/bias of the proposed method and traditional method are 1.30 g/m^3 / -0.03 g/m^3 and 1.58 g/m^3 / -0.09 g/m^3 , respectively.

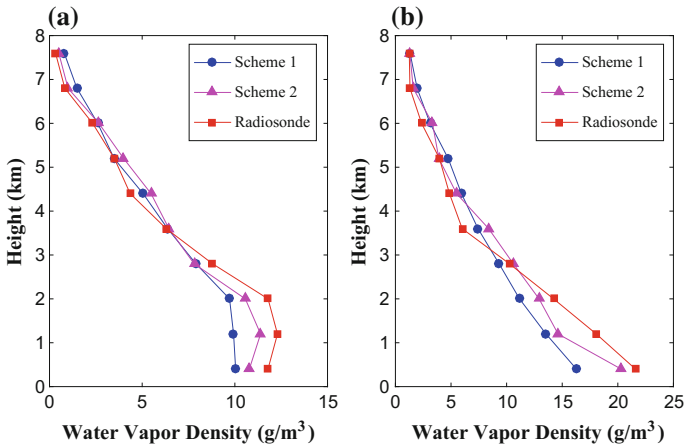


Fig. 5 Water vapor density comparison derived from different schemes at specific epochs, **a** 00:00–30:00 UTC Doy 124, 2013; **b** 12:00–12:30 Doy 137, 2013

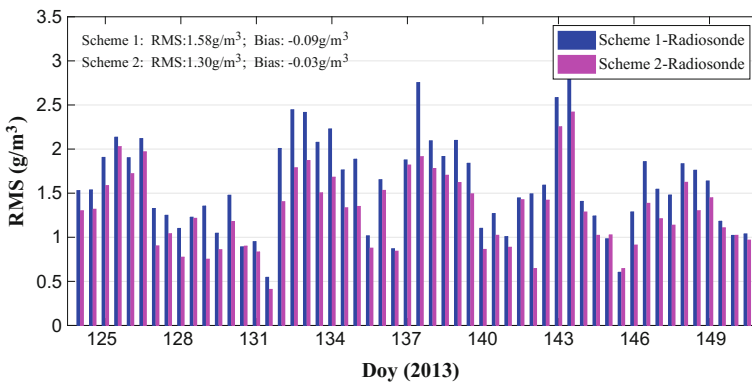


Fig. 6 Comparison of water vapor density profiles with radiosonde data for different schemes during the tested period

To further compare the relationship between the water vapor profiles and altitudes, the relative error is introduced to analyze the 27-days experimental data. Figure 7 presents the RMS error and relative error of water vapor density differences between various strategies and radiosonde data at different altitudes over the experimental period. It is clearly seen that the RMS and relative error of strategy 2 are less than that of strategy 1 at different altitudes, which is enough to manifest the better performance of the proposed method by using the signals penetrating from the model’s side face, when comparing to the traditional method.

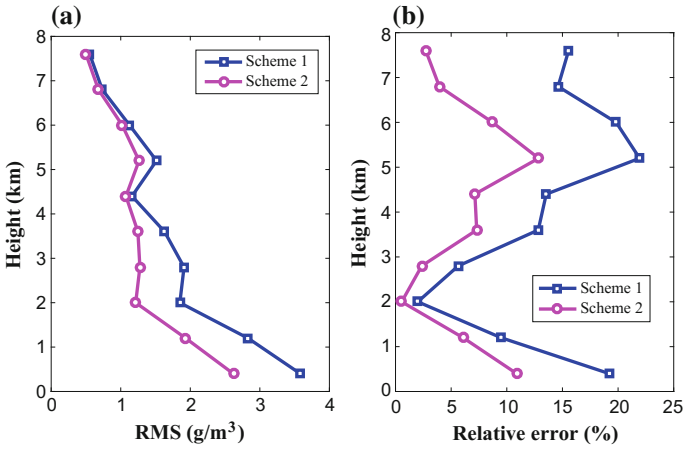


Fig. 7 RMS error and relative error of water vapor density between different schemes and radiosonde at various altitudes during the experimental period

6 Conclusion

In order to overcome the disadvantage of traditional troposphere tomography method, which cannot use the signals penetrating from side face of tomography area, a method is proposed by introducing the truncation coefficient in this paper. This method can use signals penetrating from the model's side face to build observation equation, in addition, variance component analysis is also used during the process of tomography modeling resolving. The proposed method is validated using the data from SatRef of Hong Kong over the period of Doy 124–150, 2013. Comparing to the traditional method, the RMS error of the proposed method has been improved by 17.7%, especially for the low layers of tomography area, and the quality of tomographic result has been improved significantly.

Acknowledgements The authors would like to thank IGAR for providing access to the web-based IGAR data. The Lands Department of HKSAR is also acknowledge for providing GPS data from the Hong Kong Satellite Positioning Reference Station Network (SatRef) and the corresponding meteorological data. This research was supported by the Excellent Youth Science and Technology Fund Project of Xi'an University of Science and Technology (2018YQ3-15) and the Startup Foundation for Doctor of Xi'an University of Science and Technology (2017QDJ041).

References

1. Bevis M, Businger S, Herring T, Rocken C et al (1992) GPS meteorology-remote sensing of atmospheric water vapor using the global positioning system. *J Geophys Res* 97(D14):15787–15801
2. Rocken C, Ware R, Van Hove T et al (1993) Sensing atmospheric water vapor with the global positioning system. *Geophys Res Lett* 20(23):2631–2634
3. Duan J, Bevis M, Fang P et al (1996) GPS meteorology: direct estimation of the absolute value of precipitable water. *J Appl Meteorol* 35(6):830–838
4. Emardson TR, Johansson J, Elgered G (2000) The systematic behavior of water vapor estimates using four years of GPS observations. *IEEE Trans Geosci Remote Sens* 38(1):324–329
5. Gendt G, Dick G, Reigber C et al (2004) Near real time GPS water vapor monitoring for numerical weather prediction in Germany. *J Meteorol Soc Jpn* 82(1B):361–370
6. Smith TL, Benjamin SG, Gutman SI et al (2007) Short-range forecast impact from assimilation of GPS-IPW observations into the rapid update cycle. *Mon Weather Rev* 135(8):2914–2930
7. Raja MRV, Gutman SI, Yoe JG et al (2008) The validation of AIRS retrievals of integrated precipitable water vapor using measurements from a network of ground-based GPS receivers over the contiguous United States. *J Atmos Ocean Technol* 25(3):416–428
8. de Haan S, Holleman I, Holtslag AA (2009) Real-time water vapor maps from a GPS surface network: construction, validation, and applications. *J Appl Meteorol Climatol* 48(7):1302–1316
9. Lee SW, Kouba J, Schutz B et al (2013) Monitoring precipitable water vapor in real-time using global navigation satellite systems. *J Geodesy* 87(10–12):923–934
10. Brenot H, Wautelet G, Warnant R et al (2014) GNSS meteorology and impact on NRT position. In: European navigation conference (ENC) GNSS 2014
11. Chen B, Liu Z (2014) Voxel-optimized regional water vapor tomography and comparison with radiosonde and numerical weather model. *J Geodesy* 88(7):691–703
12. Braun J, Rocken C, Meertens C et al (1999) Development of a water vapor tomography system using low cost L1 GPS receivers. In: Ninth ARM science team meeting, US Department of Energy, San Antonio, Tex, pp 22–26
13. Radon J (1917) Über die Bestimmung von Funktionen durch ihre In-te-gral-werte längs gewisser Mannigfaltigkeiten. *Comput Tomogr* 69:262–277
14. Flores A, Ruffini G, Rius A (2000) 4D tropospheric tomography using GPS slant wet delays. *Annales Geophysicae* 18(2):223–234
15. Alshawaf F (2013) Constructing water vapor maps by fusing InSAR, GNSS and WRF data. Karlsruhe, Karlsruher Institut für Technologie (KIT), Dissertation
16. Benevides P, Nico G, Catalao J et al (2015) Merging SAR interferometry and GPS tomography for high-resolution mapping of 3D tropospheric water vapour. In: IEEE international geoscience and remote sensing symposium (IGARSS), pp 3607–3610
17. Rohm W, Bosy J (2011) The verification of GNSS tropospheric tomography model in a mountainous area. *Adv Space Res* 47(10):1721–1730
18. Van Baelen J, Reverdy M, Tridon F et al (2011) On the relationship between water vapour field evolution and the life cycle of precipitation systems. *Q J R Meteorol Soc* 137(S1):204–223
19. Benevides P, Catalão J, Miranda PM (2014) Experimental GNSS tomography study in Lisbon (Portugal). *Física de la Tierra* 26:65–79
20. Hirahara K (2000) Local GPS tropospheric tomography. *Earth Planets Space* 52(11):935–939
21. Troller M, Burki B, Cocard M, Geiger A et al (2002) 3-D refractivity field from GPS double difference tomography. *Geophys Res Lett* 29:2149–2152
22. Perler D, Geiger A, Hurter F (2011) 4D GPS water vapour tomography: new parameterized approaches. *J Geophys Res* 85:539–550
23. Skone S, Hoyle V (2005) Troposphere modeling in a regional GPS network. *Positioning* 4(1&2):230–239

24. Bender M, Raabe A (2007) Preconditions to ground based GPS water vapour tomography. *Ann Geophys* 25(8):1727–1734
25. Wen D, Liu S, Tang P (2010) Tomographic reconstruction of ionospheric electron density based on constrained algebraic reconstruction technique. *GPS Solutions* 14(4):375–380
26. Engle RF, Granger CWJ (1987) Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55(2):251–276
27. Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. *J Am Stat Assoc* 74(366):427–431
28. Dickey DA, Fuller WA (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49(4):1057–1072

The Design and Implementation of a SDS-TWR Based Wireless Location System



Yao Wang, He Yang and Moyu Sha

Abstract In order to solve seamless positioning about indoor and outdoor, a SDS-TWR (Symmetric Double Side-Two Way Ranging) based wireless location system was proposed. The CSS (Chirp Spread Spectrum) signal, which has wide band characteristic, was used to obtain high-accuracy measurement of signal propagation delay. Least squares combining with linear Kalman filter algorithm was proposed in the positioning progress with linearization. Despite of increasing the complex of positioning algorithm, that improved the positioning accuracy, especially the stability of dynamic positioning situation. Finally, a hardware prototype system was exploited. And some tests were carried in the real indoor and outdoor environment. The results show that ranging error would not change significantly with distance increasing. And both of static and dynamic positioning accuracy are under one meter.

Keywords Indoor and outdoor positioning · CCS · SDS-TWR
Least square · Kalman filter

1 Introduction

Indoor and outdoor wireless positioning technology, especially for indoor positioning, has been a hot issue in recent scientific research, and as we known GNSS (Global Navigation Satellite System) positioning is mostly used all over the world. However, it is not suitable for indoor positioning field because of its high signal attenuation through the building [1]. Nowadays, it is different to balance both

Y. Wang (✉) · H. Yang · M. Sha
SKL of Satellite Navigation System and Engineering Technology, Hebei, China
e-mail: wy_hrbeu@163.com

Y. Wang
China Electronic Technology Group Corporation 54 Research Institute, Hebei, China

H. Yang
Hebei University of Science and Technology, Shijiazhuang, China

positioning accuracy and system cost. The Chirp signal is suitable for ranging in indoor environment, because of its advantages of strong multipath resolution, strong noise immunity and little frequency offset. At the same time, because of the development of surface acoustic wave device technology, the cost of Chirp signal is reduced effectively [2].

A wireless location system based on CSS technology was designed and implemented. The SDS-TWR was used so that each node didn't have to time synchronized. And the Chirp signal, which has wide band characteristic, was used to obtain high accuracy ranging. Meanwhile, the least square algorithm and linear Kalman filter algorithm were combined to realize the meter-level positioning. Finally, all the design demands were verified through the testing in actual environment.

2 Chirp Spread Spectrum Technology

The CSS technology, as a kind of communication method for spreading Chirp signals, has many advantages, such as stronger multipath resolving ability, lower transmitting power, stronger anti-frequency-offset ability, and so on. The instantaneous frequency of Chirp signal varies linearly in a signal period. So the CSS signal is also called as linear frequency modulation signal. The Chirp signal could be expressed by

$$s(t) = \cos \left[2\pi \left(f_0 t \pm \frac{kt^2}{2} \right) \right], \quad -\frac{T}{2} \leq t \leq \frac{T}{2} \quad (1)$$

In which, f_0 —signal frequency, Hz;

k —FM factor, which is Up-Chirp when k is plus, and minus for Down-Chirp.

The unit impulse response of the matched filter of the Chirp signal is $h(t)$:

$$h(t) = a \cos \left[2\pi \left(f_0 t \mp \frac{kt^2}{2} \right) \right], \quad -\frac{T}{2} \leq t \leq \frac{T}{2} \quad (2)$$

Take an Up-Chirp signal, for example, by matching the filter as follows:

$$\begin{aligned} Y(t) &= a \int_{-T/2}^{T/2} \cos \left[2\pi \left(f_0 \tau + \frac{k\tau^2}{2} \right) \right] \cos \left\{ 2\pi \left[f_0 (t - \tau) + \frac{k(t - \tau)^2}{2} \right] \right\} d\tau \\ &\approx \frac{a}{2} \int_{-T/2}^{T/2} \cos \left[2\pi \left(f_0 t + kt\tau - \frac{kt^2}{2} \right) \right] d\tau \end{aligned} \quad (3)$$

As mentioned above, after the integral and differential operations, two items are obtained, second of which contain high frequency components, so they are ignored:

$$y(t) = a \frac{\sin[\pi kt(T - |t|)]}{2\pi kt} \cos 2\pi f_0 t, \quad -T < t < T \quad (4)$$

Take $a = 2\sqrt{k}$ and $k = B/T$ into the formulation above and it will get (Fig. 1):

$$y(t) = \sqrt{BT} \frac{\sin\left[\pi Bt\left(1 - \frac{|t|}{T}\right)\right]}{\pi Bt} \cos 2\pi f_0 t, \quad -T < t < T \quad (5)$$

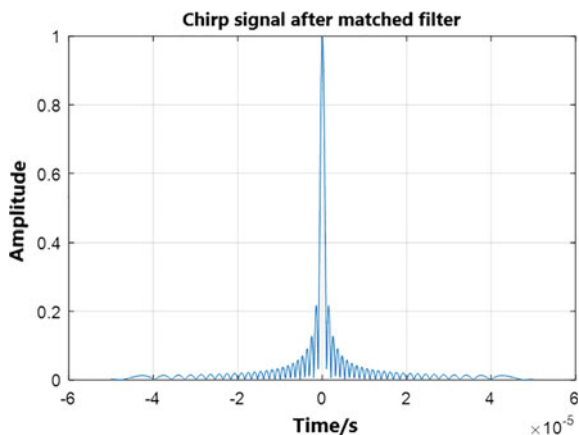
As shown in the figure above, it can be seen that the output waveform is similar to the sinc function, and the time domain characteristics are sharper. This will reduce the influence of multipath signal superposition caused by the enhanced signal multipath resolution, this property is known as pulse compression characteristics, plays a more important role in [4] application of Chirp signal in wireless positioning direction.

3 Analysis of System Positioning Algorithms

3.1 System Positioning Model

The localization model of the system needs to introduce the concept of anchor node and tag node in the system briefly. The anchor node refers to the fixed node in the wireless network, and the tag node refers to the moving target to be located. If there are N anchor nodes in the system, the location of the tag nodes can be achieved by processing the ranging values of the N anchor nodes to the tag nodes.

Fig. 1 The Chirp signal after matched filtering



For the sake of simplification, the system only considers the location in the two-dimensional plane. The number of anchor nodes involved in the system is 4, and the coordinates of the 4 anchor nodes are respectively (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) ,. For convenience, unity is recorded as (x_i, y_i) , $i = 1, 2, 3, 4$. The coordinates of the tag nodes are unknown. Set as (x, y) . Thus, the distance from the anchor node i to the tag node d_i can be expressed as:

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}, \quad i = 1, 2, 3, 4 \quad (6)$$

Taking into account the measurement error v_i of the system, the 4 measurement equations can be expressed as:

$$\hat{d}_i = d_i + v_i, \quad i = 1, 2, 3, 4 \quad (7)$$

Obviously, because of the measurement error, the set of measurement equations is overdetermined, and the coordinates of the final tag nodes can be obtained by solving the set of overdetermined equations, and the localization of the tag nodes can be achieved.

3.2 Location Algorithm

According to the positioning of the system model, the least squares method and Calman filter positioning solution based on, but because the measurement equations are nonlinear, so the need for the least squares method and Calman filter linear processing [5]. The least square method is used to linearize the measurement equation by introducing an intermediate variable in order to obtain the coordinates of the label nodes. First, the formula (7) is substituted into Eq. (6), and the following nonlinear equations are obtained:

$$\begin{aligned} \hat{d}_1 - v_1 &= \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \\ \hat{d}_2 - v_2 &= \sqrt{(x_2 - x)^2 + (y_2 - y)^2} \\ \hat{d}_3 - v_3 &= \sqrt{(x_3 - x)^2 + (y_3 - y)^2} \\ \hat{d}_4 - v_4 &= \sqrt{(x_4 - x)^2 + (y_4 - y)^2} \end{aligned} \quad (8)$$

The squares in both sides of the form are obtained:

$$\begin{aligned}
 2x_1x + 2y_1y + (x^2 + y^2) &= (\hat{d}_1 - v_1)^2 - (x_1^2 + y_1^2) \\
 2x_2x + 2y_2y + (x^2 + y^2) &= (\hat{d}_2 - v_2)^2 - (x_2^2 + y_2^2) \\
 2x_3x + 2y_3y + (x^2 + y^2) &= (\hat{d}_3 - v_3)^2 - (x_3^2 + y_3^2) \\
 2x_4x + 2y_4y + (x^2 + y^2) &= (\hat{d}_4 - v_4)^2 - (x_4^2 + y_4^2)
 \end{aligned} \tag{9}$$

Definition:

$$\begin{aligned}
 R &= \sqrt{x^2 + y^2} \\
 R_i &= \sqrt{x_i^2 + y_i^2}
 \end{aligned} \tag{10}$$

Take R and R_i into formula (9), It will get as follows:

$$\begin{aligned}
 -2x_1x - 2y_1y + R^2 &= \hat{d}_1^2 - R_1^2 + v_1^2 - 2\hat{d}_1v_1 \\
 -2x_2x - 2y_2y + R^2 &= \hat{d}_2^2 - R_2^2 + v_2^2 - 2\hat{d}_2v_2 \\
 -2x_3x - 2y_3y + R^2 &= \hat{d}_3^2 - R_3^2 + v_3^2 - 2\hat{d}_3v_3 \\
 -2x_4x - 2y_4y + R^2 &= \hat{d}_4^2 - R_4^2 + v_4^2 - 2\hat{d}_4v_4
 \end{aligned} \tag{11}$$

So, it also can be expressed as:

$$h = G\theta + v \tag{12}$$

Among them,

$$\begin{aligned}
 \theta &= [x \ y \ R^2]^T \\
 h &= \begin{bmatrix} \hat{d}_1^2 - R_1^2 \\ \hat{d}_2^2 - R_2^2 \\ \hat{d}_3^2 - R_3^2 \\ \hat{d}_4^2 - R_4^2 \end{bmatrix} \\
 G &= \begin{bmatrix} -2x_1 & -2y_1 & 1 \\ -2x_2 & -2y_2 & 1 \\ -2x_3 & -2y_3 & 1 \\ -2x_4 & -2y_4 & 1 \end{bmatrix} \\
 v &= [v_1^2 - 2\hat{d}_1v_1 \quad v_2^2 - 2\hat{d}_2v_2 \quad v_3^2 - 2\hat{d}_3v_3 \quad v_4^2 - 2\hat{d}_4v_4]^T
 \end{aligned} \tag{13}$$

So the least square solution is

$$\hat{\theta} = (G^T G)^{-1} G^T h \quad (14)$$

According to the Calman filter, because Calman filter applies only to the linear system, therefore, this system will put the solution value as a pseudo measurement Calman filtering value of least squares, using linear Kalman filtering of approximate solution for further processing, to obtain more precise positioning value [6]. The deduction of the state space model of Kalman filter is given.

Since the system is to locate the tag nodes in the two-dimensional plane, the state value $X(k)$ of the system is:

$$X(k) = \begin{bmatrix} x(k) \\ \dot{x}(k) \\ y(k) \\ \dot{y}(k) \end{bmatrix} \quad (15)$$

Among them, $x(k)$ and $\dot{x}(k)$ represent the location and speed of the tag node at the k moment, the x direction, and the $y(k)$ and $\dot{y}(k)$ indicate the position and speed of the y in the k direction at the time of the tag node.

Set the system's observations $Y(k)$

$$Y(k) = \begin{bmatrix} x_{LS}(k) \\ y_{LS}(k) \end{bmatrix} \quad (16)$$

Among them, $x_{LS}(k)$ and $y_{LS}(k)$ the estimated coordinate values obtained by the linear least squares method.

Therefore, the observation model of the system is:

$$Y(k) = HX(k) + V(k) \quad (17)$$

Among them, $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is the observation matrix, $V(k)$ represents the estimation error of the linear least squares method, assuming that it is the zero mean, and the variance σ_v^2 is the Gauss white noise.

Because the two positioning system time interval generally shorter, can within the time interval positioning motion obeys the uniformly accelerated linear motion, and the time interval is T_0 , the x direction of the movement as an example, the Newton equation of motion system subject to the following:

$$x(k+1) = x(k) + \dot{x}(k) \cdot T_0 + \frac{1}{2} \cdot a(k) \cdot T_0^2 \quad (18)$$

$$\dot{x}(k+1) = \dot{x}(k) + a(k) \cdot T_0 \quad (19)$$

In the formula, $a(k)$ by $u(k)$ maneuvering acceleration $w(k)$ and random acceleration, due to the positioning of the label node is in real time, and the trajectory of the tag nodes cannot be predicted, so the system design of maneuvering acceleration tag nodes is 0, which only consider state modeling of the system under the condition of random acceleration.

If only $x(k)$ and $\dot{x}(k)$ is considered as the system state, the equation of state of the system can be expressed as:

$$\begin{bmatrix} x(k+1) \\ \dot{x}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & T_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ \dot{x}(k) \end{bmatrix} + \begin{bmatrix} 0.5T_0^2 \\ T_0 \end{bmatrix} w(k) \quad (20)$$

When the state is extended to two dimensions, the equation of state can eventually be expressed as:

$$\begin{bmatrix} x(k+1) \\ \dot{x}(k+1) \\ y(k+1) \\ \dot{y}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & T_0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ \dot{x}(k) \\ y(k) \\ \dot{y}(k) \end{bmatrix} + \begin{bmatrix} 0.5T_0^2 & 0 \\ T_0 & 0 \\ 0 & 0.5T_0^2 \\ 0 & T_0 \end{bmatrix} W(k) \quad (21)$$

And It also can be expressed by

$$\begin{bmatrix} x_{LS}(k) \\ y_{LS}(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ \dot{x}(k) \\ y(k) \\ \dot{y}(k) \end{bmatrix} + V(k) \quad (22)$$

The above two models are the state space model of the system. After comparison, it is not difficult to get the state transition matrix Φ , noise driving matrix Γ and observation matrix H of the system, and then the linear Calman filter is carried out.

The state equation and observation equation of the system, we can see that the estimated linear least squares value as a pseudo observation Calman filter, the system can be linearized, and the linear Calman filtering is applied in the system, reducing the system by using non linear and nonlinear Calman filter complexity for real-time positioning system in terms of reducing the positioning delay, has high practical significance.

4 System Design and Implementation

4.1 System Framework

The schematic diagram of the system is shown in Fig. 2.

The anchor node (including the main anchor node and anchor node from) fixed on each known position on the label node by SDS-TWR location algorithm

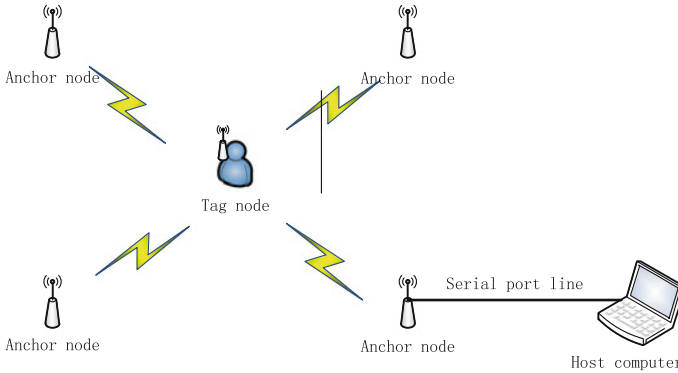


Fig. 2 Schematic diagram of system structure

respectively to each anchor node for location, location after the success of the value ranging package is transmitted to the main anchor node, main anchor node to realize data transmission and host computer the host computer through the serial port, after extracting range data, calculate the coordinates of the target node, in order to achieve the node positioning.

4.2 System Ranging Design

As a distance measurement based wireless indoor positioning system, the distance between the anchor node and the tag node is very important. The designed system avoids the complicated time synchronization problem between nodes. The SDS-TWR algorithm is used to measure the nodes. From the work, reducing the system to achieve the hardware cost [7]. This section will be in the software implementation on the basis of this distance measurement process for a detailed analysis.

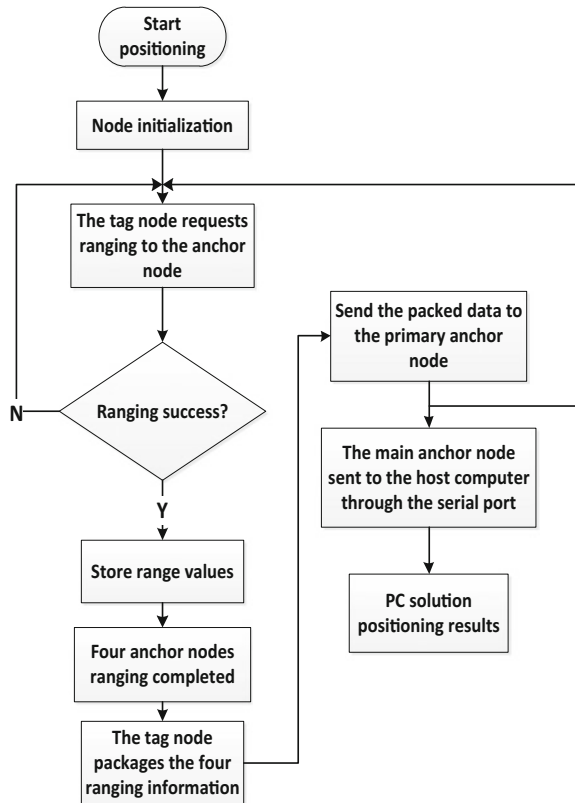
The system is designed by the label node to actively request the ranging, the anchor node passive acceptance of the ranging request. The ranging process is realized by calling the function. In the function, the following important functions are mainly included which is mainly used to request the distance measurement of the node and complete the ranging process. is mainly responsible for reading between nodes Data and the completion of feedback work, the function mainly includes two call function, mainly used to receive remote node feedback back to the data and processing, mainly for the node in the receiving distance request after the automatic reply; Mainly calculate the final range between nodes. In the process of SDS-TWR algorithm, the label node and the anchor node are mainly in the following states: RANGING_READY, RANGING_START, RANGING_ANSWER1, RANGING_ANSWER2, RANGING_SUCCESSFULLY; respectively, the preparation of ranging, starting ranging, the first distance, Secondary range, to complete the five stages of distance measurement.

4.3 System Positioning Design

In this paper, the indoor wireless positioning system includes four anchor nodes (including a main anchor node) and a label node, the specific positioning workflow shown in Fig. 3.

The positioning scheme of the system is shown in the above figure. First, the nodes are initialized. After the nodes receive the request, the tag node will get the distance between the anchor nodes and the anchor nodes through the SDS-TWR ranging algorithm, and then the four nodes The distance measurement is carried out and sent to the main anchor node. The main anchor node sends the ranging data to the host computer through the serial port. Finally, the host computer resolves the coordinates of the label node through the positioning algorithm and displays it.

Fig. 3 System positioning process flow



5 System Test and Error Analysis

5.1 System Ranging Test

Due to large indoor environment's influence on the signal transmission, will choose for outdoor and indoor two scenarios ranging experiment was carried out, with nodes of the system of ranging performance for a more comprehensive test and evaluation, to verify whether it as a theoretical analysis to obtain the desired range effect.

(1) Outdoor ranging test

The scene of outdoor ranging test is the playground of a college. The playground is more open and can basically meet the distance transmission condition of the signal. It is very suitable for outdoor ranging tests.

Two nodes are arranged in 1 m from ground test point, test between the nodes from 1 m to 70 m range value, thousands of times were carried out at every point of range test, all normal distribution fitting range value, the normal distribution of the mean as the measured distance between nodes [8]. Some measurements are shown in Table 1.

The chart shows that the outdoor environment, with the increase of the actual distance between nodes, the measurement error is fluctuated slightly, but the measuring relative error is reduced and tends to be stable, when the real distance between nodes to more than 15 m, ranging accuracy has reached a higher level.

(2) Indoor ranging test

In this part of the test, the nodes from the ground 1 m test point range, at the same time, starting from node 1 ms apart, every 1 m distance measurement experiment was carried out, thousands of times were carried out the same at every point of range and do normal distribution fitting. In addition, the maximum reliable communication distance between nodes can be obtained at the same time. The measured results are shown in Table 2.

The table above shows that when nodes are placed in from the ground 1 meter test points, the biggest distance between nodes is relatively reliable distance of 13 m, measure the absolute error between the nodes is relatively stable, part of the measured values fluctuated slightly, until the stable, range value in line with the expected effect.

Table 1 Outdoor ranging value

Actual distance (m)	Measuring distance (m)	Measuring error (m)
1	1.2067	0.2067
2	1.3554	-0.6446
4	3.3568	-0.6295
6	5.429	-0.571
...

Table 2 Indoor ranging value

Actual distance (m)	Measuring distance (m)	Measuring error (m)
1	1.2637	0.2637
2	1.8322	-0.1678
3	2.6397	-0.3603
4	3.7441	-0.2559
...

5.2 System Positioning Test

(1) Outdoor positioning test

Outdoor positioning test is in the satellite navigation system and equipment technology national key laboratory in a more open space. The system test platform consists of four anchor nodes, a label node and a host computer, each node uses three AA batteries for power supply. The anchor node coordinates in Table 3.

The outdoor single point positioning is verified and tested by least squares method. In the experiment, hundreds of positioning points were made for each point, and finally the hundreds of results were taken as the single point positioning results of the system. Outdoor single point positioning results shown in Fig. 4.

Table 3 Outdoor anchor node coordinates

Anchor node 1 (m)	(0, 0)
Anchor node 2 (m)	(9.53, 0.42)
Anchor node 3 (m)	(8.9, 11.4)
Anchor node 4 (m)	(1.26, 8.85)

Fig. 4 Outdoor static positioning test results

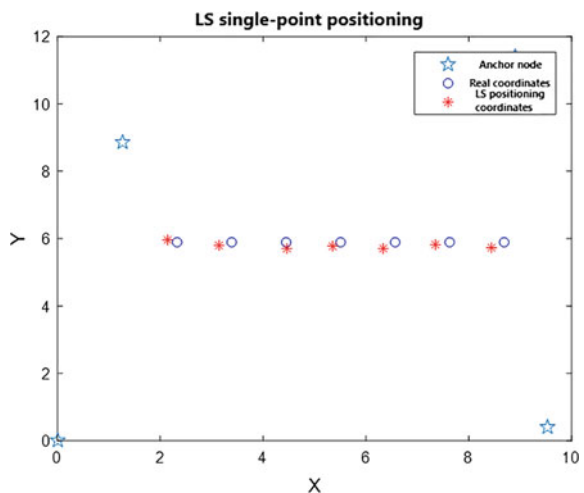


Table 4 Outdoor static positioning test results

Actual coordinate/m	LS positioning results/m	Positioning error/m
(2.32, 5.89)	(2.15, 5.97)	0.19
(3.38, 5.89)	(3.14, 5.81)	0.25
(4.45, 5.89)	(4.47, 5.70)	0.19
(5.50, 5.89)	(5.35, 5.78)	0.19
(6.56, 5.89)	(6.33, 5.70)	0.30
(7.63, 5.89)	(7.35, 5.82)	0.29

True point coordinates, LS positioning coordinates and positioning error as shown in Table 4.

The system of outdoor single point positioning results are more stable, high positioning accuracy, the average single point positioning accuracy of 0.24 m, in line with the expected design requirements.

The outdoor continuous positioning test of the system is also based on the above platform. This system will show the results of the outdoor continuous positioning test of the least squares method and the Kalman filter. The outdoor positioning result of the least squares method is shown in Fig. 5.

Kalman filter outdoor positioning results shown in Fig. 6.

As shown above, the dot is the least squares or Kalman filter positioning trajectory, the line is the real motion trajectory. It can be seen from the experimental results that the dynamic positioning results based on Kalman filter are much better than the least squares method.

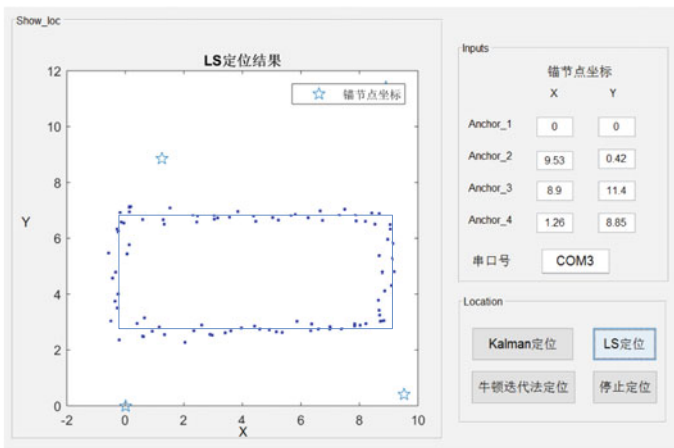


Fig. 5 The least squares outdoor positioning results

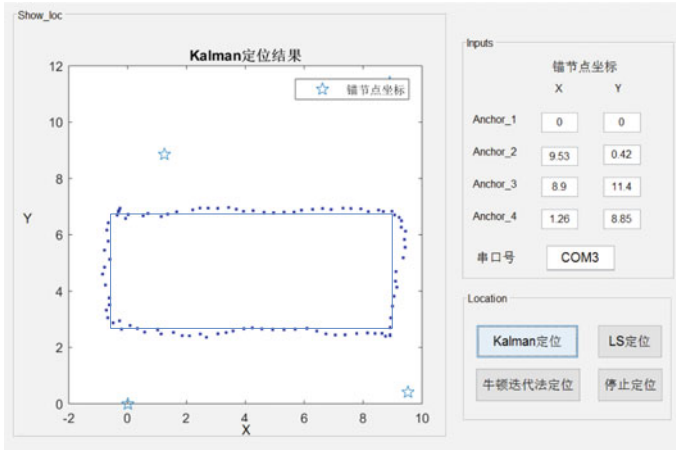


Fig. 6 Kalman filter outdoor positioning results

(2) Indoor Positioning test

Outdoor positioning test is located in the satellite navigation system and equipment technology national key laboratory indoor office area, the coordinates of the system anchor nodes are shown in Table 5.

In this paper, the indoor single-point positioning test is carried out on the built-up system. The same method is used to locate the position in the least squares method. Finally, all the positioning results are taken as the final position result. Indoor single point positioning results shown in Fig. 7.

True point coordinates, LS positioning coordinates and positioning errors are shown in Table 6.

As can be seen from the above table, indoor single point positioning accuracy is relatively low, positioning results fluctuate, not stable enough. The maximum single point positioning error is 1.19 m, the minimum is 0.09 m, the average error is 0.56 m. After the indoor single point positioning test was completed, conducted a continuous indoor positioning test to verify the system positioning effect. As in the case of outdoor continuous positioning tests, this section also makes a comparative analysis of the continuous positioning effect of least squares method and Kalman filter. The indoor positioning results of the least squares method are shown in Figs. 8 and 9.

True trajectory and positioning trajectory. As shown above, the dot is the least squares or Kalman filter positioning trajectory, the arrow straight line for the

Table 5 Indoor anchor node coordinates

Anchor node 1 (m)	(0, 0)
Anchor node 2 (m)	(3.6, 0)
Anchor node 3 (m)	(4.2, 4.2)
Anchor node 4 (m)	(0, 4.2)

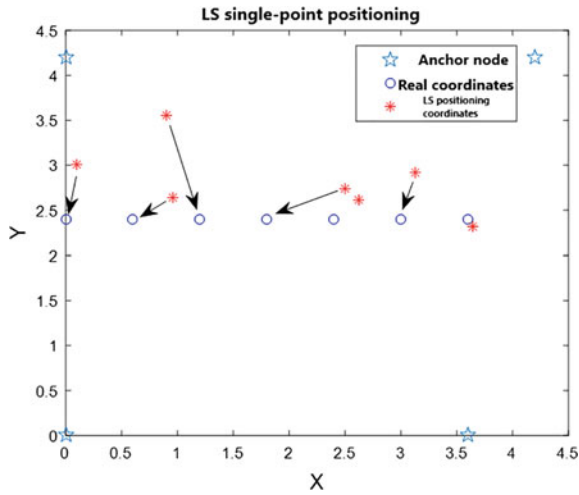


Fig. 7 Indoor static positioning test results

Table 6 Indoor single point positioning error

True point coordinates/m	LS positioning coordinates/m	Positioning error/m
(0.0, 2.4)	(0.10, 3.01)	0.62
(0.6, 2.4)	(0.96, 2.64)	0.43
(1.2, 2.4)	(0.90, 3.55)	1.19
(1.8, 2.4)	(2.50, 2.74)	0.78
(2.4, 2.4)	(2.62, 2.61)	0.30
(3.0, 2.4)	(3.13, 2.92)	0.54

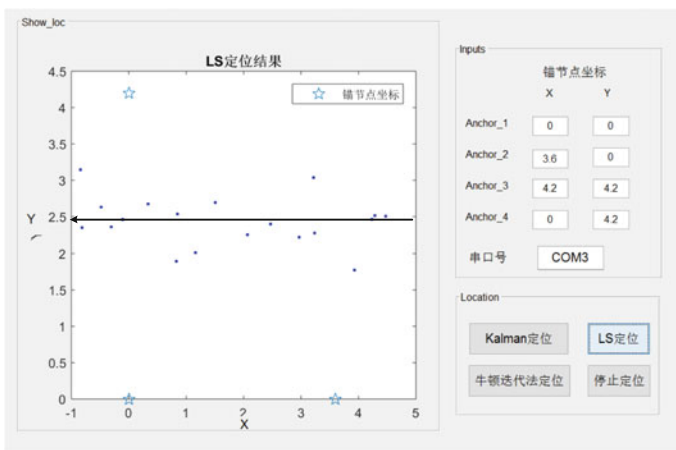


Fig. 8 Least squares indoor positioning results

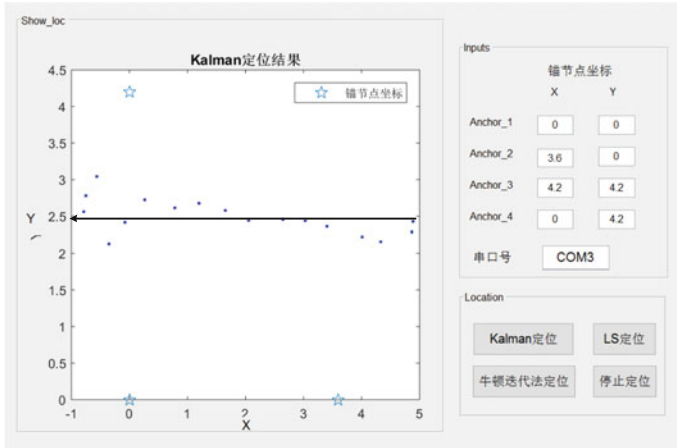


Fig. 9 Kalman filter indoor positioning results

real movement trajectory [9]. It can be seen that compared with the least squares method, the Kalman filter is better for the optimization of the positioning trajectory smoothness, which is consistent with the outdoor continuous positioning results. In addition, due to the complicated environment of the room, Kalman filter positioning accuracy than the least squares method has improved significantly, positioning effect is better [10].

6 Conclusion

For the problem about that most of current wireless positioning system could not coordinate the cost and positioning accuracy, a wireless positioning system, which was utilized the advantages of the hardware cost and Chirp signal characteristic, was designed and implemented. This system was also estimated positioning performance under indoor and outdoor situations. In the outdoor positioning test, the system had a single point average positioning error of 0.24 m. In the dynamic positioning test, it could be seen that the Kalman filter was better than the least squares method. In the indoor positioning experiment, the system had a single point average positioning error of 0.56 m. In the dynamic positioning experiment, it is proved that the Kalman filter could be realized meter-level continuous positioning accuracy. It could be applied in indoor location based service system in future, with a high practical value.

Acknowledgements Found by National key research and development project: 2016YFB0502402.

References

1. Yuhong Ji, Jianghui Wang (2010) Global navigation and positioning system GNSS technology and application. *Glob Position Syst* 5:69–75
2. Wenbo Chang, Mingbo Zhu, Diannong Liang (2002) Study on broadband Chirp signal generation technology. *Sig Process* 18(2):113–117
3. Rohrig C, Muller M (2009) Localization of sensor nodes in a wireless sensor network using the nanoLOC TRX transceiver. In: *Vehicular technology conference, 2009*, pp 1–5
4. Jiachang Gong, Jun Zhang (2011) Nanotron technology in the wireless ranging system application. *Mod Electron Technol* 34(3):56–58
5. Abatzoglou TJ, Mendel JM, Harada GA (1991) The constrained total least squares technique and its applications to harmonic super resolution. *IEEE Trans Sig Process* 39(5):1070–1087
6. Ribeiro A, Schizas ID, Roulmliotis S et al (2010) Kalman filtering in wireless sensor networks. *Control Syst IEEE* 30(2):66–86
7. Oh D, Kwak M, Chong JW (2012) A subspace-based two-way ranging system using a Chirp spread spectrum modem, Robust to frequency offset. *IEEE Trans Wirel Commun* 11(4):1478–1487
8. Xu Y, Yi W (2010) Study on the influence of ground on the characteristics of wireless antenna node antenna. *J Comput Res Dev* 47(Suppl.):16–21
9. Sheng C, Zhang J GPS real time positioning based on Kalman filter. *Radion Eng.* 46(6):41–44
10. Li S, Deng H (2015) Two-station DOA/TDOA joint location algorithm based on EKF. *Radio Commun Technol* 41(1):38–40

Research on Smoothing Filtering Algorithm of BDS/GPS Slow Deformation Monitoring Sequence



Lei Ren and Tian-He Xu

Abstract In order to improve the accuracy of slow deformation monitoring sequence, the smoothing filter is usually used. On the basis of weighted moving average method, a robust moving average method is proposed. An adaptive windowing method is developed and combined with ARMA time series analysis to improve the prediction precision of monitoring sequence. The experimental results show that, the robust moving average method can effectively weaken the error of the slow deformation monitoring sequence, and resist gross error efficiently. In addition, it can improve the accuracy and reflect the displacement trend of the slow deformation monitoring sequence more clearly with combination of ARMA method.

Keywords Slow deformation monitoring sequence · Smoothing filtering
Robust moving average · Time series

1 Introduction

Nowadays, the Global Navigation Satellite System (GNSS) is developing rapidly under the situation of multiple competition and complementary advantages [1]. Since China's BDS system is put into use, BDS and other navigation systems combined positioning technology become a hot in multi-GNSS applications [2, 3]. At present, BDS/GPS combined deformation monitoring technology has been widely used in dam, landslide, bridge, slope and other engineering deformation monitoring fields. Multi-GNSS positioning technology is often affected by the

L. Ren

School of Geology Engineering and Surveying, Chang'an University, Xi'an, China
e-mail: renchd009@163.com

T.-H. Xu (✉)

Institute of Space Science, Shandong University, Weihai, Shandong, China
e-mail: thxugfz@163.com

T.-H. Xu

State Key Laboratory of Geo-Information Engineering, Xi'an, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_3

external environment, multipath effect, atmospheric delay, receiver noise and other hardware and software influence, so that the monitoring sequence usually contains noise and outliers [4]. In order to describe the displacement of the monitoring body more accurately and predict its deformation tendency, it is necessary to smooth the monitoring sequence.

Kalman filtering is usually adopted to describe the dynamic deformation process of a monitoring body based on the state and observation equation. Using the state transition equation of system, and the predicted state according to the previous value can be obtained and the estimated state at the current moment can be calculated by principle of least squares (LS). It is often used in the observation data filtering and adjustment stage, the calculation principle is complex and requires a large amount of storage space [5, 6]. Wavelet analysis method uses a wavelet transform to extract useful deformation from local signals in time domain and frequency domain. It uses discrete wavelet function to carry out wavelet decomposition and reconstruction of the observed deformation data, and eliminates errors to obtain effective deformation characteristics and deformation trend [7]. In wavelet analysis, the choice of wave function, decomposition level and the choice of denoising methods often depend on experience. The moving average method only takes advantage of the attribute in the filtering process, which has low space complexity and time complexity [8]. Therefore it is widely used in rapid processing of deformation monitoring sequences. The conventional weighted moving average method is difficult to eliminate or weaken the influence of outliers, the moving window and weight selection still rely on experience. For this reason, this paper proposes a weighting method based on the correlation equivalent weight. It chooses the window size based on the sequence trend and the correlation between the before and after filtering, so as to adapt to the slow deformation monitoring sequence filtering under various conditions.

2 Smoothing Filtering Algorithm for Slow Deformation Monitoring Sequence

2.1 *BDS/GPS Combined Slow Deformation Monitoring Sequence*

The slow deformation monitoring sequence refers to the wide stationary sequence, whose characteristic is that the statistics of the sequence does not change with time translation, the mean and covariance do not change with the time shift, and the trend of the monitoring sequence changes little. In the monitoring time range, the sequence displacement has no sudden change and the variance of random change is small. It is reasonable to believe that the current evolution trend will continue to extend into the future, and can be assumed that the monitoring sequence is stable and slow deformation.

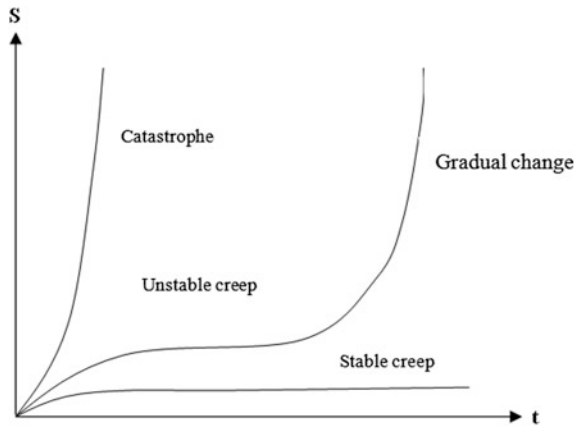


Fig. 1 Three types of time displacement curves

Geological hazard evolution is a cumulative damage with timeliness. According to the relationship between monitoring body movement and time form, geological disasters can be divided into three types: catastrophe, gradual change and stable. The gradual change can be divided into three stages: initial deformation, constant deformation and accelerated deformation (Fig. 1).

Taking the landslide disaster as an example, stable creep landslide is mainly developed in the relatively perfect geological mechanism slope. Due to some external factors, the landslide will produce deformation, but with the weakening or disappearance of the external factors, the slope will regain its stability under the influence of self weight. For the monitoring of geological disasters, the most important thing is to analyze and predict the position shift trend, so as to find disaster problems in time and reduce casualties and property losses. Therefore, the real-time monitoring and prediction of stable creep disaster is very important. It is believed that the time series monitored by the steady creep and the gradual deformation stage are the slow deformation monitoring sequences.

2.2 Robust Moving Average Method

For the structure of slow deformation geological hazard, it is considered that the change of displacement in each direction is continuous and smooth in a certain range. Therefore, the residual sequence of GNSS position in the small area mainly includes deformation tendency and high-frequency random noise. In order to analyze the deformation characteristics and strain trend of the monitoring body, it is necessary to adaptively eliminate the high frequency random noise in the sequence.

The weighted moving average model is commonly used to assign different weights in the observed value of different epoch [6]. It relies on the principle of the nearest point with the greatest impact, but the weight distribution mainly according

to the actual situation and experience with no theoretical formula of support. On the basis of the ordinary weighted moving average model, this paper improves the weighted strategy, which can not only restrain the model deviation, but resist the gross error. The improved robust moving average model can be expressed as follows:

$$\hat{y}_n = \frac{\sum_{j=0}^{m-1} \omega_j y_{n-j}}{\sum_{j=0}^{m-1} \omega_j} \quad (1)$$

where \hat{y}_n is the moving average of n data points; ω_j is the weight corresponding to the j data points; m is the window size in the robust moving average model.

The specific weight function can be expressed as [9, 10]:

$$\omega_i = \begin{cases} p_i, & |v_i| < k_0 \\ p_i \cdot \frac{k_0 (k_1 - |v_i|)^2}{|v_i| (k_1 - k_0)^2}, & k_0 < |v_i| < \omega_n S_x \\ 0, & |v_i| \geq \omega_n S_x \end{cases} \quad (2)$$

In this paper, the weight function is chosen based on IGGIII equivalent weight function, in which p_i is the original weight; k_0 and k_1 are empirical values; after a large number of experiments, k_0 generally takes 1–1.5; k_1 takes 2.5–3. It is believed that in the process of filtering epoch weights should contain three segments: When the epoch error is very small, the equivalent weight is the original weight. When the error is large but not significant, the smoothing factor of down-weighting is given according to the credibility. When the error is significantly abnormal, its weight is set to zero to eliminate it. In the process of smoothing filtering, the slow deformation sequence is divided into three segments: natural segment, suspicious segment, and eliminated segment.

Gross error can be regarded as the overall performance from other sample distribution in the sequence, thus eliminating it in order to ensure the reliability of filtering model [11]. Gross error detection usually uses three times standard deviation method, but for BDS/GPS combined slow deformation monitoring sequence, the sequence itself has tendency and periodic terms, so it is necessary to relax the tolerance limit of gross error. Chauvenet criterion is a method of mathematics, and its error tolerance is associated with sequence length, which can better estimate the intrinsic attribute of sequence itself and periodic trend etc.

Chauvenet criterion is as follows:

$$|v_i| > C = \omega_n S_x \quad (3)$$

and $\omega_n = 1 + 0.4 \ln(n)$, $S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$, where $|v_i|$ is the absolute value of epoch error, ω_n is Chauvenet coefficient, S_x is standard deviation, n is epoch number.

If the deviation after two epochs were less than $C/2$, or it is greater than the threshold value set by the Chauvenet criterion and on the same side with the current epoch value, it is judged that the epoch of mutation. Although the tolerance limit is relaxed in the process of gross error detection, the sequence suspicious segment is reduced to the weight factor in the selection process, so the anomaly of gross error detection can be eliminated.

The influence function is used to judge the sensitivity of estimated value to the outliers. It reflects the relative effect of the size of abnormal data at different reference epoch of valuation [12]. In this paper, the influence function is taken as an index to measure robustness of the model.

The influence function can be expressed as:

$$IF(L, F, \hat{X}) = \frac{\hat{X}(F_\varepsilon) - \hat{X}(F)}{s/n} \quad (4)$$

where L is the observation vector, F is the joint distribution, F_ε is the pollution distribution, s is the number of gross error, and n is the sample capacity.

2.3 Sequence Trend Analysis and Adaptive Windowing Filter

The choice of moving window m is the key to the weighted moving average model, and plays a decisive role in smoothing filter [13]. The larger the moving window is, the smoother the filtered sequence is, and the larger window width also makes filtered sequence produce a large phase lag. When the filter window is too large, the sequence will appear too smooth, so that the smoothing sequence can't truly reflect the situation of monitoring body displacement. When the filter window is small, it can ensure the real displacement of the sequence, but it may be too sensitive to the random influence, and can't filter out the non-stationary white noise sequence. Therefore, there is a close relationship between the appropriate moving window and the smoothing filter of the monitoring sequence.

In data processing strategy, the size of the moving window should be between 5 and 13. In order to reduce the influence of normalized moving average on the smoothing period, the window of the general moving average method takes odd number [13, 14]. On this basis, the optimal window is selected by iterating the correlation between the sequence trend and the model before and after filtering.

Curve regression is a commonly used method in trend analysis, and its trend model can be expressed as [15]:

$$Y_t = f(t, \theta) + \varepsilon_t \quad (5)$$

where Y_t is the model prediction object; ε_t is the prediction error; $f(t, \theta)$ is a different curve function model. Because the BDS/GPS monitoring method is less

affected by the external system error than other methods, the sequence fluctuation is relatively stable, so the monitoring sequence is usually analyzed by polynomial trend model. It can be expressed as:

$$y_t = a_0 + a_1t + a_2t^2 + \cdots + a_it^i + \varepsilon_t \quad (6)$$

The root mean square error is:

$$S = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

In order to determine the trend model order of BDS/GPS combined monitoring sequence, the least square error is selected by iterative method, and the highest order is less than 6.

By comparing the Pearson correlation coefficients of the trend model before and after filtering in different windows, the optimal filtering window is selected iteratively in the window range.

The Pearson correlation coefficient can be expressed as:

$$\rho'_{R,S} = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}} \quad (8)$$

where $\bar{R} = \frac{1}{n} \sum_{i=1}^n R_i$, $\bar{S} = \frac{1}{n} \sum_{i=1}^n S_i$, $\rho'_{R,S}$ is the correlation coefficient under t window, R represents the function value of the trend model before filtering, and S represents the function value of the filtered trend model.

When choosing the optimal window, we must consider the two basic principles that the filtered sequence is strongly correlated with the original sequence and the precision is improved. Therefore, Pearson correlation coefficient and sequence standard deviation are the criteria to evaluate the performance and stability of the algorithm in the process of algorithm verification. The following functions are introduced to quantitatively evaluate whether the window is optimal.

$$\varphi_t = \left(1 - \rho'_{R,S}\right) \times \ln(std_S) \quad (9)$$

where std_S is the standard deviation of model function value after filtering, and φ_t is defined as the optimal window function value. We know that the larger the correlation coefficient is, the stronger the sequence correlation is; and the smaller the standard deviation is, the higher the sequence precision is. So the smaller the

optimal function value of the window, the better the filtering effect of the robust moving average model.

3 Example Analysis

In this paper, the algorithm of smoothing filtering is verified by using the data of GNSS monitoring network in Ningxia. BDS/GPS relative positioning solution using single epoch ambiguity fixing method based on the experimental data. Precise clock and orbit products provided by GBM analysis center is used, and EKF method is adopted to estimate deformation parameters.

Experiment selected monitoring network in the direction of the trend evident in the station were analyzed, and the output data calculating results every 3 minutes and the total output epoch number is 3811, we get the specific sequence of monitoring stations of the diagram as following figures.

In Fig. 2, the trend of each direction sequence is stable, and there is no abnormal mutation, which can well respond to the statistical characteristics of slow deformation monitoring sequence. Due to the multipath effect, there exists a daily periodic term in each direction, and contains some white noise and gross errors, This is because that some epoch ambiguity is not completely fixed, only to get the float solution or single point solution. Before smoothing the sequence, the BDS/GPS system is used to monitor the mean value of the first day as the initial reference. After subtracting the datum from the whole 8 days, the residual sequence of coordinates is obtained, and is used as the data sequence for smoothing filtering.

Figure 3 is the sequence after filtering by the robust moving average method. It can be seen that the filtering sequence has a significant improvement in accuracy compared with the original mean sequence. The gross error of the sequence is basically filtered out, it has a good ability to resist gross error, and does not excessively smooth out the original real mutation signal. On the basis of improving the accuracy of the BDS/GPS combination monitoring sequence, the deformation tendency of the filtered sequence is consistent with the original sequence.

Table 1 shows the robust moving average window optimal value in each direction of the window range. Since the direction of the original sequence and the distribution of coarse and white noise are different, the different window sizes should be adopted in the smooth filtering in each direction. In this experiment, it can be seen that the window size of the E direction should be 9, the N direction be 5, and the U direction be 11. It shows that the method proposed in this paper can be adaptive to determined the optimal window size.

In order to verify the correctness and efficiency of the robust moving average method proposed in this paper, the different moving average method is used to compare the filtering effect of the original sequence. Figure 4 shows only the E direction, and uses simple moving average method, weighted moving average method and robust moving average method to smooth filtering under the same window condition.

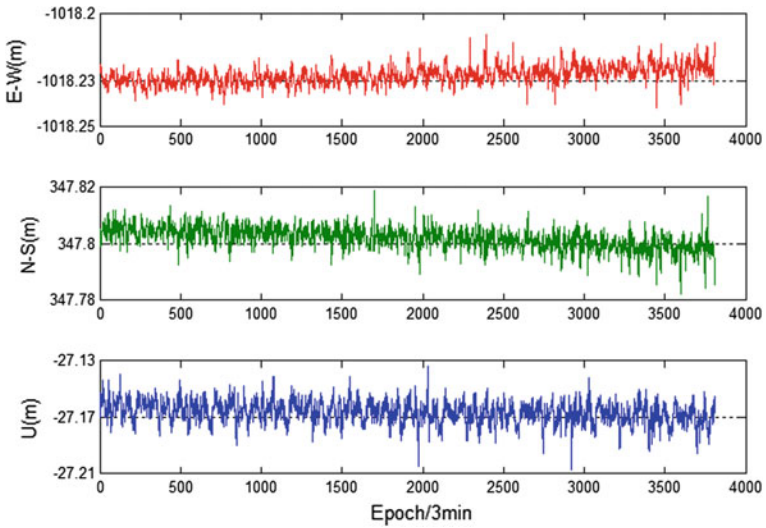


Fig. 2 Original monitoring sequence of BDS/GPS combination

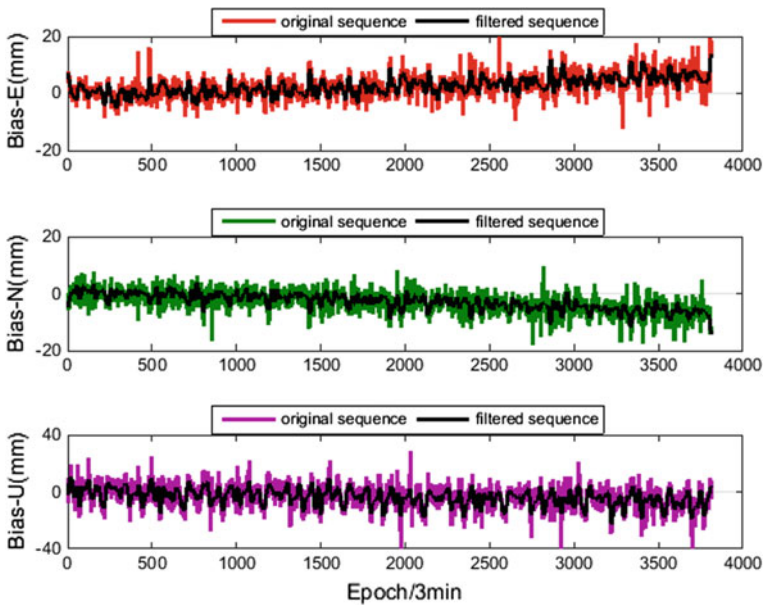


Fig. 3 Filtering sequence after robust moving average method

It can be seen from Fig. 4 that different moving average methods have a certain degree of smoothness to the original sequence, in which the smoothing effect of the robust moving average method is most prominent.

Table 1 Statistics of optimal values in each direction window

Window size	E direction window optimal value	N direction window optimal value	U direction window optimal value
5	0.667	0.073	1.401
7	0.173	0.115	0.169
9	0.116	0.124	0.189
11	0.134	0.135	0.116
13	0.163	0.142	0.209

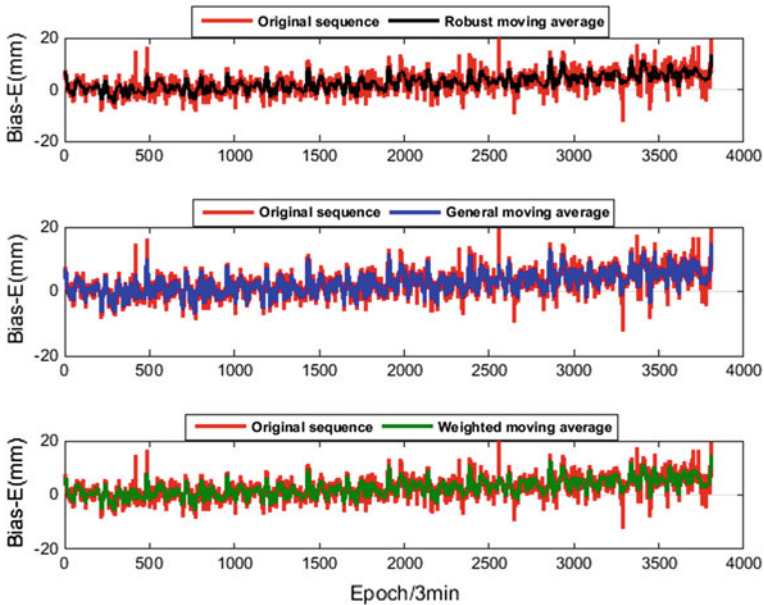


Fig. 4 Comparison of processing results with different filtering methods

Table 2 shows the standard deviation of the filtered sequences. It can be seen that the robust moving average method can improve the sequence accuracy better than the other two methods under the premise of filtering the same window, especially in the case of the large difference of U direction. The robust moving average method presented in this paper shows best ability to resist gross error.

To further verify the feasibility of this method, we propose a new combined forecasting model based on the ARMA model of time series analysis, The robust moving average method proposed in this paper is combined with the ARMA model to improve the accuracy of monitoring sequence prediction.

The monitoring sequence of the results of every 3 h in the N direction of the monitoring point is used as the experimental data. Each epoch monitoring records for a data value, and the total 64 period data is recorded. We use the pre-60 groups

Table 2 Accuracy statistics

	Original sequence STD (mm)	Robust moving average method STD (mm)	General moving average method STD (mm)	Weighted moving average method STD (mm)
E	3.71	3.01	3.30	3.13
N	3.65	2.77	3.12	2.92
U	7.78	5.84	6.74	6.29

of data to calculate and the post-4 groups of data is used as a reference for the forecast results.

It can be seen from the Fig. 5 that the fitting sequence of the two models is basically consistent with the measured sequence. The fitting results of the combined model are closer to the measured data. The robust moving average method can describe the deformation characteristics and the deformation trend more accurately. Therefore, the combined model can effectively deal with the BDS/GPS dual system combined deformation monitoring sequence, and eliminate the gross error to obtain higher fitting and prediction accuracy.

It can be seen from the Table 3 that the combined model has more obvious improvement than the ARMA model in prediction accuracy. However, with the increase of the forecast time, the historical data on the forecast is decreasing, resulting in the prediction accuracy gradually reduced. However, it can still ensure the accuracy of prediction and maintain the same order of accuracy magnitude.

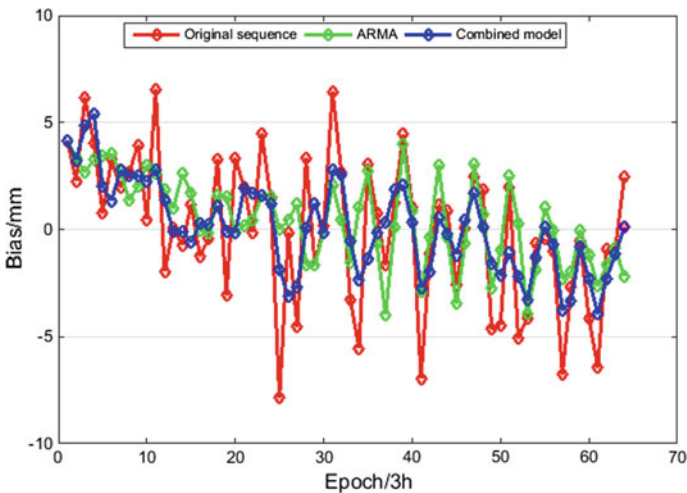


Fig. 5 Fitting data results of different models

Table 3 Comparison of prediction results

Number of periods	Measured value (mm)	Prediction value of ARMA model (mm)	Prediction residual (mm)	Prediction value of combined model (mm)	Prediction residual of combined model (mm)
61	-6.05	-3.63	-2.42	-4.93	-1.12
62	-0.95	-1.63	0.68	-1.37	0.42
63	-0.85	-1.12	0.27	-0.86	0.01
64	2.44	-2.20	4.64	0.29	2.15

4 Conclusions

The robust moving average method has a good ability to resist gross error when smoothing the BDS/GPS slow deformation monitoring sequence, and does not filter out the original mutation signal. It can improve accuracy of the filtered sequence and be adapted to different slow deformation monitoring sequences. The combination of robust moving average method and ARMA time series can improve the accuracy of prediction in short term. But the combination method is only suitable for the smoothing filter of slow deformation monitoring sequence, and the filtering effect is not ideal for the sequence with abnormal mutation and large fluctuation of monitoring. How to make robust moving average method better applied to various monitoring sequence filtering should be studied further.

Acknowledgements The research was supported by National Natural Science Foundation of China (Grant No. 41574013, 41731069) and National Key Research and Development Program of China (2016YFB0501701).

References

1. Tu R, Liu J, Lu C, Zhang R, Zhang P, Lu X (2017) Cooperating the BDS, GPS, GLONASS and strong motion observations for real-time deformation monitoring. *Geophys J Inter* 209 (3):1408–1417
2. Odolinski R, Teunissen PJG (2016) Single-frequency, dual-GNSS versus dual-frequency, single-GNSS: a low-cost and high-grade receivers GPS-BDS RTK analysis. *J Geodesy* 90 (11):1255–1278
3. Lou Y, Gong X, Gu S et al (2016) GPS/BDS mixed double differential RTK positioning method and result analysis. *J Geodesy Geodyn* 36(1):1–5 (in Chinese)
4. Li G, Dai W, Zeng F (2016) Robust moving average in GPS automatic monitoring data processing applications. *J Geodesy Geodyn* 36(1):85–88 (in Chinese)
5. Wu X (2011) Research on Kalman filtering method in dynamic deformation monitoring. Central South University (in Chinese)
6. Zhao L, Qiu H, Feng Y (2016) Analysis of a robust Kalman filter in loosely coupled GPS/INS navigation system. *Measurement* 80:138–147
7. Dong D, Wang M, Chen W et al (2016) Mitigation of multipath effect in GNSS short baseline positioning by the multipath hemispherical map. *J Geodesy* 90(3):255–262

8. Zhang J, Zhou DM (2013) Improved moving average filter for GPS disciplined. *Acta Electron Sin* 41(2):412–416
9. Yang Y (1994) Principle of equivalent weight—robust least square solution of parameter adjustment model. *Surv Mapp Bull* 6:33–35+29 (in Chinese)
10. Awange JL, Paláncz B (2016) Robust estimation. *Geospatial algebraic computations*. Springer International Publishing
11. Yang Y (1994) Anomaly influence diagnosis and robust estimation. *Surv Mapp Bull* 5:34–36 (in Chinese)
12. Zhang Q (2011) Processing and application of modern surveying data. *Surveying and Mapping Press* (in Chinese)
13. Chang BW, Gao L, Fang J et al (2014) High precision weighing improved limiting average filtering method. *Sens Technol Based Syst Res* 5:649–653 (in Chinese)
14. Sokhandan N, Curran JT, Broumandan A et al (2016) An advanced GNSS code multipath detection and estimation algorithm. *GPS Solutions* 20(4):627–640
15. Huang S (2010) Deformation monitoring data processing. *Wuhan University Press* (in Chinese)

Research on Space-Based Measurement and Control Scheme of Launch Vehicle Based on BeiDou Navigation Satellite System



Chong Li, Xianqing Yi, Yue Zhao and Zhenwei Hou

Abstract In order to solve the problem of remote monitoring and control of launch vehicles, this paper proposes that BeiDou navigation satellite system can achieve the real-time measurement and control of the launch vehicle, based on the analysis of the support of the space-based measurement and control system for the launch vehicle measurement and control tasks. And then proposes space-based measurement and control system schemes. According to the composing and task supporting ability of BeiDou navigation satellite system, the communication process of BeiDou navigation satellite system supporting the global launch vehicle space-based measurement and control is analyzed and applied. Then, according to the BeiDou global navigation satellite system, three kinds of constellation support communication schemes are designed and three networking constellations are proposed to support the space-based measurement and control tasks. Finally, the feasibility of the communication scheme is verified through simulation experiments, which provides a better solution for the space-based measurement and control system of the launch vehicle.

Keywords BeiDou navigation satellite system · Launch vehicle
Communication scheme · Space-based measurement and control

1 Introduction

In November 2017, China successfully launched the 24th and 25th BeiDou Navigation Satellites in the Xichang Satellite Launch Center, the first and second satellite of China's BeiDou III, which opened the global satellite navigation system of BeiDou a new era of networking. The increasingly perfect BeiDou satellite navigation system has provided strong technical support for China's military,

C. Li (✉) · X. Yi · Y. Zhao · Z. Hou

Science and Technology on Information Systems Engineering Laboratory,
National University of Defense Technology, Changsha 410073, China
e-mail: lichong16a@nudt.edu.cn

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_4

economy, agriculture and transportation, and also has a huge application space in space-based measurement and control of the launch vehicle.

At this stage, the measurement and control tasks of China's launch vehicles mainly rely on the traditional world-wide direct transmission and preliminary relay satellites space-based measurement and control methods, which require the launching of multiple relay satellites and setting up several ground stations in China and arranging several ocean measurement ships at sea, there is the problem of high maintenance cost, incomplete measurement coverage, low security and reliability. The BeiDou navigation satellite system is about to realize a global networking. Based on the existing BeiDou satellite resources, the launch vehicle measurement and control information relay function can not only expand the scope of measurement and control, greatly improve the real-time data transmission and monitoring data transmission reliability, but also can make full use of existing navigation satellite resources, with the least cost to obtain greater social and economic benefits.

This paper first analyses the status quo and mission requirements of the launch vehicle space-based measurement and control, and then analyses the support capability of the BeiDou system to the launch vehicle. Next, the corresponding measurement and control communication schemes of geostationary earth orbit satellites (GEO), inclined geosynchronous satellite orbit satellites (IGSO) and medium earth orbit (MEO) are proposed for different launching trajectories of the launch vehicle. Finally, STK simulation analysis of the three measurement and control communication scheme feasibility. Based on the requirements of the launch vehicle measurement and control and the actual situation of the BeiDou constellation, this paper presents a space-based measurement and control plan for the launch vehicle based on the BeiDou navigation satellite system. This paper presents a feasible plan for the wide range of applications of BeiDou navigation satellite system and the expansion of the launch vehicle space-based measurement and control.

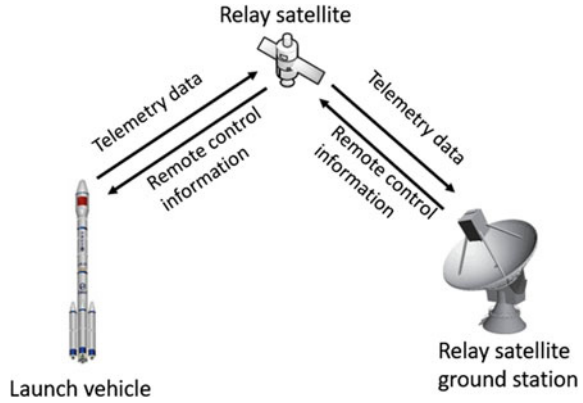
2 The Status Quo of Launch Vehicle Space-Based Measurement and Control and BeiDou System Support Ability Analysis

2.1 The Status Quo and Problem of Space-Based Measurement and Control of Launch Vehicle

The launch vehicle space-based measurement and control is a mode in which the launch vehicle relays transmission data through relay satellites during flight and communicates with the ground station. The principle diagram of the measurement and control process is shown in Fig. 1.

China's space measurement and control system can be divided into two parts of the foundation system and space-based system [1]. Measurement and control of the

Fig. 1 Traditional space-based measurement and control process schematic diagram



ground part of the launch vehicle is relatively mature, but there are also various problems. With the rapid development of the launch vehicle in our country, higher requirements have been put forward on the measurement and control system: to improve coverage adaptability and to control more diversified launch orbit; the external environment is increasingly complicated and the flight safety needs to be controlled more; the mission is more complex and need more measuring and control protection; Therefore, China now vigorously develop the launch vehicle space-based measurement and control technology to improve the real-time coverage and data transmission and improve the reliability of data transmission. At present, China is researching the key technologies of the launch vehicle measuring and control data transmission based on the Ka-band relay satellites, as well as carrying out research on the key technologies of forward-based measurement and control based on the relay satellite system to improve the coverage [2].

2.2 *Launch Vehicle Space-Based Measurement and Control Mission Needs*

The launch vehicle space-based measurement and control system relies mainly on the ground station and the measured launch vehicle established between the ground link and the inter-satellite link to achieve the measurement and control tasks. Therefore, the following capabilities are proposed for the satellite and inter-satellite links:

1. We must be able to reasonably choose relay satellites to meet the requirements of building visual links to relay and forward launch vehicle space-based measurement and control system information.
2. Such links must be established within the shortest time needed to build a chain, that is, proper linkages and link-building mechanisms should be established to meet the real-time measurement and control requirements of the launch vehicle.

3. Measurement and control process established in the ground, the inter-satellite link, in the performance indicators to meet the launch vehicle space-based measurement and control requirements, that is, the link can be relayed.
4. In the signal mechanism and protocol specification based on established inter-satellite and satellite link transmission information, the requirements of mutual compatibility and even consistency must be satisfied.

In order to complete the measurement and control tasks, the space system needs to provide the function of relay and forward when the launch vehicle go beyond the domestic visible space. The specific functions of the space system include: the ability to build full-air-space links together with the launch vehicle; the ability to build a timely and flexible chain with the launch vehicle; the signal system and protocol specification compatible or consistent with the space-based measurement and control system; ability to support the launch vehicle space-based measurement and control system coordination and task coordination.

Through the research and analysis, the BeiDou navigation system currently under construction in our country can meet the global relaying and forwarding requirements required by the space-based measurement and control system of launch vehicle and already possess the above-mentioned various requirements. The specificities of the BeiDou system's mission support capability is given below [3].

2.3 Capability Analysis of BeiDou System for Launch Vehicle Measurement and Control

The BeiDou navigation satellite system includes 32 satellites, including 5GEO satellites, 3IGSO satellites and 24MEO satellites and three ground stations, including Beijing as the main control station and injection stations as Sanya and Kashi stations. According to BeiDou constellation construction needs, BeiDou can provide the following support for the launch vehicle space-based measurement and control:

1. BeiDou system will achieve global coverage, the establishment of BeiDou global system around 2020, having the ability to measure and control launch vehicle full-time and entire airspace.
2. BeiDou system can receive, store, generate, broadcast navigation messages and can receive, compare, return precision time measurement, with the ability to provide a stable transmission of information.
3. System communication protocol to follow the CCSDS architecture, interface specification compatibility. Compatibility or signal system and protocol specification is good.
4. BeiDou system has the ability of autonomous operation of the constellation, the orbital reconstruction capabilities in the orbital as well as strong anti-interference ability, with the launch vehicle space-based measurement and control system in a timely manner, the ability to build a flexible chain.

Based on the above analysis, BeiDou system has the capability of supporting launch vehicle space-based measurement and control tasks to meet the communication requirements of mission objectives for the relay system. Due to the short duration of the launch vehicle measurement and control, the existing business impact on the BeiDou satellite system is relatively small. The BeiDou III is under construction and inter-satellite laser link may be used in it [4]. The information transmission rate is increased, and the mission support capability is enhanced. There are no theoretical problems with the BeiDou system supporting the launch vehicle's space-based measurement and control. Therefore, BeiDou navigation satellite system can be used as a relay system to complete the launch vehicle space-based measurement and control tasks.

3 Launch Vehicle Launch Orbit and Invisible Area Boundary Calculation

3.1 Launch Vehicle Launch Orbit

The launch vehicle launches from the launcher until the spacecraft into the intended orbit of the flight path known as launch vehicle launch trajectory. There are two basic forms of launch trajectory: one is the continuous thrust launch trajectory; the other is the launch trajectory with the middle orbit. Figure 2a shows a continuous thrust launch track, in the figure K0 for the launch point. K1 into the track point, the red arc as the launch trajectory. Figure 2b shows a launch trajectory with a middle orbit. As shown in the figure, K1 is located near the mid-orbit and K2 is located at the apogee of the mid-orbit, which is called the best transition.

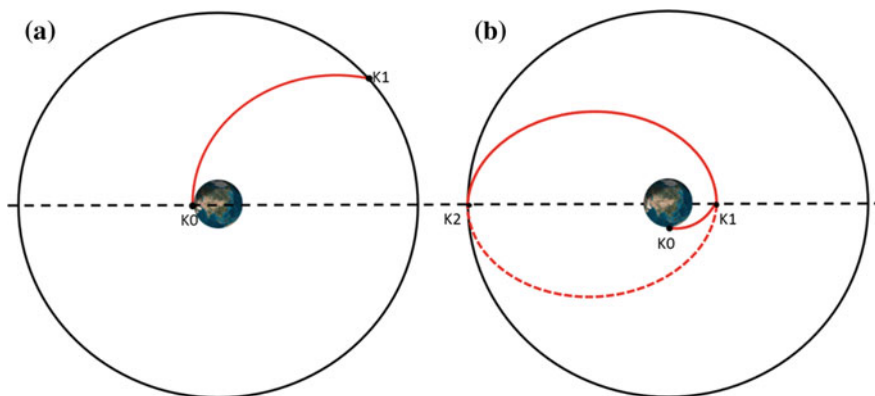


Fig. 2 Launch vehicle launch trajectory

3.2 Invisible Area Boundary Calculation

According to the capabilities analysis of the BeiDou system for the measurement and control of launch vehicles in the second section, we can see that there are three existing ground stations in China. Beijing is in the easternmost point of China, while Kashi the westernmost. It has the largest range over the longitude range.

Figure 3 shows the launching orbit of any launch vehicle (indicated by a thick red line in the figure) as an example, the longitude of point A and point B of the critical point of the range cannot be directly measured when calculating any point M on its flight to orbit.

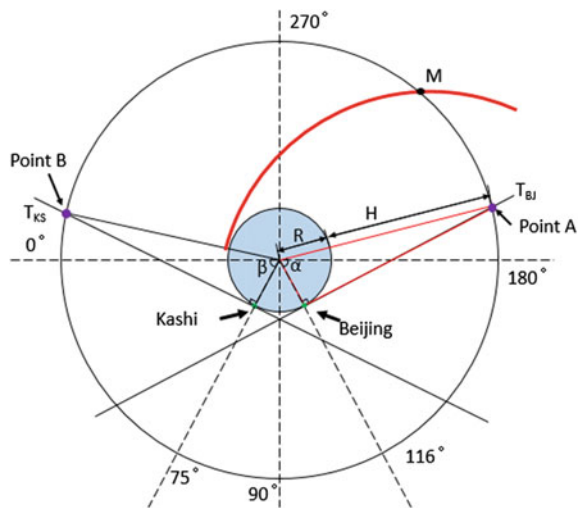
The symbols are as follows:

1. M point off the ground height H.
2. Earth's radius is R.
3. T_{BJ} and T_{KS} respectively denote the tangents to the Earth's surface at the point of Beijing and Kashi Station.
4. The blue circle in the picture shows the earth, the black circle on the outside shows a circle centered at the center of the earth and a radius of (H + R).
5. α represents the angle between the longitude of Beijing station and the point A, β is the same.
6. The Lon_{BJ} and Lon_{KS} refer to the longitude of Beijing and Kashi stations.

By the geometric relationship, A point where the longitude Lon_A is

$$\begin{aligned}
 Lon_A &= Lon_{BJ} + \alpha \\
 &= Lon_{BJ} + \arccos \frac{R}{H + R}
 \end{aligned}$$

Fig. 3 Can not directly measure the area calculation diagram



Similarly, the longitude at which point B is available is:

$$\begin{aligned} Lon_B &= Lon_{KS} - \beta + 360^\circ \\ &= Lon_{KS} - \arccos \frac{R}{H+R} + 360^\circ \end{aligned}$$

4 Measurement and Control Communication System of Launch Vehicle Based on BeiDou Satellite Navigation System Constellation

4.1 GEO Constellation Support Launch Vehicle Control Communications Scheme

BeiDou Global System GEO Constellation mission support scheme to guide the satellites GEO1 (80°E) and GEO3 (140°E) as a space-based measurement and control system in the ground station invisible target of airspace measurement and control information relay channel. The navigation satellite GEO1 undertakes the relay forwarding of the space-based measurement and control targets in the space of $290^\circ \sim Lon_B$ (cone central angle). The navigation satellite GEO3 undertakes the relay forwarding of the space-based measurement and control targets in the $Lon_A \sim 290^\circ$ (cone central angle). Because 290° the other side of the earth at 110° longitude between GEO1 and GEO3, it's the dividing line. Other airspace is the visible airspace within the territory of the ground station, from the station Kashi station or Beijing station directly landing.

When the launch vehicle is in the invisible area of domestic station, the BeiDou satellite navigation system relays satellite measurement and control information, the scheme shown in Fig. 4a, b.

1. When the inflight launch vehicle is located in the visible area ($Lon_B \sim 360^\circ (0^\circ) \sim Lon_A$ in the domestic station, orange area in the figure), it directly transmits the measurement and control information to the domestic stations.
2. When the inflight launch vehicle is in the area not directly controllable ($Lon_A \sim Lon_B$ in the green and blue areas in the figure), the measurement and control information need to be forwarded to the ground stations in the GEO1 or GEO3 for forwarding Launch vehicle real-time monitoring. When the carrier launch vehicle flies to $Lon_A - 290^\circ$ (blue area in Fig. 4a, b), the measurement and control information passes the measurement and control information via the route of the launch vehicle \rightarrow GEO3 \rightarrow domestic station to realize the indirect measurement and control of the launch vehicle. Similarly, when the launch vehicle is flying at $290^\circ \sim Lon_B$ (green area in Fig. 4a, b), the measurement and control information is transmitted via launch vehicle \rightarrow GEO1 \rightarrow domestic station.

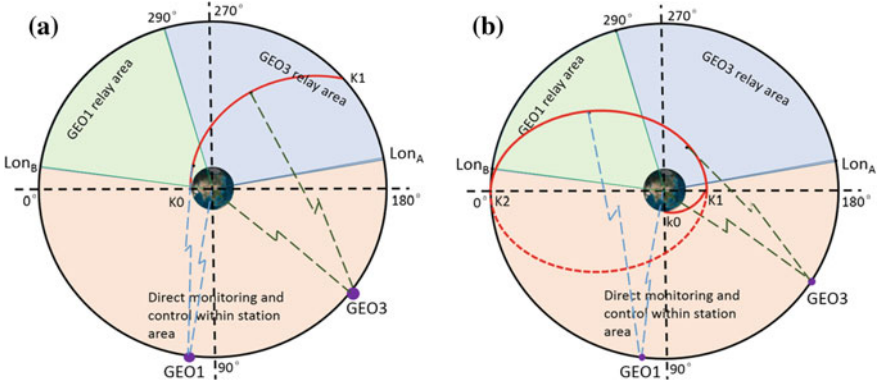


Fig. 4 GEO communication diagram

4.2 IGSO Constellation Support Launch Vehicle Control Communications Scheme

Aiming at the invisible range of $Lon_A \sim Lon_B$ (cone center angle), three satellites in the IGSO constellation of BeiDou system are used as the relay and forwarding channel of space-based measurement and control information.

Based on the orbital and operational characteristics of the IGSO satellite in the IGSO constellation, the IGSO constellation is most unfavourable when the launch vehicle is in $298^\circ E$, the other side of the Earth with longitude $118^\circ E$ of the IGSO constellation. The visual relationship of IGSO1, IGSO2, IGSO3 and launch vehicle is shown in Fig. 5 when the launch vehicle is in that position.

As shown, at least two IGSO satellites can link with the launch vehicle to achieve space-based measurement and control information relay forwarding.

When carrying out measurement and control tasks for a launch vehicle, it needs to go through the IGSO to land the monitoring and control information at its domestic stations. Through the STK simulation, three satellites in the IGSO constellation and three domestic stations (Beijing/Kashi/Sanya station) visibility relationship as shown in Fig. 6.

Visibility analysis shows that only Sanya station can directly view the whole IGSO constellation during the whole-time period, so Sanya station is selected as the uploading and landing of the launch and control information.

The following IGSO constellation support measurement and control schemes:

1. When the launch vehicle is located in the visual area (the center angle of the cone is $Lon_B \sim 360^\circ (0^\circ) \sim Lon_A$) of the domestic station, the ground station directly carries out the real-time measurement and control communication.

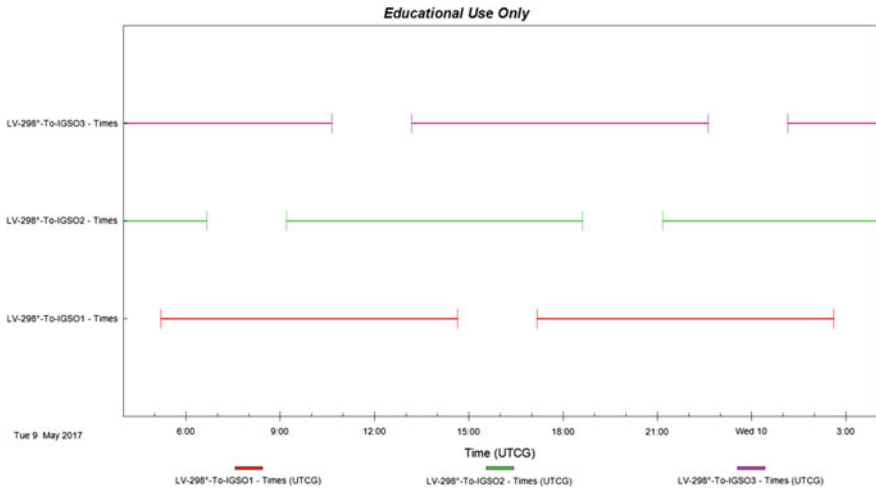


Fig. 5 Visibility analysis of IGSO constellation and launch vehicle

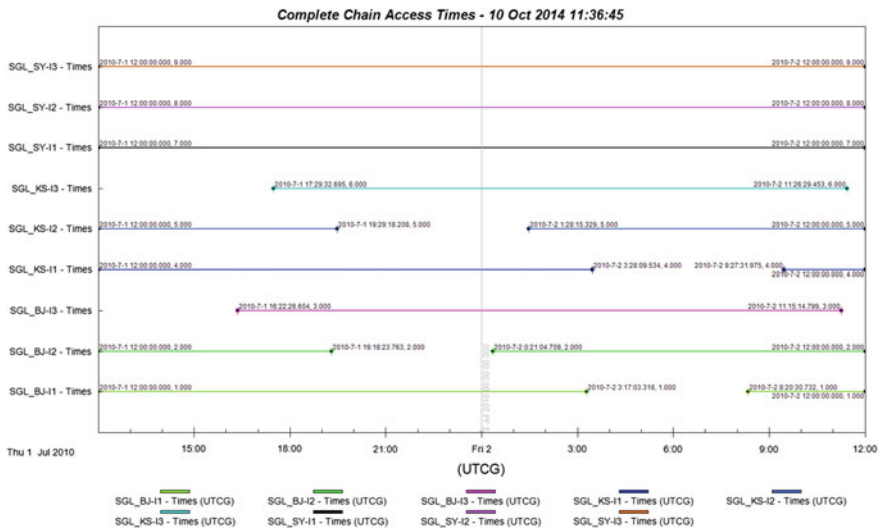


Fig. 6 IGSO constellation and Beijing station, Kashi station, Sanya station visibility analysis

- When the inflight launch vehicle is in the area not directly controllable (cone center angle $Lon_A \sim Lon_B$). Two or three of the IGSO constellations relaying, you can select IGSO1/IGSO2, IGSO1/IGSO3, IGSO2/IGSO3, or a combination of IGSO1/IGSO2/IGSO3 for relay forwarding. The inter-satellite link switch-overs according to the actual situation. That is, through the launch vehicle \leftrightarrow IGSO1 (or IGSO3...) \leftrightarrow Station.

4.3 MEO Constellation Support Launch Vehicle Control Communications Scheme

MEO constellation mainly for launch vehicles located in invisible airspace (cone center angle $\text{Lon}_A \sim \text{Lon}_B$) of the ground station of measurement and control system. Only the MEO satellites of BeiDou system relays the space-based measurement and control information.

Below one or more MEO satellites in a suitable position in the MEO constellation of BeiDou Global System can be used to select a chain, a single-hop or multi-hop selection to build a chain, and relay and forward the measurement and control information of the invisible airspace measurement and control targets in the space-based measurement and control system, the implementation of space-based measurement and control communications scheme design is as follows:

1. When the inflight launch vehicle is located in the visible area ($\text{Lon}_B \sim 360^\circ (0^\circ) \sim \text{Lon}_A$ in the domestic station, orange area in the figure), it directly transmits the measurement and control information to the domestic stations.
2. When the inflight launch vehicle is in the area not directly controllable (cone center angle $\text{Lon}_A \sim \text{Lon}_B$). The launch vehicle cannot establish a visible inter-satellite link with the domestic station and must be relayed by one or more MEO satellites of BeiDou system. When the launch vehicle has the mission demand, it establishes the inter-satellite link with the launch vehicle via the route calculation and scheduling MEO constellation (the MEO satellite establishes the inter-satellite link together with the launch vehicle based on the uplink or downlink times of the measurement and control command, the MEO operating condition and the path selection algorithm [5]. [ground station \rightarrow MEO_A \rightarrow MEO_B \rightarrow ... \rightarrow MEO_N \rightarrow launch vehicle]).

4.4 GEO, IGSO, MEO Constellations Support the Networking of Space-Based Measurement and Control Tasks

When the launch vehicle is in the visual airspace of the ground stations, selecting the direct measurement and control scheme preferentially. When the launch vehicle invisible with the ground stations, the BeiDou navigation constellation is selected according to the airspace coverage and relay hop constraints. As the inter-satellite communication link between launch vehicles and the GEO satellites is relatively mature and relay forwarding band is more convenient and topology is relatively simple, but building chain with the IGSO satellites is more complicated, and building chain with MEO satellites are more complicated. Considering the hop constraints, the priority choice of BeiDou system to support mission is GEO constellation, followed by the choice of IGSO constellation support, and finally

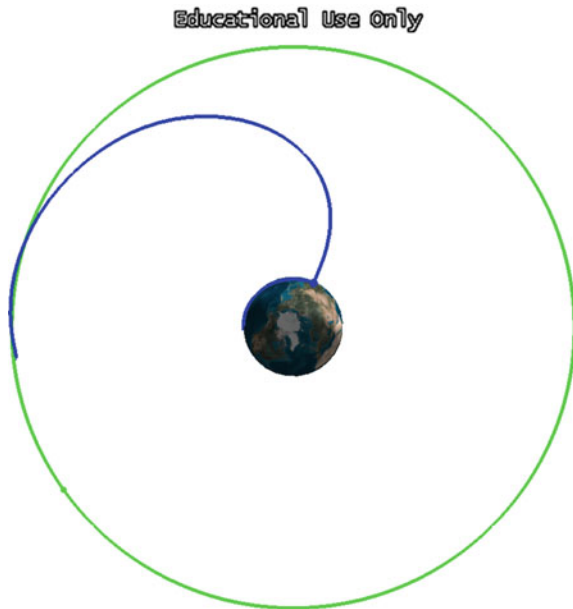
consider choosing MEO constellation support. Due to the large number of satellites in the MEO constellation and large global coverage. IGSO can cover area near China and Australia, and the GEO constellation can only cover the overlying regions in China. Considering the coverage constraint, the relay satellites gives priority to the MEO constellation in BeiDou system for mission support, secondly, IGSO is selected, and GEO is finally considered.

5 Simulation to Verify the Feasibility of Communication Solutions

The simulation software used in this paper is STK (Satellite Tool Kit). The satellites for monitoring and control is BeiDou constellations, including 5 GEO satellites, 3 IGSO satellites and 24 MEO satellites. In the simulation experiment, the launching trajectory of the launch vehicle with continuous thrust as shown in Fig. 7 is adopted. Simulation of the launching trajectory with a middle orbit is similar and will not be repeated.

To verify the BeiDou constellation’s support for the launch vehicle space-based measurement and control, that is, verify whether BeiDou constellation visible to launch vehicle, and that the launch vehicle has relay when it leaves the direct controllable area of the domestic station, which can relay information and realize the indirect the measurement and control of the launch vehicle.

Fig. 7 Continuous thrust trajectory STK simulation diagram



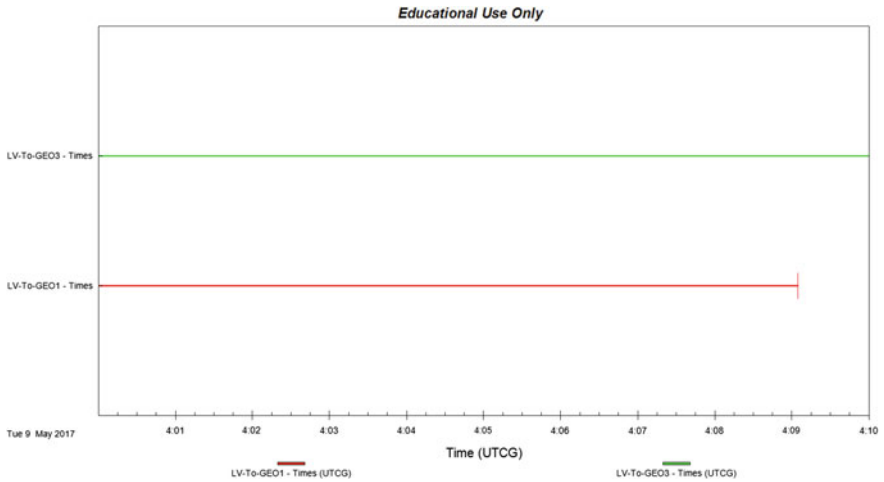


Fig. 8 Visibility analysis of continuous thrust orbit and GEO

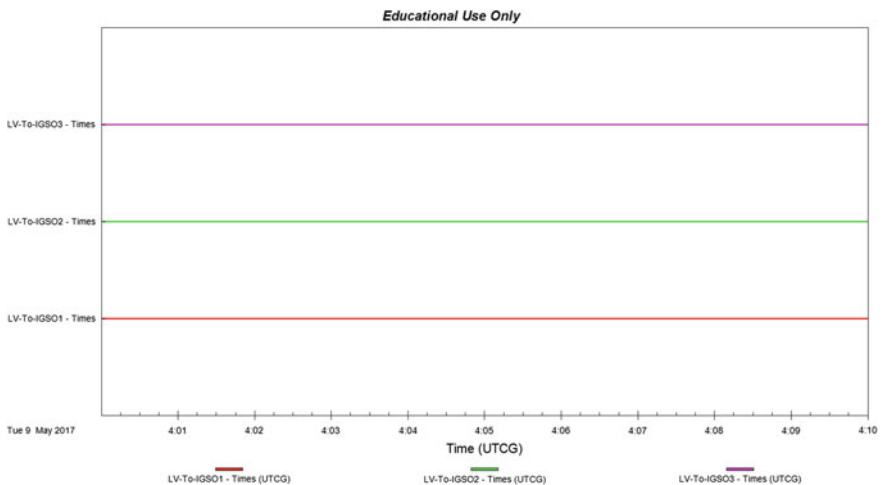


Fig. 9 Visibility analysis of continuous thrust orbit and IGSO

Taking the continuous thrust orbital launch vehicle launched from Wenchang Satellite Launch Center and finally reaching the geostationary orbit with a height of 36,000 km as an example, the visibility of the launch vehicle and the GEO, IGSO and MEO constellations is analyzed respectively, as shown in Figs. 8, 9 and 10.

As the simulation reports shown above, the visibility of the GEO constellation, the IGSO constellation, the MEO constellation and the launch vehicle are respectively calculated. It is obvious that at least one satellite and a launch vehicle target

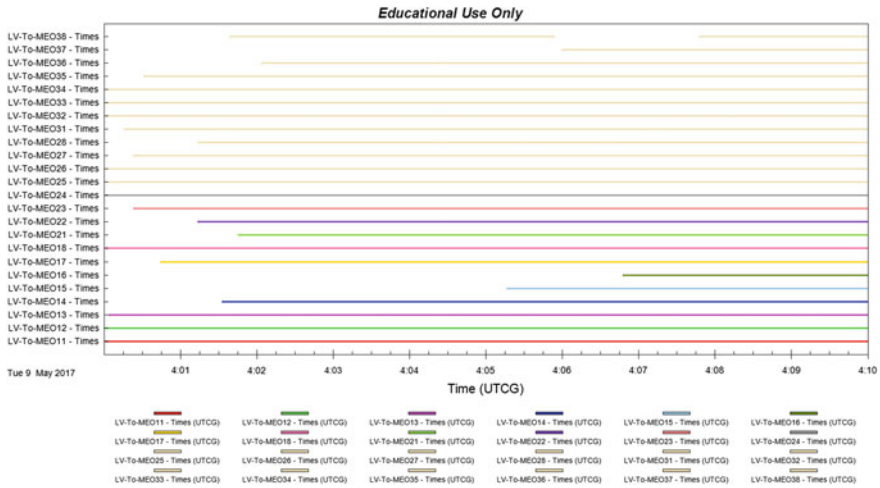


Fig. 10 Visibility analysis of continuous thrust orbit and MEO

are respectively visible in the GEO, IGSO and MEO constellations, during the launch vehicle flight, at least one satellite can be used for measuring and controlling information of its relay and realize the space-based measurement and control of the launch vehicle. Therefore, the space-based measurement and control scheme based on BeiDou system proposed in this paper is feasible, which can meet the requirements of full-time and full-space measurement and control of launch vehicles.

6 Summary

In view of the problems such as low coverage, large measurement and control cost, failure to realize full-air-space/full-time measurement and control and other issues, this paper proposes a space-based measurement and control plan for launch vehicle based on BeiDou satellite navigation system constellation. IGSO constellation and MEO constellation communication scheme is given in detail, and gives three kinds of constellations to support space-based measurement and control networking recommendations. Finally, the visibility of the three constellation satellites during the launching of the launch vehicle from our country is analyzed by simulation experiments. From the simulation results, there are always satellites and launch vehicles visible, so three kinds of measurement and control schemes are verified. It proves that space-based measurement and control system based on BeiDou system can meet the need of full-air-space and full-time measurement and control of launch vehicles and provides a feasible and better solution for the launch vehicle of space-based measurement and control technology in China. The application made a little contribution to the further application and development of BeiDou.

References

1. Liu J, Song Y (2012) Antenna cover performance analysis for space-based launch vehicle TT&C system. *J Telemet Track Command Comput Eng* 42(1):287–291
2. Wenping FAN, Jianxin LU (2017) Research on space-based telemetry data transmission technology for launch vehicle. *Radio Eng* 47(8):36–39
3. Yanhua LI, Manhong LU (2012) Trends of applications of space-based TT&C system. *J Spacecraft TT&C Technol* 31(4):1–5
4. Jing Z, Zhao S, Li Y et al (2013) Advance on data relay technology for inter-satellite laser links. *Infrared Laser Eng* 42(11):3103–3110
5. Yan H, Zhang Q, Sun Y (2015) Link assignment problem of navigation satellite networks with limited number of inter-satellite links. *Hangkong Xuebao Acta Aeronaut Astronaut Sin* 36(7):2329–2339
6. Dong G, Zhang G, Han Q (2014) Evolution of launch vehicle TT&C system technology. *J Spacecraft TT&C Technol* 33(2):93–98
7. Yachen Li, Junzhen Li, Zhilin Wang et al (2016) Measurement and control performance simulation of launch vehicle based on satellite tool kit 42(1):287–291
8. Zhang Q, Pei Y, Wang X et al (2014) Research on vehicle's telemetry command and communication system based on TDRSS. *J Telemet Track Command* 35(1):18–22

Online GPR-KF for GNSS Navigation with Unmodelled Measurement Error



Panpan Huang, Chris Rizos and Craig Roberts

Abstract To achieve the best performance for a Kalman filter (KF) for global navigation satellite system (GNSS) positioning, a comprehensive measurement model is required. However, the GNSS observations suffer from unmodelled errors resulting from multipath, interference, etc. These errors are difficult (even impossible) to model using parametric models. Inspired by Gaussian process (GP) Bayes filters with measurement and dynamic models trained with non-parametric GP regression (GPR), the unmodelled errors in the GNSS observations can be trained online based on the GPR using the measurements residuals calculated by the KF. One of the problems in using the GPR for online modelling is its high computational cost. To reduce the computational complexity, more than one forward step sliding window for the input training points and the GPR model training can be used. Furthermore, to avoid the over-prediction using the online trained GPR model, a constraint on the query point was introduced. In this study a non-linear autoregressive model was used for the online GPR model training. To verify this algorithm, both static and kinematic experiments were evaluated. The results show that the online GPR-KF algorithm can effectively improve the deteriorated GNSS positioning accuracy caused by unmodelled errors in the GNSS observations. The effectiveness of the proposed algorithm was also validated using the measurement innovation statistical test.

Keywords Gaussian process regression · Unmodelled measurement error
GNSS navigation

P. Huang (✉) · C. Rizos · C. Roberts
School of Civil and Environmental Engineering,
University of New South Wales, Sydney, NSW 2052, Australia
e-mail: p.huang@student.unsw.edu.au

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018
Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_5

1 Introduction

A precise measurement model is essential to achieve unbiased estimation for global navigation satellite system (GNSS) positioning with Kalman filtering (KF) based estimation methods. However, there are some errors in the GNSS observations, such as multipath and interference, that are difficult to model or do not cancel out through measurement differencing or cannot be mitigated using introduced bias products. To account for these errors, one commonly used strategy is to augment the state vector with the unmodelled errors so that they may be estimated along with receiver position and other relevant parameters [1, 3, 12]. However, a priori knowledge of the unmodelled errors, in terms of dynamic model or statistics, is assumed. Furthermore, if there is no error such methods could result in an unstable system. Another approach is to directly model these errors using parametric or non-parametric models [4–7, 11]. Parametric models include techniques such as autoregressive (AR), moving average (MA), autoregressive-moving average (ARMA), etc. Nonparametric models include artificial neural networks (ANNs), support vector machine (SVM), Gaussian Processes regression (GPR), and others. The disadvantage of parametric models is that substantial domain expertise is required to build these models, and that they are often simplified representations [6]. Non-parametric models are not as restrictive as the parametric model approach, as they offer a more flexible framework for approximating unknown non-linearities. The GPR method was studied in this paper as it provides not only the mean function value prediction of interest but also uncertainty estimates for the prediction [5]. In addition, it does not need the cross-validation procedure as required in SVM [11]. By taking both noise and regression uncertainty into account, the GPR can automatically increase its uncertainty when there is not enough training data. Furthermore, models such as AR, MA, and ARMA can be considered as special cases of GPR [12]. However, the GPR modelling approach usually performs model-learning offline using large volumes of training data derived from predicted measurements calculated with the known states and real observations. To adapt the GPR modelling to changes in the environment and in the system configuration, a GPR combined in the KF (GPR-KF) algorithm is proposed. The GPR model is trained in the proposed algorithm using previous measurement residuals calculated by the KF. One major limitation to training the GPR model online is its high computational cost caused by inverse matrix calculations for query point predictions. In addition, the computational burden will be increased by $O(m^3)$ when applied to GNSS positioning with m -dimensional satellite observations. To reduce the computational cost each output dimension is trained using a different GPR model, assuming the satellite observations are independent of each other. To improve reliability of the training points and further reduce the computational burden, a sliding window with more than one forward step for the input training points and model training is proposed. To avoid over-prediction of query data and to improve the robustness of the online GPR model due to its limited effectiveness, a constraint on the query data is also introduced.

The remainder of this paper is organised as follows. First, the problem of KF with unmodelled measurement error is introduced in Sect. 2. In Sect. 3, the standard GPR and the proposed online GPR-KF algorithm are described. To evaluate the proposed online GPR-KF algorithm for GNSS navigation, both static and kinematic experimental results and analyses are presented in Sect. 4. Finally, some concluding remarks are given in Sect. 5.

2 Problem Statements

Consider linear discrete systems modelled as:

$$\begin{aligned}\mathbf{x}_k &= \mathbf{\Phi}_{k|k-1}\mathbf{x}_{k-1} + \boldsymbol{\omega}_k \\ \mathbf{z}_k &= \mathbf{H}_k\mathbf{x}_k + \Delta\mathbf{z}_k + \boldsymbol{\varepsilon}_k\end{aligned}\quad (1)$$

where k denotes the epoch number; \mathbf{x} denotes the estimated state vector; \mathbf{z} denotes the measurement vector; $\mathbf{\Phi}_{k|k-1}$ and \mathbf{H}_k represent the state transition matrix and design matrix from the state to the measurement, respectively; $\boldsymbol{\omega}_k$ and $\boldsymbol{\varepsilon}_k$ denote the process noise and measurement noise with known covariances \mathbf{Q}_k and \mathbf{R}_k , respectively; and $\Delta\mathbf{z}_k$ denotes the unmodelled measurement error uncorrelated with $\boldsymbol{\varepsilon}_k$. There are several ways to take into account $\Delta\mathbf{z}_k$ in KF. For example, one common approach is to augment the state with $\Delta\mathbf{z}_k$ and allow its estimation along with \mathbf{x}_k . However this approach assumes the process model of $\Delta\mathbf{z}_k$ or a priori knowledge of its statistics is known. Furthermore, when there is no error in the measurement, this approach could lead to an ill-conditioned system. Another approach is via directly modelling of $\Delta\mathbf{z}_k$. Hence the update step in KF can be modified as follows [2]:

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_{k|k-1}\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{P}_{\Delta\hat{\mathbf{z}}_{k|k-1}}\right)^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1} - \Delta\hat{\mathbf{z}}_{k|k-1}\right)\end{aligned}\quad (2)$$

where $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ denote the predicted state vector and its associated covariance, respectively; \mathbf{K}_k is the Kalman gain matrix; $\hat{\mathbf{x}}_{k|k}$ denotes the KF estimated state; and $\Delta\hat{\mathbf{z}}_{k|k-1}$ and $\mathbf{P}_{\Delta\hat{\mathbf{z}}_{k|k-1}}$ denote the estimated unmodelled error and its associated covariance, respectively. Since the expectation of the measurement residual is:

$$\begin{aligned}E(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k}) &= \mathbf{H}_kE(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) + E(\boldsymbol{\varepsilon}_k) + E(\Delta\mathbf{z}_k) \\ &\approx E(\Delta\mathbf{z}_k)\end{aligned}\quad (3)$$

To estimate $\Delta\hat{\mathbf{z}}_{k|k-1}$, the non-parametric GPR regression is investigated in this paper as it has the advantage of not only providing the mean function value

prediction of interest but also uncertainty estimates for the prediction. This modelling method can automatically increase its uncertainty when there is not enough training data. However, the GPR modelling approach usually performs model-learning offline using large volumes of training data. In order to take full advantage of such a modelling method an online GPR combined in a KF is demanding as it must adapt to changes in the environment and in the system configuration.

3 Online GPR-KF

In this section the standard GPR is first reviewed. Then the online GPR-KF to account for unmodelled measurement error is proposed to address the shortcomings of the standard approach.

3.1 Standard GPR

A GPR can be thought of as a ‘‘Gaussian over functions’’ [8, 9]. A GP is fully specified by its mean and covariance functions. Assume a set of n training data $\{\mathbf{x}_i, y_i\}_{i=1}^n$ drawn from the noisy process. The observed targets can be described by a zero-mean multivariate Gaussian distribution:

$$\mathbf{y} \sim N(\mathbf{0}, K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}) \quad (4)$$

where \mathbf{y} is the aggregated output vector $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$; \mathbf{X} is the aggregated input matrix $\mathbf{X} = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \dots \ \mathbf{x}_n^T]^T$; $K(\mathbf{X}, \mathbf{X})$ is the covariance matrix; and is σ_n the variance of output noise. To calculate the $K_{i,j}$ elements of $K(\mathbf{X}, \mathbf{X})$, a kernel function in the form of a squared exponential (SE) is commonly used:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_s^2 \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \mathbf{W}(\mathbf{x} - \mathbf{x}')\right) \quad (5)$$

where σ_s^2 is the signal variance and \mathbf{W} is a diagonal matrix of the length scales of each input dimension. The optimal value of hyperparameters of a GP with SE can be derived by maximising the log marginal likelihood $\arg \max_{\boldsymbol{\theta}} \{\log(p(\mathbf{y}|\mathbf{X}, \boldsymbol{\theta}))\}$ [10]. The joint distribution of the observed target values and predicted value for a query point \mathbf{x}_* is given by:

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I} & k(\mathbf{X}, \mathbf{x}_*) \\ k(\mathbf{x}_*, \mathbf{X}) & k(\mathbf{x}_*, \mathbf{x}_*) \end{bmatrix}\right) \quad (6)$$

Thus, the predictive distribution over the output y_* becomes:

$$p(y_* | \mathbf{x}_*, \mathbf{X}, \mathbf{y}) \sim N\left(k_*^T (K + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}, k_{**} - k_*^T (K + \sigma_n^2 \mathbf{I})^{-1} k_*\right) \quad (7)$$

with $k_* = k(\mathbf{x}_*, \mathbf{X})$, $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$, and $K = K(\mathbf{X}, \mathbf{X})$. The mean prediction of y_* and its uncertainty are $k_*^T (K + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}$ and, $k_{**} - k_*^T (K + \sigma_n^2 \mathbf{I})^{-1} k_*$, respectively.

3.2 Online GPR-KF

The major limitation of training the GPR model online is its expensive computation of the inverse matrix $(K + \sigma_n^2 \mathbf{I})^{-1}$ which yields a cost of $O(n^3)$. For m -dimensional outputs, the computational cost would increase to $O(m^3 n^3)$. In order to reduce the computational cost, each output dimension is trained using a different GPR model, assuming the output dimensions are independent with each other. In this way the computational cost is reduced to $O(mn^3)$. To apply the standard GPR to predict the unmodelled measurement error $\Delta \mathbf{z}_k$ in the KF, the model of $\Delta \mathbf{z}_k$ used in this paper is assumed to be a non-linear autoregressive model:

$$\Delta \mathbf{z}_k = f(\Delta \mathbf{z}_{k-1}, \dots, \Delta \mathbf{z}_{k-p}) + \eta_k \quad (8)$$

where p is the number of lags used to represent the unmodelled measurement error; and η_k is white Gaussian noise. To train this model the training points of inputs \mathbf{X} and outputs \mathbf{y} can be obtained using the estimated $\Delta \mathbf{z}_k$ of previous epochs. To improve the reliability of the training points, the input training points can use a sliding window with more than one forward step:

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} \Delta \hat{\mathbf{z}}_{k-p-(n-1):s|k-p-(n-1):s} & \Delta \hat{\mathbf{z}}_{k-p-(n-1):s+1|k-p-(n-1):s+1} & \cdots & \Delta \hat{\mathbf{z}}_{k-(n-1):s-1|k-(n-1):s-1} \\ \Delta \hat{\mathbf{z}}_{k-p-n:s|k-p-n:s} & \Delta \hat{\mathbf{z}}_{k-p-n:s+1|k-p-n:s+1} & \cdots & \Delta \hat{\mathbf{z}}_{k-n:s-1|k-p-n:s-1} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta \hat{\mathbf{z}}_{k-p|k-p} & \Delta \hat{\mathbf{z}}_{k-p+1|k-p+1} & \cdots & \Delta \hat{\mathbf{z}}_{k-1|k-1} \end{bmatrix} \\ \mathbf{y} &= \begin{bmatrix} \Delta \hat{\mathbf{z}}_{k-(n-1):s|k-(n-1):s} \\ \Delta \hat{\mathbf{z}}_{k-n:s|k-p-n:s} \\ \cdots \\ \Delta \hat{\mathbf{z}}_{k|k} \end{bmatrix} \end{aligned} \quad (9)$$

where s is the forward steps. With the trained model of $\Delta \mathbf{z}_k$, the prediction of $\Delta \hat{\mathbf{z}}_{k+1|k}$ and $\mathbf{P}_{\Delta \hat{\mathbf{z}}_{k+1|k}}$ can be obtained by treating $\mathbf{x}_{*,k} = [\Delta \hat{\mathbf{z}}_{k-p+1|k-p+1} \Delta \hat{\mathbf{z}}_{k-p+2|k-p+2} \cdots \Delta \hat{\mathbf{z}}_{k|k}]^T$ as the query point. Here the training points \mathbf{X} and the query point $\mathbf{x}_{*,k}$ only represent a one-dimensional output of $\Delta \mathbf{z}_k$. This strategy increases the demand for memory. To further reduce the computational cost of $\Delta \hat{\mathbf{z}}_{k+1|k}$, the trained GPR model also uses a sliding window with q steps forward after the first modelling. During the next q epochs, $\Delta \mathbf{z}_k$ is only predicted in the KF using the last trained model for the first l epochs and for the last $q-l$ epochs KF without $\Delta \mathbf{z}_k$ correction is performed. Therefore the query points for every q epochs contains:

$$\mathbf{x}_* = [\mathbf{x}_{*,k} \quad \mathbf{x}_{*,k+1} \quad \cdots \quad \mathbf{x}_{*,k+l}] \quad (10)$$

To avoid over-prediction of l epochs query data and to improve the robustness of the online GPR model due to its limited effectiveness, a constraint is introduced:

$$\begin{cases} \Delta \hat{\mathbf{z}}_{k+1|k}, \mathbf{P}_{\Delta \hat{\mathbf{z}}_{k+1|k}} & \text{norm}(\mathbf{X}_i - \mathbf{x}_{*,j}) < d_{th} \\ \mathbf{0} & \text{otherwise} \end{cases} \quad i = 1, 2, \dots, n \quad j = k, k+1, \dots, k+l \quad (11)$$

where \mathbf{X}_i is the i th training data; $\mathbf{x}_{*,j}$ is the j th query data; and d_{th} is the threshold for Euclidean distance between the training data and the query data. When the Euclidean distance between the training data and the query data is smaller than the threshold, prediction is implemented, otherwise conventional KF is performed. When there is no unmodelled error $\Delta \mathbf{z}_k$ the online GPR-KF algorithm is reduced to the conventional KF as $\Delta \mathbf{z}_k$ derived from the measurement residual will have the same distribution as the noise $\boldsymbol{\varepsilon}_k$ in the measurement with zero-mean and covariance \mathbf{R}_k . However, directly using $\mathbf{P}_{\Delta \hat{\mathbf{z}}_{k|k-1}}$ in the GPR-KF could cause fluctuations in the estimation, and even lead to divergence due to rapid changes of the training data. Therefore in order to further improve the robustness of the online GPR-KF algorithm the Kalman gain \mathbf{K}_k can be calculated with a scale factor f as follows:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \frac{\mathbf{P}_{\Delta \hat{\mathbf{z}}_{k|k-1}} + \mathbf{R}_k}{f} \right]^{-1} \quad (12)$$

If the $\mathbf{P}_{\Delta \hat{\mathbf{z}}_{k|k-1}}$ is predicted precisely, then $(\mathbf{P}_{\Delta \hat{\mathbf{z}}_{k|k-1}} + \mathbf{R}_k)/f$ with f set as 2 is equivalent to $\mathbf{P}_{\Delta \hat{\mathbf{z}}_{k|k-1}}$ which contains both the propagated covariance of the unmodelled error and measurement noise.

The detailed procedure for the online GPR-KF algorithm is presented in Fig. 1. This method can be readily extended to non-linear KF methods such as Extended KF (EKF), Unscented KF (UKF), and Cubature KF (CKF), since the

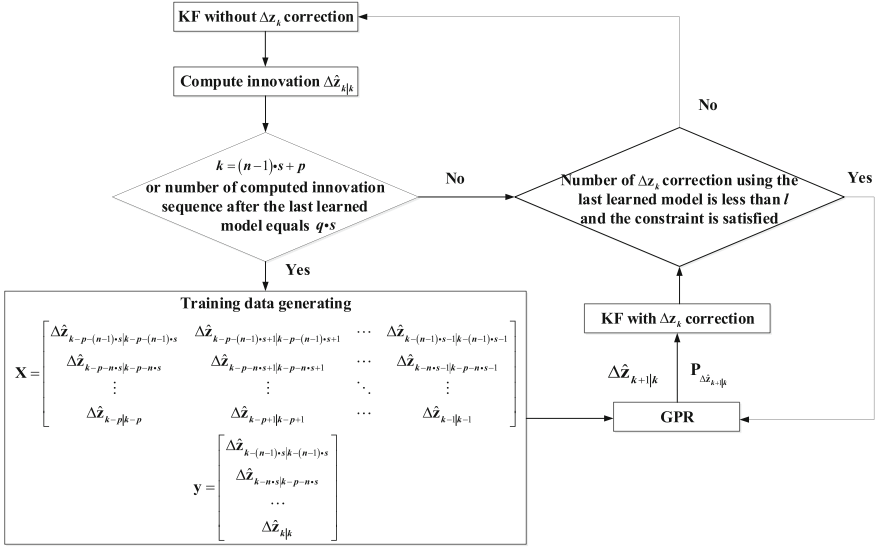


Fig. 1 Online GPR-KF algorithm flowchart

measurement residual sequences for the error model training are calculated after the update step of all the non-linear KF methods and the unmodelled error correction can be seen as new measurement noise with $\Delta\hat{z}_{k|k-1}$ mean and covariance $P_{\Delta\hat{z}_{k|k-1}}$.

4 Experiment Results and Analysis

Both static and kinematic scenarios were investigated to validate the efficiency of the online GPR-KF algorithm. GPS observations from one of the International GNSS Service (IGS) stations, with added coloured noise, were tested for the static scenario. In the case of the kinematic scenario GPS observations from a UAV experiment were used. The EKF was used in the following experiment validation.

4.1 Static Experiment

GPS measurements at 30 s intervals from the ALIC IGS station during day-of-year 277 and 278 in 2016 were randomly selected for the static experiment. The data were processed using the GNSS PPP technique with float ambiguity in the quasi-real-time mode. The parameters used in the GPR-KF were $p = 3$, $n = 15$, $q = 10$, $l = 6$, $f = 3$, and $s = 2$. GPR modelling for each observed satellite is

Table 1 Positioning accuracy with different noise magnitudes

Noise magnitudes (cm)	Positioning errors (cm)	
	Conventional PPP	PPP with GPR-KF
0	5	4
<10	8	6
<20	14	11
<30	20	16

assumed to be independent. To validate the effectiveness of the GPR-KF algorithm, all carrier phase measurements contaminated with coloured noise of different magnitudes were processed, as indicated in Table 1. The noise magnitude added to pseudorange measurements is 3 orders larger than that in the carrier phase measurements. The added coloured noise should not cause the measurement innovation sequence to diverge. Thresholds d_{ih} for pseudorange and carrier phase were set at 20 and 0.5, respectively. The station positioning errors obtained by conventional PPP, and the PPP with online GPR-KF, were compared with the IGS weekly solutions as shown in Fig. 2. The root mean square error (RMSE) of the station position estimation calculated from 10 h for both algorithms are given in Table 1. It can be seen that the positioning accuracy derived from the PPP with GPR-KF is better than that from the conventional PPP. The PPP with GPR-KF reduces the positioning fluctuations resulting from the rapid changing coloured noise, especially when the positioning accuracy has a tendency to decrease, as shown in Fig. 2.

Theoretically, if there is no unmodelled error in the measurement model, the measurement innovations should have a zero-mean Gaussian distribution. This property of the innovations is a measure of filter performance. Figure 3 shows the comparison of mean and standard deviation (STD) of GPS measurement innovations derived from these two algorithms with no added noise. The left plot in Fig. 3 is the pseudorange innovation statistics of all observed GPS satellites, and the right plot is for carrier phase innovations. It can be seen that the innovations obtained by the GPR-KF have smaller mean and STD than for the case of the conventional PPP, which are more like zero-mean normally distributed quantities. In addition, the pseudorange improvement in terms of the innovations' whiteness property is higher than the carrier phase due to its lower accuracy.

4.2 Kinematic Experiment

To investigate the efficiency of the online GPR-KF algorithm for kinematic scenarios a UAV experiment was carried out and dual-frequency GPS data collected by an OEM617 GPS receiver with sampling rate 10 Hz. Both the conventional PPP and PPP with the GPR-KF algorithms were implemented on an epoch-by-epoch

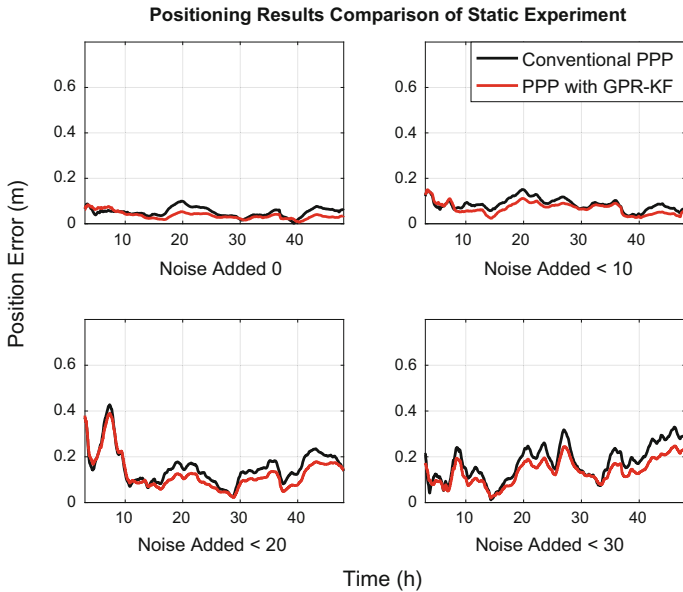


Fig. 2 Positioning errors with different noise magnitudes

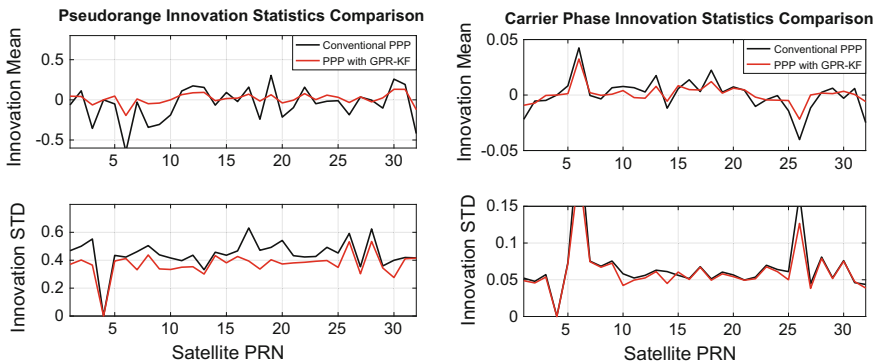


Fig. 3 Measurement innovation statistics comparison

basis in post-processing mode. The parameters used in the kinematic experiment were the same as in the static experiment. Thresholds d_{th} for pseudorange and carrier phase were set at 10 and 0.1, respectively.

To compare the performance of these two algorithms, the post-processed GPS-RTK positioning results were used as reference. Figure 4 is the comparison of the positioning error results of the two algorithms. Table 2 gives the detailed positioning accuracy comparison calculated from 20 min. It can be seen that the

Fig. 4 Positioning error comparison

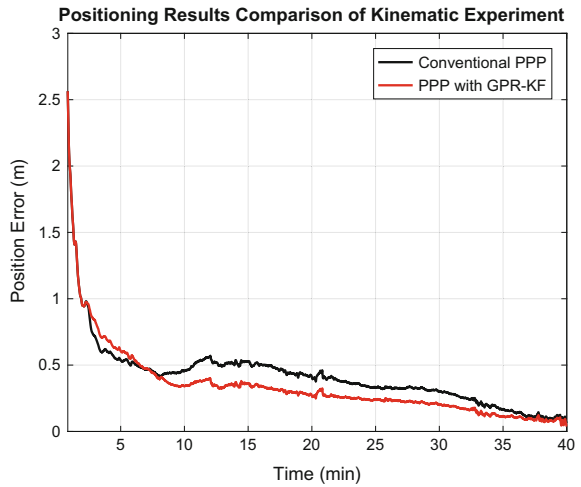


Table 2 Positioning accuracy comparison

Algorithm	Positioning errors (cm)
Conventional PPP	26
PPP with GPR-KF	21

positioning accuracy achieved by the GPR-KF algorithm is higher than that of the conventional PPP due to its ability to effectively reduce the deteriorated positioning accuracy as shown in Fig. 4 at around 8 min. The pseudorange and carrier phase innovation statistics of all observed GPS satellites from the two algorithms are compared in Fig. 5. The comparison of the measurement innovation distribution for satellite PRN 12 is shown as a typical example in Fig. 6. It can be seen that the whiteness properties of both the pseudorange and carrier phase innovations are

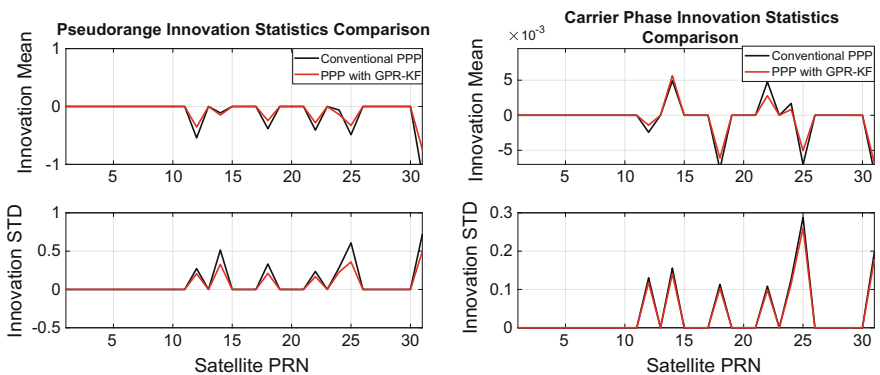


Fig. 5 Measurement innovation statistics comparison

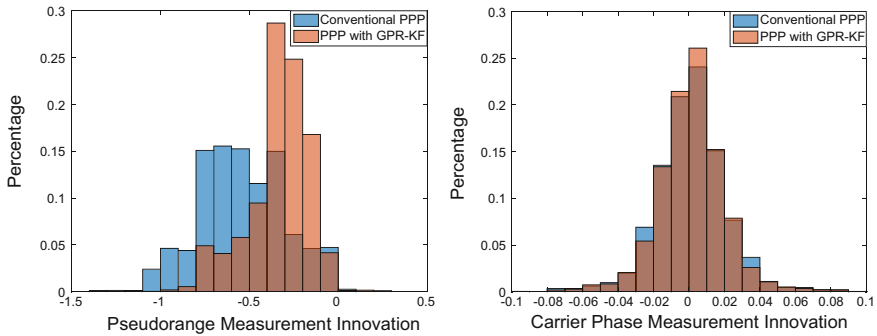


Fig. 6 Distribution of measurement innovations

improved. The improvement achieved for pseudorange measurements is higher than for carrier phase, especially in terms of STD. This is because of the lower accuracy of pseudorange measurements and their vulnerability to errors resulting from the changing environment and unstable dynamics compared with carrier phase measurements.

5 Concluding Remarks

This paper describes an online GPR-KF algorithm to deal with the unmodelled errors in GNSS observations. Independent GPR models were trained online for all observed satellites using the corresponding measurement residuals calculated by the KF. Due to the high computational cost of GPR modelling and limited training data, a more than one step sliding window was used. To avoid over-prediction using the trained GPR model, a constraint on the query point was introduced. The online GPR-KF for GNSS navigation was evaluated with both static and kinematic experiments. The results reveal that the proposed algorithm effectively reduces the effect of the unmodelled error and achieves better positioning performance compared with conventional GNSS navigation algorithms.

References

1. Chang G (2014) Alternative formulation of the Kalman filter for correlated process and observation noise. *IET Sci Meas Technol* 8(5):310–318
2. Hamilton JD (1994) *Time series analysis*. Princeton University Press
3. Hu Y, Duan Z, Zhou D (2010) Estimation fusion with general asynchronous multi-rate sensors. *IEEE Trans Aerosp Electron Syst* 46(4):2090–2102
4. Ko J, Fox D (2009) GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models. *Auton Robots* 27(1):75–90

5. Ko J, Kleint DJ et al (2007) GP-UKF: unscented Kalman filters with Gaussian process prediction and observation models. In: 2007 IEEE/RSJ international conference on intelligent robots and systems on intelligent robots and systems, San Diego, USA, Oct–Nov 2007
6. Lee K, Johnson EN (2017) State estimation using Gaussian process regression for coloured noise systems. In: 2017 IEEE aerospace conference, MT, USA, Mar 2017
7. Lv Y, Duan Y et al (2015) Traffic flow prediction with big data: a deep learning approach. *IEEE Trans Intell Transp Syst* 16(2):865–873
8. Nguyen-Tuong D, Peters JR, Seeger M (2009) Local Gaussian process regression for real time online model learning. In: *Advances in neural information processing systems*, Vancouver, British Columbia, Canada, 7–10 Dec 2009
9. Rasmussen CE, Williams CK (2004) *Gaussian processes in machine learning*. *Lect Notes Comput Sci* 3176:63–71
10. Rasmussen CE, Williams CK (2006) *Gaussian processes for machine learning*. MIT Press, Cambridge
11. Zhou Z, Li Y, Fu C, Rizos C (2016) Least-squares support vector machine-based Kalman filtering for GNSS navigation with dynamic model real-time correction. *IET Radar Sonar Navig* 11(3):528–538
12. Zhou Z, Wu J, et al (2017) Critical issues on Kalman filter with coloured and correlated system noises. *Asian J Control* 19(6):1905–1919

The Application of a New Multi-gross Errors Detection Method with the Epoch Variation Constraint in UAV



Hongyu Shan, Chunhua Li, Xiaoyu Shi and Chenggang Li

Abstract It is very important to get precise positioning information continuously in the UAV RTK. However, the complex environment often degrades the observation quality or even introduces several gross errors which will seriously affect the accuracy and reliability of the results. In this paper, a new method of multi gross error detection with the constraint of epoch variation is put forward and validated by the UAV-PPK. The result shows that it can effectively detect multiple gross errors and handle the gross-errors and cycle-slips in a different way. Besides, it can also guarantee the stability of the filtering and significantly increase the ambiguity fixing rate and positioning accuracy. Thus, the reliable and precise positioning information can be continuously supplied for UAV RTK.

Keywords UAV · Epoch variation · Multi-gross errors detection
RTK

1 Introduction

RTK (Real Time Kinematic) is a real-time kinematic relative positioning technique with the carrier phase observation [1], which can continuously provide the real-time position in centimeter level. With the development of UAV, a series of scenarios, such as plant protection, aerial survey and formation flight, have been associated with RTK [2–5] and the demand of higher accuracy and higher reliability position increases quickly. The core of these applications is how to ensure the accuracy and reliability of RTK.

On the flight of UAV, the surrounding environment is very complex and changes rapidly which makes it difficult to track the satellite continuously and stably by UAV receiver and the gross errors and cycle slips may occur frequently.

H. Shan (✉) · C. Li · X. Shi · C. Li
Hi-Target Surveying Instrument Co. Ltd, No. 555, North of Panyu Rd.,
Panyu District, Guangzhou 511493, China
e-mail: hyshan@whu.edu.cn

For the problems of the cycle slip occurred in the complex environment, many scholars have proposed a lot of methods, including ionosphere residual method, high order differenced method, kalman filter method, TurboEdit method and epoch-differenced method [6, 7]. And for the multi gross errors, the Robust Kalman Filter (RKF) is always adopted. Zhou put forward the IGG I in 1989 [8], then the model is improved by Yang et al. with the IGG II and IGG III model in succession [9]. Guo uses the IGG III model and re-initialize the observation data in the rejection region, the practical test has validated its effectiveness [10].

The sampling rate of UAV-RTK is usually high which guarantees the consistency of the observation between two epochs. Therefore, this paper put forward a scheme that first deals with the multi gross errors with the epoch-differenced model and then adopts the Robust Kalman Filter with the precise epoch-variation constraint. Besides, the cycle slips and gross errors are handled separately which guarantee the stability of the filtering even in the case of multi cycle slips and multi gross errors. Finally, the reliable and precise positioning information can be continuously supplied.

2 The Mathematical Model of Epoch-Differenced Method

In the dynamic positioning of UAV, the satellite clock bias, satellite orbit bias, ionosphere bias and troposphere bias can be eliminated by the model of epoch-differenced method. The only parameters to be estimated are the position variations between epochs and clock changes, the precise epoch differences can be easily computed by kalman filter.

Assuming the base station i and rove station j track the satellite p at the epoch t_0 and the next epoch t_1 , the atmosphere parameter and ambiguity parameter can be eliminated after epoch differenced, the complete mathematical model are as follows after linearization.

$$\begin{aligned} v_p = & l_{ij,t_1}^p dx_{t_1} + m_{ij,t_1}^p dy_{t_1} + n_{ij,t_1}^p dz_{t_1} \\ & + c \cdot dt - l^p + \gamma_{t_0,t_1}^p + \zeta_{t_0,t_1}^p \end{aligned} \quad (1)$$

where l , m , n are the coefficient matrix, dx, dy, dz are the position variations, respectively, dt is the clock variation, l is the epoch-diff constant, γ_{t_0,t_1}^p is the ignorance part of the epoch-diff equation for satellite p , ζ_{t_0,t_1}^p is the residual noise, where l and γ are as follows:

$$\begin{aligned} l^p = & (\phi_{ij,t_1}^p - \phi_{ij,t_0}^p) - (\rho_{ij,t_1}^p - \rho_{ij,t_0}^p) \\ \gamma_{t_0,t_1}^p = & (l_{ij,t_1}^p - l_{ij,t_0}^p)x_{t_0} + (m_{ij,t_1}^p - m_{ij,t_0}^p)y_{t_0} + (n_{ij,t_1}^p - n_{ij,t_0}^p)z_{t_0} \end{aligned} \quad (2)$$

The parameter ϕ is the carrier phase observations, and the ρ is the geometry of the satellite to the station. For GPS satellites, the velocity is less than 3800 m/s, for the interval of 0.1, 0.2 s, or even 0.5 s, the bias of γ is 0.84, 1.68, and 4.2 mm, the value is far less than the wavelength, so it can be completely ignored [11].

3 The Principle of Gross Error Detection with the Epoch Variation Constraint

In the dynamic situation, multi-cycle slips and multi-gross errors are usually common and it's difficult to distinguish them, the safest way is to re-initialize them. However, frequent initialization may destroy the stability of the filter, which will result in the precision worse or even filter divergence. This section is mainly focus on the above problems, put forward a new multi-gross errors detection method with the epoch variation constraint, it can effectively detect multiple gross errors and process the gross-errors and cycle-slips in a different way. The following is a detailed introduction to the principle of the algorithm.

3.1 The Principle of Multi-gross Errors Detection

The multiple gross errors detection method used in this paper is divided into two steps.

- Step 1. Data preprocessing by the epoch-differenced method, it can preliminary detect the cycle slips and multi-gross errors.
- Step 2. Robust kalman filter with the epoch variation constraint, deal with the residual cycle slips and the gross errors.

First, for the step 1, the threshold of can't be set too tight to ensure that most of the cycle slips and gross errors are detected right. In consideration of the system deviation of the modification of the equation is less than 3 cm [12], the threshold set to be 3 cm. The standardized residual value is required to be determined separately for every observation, the concrete forms are as follows:

$$\tilde{v}_i = \frac{v_i}{\sqrt{\sigma_{v_k}^2}} \quad (3)$$

Due to the error will be divided into the observed values for the least squares adjustment, so we only need to deal with the maximum value. If we can't meet the requirements after a number of eliminations, the cycle slips or gross errors of the reference satellite will be taken into consideration. The reference should be changed.

Based on the IGG III model, we can get the accurate prior information (centimeter level) with the precise epoch variation and the prior fixed position. The reliability of the robust method can obviously increase. The median is introduced to compute the normalized residual for reflecting the overall characteristics of the residual. The robust model is different for pseudo-range and carrier phase, so as follows:

$$\begin{aligned} \gamma_P &= \begin{cases} 1 & (|\tilde{v}_i| \leq k_0) \\ a|\tilde{v}_i| - b & (k_0 < |\tilde{v}_i| \leq k_1) \\ 0 & (|\tilde{v}_i| > k_1) \end{cases} \\ \gamma_\phi &= \begin{cases} 1 & (|\tilde{v}_i| \leq k_0) \\ \frac{k_0}{|\tilde{v}_i|} \left(\frac{k_1 - |\tilde{v}_i|}{k_1 - k_0} \right)^2 & (k_0 < |\tilde{v}_i| \leq k_1) \\ 0 & (|\tilde{v}_i| > k_1) \end{cases} \end{aligned} \quad (4)$$

where the γ_P, γ_ϕ is the magnifying factor of pseudo-range and carrier phase, respectively, k_0, k_1 is the threshold of the robust model, a, b can be set by the receiver quality, for this article, $a = 4.5, b = 3.5, k_0 = 1.0, k_1 = 3.0$.

3.2 The Principle of Distinguish the Cycle Slips and Gross Errors

For the carrier phase observation values which in the rejection region, the traditional way is to deal with it as gross errors and abandon it. But if it's the undetected cycle slip, we don't consider the ambiguity parameter's change, the observation may be continuously located in the rejection region. The other method is a safety way to deal with it as cycle slip, and re-initialize the ambiguity parameter. However, frequent initialization may destroy the stability of the filter, which will result in the precision worse or even filter divergence.

The gross error is an accidental error, but the cycle slip is a continuous change. So we can distinguish the cycle slips and gross errors by this characteristic.

- (1) For the epoch-differenced method, when the carrier phase observation with the same satellite is located in the rejection region between the adjacent epoch, we think it's a gross error, or it's a cycle slip, the ambiguity parameter should be re-initialized.
- (2) For the robust kalman filter, it is opposite, when the carrier phase observation with the same satellite is located in the rejection region between the adjacent epoch, we think it's a cycle slip, and the ambiguity parameter should be re-initialized, or it's a gross error.

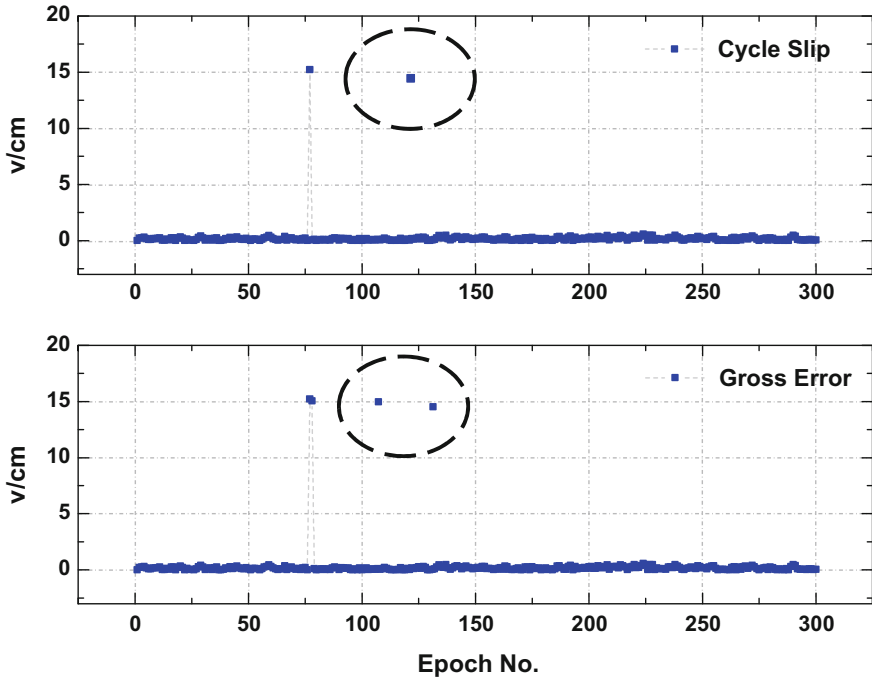


Fig. 1 Epoch-differenced test

To verify the reliability of this method, we chose a period of clean data (five minutes), add cycle slips or gross errors for G06 (without re-initialized ambiguity parameters). The results are as follows (Figs. 1 and 2).

It can be found from the figures above that if the ambiguity parameter is not re-initialized when the cycle slip happened, the observation will be continuously located in the rejection region which will affects its contribution to the filter. So we can distinguish the cycle slips and gross errors based on whether it's continuously located in the rejection region.

4 Experiment Demonstration

To validate the new multi-gross errors detection method with the epoch variation constraint, and evaluate the improvement of the RTK positioning performance of UAV. We choose some different UAV data, then deal with them by UAV-PPK V2.0.0 which is independent developed by ourselves. Then compare the results with the commercial high precision dynamic processing software GrafNav v8.6, finally, give the statistical results as follows.

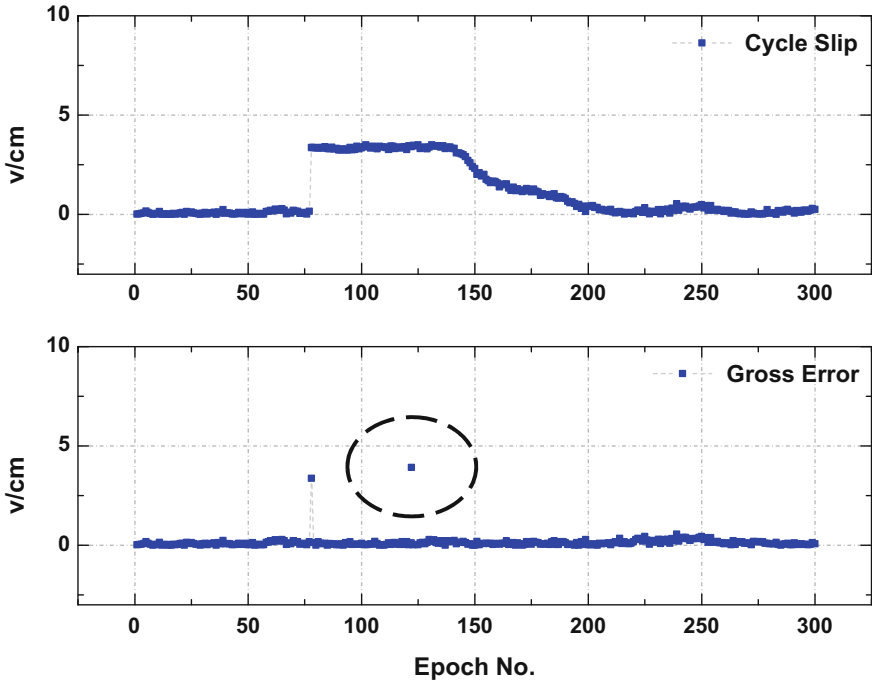


Fig. 2 Robust kalman filter test

4.1 Experiment 1

The data is collected in July 18, 2016 with three systems, the interval is 0.2 s, the observation time is about 1 h long. For detailed analyzing the data quality, the figure below shows the number of satellites and PDOP values for base and rove stations (Top: base station, Bottom: rove station, the same as follows). The figure shows the quality of the data is good, the number of visible satellites for the rove station is about 20, the PDOP is less than 2 (Figs. 3 and 4).

We deal with the data by UAV-PPK, the original kalman filter and the multi-gross errors detection method with the epoch variation constraint are presented. The results show as follows (Fix result with color blue, unfix result with color red, the same as follows) (Figs. 5 and 6).

The sequence diagram of difference with GrafNav is as follows (Figs. 7 and 8).

The specific statistics are as follows (Four columns represent the fix rate and the rms of ENU, the same as follows) (Table 1).

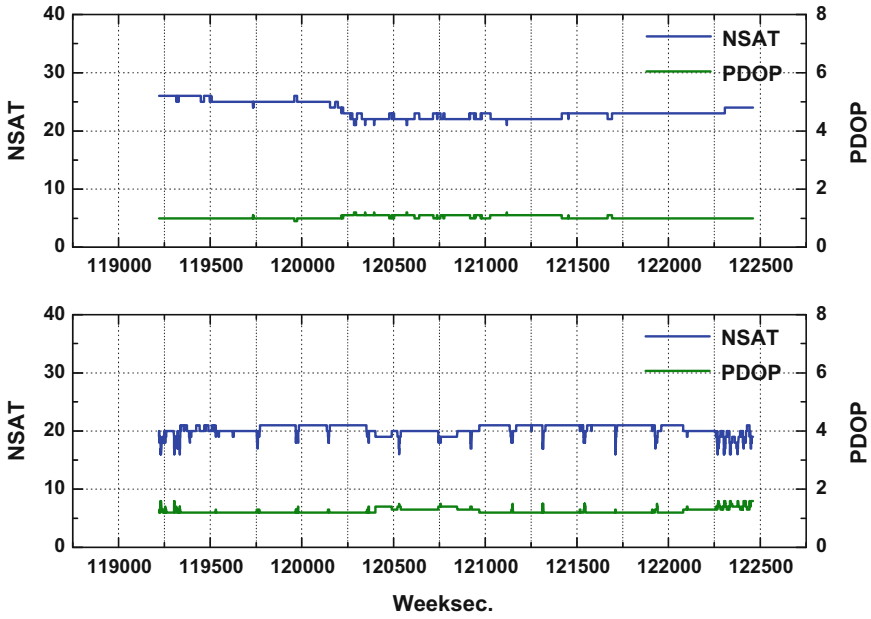


Fig. 3 NSAT and PDOP

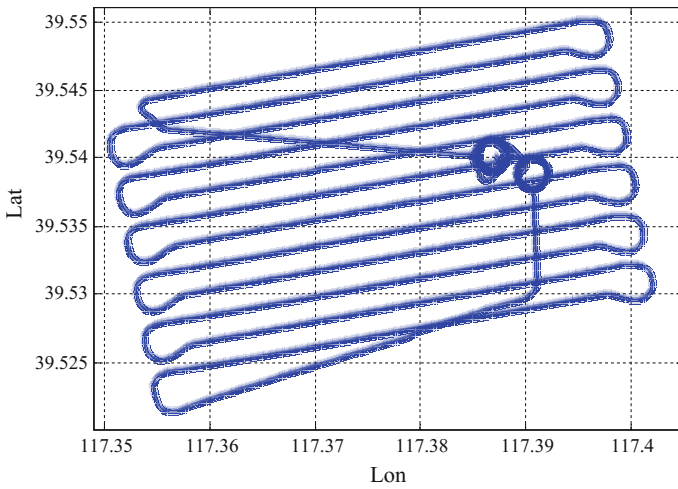


Fig. 4 Track of UAV

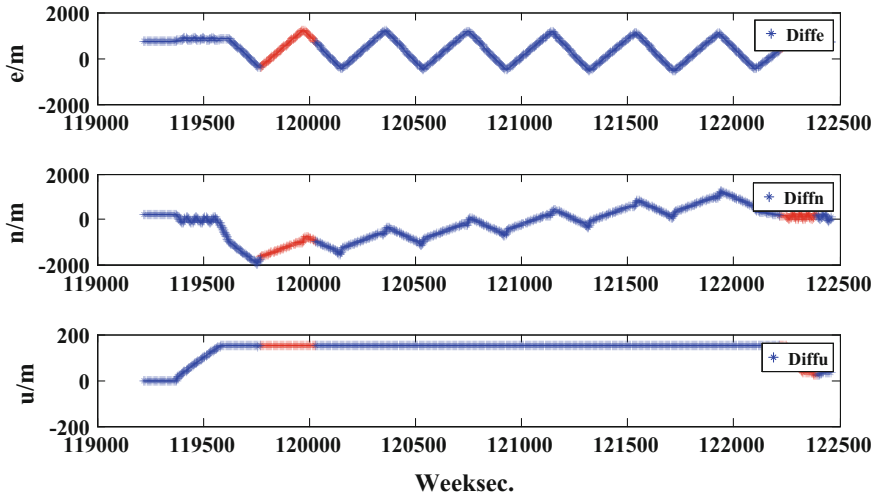


Fig. 5 Original method result

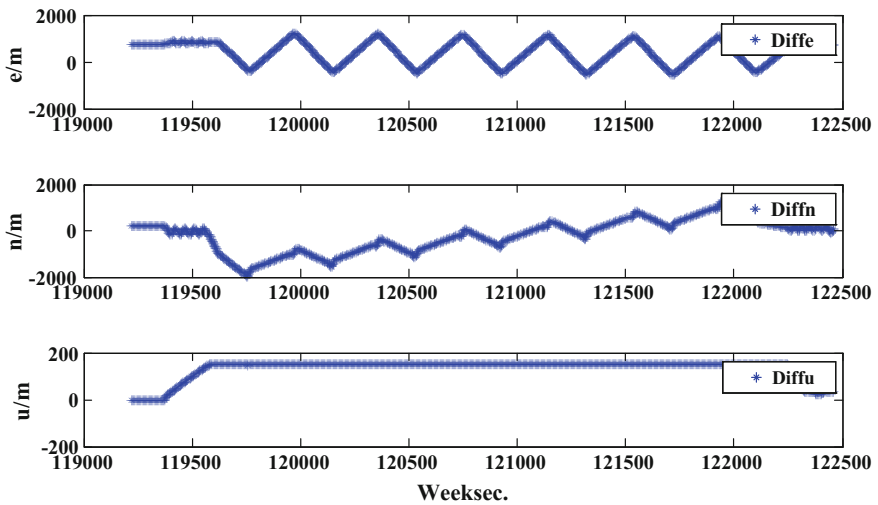


Fig. 6 Multi-gross errors detection method result

4.2 Experiment 2

The data is collected in September 8, 2016 with three systems, the interval is 0.2 s, the observation time is about 1.5 h long. For detailed analyzing the data quality, the figure below shows the number of satellites and PDOP values for base and rove stations. The figure shows the quality of the data is poor, the number of satellites for

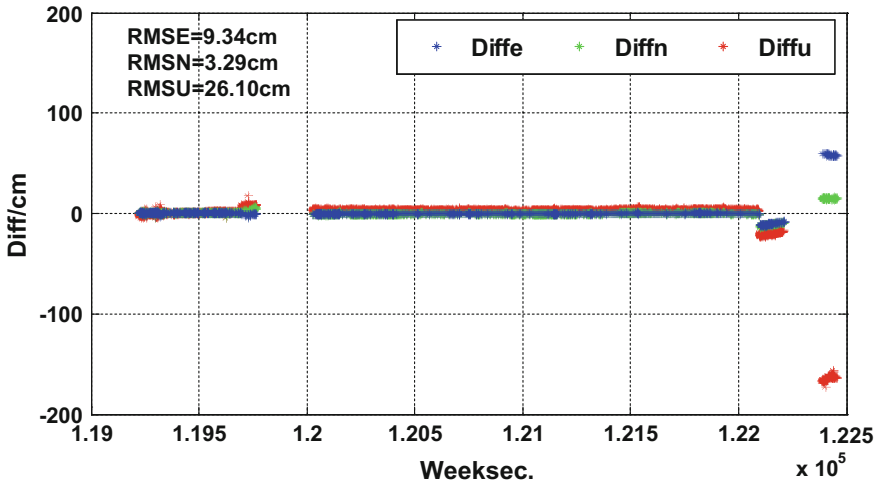


Fig. 7 Original method diff result

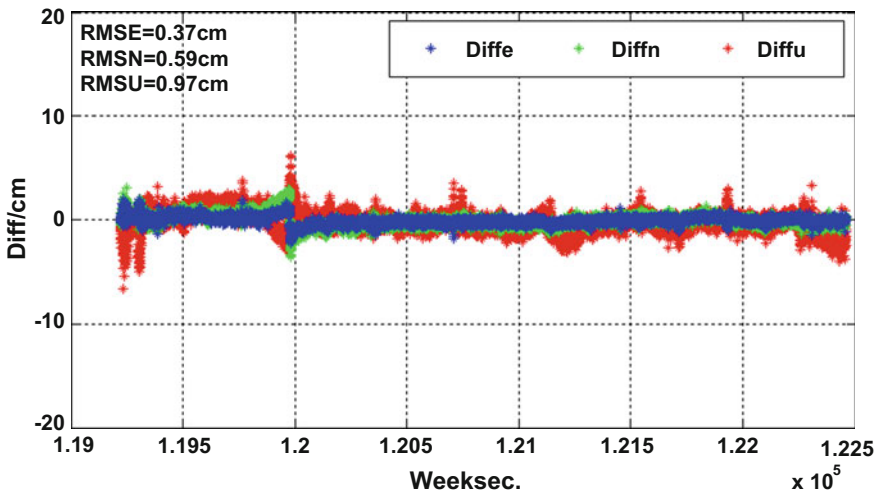


Fig. 8 Multi-gross errors detection method diff result

Table 1 Statistical results with different methods

Method	Fix rate (%)	RMSE (cm)	RMSN (cm)	RMSU (cm)
Original method	86.5	9.34	3.29	26.10
Multi-gross errors method	100	0.37	0.59	0.97

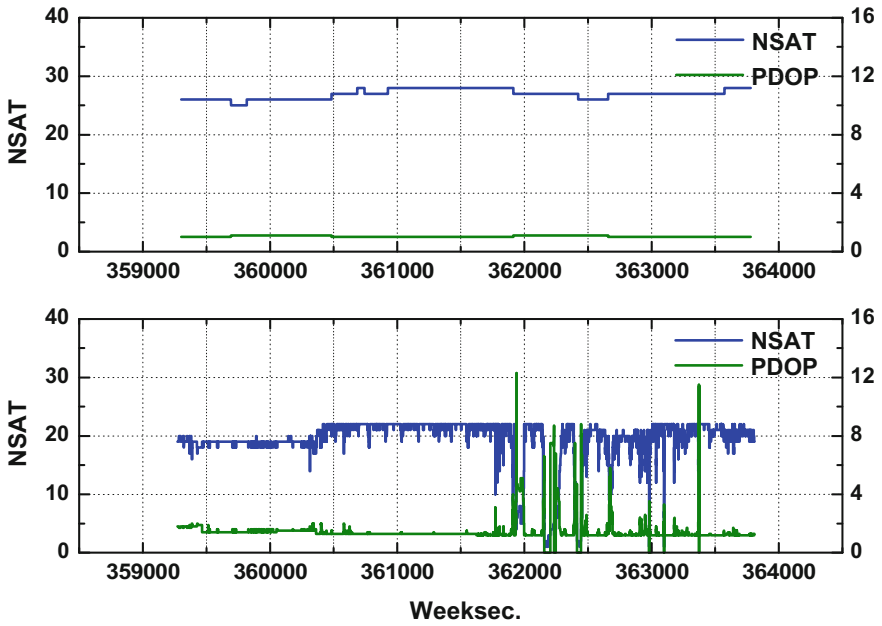


Fig. 9 NSAT and PDOP

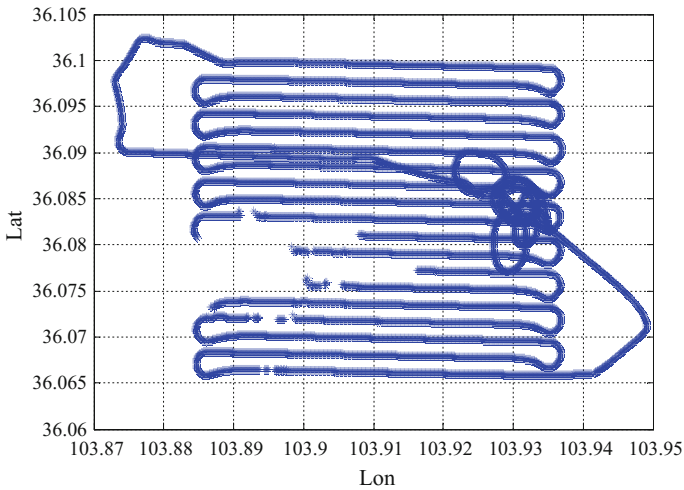


Fig. 10 Track of UAV

the rove station changes frequently, the satellites is less than 4 in some epoch. The filter is re-initialized frequently (Figs. 9 and 10).

The results with different methods show as follows (Figs. 11 and 12).

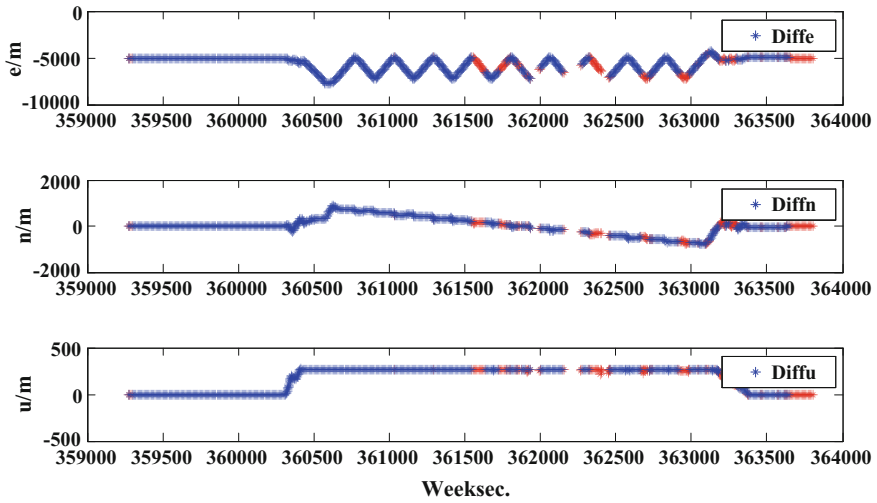


Fig. 11 Original method result

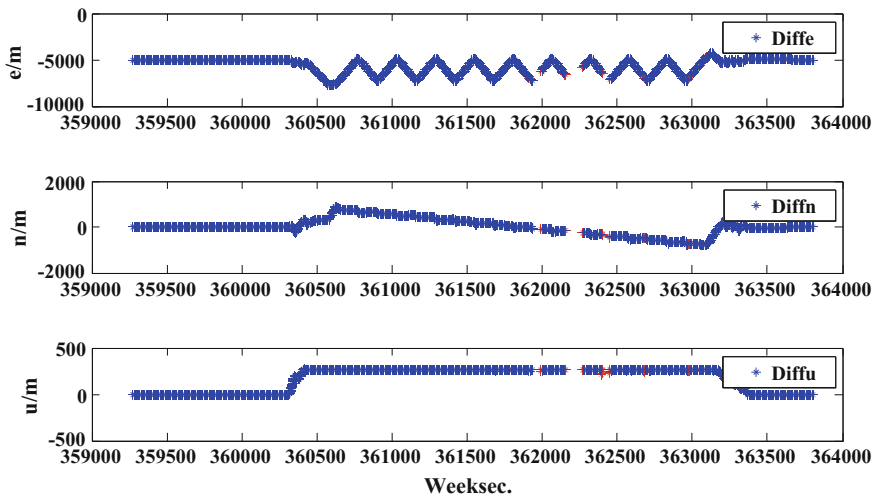


Fig. 12 Multi-gross errors detection method result

The sequence diagram of difference with GrafNav is as follows (Figs. 13 and 14). The specific statistics are as follows (Table 2).

The results of the data (Tables 1 and 2) shows that the fix rate of the original method is significant decrease when there are some undetected cycle slips and gross errors, and it's easy to fix error (Fig. 7), the accuracy and reliability of UAV-RTK will be badly influenced, even the security can't be guaranteed. But by the

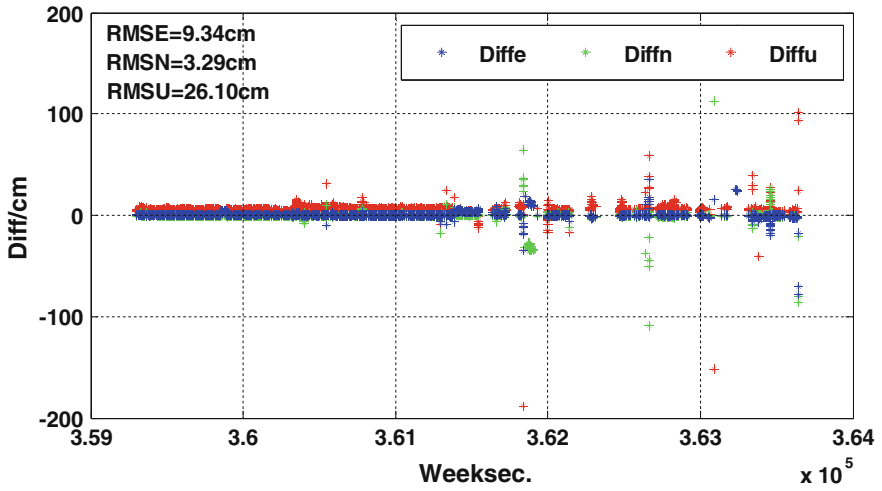


Fig. 13 Original method diff result

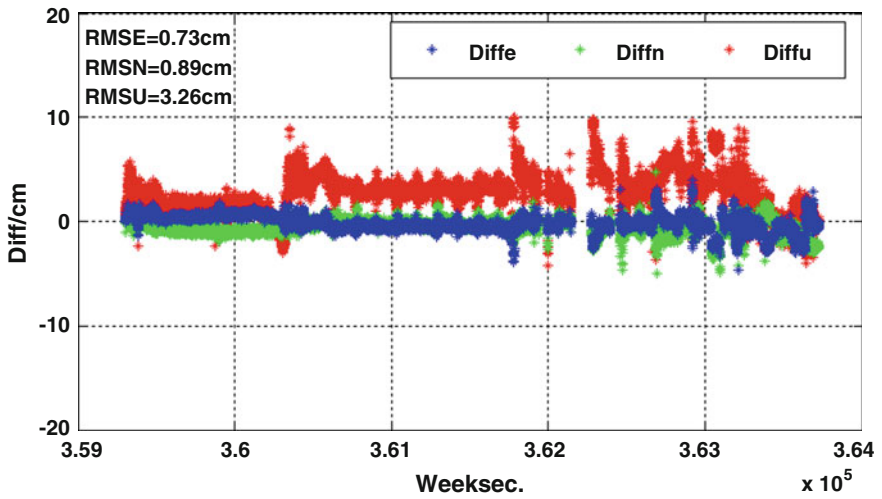


Fig. 14 Multi-gross errors detection method diff result

Table 2 Statistical results with different methods

Method	Fix rate (%)	RMSE (cm)	RMSN (cm)	RMSU (cm)
Original method	82.4	2.21	7.53	70.16
Multi-gross errors method	97.5	0.73	0.89	3.26

multi-gross errors detection method with the epoch variation which is put forward in this paper, the fix rate and the reliability can be significantly improved, avoid the appearance of outliers, Finally, the reliable and precise positioning information can be continuously supplied for UAV RTK.

5 Conclusions

A new multi gross errors detection method with the epoch variation constraint is put forward in the paper which deals with multi cycle slips and multi gross errors in two steps. First, detecting the errors by the epoch-differenced method and then applying the Robust Kalman Filter with the epoch variation constraint.

The cycle slips and the gross errors are distinguished by the parameter characteristics and processed in a different way. As a result, this method can avoid the influence of the uninitialized ambiguity when cycle slips occur and frequent re-initialized gross errors. Moreover, the method mainly focuses on the complex and changeable dynamic environment of UAV RTK and it is suitable for single frequency or multi frequency data.

The experiment results show that the method can effectively detect the multi-cycle slips and the multi-gross errors, significantly improve the ambiguity fixing rate and mitigates the impacts of outliers. Finally, the reliable and precise positioning information can be continuously supplied for UAV RTK.

Acknowledgements This work was supported by the National key Research Program of China “Collaborative Precision Positioning Project” (No. 2016YFB0501900) and The Special Project of Unmanned Intelligence Technology for Scientific and Technological Development of Guangdong Province “Research and Application of the Key Technology for Precise and Efficient Aerial Survey in UAV” (No. 2017B010117006).

References

1. Li Z, Huang J (2005) GPS surveying and data processing. WhuHan University Press, WhuHan
2. Wang J, Li C, Zhou M (2011) Application of real time kinematic technique in unmanned-aerial-vehicle image processing. In: China satellite navigation conference, CSNC 2011
3. Lu L, Geng C, Bian Y (2017) Application of agricultural plant protection UAV based on BDS-RTK. In: China satellite navigation conference, CSNC 2017
4. Sun G (2016) The application of RTK in aerial photography of low altitude unmanned aerial vehicle. Eng Technol Mag 8:00293
5. Cai C, Liu Z, Xia P et al (2013) Cycle slip detection and repair for undifferenced GPS observations under high iono-spheric activity. GPS Solutions 17(2):247–260
6. Blewitt G (1990) An automatic editing algorithm for GPS data. Geophys Res Lett 17(3): 199–202

7. Liu XD, Song LJ, Huang LY (2011) Outlier detection in detecting cycle slips based on epoch-difference location model. *Sci Surv Mapp* 36(6):58–60
8. Zhou J (1989) Classical theory of errors and robustestimation. *Acta Geod Cartograph Sin* 18 (2):115–120
9. Zhou J, Yang Y, Ou J (1992) Robustestimation paper set. Projects funded by the National Natural Science Foundation
10. Guo F (2013) Theory and methodology of quality control and quality analysis for GPS precise point positioning. WhuHan University
11. Liu XD, Song LJ, Huang LY (2011) Outlier detection in detecting cycle slips based on epoch-difference location model. *Sci Surv Mapp* 36(6):58–60
12. Chen Y, Yi C (2017) A new method to detect and repair cycle-slip based on a double-differenced model for receivers and epochs. *Geomat Inform Sci Wuhan Univ* 42 (6):845–850

Research on Multi-satellites Fusion Inversion Model of Soil Moisture Based on Sliding Window



Yueji Liang, Chao Ren and Yibang Huang

Abstract Global Positioning System Interferometric Reflectometry (GPS-IR) is a new remote sensing technique, and it can be used to estimate near surface soil moisture from signal-to-noise ratio data. Considering the advantages of multi-satellite convergence and the time and space scale of soil moisture, a multi-satellites fusion inversion model of soil moisture based on sliding window is proposed. Firstly, the direct and reflection signals of GPS satellites are separated by a low order polynomial fitting, and then the sinusoidal fitting model of reflection signals is established, the relative phase of the SNR interferogram is obtained. Finally, a sliding window is established to select the relative phases of the satellites effectively, and the best input variables of the inversion model are determined. Then, the soil moisture is retrieved by least square support vector machine. Based on the monitoring data provided by US Plate Boundary Observations Project, the feasibility and effectiveness of using single and different GPS satellites to sliding estimate soil moisture are compared and analyzed. The experimental results show that sliding window can effectively select GPS satellites and determine the best input variables for the model. The model fully takes the advantage of non-linear weight determination. The sliding type can achieve a longer time inversion. To a certain extent, this model improves the phenomenon that the inversion process is prone to jump when using single satellite inversion. Based on the inversion of more than four satellites, RMSE and MAE is less than 0.07 and 0.06, respectively, and R^2 is increased by at least 21.9% compared with inversion result of a single satellite. Therefore, the inversion of soil moisture can be treated as nonlinear event, and multi-satellites fusion inversion is feasible and effective.

Y. Liang (✉) · C. Ren · Y. Huang
College of Geomatics and Geoinformation,
Guilin University of Technology, Guilin, China
e-mail: lyjayq@glut.edu.cn

Y. Liang · C. Ren
Guangxi Key Laboratory of Spatial Information and Geomatics,
Research Center of Precise Engineering Surveying, Guilin, China

Keywords GPS-IR · Soil moisture · Multi-satellites fusion · Sliding window
Least squares support vector machine

1 Introduction

Soil moisture is an important parameter in the research of hydrology, meteorology and the agricultural environment, and it is of great significance on climate prediction, flood disasters and the water resource cycle [1]. Relative to the traditional methods of drying and weighing, time domain reflectometry, frequency domain reflectometry and other soil moisture monitoring methods [2], Global Positioning System Interferometric Reflectometry (GPS-IR) as a new means of remote sensing monitoring, with high efficiency, high resolution and other characteristics, to achieve non-contact, a wide range of soil moisture detection. GNSS-IR uses a single antenna for receiving direct and reflected signals at the same time. The soil moisture estimation is realized by using the interference of two signals at the antenna; it is considered to be the ideal method of detecting soil moisture. At present, a GNSS multipath signal has been used to estimate the environmental parameters and their changes of station, such as snow depth, seawater height, soil moisture and vegetation water [3]. According to the relationship between the soil dielectric constant and soil moisture, a simplified empirical or semi-empirical soil moisture model has been proposed to verify the effectiveness of soil moisture estimation in different environments [4]. Because these methods require numerous parameters and the modelling is complex, there are some deficiencies such as the need for manual measurement of data and the weak generalization property of the model. The inversion of the trend of soil moisture has been achieved by using the amplitude, frequency and phase of the multipath reflex component at a certain cut-off satellite height angle [5]. Moreover, the relationship between the reflected component and the reflective environment is studied and the phase observations are corrected to achieve better results [6]. Estimation of soil moisture content using delayed phase parameters is more stable than that of amplitude parameter [7]. The existing researches show that not all GPS satellite signals are suitable for inversion of soil moisture. Moreover, the effects of soil surface roughness and vegetation information on GPS signals are different, so it is more difficult to establish an accurate soil moisture inversion model directly. The existing researches are less concerned with the advantages of multi-satellite fusion estimation, and the estimation of soil moisture is affected by artificial interference; it is not conducive to the improvement of estimation accuracy. If soil moisture is changed with time scale and spatial scale, the estimation of soil moisture using the GNSS-IR technique can be regarded as a nonlinear regression problem. It is possible to establish a nonlinear estimation model of soil moisture through the phase of the SNR interferogram.

Based on the above-mentioned research, a method for estimating soil moisture is presented, and a nonlinear inversion method of multi-satellite fusion is used based on sliding window. The least squares support vector machine (LS-SVM) is

introduced into the soil moisture inversion, and the relative phase of each GPS satellite is analyzed by sliding window. Then the effective satellites are screened through the sliding window and the best input variables of the LS-SVM model are determined. The sliding is used to invert soil moisture. Based on the monitoring data provided by the PBO network, the feasibility of using single or different satellites to estimate soil moisture is compared and analysed, and the effectiveness of the sliding type estimation is studied and verified.

2 The Multi-satellites Fusion Inversion Model of Soil Moisture Based on Sliding Window

2.1 The Theory of Satellite Signal Reflection

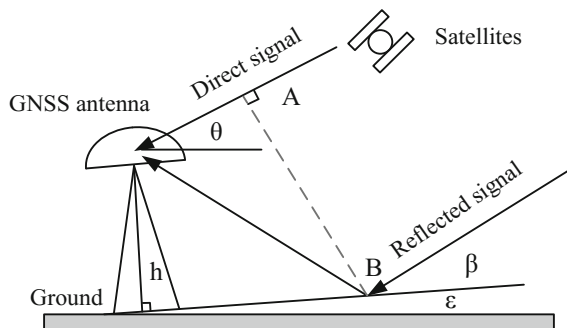
The signal-to-noise ratio (SNR) is an indicator of receiver antenna signal quality, and it is mainly affected by the antenna gain parameter, the multi-path effect and random noise in the receiver [1]. In the case of a low satellite elevation angle, the multi-path effect on SNR is more obvious. The signal received by the GNSS antenna is a composite signal of direct signal and reflected signal (Fig. 1).

From Fig. 1, θ is the satellite elevation angle of the direct signal transmitted from the satellite; h is the vertical height of the antenna to the ground; ε is the incline angle of the slope; and β is the angle between the satellite signal and the slope. When ε is small, the formula can be expressed as follows:

$$\theta = \beta + \varepsilon \approx \beta \tag{1}$$

There is a sine or cosine relationship between SNR and ψ . After removing the direct component, there is still a sinusoidal (or cosine) relationship with a fixed frequency between the multipath reflex component and $\sin \theta$ [8]:

Fig. 1 Geometric model of ground multi-path errors



$$SNR_{MP2} = A_{MP2} \cos\left(\frac{4\pi H}{\lambda} \sin \theta + \phi_{MP2}\right) \quad (2)$$

where θ , λ and H represent the satellite elevation angle, wavelength of carrier wave and GNSS antennas height, respectively; A_{MP2} is the relative amplitude of the multipath reflection components, and ϕ_{MP2} is the relative phase delay. A_{MP2} and ϕ_{MP2} are the feature parameters needed.

The relative phase delay is the best measure to estimate a change in soil moisture, and there is a theoretical linear relationship between relative phase delay and soil moisture [9]. Therefore, the estimation model of soil moisture is built based on the relative phase delay of multiple satellites in this paper.

2.1.1 The Principle of Model Inversion

(1) LS-SVM inversion model

Suppose the relative phase delay set of GNSS satellites is x , as follows:

$$\begin{aligned} X &= [x_1^0, x_2^0, \dots, x_i^0], \quad i = (1, 2, \dots, t) \\ x_i^0 &= [\phi_{MP2}^1, \phi_{MP2}^2, \dots, \phi_{MP2}^j], \quad (j = 1, 2, \dots, 32) \end{aligned} \quad (3)$$

where i is the day of year, j is the satellite vehicle's ID, and t is the length of time.

Suppose the soil moisture set corresponding to the relative phase delay set x is y . Since the relationship between input vector x and soil moisture y is non-linear. Therefore, it is necessary to map the samples in the input space to a higher dimension feature space through a nonlinear transformation $\varphi(x)$; thus, the non-linear problem in input space is transformed into a linear problem in feature space. Then, a linear LS-SVM is used to fit the sample points in the feature space. Detailed process can be reference [10].

The performance of the LS-SVM depends largely on the optimal choice of kernel function $k(\cdot)$, kernel parameter σ and regularization parameter γ . Because the radial basis function can well describe the complexity of the model and has good universality, it is selected as the kernel function of the LS-SVM. At the same time, the grid search method is adopted for parameter optimization of σ and γ , its basic principle can be reviewed in reference [11].

(2) The sliding window

Taking into account the soil moisture has a certain time and space characteristics, to a certain extent; the model input variables affect the performance of the model. The principle of the sliding window is mainly to conduct correlation R^2 analysis through linear regression models, and set thresholds, select valid satellites and determine the best input variables of the model. Assuming that the selected first x period is the

initial training sample of the LS-SVM model, and the relative phase of the period m is the longest variable of the model input. To set the sliding window decision length, the input variables x_{phi} is that:

$$x_{\text{phi}} = \{x_{m-n*b+1}, \dots, x_{m-2}, x_{m-1}, x_m\} \tag{4}$$

where $n = 1, 2, \dots, t$, $n * b \leq m$. Determine the correlation between each model input variable set and soil moisture by each judge length b , and select the relative phase of each satellite with good correlation in the same period as the best input variable.

2.1.2 The Process of Soil Moisture Inversion

The modeling process is as follows:

- (1) Separating multi-path reflection components from the GNSS signal. Using the TEQC software to solve the GNSS monitoring data, the SNR value is obtained, and the direct and reflected components are separated by low order polynomial fitting;
- (2) The resampling of multipath reflection components. Since the multipath reflex component is a value that varies with the epochs, it is necessary to convert it to the relationship between the sinusoidal values of the incident height angle by resampling;
- (3) Estimation of parameters. Fitting a sine to the component after resampling by a nonlinear least squares fitting algorithm [12], then the values of amplitude and phase could be obtained.
- (4) Selecting input variables of model and sliding inversion. Let the sliding window further determine the selected x_{phi} as the input sample and the inversion step as r to establish the LSSVM inversion model: Suppose the initial input training sample 1 is $x_i = \{x_{\text{phi}}\}$, the test input sample 1 is $y_i = x_{m+1}, x_{m+2}, \dots, x_{m+r}$; then the input training sample 2 is $x_{i+1} = \{x_{\text{phi}}, y_i\}$, the test input sample 2 is $y_{i+1} = x_{m+r+1}, x_{m+r+2}, \dots, x_{m+2r}$, and so on. The process is shown in Fig. 2.

3 Experiment Analysis

GNSS monitoring data of the P041 station, which is from the PBO network (<http://xenon.colorado.edu/portal>), is selected for the experiment. The station is located at Newcastle in Cheyenne, Wyoming, at 104.185702°W, 43.881146°N. From the LANDSAT TM multispectral image of the United States in Fig. 3, the terrain around this station is relatively flat, open and sparsely populated with vegetation

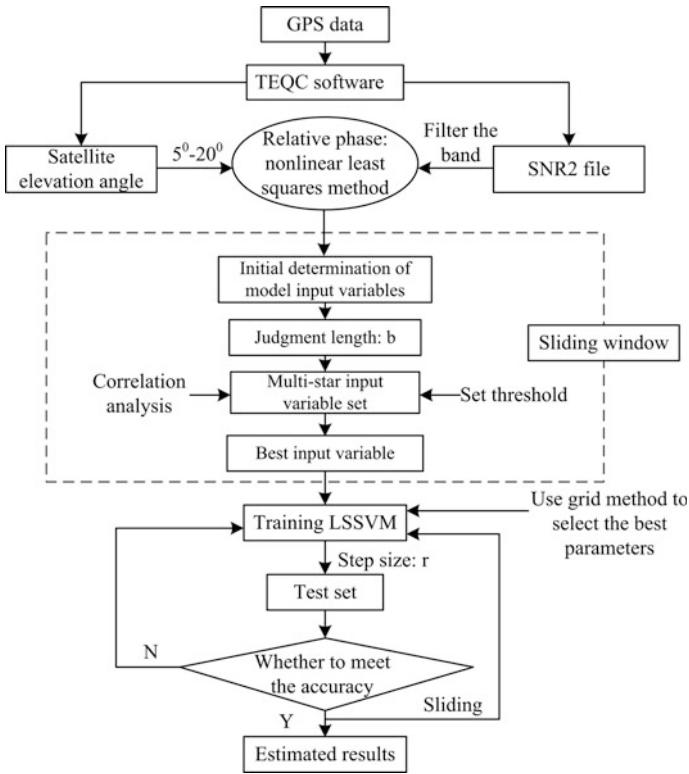


Fig. 2 The rolling estimation process of soil moisture

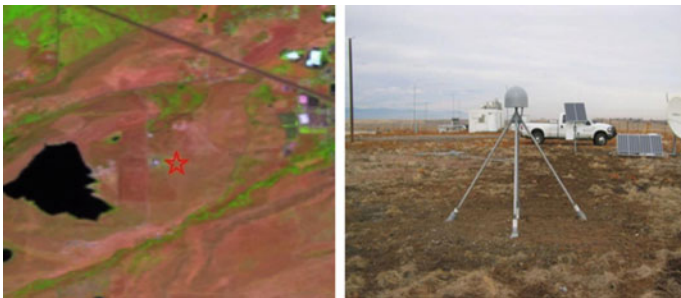


Fig. 3 The surface environment of P041 station

conductive. The change in soil moisture and rainfall at the site in the 70th–290th days of 2011 is shown in Fig. 4. Rainfall is dominated by products from the North American Land Data Assimilation System (NLDAS). Soil moisture content (SMC) is retrieved using L2C observations, based on averages of eight or more

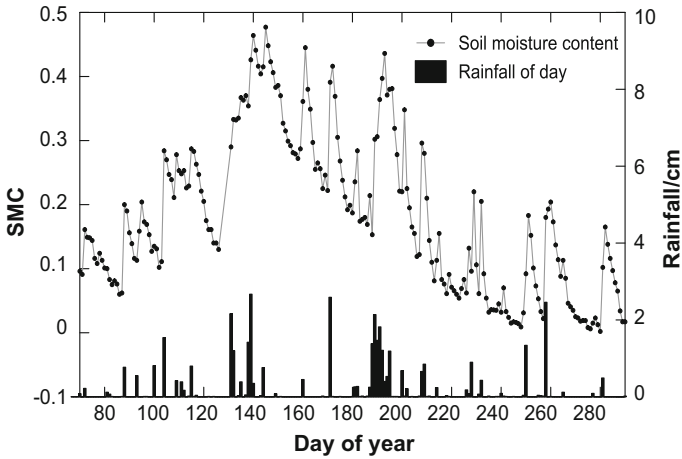


Fig. 4 Soil moisture and rainfall

satellite orbits [8, 9]. For reflection information at low satellite elevation angles, the sensing depth of the L2C signal to the soil is about 2.5 cm or less [8]. As the PBO antenna is highly inclined to be between 1.5 and 2.1 m, the sensing area for each satellite orbit is about 120 m² [13].

Figure 4 shows that during this time, there are 13 significant rainfalls, the maximum rainfall reached 2.67 cm and the continuous rainfall up to 10 days. Corresponding to precipitation, soil moisture are significantly increased, especially 131–132, 137–140 and 186–195 days. Due to the continuous rainfall, the soil moisture changes more significantly and shows strong non-linearity and randomness. When the rainfall stops, soil moisture decreases gradually. It can be seen that the site is rich in rainfall and is more representative for the study of soil moisture.

Therefore, the monitoring data from the 70st-290th day of 2011 are selected, the sampling rate is 30 Hz, and the satellite elevation is between 5° and 20°. The SNR (L2 carrier) is obtained by using TEQC to calculate the GNSS data, and separating the direct signal and reflection signal of each satellite by a quadratic polynomial separation. Then, the nonlinear least squares method is used to estimate the phase of SNR interferogram. Due to space limitations, only 10 satellites in the effective wavelength band are selected for analysis in the experiment. The relative phase of each satellite and the change of soil moisture are shown in Fig. 5.

Figure 5 shows that the relative phase of the satellites can respond when the soil humidity rises or falls. For 131–132, 137–140 and 186–195 days, the relative phase of the satellites show a large floating, and it was related to the sharp increase of soil moisture caused by continuous rainfall. The relative phase of the same GPS satellite tends to jump, and the response patterns of different satellites to soil moisture changes are not consistent. The geometric movement of the satellite with respect to the GPS antenna and the performance of the satellite itself during the observation is

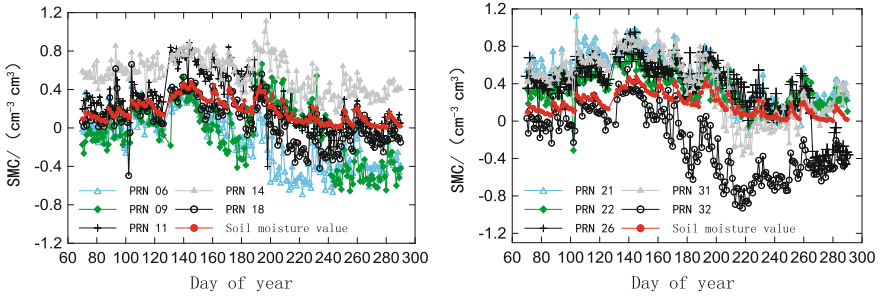


Fig. 5 Relationship between delayed phase and soil moisture

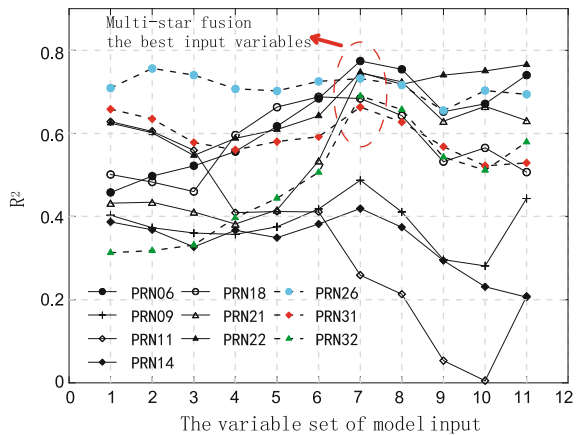
Table 1 Input variables set of LS-SVM model

Variable set	1	2	3	4	5	6
Day of year	70–189	80–189	90–189	100–189	110–189	120–189
Variable set	7	8	9	10	11	12
Day of year	130–189	140–189	150–189	160–189	170–189	180–189

not the same. Therefore, how to choose the relative phase of satellites effectively is worth discussing.

Considering the effective selection of satellites and the input variables of the model, a sliding window is proposed. The relative phases of the first 120 days of each satellite are chosen as the input variable of the LS-SVM model. The sliding window determines the length, and 12 sets of input variables are obtained, as shown in Table 1. Figure 6 shows the R^2 between different input variable sets for each satellite and soil moisture. The threshold for R^2 is set at 0.60. The PRN06, 18, 21, 22, 26, 31 and 32 satellites are preferred (as indicated by the dotted red line) and the

Fig. 6 Comparison of different input variable sets



relative phases of days 130–189 of these satellites are selected for modeling training.

In order to verify the feasibility and effectiveness of multi-satellite fusion inversion, two experimental schemes were set up: sliding inversion based on single satellite and sliding inversion based on multi-satellite fusion. To reduce the modelling error, each phase is pre-treated and normalized to $[-1, 1]$, then reduced to the original interval after the estimation. The test samples are taken from the 130th–189th days, the 190th–290th days are selected as the training samples, and the estimated step size was 1. For example, when the inversion step is 1, the 130th–189th days are selected for modeling to invert the 190th day. Then, the 130th–190th days are selected for modeling to invert the 191th day, and so on, until the 290th day is inverted. The error of model inversion is shown in Fig. 7.

Figure 7 shows that using a single satellite for soil moisture inversion makes it difficult to accurately grasp the variation of soil moisture. Moreover, the variation of inversion errors is large and the inversion result is prone to jump (blue dotted line). Comparing and analyzing the results of two- satellite, three-satellite or four-satellite inversion, the improvement is not significant. However, using more than five satellites to establish the inversion model, the fluctuation of inversion error is relatively close, and can effectively inhibit the phenomenon of jump. Visible, with the increase of the number of satellites, the inversion error is more stable.

In order to further comprehensively evaluate the performance of each scheme, R^2 , root mean square error (RMSE), average absolute error (MAE) and maximum error (MAX) are used for evaluation (As shown in Fig. 8, a total of 28 kinds of

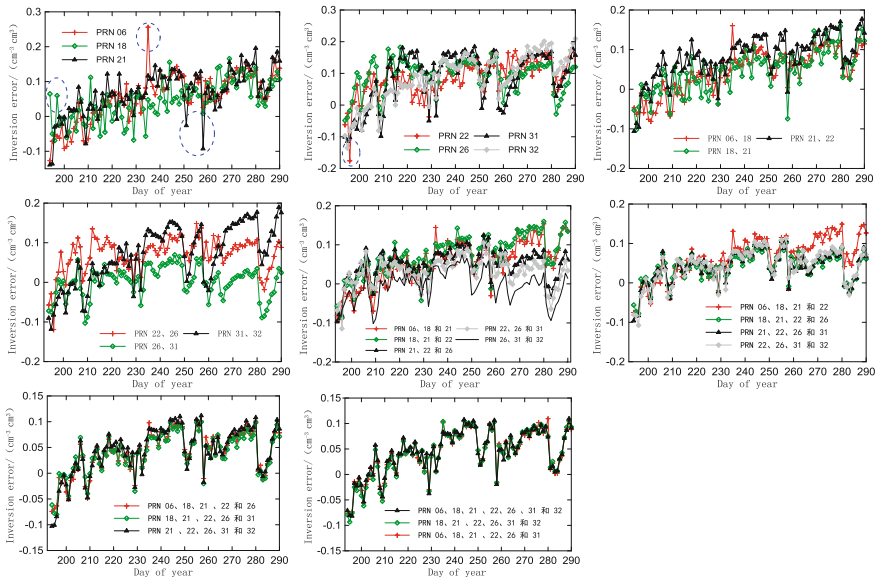


Fig. 7 Estimation error analysis of soil moisture

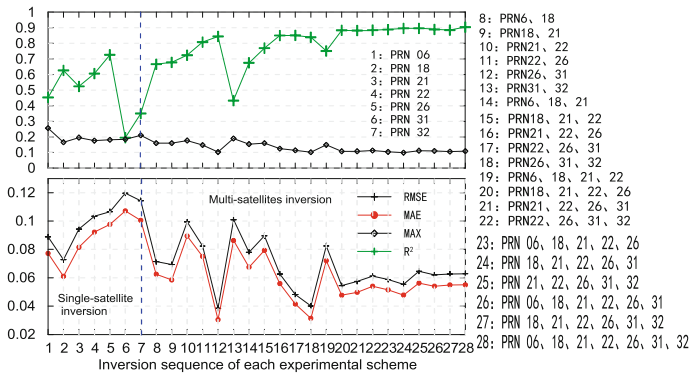


Fig. 8 Retrieval accuracy of soil moisture in each model

schemes). According to Figs. 7 and 8, it is found that the number of fusion satellites increased appropriately, and all the accuracy indexes improved. When the number of fusion satellites is more than 4, RMSE and MAE are less than 0.07 and 0.06 respectively, the maximum error of inversion is also improved, and R^2 is increased by at least 21.9% compared with that of single satellite.

4 Conclusion

This paper presents a multi-satellite fusion LS-SVM retrieval method of soil moisture based on sliding window.

The experiments show that: (1) The GNSS antenna is affected by the kinematical track and the performance of satellites during observation, and the relative phase of different satellites have different response patterns to changes in soil moisture. The sliding window can effectively select the relative phase of each satellite and determine the best input variables of the model. (2) The multi-satellite fusion inversion has more advantages than the single satellite. As the number of fusion satellites increases, the inversion error is more stable. The model gives full play to the advantages of artificial intelligence in soil moisture. It can effectively integrate the performance of the satellites, and realize the complementarities between the relative phases of the satellites through non-linear weighting. (3) When the number of fusion satellites is more than 4, RMSE and MAE is less than 0.07 and 0.06, respectively, and R^2 is at least 21.9% higher than that of a single satellite. In order to further verify the performance of this model, the data from multiple stations in different environments and the integration of vegetation, temperature, rainfall and other variables will be studied next experiment.

Acknowledgements This work was sponsored by the National Natural Foundation of China (41461089,41541032,41664002); The Natural Science Foundation of Guangxi (2014GXNSFAA1182-88); Guangxi Key Laboratory of Spatial Information and Geomatics (16-380-25-22); Basic Ability Improvement Project for Young and Middle-Aged Teachers in Universities in Guangxi (2018KY0247).

References

1. Sabater JM, Rüdiger C, Calvet JC, Fritz N, Jarlan L, Kerr Y (2008) Joint assimilation of surface soil moisture and LAI observations into a land surface model. *Agric Forest Meteorol* 148(8):1362–1373. <https://doi.org/10.1016/j.agrformet.2008.04.003>
2. Le Hégarat-Masclé S, Zribi M, Alem F, Weisse A, Loumagne C (2002) Soil moisture estimation from ERS/SAR data: Toward an operational methodology. *IEEE Trans Geosci Remote Sens* 40(12):2647–2658. <https://doi.org/10.1109/TGRS.2002.806994>
3. Hallikainen MT, Ulaby FT, Dobson MC et al (1985) Microwave dielectric behavior of wet soil-part I: Empirical models and experimental observations. *IEEE Trans Geosci Remote Sens* 1:25–34. <https://doi.org/10.1109/TGRS.1985.289497>
4. Wang Q, Guo Z, Sun G, Luo C, Liu D (2010) Construction of MGeoSALL: a hybrid BRDF model for discontinuous vegetation canopies
5. Ao S, Zhu J, Hu Y et al (2015) GPS soil moisture monitoring using SNR observations. *Geomatics Inf Sci Wuhan Univ* 40(01):117–120
6. Bilich A, Larson KM (2007) Mapping the GPS multipath environment using the signal-to-noise ratio (SNR). *Radio Sci* 42(6). <https://doi.org/10.1029/2007rs003652>
7. Zavorotny VU, Larson KM, Braun JJ, Small EE, Gutmann ED, Bilich AL (2010) A physical model for GPS multipath caused by land reflections: toward bare soil moisture retrievals. *IEEE J Select Topics Appl Earth Observ Remote Sens* 3(1):100–110. <https://doi.org/10.1109/JSTARS.2009.2033608>
8. Chew CC, Small EE, Larson KM, Zavorotny VU (2015) Vegetation sensing using GPS-interferometric reflectometry: theoretical effects of canopy parameters on signal-to-noise ratio data. *IEEE Trans Geosci Remote Sens* 53(5):2755–2764
9. Chew C, Small EE, Larson KM (2016) An algorithm for soil moisture estimation using GPS-interferometric reflectometry for bare and vegetated soil. *GPS Solutions* 20(3):525–537. <https://doi.org/10.1007/s10291-015-0462-4>
10. Suykens JA, Vandewalle J (1999) Least squares support vector machine classifiers. *Neural Process Lett* 9(3):293–300. <https://doi.org/10.1023/A:1018628609742>
11. Liu X, Jia D, Li H, Jiang J (2010) Research on Kernel parameter optimization of support vector machine in speaker recognition. *Sci Technol Eng* 10(7):1669–1673
12. Johnson ML, Correia JJ, Yphantis DA, HalvorSON HR (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys J* 36(3):575–588. [https://doi.org/10.1016/S0006-3495\(81\)84753-4](https://doi.org/10.1016/S0006-3495(81)84753-4)
13. Larson KM, Nievinski FG (2013) GPS snow sensing: results from the earth scope plate boundary observatory. *GPS Solutions* 17(1):41–52. <https://doi.org/10.1007/s10291-012-0259-7>

A Study of TEC Storm on 13 October 2016



Qingtao Wan, Jinghua Li, Xiaolan Wang, Jiangtao Fan, Jie Zhang
and Guanyi Ma

Abstract An ionospheric positive storm on 13 October 2016 was studied. The longitude effect of ionospheric storm along 40°N latitude region is studied with 23 GPS stations from 0°E to 360°E. Total electron content (TEC) was calculated with polynomial method, and TEC difference between the storm time and quiet time (Δ TEC) were calculated. The variations of TEC difference were different at different longitudes. There was a positive phase storm at the local time day side region, while the variations were not obvious at the night side longitudes. The maximum value of Δ TEC was about 25 TECU at 330°E at about 1400UT on 13 October 2016. Two Ionosonde data at EB040 (0.5°E, 40°N) and RL052 (51.5°N, 359.4°E) were also used to study the ionospheric storm. The Δ f_oF₂ of RL052 increased earlier than low latitude ionosonde of EB040. There was a second peak of Δ TEC at the local night time about 2000 LT on 13 October 2016. The latitude effect of ionospheric storm was also studied with 7 GPS stations at 0°E meridian. It was found that the positive storm was obvious at high latitude region, while it was weak at low latitude region. This phenomenon may be caused by the fountain effect. From high latitude to low latitude, Δ TEC gradually began to increase, and Δ TEC at higher latitude reached the maximum value earlier than the low latitude. These show that the positive ionospheric storm may be caused by the equatorward surge.

Keywords GPS station · Ionosonde · Total electron content · Equatorward surge

1 Introduction

The high speed solar wind from coronal mass ejection (CME) carries a lot of high-energy particles, and impacts the earth's magnetosphere, causing the earth's magnetic field to undergo dramatic changes, resulting in geomagnetic storms.

Q. Wan (✉) · J. Li · X. Wang · J. Fan · J. Zhang · G. Ma
Chinese Academy of Sciences, National Astronomical Observatories,
Beijing, China
e-mail: qtwan@nao.cas.cn

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018
Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_8

Magnetic storm is an extreme weather phenomenon. Ionospheric storms represent large global disturbance of electron density of ionosphere in response to geomagnetic storms. Increases and decreases in the ionospheric electron density during the disturbances are referred to as positive and negative storms phases, respectively [1].

An ionospheric storm consists of three phases. First, there is an initial phase that lasts for several hours after the magnetic storms starts. During the initial phase, the value of foF2 has increased relative to the quiet times. Then, it is a main phase that lasts for a long time, which from a few hours to a day for different storms. During the main phase, the foF2 is usually lower than the quiet times, which is a negative ionospheric storm. When the value of foF2 is greater than the quiet times, it is a positive ionospheric storm. After the main phase, the recovery phase is followed. At this time, the foF2 will gradually recover to the normal value of the quiet period. Usually, the recovery phase lasts a few days. In general, the positive ionospheric storms usually appear at low latitude and mid-latitude region in winter, while the negative storms are more likely at high latitude and mid-latitude in summer [2].

It is very useful to the effect of magnetic storms on the ionosphere from total electron content (TEC) data. The method of using GPS to derive TEC has the advantages of high resolution, wide space range and long duration. The use of GPS-TEC to study the ionosphere has been widely used [3].

The behaviour of TEC during ionospheric storms is basically similar to that of foF2, but it is not necessarily the same [4, 5]. There is a great advantage for combined use of GPS data and ionosonde data to study ionospheric storms.

An ionospheric storm is studied on 13 October in 2016. Section 2 describes the data used for the study. Section 3 is the results and discussions. Section 4 summarizes the results.

2 Data Used for the Study

The ionospheric storm on 13 October 2016 was studied using the data from 28 GPS stations. Figure 1 is the distribution of the GPS stations. 23 GPS stations are located at 37°N–43°N latitude region, which are used to study the longitude variation of TEC. The others are located at 0°E meridian, which are used to study the latitude variation of TEC. Ionosonde data at (0.5°E, 40°N) and (359.4°E, 51.5°N) were also used to study the storm. Figure 2 shows the interplanetary magnetic field and geomagnetic conditions during the period from 8 to 17 October 2016. The top panel shows the z-component of the interplanetary magnetic field observed by the ACE satellite. The second and third panel show the symmetric disturbance index (SYM-H) and the asymmetric disturbance index (ASY-H). The fourth panel is the Dst index. The bottom panel shows the Kp index.

The storm sudden commencement (SSC) was observed at 2100 UT on 12 October 2016. Then the Dst increased rapidly and reached the maximum value at about 2300 UT. The Dst started to decrease at 2300 UT, and reached about 0 at 0600 UT on 13 October 2016. The initial phase was from 2300 UT on 12 October

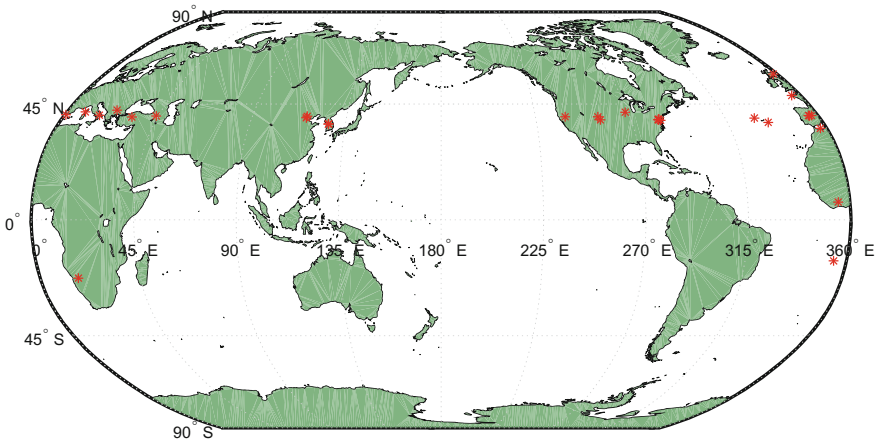


Fig. 1 The distribution of GPS stations along 40°N latitude and 0°E meridian lines

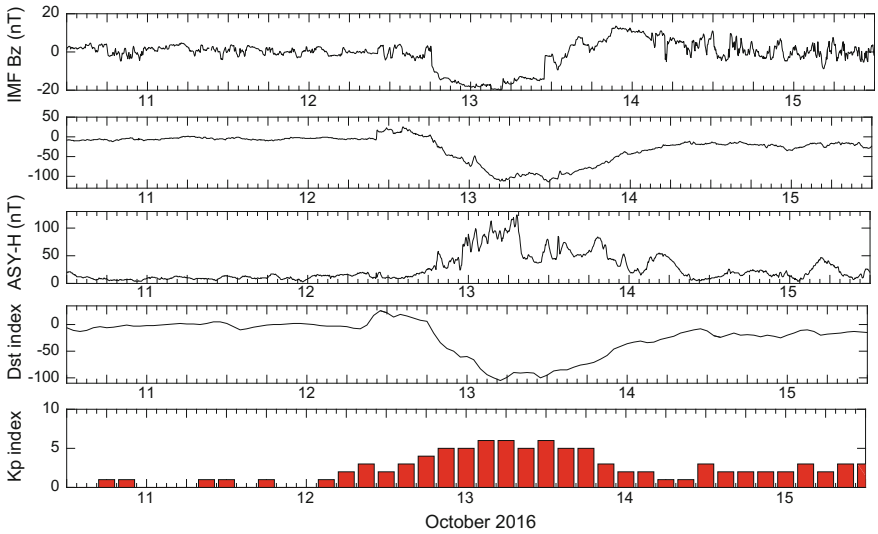


Fig. 2 The interplanetary and geomagnetic conditions during the geomagnetic storm period from 11 to 15 October in 2016. From top to bottom, they are IMF-Bz, SYM-H, ASY-H, Dst index and Kp index

to 0600 UT on 13 October. The Dst turned negative at 0600 UT and decreased rapidly. The minimum value of Dst was about -105 at about 1700 UT on 13 October 2016. The main phase lasted for 18 h. The Dst returned to normal at about 0000 UT on 15 October 2016.

3 Results and Discussions

The longitude effect of this ionospheric storm at 40°N latitude region is analyzed with 23 GPS stations. 7 GPS stations with different latitudes at 0°E meridian are used to study the latitude effect of this ionospheric storm. 2 ionosonde RL052 (51.5°N, 359.4°E) and EB040 (40.8°N, 0.5°E) are also used to study the latitude effect.

Figures 3 and 4 are the TEC and Δ TEC contour maps from 11 to 15 October in 2016. The polynomial method is used to calculate TEC. The reference value of TEC is the mean value for 4 quiet days from 8 to 11 October in 2016. The Δ TEC is the TEC difference between TEC at storm period and reference value. The solid line and dotted line are the local sunrise and sunset time, respectively.

It was shown that TEC increased rapidly at the storm time. The maximum value of TEC was about 40 TECU in the 330°E region on 13 October 2016. The positive storm was different at different longitude. The ionospheric storm is dependent on the local time. The positive storm was obvious at 0°E–45°E and 330°E–360°E, which are the local time day side region. It was not obvious at 100°E–130°E and 240°E–280°E. The TEC disturbance was obvious on the day side, while the disturbance was weak on the night side when the storm started.

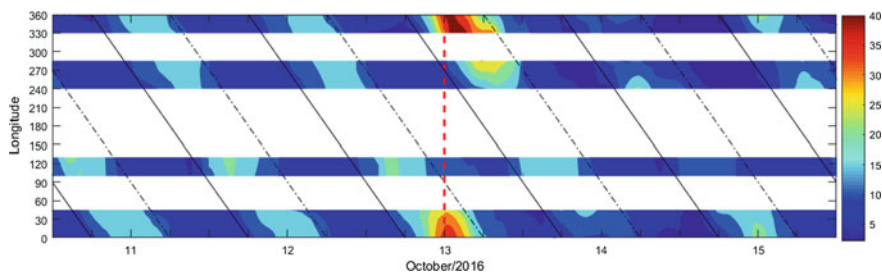


Fig. 3 The TEC contour map during the ionospheric storm period at about 40°N latitude region from 11 to 15 October in 2016

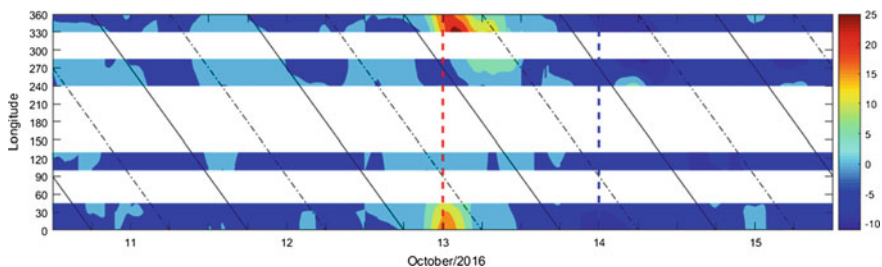


Fig. 4 The Δ TEC map during the ionospheric storm period at about 40°N latitude region from 11 to 15 October in 2016

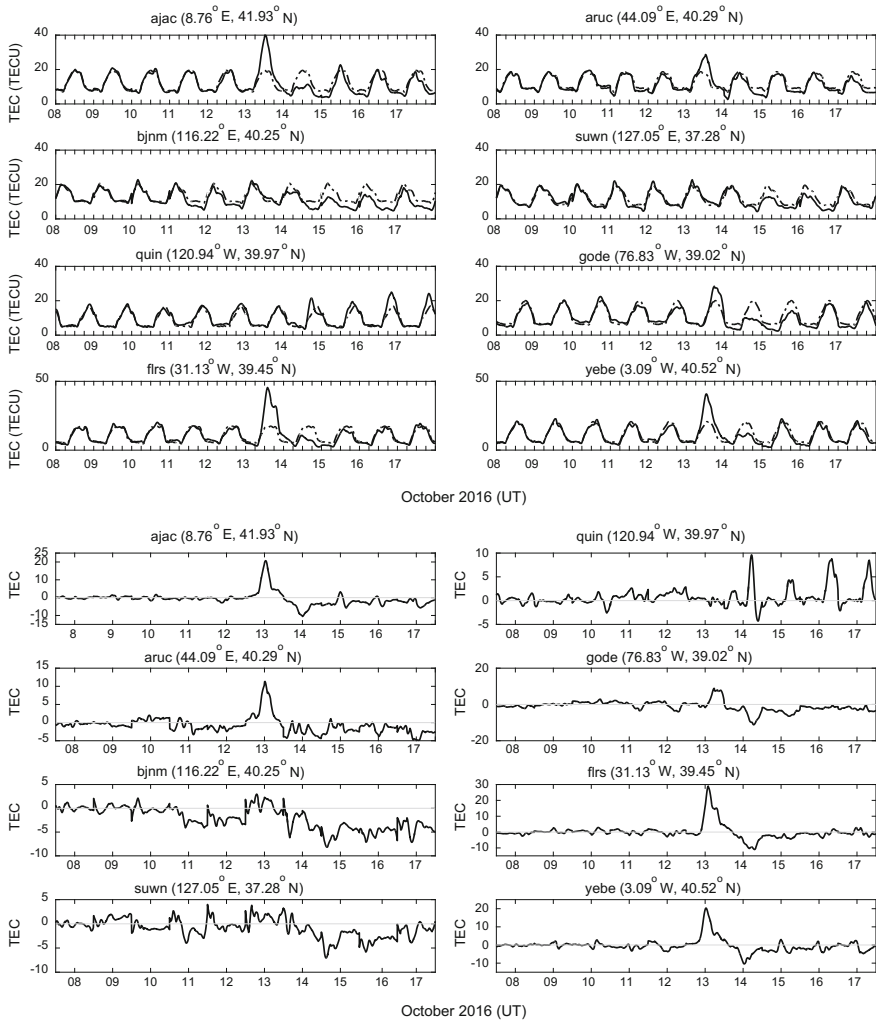


Fig. 5 The TEC and Δ TEC variations during the storm period at different longitudes

Figure 5 is about the Δ TEC variations during the storm period at different longitude GPS stations. It can be seen that positive storm is obviously at the local time day side region, while the Δ TEC variations are not obviously at the night side region.

Figure 6 is the ionosonde data (EB040, 0.5°E, 40.8°N) Δ hmF2, Δ foF2 and Δ TEC from 11 to 15 October 2016. Δ hmF2, Δ foF2 and Δ TEC have the same variation trends. Three parameters started to increase at about 0600 UT, which was the start time of main phase. The maximum value of three parameters occurred at

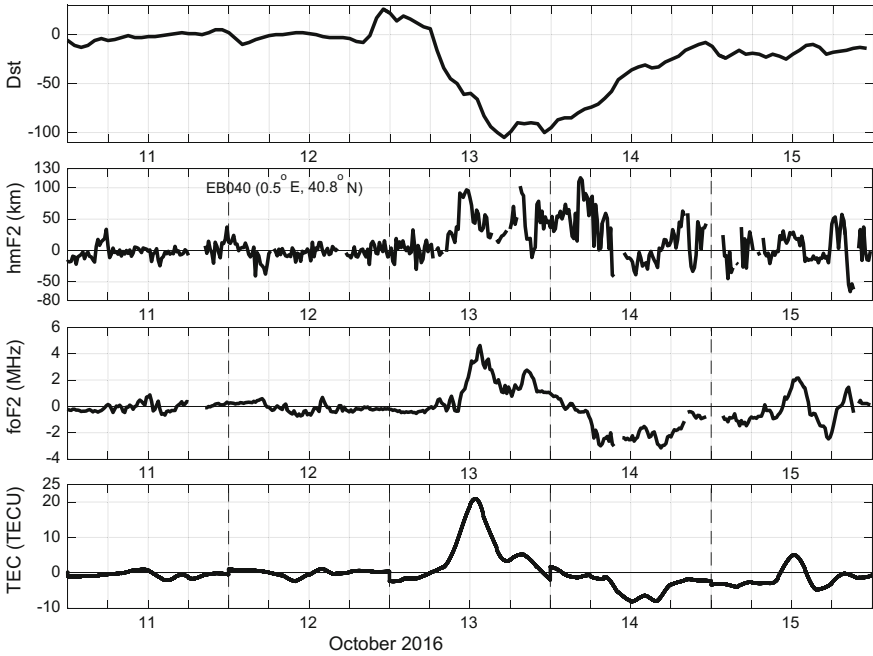


Fig. 6 The $\Delta hmF2$, $\Delta foF2$ variations at the storm period from the ionosonde data EB040 ($0.5^\circ E$, $40.8^\circ N$) and the ΔTEC variations at the same position

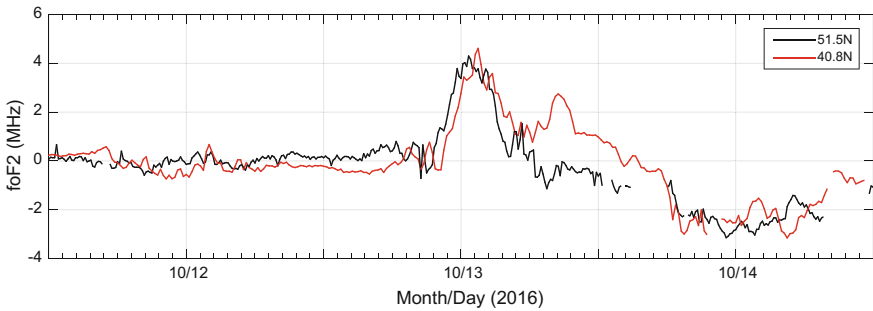


Fig. 7 The $\Delta foF2$ variations at the storm period from the ionosonde data EB040 ($0.5^\circ E$, $40.8^\circ N$) and the ΔTEC variations at the same position

about the noon time. There was a second peak of $\Delta hmF2$, $\Delta foF2$ and ΔTEC at the night time about 2000 UT.

In order to find the mechanism of the positive storm, the $\Delta foF2$ variations are studied with different latitudes' ionosonde data. Figure 7 is the $\Delta foF2$ variations from two ionosonde data RL052 ($51.5^\circ N$, $359.4^\circ E$) and EB040 ($40.8^\circ N$, $0.5^\circ E$) at the same meridian. There is a time delay for the $\Delta foF2$ variation between two

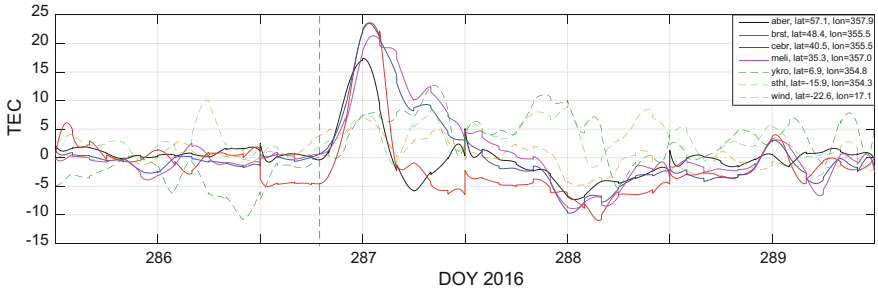


Fig. 8 The Δ TEC variations with different latitude at 0° E longitude region from 12 (DOY = 286) to 15 (DOY = 289) October in 2016

iononde at the ionospheric storm period on 13 October in 2016. The Δ foF2 of RL052 (51.5° N) increased earlier than EB040 (40.8° N) on 13 October 2016. The time difference is about 1 h, so the positive storm may be caused by equatorward surge.

Latitude effect of this positive storm is also studied. 7 GPS stations data, which are located at the different latitudes at 0° E meridian, are used to derive TEC at the storm period. Figure 8 is the Δ TEC variations of 7 stations at the 0° E meridian from 12 to 15 October in 2016. The positive storm was obvious in the mid-latitude region (35° N– 57° N), while the storm was weak in the low latitude region. This phenomenon may be caused by the fountain effect. From high latitude (57° N) to low latitude (35° N) Δ TEC gradually began to increase, and Δ TEC at higher latitude reached the maximum value earlier than the low latitude. It is also shows that the positive ionospheric storm may be caused by the equatorward surge.

4 Conclusions

An ionospheric storm is studied on 13 October 2016. The storm sudden commencement (SSC) was observed at 2100 UT on 12 October 2016. Then the Dst increased rapidly and reached the maximum value at about 2300 UT. The initial phase was from 2300 UT on 12 October to 0600 UT on 13 October. The main phase lasted for 18 h. 23 GPS stations are located at different longitudes along the 40° N latitude region are used to studied the longitude effect of TEC variation. The positive storm disturbance was obvious for the stations at the local time day side GPS stations, while it was weak in the night side region at about 40° N latitude region. The most obvious area of the positive ionospheric storm is about 330° E region. The latitude effect of this storm is also studied with 7 GPS stations data and 2 ionosonde data at 0° E meridian. The Δ foF2 of high latitude ionosonde increased earlier than the low latitude. From high latitude (57° N) to low latitude (35° N) Δ TEC gradually began to increase, and Δ TEC at higher latitude reached the

maximum value earlier than the low latitude. It was shown that the positive storm on 13 October 2016 may be caused by equatorward surge.

Acknowledgements This work is supported by National Natural Science Foundation of China (11473045, 11403045, 11503040), National Key Research Program of China (2016YFB0501909). Thanks to IGS (<ftp://cddis.gsfc.nasa.gov/gps/data/daily/>) for the GPS observation data. Thanks to GIRO (<http://giro.uml.edu/>) for the ionosonde data.

References

1. Liu L, Wan W, Ning B et al (2001) Ionospheric response at low latitude near 120°E during magnetic storm of July 2000. *Sci China (A)* 31:126–132 (in Chinese)
2. Zhang M, Radicella SM (2005) Global ionospheric response to the geomagnetic storm event of April 6–7, 2000. *Chin J Space Sci* 25(3):173–180
3. Xia C, Wan W, Yuan H et al (2005) Analysis of the intense magnetic storm of July, 2000 and of October, 2003 using the technique for nowcasting of GPS TEC data. *Chin J Space Sci* 25 (4):259–266
4. Maruyama T, Ma G, Nakamura M (2004) Signature of TEC storm on 6 November 2001 derived from dense GPS receiver network and ionosonde chain over Japan. *J Geophys Res* 109:A10302. <https://doi.org/10.1029/2004ja010451>
5. Maruyama T, Nakamura M (2007) Conditions for intense ionospheric storms expanding to lower midlatitudes. *J Geophys Res* 112:A05310. <https://doi.org/10.1029/2006JA012226>

An Alternative Approach for Estimating SNR Metrics in GPS-IR



Jizhong Wu and Wei Wu

Abstract GPS interferometric reflectometry (GPS-IR) is a bistatic radar remote sensing technique that has the potential to retrieve environmental variables such as soil moisture, snow depth and vegetation parameters. The direct and reflected signals will be mixed by omni-directional geodetic GPS antenna and the interference patterns of the both signals are recorded in signal-to-noise ratio (SNR) data. The SNR metrics, which refer to the phase, amplitude and reflector height, are the key parameters in subsequent analysis and applications of GPS-IR. Lomb-Scargle periodogram (LSP) is commonly used to estimate the dominant frequency of the detrended SNR interferograms, and the dominant frequency is converted to an effective reflector height, while phase and amplitude are finally estimated with least-squares method. In this study, the SNR metrics are determined with an alternative approach which on the basis of the least squares spectral analysis (LSSA). The core of the procedure is detecting the underlying periodicity terms in the detrended SNR interferograms, and then confirming the statistically significant spectral peaks according to the criteria formulated with Fisher distribution. The performance of the approach is examined with eight satellite tracks at the P041 site, the residuals show a 0–38% reduction in the root mean square (RMS) value compared to the LSP method and better distribution.

Keywords Global positioning system · Interferometric reflectometry
Remote sensing · SNR · Least squares spectral analysis

J. Wu (✉) · W. Wu

School of Geomatics Science and Technology, Nanjing Tech University,
211800 Nanjing, China
e-mail: jzwumail@163.com

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_9

1 Introduction

The Global Positioning System (GPS) has been proved to be a powerful tool to precisely determine the position, velocity and direction of ground stations, which are intended to help to understand the movement of the Earth's surface. Multipath occurs when signals arriving at the antenna by indirect paths due to reflection by an object or surface near the antenna and thus introduces range error to observations, and there are numerous studies on how to mitigate the effect. During recent years, several studies have demonstrated that the phase and amplitude of ground reflected signal represents the physical information of reflector, therefore the environmental parameters are potentially retrieved with reflected signal. GPS-interferometric reflectometry (GPS-IR) is such a bistatic radar remote sensing technique that using the geodetic-quality GPS receivers and antennas to estimate snow depth [1, 2], vegetation parameters [3–5], soil moisture [6–8], sea level change [9–11].

The direct and one or more ground-reflected signals are mixed in geodetic-quality antenna to create a composite signal which is the sum of the direct and ground-reflected signals. The composite signal is tracked by receiver and the oscillations of interference between the direct and indirect signals are recorded in the signal-to-noise ratio (SNR) measurements. The changes in SNR metrics, which refers to the phase, amplitude and effective reflector height, are affected to various extents by changes in reflector permittivity, height of the canopy, vegetation water content and so on. GPS-IR utilizes the SNR measurements to retrieve information about changes in the environment surrounding the antenna. This technique is currently used at many of the existing GPS stations which compose environmental sensing network to study soil moisture, snow cover, and biomass conditions. Temporal fluctuations in SNR metrics are indicative of changes in the surroundings of the GPS antenna, for instance, amplitude is the best indicator of changes in vegetation and phase is sensitive to soil moisture changes, and therefore it's possible to model the relationships between the vegetation or soil moisture parameters and corresponding SNR metrics.

Obviously, SNR metrics are extremely important for GPS-IR in order to retrieve accurate and reliable environmental parameters. However, the estimation of SNR metrics with insufficient accuracy and reliability will degrade the accuracy of environmental parameters retrieve. For the most part, the existing procedure for estimating the SNR metrics are as follows. Firstly, the SNR interferograms is fitted with a low-order polynomial to remove the direct signal. Secondly, Lomb-Scargle periodogram (LSP) is used to estimate the dominant frequency. Finally, phase and amplitude are estimated using least-square method in which the dominant frequency is substituted and treated as a constant [1, 5, 7, 8].

As described previously, the composite signal tracked by receiver actually comprises one or more reflected signals, the detection and recognition of the underlying periodicity terms in the detrended SNR interferograms should be treated with caution to provide excellent accuracy in SNR metrics estimation. In many cases, the confirmation of dominant frequency has been empirically determined by

researchers. Therefore, it's urgent to seek for a method that's sufficiently robust, reasonable to estimate the SNR metrics in various reflection scenarios. In this study, the SNR interferograms were processed using an alternative method based on the least squares spectral analysis (LSSA) technique, which has been widely used in spectral analysis for inspecting and clarifying periodic signals hidden in noisy time series with unequally spaced values and gaps [14].

2 SNR Metrics Estimation

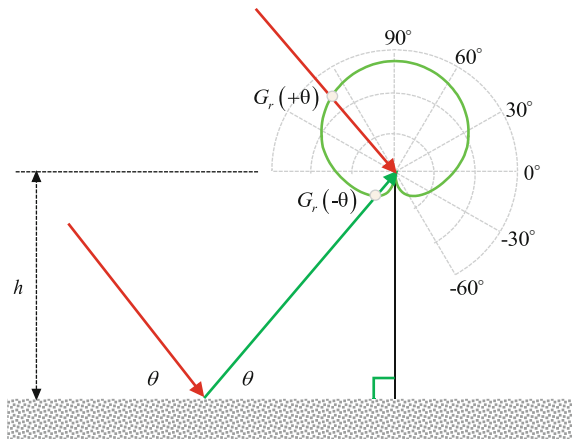
2.1 SNR Interferograms

Other than the pseudorange and carrier phase observations used in GPS applications such as high-accuracy positioning, SNR interferograms are the principal observations in GPS-IR. In addition to the signal directly from the GPS satellite (called the direct signal), a geodetic GPS receiver also records the signal reflected by ground or nearby surfaces (ground-reflected signal). The interference between the direct and ground-reflected GPS signals for a single satellite track is recorded in SNR interferograms.

The geodetic-quality GPS antennas are designed to be sensitive to the direct signal and to suppress the ground-reflected signal. For the most part, the gain patterns of geodetic antennas are designed to reject or minimize signals received from negative or low satellite elevation angles, and therefore the antenna gain pattern has higher values for the upper hemisphere than for the lower hemisphere, as shown in Fig. 1.

As receiving antennas intercept the reflected signals with lower elevation angles, the gain pattern of reflected signals is much smaller than that of direct signal, therefore, the reflected signals strength is much smaller than direct, and this means

Fig. 1 Gain patterns for a typical geodetic-quality GPS antenna



the reflected signals contribution to the SNR interferogram is small in magnitude but oscillatory while the direct signal amplitude is the dominant long-period trend in SNR interferograms. For a typical geodetic-quality GPS antenna's gain pattern, the interference between the direct and ground-reflected signals is greatest at satellite elevation angles smaller than 30° . As illustrated in Fig. 2, the raw SNR interferograms of a satellite track with elevation angle from 5° to 25° were screened and converted to a linear scale (volt/volt), it can be seen from the figure that the amplitude increases as the satellite elevation angle increases, in addition, both long-period trend and short-period oscillations are present. Based upon the above analysis, the long-period trend and short-period oscillations are corresponding to the direct and reflected signals respectively.

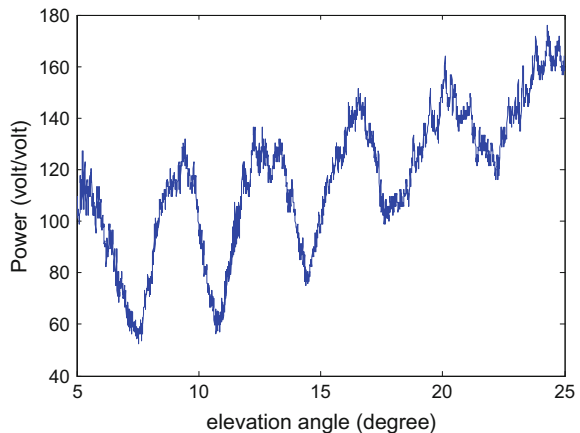
In order to isolate the reflected signals from the SNR interferograms, the direct signal part needs to be removed. This can be done either through gain pattern modeling or by fitting and removing a low-order polynomial. In practice, gain pattern modeling is complicated to implement as it requires knowledge of gain patterns for many different antenna models. In this study, a second-order polynomial fit is used to model the direct amplitude.

As an example, consider the situation when one reflected signal is present, detrended SNR interferograms can be modeled using the following equation:

$$SNR = A \cos\left(\frac{4\pi H_0}{\lambda} \sin E + \phi\right) \quad (1)$$

where A is the amplitude, H_0 is the reflector height, E is the satellite elevation angle, λ is the wavelength of the selected GPS frequency, and ϕ is a phase shift. This expression assumes that the SNR interferograms have a constant frequency which can be converted to the height of the reflector. A , H_0 and ϕ constitute the SNR metrics that are used in various GPS-IR researches and applications.

Fig. 2 Raw SNR interferograms



The commonly applied approach to estimate SNR metrics is using the Lomb Scargle Periodogram (LSP) firstly to detect the dominant frequency in detrended SNR interferograms. After having detected the dominant frequency through the LSP, the amplitude A and phase shift ϕ can be estimated using least-squares estimation under the assumption that H_0 is a known constant. Due to the presence of noise, the dominant frequency of one satellite track should be confirmed by ensuring its power is at least twice as high as the power of the second most powerful frequency in the periodogram, otherwise the satellite track will be discarded.

2.2 LSSA Method

An assumption made when detecting the peak of the spectral amplitude is that the LSP spectrums only consist of one single frequency. However, the SNR interferograms are affected by reflected signals from multiple reflectors located in the surroundings of the GPS antenna [11]. When more than one reflected signal arriving at the GPS antenna by indirect path is contained in observations, the previously mentioned method could introduce significant errors in SNR metrics estimation. An alternative method which should be robust and reliable on various occasions is necessary for estimating accurate SNR metrics.

LSSA was first developed and applied by Vaniček [12, 13] to overcome the inherent limitations of fast Fourier transform (FFT) algorithms determining the power spectrum of the data be equally spaced and equally weighted with no gaps and datum shifts. Let's consider an observed time series that is represented by $f = \{f\}, i = 1, 2, \dots, n$. time series f can be modelled by g as follows:

$$g = \Phi x \tag{2}$$

where Φ is a matrix of known base functions and x is the vector of unknown parameters. It's assumed that the values of the time series possess a fully populated covariance matrix C_f , with the standard least-squares method, it's desired to determine the model parameters with minimum difference between g and f , the model parameters can be obtained as follows:

$$\hat{x} = (\Phi^T C_f \Phi)^{-1} \Phi^T C_f^{-1} f \tag{3}$$

$$\hat{g} = \Phi \hat{x} = \Phi (\Phi^T C_f \Phi)^{-1} \Phi^T C_f^{-1} f \tag{4}$$

In spectral analysis, it is customary to search for periodic signals that are expressible in terms of sine and cosine base functions. Let's assume a set of spectral frequencies $\{\omega_i\}, i = 1, 2, \dots, n$, then we can express Φ as:

$$\Phi = [\cos \omega_i t \quad \sin \omega_i t] \quad (5)$$

Let $\hat{x} = [\hat{x}_{1i} \quad \hat{x}_{2i}]$ and then \hat{x} can be determined from Eq. (3). It should be emphasized that each frequency $\{\omega_i\}, i = 1, 2, \dots, n$, is tried independently from the rest. Then the least-squares spectrum is defined by:

$$s(\omega_i) = \frac{f^T C_f^{-1} \hat{g}(\omega_i)}{f^T C_f^{-1} f} \quad (6)$$

It's apparent that the least squares spectrum of f is the collection of the spectral values for all desired frequencies $\{\omega_i\}, i = 1, 2, \dots, n$. The greater the spectral value at a frequency ω_i , the more powerful f is at this frequency. The least squares spectrum follows the Fisher distribution with ν degree of freedom and α level of significance. Statistically significant spectral peaks satisfy the following inequality:

$$s(\omega_i) \geq \left[1 + \frac{\nu}{2} F_{\nu, 2, \alpha} \right]^{-1} \quad (7)$$

This is the level above which the detection of a frequency is statistically significant. LSSA has some strong advantages such as: (i) the systematic noise can be suppressed without causing any shift in the existing spectral peaks, (ii) time series with unequally-spaced values can also be analyzed without pre-processing, (iii) time series with an associated covariance matrix can be analyzed and (iv) statistical tests on the significance of spectral peaks can be performed, which makes the method promising and powerful for evaluating the time series [15, 16].

The interferograms were analyzed using LSSA, and the hidden periodicities of the ground-reflected signals in the interferograms were clarified. Given the number of frequencies that pass the statistical test is k , then for each determined frequency $\{\omega_i\}, i = 1, 2, \dots, k$, the matrix equation can be expressed in the form:

$$Bx = l \quad (8)$$

where: B is a matrix of coefficients, x is a vector of the terms to be computed, l is a vector which relates to the observations. B , x and l can be given as follows respectively:

$$B = \begin{bmatrix} \cos \omega_1 t_1 & -\sin \omega_1 t_1 & \cdots & \cos \omega_k t_1 & -\sin \omega_k t_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cos \omega_1 t_n & -\sin \omega_1 t_n & \cdots & \cos \omega_k t_n & -\sin \omega_k t_n \end{bmatrix} \quad (9)$$

$$x = [A_1 \cos \phi_1 \quad A_1 \sin \phi_1 \quad \cdots \quad A_k \cos \phi_k \quad A_k \sin \phi_k]^T \quad (10)$$

$$l = [f_1 \quad \cdots \quad f_n]^T \quad (11)$$

Vector x is solved by the standard least-squares method, and then desired estimates of amplitude and phase shift are derived from this vector. The amplitude estimate of the i th ground-reflected signal component is:

$$A_i = \sqrt{(A_i \sin \phi_i)^2 + (A_i \cos \phi_i)^2} \quad (12)$$

and the corresponding phase shift is given by:

$$\phi_i = \arctan\left(\frac{A_i \sin \phi_i}{A_i \cos \phi_i}\right) \quad (13)$$

3 Experimental Results

The EarthScope Plate Boundary Observatory (PBO) operates more than 1000 continuously operating GPS stations to study the motion of tectonic plates and the deformation of continental crust. The installation of PBO station P041 consists of a Trimble TRM29659.00 choke-ring antenna with a radome and a Trimble NetRS receiver. The antenna is located on bedrock with its phase center ~ 1.9 m above the ground. In this study, only the SNR data from GPS satellites which broadcast L2 C/A signals are used due to the new L2 signals have a significant improvement in recorded SNR strength compared to the old L2 signals.

The SNR interferograms of eight satellites with new L2 signals are analyzed with the method described in the previous section. It should be noticed that not all satellite tracks at this site can be utilized to retrieve SNR metrics; on the contrary, improperly selected satellite tracks will degrade the accuracy of parameter estimation. As previously described, the interference is greatest at satellite elevation angles smaller than 30° . Therefore, satellite tracks with elevation angles between 5° and 25° were used in this study. Broadcast ephemerides were used to calculate the elevation angles on each epoch.

The outputs of PRN29 on DOY 61 in 2011 are given as an example in Fig. 3 for brevity. For each satellite track that meets the requirement of elevation angle range, a second-order polynomial was fitted and removed to create detrended SNR interferograms as shown in Fig. 3a, it can be seen that the oscillations are shown as a function of sine of elevation angle. The detrended SNR interferograms track in Fig. 3a was then analyzed using LSSA at the 99% confidence level and the corresponding least-squares spectrums are shown in Fig. 3b, in which the dashed lines indicate the 99% confidence level. It can be confirmed that two spectral peaks are statistically significant as their least-squares spectrums are above the threshold

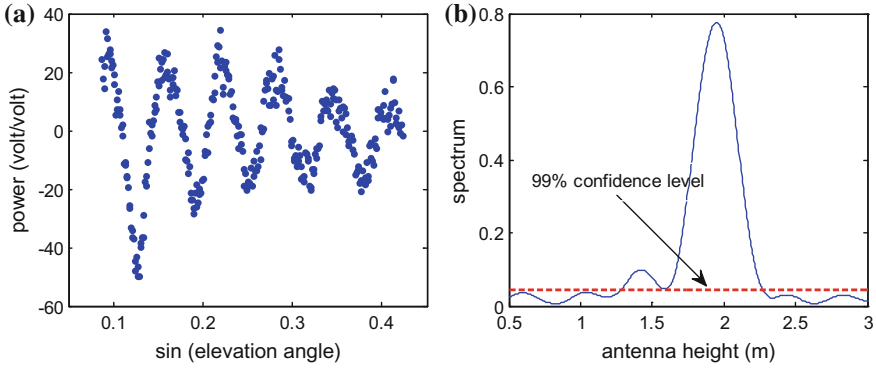


Fig. 3 Detrended SNR interferograms and corresponding least-squares spectrums

value, and the two significant frequencies will be subsequently substituted in Eq. (9) to implement the least-squares estimation.

To assess the effect of the alternative method, it's reasonable to evaluate the relative changes in the residuals instead of the absolute changes. Since the resulting residuals are practically free from periodic signals, the distribution of residuals tends to be random. In this case, the root-mean-square (RMS) was applied. The residuals for detrended SNR interferograms of PRN29 from two processing strategies, which refer to the commonly applied approach and the alternative approach suggested in this study, are shown in Fig. 4a, b respectively. As can be seen in Fig. 4a, periodic signals are still present in the residuals, while the residuals in Fig. 4b tend to random distribution. Figure 4 also gives the RMS of the two residuals time series, it can be seen that residuals from the alternative approach provide a smaller RMS with 37.6% reduction. All satellites except PRN25 were found to contain more than one significant frequency in SNR interferograms.

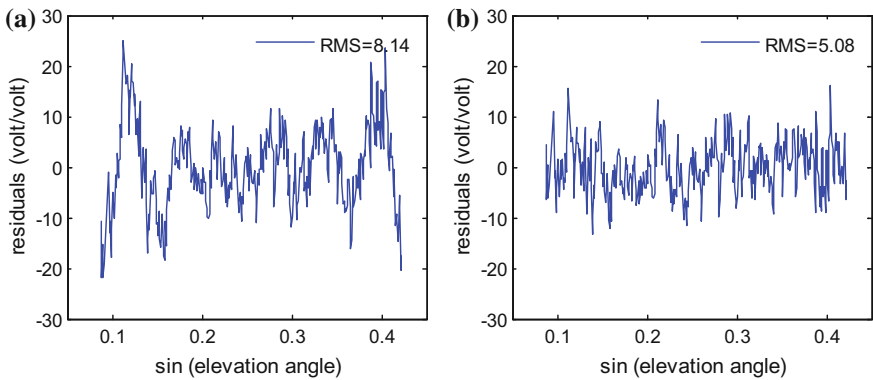


Fig. 4 Residuals for detrended SNR interferograms

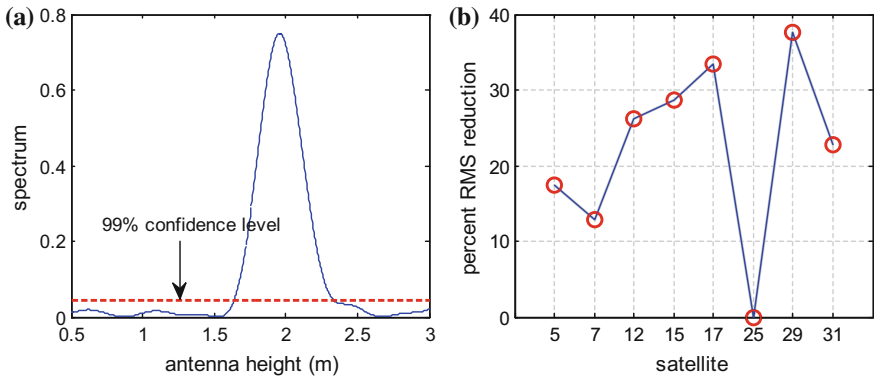


Fig. 5 **a** Least-square spectrums and significant frequency for detrended SNR interferograms of PRN25. **b** Percent reduction in residuals for each satellite

Figure 5a gives the least-squares spectrums for PRN25, the only statistically significant spectral peak in the figure correspond to the one single reflected signal, which meet the assumption of Eq. (1), and thereby the two approaches make no differences in estimation of SNR metrics as well as residuals. Additionally, Fig. 5b gives the percent reduction in residuals when the commonly applied approach served as the reference. Looking at the changes in RMS of residuals, it is apparent that the alternative approach is capable of detecting underlying periodic signals according to statistical tests and thereby results in residuals show smaller RMS values. In general, the alternative approach is preferable for its effectiveness of periodic signals detection in various reflection scenarios.

4 Conclusions

It's rather common for receivers to record SNR interferograms containing reflected signals from multiple reflectors located nearby the GPS antenna, and hence the actual data have non-constant amplitude and potentially multiple frequencies. In order to meet the assumption that the SNR interferograms only consist of one single reflected signal, it's necessary to select particular satellite tracks which have a relatively stable singular dominant frequency otherwise significant errors in subsequent phase and amplitude estimation will be introduced. In practice, it's cumbersome and time-consuming for the researcher to screen the satellite tracks.

To overcome the limitation of existing approach for estimating SNR metrics, an alternative approach which based on LSSA is suggested and investigated. The procedure of the alternative approach practically is identical to the previous one, but differs in the frequency-domain analysis of SNR interferograms using LSSA. LSSA is capable of dealing with unequally-spaced and covariance matrix associated time

series, what's more, statistical tests on the significance of least-squares spectrums can be performed at certain confidence level to detect underlying periodic signals. The clarified dominant frequencies are then substituted as constants to solve the amplitude and phase shift with least-squares method. The distribution and RMS value of residuals are examined. It's turned out that the alternative approach has remarkable improvement in terms of residuals RMS value and characteristics of distribution, furthermore, the alternative approach can take into account multiple reflections and thus can be regarded as an adaptive technique for various reflection scenarios.

However, SNR interferograms are inevitably contaminated by noise from environment and instrument, the extension of the approach will be focused on how to determine the proper confidence level and distinguish the noise from dominant frequencies in the presence of significant noise.

Acknowledgements This work is supported by the National Natural Science Foundation of China (41504024). UNAVCO's Plate Boundary Observatory (PBO) and the International GNSS Service (IGS) provided the GPS observations.

References

1. Larson KM, Nievinski FG (2013) GPS snow sensing: results from the EarthScope Plate Boundary Observatory. *GPS Solutions* 17(1):41–52
2. Nievinski FG, Larson KM (2014) Forward modeling of GPS multipath for near-surface reflectometry and positioning applications. *GPS Solutions* 18(2):309–322
3. Small EE, Larson KM, Braun JJ (2010) Sensing vegetation growth with reflected GPS signals. *Geophys Res Lett* 37(12). <https://doi.org/10.1029/2010gl042951>
4. Evans SG, Small EE, Larson KM (2014) Comparison of vegetation phenology in the western USA determined from reflected GPS microwave signals and NDVI. *Int J Remote Sens* 35(9):2996–3017
5. Wan W, Larson KM, Small EE et al (2015) Using geodetic GPS receivers to measure vegetation water content. *GPS Solutions* 19(2):237–248
6. Larson KM, Braun JJ, Small EE et al (2010) GPS multipath and its relation to near-surface soil moisture content. *IEEE J Sel Top Appl Earth Obs Remote Sens* 3(1):91–99
7. Larson KM, Small EE, Gutmann ED et al (2008) Using GPS multipath to measure soil moisture fluctuations: initial results. *GPS Solutions* 12(3):173–177
8. Chew CC, Small EE, Larson KM et al (2016) An algorithm for soil moisture estimation using GPS-interferometric reflectometry for bare and vegetated soil. *GPS Solutions* 20(3):525–537
9. Larson KM, Lofgren J, Haas R et al (2013) Coastal sea level measurements using a single geodetic GPS receiver. *Adv Space Res* 51(8):1301–1310
10. Larson KM, Ray RD, Nievinski FG et al (2013) The accidental tide gauge: a GPS reflection case study from Kachemak Bay, Alaska. *IEEE Geosci Remote Sens Lett* 10(5):1200–1204
11. Lofgren J, Haas R, Scherneck H et al (2014) Sea level time series and ocean tide analysis from multipath signals at five GPS sites in different parts of the world. *J Geodyn* 66–80
12. Vaniček P (1969) Approximate spectral analysis by least squares fit. *Astrophys Space Sci D* 4:387–391

13. Vaniček P (1971) Further development and properties of the spectral analysis by least-squares. *Astrophys Space Sci* 12:10–33
14. Erol S (2011) Time-frequency analyses of tide-gauge sensor data. *Sensors* 11(4):3939–3961
15. Erol B (2010) Evaluation of high-precision sensors in structural monitoring. *Sensors* 10(12):10803–10827
16. Pagiatakis SD (1999) Stochastic significance of peaks in the least-squares spectrum. *J Geodesy* 73(2):67–78

Research on Refined 3D Attitude Model of Smart Construction Machinery Based on Multi-sensor Fusion



Yanning Zheng, Shengli Wang, Yang Liu, Ying Xu, Xu Li
and Guiping Chen

Abstract With rapid development of technologies such as sensors and AI, smart construction machinery has gradually become research focus of construction machinery domain in recent years due to its high precision, low labor cost and high capability. As location-based services to be one of the important foundation of smart construction machinery, the design of a rigorous and convenient positioning and attitude determination algorithm has become an important foundation for the stable and agile operation of smart construction machinery. This paper combined back-hoe hydraulic excavator with GNSS positioning and orientation device and MEMS inclinometers, derived and verified a rigorous position and attitude calculation model of excavator upper structure and each part of working equipment according to the outputs of the two kinds of sensors. First, specialized algorithm to form attitude rotation matrix is deduced considering the effect characteristics of upper structure attitude on GNSS orientation baseline and output axis definition of dual-axis inclinometer, and with the GNSS antenna coordinate in topocentric coordinate system, the upper structure coordinate and attitude are obtained. Then, coordinate increment computation algorithm of key point on excavator working equipment in vehicle coordinate system considering upper structure attitude effect was deduced. Finally, by combining key point coordinate increment and upper structure coordinate and attitude, key point absolute coordinate was obtained. The correctness of this algorithm was verified by installation and test on real vehicle and

Y. Zheng (✉) · S. Wang · Y. Xu
College of Geomatics, Shandong University of Science and Technology,
Qingdao, China
e-mail: 15726247362@163.com

Y. Liu
College of Geomatics, Anhui University of Science and Technology,
Huainan, China

X. Li
College of Transportation, Shandong University of Science and Technology,
Qingdao, China

G. Chen
Shandong Astro-Compass Information Technology Co., Ltd., Jinan, China

comparison with result measured by RTK/total station. With rigorous derivation process, concise form and convenience of programming implementation, this algorithm has practical meaning in engineering.

Keywords Smart construction machinery · Back-hoe hydraulic excavator
Multi-sensor fusion

1 Introduction

In traditional working scenarios of construction machinery, to ensure construction accuracy, it's required to stake out designed scheme at construction site, let experienced operators complete the construction, and check whether desired position is achieved repeatedly during construction process. Thus, traditional construction machinery brings about complex construction process and requires the operators to be experienced. As one of the development trends of construction machinery domain, smart construction machinery has features of simplifying construction process and reducing skill requirements of operators by offering precise and real-time location-based services. As delicate modeling of construction machinery motions is required for effective location-based services, this paper studied the position and attitude calculation model of the back-hoe hydraulic excavator. A model assumed that only yaw rotation exists on the upper structure, i.e., the upper structure stays horizontal, is applied in [1–5], which is practical and easy to implement. When describing relative spatial relationships, a position and attitude calculation model based on D-H transformation matrices was introduced in [3–5], which provides a set of general expressions for robotic arms with complex structure. An attitude algorithm using IMU and potentiometer and its validation is introduced in [6]. A dynamic analysis model considering pitch and roll of upper structure is introduced in [7]. A model using UWB devices to measure position and attitude of each part of an excavator and an optimization method based on geometrics constraints is introduced in [8], which can obtain positions and attitudes of each part of an excavator simultaneously and precisely. In this paper, an excavator position and attitude model considering effect of upper structure roll and pitch is introduced. This model maintains position and attitude of the excavator upper structure and transformation between the vehicle coordinate system and the horizontal coordinate system by using GNSS dual-antenna positioning and orientation device and dual-axis inclinometer, and measures coordinates of each key point on the working equipment in the horizontal coordinate system in consideration of the upper structure attitude effect with single-axis inclinometers assembled on the boom, stick and tipping link. In this paper, the assemble scheme of hardware is introduced first. Then the main formulae of this model are given. Finally, the case of real vehicle test is introduced.

2 Installation Scheme of Hardware

As shown in Figs. 1 and 2, the vehicle coordinate system C_{veh} fixed to excavator upper structure consists of $X_{veh}, Y_{veh}, Z_{veh}$. The main antenna a_1 and the vice antenna a_2 of the GNSS positioning and orientation device are fixed on the upper structure. The dual-axis inclinometer is fixed on the upper structure and its sensitive axes are parallel to X_{veh} and Y_{veh} . Three single-axis inclinometers are applied on the lateral surfaces of the boom, stick and tipping link. Be aware that it's the tipping link instead of the bucket that the sensor is installed on to decrease damage probability of the sensor, because the bucket often touches soil and gravel.

Fig. 1 Sketch of hardware installation

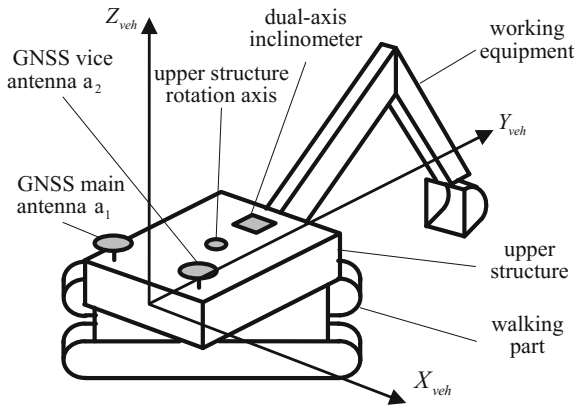
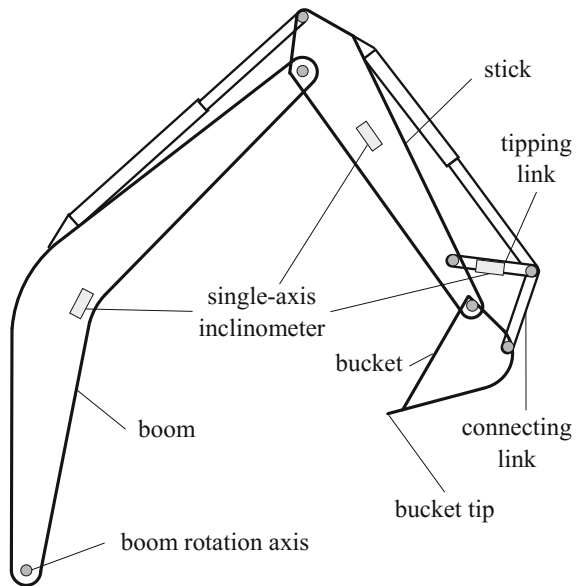


Fig. 2 Side view of working equipment



3 Overview of Algorithm

3.1 Correction of the Pitch

Figure 3 is a sketch showing the relationship between the roll and pitch output of the dual-axis inclinometer and the roll and pitch required to compose attitude rotation matrix. $X'_{veh}, Y'_{veh}, Z'_{veh}$ consist the vehicle coordinate system before applying roll and pitch rotation. $X_{veh}, Y_{veh}, Z_{veh}$ consist the vehicle coordinate system after applying roll and pitch rotation. Be aware that the attitude angle rotation order from the horizontal coordinate system to the vehicle coordinate system applied in this paper is selected to be yaw $\psi \rightarrow$ roll $\gamma \rightarrow$ pitch θ , in order to apply the pitch correction on the working equipment coordinate increment. γ is the roll output of the dual-axis inclinometer and also the roll needed in attitude rotation. θ_{cal} is the pitch needed in attitude rotation. Y_{proj} is the projection of Y_{veh} on the horizontal plane, and also the initial direction of the dual-axis inclinometer pitch output. θ_{obs} is the pitch output of the dual-axis inclinometer. According to the geometric relationship shown in the sketch, one gets

$$\theta_{cal} = \arcsin \frac{\sin \theta_{obs}}{\cos \gamma} \quad (1)$$

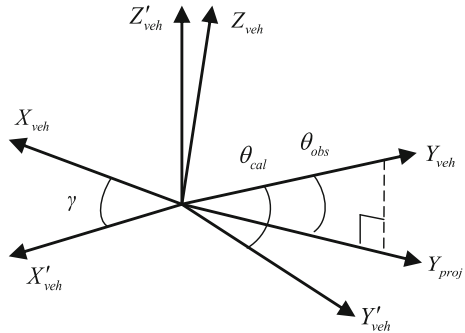
3.2 Roll-and-Pitch Rotation Matrix of the Upper Structure

The formula using γ and θ_{cal} to calculate the roll-and-pitch rotation matrix of the upper structure is

$$\mathbf{R}_{roll+pitch} = \mathbf{R}_2(\gamma)\mathbf{R}_1(\theta_{cal}) \quad (2)$$

where $\mathbf{R}_i(\text{radian})$ is the rotation matrix rotating radian around the i th axis of a cartesian system.

Fig. 3 Sketch of the dual-axis inclinometer output



3.3 Correction of the Yaw

As the baseline direction of the GNSS orientation antennas may be different from the lateral axis X_{veh} of C_{veh} , the upper structure attitude effect on the yaw measured by the GNSS device should be considered. Be aware that the yaw in this paper is defined as starting with local east direction and regarding counter-clockwise direction as positive.

$$\mathbf{B}_{rot} = \mathbf{R}_{roll+pitch}\mathbf{B} \quad (3)$$

$$\alpha_B = \arctan(y_{B_{rot}}, x_{B_{rot}}) \quad (4)$$

$$\psi_{cal} = \psi_{obs} - \alpha_B \quad (5)$$

where \mathbf{B}_{rot} is the baseline vector after applying roll and pitch rotation. \mathbf{B} is the baseline vector in C_{veh} . $\arctan(y, x)$ corresponds to azimuth calculation function $\text{atan2}(y, x)$ in the C standard library, which is used to obtain angle between (x, y) and X-axis that regards counter-clockwise as positive, and it's different with $\arctan(x)$. α_B is the baseline vector yaw after applying roll and pitch rotation. $x_{B_{rot}}, y_{B_{rot}}$ is the x and y component of \mathbf{B}_{rot} . ψ_{cal} is the yaw used in calculation. ψ_{obs} is the yaw measured by the GNSS device.

3.4 Maintenance of the Vehicle Coordinate System

The direction cosine matrix $\mathbf{R}_{attitude}$ of C_{veh} relative to horizontal coordinate system C_{hor} can be obtained by Eq. (6):

$$\mathbf{R}_{attitude} = \mathbf{R}_3(\psi_{cal})\mathbf{R}_{roll+pitch} \quad (6)$$

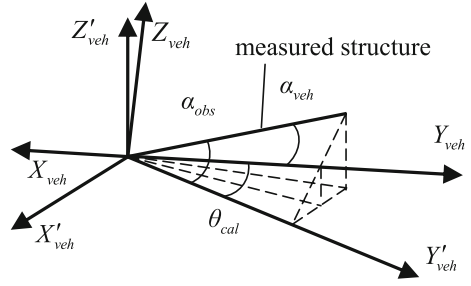
Therefore, when knowing coordinate increment \mathbf{inc}_P from GNSS main antenna to some point P in C_{veh} , upper structure attitude cosine matrix $\mathbf{R}_{attitude}$ and GNSS main antenna coordinate G in C_{hor} , the coordinate \mathbf{P} of P in C_{hor} can be denoted as

$$\mathbf{P} = \mathbf{G} + \mathbf{R}_{attitude}\mathbf{inc}_P \quad (7)$$

3.5 Analysis of Attitude Effect on Single-Axis Inclinometer Output

As shown in Fig. 4, the inclination α_{veh} of measured structure in C_{veh} does not equal to the output α_{obs} of the inclinometer. Instead, here is a relationship of $\alpha_{veh} = \alpha_{obs} - \theta_{cal}$. For convenience of computation, the working equipment

Fig. 4 Sketch of inclinometer pitch effect



coordinate increment should be converted into C_{veh} , so the pitch correction need to be applied.

Correct the outputs of the single-axis inclinometers as Eq. (8):

$$\begin{cases} \alpha_1 = \alpha_{1_{obs}} - \theta_{cal} \\ \alpha_2 = \alpha_{2_{obs}} - \theta_{cal} \\ \alpha_3 = \alpha_{3_{obs}} - \theta_{cal} \end{cases} \quad (8)$$

where $\alpha_1, \alpha_2, \alpha_3$ are the angles from Y_{veh} to the initial directions of the inclinometers in C_{veh} , and are also the values used in calculation. $\alpha_{1_{obs}}, \alpha_{2_{obs}}, \alpha_{3_{obs}}$ are the outputs of inclinometers.

3.6 Calculation of Working Equipment Key Point Coordinate

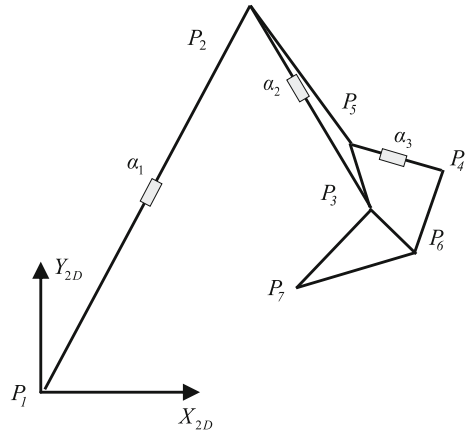
Figure 5 is the excavator working equipment sketch simplified and abstracted according to the requirements of this algorithm. Every point in Fig. 5 is a key point to be calculated in this model.

As shown in Fig. 5, any shape changes of the working equipment keep in the lateral plane, so the 2D working equipment coordinate system C_{wor} can be established as follows: suppose \mathbf{P}_1 as the origin of the system, and let X_{2D} and Y_{2D} be parallel to Y_{veh} and Z_{veh} of C_{veh} respectively. Let L_i^j denote the distance between \mathbf{P}_i and \mathbf{P}_j , $\angle \mathbf{P}_i \mathbf{P}_j \mathbf{P}_k$ denote the angle formed by the three points, and $\mathbf{P}_i = [x_i \ x_j]^T$ denote the coordinate of the point. The following formulae can be obtained:

$$\mathbf{P}_2 = L_2^1 \begin{bmatrix} \cos \alpha_1 \\ \sin \alpha_1 \end{bmatrix} \quad (9)$$

$$\mathbf{P}_3 = \mathbf{P}_2 + L_3^2 \begin{bmatrix} \cos \alpha_2 \\ \sin \alpha_2 \end{bmatrix} \quad (10)$$

Fig. 5 Simplified sketch of working equipment



$$\mathbf{P}_5 = \mathbf{P}_3 + L_5^3 \begin{bmatrix} \cos(\alpha_2 - \angle \mathbf{P}_2 \mathbf{P}_3 \mathbf{P}_5) \\ \sin(\alpha_2 - \angle \mathbf{P}_2 \mathbf{P}_3 \mathbf{P}_5) \end{bmatrix} \tag{11}$$

$$\mathbf{P}_4 = \mathbf{P}_5 + L_5^4 \begin{bmatrix} \cos \alpha_3 \\ \sin \alpha_3 \end{bmatrix} \tag{12}$$

$$\underbrace{\begin{bmatrix} 2(x_6^0 - x_3) & 2(y_6^0 - y_3) \\ 2(x_6^0 - x_4) & 2(y_6^0 - y_4) \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_6 \\ y_6 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} L_6^{32} + x_6^{02} - x_3^2 + y_6^{02} - y_3^2 \\ L_6^{42} + x_6^{02} - x_4^2 + y_6^{02} - y_4^2 \end{bmatrix}}_{\mathbf{b}} \tag{13}$$

$$\mathbf{P}_7 = \mathbf{P}_3 + L_7^3 \begin{bmatrix} \cos \alpha_3^7 \\ \sin \alpha_3^7 \end{bmatrix} \tag{14}$$

Equation (13) is the iterative formula for obtaining \mathbf{P}_6 , where x_6^0 and y_6^0 are the results of the previous iteration, x_6 and y_6 are the results of this iteration. A symmetrical point of \mathbf{P}_5 relative to the line from \mathbf{P}_3 to \mathbf{P}_4 can be used as the initial value of \mathbf{P}_6 . Thus $[x_6^0 \ y_6^0]^T$, the initial value of \mathbf{P}_6 , can be obtained by the following formula:

$$\begin{bmatrix} x_6^0 \\ y_6^0 \end{bmatrix} = \mathbf{P}_5 + 2d\vec{e} \tag{15}$$

where

$$d = \frac{|kx_5 - y_5 + y_3 - kx_3|}{\sqrt{k^2 + 1}} \tag{16}$$

$$k = \frac{y_4 - y_3}{x_4 - x_3} \quad (17)$$

$$\vec{e} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \frac{\overline{\mathbf{P}_3\mathbf{P}_4}}{\left| \overline{\mathbf{P}_3\mathbf{P}_4} \right|} \quad (18)$$

The direction α_3^7 from \mathbf{P}_3 to \mathbf{P}_7 in Eq. (14) can be obtained by the following formulae:

$$\alpha_3^6 = \arctan(y_6 - y_3, x_6 - x_3) \quad (19)$$

$$\alpha_3^7 = \alpha_3^6 - \angle \overline{\mathbf{P}_6\mathbf{P}_3\mathbf{P}_7} \quad (20)$$

3.7 Coordinate Calculation of Arbitrary Point on Working Equipment in Horizontal Coordinate System

According to the definition of C_{wor} , for a calculated coordinate \mathbf{P}_i of arbitrary key point in C_{wor} , the coordinate increment $\mathbf{inc}_{a_1}^{P_i}$ from main antenna a_1 to this key point in C_{veh} can be obtained by the following formulae:

$$\mathbf{inc}_{a_1}^{P_i} = \mathbf{inc}_{a_1}^{P_1} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{P}_i \quad (21)$$

where $\mathbf{inc}_{a_1}^{P_1}$ is the coordinate increment from the GNSS main antenna a_1 to the boom rotation axis \mathbf{P}_1 in C_{veh} .

Finally, the coordinate \mathbf{pos}_i of \mathbf{P}_i in C_{hor} can be obtained by the following formula:

$$\mathbf{pos}_i = \mathbf{G} + \mathbf{R}_{attitude} \mathbf{inc}_{a_1}^{P_i} \quad (22)$$

4 Program Implementation and Test on Real Vehicle

The correctness of this algorithm was verified by real vehicle installation, data collection, coordinate computation of the bucket middle tip, and comparison with the bucket middle tip coordinate measured by RTK and prism-free total station.

The test program which can compute the bucket tip coordinate with the sensor outputs was implemented in C++ with Qt and Eigen library. Interface for the excavator geometry parameters (such as length of boom) and the device installation calibration parameters should be designed while implementing the program to meet the needs of the field test.

In real vehicle installation, the dual-axis and single-axis inclinometers were fixed to some suitable location on the excavator with industrial tape. The two antennas of GNSS device were fixed to excavator upper structure with strong magnetic suckers. The device installation calibration parameters were measured for later use.

The test process goes as follows: drive the excavator onto the slope and turn the upper structure to make it in a non-horizontal attitude to verify algorithm correctness more effectively; read the data outputs of each sensor and apply necessary installation error compensation; enter the data to the program to get the bucket tip coordinate computed according to this algorithm; measure the bucket tip coordinate using RTK and total station; apply comparative analysis on the two sets of data.

Table 1 is the comparative analysis table composed according to actual data. The “RTK/total station results” is the difference under the main antenna topocentric coordinate system between the bucket tip coordinate measured by RTK and total station and the main antenna coordinate measured by the GNSS device. The “model results” is the coordinate increment from main antenna to bucket tip calculated by this algorithm. The “distance bias” is the length of the bias vector obtained by differencing the two sets of results. As can be seen, combinations of different yaw, roll, pitch and motion of working equipment that covered most common conditions and extreme conditions are included in the test. The geometric distance of each difference between the coordinate output from the program and coordinate measured by RTK/total station is less than 100 mm. Considering the manufacture and assembly error of construction machinery parts and the residual of the device installation calibration parameters, this test can be considered able to prove the correctness of the algorithm.

Table 1 Test data comparative analysis

γ (°)	θ_{obs} (°)	ψ_{obs} (°)	α_1 (°)	α_2 (°)	α_3 (°)	RTK/total station results (mm, mm, mm)	Model results (mm, mm, mm)	Distance bias (mm)
1.33	2.39	57.61	36.43	-7.18	14.47	-3959.90, 8079.73, 2767.94	-3883.98, 8110.27, 2778.95	82.57
1.78	1.84	261.39	41.18	-10.32	13.33	268.27, -8658.25, 2999.94	206.30, -8658.62, 2954.06	77.11
5.87	3.45	63.83	39.72	-1.14	34.81	-2998.88, 9032.69, 2462.20	-3037.94, 9034.27, 2457.39	39.39
4.70	6.01	274.48	10.26	-6.51	152.31	-2075.15, -12,832.62, 1744.81	-2155.41, -12,811.50, 1760.38	84.44
3.39	4.76	297.83	43.89	17.39	171.63	-6216.59, -8822.58, 6594.42	-6234.62, -8829.50, 6578.73	24.88
0.72	0.53	12.41	57.17	23.55	167.24	-9114.98, 3123.40, 7629.56	-9091.20, 3122.41, 7609.19	31.33
3.53	26.41	77.82	55.76	22.65	157.40	-547.24, 9846.85, 8340.42	-577.00, 9859.24, 8335.81	32.56
-25.72	4.63	351.27	47.80	13.61	133.49	-10,703.20, -3396.65, 5767.15	-10,744.10, -3323.41, 5738.37	88.69
-15.31	17.62	28.93	51.31	9.89	138.78	-9780.57, 4535.81, 6754.57	-9785.52, 4483.34, 6745.32	53.51
-15.31	17.63	28.92	37.91	-103.22	-83.64	-5866.51, 3323.11, 3631.12	-5840.37, 3283.40, 3606.04	53.75

5 Conclusions

By considering the effect of upper structure attitude on key point coordinate calculation while deriving, a rigorous algorithm for excavator upper structure position and attitude and working equipment key point coordinate using GNSS positioning and orientation device and MEMS inclinometer is derived in this paper, and the refined attitude model of smart construction machinery is established. This model have the following advantages: (1) With a rigorous derivation process, this model can describe the working scene more accurately. (2) As this model is simple in form and convenient for program implementation, it is easy to be applied widely in smart construction machinery design. (3) This model is stable and reliable, which can improve safety of construction process. Thus, this model can be applied in actual production activities to greatly reduce the complexity of site construction.

References

1. Bradley DA, Seward DW (1998) The development, control and operation of an autonomous robotic excavator. *J Intell Robot Syst* 21(1):73–97
2. Jin SM, Choi JJ, Lee DY, Yang SY (2008, April) Development of remote control system for field robot. In: International conference on smart manufacturing application, 2008. (ICSMA 2008). IEEE, pp 428–432
3. Wang D, Zheng L, Yu H, Zhou W, Shao L (2016) Robotic excavator motion control using a nonlinear proportional-integral controller and cross-coupled pre-compensation. *Autom Constr* 64:1–6
4. Jun Y, Bo L, Qunzhang T, Gang G, Yonghua Z (2011, January) Automatization of excavator and study of its autocontrol. In: 2011 Third international conference on measuring technology and mechatronics automation (ICMTMA), vol 1. IEEE, pp 604–609
5. Xu GN, Lu FY, Hou PL (2016) Kinematic analysis of hydraulic excavator working device based on DH method. *DEStech Trans Mater Sci Eng (ammme)*
6. Bender FA, Göltz S, Bräunl T, Sawodny O (2017) Modeling and offset-free model predictive control of a hydraulic mini excavator. *IEEE Trans Autom Sci Eng*
7. Vujic D, Lazarevic O, Batinic V (2017) Development of dynamic-mathematical model of hydraulic excavator. *J Cent S Univ* 24(9):2010–2018
8. Vahdatikhaki F, Hammad A, Siddiqui H (2015) Optimization-based excavator pose estimation using real-time location systems. *Autom Constr* 56:76–92

Feasibility Study of Low Cost Receiver for Deformation Monitoring



Hongke Hou, Cuilin Kuang, Yaozong Zhou and Yonglin Zhang

Abstract In the satellite navigation and positioning process, the performance of the receiver directly impacts on the quality of observation data, and affects the positioning accuracy ultimately. Therefore, the receiver is sufficiently important for the whole navigation and positioning process. Deformation monitoring belongs to precise engineering survey and it often needs to reach the millimeter-level accuracy, so that the geodetic receiver with excellent performance and high price is often used in this survey. Under this background, improving the positioning accuracy of low cost receiver and applying it to deformation monitoring can effectively reduce the engineering cost, which has great practical significance and research value. This study is based on the low cost receiver EVK-M8T from ublox company. Firstly, the receiver performance is evaluated from signal to noise ratio (SNR), receiver clock and accuracy of standard point positioning (SPP) through comparing with geodetic receiver Trimble NetR9. Then through the experiment simulation of deformation monitoring, the feasibility of deformation monitoring using low cost receiver is analyzed. The numerical results show that the location accuracy of low cost receiver can basically keep in millimeter level when the observation period reaches more than 2 h, and it can correctly detect the movement of the point. Hence low cost receiver can be applied to the deformation monitoring under the premise of improving its stability.

Keywords GNSS · Low cost receiver · Performance evaluation
Deformation monitoring · Feasibility



H. Hou · C. Kuang (✉) · Y. Zhou · Y. Zhang
School of Geosciences and Info-physics, Central South University,
Changsha 410083, China
e-mail: cuilinkuang@163.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018
Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_11

1 Introduction

In engineering survey, GNSS receivers have direct influence on the quality of collected data, and affect the accuracy of positioning ultimately. Though the geodetic receiver can obtain accurate positioning at millimeter level, it is expensive and not suitable for promotion in some projects. Therefore, it has great practical significance and research value to improve the accuracy of low cost receiver and then apply it to engineering survey. Table 1 performs the parameters comparison between the geodetic and low cost receivers, showing different characteristics of these two receivers. Realini et al. carried out a series of research on low cost receiver with goGPS, and verified that location accuracy of single frequency and low cost receiver processed by goGPS has been greatly improved [1–3]. Takasu et al. had evaluated the performance of low cost receivers and achieved precise result with low cost receivers by RTKLIB [4]. Marco et al. studied on the landslide monitoring using the low cost receiver and investigated the effect of baseline length on the result [5, 6]. Caldera et al. based on a few open source and commercial processing softwares to process data from low cost receiver, and achieved millimeter accuracy at short baseline [7]. Deformation monitoring is precise survey in engineering and the accuracy needs to reach millimeter level so that displacement can be detected accurately. It is feasible and meaningful to study the deformation monitoring using low cost receiver, which can reduce the cost of engineering project.

Table 1 Parameters comparison between geodetic and low cost receivers

		
Constellation	GPS, GLNASS, Galileo, BDS, QZSS	GPS, GLNASS, Galileo, BDS, QZSS
Frequency	L ₁ , L ₂ , L ₅	L ₁
Cost	10,000–200000 RMB	100–3000 RMB
Weight	>2 kg	<500 g
Power consumption	5 W	120 mW
External antenna	Yes	Yes
Post-processing accuracy	1–10 mm	1–25 mm

In this paper, we analyzed the feasibility of low cost receiver EVK-M8T for deformation monitoring. Firstly, we evaluate performance of the receiver, including signal to noise ratio, receiver clock performance and the accuracy of SPP, and compare these with the geodetic receiver. Then the simulation experiment of deformation monitoring is carried out.

2 Performance Evaluation of Low Cost Receiver

With the collected data from low cost receiver EVK-M8T, the performance is evaluated from the perspective of data quality and the receiver clock [8]. The data quality is checked through signal to noise ratio and the accuracy of SPP positioning. The frequency stability of the receiver clock and the clock error bias between different systems is also analyzed. Meanwhile, these results above are compared with that of the geodetic receiver.

2.1 Signal to Noise Ratio

Signal to noise ratio (SNR) refers to the proportion of signals and noises in an electronic device or system, which is usually expressed as the carrier noise power density ratio (C/N_0) with unit dB/Hz. The relevant formula is as follows:

$$SNR = 10 * \lg(P_s/P_n) \quad (1)$$

P_s and P_n represent the effective power of signal and noise, respectively.

In the process of satellite navigation and positioning, the higher the SNR, the smaller the proportion of the noise and the better the signal quality and accuracy of the observation [9]. At the same observation conditions, the higher the SNR, the stronger receiver's anti-disturbance ability and the better performance. Therefore, the geodetic receiver and low cost receiver are placed in the same observation condition for data collection, and then SNR of each observation epoch of all satellite is extracted directly from the RINEX file.

As shown in Fig. 1, the relationship between the SNR and the elevation angle for two receivers is presented, and it can be seen that the SNR extracted from the low cost receiver is all integer. When the elevation angle is higher than 40° , the SNR of geodetic receiver is better than low cost receiver, but the gap is small. At low elevation angles (less than 40°), the SNR of the low cost receiver fluctuates intensely, and its stability is poor. The difference between these two receiver is obvious. Then the mean of SNR of different satellite systems for two receivers is

Fig. 1 The relationship between SNR and the elevation angle of GPS

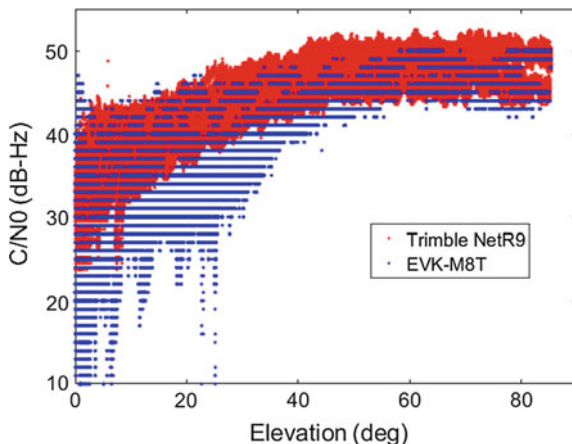
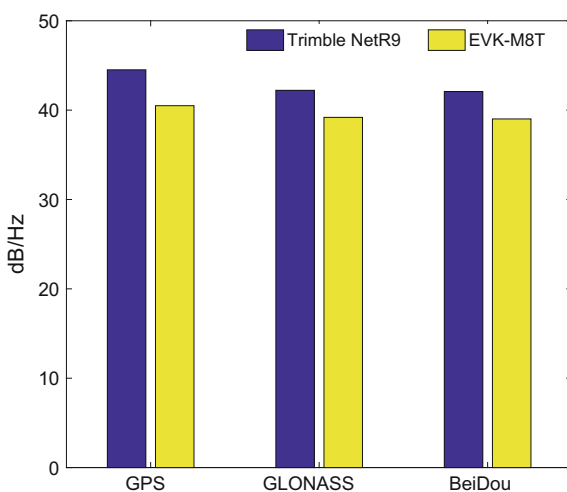


Fig. 2 The mean of SNR between geodetic and low cost receivers



shown in Fig. 2. The results show that the SNR of each satellite system of the geodetic receiver is higher than that of the low cost receiver, so that the anti-disturbance ability of geodetic receiver is stronger.

2.2 The Performance of Receiver Clock

When the receiver's position is calculated, the receiver clock offset is also obtained simultaneously. The calculation based on single point positioning (SPP), and then two types of receivers clock offset are extracted in each epoch. The clock offset of

the geodetic receiver is at 10^{-8} – 10^{-9} s, and that of the low cost receiver is at 10^{-3} – 10^{-4} s.

Frequency stability is the most important indicator to describe the performance of the clock. It indicates the ability of the oscillator to generate the same time and frequency in a certain time. In general, there are some kinds of variances used to describe the wave characteristics of the frequency source. In this paper, the Hadamard variance is adopted, which is a three times sampling variance and not affected by the influence of frequency drift [10]. The calculation formula is as follows:

$$H\sigma_y^2(\tau) = \frac{1}{6\tau^2(N' - 3)} \sum_{i=1}^{N'-3} [x_{i+3m} - 3x_{i+2m} + x_{i+m} + x_i]^2 \tag{2}$$

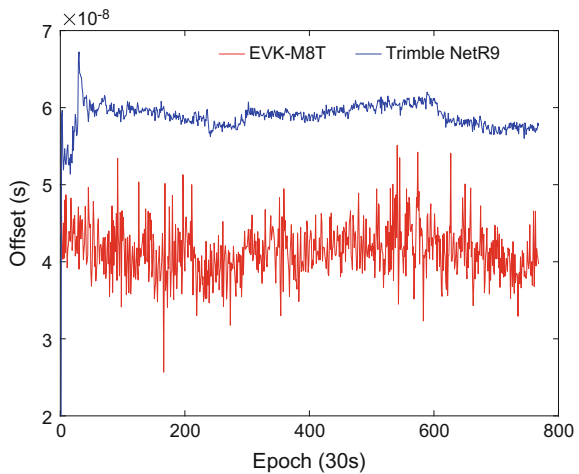
where $\{x_n, n = 1, 2, \dots, N\}$ is a clock error sequence, N is the number of the clock error data, m is the smoothing factor, τ_0 is the sampling interval and τ is the measurement interval. N' is the number of clock error in the measurement interval.

According to the clock offsets extracted, we calculate the hour stability of two types of receiver clocks. The hour stability for Trimble NetR9 is 1.71×10^{-9} , and that of EVK-M8T is 8.45×10^{-5} . The results show that the low cost receiver clock is less stable than the geodetic receiver clock.

With the construction of the four satellite navigation systems, Compatibility and interoperability between systems becomes more and more important. Whether the receiver clock is consistent with different systems can reflect the performance of the receiver clock. At present, the bias of receiver clock offset is usually used to reflect this consistency.

EVK-M8T can receive three kinds of satellite signals at the same time, so GPS and BDS data are collected to calculate the BDS-GPS clock offset's bias of the

Fig. 3 BDS and GPS receiver clock offset's bias



receiver. From Fig. 3, we can see that the clock bias of the two receivers is all in the order of 10^{-8} , but the clock offset's bias of the low cost receiver is higher than that of the geodetic receiver, and the stability is poor. Therefore, the clock of the geodetic receiver is better for the consistency of different satellite systems.

2.3 Standard Point Positioning

In this paper, the low cost receiver navigation performance evaluation is based on the standard point positioning (SPP) of GNSS pseudo-range observation. Considering some important errors, the GNSS pseudo-range measurements can be written as:

$$\tilde{\rho}_i = \sqrt{(X^i - X)^2 + (Y^i - Y)^2 + (Z^i - Z)^2} - cV_{tr} + cV_{t_i}^s - (V_{ion})_i - (V_{trop})_i + \varepsilon \quad (3)$$

where X^i , Y^i and Z^i are the three dimensional positions of the satellite, $V_{t_i}^s$ is the satellite clock error, which can be obtained from the satellite ephemeris. X , Y and Z are the receiver positions, and V_{tr} is the receiver clock error. The ionosphere delay V_{ion} can be eliminated by the Klobuchar model and the troposphere delay V_{trop} can be eliminated by the Saastamoinen model. And the pseudo-range measurements noise ε can be ignored.

The receivers are placed in the same observation environment to collect data, and the data are processed by SPP mode. The result shows that the location deviation of the low cost receiver fluctuates greatly and the stability is poor, as shown in Fig. 4. Then the location error RMS mean value is calculated respectively. As shown in Table 2, the SPP positioning accuracy of the low cost receiver

Fig. 4 The comparison of positioning errors between geodetic and low cost receivers (SPP)

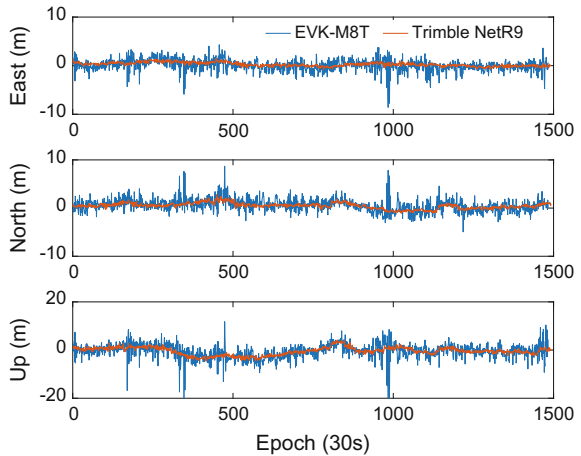


Table 2 RMS mean of positioning errors (SPP)

	EVK-M8T	Trimble NetR9
East (m)	1.22	0.49
North (m)	1.44	0.76
Up (m)	3.15	1.69
3-D(m)	3.67	1.92

Fig. 5 Base station CREE



is better than 4 m, while that of the geodetic receiver is better than 2 m, thus the SPP location accuracy of the low cost receiver is worse than that of the geodetic receiver evidently.

The results above show that the low cost receiver has a certain gap in performance with the geodetic receiver. Therefore, a good observation condition was selected in the following experiment, post-processing differential positioning mode was used in data processing, and a variety of error correction models were adapted to compensate for the low-cost receiver defects.

3 Deformation Monitoring Experiment Simulation

The experimental field is selected at the top of a graduate building in a university, where is no shelter and wide field of vision around. In order to reduce the interference of baseline length to experimental results, the baseline length is controlled



Fig. 6 Monitoring station CREF

to about 5 m. We set up the base station CREE and the monitoring station CREF on the top of the building. Among them, the base station CREE uses the Trimble geodetic antenna and the Trimble NetR9 receiver, as shown in Fig. 5. The monitoring station uses the low cost antenna and the EVK-M8T receiver, and the controlled mobile device is applied to simulate the deformation, as shown in Fig. 6.

The experiment was carried out for 5 days. The receiver collected data in stationary state at first day and the simulated deformation monitoring was carried out at the following days.

The static data on the first day are processed, and static solutions are generated every two hours. The result is subtracted from the real coordinates, and then transformed to the corresponding East (E), North (N) and Up (U) directions. Finally

Fig. 7 Time sequence for residuals, 2 h sessions (static)

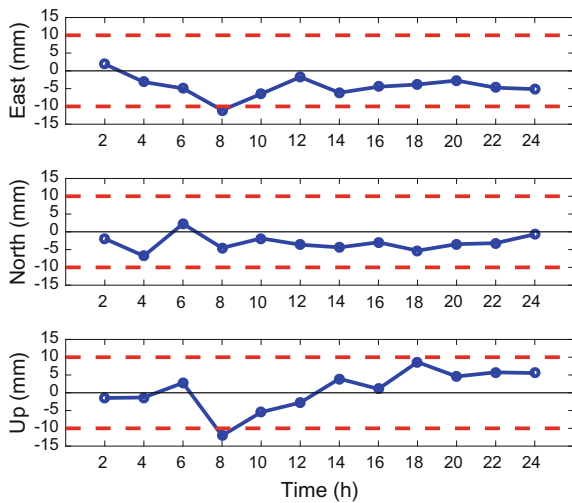


Table 3 Statistics of residuals (static)

	East	North	Up
Max (mm)	-11.07	-6.78	-12.06
Min (mm)	-1.75	-0.70	-1.32
RMS (mm)	5.28	3.77	5.55

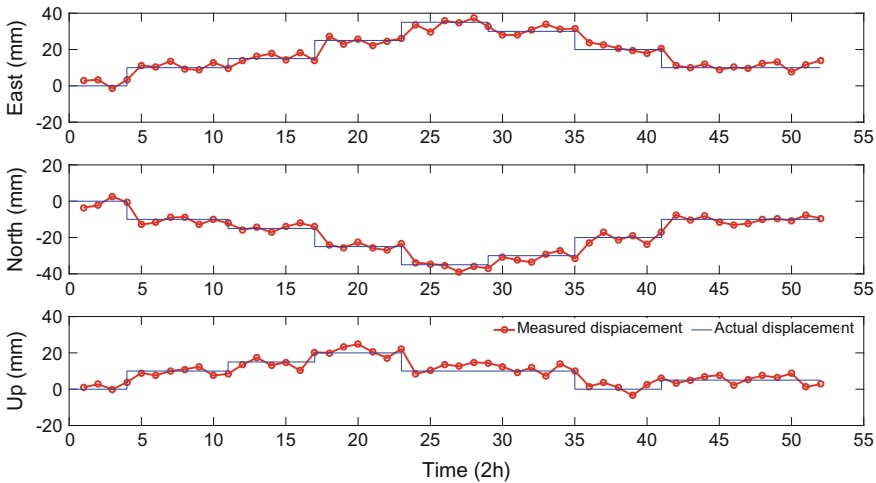


Fig. 8 The comparison between measured displacement (in red) and actual displacement (in blue)

residuals sequence is obtained, as shown in Fig. 7. In order to analyze the positioning accuracy of the low cost receiver in the static state, the RMS values of three directions (as shown in Table 3) are computed respectively.

The following deformation monitoring simulation data are processed in the same way. As shown in Fig. 8, the blue line represents the manual movement and the red line is the measured deformation by low cost receiver.

The results show that the positioning accuracy of the three directions basically keeps in millimeter, but there are several abnormal values in the coordinate series. The deformation monitoring experiment simulation shows that the low cost receiver can accurately detect the movement of the antenna.

4 Conclusions and Prospect

The performance evaluation and precision test of low cost receiver are carried out in this paper, finally the deformation monitoring experiment simulation is implemented. The results show that there is a gap between the low cost receiver and the geodetic receiver in performance. During data processing, mature software and corresponding error models are used to make up for the performance defects of low

cost receiver. The tests show that the accuracy of the low cost receiver basically keeps in millimeter and it can detect the movement of the point correctly.

In the next work, the performance of the receiver needs to be further evaluated, including the internal noise of the receiver, the multipath error, and so on. In addition, with regard to the deformation monitoring in the engineering project, the base station is usually set up in the stable area, and the length of baseline is often more than 100 m. So we will further discuss the influence of the baseline length on the positioning accuracy and the corresponding improvement methods.

Acknowledgements The financial supports from National Key R&D Program of China (No. 2016YFB0501803) and National Natural Science Foundation of China (No. 41774040) are greatly appreciated.

References

1. Realini E (2009) GoGPS free and constrained relative kinematic positioning with low cost receivers. Politecnico di Milano, Milan
2. Realini E, Yoshida D et al (2012) Enhanced satellite positioning as a web service with goGPS open source software. *Appl Geomat* 4(2):135–142
3. Herrera A, Suhandri H, Realini E et al (2016) goGPS: open-source MATLAB software. *GPS Solut* 20(3):595–603
4. Takasu T, Yasuda A (2009) Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. In: International symposium on GPS/GNSS, Jeju South Korea
5. Marco P (2015) Low-cost GNSS receiver for geodetic local monitoring. Politecnico di Milano, Milan
6. Alberto C, Marco P (2015) Performance of low-cost GNSS receiver for landslides monitoring: test and results. *Geomat Nat Hazards Risk* 6(5–7):497–514
7. Caldera S, Realini E, Barzaghi R et al (2016) Experimental study on low-cost satellite-based geodetic monitoring over short baselines. *J Surv Eng* 142(3):04015016
8. Li J, Yang K (2015) Performance evaluation of self-studied BDS/GPS dual mode high precision receiver. *Glob Position Syst* 40(4):46–49
9. Dou S, Kuang C, Zhou Y (2017) Analysis of signal quality and navigation performance for beidou system. In: China satellite navigation conference (CSNC) 2017 proceedings, vol I. Springer, Singapore, pp 671–681
10. Guo H (2006) Study on the analysis theories and algorithms of the time and frequency characterization for atomic clocks of navigation satellites. PLA Information Engineering University, Zhengzhou

Algorithm and Performance of Precipitable Water Vapor Retrieval Using Multiple GNSS Precise Point Positioning Technology



Peng Hu, Guanwen Huang, Qin Zhang, Xiaolei Wang and Min Mao

Abstract It is of great importance to estimate zenith tropospheric delay (ZTD) and retrieve precipitable water vapor (PWV) using ground-based GNSS remote sensing technology. At present, the ground-based GNSS technology is usually studied based on single GPS system. In this paper, the precise point positioning technology for combined multi-GNSS is carried out using the observation data of 30 MGEX tracking stations for one month. The ZTD and PWV results obtained from individual GNSS and multi-GNSS are carefully compared and analyzed. The performance of multi-GNSS data for ZTD/PWV retrieval is also assessed and the accuracy is verified by CODE tropospheric products and Radiosonde observations. The statistical results show that: (1) There are significant differences in the PPP ZTD results obtained by different navigation systems, and the more stable ZTD results can be obtained from the multi-GNSS observations. (2) The ZTD series obtained from single GNSS and multiple GNSS show good agreement with CODE ZTD series. Compared with that of GPS, GLONASS, Galileo and BDS, the solution precision of combined GNSS is the highest, which is improved by 10.91, 19.04, 33.21 and 70.16% respectively. (3) The accuracy of atmospheric vapor in the PPP data processing can reach the meteorological requirements such as numerical weather prediction model. The performance of combined GNSS compared with sounding data is the best, the accuracy of which is improved than that of GPS, GLONASS, Galileo and BDS by about 3.45, 16.16, 16.45 and 41.78% respectively. Furthermore, the results reveal that the multi-GNSS combined PPP technology can significantly improve the accuracy and reliability of ZTD/PWV series, which can support for meteorological applications such as weather monitoring and forecasting.

Keywords GNSS · Precise point positioning · Precipitable water vapor
Zenith tropospheric delay

P. Hu · G. Huang (✉) · Q. Zhang (✉) · X. Wang · M. Mao
College of Geology Engineering and Geomatics, Chang'an University,
Xi'an 710054, Shanxi, China
e-mail: huang830928@163.com

Q. Zhang
e-mail: zhangqinle@263.net.cn

1 Introduction

Atmospheric water vapor is one of the main gases that cause the greenhouse effect and that plays a very important role in weather prediction, micro-meteorology and global climate change [1]. Some traditional means of measuring water vapor, such as microwave radiometers and radio sounding balloons, are limited in their application to modern meteorology, due to their high operating cost, sparse stations distribution and the inability to all-weather availability. With the advent of GPS/MET technology [2] and the continuous improvement of GPS technology, the GPS ZTD/PWV products have been significantly improved in terms of accuracy, timeliness and time resolution, which is of great significance for the development of meteorological applications. At the same time, precise point positioning technology plays an important role in the field of ground-based meteorology due to the advantages of single station operating, flexible, free from operating distance, low operating cost and directly access to absolute delay [3–5].

At present, with the gradual modernization of American GPS and Russian GLONASS, Chinese BDS and European Galileo have also been initially equipped with navigation and positioning capabilities. The single GPS navigation satellite system has gradually developed into a multi-GNSS (multi-constellation global navigation satellite systems) [4, 5]. The multi-GNSS combination has become one of the research topics in the field of tropospheric zenith delay estimation and high-precision water vapor retrieval. Lu et al. [6–8] obtained ZTD series from GPS and GLONASS observations and found that the GPS/GLONASS combination can improve the solution accuracy. Fan et al. [9, 10] estimated PPP ZTD from GPS and BDS mixed observations and verified the ZTD/PWV results have higher stability. Li et al. [11] adopted multi-GNSS real-time precise point positioning technology and concluded that the multi-GNSS combination can improve the accuracy of water vapor retrieval. Ding et al. [12] performed real-time ZTD estimation based on the modified PPP-wizard software and demonstrated that applying ambiguities resolution and multi-GNSS observations can significantly contribute to accuracy of the ZTD solution.

However, most of the existing research results mainly focus on the analysis of tropospheric delay by the integration of GPS, GLONASS and BDS navigation systems with the precise point positioning technique, so far only limited studies related to the contribution of Galileo system and the performance differences of the tropospheric delay and water vapor retrieval in the GNSS four-system combination. In order to analyze and reveal the accuracy difference of GNSS four systems combined PPP technology and the performance improvement of combined positioning, this paper selected 30 globally distributed stations which can simultaneously receive multi-GNSS satellite signals and all the observation data during the period of March 1–31, 2017, are processed in PPP mode to estimate ZTD/PWV. Firstly, the ability of ZTD/PWV results derived from different constellations are analyzed and compared. Then CODE tropospheric products and Radiosonde data,

as an external reference result, are used to assess the precision magnitude of GNSS four-system combined solution.

2 Multi-GNSS Retrieval Precipitable Water Vapor

In this paper, the precise point positioning technology is used for data processing with taking GPST as reference time. GNSS PPP processing utilizes the ionosphere-free combinations of dual-frequency pseudorange and carrier-phase observations, data preprocessing by setting threshold for the observation data to test and eliminate the gross error, while using ionosphere-free linear combinations, geometry-free linear combinations and M-W linear combinations for cycle slip detection. Furthermore, there are a number of error sources corrections include antenna phase center correction, antenna phase wind-up correction, relativistic effect correction, differential code deviation correction, solid tide correction, ocean loading tidal correction, polar tide correction and so on [13]. The individual GNSS PPP observation equation can be expressed as following:

$$\begin{cases} L = \rho + c(dt_r - dt_s) + T + \lambda N + \varepsilon_L \\ p = \rho + c(dt_r - dt_s) + T + \varepsilon_P \end{cases} \quad (1)$$

where L and P denote the dual-frequency non-ionosphere carried-phase and the pseudorange observations, respectively, ρ is the geometric distance from the satellite to the station, dt_r and dt_s are the receiver-satellite errors, c is the speed of light, T is the tropospheric delay, N is the ambiguity parameter, ε_L and ε_P are the errors that do not eliminate the multipath and measurement noise in the carrier-phase and pseudorange observations. In multi-GNSS combined PPP processing, we must consider the receiver signal delay bias between different channels; the GNSS four systems combined observation equation is given:

$$\begin{cases} p^G = \rho^G + c(dt_r - dt_s^G) + T^G + \varepsilon_P^G \\ p^{R_K} = \rho^R + c(dt_r - dt_s^R) + B_P^{R_K-G} + T^R + \varepsilon_P^R \\ p^E = \rho^E + c(dt_r - dt_s^E) + B_P^{E-G} + T^E + \varepsilon_P^E \\ p^C = \rho^C + c(dt_r - dt_s^C) + B_P^{C-G} + T^C + \varepsilon_P^C \\ L^G = \rho^G + c(dt_r - dt_s^G) + T^G + \lambda^G N^G + \varepsilon_L^G \\ L^{R_K} = \rho^R + c(dt_r - dt_s^R) + B_L^{R_K-G} + T^R + \lambda^{R_K} N^R + \varepsilon_L^R \\ L^E = \rho^E + c(dt_r - dt_s^E) + B_L^{E-G} + T^E + \lambda^E N^E + \varepsilon_L^E \\ L^C = \rho^C + c(dt_r - dt_s^C) + B_L^{C-G} + T^C + \lambda^C N^C + \varepsilon_L^C \end{cases} \quad (2)$$

where G, R, E and C respectively represent GPS, GLONASS, Galileo and BDS, K in R_K represents GLONASS satellite number, B_P and B_L represent inter-system biases of GLONASS, Galileo and BDS relative to GPS.

In the precise point positioning, the delay caused when the electromagnetic wave signal propagates through the troposphere is called the Slant Total Delay (STD), which is equal to the Zenith Total Delay (ZTD) multiplying the corresponding mapping function. The tropospheric zenith delay can be divided into two parts: Zenith Hydrostatic Delay (ZHD) and Zenith Wet Delay (ZWD) [14]. ZHD is easy to model due to their low variability, while the ZWD can't be accurately corrected and is estimated as a parameter from the observations. Meanwhile, the tropospheric horizontal gradient parameters need to be estimated in consideration of the azimuthal anisotropy of the real atmosphere [15]. The specific formula is:

$$T(e) = Z_h \cdot MF_h(e) + Z_w MF_w(e) + MF(e)_{azi} \cdot (G_N \cos \phi + G_E \sin \phi) \quad (3)$$

where Z_h and Z_w represent the ZHD and ZWD, $MF(\cdot)$ is the mapping function, e is the satellite elevation angle, ϕ is azimuth, $MF(e)_{azi}$ denotes the gradient mapping function, G_N is the gradients parameter of north-south direction, G_E is the gradients parameter of east-west direction.

In the processing of obtaining PPP ZTD/PWV, the precise ephemeris and clock products provided by GFZ, together with the ionosphere products provided by CODE, are used. Since the information of Galileo and BDS satellite antenna and receiver antenna is unknown, Galileo antenna corrects will refer to GPS antenna information; BDS antenna correction will use recommended value of IGS08.ATX. The PPP processing strategies are listed in Table 1 in detail. For multi-GNSS combination PPP, we must consider the differences among different systems. In order to rationally allocate the weight for different systems, this paper determines the weight ratio of GPS, GLONASS, Galileo and BDS to be 6:4:4:3 according to the method of Helmet variance component estimation [16–18].

Table 1 PPP processing strategies

Item	Models/strategies
Constellations	GPS/GLONASS/BDS/Galileo
Observation	PC, LC
Frequency	GPS/GLONASS:L1&L2; BDS:B1&B2; GAL:E1&E5a
Sampling rate	30 s
Elevation cutoff	5°
Weighting strategy	Elevation-dependent weighting
Satellite antenna	GPS/GLONASS:PCV.I14; GAL/BDS:PCO
Receiver antenna	GPS/GLONASS/GAL: PCV.I14; BDS:PCO
Phase wind-up	Corrected
Tidal effects	Solid tides, ocean loading and polar tides
Station coordinates	Constrained (5, 5, 20 mm)
Zenith tropospheric delay	A priori model:DRY GMF; Mapping function:WET GMF; Estimated with piecewise linear method; 2 h intervals
Horizontal gradients	CHENHER Gradients model; 12 h intervals

Once the zenith wet delay (ZWD) is estimated, the PWV can be calculated by using the atmospheric pressure, relative humidity and temperature information over the station:

$$\text{PWV} = \text{ZWD} \cdot \Pi(T_m) \quad (4)$$

$$\Pi(T_m) = \frac{10^6}{\rho_w R_v [(k_3/T_m) + k_2]} \quad (5)$$

where the parameter $\Pi(T_m)$ varies as a function of the weighted mean temperature of the atmosphere T_m , $\rho_w = 999.97 \text{ kg m}^{-3}$ denotes the density of liquid water, $R_v = 461.51 \text{ JK}^{-1} \text{ kg}^{-1}$ is the specific gas constant of water vapor, K_2 and K_3 are atmospheric refraction constants.

3 Experiment and Analysis Results

3.1 Data Sources

To study emerging satellite navigation systems, IGS(the International GNSS service) has established a multi-GNSS experiment(MGEX) to track, collect and analyze data from GPS, GLONASS, Galileo, BeiDou, QZSS and any space-based augmentation systems of interest [6, 11]. This article selects 30 MGEX tracking stations that can simultaneously receive GREC(G: GPS, R: GLONASS, E: Galileo and C: BDS) four systems signals during the period of 60–90, 2017. Station distribution is shown in Fig. 1.

3.2 Comparative Analysis of GNSS PPP ZTD

In order to analyze the differences of zenith tropospheric delay between single GNSS and multiple GNSS, and to assess the performance of ZTD in the joint estimation of single GNSS observations, observations from 30 MGEX stations, where can simultaneously receive multi-GNSS satellite signals, are processed by precise point position technology during the period of 60–90, 2017. The PPP ZTD results of the individual GNSS and multi-GNSS combination are obtained. First of all, the GNSS ZTD series, which are obtained from the observation of the stations REYK and CHTI by using the PPP technique, are shown in Fig. 2.

We can notice that the different single GNSS solution are all show the same trend as the GNSS combination solution, but there are obviously more gross errors in the ZTD sequences estimated by the BDS single system, this is probably due to the low precision of the BDS error model. In order to fully demonstrate the



Fig. 1 The distribution of the stations from MGEX (Red dots indicate stations for specific example analysis)

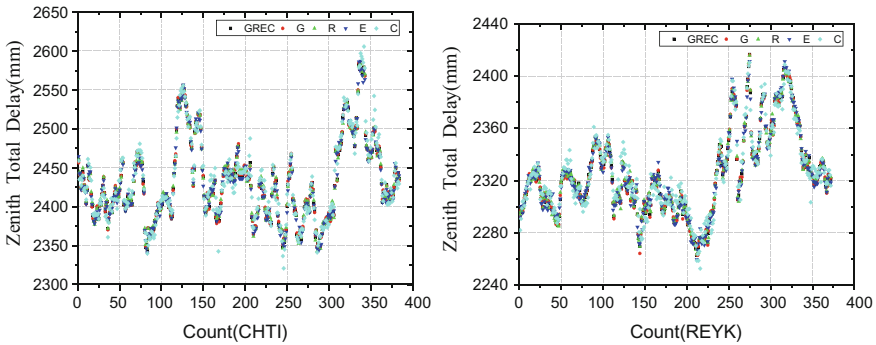


Fig. 2 PPP-ZTD series derived from individual GNSS and multiple GNSS

differences among different GNSS in estimating tropospheric zenith delay, this paper will illustrate the two aspects of linear correlation coefficient and deviation distribution. Figure 3 shows the linear correlation between the PPP ZTD series of the GNSS single system and the multi-GNSS combination at the stations REYK station and CHTI.

It can be found from Fig. 3, the GPS solution show a highly correlation with the GNSS combined solution, the correlation coefficients are 99.94 and 99.90% at the stations CHTI and REYK; The performance of the GLONASS solution is similar to

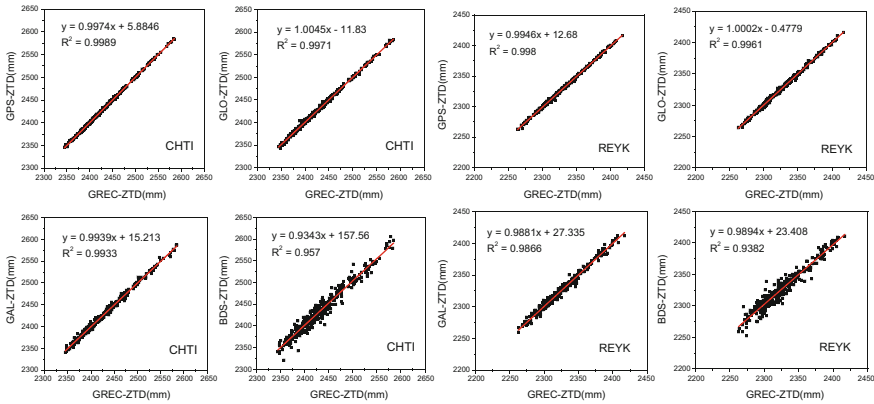


Fig. 3 Correlation of ZTD results between individual GNSS and multiple GNSS

that of GPS, the correlation coefficients are 99.86 and 99.81% respectively; Galileo solution also has good correlation, but slightly worse than GPS and GLONASS, the correlation coefficients are 99.66 and 99.33%, respectively; while BDS solution has relatively poor performance, the correlation coefficients are 97.83 and 96.86% respectively. In Fig. 4, the distribution of differences between individual GNSS solution and multi-GNSS solution is comparatively analyzed.

Figure 4 presents the distribution of ZTD differences between individual GNSS and multiple GNSS at the stations CHTI and REYK. As we can notice, the frequent counts of ZTD differences are all close to normal distribution, which means there is no significant systematic deviation in the solution. Among them, the GPS ZTD has the most concentrated deviation distribution and the average distribution is ± 5 mm, The deviations distribution of GLONASS and Galileo ZTD are ± 7.5 and ± 12 mm respectively; BDS ZTD deviation distribution is more dispersed, evenly distributed

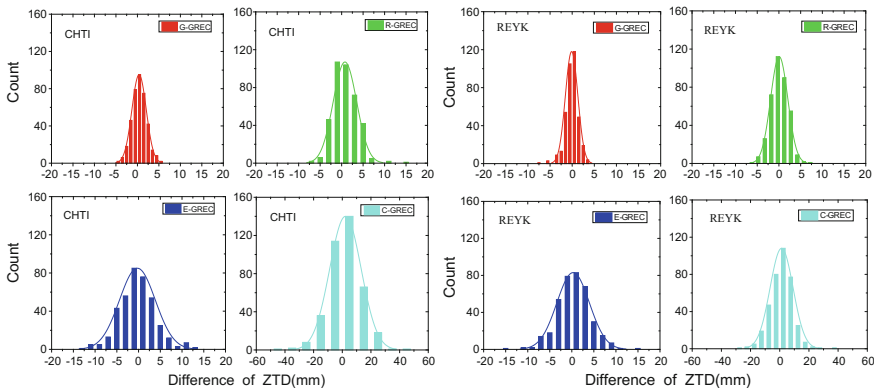


Fig. 4 Distribution of ZTD differences between individual GNSS and multiple GNSS

between ± 20 mm. In summary, there are obvious differences in the tropospheric zenith delay estimation for each GNSS system, of which GPS performs relatively best, GLONASS and GPS, are of similar precision and very close to the results of GNSS combinatorial solution. Galileo ZTD is slightly worse than the GPS ZTD and the GLONASS ZTD, which may be related to the current Galileo system has only a few visible satellites; Because of the special constellation configuration of “5 GEO + 5 IGSO + 4 MEO”, BDS ZTD performs the worst. In addition, a more stable ZTD series can be obtained by using multi-GNSS observations, which is due to the elimination of outliers and short-term fluctuation that were originally found in individual GNSS ZTD series in the multi-GNSS combination solution.

3.3 Precision Validation of GNSS PPP ZTD

In order to verify the accuracy of the tropospheric zenith delay results from the individual GNSS and multiple GNSS, the tropospheric products provided by CODE are used as reference values for this analysis. The final tropospheric products released by CODE institutions are obtained by processing more than 200 GPS stations using Bernese software PPP technology and CODE final ephemeris and clock products, the Vienna Mapping Function(VMF) is applied and the estimated interval is 2 h [19]. Figure 5 shows the results of GNSS-ZTD obtained by using the PPP technique from KIRU and HARB stations.

In Fig. 5, the CODE reference series is shown in black, and the GNSS four-system combination ZTD results are shown in red. From the comparison between the two, it can be found that the ZTD sequences obtained by the GNSS four-system combination show good consistency with the CODE reference series with a difference of millimeter. Individual GNSS solution are also compared, The solution for GPS, GLONASS, Galileo and BDS were shown by green, dark blue, blue and purple symbols. It can be seen that the ZTD series, which is solved by GPS, GLONASS and Galileo single system, have abnormal jitter at some epochs,

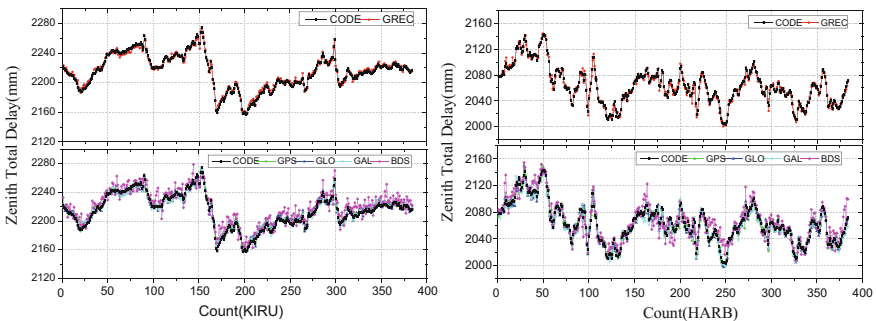


Fig. 5 GNSS ZTD results obtained from stations KIRU (left) and HARB (right)

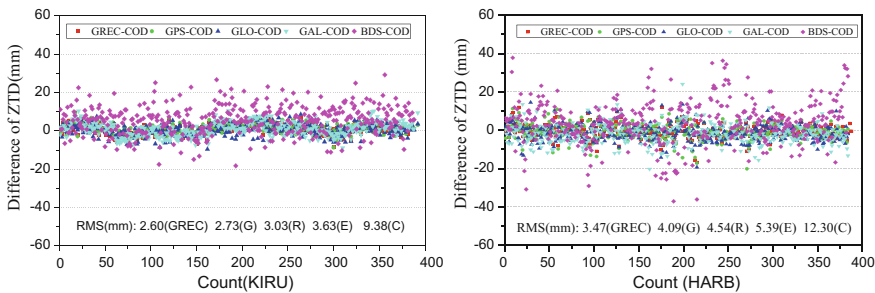


Fig. 6 Differences between GNSS ZTD and CODE ZTD at stations KIRU (left) and HARB (right)

but they still have good consistency with the CODE reference series. However, the BDS-ZTD series show large differences with the reference sequence in some periods, resulting in relatively poor consistency between the two. In order to display the deviation between the GNSS solution and the CODE reference value, this paper presents a sequence of the ZTD difference of different GNSS solution with respect to CODE ZTD in Fig. 6.

Figure 6 shows the difference between the GNSS ZTD series and the CODE reference value at the stations KIRU and HARB. It can be seen that the ZTD difference sequences obtained from the GNSS combination and each single GNSS mainly concentrate on ± 20 mm, the difference between the GNSS combination solution and the CODE reference value is the smallest, and the corresponding difference sequence is also smoother, the corresponding Root-Mean-Square (RMS) values are 2.60 and 3.47 mm respectively. BDS solution is obviously worse than other results in accuracy and stability, and the corresponding RMS values are 9.38 and 12.30 mm respectively. Moreover, there are appear several large gross error in the results of GPS, GLONASS and Galileo, which affects the stability of the ZTD series, to a certain extent, which reduces the accuracy of the solution results; the statistic results are given in Fig. 7.

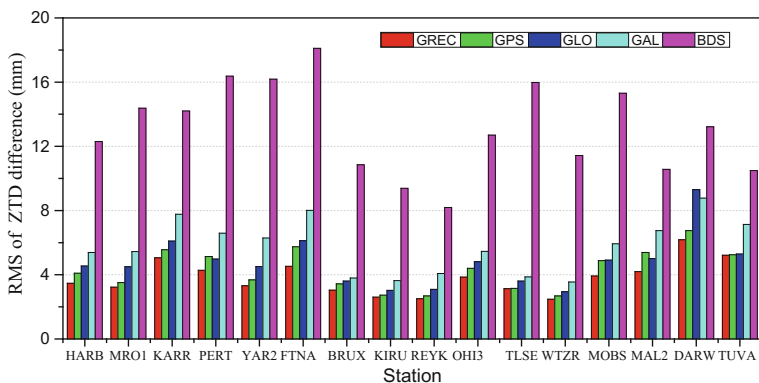


Fig. 7 RMS values of the ZTD differences of GNSS ZTD with respect to the CODE ZTD

From the statistic results given in Fig. 7, the multi-GNSS combination has the best solution, while the BDS solution is the worst. At the same time, it can also see that the three single systems of GPS, GLONASS and Galileo, as well as the multi-GNSS combination, can all obtain the millimeter accuracy. The average RMS of the GNSS four-system combination is 3.81 mm. GPS and GLONASS have similar accuracy, the corresponding RMS mean value are 4.32 and 4.77 mm respectively; despite of fewer global visible satellites, Galileo has a precision of 5.78 mm, which may be related to the stability of the Galileo satellite, while the accuracy of the BDS solution has obvious deviation with the RMS value range from 8 to 18 mm; However, though the obvious difference exist in the zenith tropospheric delay obtained from different GNSS systems, the multi-GNSS combination PPP ZTD results are significantly improved accuracy of 10.91, 19.04, 33.21 and 70.16% with respect to GPS, GLONASS, Galileo, and BDS results, respectively, This is due to the GNSS combination having greater numbers of visible satellites and more optimized satellite geometry. Therefore, this implied that multi-GNSS PPP technology has greater potential for meteorological applications such as weather monitoring and forecasting.

3.4 Precision Validation of GNSS PPP PWV

In order to verify whether the vapor products retrieved from the multi-GNSS observations can meet the requirements of modern weather applications; Radiosonde (RS) is used to evaluate the accuracy of GNSS retrieval of atmospheric water vapor results. Radiosonde data is one of the reliable methods for measuring atmospheric pressure, temperature and water vapor in situ currently, and the accuracy of the Radiosonde PWV is better than 1.5 mm. Due to the difference of the geographical location of the radiosonde station and sounding station, it can be used as an external reference to verify the accuracy of GNSS-PWV [6] when the horizontal distance between the two is within 50 km. Figure 8 plots the GNSS PWV series for the WTZR station in Europe and the DARW station in Australia.

It can be clearly seen from Fig. 8 that the PWV series obtained by the GNSS four-system combination has a good agreement with the RS reference series, and the difference is only a millimeter. In order to further analyze the GNSS PWV retrieval accuracy and taking into account the difference of sampling interval between the sounding data and the GNSS vapor, we compare the GNSS retrieval results with the same epoch of the sounding results.

According to the statistical results given in Fig. 9, we can see that the GNSS combined PWV series has a good agreement with the sounding data, with an average RMS of 1.58 mm, while the BDS PWV series has the worst performance with an average RMS of 2.52 mm. The accuracy of GPS, GLONASS and Galileo PWV retrieval is quite good, and the average RMS values are 1.63, 1.86 and 1.88 mm respectively. It is shown that the retrieval of atmospheric water vapor

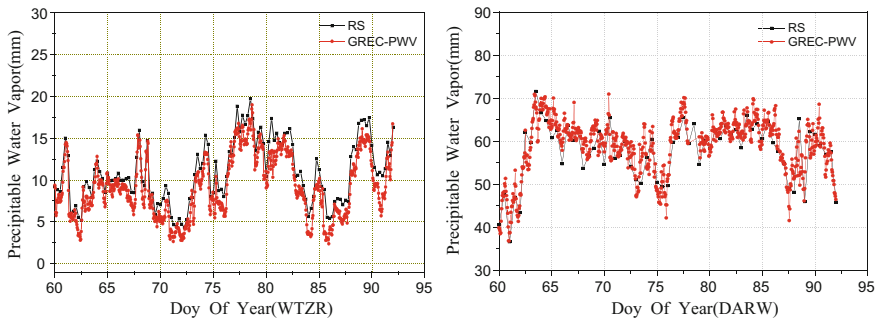
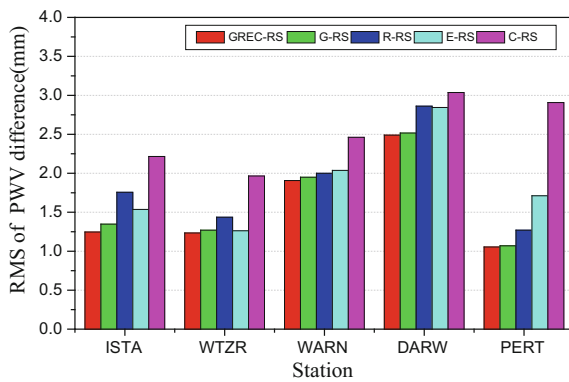


Fig. 8 PWV retrieved from multi-GNSS combination and radiosonde solutions at stations WTZR (left) and DARW (right)

Fig. 9 RMS of PWV differences of GNSS PWV with respect to the radiosonde solutions



using GNSS observations has high precision and can meet the requirements of numerical weather forecasting model and other applications. The accuracy of GNSS combination retrieval is improved by 3.45, 16.16, 16.45 and 41.78% respectively on the basis of GPS, GLONASS, Galileo and BDS. It is proved that the combined observation and PPP technology can significantly improve the accuracy of GNSS retrieve water vapor ability and credibility. Therefore, the retrieval of water vapor using multi-PPP technology can provide more meaningful reference for modern meteorological applications.

4 Conclusions

In this paper, the precise point positioning technology for combined multi-GNSS is carried out using the observation data of 30 MGEX tracking stations during the period of (DOY) 60–90, 2017. The ZTD and PWV results derived from different constellations are carefully compared and analyzed. The performance of multi-GNSS acquisition of ZTD/PWV series is also evaluated and its accuracy is

validated by applying CODE tropospheric products and Radiosonde data. The main conclusions are summarized as follows:

- (1) Compared with the multi-GNSS solution, the differences of individual GNSS ZTD series are analyzed from correlation coefficient and deviation distribution. Among them, the result of GPS ZTD series is relatively better, the accuracy of GLONASS solution is quite similar to GPS solution, which is very close to the multi-GNSS solution, and there is only several millimeter differences between Galileo ZTD series and GPS ZTD series, whereas BDS ZTD series presents the worst agreement. However, until Galileo and BDS have complete constellation configurations, the difference in tropospheric performance will be greatly improved. In addition, compared with the short-term fluctuations and outliers exist in the ZTD series of individual GNSS, the multi-GNSS solution is more reliable and robust.
- (2) In contrast with the CODE tropospheric products, The ZTD series obtained from single GNSS and multiple GNSS all show good agreement, among which, the gap among GPS, GLONASS, Galileo, GNSS four-system combination solutions and the CODE products are in the millimeter magnitude, while BDS solution has poor accuracy, the accuracy ranges from 8 mm to 18 mm. Meanwhile, the multi-GNSS combination has the highest solution accuracy with 10.91, 19.04, 33.21 and 70.16% increase over the GPS, GLONASS, Galileo and BDS single system respectively, due to the fusion of multi-GNSS has more visible satellites number and more optimized satellite geometry.
- (3) By using radiosonde data to verify the retrieval accuracy of GNSS PWV, we find that the GNSS combined PWV series has the highest coincidence, with an average RMS of 1.58 mm, while the BDS-PWV series present the worst, with an average RMS of 2.52 mm, which shows that using different GNSS data to retrieve atmospheric water vapor can have high accuracy and can meet the requirements of numerical weather forecasting model and other applications. Compared with GPS, GLONASS, Galileo and BDS solution, the results of GNSS combination solution are the best, which are improved by 3.45, 16.16, 16.45 and 41.78%, respectively. This also verifies that GNSS combination observation can significantly improve GNSS water vapor performance and accuracy.

This paper reveals that the individual GNSS has the ability to retrieve water vapor independently, but the accuracies from different individual GNSS results are different. The accuracy of multi-GNSS combined PPP technology is better, which has greater potential in weather monitoring and forecasting, and which may provide more reference for modern weather applications.

Acknowledgements The author would like to thank the Programs of the National Natural Science Foundation of China (41774025, 41731066), the Natural Science Foundation of Shaanxi Province (2016JQ4011), the Special Fund for Basic Scientific Research of Central Colleges (Grant No. CHD310826171004, Chang'an University), and the Grand Projects of the Beidou-2 System (GFZX0301040308). The authors would like to thank GFZ for providing multi-GNSS precision ephemeris and clock products, CODE for Ionosphere products and tropospheric products, American University of Wyoming for Radiosonde data and IGS MGEX for multi-GNSS observations.

References

1. Li H, Yan W, Wang QQ (2014) Application of ground-based GPS observations in water vapor monitoring. *J N China Uni Water Resour Electr Power (Nat Sci Ed)* 35(2):89–92
2. Bevis M, Businger S, Herring TA et al (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. *J Geophys Res Atmos* 97 (D14):15787–15801
3. Li X, Ge M, Dai X et al (2015) Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou and Galileo. *J Geod* 89(6):607–635
4. Ren X, Zhang K, Li X et al (2015) Precise point positioning with multi-constellation satellite systems: BeiDou, Galileo, GLONASS, GPS. *Acta Geod Cartogr Sin* 44(12):1307–1313, 1339
5. Hu L, Lin Pan (2017) Precise point positioning with combined GPS/GLONASS/BDS/Galileo. *J Navig Position* 5(1):86–90
6. Lu C, Li X, Ge M et al (2016) Estimation and evaluation of real-time precipitable water vapor from GLONASS and GPS. *GPS Solut* 1–11
7. Cai CS, Xia PF, Shi JB et al (2013) Estimating zenith tropospheric delay by using GPS/GLONASS combined with precise point positioning. *J Geod Geodyn* 33(2):54–57
8. Zhou F, Gu S, Chen W et al (2017) Comprehensive assessment of positioning and zenith delay retrieval using GPS + GLONASS precise point positioning. *Acta Geodyn Et Geomater* 14(3):317–326
9. Fan L, Zhong SM, Tan BF et al (2015) Comparison of zenith tropospheric delay estimation based on BDS and GPS precise point positioning. *J Geod Geodyn* 35(1):000067–000071
10. Han Y, Zhiwei LV, Jian XU et al (2017) Retrieval of precipitable water vapor from BDS and GPS observations. *J Navig Position* 5:39–45
11. Li X, Dick G, Lu C et al (2015) Multi-GNSS Meteorology: real-time retrieving of atmospheric water vapor from BeiDou, Galileo, GLONASS, and GPS observations. *IEEE Trans Geosci Remote Sens* 53(12):6385–6393
12. Ding W, Teferle FN, Kaźmierski K et al (2017) An evaluation of real-time troposphere estimation based on GNSS precise point positioning. *J Geophys Res Atmos* 122
13. Yuan Y, Zhang K, Rohm W et al (2015) Real-time retrieval of precipitable water vapor from GPS precise point positioning. *J Geophys Res Atmos* 119(16):10044–10057
14. Bałdysz Z, Nykiel G, Figurski M et al (2015) Investigation of the 16-year and 18-year ZTD time series derived from GPS data processing. *Acta Geophys* 63(4):1103–1125
15. Davis JL, Elgered G, Niell AE et al (1993) Ground-based measurement of gradients in the “wet” radio refractivity of air. *Radio Sci* 28(6):1003–1018
16. Gao X, Dai WJ (2014) Application of robust helmert variance component estimation to position in combination of GPS and BDS. *J Geod Geodyn* 34(1):173–176
17. Zhang Y, Tian LY, Xu JM et al (2013) Applications of robust variance component estimation in GPS/GLONASS integrated navigation. *J Geomat Sci Technol* 30(2):132–135
18. Zhang CY, Zhao XW (2017) Unified representation of GNSS observations and analysis of PPP performance. *J Geomat Sci Technol* 34(1):10–14
19. Teke K, Böhm J, Nilsson T et al (2011) Multi-technique comparison of troposphere zenith delays and gradients during CONT08. *J Geod* 85(7):395
20. Shoji Y, Kunii M, Saito K (2011) Mesoscale data assimilation of Myanmar cyclone Nargis—part II: assimilation of GPS-derived precipitable water vapor. *J Meteorol Soc Jpn* 89(1):67–88
21. Dach R, Hugentobler U, Fridez P et al (2007) Bernese GPS software version 5.2. University of Bern, Bern

Retrieval of PWV Based on GPS and Multi-level Isobaric Surface Data



Hongkai Shi, Xiufeng He and Xinyuan Wang

Abstract In GPS meteorology, the pressure and temperature parameters of GPS station are very important when obtaining accurate precipitable water vapor (PWV). However, there are few GPS stations equipped with co-located sensors for these meteorological parameters. In order to solve the problem, two methods, which is Parameter Conversion Method (PCM) and Layer Interpolation Method (LIM), were analysed and compared in this paper. The mean sea level products and multi-level isobaric surface products provided by NCEP for the period from Mar to June 2017 were used to two methods, respectively. Based on 36 GPS stations around the world which contain meteorology file and co-located radiosonde data, the experiment verified the availability and accuracy of two methods in obtaining meteorological parameters and further, PWV. Results show that: (1) both LIM and PCM works well when station height below 1600 m, the average bias of LIM is 0.67 hPa and 1.12 K compared with 0.9 hPa and 1.65 K of PCM; (2) with increasing of the station height, the difference between observed values and calculated values of pressure become larger, while the LIM has better accuracy than PCM and the robustness is better, the RMS is 2.54 and 2.91 hPa, 3.53 and 4.69 K, respectively; (3) The experiment results and analysis shown the validity of LIM and PCM, and the estimated PWV shows a higher accuracy using LIM.

Keywords GPS meteorology · PWV · NCEP · Interpolation

1 Introduction

Global positioning system (GPS) has been playing one of the most important role in water vapor detection with its all-weather, all-time capabilities and high resolution features, which is important for weather forecasting and nowcasting. In order to

H. Shi · X. He (✉) · X. Wang
School of Earth Science and Engineering, Hohai University,
8 Fochengxi Road, Nanjing 211100, China
e-mail: xfhe@hhu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_13

retrieve high accuracy result of precipitable water vapor (PWV), the pressure (P) and temperature (T) are two parameters of great relevance to PWV retrieval from GPS data [1]. However, information about the P and T cannot be available for those GPS stations that are not co-located with meteorological sensors, which means empirical models or reanalysis data are needed.

The NCEP Climate Forecast System Version 2 (CFSv2) dataset provides various parameters, such as temperature, dew point, mean sea level pressure, pressure, and other observed elements. The dataset has a vertical resolution of 37 pressure levels ranging from 1000 to 1 hPa, a horizontal resolution of $0.5^\circ \times 0.5^\circ$ and a temporal resolution of six hours (namely, at 0, 6, 12, 18 UTC), respectively [2, 3]. The data are organised in full 360° latitude circles beginning at 90° N and stepping southward to 90° S. Schueler took experiment based on NCEP data, the results verified the possibilities of estimating the meteorological through weather models [4]; Wang obtained global PWV data based on IGS troposphere product and NCEP interpolation meteorological data [5]; Jade used 6 GPS stations of India, combined NCEP/CEAR data to inverse the PWV of India area [6].

In this study, two different products, which is the mean sea level (MSL) products and multi-level isobaric surface products, were used to investigate the optimal strategy of metrology parameters interpolation. The Parameter Conversion Method (PCM) and Layer Interpolation Method (LIM) was implemented to the two products mentioned above, respectively. The products provided by NCEP for the period from Mar 2017 to June 2017 were used. And based on 36 IGS stations around the world which contain meteorology files data, the experiment verified the availability and accuracy of two methods in obtaining meteorological parameters and further, PWV.

2 Data and Method

2.1 Data and GPS Sites Distribution

In this study, a uniformly distributed GPS dataset (Mar. 2017–Jun. 2017) with a sampling rate of 30 s obtained from 29 stations globally are used to estimate ZTD parameters by GAMIT10.61 and 14 stations which contain meteorology files are used to verify pressure and temperature parameters results. The T_m parameter is calculated by GPT2w [7] model directly, which is an important factor for PWV calculation. Besides, the co-located meteorology files and PWV time series provided by SUOMINET is also available for the validation. The station details and data availability is shown in Fig. 1 and Table 1.

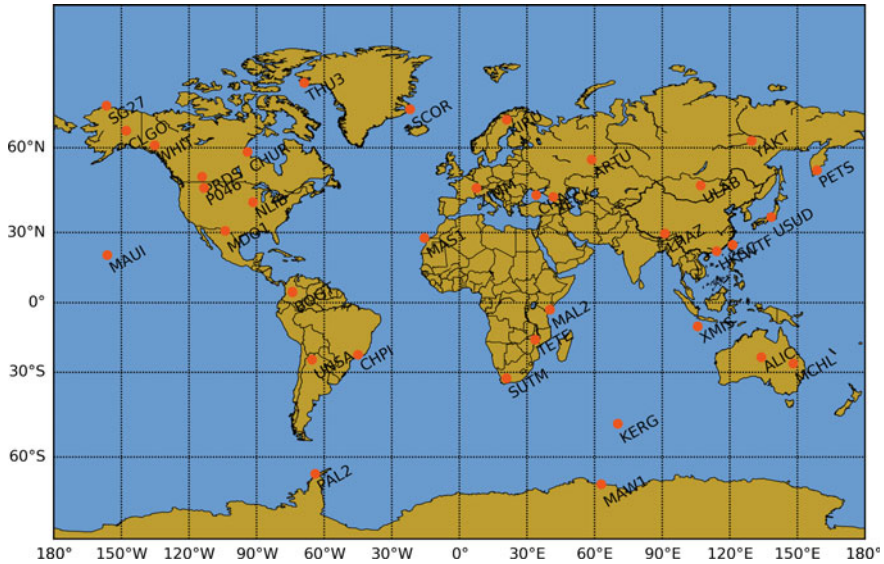


Fig. 1 Distribution of GPS stations

2.2 Layer Interpolation Method

NECP multi-level isobaric surface data provides 37 pressure layers data in each grid point including pressure, temperature and geopotential height parameters, which ranges from 1000 to 1 mbar [3, 8]. The layer interpolation method (LIM) based on pressure layer data is proposed and the steps are as follow:

1. Find the longitude and latitude data of the nearest four grid points around target point S_0 , get temperature and height data of first 16 pressure layer respectively, which ranges from 1000 to 500 mbar with the step of 25 mbar.
2. Interpolate pressure data by cubic hermite interpolation, the pressure P_{oi} of nearest four grid points S_{oi} ($i = 1, 2, 3, 4$) at the same height as S_0 are obtained. Using cubic hermite interpolation to interpolate temperature data and then temperature T_{oi} of corresponds to pressure P_{oi} are calculated.
3. Using IDW to interpolate pressure and temperature in the horizontal direction, getting the pressure and temperature values of GPS station at standard time (UTC 0, 6, 12, 18 h).
4. Basing the pressure and temperature values at standard time of GPS station, cubic spline interpolation is used to get the temperature and pressure at any time.

Table 1 GPS stations and data availability

Site	Lat (°)	Lon (°)	H (m)	GPS observation time (DOY)	Site	Lat (°)	Lon (°)	H (m)	GPS observation time (DOY)
ALIC	133.88E	23.670S	585.55	60–152	PRDS	114.29W	50.871N	1263.47	(1 day missing)
ARTU	58.560E	56.429N	253.72	(5 days missing)	SCOR	21.950W	70.485N	72.16	60–152
BOGT	74.080W	4.640N	2553.47	(1 day missing)	SG27	156.61W	71.322N	8.48	(1 day missing)
CHPI	44.985W	22.687S	619.52	(46 days missing)	SUTM	20.810E	32.381S	1759.89	60–152
CHUR	94.088W	58.759N	28.81	60–152	THU3	68.825W	76.537N	19.05	(3 days missing)
CRAO	33.990E	44.413N	340.31	60–152	TWTF	121.16E	24.953N	182.93	60–152
KERG	70.255E	49.351S	32.51	(1 day missing)	ULAB	107.05E	47.865N	1611.59	(2 days missing)
KIRU	20.968E	67.857N	361.40	60–152	UNSA	65.407W	24.727S	1223.42	(1 day missing)
MAL2	40.194E	2.9960S	8.95	(3 days missing)	USUD	138.36E	36.133N	1465.41	60–152
MAS1	15.633W	27.763N	153.95	60–152	WHIT	135.22W	60.750N	1419.42	(20 days missing)
MAUI	156.25W	20.706N	3046.51	60–152	XMIS	105.68E	10.449S	264.58	60–152
MAW1	62.870E	67.604S	29.44	60–152	YAKT	129.68E	62.030N	108.47	60–152
NLIB	91.574W	41.771N	239.43	60–152	ZECK	41.565E	43.788N	1145.25	60–152
PAL2	64.051W	64.775S	14.66	60–152	ZIMM	7.4652E	46.877N	907.12	60–152
PETS	158.65E	53.023N	78.83	(1 day missing)	HKSC	114.14E	22.322	20.227	60–152
LHAZ	91.104E	29.657N	3622	60–152	MDO1	104.01W	30.68	2004.48	60–152
MCHL	148.145E	26.3589S	534.593	60–152	TETE	33.5764E	16.1472S	150.759	60–152
P046	113.33W	47.029N	1290.302	60–152	CLGO	147.8605W	64.8738N	196.077	60–152

2.3 Parameter Conversion Method

As we retrieve temperature and pressure parameters through mean sea level (MSL) products of NCEP, there will be a serious system error if not considering the difference between mean sea level and ellipsoidal height. The MSS_CNES_CLS2011 (MCC) model is the latest mean sea level model provided by CNES/CLS which used satellite altimetry data. This model covers from 80° S to 84° N and has a 2 min resolution, the model error is less than 6 cm [9]. In this study we used 6 hourly 0.5° mean sea level products to retrieve meteorology parameters and use MCC model for the altitude correction. The specific steps are described as follows.

1. Find the longitude and latitude data of the nearest four grid points G_i ($i = 1, 2, 3, 4$) around target point G_0 and convert T from ground surface to MSL:

$$T_{MSL} = T_s + 0.0065H \quad (1)$$

where H is the geopotential height of G_i , T_s and T_{MSL} is the ground temperature and temperature of MSL, G'_i , G'_0 is the projection point of MSL of G_i and G_0 , respectively.

2. Use inverse distance weighted method to G'_i on MSL to get the temperature and pressure result of G'_0 :

$$\begin{cases} P'_{GPS} = \frac{\sum_{i=1}^4 (P'_i/d_i^2)}{\sum_{i=1}^4 (1/d_i^2)} \\ T'_{GPS} = \frac{\sum_{i=1}^4 (T'_i/d_i^2)}{\sum_{i=1}^4 (1/d_i^2)} \end{cases} \quad (2)$$

where d is the distance between G'_i to G'_0 , P'_{GPS} and T'_{GPS} is pressure and temperature of G'_i , respectively.

3. Use MCC model for the altitude correction. Get the MSL height of GPS station from MCC model using IDW then get the potential height of GPS station:

$$h = H_{GPS} - h_{mss} \quad (3)$$

4. Convert the meteorology parameters from MSL to GPS station:

$$\begin{cases} P_{GPS} = P'_{GPS}(1 - 2.26 \times 10^{-5} h)^{5.225} \\ T_{GPS} = T'_{GPS} - 0.0065 h \end{cases} \quad (4)$$

5. Use cubic spline interpolation to get a 10 min resolution result of meteorology parameters.

3 Retrieval Pressure and Temperature by NCEP Products

Meteorological data from LIM and PCM are obtained respectively and measured meteorological data are available in 14 GPS sites. Some of the interpolation results of ZIMM station are shown in Fig. 2. Results show that the parameters retrieved from LIM and PCM have a good consistency compared with the validation data. PCM shows a better accuracy than LIM in pressure of ZIMM but decrease in the last 20 days.

The average bias, RMS and STD (Standard Deviation) for pressure and temperature of 14 stations is shown in Figs. 3 and 4, which is fitted with altitudinal gradient. As for pressure parameter, result shows that the average bias of LIM and PCM is between ± 6 and ± 7.5 mbar, respectively. Accuracy of both LIM and PCM decreases as the increase of stations' altitude, the biggest bias of LIM appears in MCHL, ZIMM and UNSA, the bias of PCM increase rapidly while station altitude over 1600 m which indicates that PCM is more vulnerability to altitude. However the STD of LIM is better than PCM in almost every station. And the RMS shows that PCM has a better robustness than LIM while station altitude below 1600 m. As for temperature parameter, LIM show a better accuracy and robustness than PCM, the average bias of LIM is 0.67 hPa and 1.12 K compared with 0.9 hPa and 1.65 K of PCM.

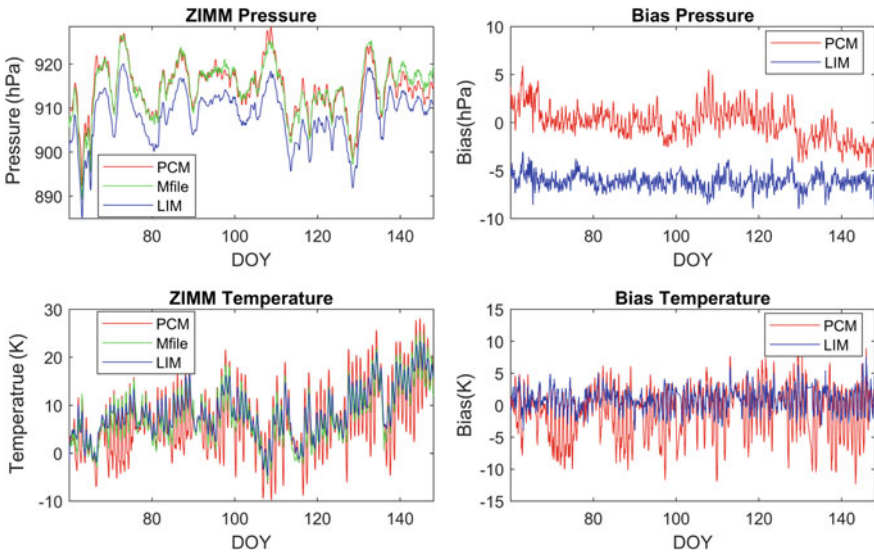


Fig. 2 Interpolation result and bias time series of pressure and temperature parameters in ZIMM station during Mar. to June 2017

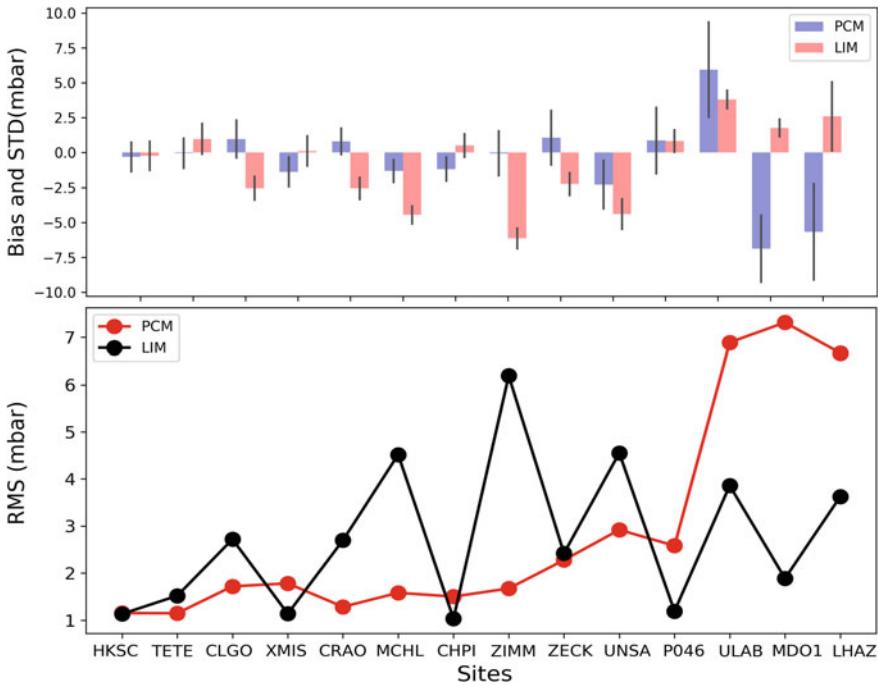


Fig. 3 Average bias, STD and RMS results of 14 GPS stations for pressure

4 Retrieval PWV by NCEP Products

29 GPS stations which distributed globally were used to estimate ZTD parameter by GAMIT10.61 because of the data availability, and considering the influence of affiliating scale average temperature (T_m), the GPT2w model was introduced to calculate this parameter. 30 min time resolution PWV results were retrieved with the usage of temperature and pressure parameters interpolated by LIM and PCM, respectively. The average bias and STD compared with validation data provided with COSMIC SUOMINET is shown in Fig. 5. Figure 6 is shown in order to investigate the relationship between retrieval PWV and geographical position. Result shows that (1) average bias of LIM is better than PCM, which is in range -3 to 2 mm and -6 to 4 mm, respectively. (2) With the increase of longitude and latitude the accuracy decreases generously, however LIM shows higher accuracy at high latitude and altitude area. The biggest bias of PCM appears in high latitude areas such as KIRU and THU3 and high attitude area such as ZIMM, which indicates that PCM is more susceptible to altitude and latitude. (3) However LIM always shows a smaller and stable bias which indicate that LIM has a better robustness for PWV retrieval.

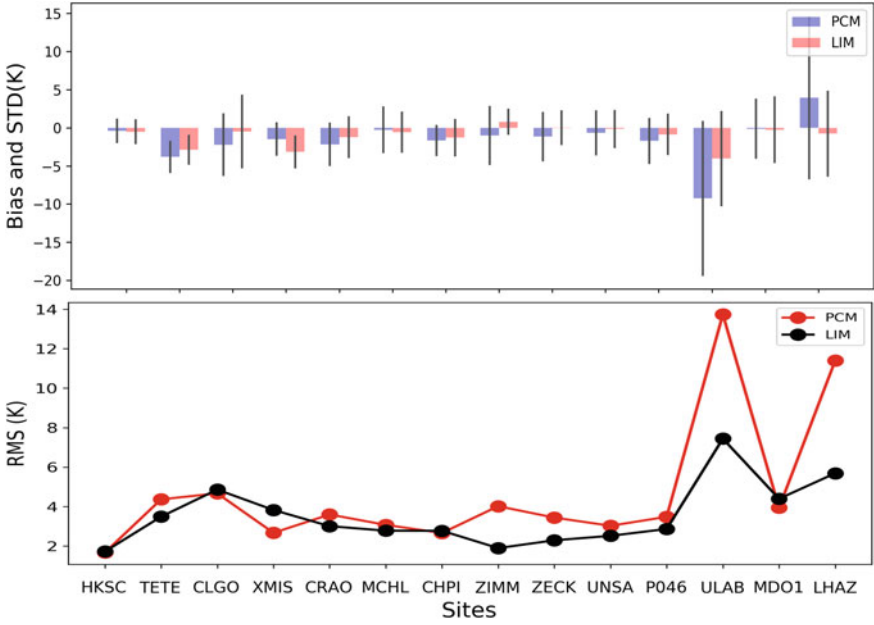


Fig. 4 Average bias, STD and RMS results of 14 GPS stations for temperature

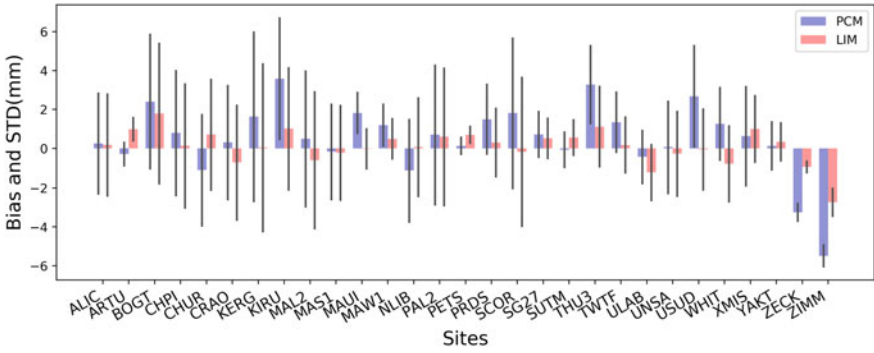


Fig. 5 Average bias and STD results of 29 GPS stations for PWV

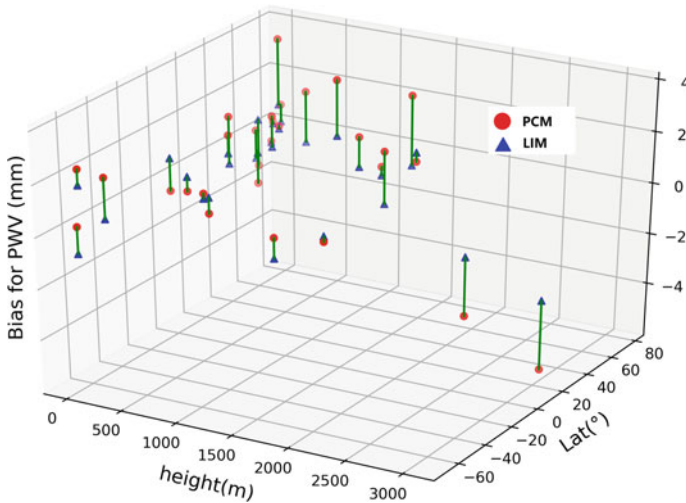


Fig. 6 Accuracy comparison between PCM and LIM

5 Conclusions

The pressure and temperature are two parameters of great relevance to PWV retrieval from GPS data. However, there are few GPS stations equipped with co-located sensors for these meteorological parameters, which leaves amount of researches in this area. In this study, we investigated the optimal GPS PWV retrieval method using MSL data and multi-level isobaric surface data provided by NECP CFSv2 with two different interpolation strategy, PCM and LIM. The accuracy and stability of each product was explored with the usage of GPS data (Mar. to June 2017) obtained from 29 GPS stations distributed globally. The T_m parameter was considered by using GPT2w model. The co-located meteorology data and PWV time series provided by COSMIC SUOMINET was used as validation. The results shows that: (1) Both LIM and PCM work well for meteorology parameters when station height below 1600 m, the average bias of LIM is 0.67 hPa and 1.12 K compared with 0.9 hPa and 1.65 K of PCM; (2) With increasing of the station height, the difference between observed values and calculated values of pressure become larger, while the LIM has better accuracy than PCM and the robustness is better, PCM is more vulnerability to altitude and latitude, the RMS is 2.54 and 2.91 hPa, 3.53 and 4.69 K, respectively. (3) The estimated PWV shows a higher accuracy using LIM.

References

1. Zhang H, Yuan Y, Li W, Ou J, Li Y, Zhang B (2017) GPS PPP-derived precipitable water vapor retrieval based on Tm/Ps from multiple sources of meteorological data sets in China. *J Geophys Res Atmos* 122(8):4165–4183
2. Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou Y, Chuang H, Iredell M, Ek M, Meng J, Yang R, Mendez M, Van Den Dool H, Zhang Q, Wang W, Chen M, Becker E (2014) The NCEP climate forecast system version 2. *J Clim* 27(6):2185–2208
3. Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou Y-T, Chuang H-Y, Iredell M, Ek M, Meng J, Yang R, Mendez MP, van den Dool H, Zhang Q, Wang W, Chen M, Becker E (2011) NCEP climate forecast system version 2 (CFSv2) 6-hourly products. Research data archive at the national center for atmospheric research, Computational and Information Systems Laboratory, Boulder, CO. Available online at <https://doi.org/10.5065/D61C1TXF>
4. Schueler T (2001) On ground based GPS tropospheric delay estimation. Dissertation for the Doctoral Degree, University der Bundeswehr München, Neubiberg
5. Wang J, Zhang L, Dai A, Hove TV, Baelen JV (2007) A near-global, 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements. *J Geophys Res Atmos* 112(D11):D11107
6. Jade S, Vijayan MSM (2008) GPS-based atmospheric precipitable water vapor estimation using meteorological parameters interpolated from NCEP global reanalysis data. *J Geophys Res Atmos* 113(D3):D03106
7. Böhm J, Möller G, Schindelegger M, Pain G, Weber R (2015) Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solut* 19(3):433–441
8. Wang Y, Yang K, Pan Z, Qin J, Chen D, Lin C, Chen Y, Lazhu, Tang W, Han M, Lu N, Wu H (2017) Evaluation of precipitable water vapor from four satellite products and four reanalysis datasets against GPS measurements on the Southern Tibetan Plateau. *J Clim* 30(15):5699–5713
9. Schaeffer P, Faugere Y, Legeais J, Ollivier A, Guinle T, Picot N (2012) The CNES_CLS11 global mean sea surface computed from 16 years of satellite altimeter data. *Mar Geodesy* 35:3–19

Multipath Mitigation Method in GPS/BDS Deformation Monitoring Based on Observation Domain



Qinggen Yi, Guoli Lin, Wanke Liu and Jiuhui Pan

Abstract In GNSS real-time high-precision monitoring, the multipath spatial correlation is weak, it is difficult to eliminate the multipath error by double difference, and the traditional method based on the sidereal filtering in the coordinate domain to mitigate multipath error of MEO satellite will no longer be applicable. Aiming at the above problem, this paper storage single difference observation residuals after ambiguity resolved in real-time baseline resolution, where storage residual information for 1d in GPS, BDS GEO, IGSO, and BDS MEO satellite for 7d. When the residual information database initialization is complete, each satellite multipath error sequence model will be established from the residual database by using the low-pass filter method, then mitigating multipath error of the corresponding satellite carrier phase observations in real time. The measured data show that, based on the mitigating multipath of observation domain, the real-time monitoring accuracy of the 3D position can be improved by about 50% after mitigating multipath error in real-time.

Keywords GPS/BDS · Deformation monitoring · Multipath · Observation domain

1 Introduction

Most of the strong correlation errors in deformation monitoring can be eliminated by means of double difference. However, multipath errors in propagation process depend mainly on the spatial geometry of the stations, satellites and peripheral reflectors, making the multipath of the stations have do not with the correlation,

Q. Yi (✉) · G. Lin · J. Pan
Hi-Target Positioning Technology Co., Ltd., Guangzhou 511400, China
e-mail: sggysq2014@qq.com

W. Liu
School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
e-mail: WkLiu@sgg.whu.edu.cn

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_14

due to the multipath still remains in the observation after double difference, it will weaken the monitoring accuracy in real-time.

Aiming at the multipath error in GNSS deformation monitoring environment, the method of data processing mitigates multipath errors mainly from the observation domain and the coordinate domain. In the coordinate domain, sidereal day filtering method is more mature; Instead of the conventional sidereal day period, Choi, according to the broadcast ephemeris to obtain the average period of observation satellite as the sidereal day filtering translation period, which is better applied to mitigate multipath in high-frequency coordinate sequence [1]. Axelrad et al., according to the minimum direction cosine of satellites relative to the station in the adjacent day to determine the repetition period of each satellite, this method can mitigate multipath with the sidereal day repetition characteristics and achieve a better result than the average period filtering method [2]; The key of the sidereal day filtering method is to determine the translation period and the filtering de-noising algorithm, while the filtering de-noising process is more complicated and there is no common method [3]; In addition, the sidereal day filtering method can only process the satellite with a sidereal day of repetition period. For example, the method of mitigating multipath is no longer applicable for the BDS MEO satellite with orbit repeat period of 7 sidereal days [4]; Sidereal filtering method do not specifically consider the each satellite multipath repetition period. Therefore, some scholars proposed to mitigate multipath errors in the observation domain. In observation domain, Ragheb et al., Zhong et al., Lau et al., Dong et al., They established multipath error time series model of each satellite from the satellite observation residuals which are extracted in the reference day, The model can mitigate the carrier phase multipath error of each satellite in the repetition orbit period in real time [4–7]. In the field of deformation monitoring, the methods of mitigating multipath error in observation domain are mostly applied to the single-system level of GPS and BDS, but not the combined system of GPS and BDS. This article will start with the deformation monitoring model, based on the observation residuals, we will focusing on the analysis of the characteristics of GPS and BDS multipath errors. Moreover, based on the single difference residuals observations, the low-pass filtering method is used to establish the GPS/BDS deformation monitoring multipath error mitigation model in observation domain. It will further improve the accuracy of GPS/BDS real-time monitoring.

2 Mathematical Model

The essence of deformation monitoring is short baseline resolution, most of its basic mathematical model are based on the double-difference within the frequency and system, the following formula briefly gives its carrier phase double difference observation equation, the specific formula is expressed as:

$$\lambda_f \nabla \Delta \Phi_{ij}^{sk} = \nabla \Delta \rho_{ij}^{sk} + \nabla \Delta M p_{ij}^{sk} + \lambda_f \nabla \Delta N_{ij}^{sk} + \varepsilon_{ij}^{sk} \quad (1)$$

In the above formula, λ_f is carrier phase wavelength, $\nabla \Delta \Phi$ is double difference carrier phase observation, $\nabla \Delta M p$ is relative multipath error of double difference, $\nabla \Delta N$ is double difference ambiguity, ε is measure noise; troposphere, ionosphere and other strong correlation error can be eliminated completely by double-difference. Due to deformation monitoring body has a strong constraint between the continuous observation epoch, the parameter estimation model will use the least-squares filter model based on prior information, the specific formula is expressed as:

$$\begin{aligned} V &= Ax - L, & P_L \\ U &= W - Bx, & P_W \end{aligned} \quad (2)$$

In the above formula, A is design matrix, V and U is residual vector, L is observation subtracts calculation, W is constant vector, P_L is the matrix of current epoch, P_W can be obtained by inversing the variance-covariance matrix of the parameter from previous epoch, B can calculate by the relationship of estimated parameters between the current and previous epoch in real time. According to the formula (2), it can get the monitoring body position information in real time [8], the solution of (2) is expressed as:

$$x = (A^T P_L A + B^T P_W B)^{-1} (A^T P_L L + B^T P_W W) \quad (3)$$

2.1 Obtained Observed Residuals

After obtaining the coordinate and ambiguity parameters from the formula (3) in each epoch, next, it can obtain the double difference residuals of observed satellite, the specific formula is expressed as:

$$v = \Delta l_j^{sk} d\Delta x_{ij} + \Delta m_j^{sk} d\Delta y_{ij} + \Delta n_j^{sk} d\Delta z_{ij} + \lambda_f N_{ij}^{sk} - \lambda_f \nabla \Delta \phi_{ij}^{sk} + M p_{ij}^{sk} + \varepsilon_{ij}^{sk} \quad (4)$$

3 Multipath Modeling

In the case of double difference multipath which is based on double difference satellite pairs, when the reference satellite PRN of the current epoch is different from the reference satellite PRN before its repetition period in the actual solution, the actual correction process needs to make a reference transformation to the double difference multipath which will match the current epoch reference satellite, that making it more difficult to mitigate the double difference multipath error. Therefore,

this article further to obtain single difference multipath mitigation model based on each satellite respectively, according to the algorithm of double difference residual transforming to single difference residual where the constraint that the square sum of the residuals of the single difference equal to zero is introduced [6], double difference observation residuals can be converted to the single difference observation residuals, the transform formula is expressed as:

$$\begin{bmatrix} w_1 & w_2 & w_3 & \cdots & w_n \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ & & & \cdots & \\ 1 & 0 & 0 & & -1 \end{bmatrix} \begin{bmatrix} S_{AB}^1 \\ S_{AB}^2 \\ S_{AB}^3 \\ \cdots \\ S_{AB}^n \end{bmatrix} = \begin{bmatrix} \sum w_i S_{AB}^i \\ DD_{AB}^{12} \\ DD_{AB}^{13} \\ \cdots \\ DD_{AB}^{1n} \end{bmatrix} \quad (5)$$

where w_i is the elevation weighting function value. when the number of observing satellites is enough, multipath errors in the observation can be considered as high-frequency random noise, the weighted average of single difference observation residual should be zero theoretically [9], That it means $\sum w_i S_{AB}^i = 0$. However, the number of observed satellite is limited in actual, further verification will be carried out in this article. Since the deformation monitoring model in this paper is based on the double difference model, therefore, weighted square sum of single difference residual for each system is statistics separately. The result is shown in Fig. 1, we can see the magnitude of the weighted sum of squares of the single difference residuals is 10^{-17} which is closed to zero. So, we can extract the single difference residuals from the formula (5), this single difference residual can be considered as only containing the multipath error and random noise of the satellite. In addition, multipath errors of single difference residuals do not depend on reference satellites and mitigate better than double difference multipath errors theoretically.

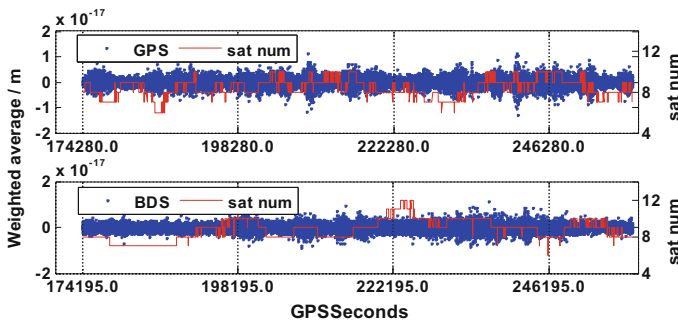


Fig. 1 Weighted average time series of the single difference residual

3.1 Multipath Error Analysis

In this paper, an accurate multipath mitigation model is established based on the observation residuals of each satellite, and then multipath sequence in the repeated period is analyzed in the sky map, so as to verify the repeated characteristics of multipath errors in deformation monitoring and further to verify the theoretical feasibility of mitigating multipath error based on observation domain. In this paper, the average filter de-noising method is used to establish a low frequency multipath mitigation model for removing high frequency random noise from the satellite residuals. According to the above method, the single difference residuals of two consecutive of two days which interval is 7 days from engineer monitoring data of bridge pier in China are subjected to de-noising high frequency noise, and the extracted low frequency multipath error time series are mapping in sky map. GPS G01, BDS GEO C01, and MEO C12 are analyzed here, the experimental results are as follows: it can be seen from Fig. 2, that the variation of the adjacent multipath of GPS satellite tends to be consistent basically; In Fig. 3, although the BDS GEO satellite is stationary with respect to the earth, it still has relatively small fluctuations with respect to the station. The change of single difference multipath is generally flat and have a strong correlation in adjacent days. As can be seen in Fig. 4, the adjacent days for MEO satellites do not have repetitive characteristics, but show a

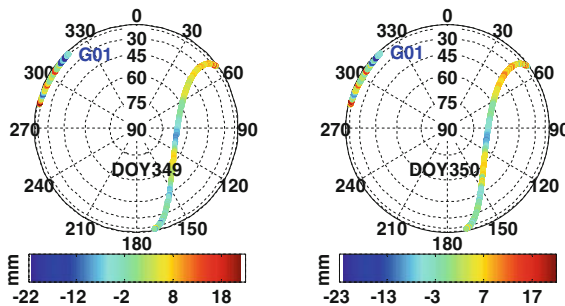


Fig. 2 The sky map of G01 single difference multipath adjacent days sky view

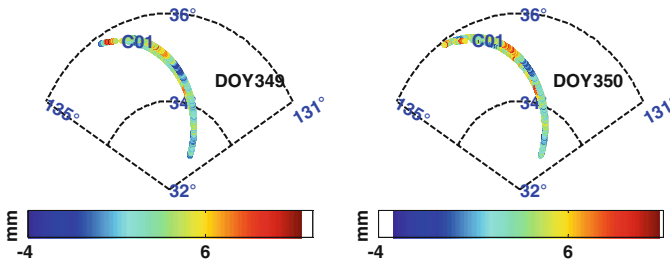


Fig. 3 The sky map of C01 single difference multipath in adjacent days

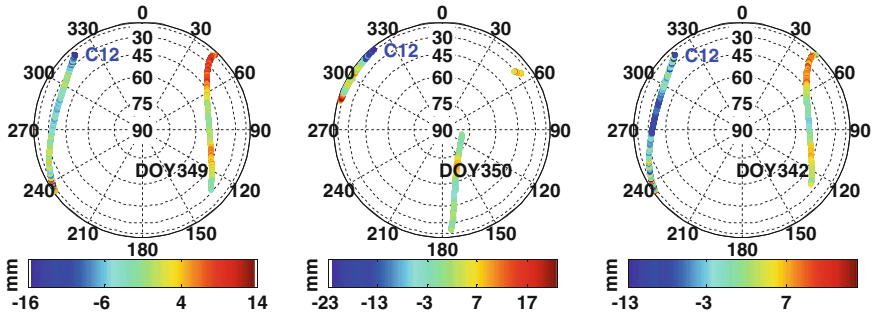


Fig. 4 The sky map of C12 single difference multipath in adjacent days

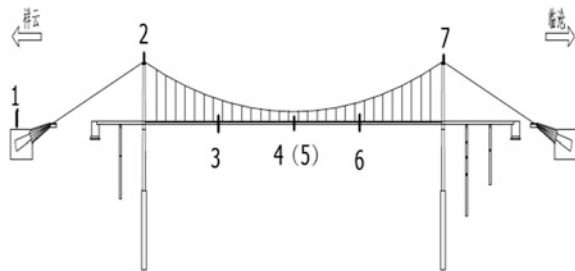
strong correlation between DOY342 and DOY349, which is consistent with the fact that 13 times orbital period of MEO satellite is equal to the seven sidereal day, that is to say, the period of repetition of the MEO satellite relative to the fixed station is seven sidereal days.

4 Experiment and Analysis

4.1 Experimental Data

Based on the Hi-Monitor of self-developed real-time dynamic deformation monitoring software, this paper adds the multipath error mitigation module in the real-time resolution algorithm to verify the correctness of the proposed algorithm. Experimental data is received from the data of bridge health monitoring in real time, the actual scene of the bridge monitoring project shown in Fig. 5, the whole project has six monitoring points and one reference point, 2# and 7# are the pier monitoring point, we choose 2# pier monitoring points as the experimental data to verify algorithm of this paper, the base station of project located in the top of the bridge management station which is about 0.1 km of distance from the bridge, that is 1# location in the Fig. 5; The experimental data sampling rate is 15 s, the cut off

Fig. 5 Distribution of the bridge monitoring points



elevation angle is 10° , the receiver type is BX dual-frequency, and the base station and monitoring station of height difference is about 36 m, moreover, the current data is analyzed for 2 h period post-processing resolution by high-precision data commercial software, the results show that the bridge pier has been not deformation basically of the current day.

4.2 Experimental Program

Multipath error mitigation algorithm verification process is as follows: First, we established the MEO satellite multipath sequence model corresponding to the 7 days ago and the GPS satellites, BDS GEO and IGSO satellites of previous day respectively. Which are used as the multipath mitigation models in real-time resolution of the current day. Secondly, finding the corresponding multipath error correction values from the single difference satellite multipath time series model database according to the PRN number of the synchronous observation satellite and its repetition period, then, mitigating it to the carrier phase observation of the current epoch in real time resolution. Finally, the calculating standard deviations of the coordinate bias and the observation residuals sequence before and after mitigating multipath are used as the verification standard of the reliability and correctness of the algorithm.

4.3 Result Analysis

Figure 6 shows the coordinate bias time series before and after mitigating the multipath error, it can be seen from the figure that the time sequence of the coordinate residuals mitigated by the single difference multipath model is more gradual,

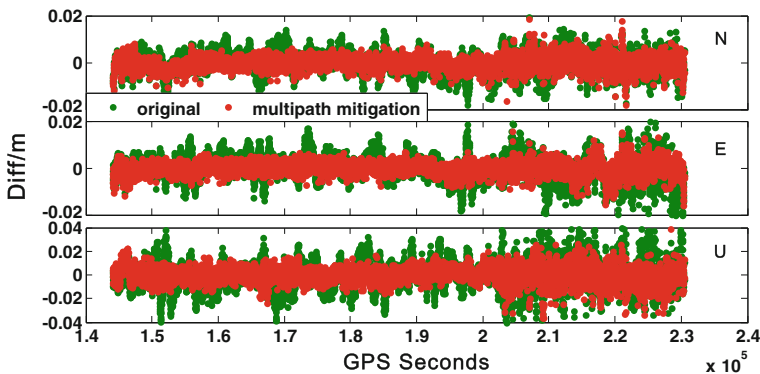


Fig. 6 Coordinate sequences before and after mitigating multipath based on observation domain

Table 1 Accuracy analysis and Correction effect before and after mitigating multipath (units: mm)

	Original	Multipath mitigation
N	4.67	2.81 (40%)
E	5.75	2.84 (51%)
U	13.16	6.90 (48%)
3D	15.10	7.80 (48%)

and the jitter epoch caused by multipath is reduced entirely. In addition, in order to more intuitively analysis the coordinate accuracy and correction performance after mitigating the multipath, we analysis the statistical accuracy of coordinates before and after mitigating multipath. The standard deviation of the coordinate sequence in Table 1 shows: in this short baseline solution after mitigating multipath by the single difference multipath model of the observation domain in real-time, the STDs in the horizontal directions N and E and in the direction U can be increased by 40, 51 and 48% respectively; The above results show that the multipath mitigating algorithm based on single difference mode is correct and reliable, which can mitigate the multipath in the coordinate sequence well.

In order to further validate the effect of mitigating multipath errors in real-time, this paper analyzes the precision of residual sequences after mitigating multipath based on observation domain. The blue curve on the left in Figs. 7 and 8 represents the uncorrected multipath and the red curve on the right represents the correction, as can be seen in Table 2, the standard deviation of the GPS satellite observations residual sequence after mitigating multipath can be reduced by about 40% entirely. For the BDS GEO satellite in Fig. 8, due to the small change in geometric configuration with respect to the station space, the multipath error caused by this project environment is relatively small, so the correction of magnitude is correspondingly smaller. The multipath of MEO satellite observation residuals has also been weakened well, it shows the multipath model of MEO satellite which

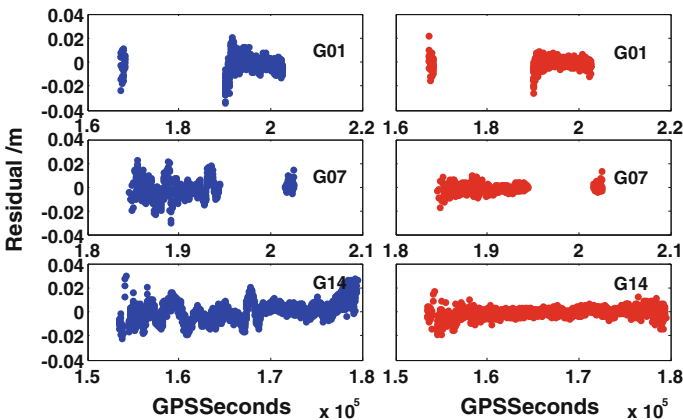


Fig. 7 Residual sequence of GPS observations before and after mitigating multipath

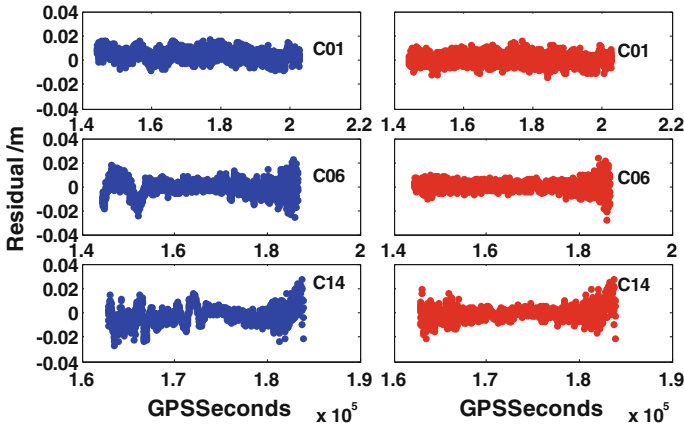


Fig. 8 Residual sequence of BDS observations before and after mitigating multipath

Table 2 Accuracy analysis and Correction effect before and after mitigating multipath (units: mm)

Sat ID	Original	Multipath mitigation	Mitigation effect (%)
G01	3.92	2.46	37
G07	8.27	6.43	22
G14	7.82	3.69	53
C01	3.56	3.15	12
C06	5.90	3.68	38
C14	7.37	5.67	23

established by a period of seven sidereal days is feasible. In addition, the magnitude of multipath mitigation for BDS GEO satellites is smaller than that of GPS satellites, which may be due to the large observation noise of GEO satellites and GEO satellite orbit accuracy is low relatively [10], it is difficulty to establish multipath model, moreover, the multipath caused by GEO in the observation environment is relatively small, so that the correction amplitude is correspondingly smaller.

5 Conclusion

Based on the Hi-Monitor which is a high precision and real time deformation monitoring software independently developed by the author, the multipath mitigation algorithm based on the observation domain is added in this monitoring algorithm. Based on the precise multipath model extracted before each satellite’s repetition period, so as to mitigate the multipath on the carrier phase observation of each satellite in real-time. According to the measured data verify, the following two conclusions can be drawn:

The multipath mitigating algorithm in the single difference observation domain is more rigorous than the traditional sidereal day filtering algorithm in the coordinate domain theoretically, not only multipath mitigation can be performed on satellites with a repetition period of a sidereal day, but also MEO satellites with a repetition period of seven sidereal days, which can handle the multipath errors of BDS satellites well.

In the future process of BDS globalization, it is planned that BDS will have 27 MEO satellites by the end of 2020. Therefore, establishing multipath mitigation model of each satellite in the observation domain and mitigating multipath errors in real-time deformation monitoring will be a new trend application of fusion BDS in high precision of deformation monitoring.

Acknowledgements This work was supported by the “National key R&D project” (No. 2017YFD0700400, No. 2017YFD0700402) and Applied “Technology R&D special” (No. 2015B010131009).

References

1. Choi K (2004) Modified sidereal filtering: implications for high-rate GPS positioning. *Geophys Res Lett* 31(22):L22608
2. Axelrad P, Larson K, Jones B (2005) Use of the correct satellite repeat period to characterize and reduce site-specific multipath errors. In: *Proceedings of the ION GNSS*
3. Lau L (2013) Comparison of measurement and position domain multipath filtering techniques with the repeatable GPS orbits for static antennas. *Surv Rev* 44(324):9–16
4. Ye S, Chen D, Liu Y et al (2015) Carrier phase multipath mitigation for BeiDou navigation satellite system. *GPS Solut* 19(4):545–557
5. Ragheb AE, Clarke PJ, Edwards SJ (2007) GPS sidereal filtering: coordinate- and carrier-phase-level strategies. *J Geodesy* 81(5):325–335
6. Zhong P, Ding X, Yuan L et al (2010) Sidereal filtering based on single differences for mitigating GPS multipath effects on short baselines. *J Geodesy* 84(2):145–158
7. Dong D, Wang M, Chen W et al (2016) Mitigation of multipath effect in GNSS short baseline positioning by the multipath hemispherical map. *J Geodesy* 90(3):255–262
8. Cui X, Yu Z, Tao B et al (2005) *Generalized surveying adjustment*. Press of Wuhan University
9. Xue Z (2012) *A study of key technology for dynamic deformation monitoring using GNSS*. PLA Information Engineering University, Zhengzhou
10. Liu W, Ren J et al (2016) Accuracy assessment of BDS signal-in-space range errors in 2013–2015. *J Natl Univ Def Technol* 38(3):1–6

Analysis of Sea Surface Multipath and Impact on GNSS Precision Positioning



Zhengxun He, Shuangcheng Zhang, Yang Nan, Kai Liu, Wei Qu, Huilin Wu and Qi Liu

Abstract With the rapid development of GNSS technology in the field of atmospheric and ocean remote sensing, especially the rise of GNSS remote sensing technology such as GNSS-R, which has promoted the wider application of coastal GNSS stations, such as crustal deformation, absolute sea level change, tide level changes monitoring, air-sea exchange and other fields. In view of the current research on the sea surface multipath and its impact on the precision of GNSS positioning is relatively few, in this paper, we selected the United States PBO network of coastal CORS station SC02 and its surrounding three PBO stations ALBH, P438 and P439 to analyze the characteristics of multipath effects of the coastal GNSS stations and their time series features. Then the stations are grouped and GAMIT is used for baseline solution. By setting different cutoff elevation angles, comparing the results of baseline solution and the residuals of multipath effects to explore whether the multipath effect of sea surface has a significant impact on the precision positioning of coastal GNSS stations. The result shows that when the cutoff elevation angle is selected from 10° to 15° , the result of the baseline solution is the best, and the effect of sea surface multipath on GNSS precision positioning is the least. Through the experimental, it is hoped to provide a reference for the future deployment of GNSS stations and GNSS for tide level real-time monitoring in coastal areas in the future, and to further promote the integrated application of coastal GNSS stations.

Keywords GNSS remote sensing · Sea surface multipath effect
Coastal stations · GNSS precision positioning

Z. He · S. Zhang (✉) · K. Liu · W. Qu · H. Wu · Q. Liu
College of Geology Engineering and Geomatics, Chang'an University,
Xi'an, China
e-mail: shuangcheng369@vip.163.com

Y. Nan
Research Center of GNSS, Wuhan University, Wuhan, China

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_15

1 Introduction

The Global Navigation Satellite System (GNSS) is widely used in positioning, navigation and timing (PNT) for its all-weather, near-real-time, high-precision and sustainably transmitted L-band signals. With the gradual improvement of navigation satellite systems, the increase of constellations and observation stations, its application fields are more and more widely. GNSS technology has been proved can not only for positioning, timing and navigation, but also for remote sensing with reflection signals, the potential application value of GNSS technology in the field of atmospheric and ocean remote sensing has become increasingly prominent [1, 2].

The concept of GNSS reflectometry was first proposed by Martin-Neira of the European Space Agency (ESA) in 1993. Its basic principle is that GNSS surface reflectance signals are received by receivers together with direct signals and the delay between them can be used for interferometric measurements, which namely PARIS (passive reflectometry and interferometry system) [3]. In recent years, the emerging remote sensing methods represented by the GNSS-R (GNSS Reflection) technology have been proved to be useful in the field of crustal deformation monitor, absolute sea level change, tide level monitoring and air-sea exchange [4–7].

Among various error sources of high-precision GNSS positioning technology, many errors such as ephemeris error, satellite clock error, atmospheric refraction error and receiver clock error can be eliminated or reduced by correcting the model or the difference technique. However, the multipath effect could not be eliminated or reduced by the above method for there is no correlation between the multipath effects at both ends of the baseline, which makes the multipath effect a major source of error for high-precision GNSS measurements [8, 9]. When the GNSS station is located at the onshore, the accuracy of GNSS positioning is inevitably affected by the multipath effect due to factors such as high reflectance of sea surface and complex sea environment. At present, the researches on the multipath effects of the sea at home and abroad are mainly from two aspects - digital signal processing and multipath modeling. The research methods are mainly based on the repeatability of multipath effects in time domain, which has some limitations. Moreover, its research object is not coastal GNSS stations [10–14].

With the development of GNSS technology in the field of ocean remote sensing, the application of coastal GNSS stations will continue to expand, and the deployment of coastal GNSS stations will surely increase day by day. By then, a high-density, large-scale global GNSS coastal observatory network will be formed. In view of the current situation that there is relatively few research on sea surface multipath effect and impact on GNSS precision positioning, in this paper, we selects the coastal CORS station SC02 in the PBO network and three other PBO stations ALBH, P438 and P439 in the surrounding area to analyze the sea multipath effect characteristics and time series features and to explores the effect of the multipath effect of the sea surface on the precision positioning of the GNSS, which is expected to provide some reference for the future deployment of coastal GNSS stations and GNSS for real-time monitoring of tide levels, and further promote the development of integrated applications of coastal GNSS stations.

2 GNSS Multipath Effect

The GNSS receiver antenna receives more than satellite direct signals during the observation process and also receives reflection signals from various media around the station, such as buildings, the reflection of water surface. These reflected signals and direct signals interfere with each other and the observed values deviate from the true values to produce so-called “multipath errors”. The effect of the interference delay caused by multipath signal propagation is called multipath effect (see Fig. 1).

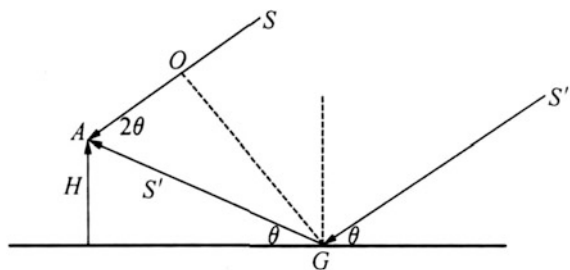
As shown in Fig. 1, if the antenna receives both the direct signal S from the satellite and the reflected signal S' which reflected by the ground simultaneously. Obviously, the path of these two signals through is different, the extra path through by reflecting signal recorded as Δ , which can be seen from Fig. 1:

$$\begin{aligned} \Delta &= GA - OA = GA(1 - \cos 2\theta) \\ &= \frac{H}{\sin \theta} (1 - \cos 2\theta) = 2H \sin \theta \end{aligned} \tag{1}$$

In Eq. (1), GA is the path length of the reflected signal S' , OA is the path length of the direct signal S , H is the elevation of the antenna from the ground, and θ is the incident angle of the reflected signal.

The effects of multipath effects on pseudorange observations can reach tens of meters, while the impact on carrier phase observations is two orders of magnitude smaller. In high-intensity reflection environment, multipath error can easily lead to cycle slip, and even signal loss of lock. The effect of multipath effect is related to the reflection coefficient. Different reflection surfaces have different reflection coefficients. The multipath effect is also related to the distance between the antenna and the reflection source. Since the electromagnetic wave propagates in the atmosphere, there will be energy attenuation and the error of multipath effect will decrease as the distance increases. In addition, the multipath effect is also related to the satellite elevation angle. The lower the satellite elevation angle, the greater the multipath effect will be [13, 15].

Fig. 1 Diagram of multipath effect



3 Analysis of Coastal GNSS Multipath Effect

When the GNSS station is deployed onshore, the signal received by the station is a composite signal which consists of a direct signal and a reflected signal reflected by the sea surface. Therefore, there is inevitably existence sea surface multipath effect. In order to explore the multipath effect characteristics of coastal GNSS, we selected four stations with different station environments in the PBO network in the United States—coastal station SC02 and three stations ALBH, P438 and P439 in the periphery to conduct research and analysis (see Fig. 2).

Figure 2 followed by SC02, ALBH, P438 and P439 station environment diagram of the station. Among them, the SC02 station is located onshore, which can receive reflected signals from the sea surface in a wide range. ALBH station is located in the plains area, the northeast direction of the trees shelter serious. The P438 station is located on hard rock with low grass around and observation environment is good. The P439 station is around the soil, which surrounded by no obvious obstruction.

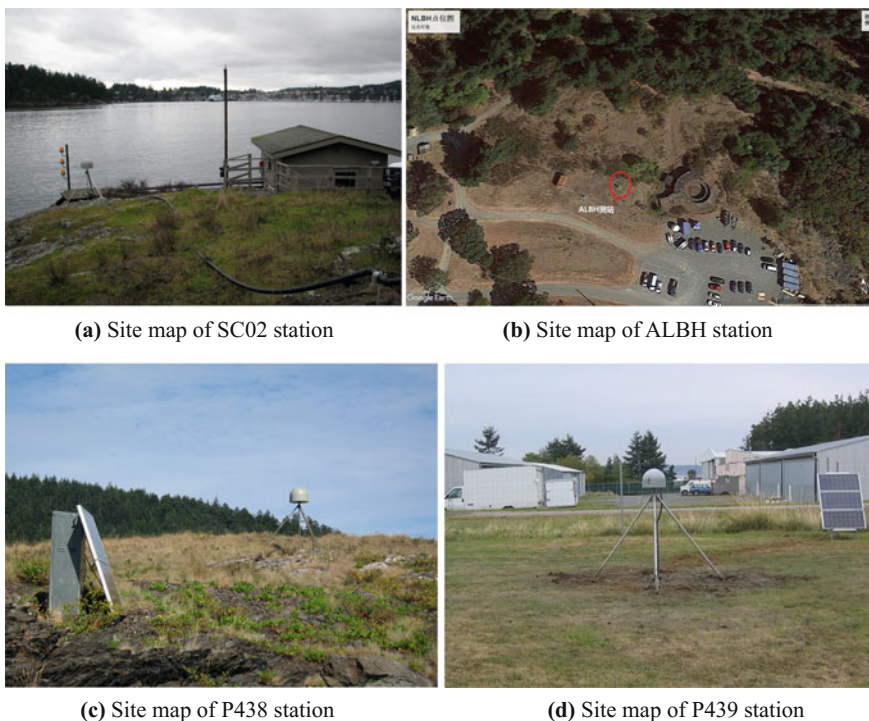


Fig. 2 Diagram of each station environment

In order to explore the characteristics of multipath effects of sea surface, we calculate the multipath effect MP1 value of satellite SC02 station by the RTKLIB software firstly, and then we draw the corresponding relationship between MP1 value and elevation angle (see Fig. 3).

In Fig. 3, the horizontal axis is the epoch time, the vertical axis from top to bottom were satellite elevation angle and multipath effect MP1 value changes. It can be seen from Fig. 3 that the trend of multipath effect is obviously different at low elevation angle and high elevation angle. The MP1 value oscillation of multipath effect is obviously at the elevation angle of 0° to 15° . In high elevation angle term, the MP1 value sequence of multipath effects tends to be stable, indicating that the multipath effect mainly occurs in the range of 0° to 15° .

To further explore the characteristics of multipath effects on the sea surface, Fig. 4 shows the MP1 time series of multipath effects for the descent phase of G02 from 092 days to 121 days on the day of 2016 in SC02 station.

In Fig. 4, the horizontal axis is the epoch time and the vertical axis is the multipath effect MP1 value (in order to facilitate the display, the MP1 values of all days are sequentially added with constants). It can be seen from Fig. 4 that the trend of MP1 values of multipath effects on G02 is basically the same, indicating that the multipath effect has obvious periodicity. It can also be seen that MP1 value oscillates seriously at low elevation angle term, which shows that multipath effect is serious at low elevation angle.

In order to further analyze the multipath effects caused by different GPS station environments, Fig. 5 shows the comparison of single-epoch MP1 values of multipath effects of G02 satellite at 092 days on the day of year 2016 between the SC02 station and the other three stations.

In Fig. 5, the horizontal axis is the epoch time, and the vertical axis is the comparison of multipath effect MP1 between SC02 station and ALBH, P438, P439 station from top to bottom. As we can see from Fig. 5, the multipath effect of SC02 station in

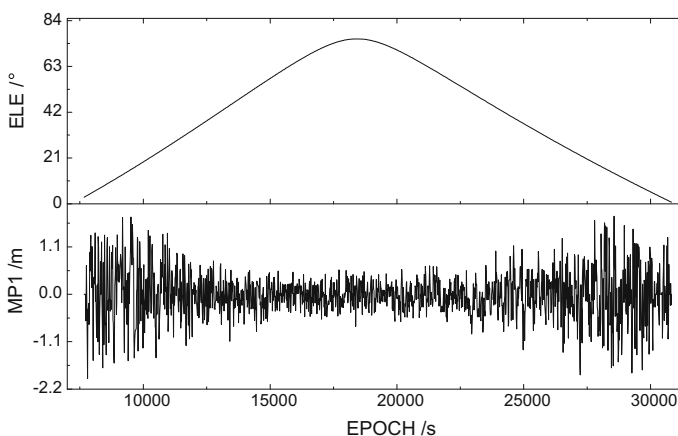


Fig. 3 Multipath effects versus elevation angle

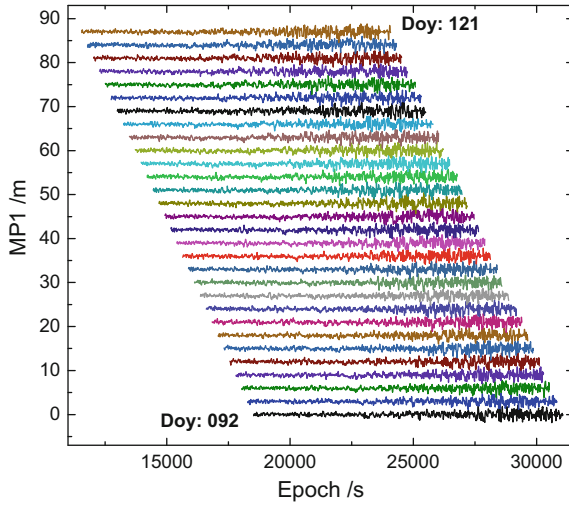


Fig. 4 Time series of G02 satellite multipath effect

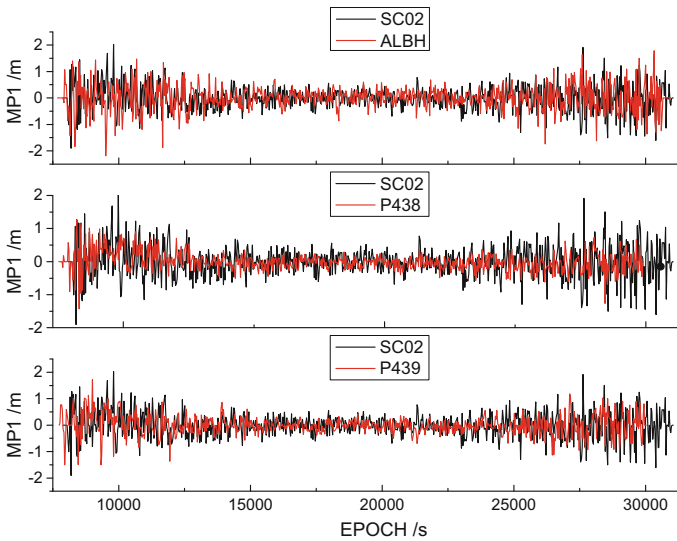


Fig. 5 Comparison of multipath effects at each station

coastal station is comparable to the multipath effect of sheltered ALBH station, while the multipath effect of station SC02 is obviously worse than P438 station and P439 station with good observational environment. The results show that the multipath effect is closely related to the station environment. When the station is located onshore or blocked by obstructions, the multipath effect is obviously more severe than the station which is located in rock or soil with no obvious obstruction around.

Through the above experimental analysis, we can initially obtain some characteristics of the sea surface multipath effect:

1. There is a clear correspondence between sea surface multipath effect and elevation angle, and multipath effect is serious at low elevation angle. As the elevation angle increases, the multipath effect gradually decreases and tends to be stable;
2. The multipath effect of the sea surface has obvious periodicity, which is beneficial for studying and modeling the sea surface multipath effect so as to weaken the multipath effect;
3. The multipath effect is closely related to the environment of the station, sea surface multipath effect is serious due to its high reflection coefficient and complex environment.

4 Analysis of Sea Surface Multipath Effect on GNSS Precision Positioning

In order to study and analyze the effect of sea surface multipath effect on GNSS precision positioning, four stations of SC02, ALBH, P438 and P439 in PBO network were selected for experiment. Four stations were selected for 2016 092 days to 121 days of observation data, and use GAMIT for baseline solution. In this paper, the solution strategy is to change the cutoff angle to 10° , 15° and 20° in turn, and the rest of the parameters are set by default and remain unchanged. By comparing baseline NRMS values, baseline repetition rates and multiplicative RMS residuals to explore the effect of surface multipath effects on GNSS precision positioning [16].

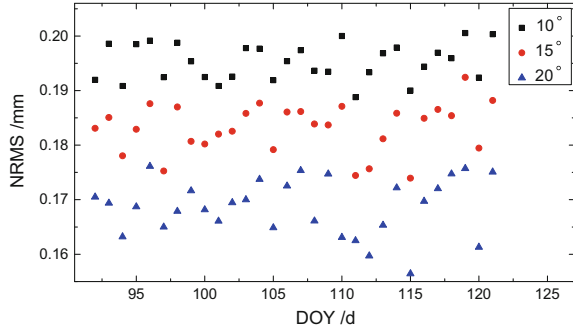
4.1 Analysis of GNSS Baseline Solution Accuracy

The criteria for evaluating the baseline solution are generally NRMS values and baseline repetition rates for baseline resolution. In general, a smaller NRMS value indicates a higher baseline solution accuracy and the value should generally be less than 0.3. Figure 6 shows the comparison of the baseline NRMS values solved at different cutoff elevation angles.

In Fig. 6, the horizontal axis represents the day of year and the vertical axis represents the NRMS value. As Fig. 6 shown, the baseline resolved NRMS values decrease with the elevation angle increases, but all of the three groups are less than 0.3, indicating that all three baseline solutions are eligible.

The baseline repetition rate is relative to several periods' solution in baseline solution, which is the reflection of the accuracy of baseline solution. The smaller the

Fig. 6 Comparisons of baseline solution NRMS values



value, the higher the accuracy of the baseline, the better the baseline quality, and on the contrary, the worse the baseline quality is.

Figure 7 shows the comparison of the baseline repetition rates in different directions at different cutoff altitudes. Since the purpose of this paper is to study the effect of the multipath effects of sea surface on the precise positioning of GNSS, therefore, only the comparison of the baseline repetition rate in each direction with the three baselines connected to the SC02 station is given.

In Fig. 7, the horizontal axis represents the baseline and the vertical axis represents the baseline repetition rate in each direction. As we can see from Fig. 7, the baselines in the three directions N, E, L of the baseline repetition rate is low, indicating that the base line in these three directions to calculate the accuracy is better, and it can also be seen that the change of cutoff elevation angel has no significant effect on the baseline repetition rate in three directions of N, E and L; it can also be seen that the baseline repetition rates of each baseline in the U direction

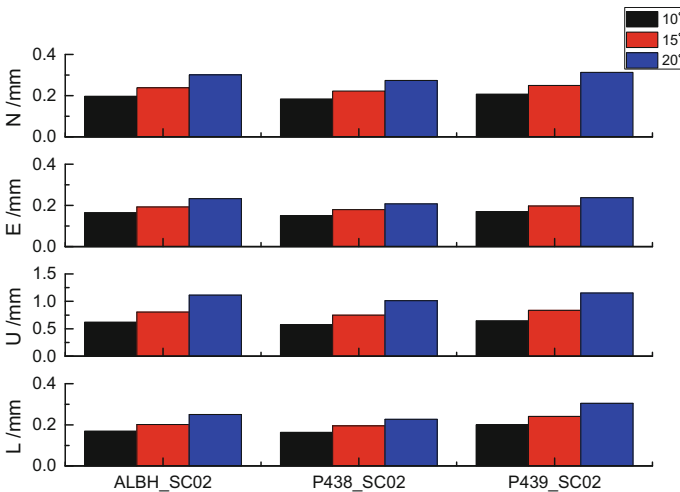


Fig. 7 Comparison of baseline repetition rates in all directions

are significantly higher than those in N, E and L, indicating that the baseline resolution in the U direction worse than N, E, L direction. Moreover, when the cutoff elevation angle is increased to 20° , the baseline repetition rate in the U direction is significantly deteriorated compared to when the cutoff elevation angle is 10° and 15° , which is probably because when the elevation angle is raised to 20° , a large number of original observation data were deleted, resulting in reduced observation accuracy. The result indicated that when the cutoff elevation angle is selected from 10° to 15° , the baseline repetition rate is generally good and the baseline solution accuracy is high.

4.2 Analysis of Multipath Effects Residual

To further explore the effect of multipath effects on GNSS precision positioning, Fig. 8 shows the multipath effect residual sky maps of each station at the cutoff elevation angle of 15° between 8:00 am to 12:00 pm at 92 days of 2016.

In Fig. 8, plots of the multipath effect residual sky followed by stations ALBH, SC02, P439 and P438 are shown. Among them, the red line in Fig. 8 represents the trajectory of the satellite, the green line and the yellow line represent the positive and negative of the multipath residual values respectively and the serrated more clearly represents the larger the multipath effect residual value. As we can see, the overall residual multipath effects at each station are small, indicating that the multipath effect is effectively attenuated at elevation angle of 15° . We can also clearly see in Fig. 8, at the low altitude, the serrated is obvious, and the serrated becomes smaller as the elevation angel increases, which indicates that the multipath effect is more serious at the low elevation angle, and the multipath effect is obviously improved with the elevation angle increases.

In order to more intuitively explore the multipath effect residuals of baseline solutions at different cutoff elevation angles, Fig. 9 shows the comparison of multipath effect residual RMS values at SC02 stations with different cutoff angles.

In Fig. 9, the horizontal axis is the day of year and the vertical axis is the RMS value of the multipath effect residual. As shown in Fig. 9, as the cutoff elevation angle increases, the RMS value of the multipath effect residual tends to increase as a whole. When the cutoff elevation angle is 10° and 15° , the RMS values of the multipath effect residuals are basically close, which are obviously smaller than 20° . The results indicate that when the cutoff elevation angle is selected to be 10° – 15° , the residual value of multipath effects on the sea surface is lower, and the multipath effect is weaker. Correspondingly, the effect of sea surface multipath effect on GNSS precise positioning is less affected at this time.

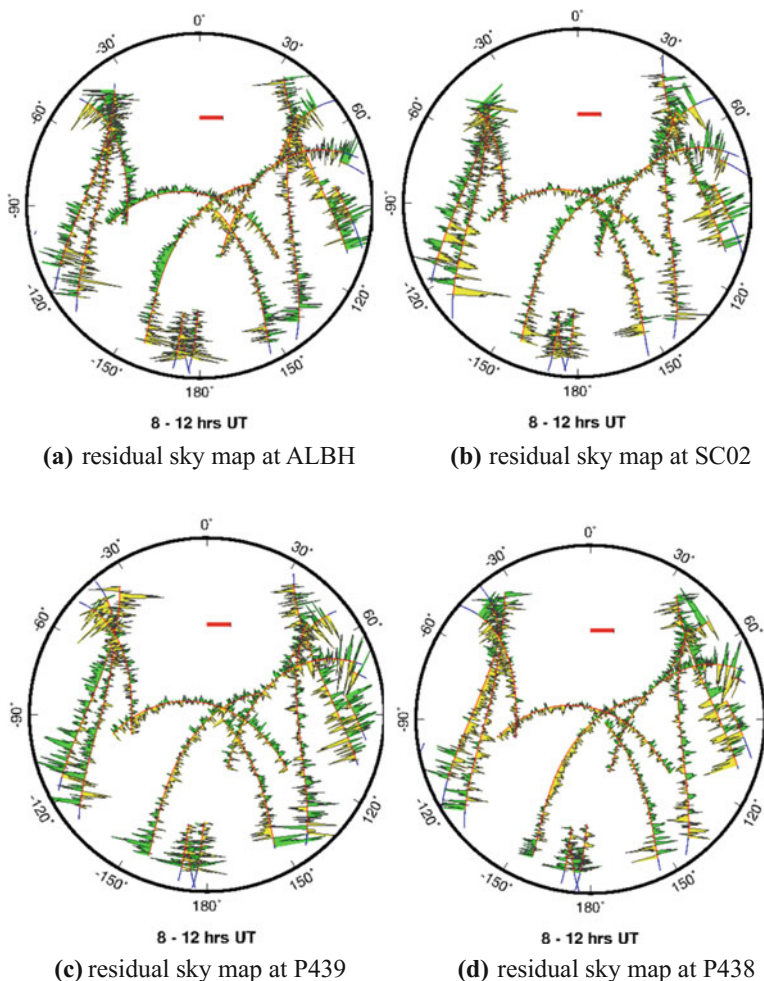


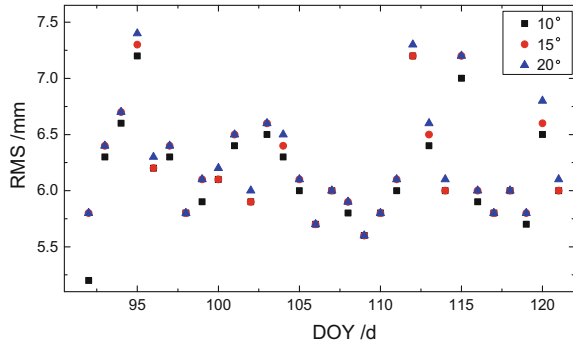
Fig. 8 Multipath effects residual sky map at each station

4.3 Comprehensive Analysis of Experimental Results

Based on the above experimental analysis, we can initially come to the following conclusion:

1. By setting different cutoff elevation angles of 10° , 15° and 20° , the NRMS values calculated from the baseline comparison can be drawn. As the elevation angle increase, the NRMS values for the baseline solution decrease, and all less than 0.3, indicating baseline solution are qualified;
2. By comparing the baseline repeat rates in different directions at different cutoff angles, it can be concluded that the baseline repetition rate in the U direction is

Fig. 9 Comparison of multipath effects residuals



significantly worse than N, E and L directions. Moreover, when the cutoff elevation angle is raised to 20°, the baseline repeat rate in the U direction is seriously worse than 10° and 15°, the result shows that when the elevation cutoff angle is selected from 10° to 15°, the accuracy of baseline solution is good, and the multipath effect of the sea surface has little effect on the precision positioning of GNSS;

- By comparing RMS values of multipath residuals calculated at different cutoff elevation angles, it can be concluded that when the cutoff elevation angle is 15°, the residual value of multipath effects on the sea surface is low, indicating that multipath effect weakening effect is better, and the effect of sea surface multipath effect on GNSS precision positioning is the least.

5 Conclusion

As more and more multi-frequency GNSS navigation constellations and space-based augmentation systems continue to evolve, as well as the implementation of a series of space-based GNSS reflectometry tasks (such as the next FORMOSAT-7/COSMIC-2 and CYGNSS missions). In addition, researchers are developing more advanced GNSS receivers and data processing algorithms such as the next generation of Tri-GNSS receivers with multi-mode GNSS reflection and refraction technology. GNSS has broad prospects for development in the field of ocean remote sensing in the future. In the coming years, large-scale coastal GNSS observational networks are bound to be built in the world. China will also strengthen the establishment of coastal GNSS observing facilities and provide a platform for China’s maritime Silk Road strategy, marine resources development, protection of maritime rights and interests to provide protection. Therefore, this paper studies and analyzes the multipath effects of sea surface and impact on the precision positioning of GNSS. Based on the measured data of SC02 station and three surrounding stations in PBO network, the characteristics of the multipath

effect on the sea surface are analyzed firstly. It is concluded that the sea surface multipath effect is serious due to the high reflection coefficient of the sea surface, and the multipath effect on the sea surface has obvious periodic characteristics at low altitude. Then, four stations are networked and the GAMIT baseline solution is adopted. By setting different cutoff elevation angles of 10° , 15° and 20° , compared the baseline solution accuracy and multipath effect residuals, we concluded that when the cutoff elevation angle is selected from 10° to 15° , the baseline solution results are the best. At this time, the multipath effect of the sea surface has no significant influence on the precision positioning of the coastal GNSS stations. The coastal GNSS station can be used for the research of crustal deformation monitoring, atmospheric and ocean remote sensing and other fields simultaneously. Of course, the experimental results obtained in this paper are only preliminary conclusions. There are still many deficiencies in the experiment. For example, only one coastal station was chosen for the experiment. The precision of the GNSS in coastal stations was also affected by factors such as ocean tides and sea breeze waves. Experimental design and evaluation of multipath effects on GNSS precision positioning still need further study.

Acknowledgements Thanks to the United States NSF, USGS, and NASA who launched the Earthscope and authorized the use of GPS data. This research is supported by the National Natural Science Foundation of China (41104019; 41674001; 41731066); Fundamental Research Funds for Research Funds of Central Universities (310826172202). Thanks to the editors and anonymous referees provide for the valuable comments and suggestions in this paper!

References

1. Yang Y (2016) Complex PNT system and its key techniques. *J Mapp* 45(5):505–510
2. Jin S, Zhang Q, Qian X (2017) New progress and application prospects of global navigation satellite system reflectometry (GNSS + R). *Acta Geodaetica Cartogr Sin* 46(10):1389–1398
3. Martin-Neira M (1993) A passive reflectometry and interferometry system (PARIS) application to ocean altimetry. *ESA J* 17(4):331–355
4. Larson KM, Ray RD, Nievinski FG et al (2013) The accidental tide gauge: a GPS reflection case study from Kachemak Bay, Alaska. *IEEE Geosci Remote Sens Lett* 10(5):1200–1204
5. Löfgren JS, Haas R, Scherneck H-G et al (2016) Three months of local sea level derived from reflected GNSS signals. *Radio Sci* 46(6):1–12
6. Watson C (2017) Remote leveling of tide gauges using GNSS reflectometry: case study at Spring Bay, Australia[M]. Springer, New York
7. Wöppelmann G, Marcos M (2016) Vertical land motion as a key to understanding sea level change and variability. *Rev Geophys* 54(1):64–92
8. Li Z, Huang J (2016) GPS measurement and data processing. Wuhan University Press, Wuhan
9. Zhang Q (2001) Global positioning system (GPS) measurement principle and data processing basis. Xi'an Map Publishing House, Xi'an
10. Cohen CE, Parkinson BW (1991) Mitigating multipath error in GPS based attitude determination. In: Guidance and control, pp 53–68
11. Hannah BM, Walker RA, Kubik K (1998) Towards a complete virtual multipath analysis tool

12. Jia M, Tsakiri M, Stewart M (2000) Mitigating multipath errors using semi-parametric models for high precision static positioning. *Geodesy Beyond* 393–398
13. Zhou D (2009) Multipath path error based on wavelet transform. Jiangxi University of Science and Technology
14. Liu Y (2011) Study on modeling and eliminating methods of GNSS multipath in dynamic marine environment. Shandong University of Science and Technology
15. Fan X, Zhou Z (2010) A review of multipath effects in GPS surveying. *J Eng Geophysics* 07 (3):382–386
16. Huang G, Wang B, Wang Y (2011) Validity repeatability test and quality analysis of high precision GPS control network. *Chin J Mapp* 7:9–11

HECTOR for Analysis of GPS Time Series



Yuefan He, Shuangcheng Zhang, Qianyi Wang, Qi Liu, Wei Qu
and Xiaowei Hou

Abstract High precision GPS coordinate time series has become a rich source of data in many fields of research, such as studying the slow deformation of the earth's surface, establishing and maintaining a regional or global reference frame, and studying the deformation process of earthquake pregnancy. In view of the current GPS time series analysis software in the processing of data is slow, inefficient, less choice of combination model, this paper studies and analyzes a new time series analysis software HECTOR. Firstly, the function and characteristics of HECTOR software are described in detail; Then, the software is used to obtain the periodic items and trend items in the three directions of the GPS time series, and compared with the parameters obtained by the CATS software. Secondly, the data of GPS time series in the study area are analyzed by using different noise models; Finally, BIC numerical analysis based on the maximum likelihood estimation and spectrum analysis are used to compare the results of different combinations of noise models. The results show that the HECTOR software can be used to obtain the parameters quickly and further. At the same time, GPS data for most of the selected study areas, white noise + power law noise for the optimal noise model. For GPS data of a few study areas, white noise + generalized Gaussian Markov noise, white noise + ARMA (5) noise are better noise models. Finally, it is found that the best noise model obtained in different directions of the same site is not the same, which can provide certain reference meaning for the future research in this direction.

Keywords HECTOR · Noise model · GPS time series · Maximum likelihood estimation

Y. He · S. Zhang (✉) · Q. Wang · Q. Liu · W. Qu · X. Hou
College of Geology Engineering and Geomatics, Chang'an University, Xi'an, China
e-mail: shuangcheng369@vip.163.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_16

1 Introduction

The global positioning system (GPS) base station coordinate time series provides valuable basic data for geodesy and geodynamics research [1, 2]. However, the various types of noise contained in the GPS time series restrict the potential application value of GPS technology in the above fields. Therefore, it is a hot topic that scholars study now that the types of noise and the influence of noise on time series are studied.

The current domestic and foreign scholars generally agree that GPS location optimal noise model of time sequence is white noise + flicker noise [3–7]. However, this is not the case in the strict sense of the word, and the noise characteristics of GPS base stations are complex. The literature [8] in 2008 estimated the southern California and southern Nevada run of 236 consecutive GPS station noise model, then the best noise model of 50–60% for the Flicker noise (Flicker noise, FN) or Random Walk noise (Random Walk noise, RW), 25–30% stations show the FN + RW, or non-integer spectral exponential power law noise (Power-law noise, PL), 15% of the stations were Band Pass + Power Law noise (BPPL) and first-order Gaussian Markov + random walk noise. In [9], Tian Yunfeng analyzed the correlation noise characteristics of the time series of GPS base stations in China Crustal Movement Observation Network in 2010 and investigated various noise models. It was found that in addition to the flicker noise, the east component of 50% of the stations Related noise can be described by “Random walk noise + first-order Gaussian Markov noise” model. The literature [10], in 2012, Li Zhao analyzed the IGS baseline time series noise model in China. The results show that the noise models of IGS base stations in China are diverse and have different noise characteristics. The main manifestations are flicker noise + White noise and bandpass power law noise + white noise, and the noise model is affected by the surface mass load and the terrain in the area.

The current time-series analysis software for noise analysis of GPS observations is CATS software developed by Simon Williams [11], Tian Yunfeng’s iGPS software and so on. Most of them use CATS to analyze and compare the noise models of different regions and obtain relevant parameters for further study. However, because the combined noise model supported by the CATS is relatively small and the data processing of the noise combination model is slow. Therefore, this paper introduces another more optimized time series analysis software HECTOR, and separately based on the maximum likelihood estimation BIC (The Bayesian Information Criteria) numerical analysis and spectrum analysis in the frequency domain GPS observations on the noise analysis.

2 HECTOR Software and Its Basic Theory

2.1 HECTOR Software and Functions

HECTOR is a time series analysis software developed by Machiel Bos and Rui Fernandes (Portugal) based on c++ language. It can run on Unix system and has fast running speed. The earliest version was HECTOR1.1, which has been continuously improved and has now been updated to HECTOR1.6. It mainly includes 5 parts [12]

- (1) *estimatetrend*—Main program to estimate a linear trend.
- (2) *estimatespectrum*—Program to estimate the power spectral density from the data or residuals using the Welch periodogram method.
- (3) *modelspectrum*—Given a noise model and values of the noise parameters, this program computed the associated power spectral density for given frequency range.
- (4) *removeoutliers*—Program to find offsets and to remove outliers from the data.
- (5) *simulatenoise*—Synthetic colored noise files.

2.2 HECTOR Noise Analysis Method

At present, the main methods of noise analysis are maximum likelihood estimation and spectrum analysis [5, 13]. HECTOR software has two kinds of analysis method of noise, one is based on the maximum likelihood estimator of AIC (The Akaike Information Criteria) and BIC numerical analysis method, both with maximum likelihood estimator as the starting point, but in order to avoid excessive fitting problem, increase the measures to add parameters. The other is a conventional spectrum analysis method.

2.2.1 AIC and BIC Numerical Analysis

HECTOR software allows the estimation of linear trend items and combinations of higher order polynomials, seasonal items and other periodic signals and various noise models. HECTOR selects the best model to use AIC and BIC information standards [14, 15]. The formulas for analyzing different noise models using AIC and BIC standard are as follows:

$$\ln(L) = -\frac{1}{2} [N \ln(2\pi) + \ln \det(C) + r^T C^{-1} r] \quad (1)$$

where N refers to the actual number of observations, the covariance matrix C can be decomposed into

$$C = \sigma^2 \bar{C} \quad (2)$$

\bar{C} refers to the sum of all kinds of noise model, σ refers to the standard deviation and it can be estimated by residual error.

$$\sigma = \sqrt{\frac{r^T \bar{C}^{-1} r}{N}} \quad (3)$$

In the light of $\det cA = c^N \det A$, So you can get

$$\ln(L) = -\frac{1}{2} [N \ln(2\pi) + \ln \det(\bar{C}) + 2N \ln(\sigma) + N] \quad (4)$$

Parameter k is the sum of design matrix H and noise model parameters and driving white noise variance. For example, using white noise + power law noise model to estimate the linear trend items including five arguments, the normal offset, linear trend, the difference between power law noise and white noise, power spectrum index and drives the white noise variance ($k = 2 + 2 + 1 = 5$)

So:

$$AIC = 2k + 2 \ln(L) \quad (5)$$

$$BIC = k \ln(N) + 2 \ln(L) \quad (6)$$

The preferred model is the one with the minimum AIC or BIC value. Note that these are relative measures between various choices, not absolute criteria.

2.2.2 Spectral Analysis Method

Spectrum analysis is a method of analyzing signals in the frequency domain. The power spectrum of the noise time series in the spectral domain can be expressed in the form of Power Law [16].

$$P(f) = P_0 f^{-a} \quad (7)$$

In the formula, a is the spectrum index; P_0 is constant. The larger a , the higher the time correlation of the noise sequence.

The formula (7) takes logarithmic curve fitting, and the P_0 and a value can be obtained by using least square method. a is usually an arbitrary real number between -1 to 3 , where integer a represents some special noise type: When $a = 0$, which is

standard white noise; When $a = 1$, it is standard flashing noise; At $a = 2$, the standard random walk noise. In addition, in addition to the standard white noise, the rest is referred to as colored noise.

For the combined model, assuming that the noise in the GPS coordinate sequence is composed of white noise and colored noise, the power spectrum is [17].

$$P(f) = P_0(f^{-a} + f_0^{-a}) \tag{8}$$

where f_0 is the crossover frequency; P_0 , f_0 and a are parameters to be called.

3 HECTOR for the Analysis of GPS Time Series

3.1 Data Sources

This paper selects PBO website and based on the GPS time series data of 12 IGS sites under the ITRF08 framework, the longest one is SC02 station (16.5 years), and the shortest is AB43 station (7.1 years). Time series analysis and Fig. 1 is the site map selected.

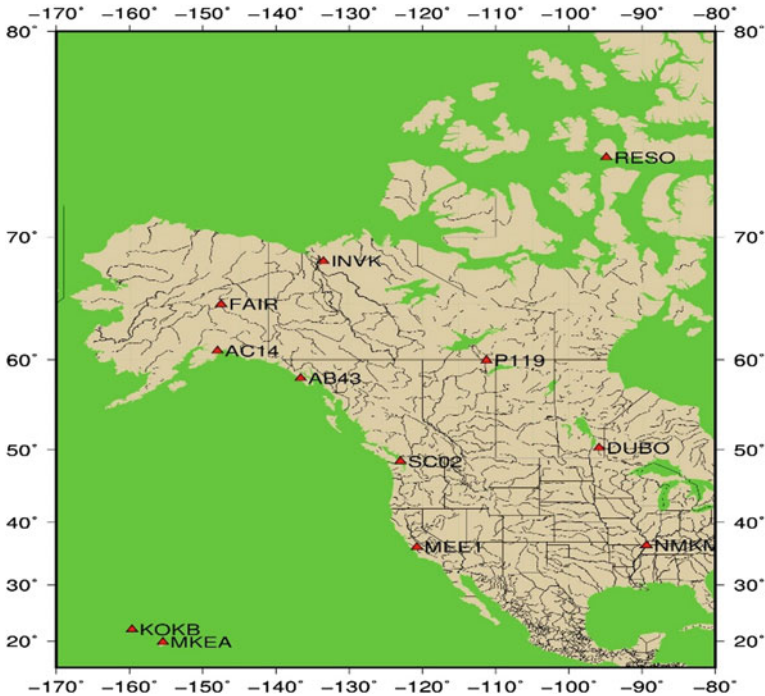


Fig. 1 GPS Site location

3.2 Extraction of Trend Items and Cycle Items

The trend changes mainly reflect the inheritance tectonic movement under the control of tectonic stress field in the same area [18] and it is of great significance and application value to analyze the trend items and period items of GPS coordinate time series. This section mainly uses the HECTOR software to extract trends in the GPS time series for analysis. Take the SC02 station with the longest time span (16.5 years) as an example.

As shown in Fig. 2a–c, the horizontal axis to MJD time, longitudinal axis respectively, E, N, U direction trend fitting figure. Blue as the original time series, red

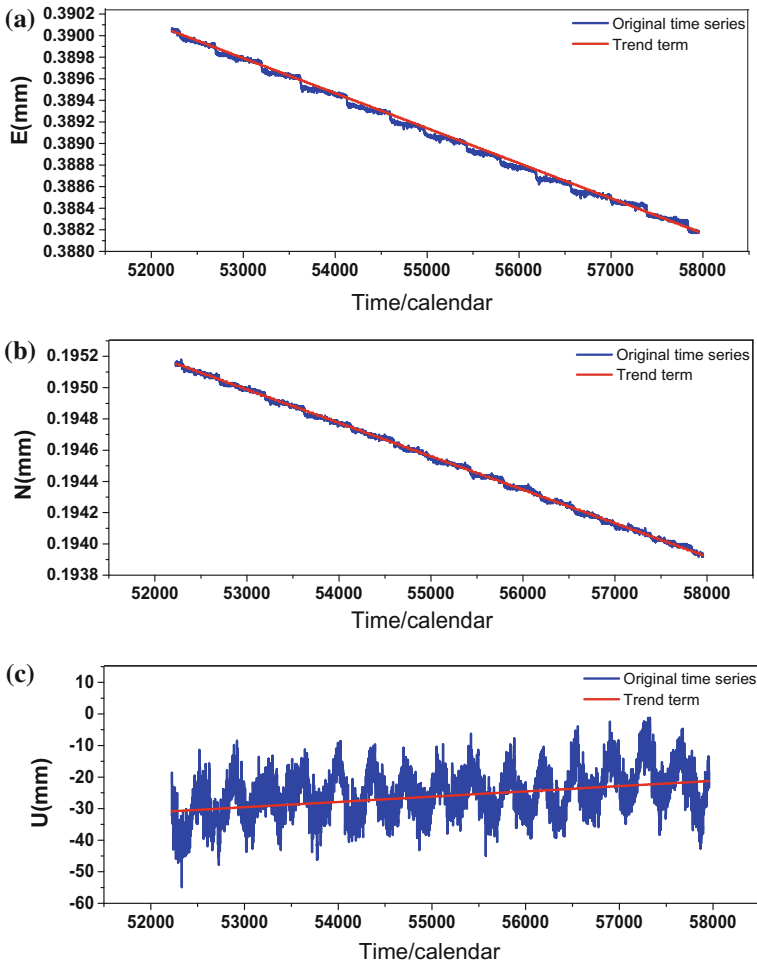


Fig. 2 a The trend in the direction of E, b the trend in the direction of N, c the trend in the direction of U

Table 1 Comparison of CATS with HECTOR software

SC02		Slope	Cos yearly	Sin yearly	Cosh yearly	Sinh yearly
E	C	1.1585	-0.091	-0.1524	0.4432	0.1084
	H	1.195	-0.135	-0.217	0.414	0.114
	RMS	0.0365	-0.044	-0.0646	-0.0292	0.0056
N	C	0.3178	-1.1052	0.0369	0.1752	-0.0115
	H	0.084	-1.101	0.036	0.170	-0.012
	RMS	-0.2338	0.0042	-0.0009	-0.0052	-0.0005
U	C	0.9166	-2.6930	-7.9512	-1.1183	0.5672
	H	0.051	-2.742	-7.906	-1.115	0.585
	RMS	-0.8656	-0.049	0.0452	0.0033	0.0178

red is obtained with HECTOR software trend item, you can see it can be a good development trend. Same as shown in Table 1, you can use the HECTOR software for time series of the slope, year cycle, half a cycle, and compared with CATS software. It was found that the RMS of both the trend and the period of the two software were less than 1 mm and it can further prove the correctness of using HECTOR software to obtain time series parameters.

3.3 Analysis of Optimal Noise Model

In this section, six noise models of HECTOR software were used to analyze the time series of the selected 12 IGS stations. The noise models used are white noise (W), white noise + power law noise (W + P), Generalized Gaussian Markov Noise (G), White Noise + Generalized Gaussian Markov Noise (W + G), white noise + ARMA (1) noise (W + A (1)), white noise + ARMA (5) noise (W + A (5)). Because the selected sample is small, the optimal noise model is better by using BIC numerical method based on maximum likelihood estimation and the optimal model is identified from the frequency domain using the spectral analysis method.

3.3.1 Analysis of Numerical Results Based on BIC

The BIC values of the three components of each of the 12 stations are calculated. And then find the smallest BIC value corresponding to each station in the same component, then the noise model corresponding to this BIC value is the best noise model for this station. Finally, calculate the percentage of the total number of the best model based on the number of directions in each direction.

The largest proportion of the optimal noise model in E direction is white noise plus noise model ARMA (5), accounting for 33.13% of all sites, then followed by

white noise + power law noise, White noise + generalized Gaussian Markov noise and so on. The largest proportion of the optimal noise model in N direction is white noise model + power law noise, accounted for 41.41% of all sites, then followed by white noise + ARMA (5) noise, White noise + generalized Gaussian Markov noise and so on. The largest proportion of the optimal noise model in U direction is white noise + power-law noise, accounting for 42.04% of all sites, followed by white noise + ARMA(5) noise model. It can be seen that the proportion of optimal noise model is different in different directions. But overall for most points, white noise + power law model for better noise model, for a few points, white noise + ARMA (5) noise, White noise + generalized Gaussian Markov noise a model is better.

3.3.2 Analysis Result Based on Spectrum Analysis Method

Because of the limited space, the spectral analysis is carried out by the white noise + power law noise and the white noise + generalized Gaussian Markov noise model. The spectrum index a was obtained from the 12 IGS sites using HECTOR software to further verify the optimal properties of white noise + power-law noise.

From Tables 2 and 3, it can be seen that the larger a value, the higher the time correlation of the noise sequence. It can be seen that the results obtained by using the spectrum index are consistent with that of the BIC numerical analysis based on the maximum likelihood estimate. It can be clearly seen that for most stations, white noise + power-law noise model is advantageous to white noise + Generalized Gaussian Markov Noise model.

Many scholars have done many different combinations of noise models before comparisons, but most are using CATS software for data processing and analysis of contrast, compared to CATS software, HECTOR software to process data faster and

Table 2 a value under the W + P model

站点	E	N	U
DUBO	0.4793	0.5899	0.5791
FAIR	0.7669	0.8500	0.5438
INVK	0.6702	0.7372	0.7818
KOKB	0.5027	0.4895	0.5733
MKEA	0.5976	0.4350	0.5341
RESO	0.6563	0.7578	0.4258
SC02	0.8500	0.8162	0.4534
AB43	0.5892	0.4225	0.4209
AC14	0.7265	0.7282	0.7132
MEE1	0.8500	0.8307	0.3957
NMKM	0.8500	0.2313	0.4424
P119	0.8500	0.8500	0.4097

a Spectrum index

Table 3 α value under the W + G model

站点	E	N	U
DUBO	0.4675	0.4704	0.5789
FAIR	0.4808	0.4808	0.4493
INVK	0.7647	0.8129	0.4994
KOKB	0.4646	0.4915	0.5774
MKEA	0.6042	0.6042	0.5429
RESO	0.7257	0.8241	0.4254
SC02	0.5029	0.9305	0.4687
AB43	0.4221	0.3271	0.4244
AC14	0.7625	0.4853	0.8633
MEE1	0.7231	0.9314	0.3997
NMKM	0.4988	0.2371	0.4672
P119	0.7520	0.4990	0.4221

more efficient, you can choose the noise combination model is more and the result is consistent with CATS. But HECTOR software also has some problems, because the software has not been widely used, so it is not enough mature and perfect, more with some missing data in the station, using a combination of some cannot be output to the parameters of the model results, automatically interrupt handling. More scholars are needed to further study.

4 Conclusion

This paper makes detailed introduction to HECTOR software, HECTOR was used to analyze the time series of GPS points. Finally, the comparison and analysis the dimension noise model of GPS point of different plates. The following conclusions can be obtained:

- (1) HECTOR can accurately obtain the station kinematics parameters of the earth station, and get the correct site kinematics model under the premise that can lay the foundation for the next step of the study.
- (2) Using the HECTOR software to analyze the time series of the selected 12 IGS sites by using 6 noise models. For the selected area, white noise + power law noise for better noise model, In addition, White Noise + Generalized Gaussian Markov Noise Model, white noise + ARMA (5) and other groups and noise models are good noise models.
- (3) Experimental analysis it is concluded that the noise model for the most points is not a single noise. Instead, the noise model is composed of different noises, and the optimal noise model is different in different directions.

Acknowledgements Thanks PBO website: (<http://www.unavco.org/data/gps-gnss/data-access-methods/dai2/app/dai2.html>) provides the time series data. This research is supported by the National Natural Science Foundation of China (41104019; 41674001; 41731066); Fundamental Research Funds for Research Funds of Central Universities (310826172202). Thanks to the editors and anonymous referees provide for the valuable comments and suggestions in this paper!

References

1. Chen J, Zhang P, Wu J (2007) On Chinese national continuous operating reference station system of GNSS. *Acta Geodaetica Cartogr Sin* 36(4):360–366
2. Chen J, Xu H, Hu J et al (2004) The latest advances in geodesy reported in the 23rd General Assembly of IUGG. *Acta Geodaetica Cartogr Sin* 33(1):12–21
3. Williams SDP (2003) The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. *J Geodesy* 76(9–10):483–494
4. Mao A, Harrison CGA, Dixon H (1999) Noise in GPS coordinate time series. *J Geophys Res* 104(B2):2797–2816
5. Huang L (2006) Noise properties in time series of coordinate component at GPS fiducial stations. *J Geodesy Geomatics* 26(2):31–38
6. Zheng G, Wang H, Wright TJ (2017) Crustal deformation in the India-Eurasia collision zone from 25 year of GPS measurement. *J Geophys Res Solid Earth* 122
7. Zhu W, Fu Y, Li Y (2003) Global height vibration and its seasonal variation induced by GPS height. *Sci China Ser D* 33(5):470–481
8. Langbein J (2008) Noise in GPS displacement measurements from Southern California and Southern Nevada. *J Geophys Res* 113(B5):1–12
9. Tian Y, Sheng Z (2010) Analysis of correlation noise in continuous GPS observation. *J Earthquake* 36(6):696–704
10. Li Z, Jiang W (2014) Establishment and analysis of the noise model of the coordinate time series of IGS base station in China. *Acta Geodaetica Cartogr Sin* 41(4):496–503
11. Ye J (2015) The GPS time series analysis based on the CATs software. *Geospat Inf* 13(6):121–123
12. Bos MS, Fernandes RMS, Williams SDP, Bastos L (2013) Fast error analysis of continuous GNSS observations with missing data. *J Geodesy* 87(4):351–360
13. Langbein J, Johnson H (1997) Correlated errors in geodetic timeseries: implications for time-dependent deformation. *Geophys Res* 102(1):591–603
14. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19(6):716–723
15. Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6(2):461–464
16. Agnew DC (1992) The time-domain behavior of power law noises. *Geophys Res Lett* 19(4):333–336
17. Tian Y, Shen Z (2009) The research progress of the method of nonstructural noise elimination in GPS coordinate time series. *J Earthquake* 31(1):68–81
18. Jiang W, Xia C, Li Z et al (2014) Analysis of environmental loading effects on regional GPS coordinate time series. *Acta Geodaetica Cartogr Sin* 43(12):1217–1223

Inversion of Water Vapor Variation During Typhoon by Quad-Constellation GNSS Tomography



Yonglin Zhang, Changsheng Cai, Xiaotao Bai, Hongke Hou
and Xianqiang Cui

Abstract The ground-based GNSS tomography technology can obtain three-dimensional distribution of water vapor. However, many voxels may not be passed through by satellite rays because of the limitation of satellite orbits and the topological distribution of ground GNSS stations, which results in the rank defect of tomography equations. In this paper, the geometric relationship between the elevation angle of the satellites and the satellite ray height intersecting with tomography boundary is deduced. The distribution of the GNSS satellite signals that can be used for the tomography at different stations is shown for the selected tomography area. In order to increase more usable GNSS rays and get stable and reliable results, the quad-constellation GNSS including GPS, GLONASS, Galileo and BDS (BeiDou navigation satellite system) systems are used simultaneously. Further, the Gaussian function is used as horizontal restriction while the mean value of three-day radiosonde profiles prior to the experiment is used as vertical restriction. To validate the four-constellation integrated tomography method, the raw GNSS observations were collected from 26 August, 2017 to 28 August, 2017 in Hong Kong to carry out the tomography experiment. During this period, there was a typhoon named Pakhar to pass the tomographic area, causing water vapor to vary dramatically. This process was described and captured by the presented tomographic technique. The result indicates that the current quad-constellation GNSS improves the tomographic accuracy by 5% compared with the GPS-only approach.

Keywords Tomography · GNSS · Water vapor · Typhoon

Y. Zhang · C. Cai (✉) · X. Bai · H. Hou · X. Cui
School of Geosciences and Info-Physics, Central South University,
Changsha, People's Republic of China
e-mail: cscai@hotmail.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_17

1 Introduction

As an important component of the atmosphere, water vapor is deeply involved in the global water cycle. In summer, all kinds of rainstorms, droughts and typhoons are closely related to the variation of water vapor [1, 2]. The study of the variation of water vapor is vital for predicting the bad weather. Unlike the traditional methods such as the radiosonde or water vapor radiometer to detect the water vapor of the atmosphere, using the global navigation satellite system (GNSS) to monitor the water vapor has the advantages of low cost and high spatial resolution [3]. Since the GNSS method was first proposed by American scholar Bevis in the 1990s, it has become an important method to detect water vapor content in the atmosphere [4]. Nowadays, many countries use GNSS continuously operating reference stations (CORS) to monitor the nationwide water vapor variation and the results are well consistent with the numerical weather model (NWM) [3, 5, 6]. Further, the four-dimensional distribution of water vapour can be obtained based on a tomography technique, which can be used to acquire the temporal and spatial variation of water vapor under different weathers. Although the GNSS tomography method cannot predict precipitations directly, it has the ability to improve accuracy of the NWM. Because of the limitation of satellite orbits and the topological distribution of ground GNSS stations in the tomography area, many voxels may not be passed by GNSS rays, which results in unreliable tomography solutions. Adding the extra observations and constraints are widely used to solve the problem. At present, most tomography experiments are based on GPS observations [3–6]. The introduction of GLONASS can increase the amount of available observation data and thus improve the accuracy of slant water vapor (SWV) [7]. With the rapid development of Galileo and BeiDou navigation satellite system (BDS), quad-constellation integrated GNSS tomography can effectively decrease the number of voxels without passing GNSS signals and improve the resolution of tomography [8, 9].

The main structure of this paper is as follows. Section 2 introduces the basic principle of ground-based GNSS tomography, deriving the geometric relationship between the elevation angle of the satellite and the satellite ray height intersecting with tomography boundary. Section 3 shows the available distribution of GNSS satellite signals at different stations in the tomography area. The GNSS tomography results indicate that the water vapor variation caused by the typhoon in Hong Kong can be successfully captured.

2 The Method of Ground-Based GNSS Tomography

2.1 The GNSS Tomography Model

GNSS signals are delayed when they pass through the Earth atmosphere due to the refraction of the ionosphere and troposphere. Since the troposphere is neutral, the tropospheric delay is independent of the signal frequency and cannot be eliminated

by multi-frequency combinations like ionosphere delay. In order to mitigate its effect, the zenith troposphere delay (ZTD) is commonly estimated as an unknown parameter in precise point positioning (PPP), single-difference or double-difference positioning techniques [3, 10]. The ZTD can be divided into zenith dry delay (ZHD) and zenith wet delay (ZWD) [11] with the relationship as shown by Eq. (1).

$$ZTD = ZHD + ZWD \tag{1}$$

The slant path tropospheric delay (STD) corresponding to each satellite can be further obtained as [8, 12]:

$$STD = m_h(\varepsilon) \cdot ZHD + m_w(\varepsilon) \cdot [ZWD + \cot(\varepsilon) \cdot (G_N \cos \phi + G_E \sin \phi)] \tag{2}$$

where ε and ϕ represent the elevation and azimuth of satellites, respectively, m_h and m_w represent the dry and wet mapping functions, G_N and G_E are north and east components of the gradient, respectively.

In general, the ZHD can be computed precisely using the tropospheric model based on ground meteorological observations. Thus, the ZWD can be obtained by applying Eq. (1). The ZWD is closely related to the content of atmospheric water vapor and can be converted to the precipitable water vapor (PWV) by multiplying a conversion factor as shown in Eq. (3). The slant wet delay (SWD) can also be converted into the slant precipitable water vapor (SWV) similar to the conversion from the ZWD to the PWV.

$$PWV = \Pi \cdot ZWD \tag{3}$$

It is generally considered that the SWV is equal to the integral of the water vapor density (WVD) along the path of the GNSS signal propagation. To simplify the calculation, the tomography area is divided into $m \cdot n \cdot l$ voxels from the longitude, latitude and elevation directions, respectively. The WVD is considered as a constant within each voxel. Thus, the SWD is equal to the sum of the product of the signal intercepted distance by the corresponding water vapor density in each voxel, as shown in Eq. (4).

$$SWV = \int_S x ds = \sum a_{i,j,k} x_{i,j,k} \tag{4}$$

where S represents the propagation path of GNSS rays, x and a represent the WVD and intercepted distance in each voxel.

Due to the limitation of GNSS satellite orbits and the distribution of ground GNSS stations in a tomography area, there are many voxels that is probably not passed by any GNSS ray during the entire observation session. In order to obtain

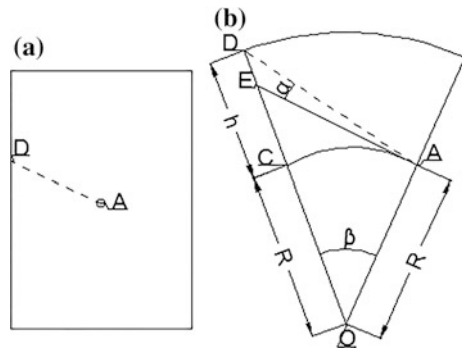
stable and reliable three-dimensional distribution of the WVD, it is necessary to add some constraints in horizontal and vertical directions or use the algebraic reconstruction techniques to solve the rank deficiency issue of the tomography equation [8].

2.2 The Relationship Between Satellite Elevation and the Satellite Ray Height Intersecting with Tomography Boundary

In the tomography model, only the GNSS signals that enter from the top of the tomography grid can be used since the PWV is an integral of water vapor density in different layers along the slant path of the satellite signal. From this point of view, the GNSS stations located in the center of the tomography area may utilize the GNSS signals in a largest degree whereas those stations that are located near the tomography boundary area have to discard many signals that cannot pass the tomography grid from its top. In this section, the geometric relationship between the elevation angle of the satellites and the height of points at which the satellite rays intersect with tomography boundary is established.

As shown in Fig. 1a, the rectangle represents a tomography area. The cycle point ‘A’ represents a GNSS station. The line between ‘A’ and ‘D’ represents a beam of GNSS signals received by the station ‘A’. ‘D’ is the intersection point of the satellite signal and the tomography boundary. Figure 1b is a sectional view of Fig. 1a. In Fig. 1b, ‘AC’ is a circular arc, ‘R’ is the radius of Earth, ‘O’ is the Earth center, ‘C’ is the intersection point of AC and OD, h represents the length of CD, AE represents the tangent line at the point A, α is the satellite elevation angle at station A and β represents the central angle of arc AC. Considering that the Earth’s oblateness is only about 0.003, the earth in the tomography area is approximated as a sphere. Thus, ‘h’ is the ellipsoidal height of point ‘D’.

Fig. 1 The useful satellite signals at a station for a certain tomography area



Based on the triangular relations, it is easy to know geometric relations below:

$$\begin{aligned} \angle OAE &= 90^\circ \\ \angle ADO &= 180^\circ - (\beta + \angle OAE + \alpha) = 90^\circ - \alpha - \beta \end{aligned}$$

Using the Sine theorem in $\triangle ADO$:

$$\frac{R+h}{\sin(90^\circ + \alpha)} = \frac{R}{\sin(90^\circ + \alpha - \beta)} \tag{5}$$

Simplifying Eq. (5), the following equations can be obtained:

$$h = \frac{R \cos(\alpha) - R \cos(\alpha + \beta)}{\cos(\alpha + \beta)} \tag{6}$$

Or

$$\alpha = \arctan\left(\frac{R - (R+h) \cos(\beta)}{(R+h) \sin(\beta)}\right) \tag{7}$$

Because the earth is actually an ellipsoid, the radius ‘R’ varies with latitude and their geometric relationships are expressed below.

$$R = \frac{MN}{N \cos^2(\lambda) + M \sin^2(\lambda)} \tag{8}$$

where

$$\begin{aligned} M &= \frac{a(1 - e^2)}{W^3}, \quad N = \frac{a}{W} \\ e^2 &= \frac{a^2 - b^2}{a^2}, \quad W = \sqrt{1 - e^2 \sin^2(B)} \end{aligned}$$

where M and N represent the radii of meridian and prime vertical circle at point A, respectively, λ represents the azimuth of the arc AC, a and b are the semi-major axis and semi-minor axis of the ellipsoid, respectively, B represents the latitude of A. Because the boundary of tomography is set in advance, the Eq. (7) can be used to calculate the lowest elevation angles of usable satellite signals for tomography resolution.

3 Realization of Quad-Constellation GNSS Tomography at Hong Kong

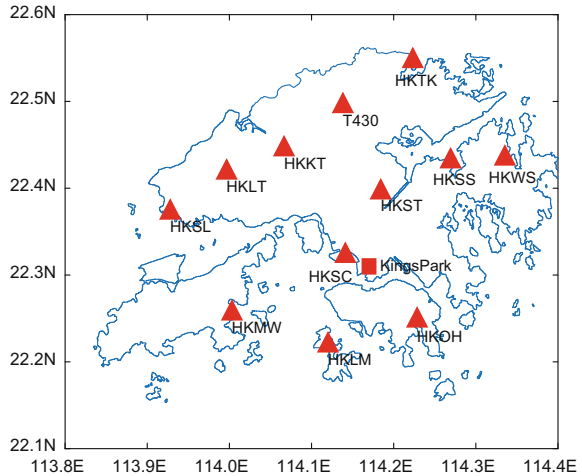
At about UTC 12:00 on 24 August, 2017, a typhoon named Pakhar was generated in the Western Pacific and began to move toward west northwest of China. Then it landed on the southeast coast of Taishan City, Guangdong Province, China at UTC 1:00 on 27 August, 2017. Hong Kong was also influenced by it as shown in Fig. 2 [13].

The quad-constellation GNSS observations at 12 stations from Hong Kong Satellite Positioning Reference Station Network (SatRef, seen Fig. 3) were collected with period from 26 August, 2017 to 28 August, 2017. The tomography area covers latitudes from 22.16°N to 22.56°N, longitudes from 113.87°E to 114.35°E and the vertical boundary is 10 km in height to Earth surface [14]. The voxel division is 0.05°, 0.06° and 500 m in latitude, longitude and height, respectively. The total number of voxels is $8 \times 8 \times 20$. Finally, the water vapor variation during typhoon was retrieved by the quad-constellation GNSS tomography whose results are compared with the radiosonde data collected from the radiosonde station of Kings Park (seen Fig. 3).



Fig. 2 The track of typhoon Pakhar

Fig. 3 The distribution of GNSS stations and radiosonde station in Hong Kong



3.1 The Distribution of Available GNSS Signals

As described above in Sect. 2.2, due to the limitations of the tomography model, GNSS signals received at stations cannot be utilized completely, and the usable signals at different locations vary largely. In this section, the distribution of usable signals using quad-constellation GNSS observations at stations HKST, HKSL, HKOH and HKWS (seen Fig. 3) on 26 August is analysed.

Based on the ellipsoid parameter of WGS-84 and Eq. (7), for the vertical tomography boundary of 10 km, the distribution (seen Fig. 4) of available satellite signals is obtained. Only green satellite signals which are surrounded by blue curve are used for tomography modeling. By contrast, those grey satellite signals out of blue cycle will be discarded since they enter from the side face. As seen from Fig. 4, the number of usable signals at the station in the center of tomography area is significantly greater than those stations located near the tomography boundary such as the stations of HKST and HKWS. This also explains to a certain extent that the water vapor inversion accuracy at the boundary is lower than that at the tomography center.

3.2 Inversion of Water Vapor Content Based on Quad-Constellation Tomography

With the rapid development of BDS and Galileo systems, quad-constellation GNSS tomography has been available, which can effectively increase the observations and decrease the number of voxels lacking passed satellite rays [9]. In this study, the open-source RTKLIB software is used to process quad-constellation GNSS

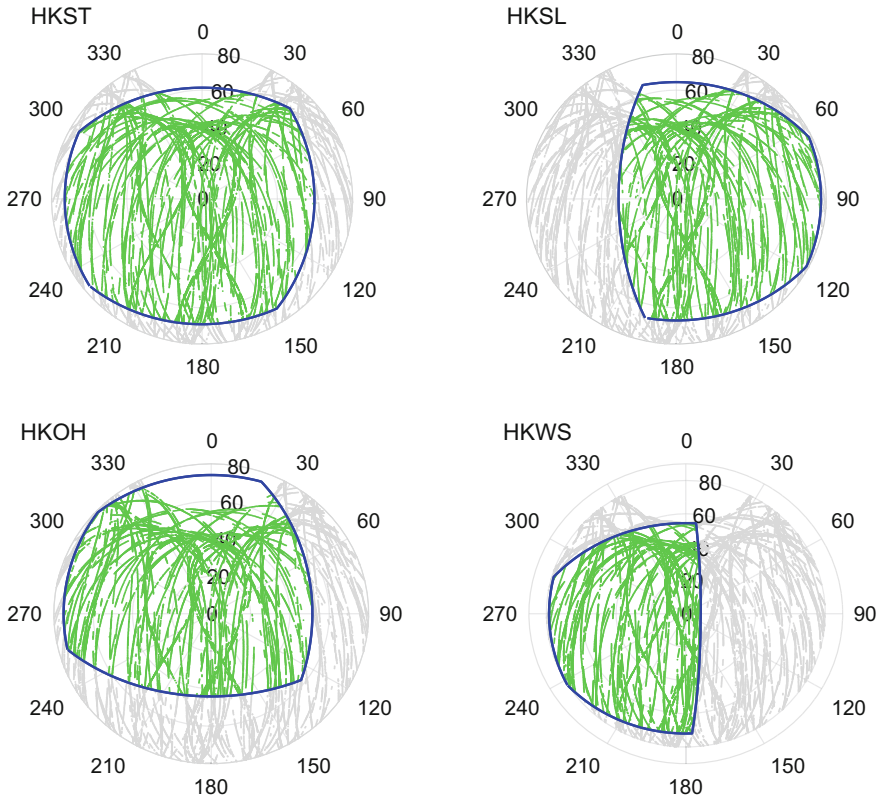


Fig. 4 The distribution of usable multi-GNSS satellite signals at different stations for the tomography vertical boundary of 10 km

observations in a static mode. The elevation mask angle is set to 10° . The Global Mapping Function (GMF) is used to convert from ZWD to SWD. The weighted mean temperature is obtained from GPT2w [15] and the length of tomography time window is set to 30 min. Due to the relatively poor orbit accuracy of BDS GEO satellites, this study just uses the MEO and IGSO satellites of BDS in conjunction with other satellite systems to calculate the ZTD and gradients. Figures 5 and 6 show the change of the number of GNSS satellites and the rate of voxels without passing GNSS signals during the entire tomography session. With the joint use of multiple GNSS systems, the number of visible satellites increases greatly and consequently the voxels lacking passed GNSS signals decrease compared with GPS single constellation. Before the tomography, the PWV acquired from GPS and quad-constellation GNSS are compared and the results show a good agreement with each other (seen Fig. 7), which indicate that the GPS combined with GLONASS, Galileo and BDS tomography is feasible. Also, it can be seen from Fig. 7, the blue histogram below the PWV curve represents the hourly precipitation at Sha Tin

District obtained from the Hong Kong Observatory. The absence of a few bars is because of no precipitation during this period. In addition, we can find a sharp rise of PWV before beginning to raining. When the PWV is high, it agrees with a time when precipitation concentrates. When the PWV decreases gradually, the precipitation also decreases correspondingly. This result shows a good correlation between PWV and precipitation and indicates that the PWV can provide an evidence for monitoring the typhoon movement and forecasting precipitation.

In order to solve the rank defect issue in the tomography model, the mean value of three-day radiosonde profiles prior to tomography is used as vertical restraint while the Gaussian function is used as horizontal restraint. The relaxation parameter λ is set to 0.2. The multiplicative algebraic reconstruction techniques (MART1) is

Fig. 5 The number of satellites at HKST

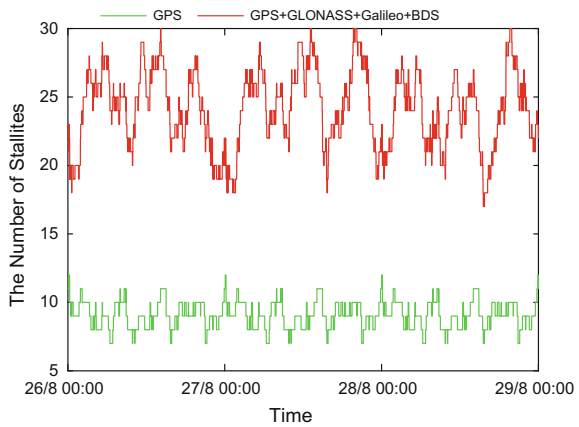
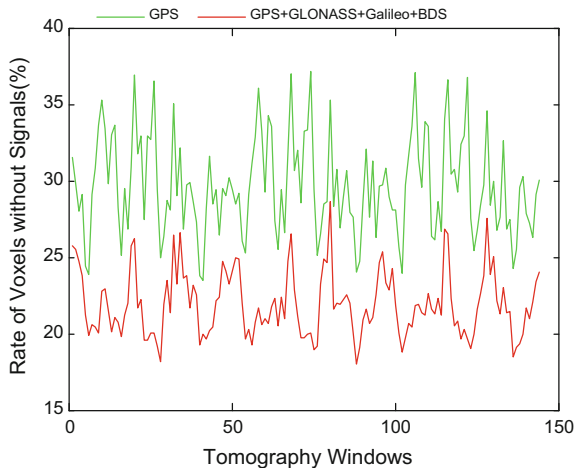


Fig. 6 The rate of voxels without passing GNSS signals for different tomography windows



adopted [8]. Figure 8 shows the time series of tomography profiles at HKST. It can be found that the variation of PWV is mainly concentrated in the area below 4 km in height, which is approximately consistent with PWV in Fig. 7. Figure 9 shows the tomographic sections along the 114.200°N and 22.435°E during typhoon landing. It can be found that the WVD kept steady at 12:00 on 26 August before typhoon approaching. Then the content of WVD increased dramatically when typhoon landing. Finally, the content gradually decreased and returned to the level prior to landing. In order to verify the tomography results, the radiosonde data from 27 August to 28 August, 2017 is used to compare with the GNSS tomography results. Figure 10a, c show the water vapor profiles from radiosonde data, GPS, GPS + GLONASS, GPS + GLONASS + Galileo, GPS + GLONASS + Galileo + BDS tomography, respectively on 27 August. Figure 10b, d show the offset between different GNSS combined results and the radiosonde profiles. It is obvious that the GNSS tomography results have good agreement with the sounding profiles. Compared with the GPS tomography, the multi-GNSS improves the accuracy at

Fig. 7 The time series of PWV during typhoon at HKST

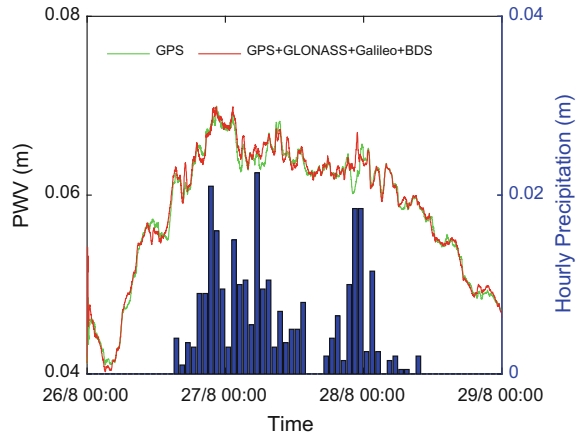


Fig. 8 The time series for tomography profiles at HKST

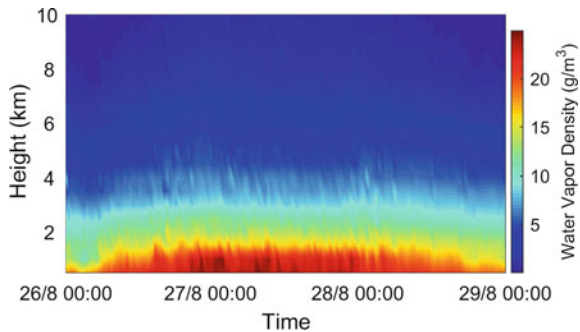
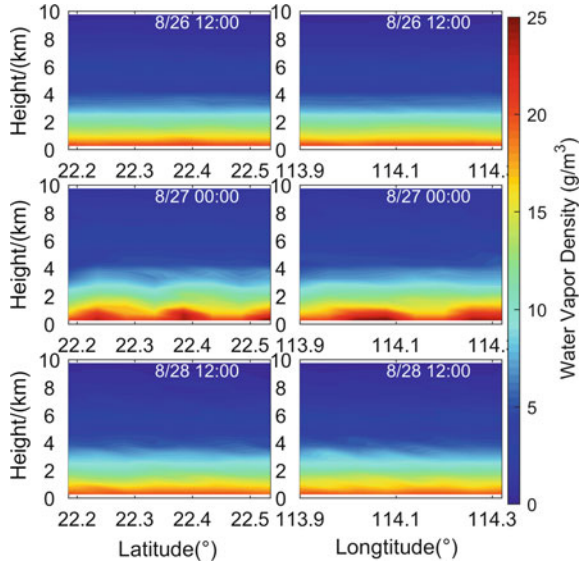


Fig. 9 The tomographic sections along the 114.200°N (left figure) and 22.435°E (right figure) during typhoon landing



middle and high layers, but no significant improvement is available at bottom layers in the tomography area. The details of RMS and Bias of two-day GPS and different GNSS combination tomography results are listed in Table 1 with respect to the radiosonde results. The results show that the GPS combined with GLONASS mainly reduce the bias by 31%. With the employment of Galileo and BDS, the RMS values obviously are reduced. The quad-constellation GNSS improve the accuracy by 5, 35% for RMS and Bias over the GPS tomography, respectively.

4 Conclusions

The basic principle of three-dimensional water vapor tomography is summarized and the geometric relation between the elevation angle of the satellites and the satellite ray intersecting point height with tomography boundary is deduced. Based on the GNSS stations network in Hong Kong, results indicate that the distribution of GNSS signals that can be used for tomography is quite different at different stations. For stations located at boundary of tomography area, about a half of signals cannot be used for tomography.

Using GPS and quad-constellation GNSS (GPS + GLONASS + Galileo + BDS) tomography technique, the variation of water vapor in Hong Kong during typhoon is successfully retrieved. For the quad-constellation tomography, the number of visible satellites is much more than GPS single constellation. As a result, the number of voxels without passing GNSS rays significantly reduces. Consequently, the tomography accuracy is improved by 5% compared with

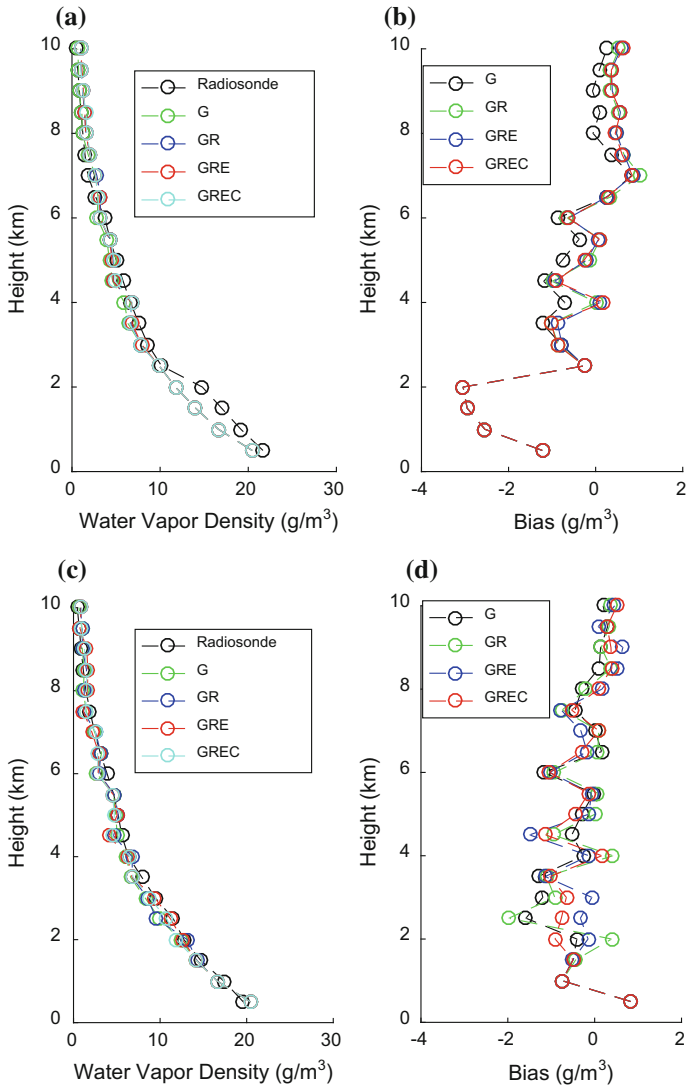


Fig. 10 The comparison of water vapor profiles (a, c) and bias (b, d) between GNSS tomography and radiosonde at 00:00–00:30 and at 12:00–12:30 on 27 August, 2017 respectively (G: GPS, R: GLONASS, E: Galileo, C: BDS)

Table 1 Comparison between GPS, multi-GNSS tomography and radiosonde profiles (g/m^3)

Comparison	RMS	Bias
GPS	0.82	-0.32
GPS + GIONASS	0.82	-0.21
GPS + GLONASS + Galileo	0.80	-0.19
GPS + GLONASS + Galileo + BDS	0.77	-0.21

GPS-only tomography. Due to the limitation of conventional tomography model and distribution of GNSS stations in a network, some disadvantages still exist for multi-GNSS tomography. For instance, there is no significant decrease of the number of voxels without passing rays at the bottom of tomography area since few signals are available at tomography boundary. Further study is needed to improve the tomography accuracy at the bottom of the atmospheric layer in the future.

Acknowledgements The contribution of data from Survey and Mapping Office, Lands Department, Hong Kong (<https://www.geodetic.gov.hk/smo/index.htm>) and GeoForschungs-Zentrum Potsdam (GFZ: <http://wzw.cn/index.php/Home/DataProduct/mgex.html>) are greatly appreciated. The software provided by Tomoji Takasu (<http://www.rtklib.com/>) is also gratefully acknowledged. The financial support from National Key R&D Program of China (No. 2016YFB0501803) and National Science Foundation of China (No. 41674039, No. 41674012) are greatly appreciated.

References

1. Li YH, Xu HM, Gao YH, Li Q (2017) The characteristic of moisture transport associated with drought/flood in summer over the east of the southwestern China. *Acta Meteorologica Sinica* 68(6):932–943
2. Li L, Jiang T, Tian Y et al (2017) Analyzing temporal-spatial variation of water vapor field for a local rainstorm in northern Hunan Province using CORS network. *J Geodesy Geodyn* 37(8): 835–840
3. Zhang K, Manning T, Wu S et al (2015) Capturing the signature of severe weather events in Australia using GPS measurements. *IEEE J Sel Top Appl Earth Observ Remote Sens* 8 (4):1839–1847
4. Bevis M, Businger S, Herring TA et al (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. *J Geophys Res Atmos* 97(D14): 15787–15801
5. Troller M, Geiger A, Brockmann E et al (2006) Tomographic determination of the spatial distribution of water vapor using GPS observations. *Adv Space Res* 37(12):2211–2217
6. Bender M, Dick G, Wickert J, et al (2009) Estimates of the information provided by GPS slant data observed in Germany regarding tomographic applications. *J Geophys Res Atmos* 114(D6)
7. Dong ZN, Cai CS (2016) Analysis of three-dimensional water vapor tomography using integrated GPS/GLONASS observations. *Journal of Geodesy and Geodynamics* 36(10): 889–891
8. Bender M, Dick G, Ge M et al (2011) Development of a GNSS water vapour tomography system using algebraic reconstruction techniques. *Adv Space Res* 47(10):1704–1720
9. Wang W, Song S, Wang J et al (2016) Distribution analysis of multi GNSS slant delays and simulated water vapor tomography in Yangtze River delta. *Acta Geodaetica Cartogr Sin* 45(2):164–169
10. Lee SW, Schutz B, Kim DH et al (2013) Monitoring precipitable water vapor in real-time using global navigation satellite systems. *J Geodesy* 87(10–12):923–934
11. Yuan Y, Zhang K, Rohm W et al (2014) Real-time retrieval of precipitable water vapor from GPS precise point positioning. *J Geophys Res Atmos* 119(16):10044–10057

12. MacMillan DS, Ma C (1997) Atmospheric gradients and the VLBI terrestrial and celestial reference frames. *Geophys Res Lett* 24(4):453–456
13. Chinese Weather. <http://typhoon.weather.com.cn>. Accessed 6 Jan 2018
14. Yao Y, Zhao Q (2016) Maximally using GPS observation for water vapor tomography. *IEEE Trans Geosci Remote Sens* 54(12):7185–7196
15. Böhm J, Möller G, Schindelegger M et al (2015) Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solutions* 19(3):433–441

GPS-MR for Altai Snow Depth Monitoring



Chenglong Zhang, Shuangcheng Zhang, Tao Che, Yajie Wang,
Ning Zhang, Wei Qu and Tianhe Wan

Abstract Snowfall as an important global freshwater resource, especially for Xinjiang Altai region in the snow all the year round, accurate and convenient snow thickness monitoring is particularly important, at present, the snow monitoring is divided into the time and space analysis of the macro snow cover and the micro snow depth monitoring. Macro monitoring are mainly composed of MODIS snow product, the micro snow depth monitoring from use of discrete meteorological point monitoring initially, laser detection, with the development of technology, satellite remote sensing monitoring has become the main trend. With the development of technology, satellite remote sensing monitoring has become a major trend. With the continuous improvement of GPS-MR snow remote sensing theory in recent years, in this paper, Altai GPS monitoring station as a demonstration using GPS-MR technology for snow monitoring research. Using the monitoring data from January 1 to March 31, 2017 for statistical analysis. The selection of satellite height Angle increase, due to the multipath effect will gradually reduce with the increase of altitude Angle, higher SNR ratio will lead to the inversion precision is decreased obviously. GPS data sampling rate has little effect on the inversion accuracy. Based on Altai GPS snow monitoring experimental station, Satellite elevation angle selection 5° – 20° , the sampling rate is set to 15 s, the GPS-MR snow depth obtained is better than 0.027 m. Therefore, it is possible to make full use of the ground-based GNSS water vapor monitoring station for snow depth detection, GNSS remote

C. Zhang · S. Zhang (✉) · Y. Wang · N. Zhang · W. Qu
College of Geology Engineering and Geomatics, Chang'an University,
Xi'an, China
e-mail: shuangcheng369@vip.163.com

S. Zhang
China Meteorological Administration Urumqi Institute
of Desert Meteorology, Urumqi, Xinjiang, China

T. Che
Northwest Institute of Eco-Environment and Resources, CAS,
Lanzhou, China

T. Wan
Xinjiang Altai Meteorological Bureau, Altai, China

sensing application potential will be brought into full play in the future environmental meteorological monitoring in our country.

Keywords Snow monitoring · Altai · GPS-MR · SNR · Snow depth

1 Introduction

Altay region of Xinjiang as one of the three major snowstorms in our country, causing great troubles to the life and production of local people, seriously restricting the development of the local economy. Snow as the main freshwater resources in arid northwest region of Xinjiang, in the local agriculture, animal husbandry plays an important role, Altay as one of the most abundant snow-covered areas in our country, most of the river's northern Xinjiang source from snowmelt runoff [1]. Therefore, it is more and more important to monitor and predict the snowfall and snow cover in the Altay region of Xinjiang.

With the launch of global remote sensing satellites, the use of satellite remote sensing to monitor snow depth has become the mainstream trend. In 2004, Che Tao and others used the improved SSM/I algorithm to estimate the snow depth distribution in China [2]. In 2005, Liu and others took the northern slope of Tianshan Economic Belt as an experimental area, based on winter MODIS 1B data for snow depth inversion, the highest accuracy is about 80% of the two channels, the rest 58–68%, the accuracy is too low [3]. In 2007, Larson was first proposed method of utilizing existing GPS continuously operating observatory soil moisture measurement, namely GPS-MR technology, and using 70-day data series to prove this method is feasible [4]. In 2009 Larson et al. [5] and Nievinski and Larson [6] used direct and reflection multipath signals in 2014 to compose time series of observed signal-to-noise (SNR), the difference of multi-path signals caused by the snowfall thickness is applied to measure the snow depth, and the change of the snowfall thickness is successfully monitored. The growth of the corn is also studied by the same method. Subsequently, he utilized multi-path reflection components of SNR observations to monitor other surface environments such as sea level, vegetation changes and volcanic eruptions, and obtained a series of extremely valuable research and application results. In 2012, Ao et al. used the multi-path reflection component in the observed GPS signal-to-noise ratio to simulate the trend of soil moisture change. And in the subsequent study, the quality of the data was evaluated, calculation of how to select the appropriate satellites to do the analysis and evaluation of the results of the inversion quality [7]. In 2012, Wu et al. gave a method of measuring the height of a water surface using a single measurement type GPS receiver reflection signal, and the use of experiments to verify the results of the tide gauge and its comparison can be a good echo [8]. In 2014, Yang et al. analyzed and evaluated the GPS observation data of Jiufeng station in Wuhan in recent ten

years, and founded that there was a consistent annual variation rule between GPS multipath effect and vegetation coverage around stations, it was concluded that the multi-path effect of GPS can be used to study the trend of vegetation coverage [9]. In 2015 Zhang in Chang'an University and others began to use GPS signal to noise ratio for snow monitoring, using Altay snow data inversion, accuracy has reached about 7 cm. It provides a powerful explanation for GPS-MR technology applied to snow monitoring [10]. In 2012, Ozeki and Heki proposed a combination of dual-frequency phases based on the difference between GPS L1 and L2 signals, as a new kind of snow depth estimation method, namely L4 phase method, they are in Japan's Hokkaido region, using GPS data from January to April 2009, using L4 and signal-to-noise ratio, respectively, to do a comparison [11]. In 2013, Li et al. used the wind-three microwave brightness temperature data to conduct snow depth inversion in northern Xinjiang. When the measured snow depth is greater than or equal to 15 cm on the northern slope of Tianshan Mountains, the minimum percentage of error is 8% [12]. In 2015, Zhang in Wuhan University and others proposed a new snow depth estimation method, combined the multipath reflectance measurements with the phase measurements of GPS tri-band signals (L1, L2, and L5) linearly. Using two sets of experimental data in different environments, the existing improvements have been made, without being affected by ionospheric delay [13]. In 2015, Shao and others used GNSS-R method to study Antarctic dry-snow data, in the dry snow depth of 10, 85, 135, 225 cm and the real data are in good agreement [14]. In December 2016, Li et al. used the single-antenna GNSS inversion principle to analyze snow cover data from September 7, 2010 to July 16, 2011 in Colorado, the correlation between the results and the actual data was 95.99% [15]. Hou et al. used the FY-3 microwave imager to retrieve snow depth, US IMS snow cover area data, as well as Altay and surrounding areas measured snow depth inversion of snow depth data, making the microwave and the average snow depth error reduced from 21.7 to 12.1 cm before correction to a corrected 3.7–1.5 cm [16].

At present, Altai Prefecture in Xinjiang mainly uses remote sensing monitoring methods. The accuracy of coverage of snow is high, however, the accuracy of snow depth monitoring needs to be further improved. In recent years, based on GPS-MR technology for snow detection has become a research hot spot, But still in the research stage. In this paper, Altay Meteorological Bureau of Xinjiang GPS snow monitoring demonstration station for the study, using GPS-MR technology for monitoring research, the detection accuracy of GPS-MR in this area is analyzed from satellite elevation. The GPS-MR snow depth obtained is better than 0.027 m. With a view to building ground-based GNSS snow monitoring system as an effective complement to snow cover monitoring means of meteorological stations in Xinjiang.

2 GPS-MR Detection of Snow Depth of the Basic Theory

The measurement type GPS receiver receives the direct signal from the satellite and the satellite signal reflected by the surface environment, reflection signals produces interference, produce multipath effect, although all are now studying how to suppress multipath effects, however, when the satellite elevation is low, multi-path affection is still obvious. The SNR value represents the ratio of signal strength to noise strength, signal to noise ratio and multipath have obvious negative correlation, the signal-to-noise ratio rises as the satellite elevation increases. The essence of GPS-MR technology is the use of spectral analysis of signal to noise ratio to obtain the surface environmental parameters that cause multipath effects, snow is also one of the surface environment [17].

$$SNR \propto P_d + P_r + \sqrt{P_d P_r} \cos \varphi \tag{1}$$

P_d In Eq. (1) is the direct signal energy, P_r is the reflected signal energy, φ is the angle between the direct signal and the reflected signal. In order to obtain the change information of GPS multipath caused by surface reflection in SNR, multi-path effects need to be separated from the received SNR observations (Fig. 1), since P_d and P_r are numerically large, the reflected signal energy can go through the formula

$$SNR(Volts/volts) = 10^{\frac{SNR(dB-Hz)}{20}} \tag{2}$$

After linearization, it is extracted by means of eliminating trend items. The residual sequence amplitude of the multipath reflected signal (Fig. 2) after removal of the direct signal trend term can be expressed as:

$$dSNR = A \cos(4\pi H \lambda H^{-1} \sin E + \phi) \tag{3}$$

Fig. 1 SNR variations of GPS signal

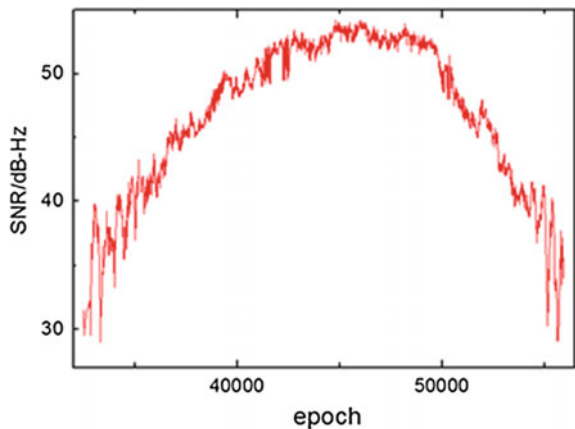


Fig. 2 SNR variations after detrending

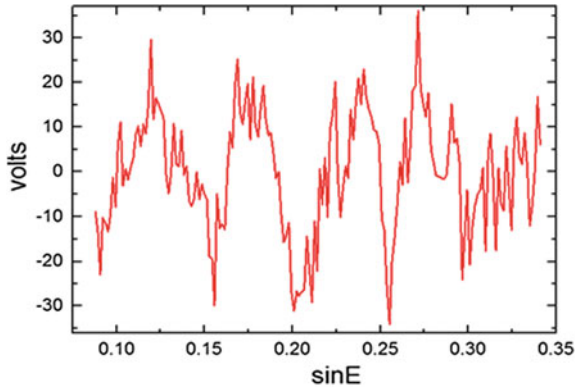
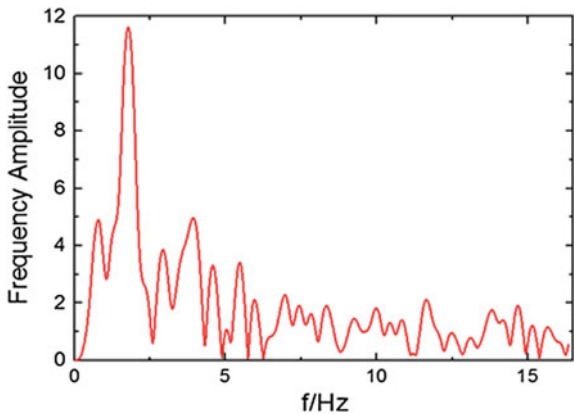


Fig. 3 Lomb-Scargle periodogram of multipath pattern



In Eq. (3), A is the amplitude, H is the reflection height, λ is the wavelength of the GPS carrier, E is the elevation angle of the satellite, and ϕ is the phase. $t = \sin E$, $f = 2H/\lambda$ the frequency can be obtained by Lomb-Scargle spectrum analysis, the vertical reflection distance H is obtained from $f = 2H/\lambda$ (Fig. 3), further, the difference between the vertical height of the station and the reflection is used to obtain the snow thickness.

3 Construction of Altai GPS Snow Monitoring Station

3.1 GPS-MR Technology to GPS Tracking Station Construction Site Requirements

- (1) Using ordinary measurement receiver placed in the open around the natural surface.

- (2) GPS receiver can output signal to noise ratio signal.
- (3) GPS receiver sets lower satellite elevation.

3.2 Altay GPS Snow Monitoring Station Layout

On December 28, 2016, the Altay GPS Snow Cover Monitoring and Testing Station (ALTA) was set up at the snow station of Altay Meteorological Station in Xinjiang at $47^{\circ} 44' 26''$ North latitude, $88^{\circ} 4' 17''$ East longitude and an average altitude of 750 m, this is currently the first snow monitoring station in China, GPS snow monitoring station using TRIMBLE NETR9, the antenna type is TRIMBLE59900, January 1, 2017 formally collected data, satellite data sampling interval is 1 s, satellite elevation is 0° .

4 Altay GPS-MR Snow Detection Experiment Analysis

First of all, this paper analyzes the feasibility of GPS-MR technology, then interception of different satellite elevation accuracy of the results of the analysis, and compared with the measured data of Altay Meteorological Bureau in Xinjiang.

4.1 Analysis of Altay Snow Depth Accuracy by Ground Based GPS

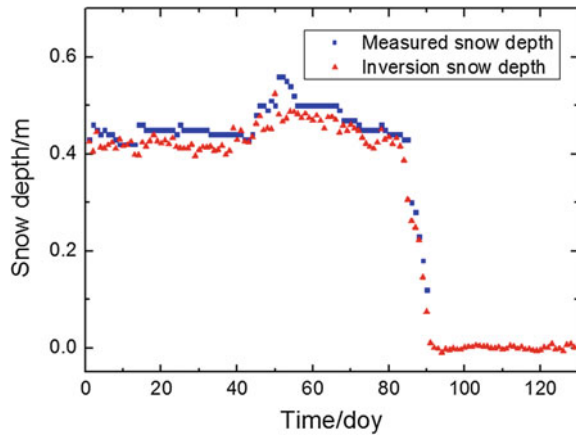
In order to verify the feasibility and accuracy of GPS-MR technology for Altay snow monitoring, firstly, the GPS data of 129 days from January 1, 2017 on the satellite whose satellite elevation is 5° – 20° and the sampling rate is 30 s are explored.

It can be seen from the Fig. 4 above that the depth of snow depth measured by GPS-MR is very close to the measured snow depth on the order of magnitude and trend coincidence, the correlation coefficient is 0.94, the experimental results show that GPS-MR technology has a higher accuracy of snow depth detection.

4.2 Monitoring and Analysis of GPS-MR Snow Depth in Satellite Elevation

At low satellite elevation, due to the obvious multipath effect of surface environmental parameters, the data of low satellite elevation are processed to obtain the

Fig. 4 ALTA station
GPS-MR snow depth contrast
Map



detected snow depth. Whether the accuracy of the detection accuracy is affected by the selection of the length of the data sequence at a low satellite elevation, next, three different satellite elevations of 5° – 20° , 5° – 25° and 5° – 30° are used to explore the accuracy of GPS-MR technology in detecting snow depth, due to the current GPS tracking station is mainly used for tracking stations and crustal deformation monitoring, in order to reduce the impact of multipath effects, satellite elevation is generally set at 5° , the data sampling rate is generally 30 s, in order to get the result with universal reference value, this paper also collected the 129 days GPS data of different satellite elevation angle interval selection using the same settings for analysis.

Figure 5 blue represents the Altay meteorological data provided by the measured data, red is the satellite elevation interval 5° – 20° , Purple is the satellite elevation of 5° – 25° , green is the satellite elevation of 5° – 30° GPS-MR inversion snow depth. Satellite altitude range. It is apparent from Table 1 and Fig. 5 as the satellite

Fig. 5 Different satellite
elevation of snow depth and
measured snow depth map

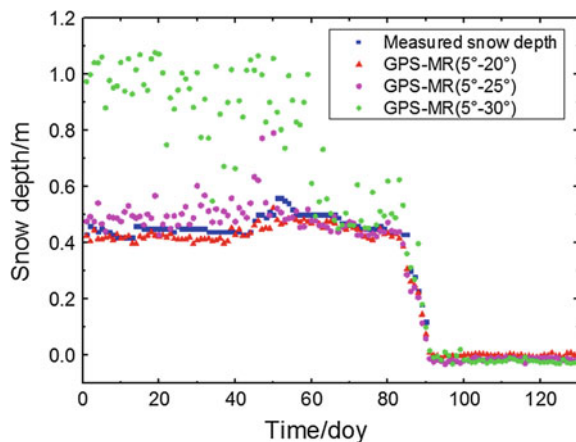


Table 1 Statistical analysis of snow depth and measured snow depth on different height angles at ALTA station

	5°–20°	5°–25°	5°–30°
Pearson's r	0.944	0.771	0.412
Poor mean (m)	0.027	0.053	0.331
RMSE (m)	0.022	0.068	0.232

elevation interval increasing, signal to noise ratio residual sequence will become longer, however, the correlation coefficient is getting smaller and smaller, RMS bigger, the difference with the measured snow depth and the trend is more and larger.

5 In Conclusion

This article uses the GPS observation data independently arranged in Altay, Xinjiang as the original data to be processed, and the use of different satellite elevation on the accuracy of GPS-MR technology to explore the following conclusions drawn:

1. Through the experimental verification of GPS-MR technology, it can effectively detect the snow depth of Altay, Xinjiang, the GPS-MR snow depth obtained is better than 0.027 m.
2. The accuracy of GPS-MR technology has a bigger impact on different satellite elevation, the distribution of data performance is also gradually unstable, with the increase of the satellite elevation, and Multi-path effect is obviously weakened, and residual sequence oscillation characteristics become insignificant, lack of sufficient information to obtain inversion parameters.

In this paper, Altay area by building ground-based GPS snow monitoring demonstration station, and the use of GPS-MR technology for detecting the validity of Altay snow depth of snow; And the precision of GPS-MR is analyzed from different satellite elevation, gradually improved the GPS-MR experimental verification of practice, broaden the scope of CORS network applications, give full play to the important role of ground-based GPS stations in global environmental change monitoring, in order to be able to serve as an effective supplement to the Altay Meteorological Bureau's foundation environmental monitoring.

Acknowledgements Thanks to the Northwest Institute of Ecology and Environmental Resources, Chinese Academy of Sciences and the Altay Meteorological Bureau in the construction of snow monitoring station to give the support and assistance. This research is supported by the National Natural Science Foundation of China (41104019; 41674001; 41731066); Fundamental Research Funds for Research Funds of Central Universities (310826172202). Thanks to the editors and anonymous referees provide for the valuable comments and suggestions in this paper!

References

1. Xu J, Shu H, Liu Y (2014) 2000–2010 snow disaster temporal and spatial autocorrelation analysis of Xinjiang. *Disaster Sci* 29(01):221–227
2. Che T, Li X (2004) Retrieval of snow depth and its accuracy evaluation by using passive microwave remote sensing data. *Sens Technol Appl* 05:301–306
3. Liu Y, Zhang P, Li Y, Li G (2005) Snow depth inversion based on MODIS data—a case study of the North Tianshan economic belt. *Geogr Geo Inf Sci* 06:41–44
4. Bilich A, Larson KM (2007) Mapping the GPS multipath environment using the signal-to-noise ratio (SNR). *Radio Sci* 42(6):1–16
5. Larson KM et al (2009) Can we measure snow depth with GPS receivers? *Geophys Res Lett* 36(17):L17502
6. Nievinski FG, Larson KM (2014) Inverse modeling of GPS multipath for snow depth estimation. *IEEE Trans Geosci Remote Sens* 52(10):6561–6573
7. Ao S, Hu Y, Liu Y et al (2012) Inversion of soil moisture trend observed by GPS signal-to-noise ratio. *Chin J Surv Mapp* 29(2):140–143
8. Wu J, Yang R (2012) Measurement of water surface height using GPS receiver reflections. *J Geodesy Geodyn* 32(6):135–138
9. Yang Y, Hao X, Peng B (2014) Research on vegetation coverage based on multi-path effect. *Proc Geom* 39(2):52–54
10. Zhang S, Dai K, Liu K, Hou X, Zhao Y (2016) Study on snowfall monitoring using GPS-MR technology. *Progr Geophys* 04:1879–1884
11. Ozeki M, Heki K (2012) GPS snow depth meter with geometry-free linear combinations of carrier phases. *J Geodesy* 86(3):209–219
12. Li Y, Zhang P, Liu Y, Lu X, Liyun M (2013) Research on snow depth inversion using three winds of microwave radiation data in North Xinjiang. *Soil Water Conserv Res* 20(04):166–171
13. Yu K, Ban W, Zhang X, Yu X (2015) snow depth estimation based on multipath phase combination of GPS triple-frequency signals. *IEEE Trans Geosci Remote* 53(9):5100–5109
14. Shao L, Zhang Y, Meng W, Yuan G (2015) Research on dry snow depth detection based on GNSS-R. *Electron Design Eng* 23(17):9–12. <https://doi.org/10.14022/j.cnki.dzsjgc.2015.17.003>
15. Li B, Zhang Y, Yang S, Hong Z, Han Y (2016) Inversion of snow cover thickness of single antenna GNSS reflection signal. *JPG* 41(06):37–41. <https://doi.org/10.13442/j.gnss.1008-9268.2016.06.008>
16. Hou X, Zhang P, Zheng Z, Li S (2015) Snow depth inversion in Aletai area based on multi—source data. *Deep Invers Sens Technol Appl* 30(01):178–185
17. Larson KM, Nievinski FG (2013) GPS snow sensing: results from the earth scope plate boundary observatory. *GPS Solut* 17:41–52

Snow Depth Detection Based on L2 SNR of GLONASS Satellites and Multipath Reflectometry



Wei Zhou, Lilong Liu, Liangke Huang, Jun Chen and Songqing Li

Abstract With the continuous improvement of Global Navigation Satellite System (GNSS) theory, the signal-to-noise ratio (SNR) of a GNSS reflectometry signal can be applied to detect snow depth. However, there are still some issues, such as insufficient observation data, low detection precision, etc. To solve the above problems, this paper uses the SNR reflectometry data of GLONASS L1 and L2 to detect snow depth in Yellowknife, Canada, from July 2015 to June 2016. Then we analyzed the L1 and L2 SNR-derived snow depths and the average snow depths. The results show that snow depth detected using the two SNR signals can reach centimeter level. There is weak bias and strong correlation when comparing the detected snow depth based on which single-frequency SNR observations with in situ measurements are used. For L1 and L2 SNR-derived snow depths in a separate 365-day campaign, the former bias is superior to the latter. Both RMSE values are 4.5 and 2.6 cm. The stability of L2 SNR-derived snow depth was improved by over 40% than that of L1 SNR. The average snow depths detected using the two SNR signals have no significant improvement on the precision, but improve the spatial resolution because of more satellites. The experimental results show that the applications of GNSS technology can be further extended by GLONASS-MR technology based on L2 reflectometry signals.

Keywords Global navigation satellite system · Multipath reflectometry
Signal-to-noise ratio · Snow depth detection

W. Zhou · L. Liu (✉) · L. Huang (✉) · J. Chen · S. Li
College of Geomatic Engineering and Geoinformatics,
Guilin University of Technology, Guilin, China
e-mail: hn_liulilong@163.com

L. Huang
e-mail: lkhuang666@163.com

W. Zhou · L. Liu
Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin, China

L. Huang
GNSS Research Center, Wuhan University, Wuhan, China

1 Introduction

Snow is one of the most essential elements affecting climate change in the cryosphere of the earth, which plays a crucial role in the Earth's hydrology and surface energy balance [1, 2]. In atmospheric circulation, snow is a key storage method for freshwater, which affects climate through a series of interactive feedback mechanisms [3]. Therefore, it is important to analyze this variability in order to identify possible trends. Since the L-band carrier waves of GNSS satellites can weaken the attenuation through the atmosphere and change sensitively for water content variations on the surface, GNSS-R (GNSS Reflectometry) technology has been widely used in the sea-surface altimetry measurement [4–6], sea ice [7–9], sea-surface wind [10, 11], soil moisture [12, 13] and other fields. Nowadays, some researchers have proposed to detect snow depth based on GNSS-R [14–17] technology.

Based on this relevant research, domestic and foreign researchers have obtained some basic achievements. Larson et al. [14, 18] used reflectometry signals affected by multipath to be as a signal source, and established the functional relationship between surface reflectivity and soil moisture to estimate the variations. Then this algorithm is used to verify whether it is suitable for snow depths detected using reflectometry signals of CORS network. Ozeki et al. [19] proposed the algorithm of snow depth detection using the ionospheric geometrical linear combination (L4) of GPS reflectometry signals and obtained results similar to GPS SNR. Yu et al. [20] combined SNR observations of GPS triple-frequency signals to obtain snow depths. Jin et al. [21] analyzed the results from using GPS L2P SNR to retrieve snow depth in Alaska and obtained even better correlation coefficients compared with L1 C/A.

To furtherly expand more GNSS observations and improve the precision of snow depth detection, in this paper, we explore to use the SNR reflectometry data of GLONASS L1 and L2 to detect snow depth in the same area, respectively. Additionally, the detected snow depths were separately compared to in situ measurements. Finally, we try to analyze the accuracy and reliability of snow depth detected using GLONASS-MR (GLONASS Multipath Reflectometry) with two different SNR signals in a fixed area.

2 GLONASS-MR Model and GLONASS Data Processing

Since there is shortage of snow depth monitoring station, its distribution is not balanced, and in situ snow depth measurements are discontinuous. Additionally, the remote sensing technology can only observe the global snow cover with high coverage but cannot feedback the depth information of a wide range of snow cover. Therefore, it is proposed to detect snow depth based on GLONASS-MR technology. Figure 1 shows the basic structure diagram of monitoring snow depth based on GPS-MR technology.

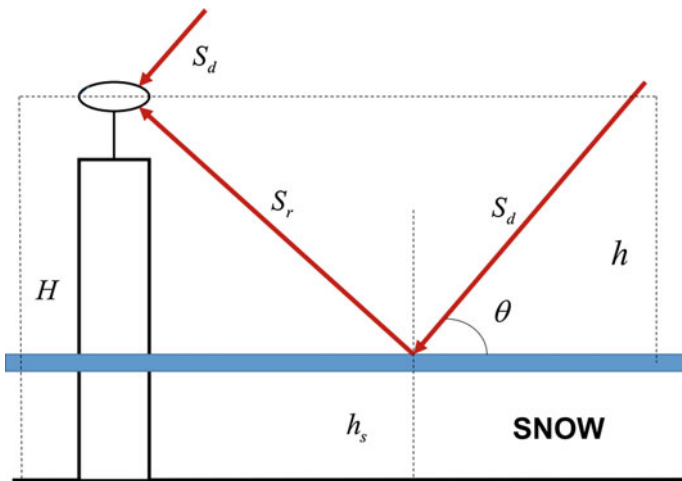


Fig. 1 The structure diagram of monitoring snow depth based on GLONASS-MR technology. The structure is meant to capture the different carrier signals through the GNSS receiver antenna

Since the reflectometry signals are greatly affected by the surface medium, phase and amplitude have changed. We can obtain snow depths on the surface when measure the above characteristics. The signal received by the GPS receiver includes the direct signal (S_d) and the reflected signal (S_r) through the surface medium. Both GLONASS signals will interfere with each other to produce the composite signal (S_c). The mathematical formula of the composite signal [22] is as follows:

$$SNR \approx S_c = S_d + S_r = A \sin(\omega_0 t) + \sum_{i=1}^n \alpha_i A \sin(\omega_0 t + \Delta_i) \tag{1}$$

where A is the carrier amplitude value, α_i is the reflection coefficient, ω_0 is the carrier angle frequency of the received reflected signal, Δ_i is the phase delay of the signal, and t is the observation period.

For the multipath effect, the mathematical formula for the distance from the reflector to the receiver antenna, the phase offset of the multipath signal and the incident angle of the direct signal are shown below:

$$\Delta_1 = \frac{2h}{\lambda} \sin\theta \tag{2}$$

where Δ_1 is the phase offset of the multipath signal, h is the vertical distance from the GPS receiver antenna to the snow surface, λ is the carrier wavelength, and θ is the incident angle of the direct signal.

As for the GPS satellites, the surface reflection layer and the receiver antenna change the geometric position, which will make the reflected angle and multipath

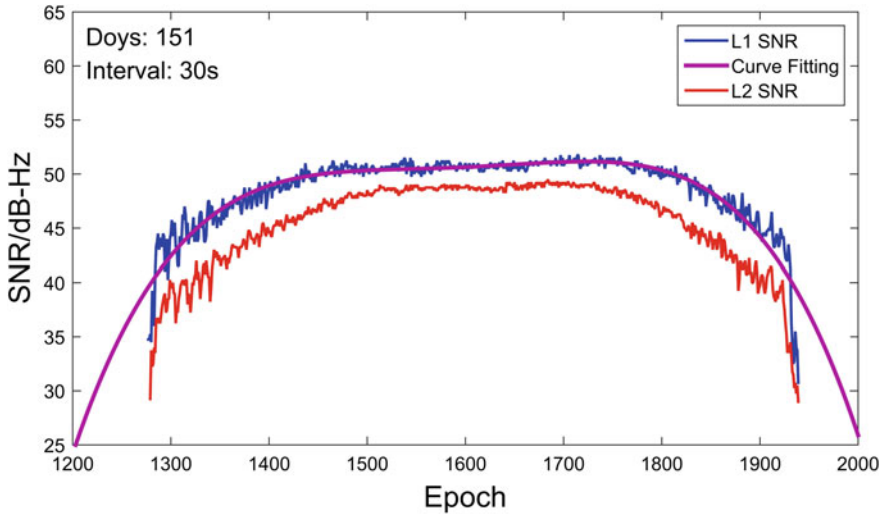


Fig. 2 Observed series of SNR1 and SNR2 (transformed to linear scale) for GLONASS satellites (elevation angle from 0° to 25°) for one successive day

error simultaneously occur with seasonal changes. For the general GNSS receiver, $S_r \ll S_d$ should be satisfied to suppress the multipath error caused by surface reflection. The received composite signal contains the required SNR data. In order to obtain the variation of the GPS SNR signal caused by the multipath process, the SNR data of the direct signal and the reflected signal should be separated. Generally speaking, the low-order polynomial interpolation fitting method is used to eliminate the trend of the composite signal, which is the direct signal. Figure 2 shows the trend of the SNR signal from GLONASS satellites in the YEL2 observation station. The sampling rate of the GPS receiver is 30 s.

In the above Fig. 2, the overall trend of the SNR data for the two L-band reflected signals shows a parabolic shape. The SNR data in the middle is stable, and the residual value is low. Because the receiver receives the signal in a large elevation angle, which is less affected by the multipath, the large fluctuations of SNR data at both ends are mainly caused by the influence of the larger multipath at the low elevation angle. The reflected signals contain a large number of surface object information, resulting in the instable phase and amplitude of the satellite signal. The phase and amplitude have been instable. The stability of the residual values sharply changed, which shows the existence of heteroscedasticity, and a non-linear interpolation method can be used to fit the direct signal.

Removing the overall trend of the SNR data in Fig. 2, and the SNR data are obtained at low altitude angles. The above obtained SNR data are residuals sequence, elevation range 0° – 20° . For the above reflected signal after removing the direct component, the mathematical formula of the reflected signal is as follows:

$$S_r = A \cos\left(\frac{4\pi h}{\lambda} \sin E + \varphi\right) \quad (3)$$

where A and λ are the carrier amplitude and wavelength, respectively, E is the satellite elevation angle, h is the effective vertical reflection distance, and φ is the phase value less than one integer period.

It is generally difficult to use the fast Fourier transform to solve the problem that using the non-equal interval method to resample the satellite elevation angle cannot intercept the entire SNR residual sequence. Since the Lomb-Scaglie (L-S) transform can extract the weak periodic signal of the time series, it can partially weaken the false signal caused by the non-uniform time series. Therefore, the problem can be solved by L-S spectral analysis. We used L-S spectral analysis to analyze formula (3), and the highest frequency (f) is obtained. Then we can accumulate the vertical reflected distance (h) for the formula ($f = 2h/\lambda$), and the detected snow depth (h_s) can be calculated for the formula ($h_s = H - h$). Therefore, Fig. 3 shows the basic flow diagram of snow depth detected using GLONASS-MR technology.

3 Results and Discussion

To verify whether the reliability of the detection precision when using the SNR signals of the GLONASS two-frequency reflected signals to detect snow depths, the experiment uses GLONASS observations of the YEL2 station from July 2015 to June 2016 to detect snow depths. Meanwhile, Then, the Root-Mean-Square Error (RMSE), Bias, and R-square (R^2) are used to evaluate the relative precision of snow depth. For all of the above, they are expressed by the mathematical formulas (4):

$$\left\{ \begin{array}{l} \text{Bias} = |X_{model,i} - X_{obs,i}| \\ \text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \\ \text{R}^2 = 1 - \frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{\sum_{i=1}^n (X_{obs,i} - \bar{X}_{obs,i})^2} \end{array} \right. \quad (4)$$

For formula (4), n is the total amount of experimental data; i is the number of experimental data, where $i = 1, 2, 3 \dots n$; $X_{obs,i}$ and $\bar{X}_{obs,i}$ are, respectively, in situ measurements and its mean; and $X_{model,i}$ is snow depth detected using GLONASS-MR model.

In order to determine the influence of different snow depths, we randomly selected GLONASS reflectometry data of the YEL2 observatory station for two days. Components (a) and (b) of Fig. 5 separately show the residual sequence of the GLONASS SNR (L1 and L2) signal and its L-S spectral analysis. For Fig. 4a, the horizontal axis represents the sine of the GPS satellite elevation angle, and the vertical axis represents the SNR signal with the removal of the direction signal.

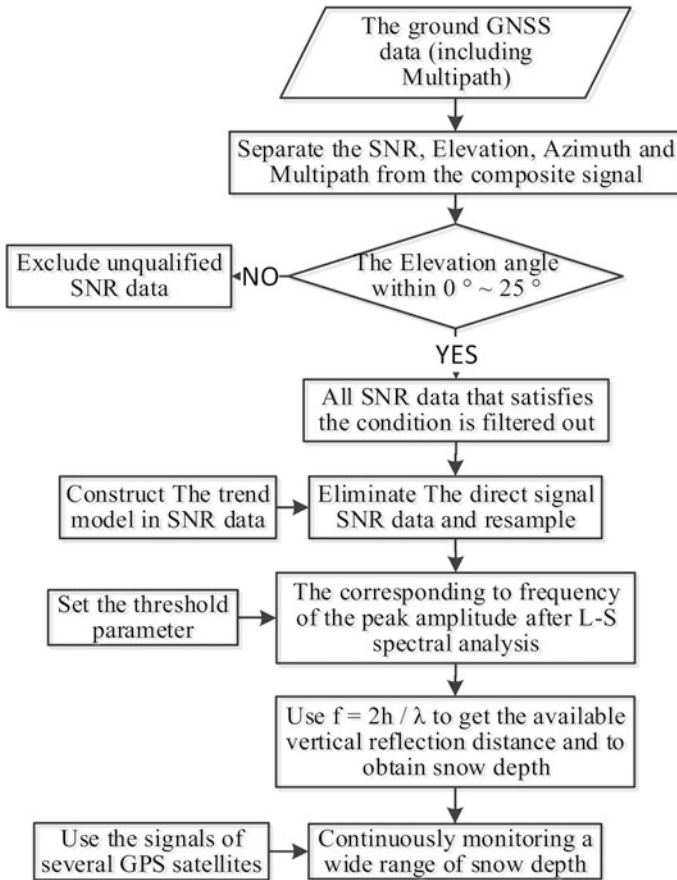


Fig. 3 The basic flow diagram of monitoring snow depth based on GLONASS-MR technology

In Fig. 4b, the horizontal axis represents the vertical reflection distance from the center of the GPS receiver antenna to the snow surface, and the vertical axis indicates the spectral amplitude intensity of the SNR signals. The vertical reflection distance (h) is obtained through L-S spectrum analysis, and the horizontal axis value that the highest amplitude of the spectral amplitude corresponds to is h . Then snow depths can be obtained.

For Fig. 5a, with snow depth gradually increasing, the fluctuation range of the SNR signal residual sequence are not obvious but the oscillation frequency is also increasing. The surface is smoother and smoother due to the increase of snow depths, a specular reflection has happened for the reflected signal, and the increase in the elevation angle weakens the diffuse reflection on the ground surface. Meanwhile, component (b) shows that the spectrum amplitude peak values of GLONASS L1 and L2 gradually move to the right when snow depths are

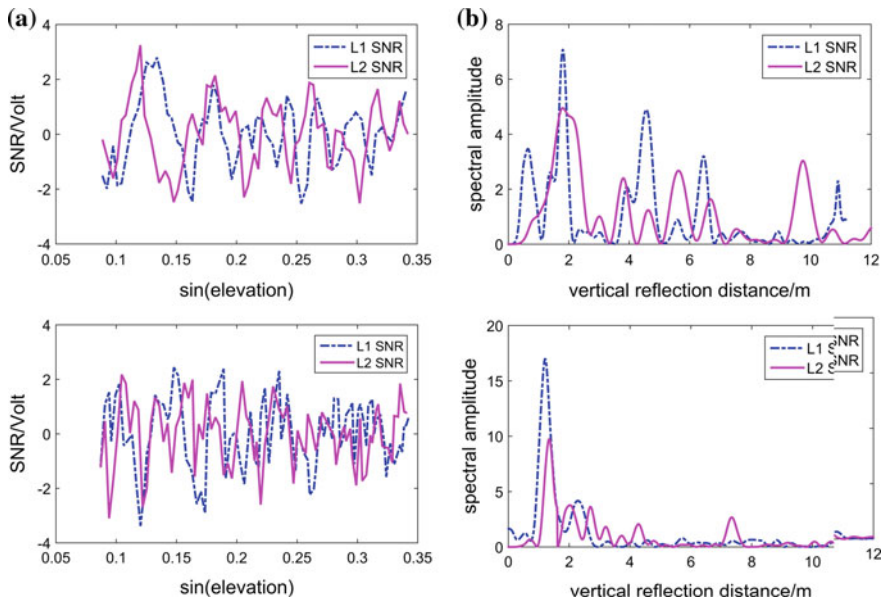


Fig. 4 The residual sequence and L-S spectral analysis of GLONASS SNR on different days. **a** is the residual sequence values of SNR data; and **b** is the figures of L-S spectral analysis

increasing, which have a significant difference that increase in snow depth makes the spectrum amplitude peak values have similar the changed trend. What’s more, the spectrum amplitude peak values of L1 SNR are more than that of L2 SNR, whose impossible reason is that the frequency of GLONASS L1 and L2 carriers. The result shows that the variations of different snow depths are closely related to the dynamic changes in the energy spectrum of low-elevation reflection components, and the changes in snow depth produce some regular changes to the SNR observations of the reflected signal.

We furtherly analyze snow depth detected using the SNR data of L-band reflectometry signals for the 365 consecutive days, which is shown in Fig. 5. Where the horizontal axis represents the continuous time series of the detected snow depth, and the vertical axis represents the depth of snow cover. The legend shows that the dots are in situ measurements, the star and the triangle points are the detected snow depth based on SNR of GLONASS L1 and L2, respectively, and the rhombus points are the mean of the detected snow depths. Figure 5 reflects the overall trend between the detected snow depth based on GLONASS-MR and in situ measurements, which has a high consistency and diversity in the continuous time series. However, the point distribution of L2 SNR-derived snow depths is obviously concentrated compared with in situ measurements than that of L1 SNR, and the tendency of the mean is similar to that of L2 SNR.

This experiment made a comparison of the bias between the detected snow depth and in situ snow depth measurements for the 365 consecutive days, and the results

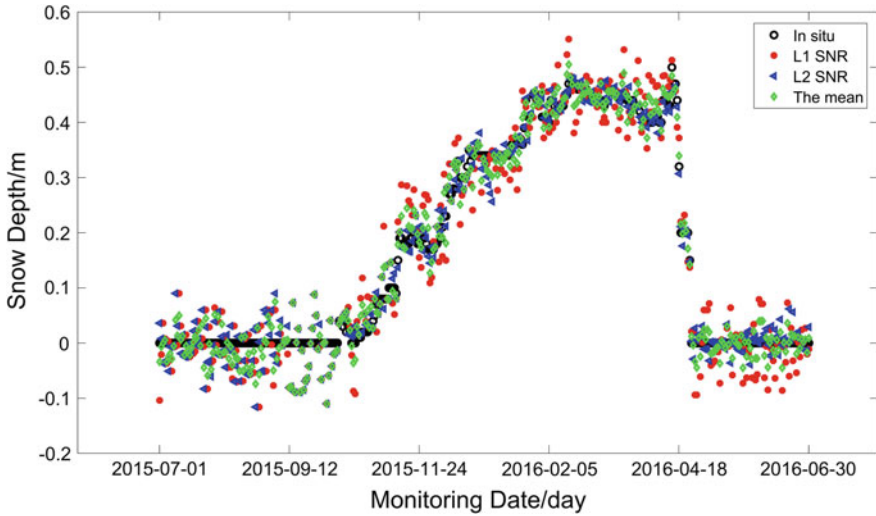


Fig. 5 The comparison between the detected snow depths based on GLONASS-MR technology and in situ measurements

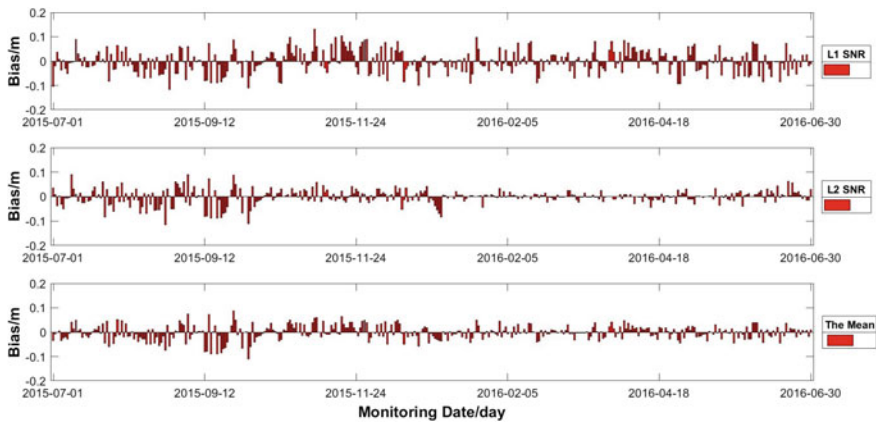


Fig. 6 The bias values between the detected snow depth and in situ measurements for the 365 consecutive days

are shown in Fig. 6 and Table 1. For Fig. 6, the horizontal axis represents the continuous time series of the detected snow depth, and the vertical axis represents the bias between the detected snow depth and in situ snow depth measurements. The bias values between snow depths based on GLONASS-MR over a long period and in situ snow depth measurements are significantly different. The L1 SNR is the least effective for all detected snow depth, and L2 SNR is better than that of L1. Both RMSE values are 4.5 and 2.6 cm, and their correlation coefficient are 0.948

Table 1 The bias precisions of the detected snow depth based on L1 and L2 SNR and the in situ snow depth measurements

The detected snow depth	The average bias/cm	RMSE/cm	The improvement rate	R ²
L1 SNR	3.5	4.5	–	0.948
L2 SNR	2.0	2.6	42.2%	0.978
The average value	2.1	2.8	37.8%	0.977

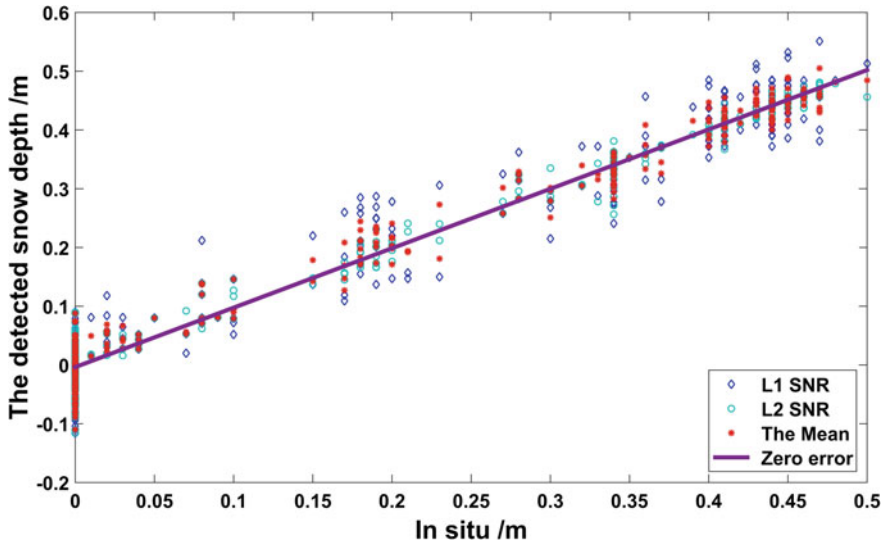


Fig. 7 The correlation analysis between the detected snow depths based on GLONASS-MR and the in situ snow depth measurements

and 0.978. Compared with L1 SNR-derived snow depth, which improved by over 40% of the bias precision for L2 SNR-derived snow depth. We calculated the mean of snow depth detected using L1 and L2 SNR, which has no significant difference. It can be seen that a worse match occurs at the beginning and ending of the snowfall process (early November and late April), whereas the detection error is relatively small over the middle period of the snowfall season. The main reason would be that the snow depth is virtually zero in early November and late April, and other information is added to the reflectometry signal.

Finally, the experiment analyzed the correlation and showed the result in Fig. 7. The horizontal axis shows the in situ snow depth measurements, and the vertical axis shows the snow depths based on GLONASS-MR technology. The trend shows a clear linear relationship. From Table 1, the correlation coefficients between the detected snow depth based on all of the above SNR signals from GPS satellites and in situ snow depth measurements are more than 90%. The experimental result declared the superiority of GLONASS-MR based on the L2 SNR in detecting snow depths.

4 Conclusion

In this paper, we mainly discussed the basic theory of GNSS-R and explained how to use the SNR data of GLONASS reflectometry signals to detect snow depth. Additionally, analyzed the bias precisions for snow depth detected using L1 and L2 SNR data. Finally, the experiment used the GLONASS observation data to verify the above GLONASS-MR model. Based on the above theoretical analysis, the experimental results are shown as follows. Firstly, snow depths detected using GLONASS-MR based on L1 and L2 SNR data have the centimeter-level precision. Secondly, for all detected snow depth, the precision of L2 SNR is obviously better than that of L1. Compared with in situ snow depth measurements, the correlation coefficient reaches 0.978. The stability of L2 SNR-derived snow depth was improved by over 40% than that of L1 SNR. The thirdly, the average bias values of snow depths detected based on GLONASS L1 and L2 SNR are 3.5 and 2.0 cm, respectively. What's more, by comparing the bias and RMSE for snow depth detected using GLONASS-MR and the mean value, the mean value is better than the snow depth detection of L1 SNR data but has no significant difference with that of L2 SNR. The experimental results show that GLONASS-MR technology based on L1 and L2 SNR has high feasibility in detecting the dynamic variations of snow depth. Meanwhile, it shows the superiority of GLONASS-MR based on L2 reflectometry signals, which provides a foundation for accurate measurement in the future. However, due to the limited experimental data used in this experiment, it still needs to be verified whether GLONASS-MR method has a good applicability in other GNSS observatory stations.

Acknowledgements This work was sponsored by the National Natural Foundation of China (41704027; 41664002); the “Ba Gui Scholars” program of the provincial government of Guangxi; Guangxi Natural Science Foundation of China (2017GXNSFBA198139; 2017GXNSFDA198016); the Guangxi Key Laboratory of Spatial Information and Geomatics (16-380-25-01); and the basic ability promotion program for young and middle-aged teachers of Guangxi (KY2016YB189). The authors would like to thank the International GNSS Service Center (IGS) for providing the GPS observation and the National Climatic Data Center (NCDC) for providing the in situ snow depth data.

References

1. Frei A et al (2012) A review of global satellite-derived snow products. *Adv Space Res* 50 (8):1007–1029
2. Gleason S (2010) Towards sea ice remote sensing with space detected gps signals: demonstration of technical feasibility and initial consistency check using low resolution sea ice information. In: AGU Fall Meeting, 2010
3. Solomon S (2013) Climate Change 2007: the physical science basis: contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, 2013, pp 159–254

4. Wang X, Zhang Q, Zhang S (2017) Water levels measured with SNR using wavelet decomposition and Lomb-Scargle periodogram. *GPS Solut* 22(1):22
5. Jin S, Qian X, Wu X (2017) Sea level change from BeiDou navigation satellite system-reflectometry (BDS-R): first results and evaluation. *Glob Planet Change* 149:20–25
6. Yu K (2016) Weak tsunami detection using GNSS-R-based sea surface height measurement. *IEEE Trans Geosci Remote Sens* 54(3):1363–1375
7. Zhu Y et al (2017) Sea ice detection based on differential Delay-Doppler maps from UK TechDemoSat-1. *Sensors* 17(7):1614
8. Yan Q, Huang W, Moloney C (2017) Neural networks based sea ice detection and concentration retrieval from GNSS-R Delay-Doppler maps. *IEEE J Sel Top Appl Earth Obs Remote Sens* 10(8):3789–3798
9. Strandberg J, Hobiger T, Haas R (2017) Coastal sea ice detection using ground-based GNSS-R. *IEEE Geosci Remote Sens Lett* 99:1–5
10. Jing C et al (2016) Retrieval of sea surface winds under hurricane conditions from GNSS-R observations. *Acta Oceanol Sin* 35(9):91–97
11. Zhou X (2012) Simulation techniques of GNSS-R sea surface wind field retrieval from airborne remote sensing. *J Remote Sens* 16(1):143–153
12. Yan S et al (2017) Feasibility of using signal strength indicator data to estimate soil moisture based on GNSS interference signal analysis. *Remote Sens Lett* 9(1):61–70
13. Wan W et al (2015) Initial results of China's GNSS-R airborne campaign: soil moisture retrievals. *Sci Bull* 60(10):964–971
14. Larson KM et al (2009) Can we measure snow depth with GPS receivers? *Geophys Res Lett* 36(17):L17502
15. Fabra F et al (2010) Monitoring sea-ice and dry snow with GNSS reflections 38(5):3837–3840
16. Jin S, Komjathy A (2010) GNSS reflectometry and remote sensing: new objectives and results. *Adv Space Res* 46(2):111–117
17. Jin S, Feng GP, Gleason S (2011) Remote sensing using GNSS signals: current status and future directions. *Adv Space Res* 47(10):1645–1653
18. Larson KM et al (2008) Using GPS multipath to measure soil moisture fluctuations: initial results. *GPS Solut* 12(3):173–177
19. Ozeki M, Heki K (2012) GPS snow depth meter with geometry-free linear combinations of carrier phases. *J Geodesy* 86(3):209–219
20. Yu K et al (2015) Snow depth estimation based on multipath phase combination of GPS triple-frequency signals. *IEEE Trans Geosci Remote Sens* 53(9):5100–5109
21. Jin S, Qian X, Kutoglu H (2016) Snow depth variations estimated from GPS-reflectometry: a case study in Alaska from L2P SNR Data. *Remote Sens* 8(1):63
22. Huang DF et al (2001) Wavelet Filters Based Separation of GPS multi-path effects and engineering structure vibrations. *Acta Geodaetica Et Cartographic Sinica* 30(1):36–41

Analysis of Ionospheric Anomalous Disturbance During a Heavy Rainfall



Lulu Shan and Yibin Yao

Abstract The global navigation satellite system (GNSS) has the advantages of high precision, high temporal resolution and low cost. The total electronic content (TEC) can be calculated by GNSS observations. The TEC sequence can be used to reflect the disturbances of the ionosphere, and it can detect the ionospheric anomalies caused by strong convective weather which plays an important role in studying the coupling mechanism between the ionosphere and the troposphere. In this paper, the GNSS observations during a heavy rainfall at July 9, 2013 in Sichuan Province were selected, and the de-trend TEC sequence was obtained. It can be seen from the dTEC sequence that there are obvious anomalies in the ionosphere when the rainfall occurred, and the abnormal times between the different stations are slightly different which show the abnormal propagation process. In this paper, we argue that the ionospheric anomalies are related to the acoustic gravity waves (AGWs). Due to the special terrain of Sichuan, when the waves encountered the mountain, they would spread upward, and then reach the ionosphere and cause ionospheric anomalies. According to its propagation velocity, Sichuan Province is divided into uniform grid and the approximate location of the excitation AGWs can be found. Finally, based on the analysis of the heavy rainfall event, the coupling mechanism between ionosphere and troposphere is discussed.

Keywords GNSS · TEC · Heavy rainfall · Ionospheric disturbance
Acoustic gravity waves (AGWs) · Ionosphere-troposphere coupling mechanism

L. Shan (✉) · Y. Yao
School of Geodesy and Geomatics, Wuhan University,
129 Luoyu Road, Wuhan, China
e-mail: llshan@whu.edu.cn

Y. Yao
Key Laboratory of Geospace Environment and Geodesy,
Ministry of Education, Wuhan University, 129 Luoyu Road, Wuhan, China

1 Introduction

There are a large number of ions and electrons in the ionosphere. Anomalous ionosphere can have an important harm on radio technology and spacecraft. Therefore, the ionospheric disturbance research is of great significance [1]. Studies have shown that the ionosphere is not only affected by solar radiation and geomagnetic changes, but also by changes in the troposphere, such as atmospheric disturbances and atmospheric circulation [2]. Severe weather such as thunderstorms, typhoons, cold spells, tsunamis, tornadoes and so on will cause varying degrees of impact on ionosphere, then causing ionospheric anomalies [3–5].

Many scholars pointed out that the weather system in the troposphere can have a significant impact on the ionosphere. In the 1950s, Bauer studied the ionospheric responses to the typhoon and cold front during the transit. It was found that f_0F_2 reached the maximum value within 1–3 h after the typhoon center registered the coastline [6]. Shen analysed a dozen typhoons passing through Hainan Island and found that it had a significant effect on the ionospheric f_0F_2 layer [7]. Mao et al. analysed the effect of typhoon ‘Matsa’ on the Total Electron content (TEC) of the ionosphere, and described in detail the changes of TEC before and after the typhoon landed [8]. Liu et al. found that the Indian tsunami of 2004 triggered atmospheric anomalies near the sea surface and propagated upward to the ionosphere, causing a disturbance in the ionospheric electron density [9].

So far, there has been no comprehensive explanation of the coupling mechanism between the ionosphere and the troposphere. A common view is that various meteorological processes generate gravity waves in the lower atmosphere, which then propagate upward to the ionosphere and cause ionospheric anomalies [10]. Tsutsui et al. confirmed by HF Doppler observations that typhoon-induced gravity waves can produce mesoscale Traveling Ionospheric Disturbances (TIDs) [11]. Xiao et al. studied the entire process of ionosphere responses to typhoons, and pointed out that the Acoustic gravity waves (AGWs) produced during the typhoons can propagate to the ionosphere and affected the ionosphere [12]. Using GNSS multi-station observations, Kong et al. analysed the ionospheric TEC during the transit of Typhoon ‘Dujan’, and calculated the horizontal velocity and radial velocity of the AGWs propagation, then obtained the approximate location where excited AGWs [13].

The specific mechanism of the coupling between low, middle atmosphere and high-level atmosphere is exactly the problem that needs to be further studied and solved. This paper analysed the ionospheric anomalies from a heavy rainfall event in Sichuan Province and discussed possible coupling mechanisms between the ionosphere and the troposphere.

2 Data Selection and Processing

On the evening of July 9, 2013 (DOY190), Sichuan Province suffered heavy rainfall in succession, some areas are severely affected, such as Chengdu, Mianyang and Ya'an. As at 16:00 on July 10, the torrential rainfalls caused 9 deaths, 62 missing and more than 1.45 million people were affected. In order to understand the ionospheric disturbance caused by heavy rainfalls, this event was selected for analysis.

Selecting the GNSS observations of Crustal Movement Observation Network in Sichuan Province, a total of 23 stations, the distribution of stations as shown in Fig. 1, the coordinates of the station in Table 1. The date was one day on July 9, 2013 (DOY 190) with a temporal resolution of 30 s.

The TEC can be calculated by using the dual frequency pseudo-range observations and the dual-frequency phase observations respectively [14] as shown in Eq. (1):

$$\begin{cases} TEC_R = \frac{f_1^2[(P_1 - P_2) - B_i - B^p]}{40.28(1 - \gamma)} \\ TEC_\varphi = \frac{f_1^2[(\lambda_1 \varphi_1 - \lambda_2 \varphi_2) - (\lambda_1 N_1 - \lambda_2 N_2) - b_i - b^p]}{40.28(\gamma - 1)} \end{cases} \quad (1)$$

where, TEC_R is the TEC calculated using dual frequency pseudo range observations, and TEC_φ is calculated using dual frequency phase observations; $f_1 = 1575.42$ MHz is one of the frequencies of GPS signals; P_1 and P_2 are pseudo range observations in L_1, L_2 , respectively; B_i and B^p are the hardware delays difference between the receiver and satellite of pseudo range observations, respectively; φ_1 and φ_2 are the observations of phase measuring in two frequencies, respectively; λ_1, λ_2 and N_1, N_2 are the wavelengths of carrier signals and their

Fig. 1 Distribution of crustal movement observation network in Sichuan

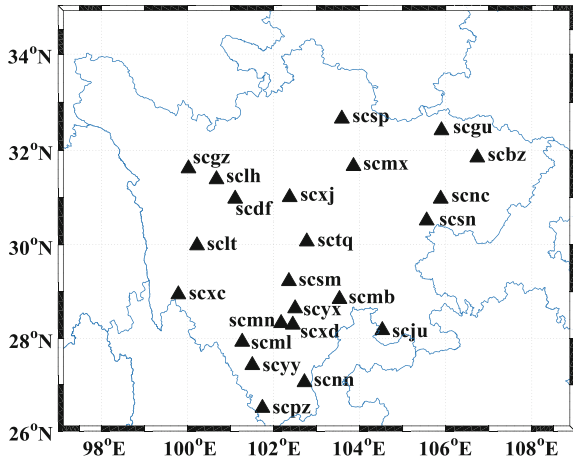


Table 1 Coordinates of crustal movement observation network in Sichuan

测站	经度	纬度	测站	经度	纬度
scbz	106.74	31.84	scnn	102.72	27.06
scdf	101.12	30.98	scpz	101.74	26.50
segu	105.90	32.42	scsm	102.35	29.23
segz	100.02	31.61	scsn	105.56	30.51
seju	104.52	28.18	scsp	103.58	32.65
selh	100.67	31.39	sctq	102.77	30.07
selt	100.22	29.99	scxc	99.80	28.94
scmb	103.53	28.84	scxd	102.44	28.30
scml	101.28	27.93	scxj	102.37	31.00
scmn	102.17	28.33	scyx	102.51	28.65
scmx	103.85	31.67	scyy	101.51	27.43
scnc	105.88	30.98			

corresponding starting integer ambiguities of the two frequencies, respectively; b_i and b^p are the hardware delays difference between the receiver and satellite of phase observations, respectively; γ is a conversion factor related to frequency, and

$$\gamma = \left(\frac{f_1}{f_2}\right)^2 = \left(\frac{1575.42}{1227.60}\right)^2 \approx 1.647.$$

Because the accuracy of TEC_R is poor, and TEC_φ is a relative amount, therefore, two types of TEC results can be complemented, which is mean that using the phase values TEC_φ to smooth the pseudo range observations TEC_R , so that a high accuracy TEC is obtained.

In the n th epoch, the difference of TEC between two types of observations are calculated:

$$\Delta TEC_n = TEC_{R,n} - TEC_{\varphi,n} \quad (2)$$

ΔTEC_n can be obtained in each epoch, and when the integer ambiguity is a constant, it should be a constant, in theory. An average of ΔTEC_n obtained from each epoch yields higher precise ΔTEC_N :

$$\begin{aligned} \Delta TEC_N &= \frac{1}{N} \sum_{n=1}^N \Delta TEC_n \\ &= \frac{1}{N} (TEC_{R,n} - TEC_{\varphi,n}) \frac{1}{N} [(N-1)\Delta TEC_{N-1} + (TEC_{R,N} - TEC_{\varphi,N})] \end{aligned} \quad (3)$$

Adding ΔTEC_N to TEC_φ yields a high precise and absolute TEC, which can be calculated as follows:

$$TEC_{SM,N} = TEC_{\varphi,N} + \Delta TEC_N \tag{4}$$

High precise TEC can be obtain using the pseudo-range and phase observations after calculated the satellite and receiver hardware delays [15].

It is a good way to detect the ionospheric anomalies by moving an average from the TEC sequence. The basic idea of this method is to detect the ionospheric variation of the current epoch by using the TEC average of 20 epochs before current epoch. After removing the normal change TEC from the current epoch, the detrend TEC (dTEC) is obtained and the abnormal part will be detected accurately. Compared with the second-order difference de-trend method, the moving average can detect slow changes of TEC anomalies.

3 Experiment Analysis

Using the method proposed in Sect. 2, dTEC time sequence can be obtained. The dTEC time sequence of 23 stations was drawn, as shown in Fig. 2, and it can be found that there were significant disturbances between 14 and 16 h. Compared with the steady state (fluctuated around 0), the dTEC of most stations fluctuated obviously during the period of 14–16 h with a maximum of 0.3TECU. The abnormal

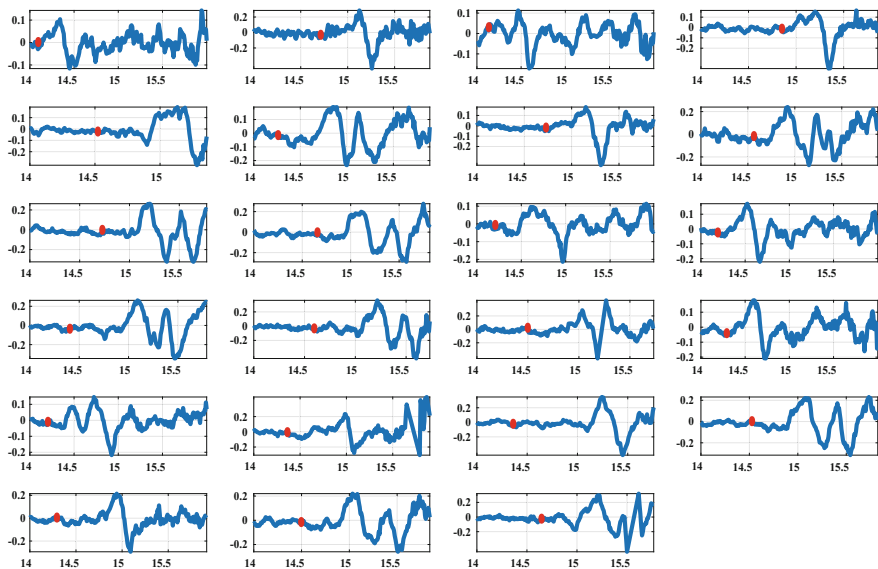


Fig. 2 dTEC time sequence of PRN10

time of each station was different because the location of the Ionospheric Pierce Point (IPP) is different, and abnormalities spreads in the ionosphere, which results in differences in the timing of abnormalities starting.

As solar activity and geomagnetic activity are also responsible for the abnormal changes in the ionosphere, the effects of these two activities can be assessed using the three indexes of Kp, Dst and F10.7. An time series figure before and after the heavy rainfall was made (DOY182–202), as shown in Fig. 3. As can be seen from Fig. 3, before the rainfall, at DOY182, the $F10.7 > 135$, $Dst < -70nT$, and $Kp > 4$, which means that day the solar activity and geomagnetic activity is more intense, and other times are relatively calm. At DOI190, $F10.7 < 130$, and Dst index is at a low level, which indicates that solar activity has a smaller effect. And $Kp < 3$, which indicates that the impact of geomagnetic activities is small. This shows that we can ignore the ionospheric anomalies caused by solar activities and geomagnetic activities, and analyze the relationship between rainfall and ionospheric anomalies, then study on the mechanism in ionosphere-troposphere coupling.

4 Discussion on the Triggering Mechanism of Ionosphere Anomalous

This paper argues that the main reason of the ionospheric anomalies is the AGWs caused by various atmospheric processes in the troposphere. Due to the influence of the terrain and wind, the waves reach the ionosphere and cause ionospheric disturbances. Based on this point of view, the trigger area of AGWs can be roughly found.

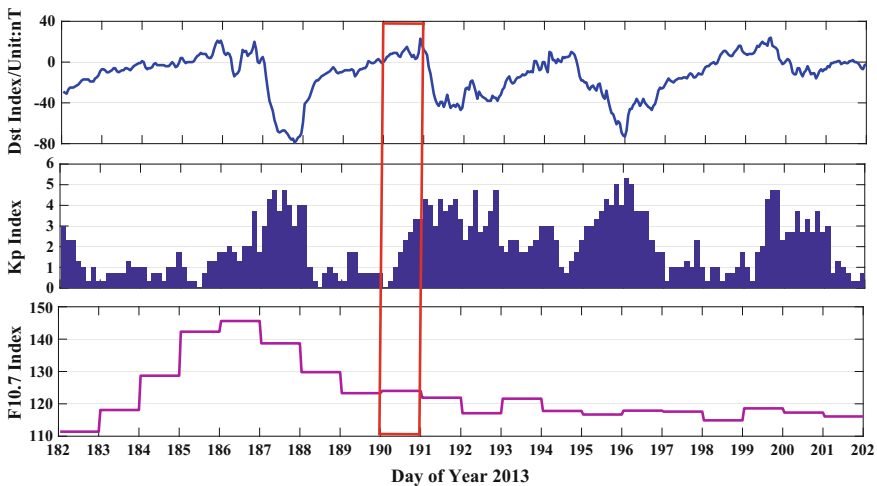


Fig. 3 Kp, Dst, and F10.7 indexes covering from DOY 182 to DOY202

According to the dTEC sequence, the start time of disturbances of each station can be judged roughly, as indicated by the red dots in Fig. 2. Then, using the distance between any two IPPs to divide by the time difference between the two stations, and the horizontal velocity of the AGWs propagating in the ionosphere can be roughly obtained. As shown in the following formula:

According to the two IPPs latitude and longitude (B1, L1), (B2, L2), the distance between two points can be calculated:

$$D = 6,371,004 * \text{Acos}((\sin(B1) * \sin(B2) + \cos(B1) * \cos(B2) * \cos(L2 - L1))) \quad (5)$$

where, 6,371,004 is the radius of the earth, and the unit is m.

The time difference t can be obtained from Fig. 2, and the propagation velocity of AGWs in the ionosphere can be calculated:

$$V = D/t \quad (6)$$

Since the direction of propagation of the AGWs are arbitrary, any one of the two stations can determine a velocity. 23 stations has the total of $C_{23}^2 = 253$ velocities, and take the average value as the propagation velocity of AGWs in the ionosphere. The horizontal propagation velocity is 271.63 m/s.

Sichuan Province (east longitude 97° – 109° , north latitude 26° – 35°) was divided into $1^\circ * 1^\circ$ grid, and the trigger time and trigger point was searched. Only take the PRN10 star as an example, the search results can be seen in Fig. 4, and the contour value is the time residual. It can be seen that at the western boundary of Sichuan Province, the absolute value of the residual value is the smallest (around 0), which indicated that the approximate location of the trigger source is at the boundary of Sichuan Province.

The residuals obtained from only one star are concentric circles, and when the results of all the stars intersect at this time, a more accurate trigger source can be obtained. This research will be done in future work.

Sichuan Province acrosses the Qinghai-Tibet Plateau, Hengduan Mountains, Yunnan-Guizhou Plateau, Qinba Mountains, Sichuan Basin and other major geomorphological units, and the topography of it shows west high while east low.

The horizontal wind field information was obtained from the ECMWF, and a wind map at 12:00 on July 9 was made, as shown in Fig. 5. It can be seen that at 12 o'clock the wind blow to the boundary of Sichuan Province where the wind speed was small, which is due to the blocking effect of the mountain, then the horizontal wind changed to vertical wind, and carried AGWs upward until to the ionosphere and caused the ionospheric anomalies.

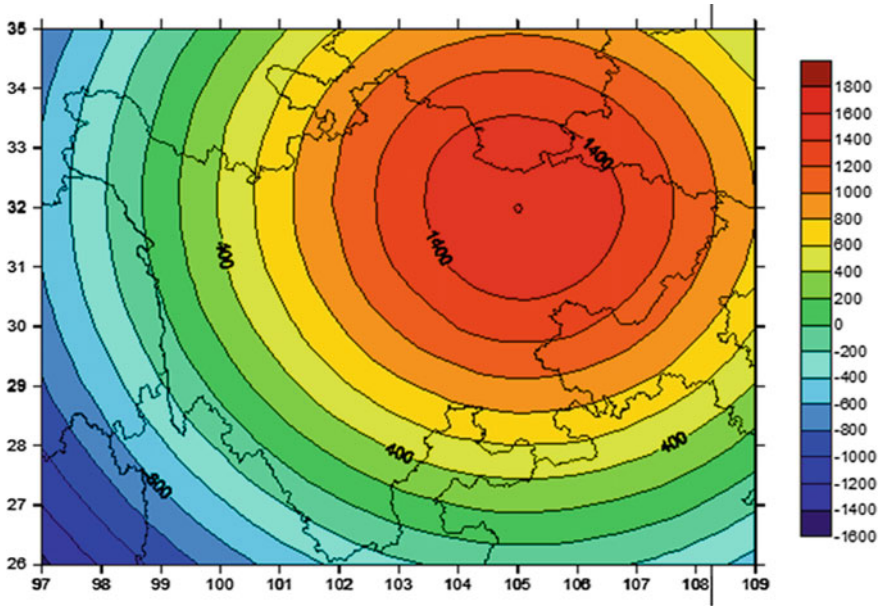


Fig. 4 Time residual distribution in searching process of PRN10 (unit: s)

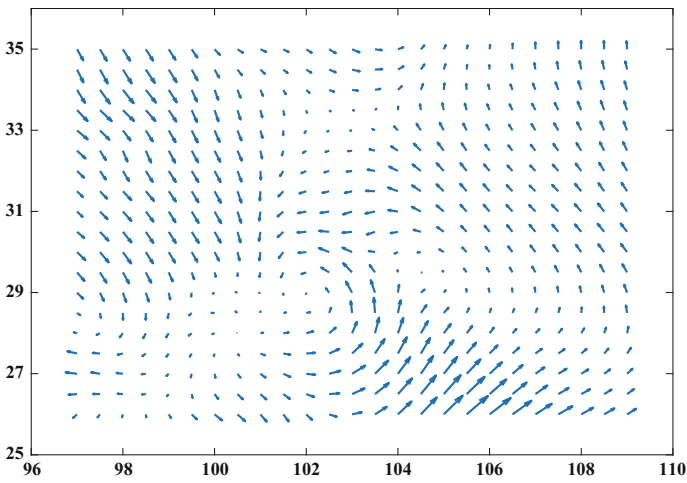


Fig. 5 The wind map at 12:00 of 9, July

5 Conclusion and Outlook

Waves with a period longer than Brunt-Vaisala are called gravity waves and can be stimulated by many activities, such as meteorological activities, aurora activities, earthquakes, volcanic eruptions and others [16]. Ionospheric disturbances often occur when a typhoon approaches or lands, as the rapid dissipation of momentum and viscous interactions at this time, which can excite multiple frequencies of AGWs [17]. These AGWs can reach to higher atmosphere due to the relatively large wavelength.

In this paper, the disturbance caused by a heavy rainfall in Sichuan Province is discussed. It can be clearly observed by dTEC sequence has an obvious abnormality before the rainfall than steady state. The ionospheric disturbance was caused by AGWs which excited by atmospheric activity in the lower atmosphere, and it would propagate with wind. Terrain is also an important factor affecting the propagating of AGWs. At the southwestern boundary of Sichuan Province, there are the Qinghai-Tibet Plateau and the Hengduan Mountains. When the air flow passes through the mountainous regions, the lower air pressure will form strong convective weather, and at this moment, it is easy to excite the AGWs on the leeward surface of the mountain. At the same time, due to the ups and downs of the mountains and the wind, the AGWs will propagate to the ionospheric height and cause disturbance to the electron density. According to the linear propagation theory of sound waves, the horizontal propagation velocity can be roughly calculated as 271 m/s, and the trigger source for AGWs can be roughly found at the southwestern boundary of Sichuan Province, which is in good agreement with the terrain.

A preliminary discussion of the coupling mechanism between the ionosphere and the troposphere is that TID is caused by the propagation of AGWs. Due to the traction of mountains and changes in atmospheric factors, AGWs are excited somewhere and propagate to the upper atmosphere. However, the sound waves are not free to propagate in the atmosphere due to the wavelength dissipation, the absorption of the critical layer, and the dissipation in the stratosphere [18].

In future work, to look for the exact trigger source location, the search for the trigger source and time of the other stars will be done. Other heavy rainfall events happened in other regions will be analyzed to further explore the coupling mechanism between the ionosphere and the troposphere.

References

1. Cang Z, Cheng G, Cheng G (2015) Ionospheric anomalous disturbance during the tropospheric strong convective weather. *J Atmos Solar-Terr Phys* 129:55–61
2. Lastovicka J (2009) Lower ionosphere response to external forcing: a brief review. *Adv Space Res* 43:1–14
3. Kazimirovsky ES, Herraiz M, De la Morena BA (2003) Effects on the ionosphere due to phenomena below it. *Surv Geophys* 24:139–184

4. Rishbeth H (2006) F-region links with the lower atmosphere. *J Atmos Sol-Terr Phys* 68: 469–478
5. Lastovicka J (2006) Forcing of the ionosphere by waves from below. *J Atmos Sol-Terr Phys* 68:479–497
6. Bauer SJ (1957) A possible troposphere-ionosphere relationship. *J Geophys Res* 62:425–430
7. Shen C (1982) The correlations between the typhoon and the f_0F_2 of ionosphere. *Chin J Space Sci* 2(4):335–340
8. Mao T, Wang J, Yang G et al (2010) Effects of typhoon Matsa on ionospheric TEC. *Chin Sci Bull* 55:712–717
9. Liu J et al (2006) Ionospheric GPS total electron content (TEC) disturbances triggered by the 26 December 2004 Indian Ocean Tsunami. *J Geophys Res* 111:A05303
10. Artru J et al (2005) Ionospheric detection of gravity waves induced by tsunamis. *Geophys J Int* 160(3):840–848
11. Tsutsui M, Ogawa T (1973) HF Doppler observation of ionospheric effects due to typhoons. *Rep Ionos Space Res Jpn* 27(3):121–123
12. Xiao S, Hao Y, Zhang D, Xiao Z (2006) A case study on whole response processes of the ionosphere to typhoons. *Chinese J Geophys* 49(3):623–628 (in Chinese)
13. Kong J, Yao Y, Xu Y et al (2017) A clear link connecting the troposphere and ionosphere: ionosphere response to the 2015 Typhoon Dujuan. *J. Geod.* 91:1087–1097
14. Liu Z, Gao Y, Skone S (2005) A study of smoothed TEC precision inferred from GPS measurements. *Earth Planets Space* 57(11):999–1007
15. Chen P (2012) Research on GNSS-based ionospheric tomography and pre-earthquake ionospheric anomaly. Wuhan University
16. Sindelarova T, Buresova D, Chum J, Hruska F (2009) Doppler observations of infrasonic waves of meteorological origin at ionospheric heights. *Adv Space Res* 43:1644–1651
17. Xiao Z, Xiao S, Hao Y, Zhang D (2007) Morphological features of ionospheric response to typhoon. *J Geophys Res Sp Phys* 112(A4)
18. Vadas SL, Fritts DC (2009) Reconstruction of the gravity wave field from convective plumes via ray tracing. *Ann Geophys* 27(147):177

Water Level Changes Measurement of Reservoir Using Multi-GNSS Reflectometry



Kai Liu, Cuilin Kuang, Wei Zhang, YaXin Wen and Chunlong Gong

Abstract Water level change is the crucial disastrous factor of reservoir deformation. This paper derives the water level changes of reservoir based on GNSS-R technology by using the multi-GNSS deformation monitoring observation data of Xikeng dam in Shenzhen, and compares the results with the artificial observation data. The numerical results show that the root mean square error (RMS) of GPS-R results is 7 cm with the accuracy of BDS-R and GLONASS-R is slightly lower than GPS-R, and the range of RMS is 8–10 cm. However, the changes are consistent with the artificial observation data, and the correlation coefficients between the experimental results of three system and the artificial observation data are better than 0.9. This research verifies the validity and reliability of GNSS-R technology used to water level changes measurement, and expands the application areas of the GNSS deformation monitoring system, which also indicates that multi-GNSS system can greatly improve the temporal resolution of single-system water level measurement.

Keywords GNSS-R · BDS, Signal-to-noise ratio · Water level changes measurement

1 Introduction

China's number of reservoirs is the largest in the world. The water level change is the crucial factor that causes the deformation of reservoir. Therefore, it is imperative to measure and monitor the water level of reservoirs effectively and timely.

K. Liu (✉) · C. Kuang · Y. Wen
School of Geosciences and Info-Physics, Central South University, Changsha, China
e-mail: liukainew@163.com

W. Zhang
GNSS Research Center, Wuhan University, Wuhan, China

C. Gong
Shenzhen Water Planning and Design Institute Co., Ltd., Shenzhen, China

However, as most reservoirs are located in the suburbs, it is inconvenient and inefficient for manpower to measure the water level changes. In recent years, a variety of sensors, such as ultrasonic detector, differential pressure level sensor and float level sensor are widely used in monitoring the water level change of reservoir [1]. But these sensors are defective, which require high installation accuracy and the system error is larger and so on.

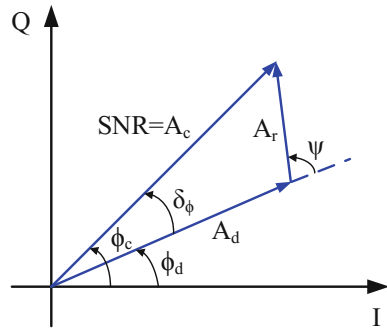
With the development and application of GNSS, scholars have found that satellite reflected signals can be used to monitor the environment changes in the earth [2, 3]. GNSS receiver placed on the coast can receive the direct satellite signals and the reflected signals (also called multipath signals) reflected off the sea surface with the interference of the multipath signals recorded in the signal-to-noise ratio (SNR) data. Bilich et al. [4] conducted a research on the relationship between multipath environment and SNR, laying a solid foundation for SNR applied in monitoring earth's surface environment. The concept of using the freely available GNSS signals for environmental sensing was first introduced by Martin-Neira [5] in 1993, initiating the field of GNSS Reflectometry (GNSS-R). Larson et al. [6] derived sea level time series by using the data of Onsala Space Observatory (OSO) and compared the results with local tide gauge, the root mean square is 4.5 cm. Zhang et al. [7] used GPS-R technology to monitor the sea level of Friday Harbor in the United States with the root mean square is less than 10 cm, and the correlation coefficient is better than 0.98.

This article first extends the application of GNSS-R to reservoir water level observation, which can make up the defects of sensors. The advantage of GNSS-R technology is that the receivers originally installed for deformation monitoring can receive a large number of satellite signals from GPS, GLONASS, BDS and GELILEO. Therefore, multiple satellite signals are simultaneously received by one observation station for observing water level changes of the reservoir, which can improve the temporal resolution of the results. At the same time, these observation data can not only be used to monitor the deformation of reservoirs, but also observe the changes of water level in reservoirs, contributing to the utilization of these data is greatly improved. Thus, GNSS-R technology will play an increasingly important role in observing reservoir water level changes.

2 The Basic Theory of Measuring Water Level Change Based on GNSS SNR

The general GNSS receivers are designed to amplify direct signals and suppress reflected signals, but some reflected signals still can be received and affect navigation accuracy, known as multipath effects. The signals received by GNSS receiver antenna are the composite signals including direct and reflected signals. SNR is equivalent to the composite signals amplitude and it can be gotten from the

Fig. 1 Phasor diagram describing carrier tracking loop of GNSS receiver



carrier tracking loop of GNSS receiver. The carrier tracking loop is described as the relationship between in-phase (I) and quadrature (Q) channels.

The relationship between in-phase (I) and quadrature (Q) channels is described in Fig. 1. On the condition of multipath-free, the loop will only track the amplitude A_d and phase ϕ_d of direct signals. When the reflected signals (amplitude A_r and phase ψ) are received, it will cause the phase deviation (δ_ϕ), and the receiver records the composite signals (A_c) and its phase ϕ_c instead. According to the above figure, we can get that:

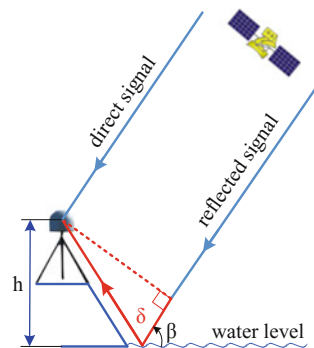
$$SNR^2 \equiv A_c^2 = A_d^2 + A_r^2 + 2A_dA_r \cos \psi \tag{1}$$

Generally, we can remove the trend by the low-order polynomials, and the residual sequence of reflected signals amplitude can be expressed as:

$$\delta(SNR) \equiv A_r = A \cos(\psi + \varphi) \tag{2}$$

The principle of water surface measurement by GNSS-R was described in Fig. 2. The reflected signal is longer than the direct signal by an additional distance as the path delay δ . Where h is for the distance from receiver antenna phase center to water surface, β is the satellite elevation. From the Fig. 2, we can get:

Fig. 2 Principle of water surface measurement by GNSS-R



$$\delta = 2h \sin \beta \quad (3)$$

The phase delay between direct signal with the reflected signal is ψ , so:

$$\psi = \frac{2\pi}{\lambda} \delta = \frac{4\pi h}{\lambda} \sin \beta \quad (4)$$

Due to the multipath change over time, the phase delay of multipath is a function of time, and the carrier phase and the error of pseudorange multipath is sine or cosine function of the phase delay. They all have the same cyclical or has a certain frequency. So taking the time-derivative of ψ :

$$f_{\psi \Rightarrow t} = \frac{d\psi}{dt} = \frac{4\pi \hat{h}}{\lambda} \sin \beta + \frac{4\pi h}{\lambda} \cos(\beta) \hat{\beta} \quad (5)$$

\hat{h} ($= \frac{dh}{dt}$) defines the vertical velocity and $\hat{\beta}$ ($= \frac{d\beta}{dt}$) defines the elevation angle velocity. Assuming the sea level is static ($\hat{h} \approx 0$) during the satellite arc, the frequency ($f_{\psi \Rightarrow t}$) of the multipath oscillation is a constant and directly proportional to the height (h).

If set $x = \sin \beta$ and $\hat{h} = 0$ (assume the h is constant), then

$$f_{\psi \Rightarrow x} = \frac{d\psi}{dx} = \frac{4\pi}{\lambda} h = 2\pi \frac{2h}{\lambda} \quad (6)$$

According to formula (2) and (6), we can get:

$$f_{\delta(SNR) \Rightarrow x} = \frac{d\delta(SNR)}{dx} = \frac{f_{\psi \Rightarrow x}}{2\pi} = \frac{2h}{\lambda} \quad (7)$$

We can derive the dominant multipath frequency ($f_{\psi \Rightarrow t}$) by spectral analysis $\delta(SNR)$. Because the SNR data are evenly sampled in time, but $\sin \beta$ is not, and the Fast Fourier Transformation (FFT) cannot solve the problem of unequal sampling interval. So we analyze the $\delta(SNR)$ data by Lomb-Scargle (LS) spectrum analysis method usually, since it can handle unevenly spaced samples. The vertical reflection distance from the sea surface can be gotten by:

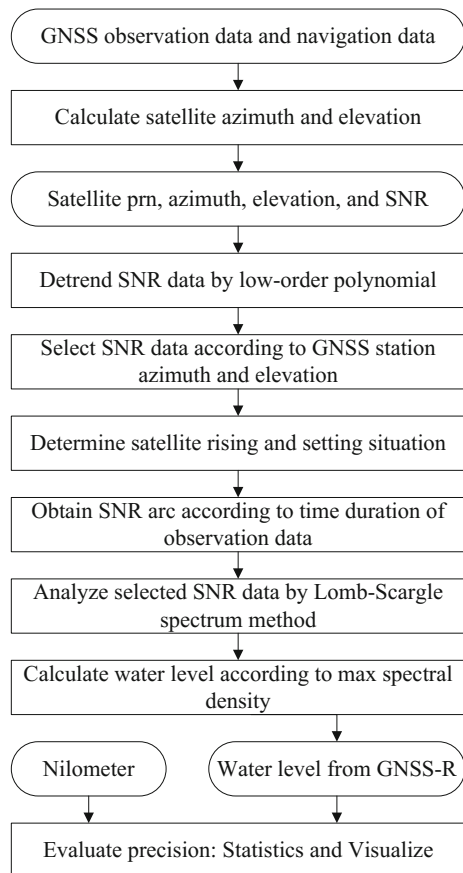
$$h = \frac{\lambda f_{\delta(SNR) \Rightarrow x}}{2} \quad (8)$$

3 Analysis Methods

The data processing flow of deriving the water level changes as shown in Fig. 3. According to the station coordinates and Navigation Message, the azimuth and elevation of satellite are calculated firstly, and the SNR data are extracted from the observations file, then the available data need to be selected from the extracted SNR data according to the reflection environment of the station. Since the station is located in the east of the water area, the azimuth parameter is set as 200° – 340° , and the multipath effect is more significant at low satellite elevation, so the elevation parameter is set as 5° – 25° . However, it is important to realize the selected available SNR data may be observed respectively when the same satellite rising and setting. In this case, the SNR data should be divided into two satellite arcs.

Taking the data of doy 11 in 2016 as an example, Fig. 4a shows the SNR of the MEO satellite with a PRN of 14 at azimuth of 305° . It is hard to see the influence of the water surface reflection directly from the SNR, because of the direct signal in the

Fig. 3 Flowchart of water level calculation by SNR based BDS-R



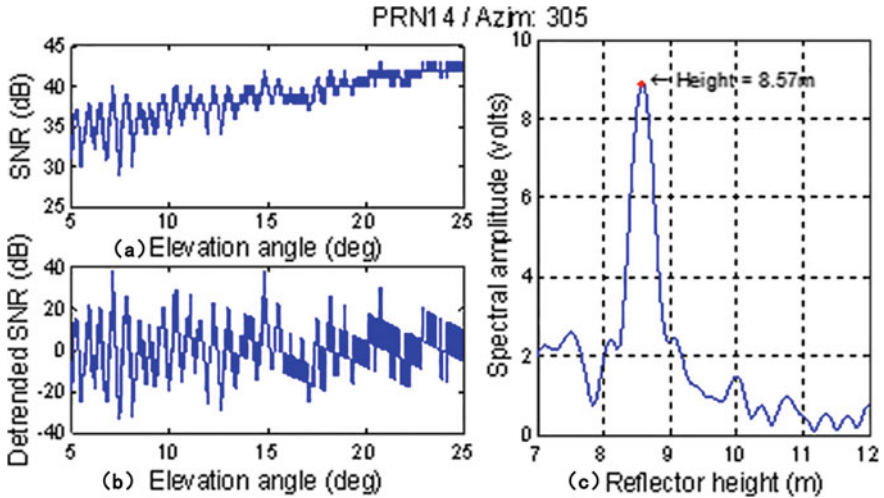


Fig. 4 SNR and Lomb-Scargle periodograms

composite signal received by the receiver is dominant. Therefore, the second-order polynomial is used to remove the trend term of the signal. As shown in Fig. 4b, the remaining sequence of SNR residuals reflects the multipath effect of satellite signals due to surface reflections. In order to obtain the frequency of the SNR residual sequence, spectral analysis can be performed using the Lomb-scargle spectral method. The required vertical reflection distance (h) can be obtained from the maximum spectral density, as shown in Fig. 4c. Due to the results from LS spectrum analysis exist abnormal value, so the water level change time series can be obtained from the reflection height after eliminate outliers. The accuracy of GNSS-R was Assessed through the residuals, RMS, and correlation coefficient by comparing with the data from manual observation.

4 Data Analysis

4.1 Data Source

Shenzhen Xikeng Reservoir is located in the upper stream of Xikeng River with the area of 4.98 km^2 and the normal water level is 71.0 m. In order to monitor the deformation of Xikeng dam, the all-weather GNSS deformation monitor system of Xikeng Reservoir is set up on the dam, and the system, which consists of 12 GNSS monitoring points and 1 datum point (Fig. 5). Data from BDS, GPS, Galileo and GLONASS are collected through multi-constellation GNSS receivers, among which XK03 station is located in the middle of Xikeng dam, facing the reservoir in larger azimuth, and the reflected signals received by the GNSS receiver mainly

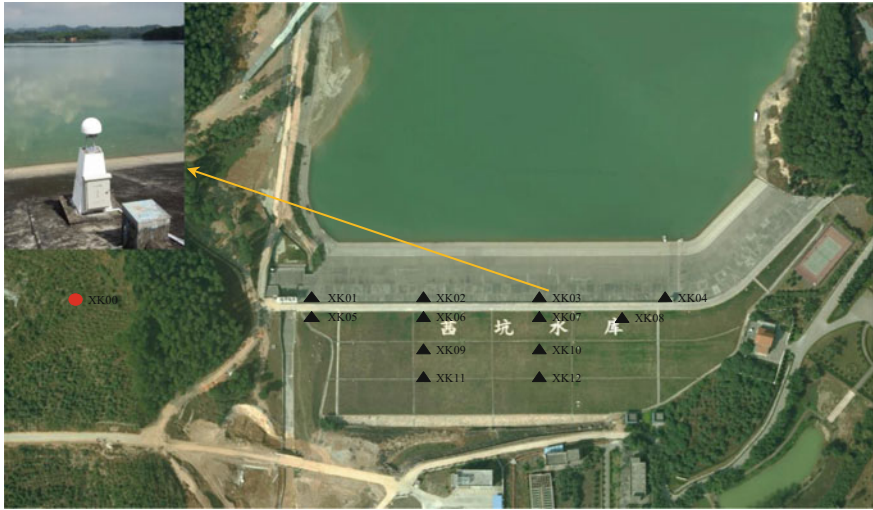


Fig. 5 GNSS deformation monitoring system of Xikeng dam

reflect off the water surface, while it is less affected by other multipaths such as surrounding vegetation. So, in this article, the change of reservoir water level is derived on the basis of GNSS-R technique with the SNR data of XK03 station. Before the research was carried out, the observation of the water level in the reservoir was conducted manually at 8 o'clock every morning. This method was inefficient with low temporal resolution. However, the reservoir is located in the mountainous areas and there are usually no large waves, so the artificial observation of the water level data is relatively accurate, and this article uses artificial observation as reference data to evaluate the GNSS-R experimental results.

4.2 Experimental Results

According to the data processing procedures given above, the water level change was derived based on the L1 band SNR data of 2016 from GPS, BDS and GLONASS of XK03 station. The inversion results of water level data were compared with the data from manual observation. Since the water level derived by GNSS-R is relative and there are systematic deviations in inversion results with manual observation data, we systematically shifted the GNSS-R results prior to the comparison. There is a missing part of the GNSS data because we maintained and updated the monitoring system in May, 2016 and other dates, so the corresponding GNSS-R results were not obtained.

Figure 6 shows a comparison of the GPS-R results with the manually observed water level, where the red line indicates the manually observed water level with a

Fig. 6 Water level changes time series from GPS-R and manual measurement

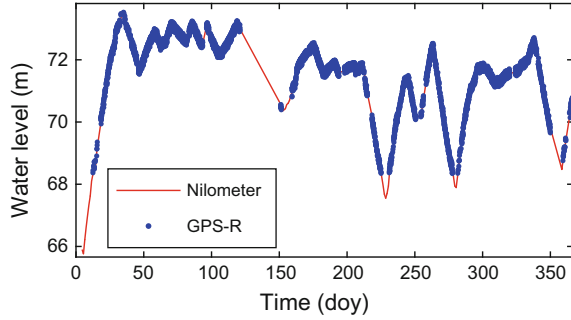
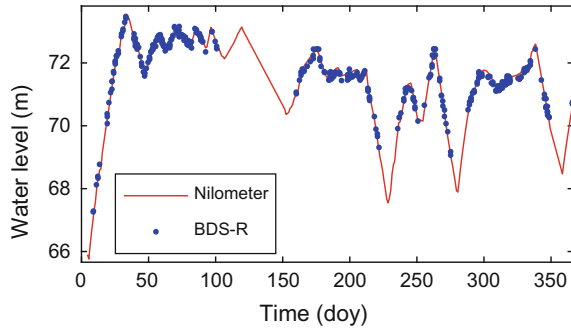


Fig. 7 Water level changes time series from BDS-R and manual measurement



temporal resolution of 1 day and the water level derived by the GPS-R is indicated by the blue dot. As can be seen from the figure, the results of GNSS-R have higher temporal resolution, and there is a good consistency in water level variation values with manual observation. The root mean square error (RMS) of the difference between two methods is 7.8 cm.

The water level time series inversion by BDS-R is shown in Fig. 7. Since only the MEO satellites can be used to observe the water level changes in the BDS system and only three MEO satellites are in orbit in 2016, the experimental results of BDS-R are worse than that of GPS-R. The RMS of this result is 9 cm, but there still has a good agreement with the manually observation generally. At the same time, Fig. 8 shows the GLONASS-R inversion of water level time series. However, it has lower temporal resolution than GPS-R, because of poor quality of the observed data in some days. The RMS of this result is 8.3 cm, which is better than BDS-R.

It is unable to inverse water level using observation data from single-system effectively and completely. The temporal resolution can be greatly improved by combining multiple GNSS system data (as shown in Fig. 9). Figures 10 and 11 respectively describe the residual statistics and correlations of the three GNSS-R systems experimental results compared with the manually observed water level data.

Fig. 8 Water level changes time series from GLONASS-R and manual measurement

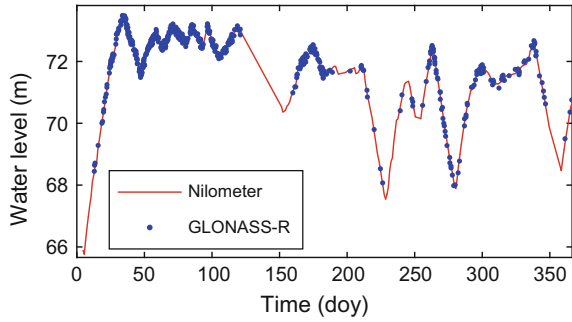
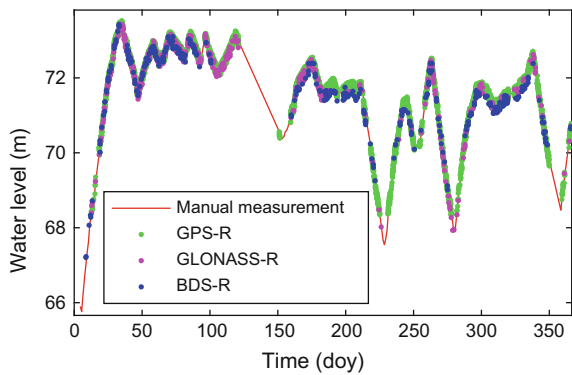


Fig. 9 Water level changes time series from GNSS-R and manual measurement



In Fig. 10, the horizontal axis represents the difference between the time series of water level inversion by GNSS-R compared to the manually observed water level data, and the vertical axis represents the statistics falling in different difference intervals. Among them, the experimental errors of GPS-R and GLONASS-R are within 18 cm, while the experimental errors of BDS-R are slightly larger and some errors are more than 20 cm. It can be concluded from the figure that the statistics are generally obeyed the normal distribution with all the differences are concentrated around zero, and the statistics gradually decrease as the error increases, which means that the GNSS-R technique can effectively invert the water level of the reservoir.

A correlation analysis of the GNSS-R experimental results with the data from nilometer is shown in Fig. 11, where the red solid line is the reference line $y = x$ and the blue dotted line describes the linear fit of the two results. The numerical results show that the water level retrieved by GPS-R is in good agreement with the observed water level, and the least-squares correlation coefficient is as high as 0.95. However, the experimental results of BDS-R and GLONASS-R are slightly lower than those of GPS-R. Correlation coefficients were 0.9 and 0.92, respectively.

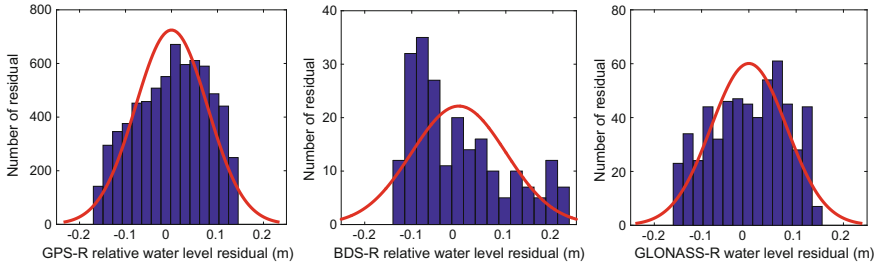


Fig. 10 The residues of water level between GNSS-R and manual measurement

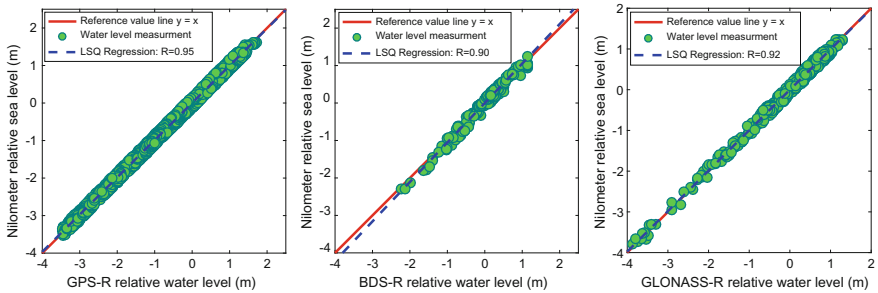


Fig. 11 The correlation analysis of water level between GNSS-R

5 Conclusion

For Xikeng reservoir plays a significant role in water supply and regulation, it is necessary to grasp its water level changes. The manual observation method is time-consuming and labor-intensive and gets results with low temporal resolution. This paper proved the feasibility and reliability of GNSS-R technique used in observing the reservoir water level changes, and extended the application range of the GNSS deformation monitoring data. Firstly, the relationship between SNR and multipath is established in this paper to derive the basic formulas by using GNSS SNR data to derive water level series. Then, the water level time series in 2016 of this reservoir is derived by the self-development software. By comparing with the data from artificial observation, the numerical results show that the root mean square error (RMS) of GPS-R results is 7 cm, the accuracy of BDS-R and GLONASS-R is slightly lower than GPS-R, the RMS is 8–10 cm, and the correlation coefficient between the experimental results and the artificial observation data is better than 0.9. This research verified the feasibility and reliability of GNSS-R used in observing the water level of reservoir, which also shows that the temporal resolution can be greatly improved based on multi-GNSS data. At present, due to

the small number of MEO satellites in Beidou navigation system, there are relatively few data available in BDS-R. However, with the development of BDS, there will be more MEO satellites available for this technology.

Acknowledgements The financial supports from National Natural Science Foundation of China (No. 41774040) are greatly appreciated.

References

1. Leng Z (2013) Research on FBG sensor used for real-time monitoring of reservoir water level. Wuhan University of Technology
2. Bilich A, Larson KM (2007) Mapping the GPS multipath environment using the signal-to-noise ratio (SNR), *Radio Sci* 42:RS6003
3. Larson KM, Ray RD, Nievinski FG et al (2013) The accidental tide gauge: a GPS reflection case study from Kachemak Bay, Alaska. *IEEE Geosci Remote Sens Lett* 10(5):1200–1204
4. Bilich AL, Axelrad P, Larson KM (2007) Scientific utility of the signal-to-noise ratio (SNR) reported by geodetic GPS receivers. In: Proceedings of the 20th international technical meeting of the satellite division of the Institute of Navigation ION GNSS 2007, Fort Worth, Texas, pp 1999–2010
5. Martin-Neira M (1993) A passive reflectometry and interferometry system (PARIS): application to ocean altimetry. *ESA J* 17:331–355
6. Larson KM, Small EE, Gutmann ED et al (2008) Use of GPS receivers as a soil moisture network for water cycle studies. *Geophys Res Lett* 35(24):L24405
7. Zhang S, Nan Y, Li Z, Zhang Q, Dai K, Zhao Y (2016) Analysis of tide variation monitored by GNSS-MR. *Acta Geodaetica et Cartographica Sinica* 45(9):1042–1049

The Impact of Atmospheric Mass Loading on Precise Positioning of National GNSS Stations



Yangyang Liu, Yamin Dang and Changhui Xu

Abstract In the calculation of high-precision GNSS data, solid tides, tides and atmospheric tides are not negligible factors in the solution process. In the present study, a large number of studies have been conducted on the effects of tidal load. The results show that the tide load The impact can reach the magnitude of cm. Relatively speaking, ATML is weak and there are few studies, but it is still a non-negligible interference factor, especially for the data calculation of national stations For example In this paper, GAMIT/GLOBK is used to solve the data of national stations in some provincial areas in 2017. The differences between ATML correction models and different stations are analyzed. The influence of ATML on the accuracy of GNSS baseline solution and station displacement is compared. The results show that the influence of ATML correction model on the baseline solution of regional GNSS is not related to the station location. ATML is insensitive to the influence of the baseline length, but it has an impact on station positioning and can bring about 1–2 mm, and mainly in the U direction, N, E direction of the small, can be ignored. The results also show that the impact of ATML on national stations in solving data is negligible, as is the case in engineering applications and in general scientific research, but corrections are necessary for solutions with higher precision requirements, however At present, there are few models of atmospheric mass load correction applied to baseline solution, so it is worth further study to determine the optimal correction model and eliminate its influence.

Keywords Country stations · Atmospheric mass loading · Baseline Displacement of stations

Y. Liu (✉) · Y. Dang · C. Xu
Chinese Academy of Surveying and Mapping, Beijing 100830, China
e-mail: liuyang1201@126.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_22

1 Introduction

Since the global positioning system (GPS) running. The project of geoscience had revolutionary development, at the same time, there are so much high precision processing software, and the most representative software is named GAMIT/GLOBK, which developed by The Massachusetts Institute of Technology (MIT) and The Scripps Institution of Oceanography (The SIO), The calculating relative accuracy of long baseline can reach 10⁻⁹ level, and short baseline accuracy is better than 1 mm [1, 2], meanwhile, its code is free of charge.

In study to the station time series, environmental loading factors, including ocean loading, atmospheric loading and hydrological loading, there are previous studies have shown that the influence caused by the ocean tide loading attain to cm level in the offshore area vertical crust deformation [3–5] At the same time, the Atmospheric Mass Loading (ATML) would also make position displacement, it is an important factor to seasonal fluctuations of GPS coordinates, especially in the vertical direction, it could cause anniversary of the movement of mm level [6–8], Atmospheric Mass Loading could make effects on inland plate deformation, especially it is more significant in the vertical direction, we also should make analysis to it [9]. Now, in GAMIT, there is specialized ATML correction model, its great significance for precise positioning in the future, but it's a question whether it is more precision than no-ATML in all kinds of baselines, which need to be discussed. Now the country stations had been put into used, but there is no clear research about the influence caused by ATML model in data calculating. This paper chooses partly country stations, comparison and analysis the coordinate time series difference between whether using the ATML correction model or not. And finally determined the influence of atmospheric mass loading to the country stations date processing.

2 Theoretical Model

The paper [10, 11] had studied the influence of atmospheric pressure on the vertical deformation of the crust, and the simplified formula is following

$$\Delta v = -0.35P - 0.55\bar{P} \quad (1)$$

In this formula, Δv is symbol of the vertical deformation of the crust; P is symbol of the instantaneous pressure at the station; \bar{P} is meaning for the average pressure in the 2000 km range around the station. And the formula that could get \bar{P} is follow

$$\bar{P} = \left[\iint p(x, y) dx dy \right] / \left[\iint dx dy \right] \quad (2)$$

$$P(x, y) = A_0 + A_1x + A_2y + A_3x^2 + A_4xy + A_5y^2 \quad (3)$$

$$R = x^2 + y^2 \quad (4)$$

Among the formula (2) and (3), x , y are the distance between the sampling point to the stations in direction of N, E, A_0 – A_5 are needed, which means 6 pressure data fitting coefficient. The last formula is the second curve used to describe the pressure distribution around the station.

The GAMIT has an especially module for ATML correction, and now, the model named ATMDISP is commonly be used, it has two ways to correct depend on atmospheric non-tidal correction: (1). The error correction model of the earth integral mass center in the observation data processing stage; (2). The error correction model of the solid earth part quality center in the post-processing stage of the ground solid coordinate system [12], and we are used to using the first kind of error correction model. Yet the ATMDISP module can't meet the real-time application requirements, there is a month or so lag [13], at present in application, according to the calculating time, we should chose the correction model table of the year.

3 The Experimental Data and Scheme

According to the quality of daily observed data, but also taking into account the data processing workload, this paper choose partly country stations and the value of doys is range from 90 to 109 representatively, and join IGS station which named BJFS, SHAO, URUM to constraint this processing, site distribution as shown in Fig. 1. This paper uses GAMIT 10.5 to get data solution in process, and in order to explore the impact of atmospheric mass loading better, so we set two kinds of programs in the process:

- (1) adopting the ATML correction model;
- (2) no-ATML.

Configuring a Table file which named sestbl. in both solutions, whether to use the ATML model, setting “Y/N”, other default settings of the same parameters, The main parameters set in Table 1 following.

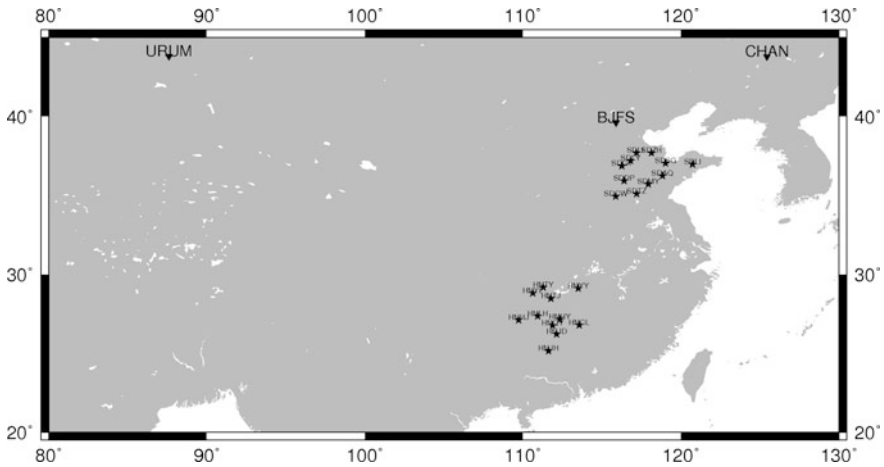


Fig. 1 The selected station distribution

Table 1 The main parameter settings in solving process

Parameter category	Model and parameter
Choice of observation	LC_AUTCLN
Choice of experiment	baseline
Sampling interval	30 s
Ambiguity resolution	fixed
Ocean tide loading	FES2004
Elevation cutoff	10°
Earth gravitational field	EGM
Atmospheric light pressure	BERNE
Atmospheric tide model	atl.grid
Troposphere mapping function model	VMF1

4 Analysis the Solution

4.1 Influence to Baseline Solution

Normalized Root Mean Square (NRMS) is a measure of the quality of the GAMIT solution. NRMS is an important index to measure the GAMIT calculation results, In general, the smaller the NRMS value and the lower the precision, the higher the baseline estimation accuracy; otherwise, the accuracy is lower. The result is then measured and compared by the NRMS value in Q files [14], the calculation formula is as follow:

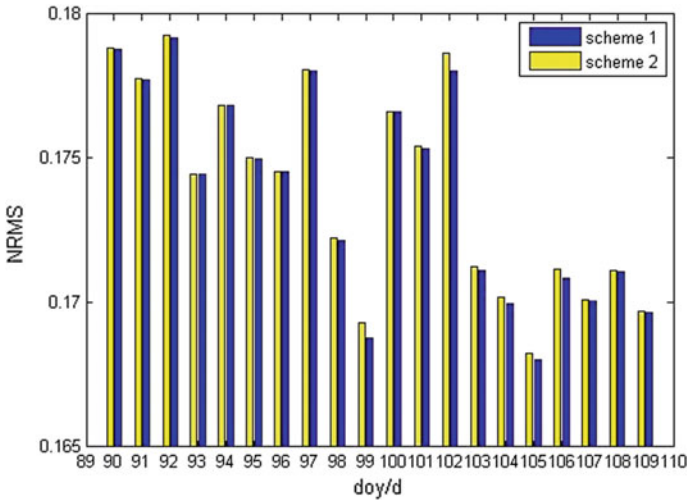


Fig. 2 NRMS of different scheme

$$NRMS = \sqrt{\frac{1}{N} \sum_{i=1}^n \frac{(Y_i - Y)^2}{\delta_i^2}} \tag{5}$$

Figure 2 shows the NRMS values respectively obtained according to two solution strategies, it can be seen from Fig. 1 that the results of the two solutions are different. The NRMS value of single-day solution and the effect of adding ATML correction are better than those without the correction model. However, the NRMS value of single-day solution is discontinuous and unstable, there is a problem with the quality of the data. Therefore, when processing the data of the country stations, we should take the initiative to join the ATML correction model and set the correct model as the year model of ATMDISP.

In order to further analyze the effect of atmospheric load correction model on baseline solution, the baseline length and the calculated baseline Root Mean Square (RMS) value were analyzed to analyze the impact on baseline with different lengths.

As can be seen from Fig. 3, the two solutions to calculate the long baseline RMS changes in the same trend, indicating that as the baseline length increases, the baseline solution RMS value also increases; although the station spacing from 60 to 3000 km, but did not change due to the length of the two programs RMS value of the difference has changed, but showed the same increasing trend, indicating that the impact of atmospheric quality load on the baseline solution is not sensitive to the baseline length; from the map can It is obvious that the RMS values of different schemes are different. The RMS value of adding ATML correction is obviously smaller than the RMS value without model correction, about 14.6%, which shows that the correction model of atmospheric mass load has a good correction.

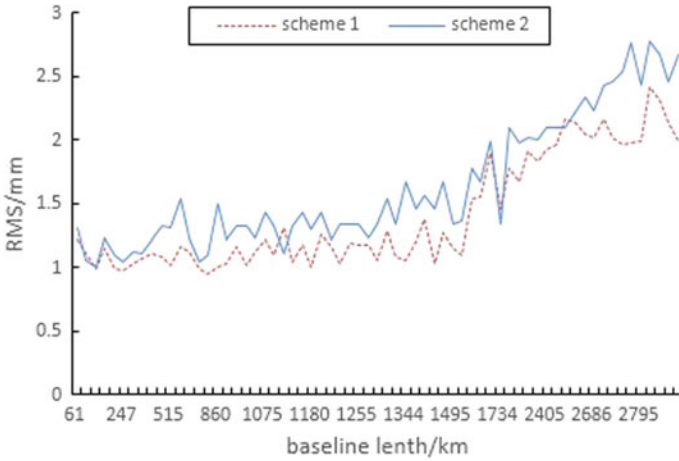


Fig. 3 RMS values of the solution length for different scheme

4.2 Influence to Stations Coordinate

According to FARELL’s loading theory [15], the effect of vertical displacement due to surface loading on station (θ, λ)

$$U(\theta, \lambda, t) = R^2 \int_0^{2\pi} \int_0^{2\pi} G(\psi)H(\psi, A, t) \sin\psi d\psi dA \tag{6}$$

where U is the vertical deformation caused by the load, R is the mean radius of the earth, $G(\psi)$ is the Green’s function of load, which satisfies:

$$G(\psi) = \frac{R}{m_e} \sum_{n=0}^{\infty} h_n p_n(\cos\psi) \tag{7}$$

This is a function of the spherical distance ψ between the load source and the station, m_e is the Earth’s mass, h_n is the load Love number, which can be calculated from a given Earth model, and p_n is a Legendre polynomial. $U(\theta, \lambda, t)$ is the thickness of the (equivalent) load mass at time t , and A is the azimuth from the station to the load source; the magnitude of the load effect can be theoretically calculated from the H provided by the global model.

In order to analyze the correction effect of ATML, the article directly adjusts the baseline solution by using GLOBK software. The coordinate frame of the control station used is ITRF2008, and the reference epoch is March 31, 2017, the coordinate results can be directly analyzed. The article draws a time series of a station, and draws the two schemes as shown in Fig. 4 respectively.

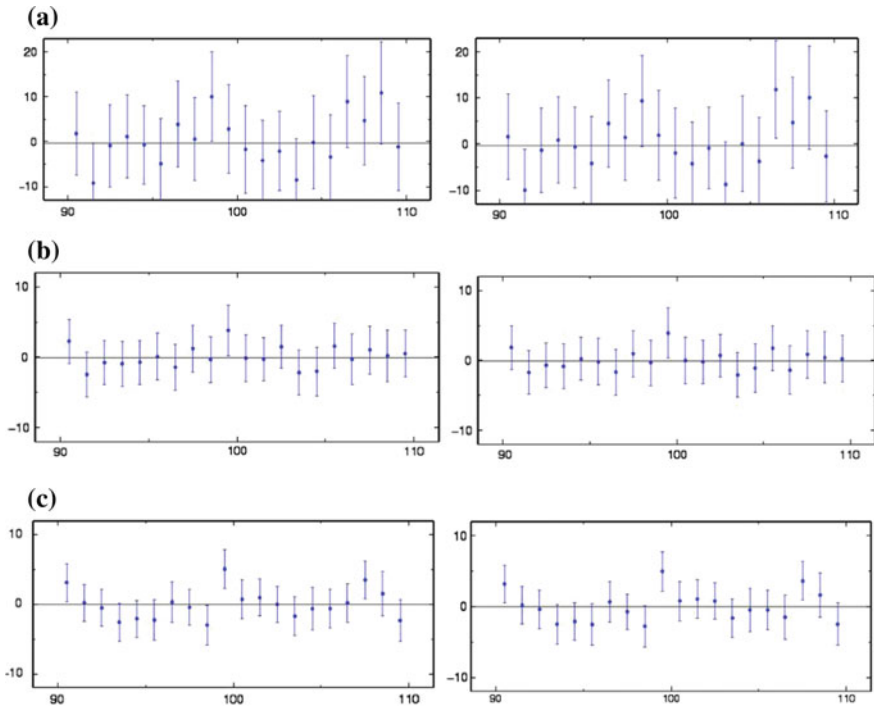


Fig. 4 **a** The time series of adding ATML corrections and with no ATML corrections in the N direction. **b** The time series of adding ATML corrections and with no ATML corrections in the E direction. **c** The time series of adding ATML corrections and with no ATML corrections in the U direction

It can be seen from Fig. 4 that the sequence diagram of scheme 1 is more stable than the sequence diagram of scheme 2, and the separation span is less than that of scheme 2, indicating that the data solution at the country station Counting, adding ATML correction of the coordinate displacement error is less than without adding the correct error.

Further analysis of the impact of different ATML corrections on the coordinates, the coordinates of the different solutions between the stations to make poor, with station serial number as the horizontal axis, the difference between the direction components coordinates of the vertical axis (mm), the statistics shown in Fig. 5.

Figure 5 shows the displacement values in different directions under different schemes, the value of the displacement can be compared with each other alone and found that in the N, E direction of the two programs of displacement are basically the same, the difference between the different programs up to 1 mm, and in the data of more than 1 mm in the 20-day data is only one day, which it may be caused by data or solve. In the U direction, the maximum difference to 2 mm, and more than 1 mm days accounted for 25% of the total number of days, and in the general direction, the difference between the U direction than the N, E direction showed a

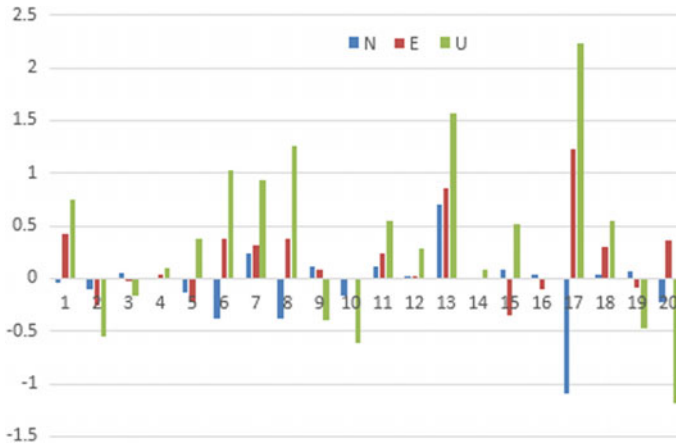


Fig. 5 The coordinate difference between the two programs in the directions

significant trend of the dominant trend, the results show that the atmospheric mass loading can cause the displacement error of mm level in the calculation of the country stations data, and mainly in the U direction. In the overall coordinate shift, adding ATML correction can bring 5% RMS improvement.

5 Conclusion

Based on the previous research results and combining GAMIT/GLOBK with two schemes to solve the problem of country stations in some provincial regions, this paper expounds the atmospheric mass loading on the national benchmark data processing based on baseline solution and station coordinates Influence, come to the following conclusion:

- (1) Atmospheric mass loading correction in the country stations data baseline solution, the baseline length is insensitive to thousands of kilometers for the baseline and tens of kilometers of short-range baseline effect of the same, generally affecting the level of 1 mm, the impact of atmospheric mass loading accounted for 5% in the total baseline solution error.
- (2) The atmospheric mass loading has little effect on the NRMS value of the single-day baseline solution. After adding the ATMDISP model, the single-day solution is better. Therefore, the processing of the country stations data should be added to the current ATMDISP correction model.
- (3) Atmospheric mass loading will have an impact on the station coordinate, but the effect is not obvious. The influence of ATML correction on station coordinate displacements is about 1–2 mm, and mainly affects in the direction of U, and the effect of N, E directions are Negligible.

- (4) For the atmospheric mass loading correction model, the ATMDISP model is mainly used in the GAMIT currently. Further exploration should be to build the ATML correction model in China so that it would be better to solve the station calculation in China.

Acknowledgements This study is funded by National Basic Surveying and Mapping Technology Project of China (2017KJ0205). This work is also supported by the National Science Foundation of China (41474011; 41104018; 41404034).

References

1. Li Z, Zhang X et al (2009) New techniques and precise data processing methods of satellite navigation and positing. Wuhan University Press, Wuhan
2. Wang C, Guo J et al (2011) The methods of batch process of GPS baseline solution and its implementation based on GAMIT software. In: China satellite navigation conference, 2011, pp 71–75
3. Zheng Y, Wu J, Wang J et al (2003) Ocean tidal displacement corrections in GPS precision positioning. *Geomat Inf Sci Wuhan Univ* 28(4):405–408
4. Zhao W, Dang Y, Cheng Y (2013) Impact analysis of ocean tide and its parameters selection of GPS baseline solution. *J Navig Position* 1(2):42–45
5. Schreneck HG (1990) Loading Green's functions for a continental shield with a Q-structure for the mantle and density constrains from the geoid. *Buuld' Inform Marees Terrestres* 108 (7):757–792
6. Dong D et al (2002) Anatomy of apparent seasonal variations from GPS-derived site position time series. *J Geophys Res* B4:2075
7. Zhang FP et al (2002) Seasonal vertical crustal motions in China detected by GPS. *Chin Sci* 47(21):1772–1780
8. Tian Y (2010) Effect of atmospheric mass loading on GPS fiducial stations. *J Geod Geodyn* 30(5):0064–0067
9. Wang M et al (2005) Effects of non-tectonic crustal deformation on continuous GPS position time series and correction to them. *Chin J Geophys* 48(5):1045–1052
10. Liu L (2011) The effect of tide loading on precise positioning in China coastal areas. Chinese Academy of Surveying and Mapping, Beijing, pp 16–17
11. Cheng Y, Hu J (1998) Using GPS to monitoring atmospheric loading model for land and sea vertical movement. *Dev Surv Mapp* 4:40–42
12. Tregoning P, van Dam T (2005) Atmospheric pressure loading corrections applied to GPS data at the observation level. *Geophys Res Lett* 32(22):310–322
13. Jiang W, Zou X (2008) Research and analysis on error of atmospheric models in GPS precise positioning. *Geomat Inf Sci Wuhan Univ* 33(11):1106–1109
14. Cao B, Cheng Y, Xu C (2015) Discussion of CORS data-solving influenced by ocean tide loading model. *Sci Surv Mapp* 40(12):108–111
15. Farrell E (1972) Deformation of the Earth by surface loads. *Rev Geophys Space Phys* 10:761–797

Part II
Navigation and Location Service

A Unidirectional Communication Architecture for Extended Location-Based Services



Yanming Feng and Charles Wang

Abstract Location-Based Service (LBS) is an information service based on locations of mobile users and two-way interactions. Most LBS communications follow the client-server architecture which enables bidirectional communication between the server and a client. The paper examines the limitations of the current client-server LBS framework for time critical LBS application scenarios and their system requirements for networking and computing. It presents the publish/subscribe communication paradigm for unidirectional connections that support decentralised device-to-server and device-to-device data services with low-latency and high scalability. A lightweight publish/subscribe application protocol, namely Message Queue Telemetry Transport (MQTT), is also introduced for high-rate vehicle RTCM data exchanges for timely-critical connected vehicle applications. To prove the concept of the extended LBS, experiments are performed to demonstrate how well the MQTT can support real time location-based data transfer. Results show that the MQTT Broker overwhelmingly outperforms the NTRIP Caster in delivering RTCM correction data in terms of CPU usage of the network computer server and performs similarly well in latency and packet losses. MQTT is also tested for device to device GNSS data exchanges. Results show that two mobile devices can exchange RTCM v3 data with the reliability of 99% within the latency of 1 s under 4G cellular connectivity. The results also show that the MQTT Quality of Services mode 1 can support the 10 Hz data transfer at the latency of less than 100 ms and the reliability of above 99%, showing the potential in support of connected vehicle safety applications.

Keywords Location-based services · Publish/subscribe · Connected vehicles Device-to-device · Real time data distributions · Internet of things

Y. Feng (✉) · C. Wang
School of Electrical Engineering and Computer Science,
Queensland University of Technology, Brisbane, QLD, Australia
e-mail: y.feng@qut.edu.au

C. Wang
e-mail: cc.wang@qut.edu.au

1 Introduction

Location-based services (LBS) is additionally defined as an information service which is accessible with mobile devices through mobile network and utilizes the ability to make use of the geographical position of the mobile device [1]. Today, user's ability to generate contents is core to the services or applications. In the work [2], LBS is defined as a service where

- *The user can determine their location.*
- *The information provided is spatially related to the user's location.*
- *The user is offered two-way interactions with the location information or content.*

Over the last ten more years, various mobile location-based services have achieved mainstream acceptance. Consumer LBS applications, such as road and personal navigation, location-based social networks, mobile location-based gaming, entertainment services, have experienced rapid growth. Commercial or business-to-business or business to consumer services, including tracking, monitoring, emergency, marketing, information services and augmented reality, also encounters the strong market growth.

These traditional LBS applications have some common characteristics. Firstly, the most LBS communications follows the client-server architecture. It is point-to-point data connection allowing for tightly coupled two-way interactions. The well-known application protocol, Hypertext Transfer Protocol (HTTP), functions as a request/response protocol in the client-server computing model. To establish the connection between a client and a server, the client begins to send a request to the server, then the client waits the server response. The server sends response back by re-establishing the connection with the client side. In this centralised structure, the communication load will be multiplied by the number of clients or apps. The loads grow with the number of connections quickly. Also because of HTTP overheads, the request/response communication paradigm is not suitable for low-latency and high-timeliness applications. Secondly, the computation is mostly completed at the LBS server end, such as management of users and dynamic entities, rendering map titles, searching the best travel route for navigation, providing real-time traffic condition and emergency information. Time to compute and analyse at the LBS server usually takes seconds to minutes. Thirdly, creating and developing a successful LBS system is very a challenging task for two main reasons [3]: (1) there are many components, players, and factors involved; and (2) creating and maintaining interoperability and cooperation of LBS components requires from service developers using several standards. In addition, marking the mass-market LBS work effectively and universally is also challenging [2].

On the other hand, not all LBS deployments have to follow the same communication mode. In many scenarios, one-way communication from a mobile platform to a LBS server (the application), or from the server to a mobile user is required. Vehicle or asset tracking is a one-way communication example. There are more

applications that are truly location-based but have not been specified in the current LBS deployments. For instance, a mobile platform/device provides data streams to the servers instead of text messages. The computation or decision making may be completed at the mobile platform edge in real time. But their requirements for communication timeliness, latency and user scalability are very high. For instance, vehicle-to-vehicle and vehicle-to-roadside road safety applications are based on the 10 Hz update for absolute or relative states of vehicles. For this, the request/response paradigm introduces overheads of the unneeded messages running in the reverse direction. These challenges cannot be properly addressed without fundamentally changing the widely-deployed client-server LBS communication architecture.

In this work we will review the current LBS architecture based on the request/response communication paradigm and identify new LBS application scenarios and system requirements. We then present a new LBS framework based on the publish/subscribe architecture and extend the scope of the current consumer and business LBSs to the professional class. The publish/subscribe paradigm supports low-latency, high-reliability and high scalability data exchanges. We introduce the Message Queue Telemetry Transport (MQTT) middleware, which is a lightweight publish-subscribe protocol [4]. Experiments are designed to use MQTT to publish and subscribe high-rate vehicle locations and/or RTCM data messages for time critical connected vehicle applications. Experiments aim to show how well MQTT can perform in delivering the RTCM data with massive number of connections. Results will also demonstrate to what degree the MQTT Quality of Services (QoS) can meet the latency requirements for device-to-device (D2D) data exchanges. Overall, the experimental results confirm the benefits and feasibility of the new LBS architecture for time and location critical mobile applications.

2 Related Work on LBS Architectures and Services

A LBS architecture mainly refers to the communication paradigm that connects all the components effectively to offer certain location-based services and applications. This section will review the LBS components, centralised client-server architecture. The limitations of the current LBS framework will be discussed, providing the basis for the new LBS framework in the next section.

2.1 *LBS Components and Services*

A Location-Based Service generally requires several components. A more complete LBS model with six components which include five technological and one human components was proposed in [3], including

- Positioning Systems
- Communication Networks
- Service and Application Provider
- Data and Content Provider
- Mobile Devices
- User (people)

Figure 1 illustrates the diagram of the six component LBS model, showing the two-way client-server interactions that support push, pull, upload and download mechanisms and track and monitor data flows. Push and pull mechanisms are used to deliver information between users and servers. Push mechanism distributes automatically data/information to users based on their location through mobile devices periodically or on a event basis, whilst pull service transfers data/information which is requested from users. For instance, a navigation LBS server may push the turn-by-turn instructions and map titles to the mobile users. A mobile may browse the server to pull travel routes or points of interests (POI). Upload functions allow LBS apps to report user generated contents (UGC), while the download services enable mobile users to access multimedia in near real time. Track services mean to follow the tracks of a mobile device carried by a person, or on an asset or vehicle. An application can then find them or note their location at various points along a course. Location tracking is concerned with the raw historical location data. Monitor services mean to observe and check the progress of something over time, maintain regular surveillance and to listen and report on. Location monitoring has more to do with knowledge such as exceptions, alerts and warnings based on a set of rules, normally bound by space and time [5, 6]. It is noted that LBS applications may reside on mobile terminals. This reference architecture is re-explained as two one-way publish/subscribe data flows between server and mobile devices and between mobile devices for the extended LBS concept as a contribution of this work.

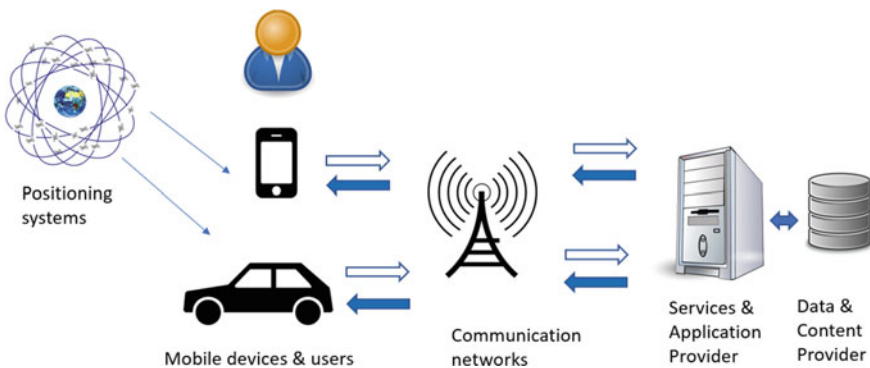


Fig. 1 Diagram of location-based service components and reference architecture

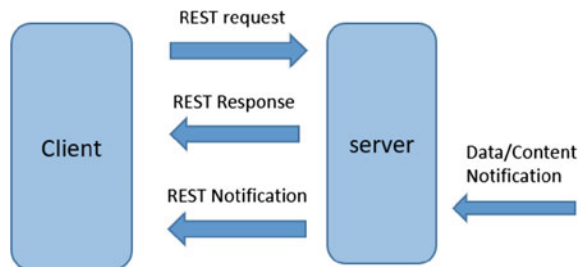
2.2 LBS Request/Response Communication Paradigm and Application Protocols

The client-server LBS architecture for the above services mostly follow the Request/Response paradigm. The Request/Response paradigm enables bidirectional communication between server and client endpoints. The arrows represent the interactive two-way communication: both endpoints have information to send to the other side. In other words, both endpoints can request and respond. The ways of communicating between a LBS server and LBS client is reviewed as follows. In Fig. 1, the application or mobile device is the client. The server is hosted at the computer of the LBS service and application provider, which may include several connected servers in the Cloud. The LBS server functionality is usually accessed through a HTTP Application Programming Interface (API) call. HTTP is an application protocol for distributed, collaborative, and hypermedia information systems. HTTP functions as a request/response protocol in the client-server computing model. The way the HTTP API is structured is called RESTful. Representational state transfer (REST) is a software architecture consisting of clients and servers, as shown in Fig. 2 [2].

The client submits an REST *request* message to the server. The server, which process the request, or performs other functions on behalf of the client, returns a REST *response* message to the client. The response contains complete status information about the request and may also contain requested data and content in its message body. LBS services, push, pull, download and upload are all communicated this way. In the track and monitor services, the server submits an REST *request* message to the mobile client, which return a REST response message to the server, and return with requested data such as vehicle location data streams.

The above client-server communication architecture is tightly coupling by space (where and who), structure (what) and time (when) constraints. It inherently supports one-to-one connection, although the server can push messages to multiple clients. It is a centralised structure and the communication load will be multiplied by the number of clients or apps, thus not being easily scalable. The centralised LBS infrastructure inherently contradicts the preservation of users' privacy and declines the peer-to-peer protocols. Also because of HTTP overheads, the request/response communication paradigm has disadvantages in serving massive number of

Fig. 2 Applications and servers exchange data using RESTful



connections simultaneously. In addition, the deployment of LBSs based on client-server communication architecture is complex as it involves many factors [3]. As a result, the client-server communication architecture significantly limits the potential of LBS in closely related application domains, such as precise positioning correction services, connected and automated vehicles (CAV) and Internet of Things (IoT).

3 Extended LBS Application Scenarios and System Requirements

The existing LBS applications largely rely on access of data sets in the backend, including spatial databases, mapping APIs and POIs. This is so-called data at rest. The two-way interactions between app and server are human-to-human or human-to-machine, thus tolerating the slow server response. For instance, a delay of a few seconds in providing turn-by-turn instructions, or traffic information does not affect much about driving experience. The update of vehicle locations at the rates of 15–30 s in fleet management, taxi and public transport operations may be sufficiently high for managerial services. Business intelligence decision are often made offline or within the time of seconds to minutes. However, the new data that will be generated by location sensors or mobile devices is real time data that fits into a broad category called “data in motion”. This refers to the constant data stream of machine-generated data, instead of traditional UGCs for text messages, images and videos. The traditional processes for UGC capture, distribution, storage, and analysis are no longer able to handle real time data streams. We recognise the opportunities for use of LBS in several time critical applications which have all the location-based, or location-aware characteristics, but have not been articulated in the existing LBSs. These applications require device to device or machine to machine interactions and are classified into the professional category. The target applications include, but not to limit to, the following three scenarios:

The first scenario is location-based data streaming from static and dynamic data sensors. One good example is streaming of raw GNSS data from hundreds to thousands of continuously operating reference stations (CORS) deployed locally and regionally. The data streams are sent to local or regional servers at 1 Hz rate, then processed to generate GNSS related products, such as Radio Technical Commission for Maritime Services (RTCM) corrections to support various precise positioning and scientific services. Another example is collecting, addressing and transmitting of data from a massive number of the Internet of Thing (IoT) sensors and wireless sensor networks. Certain IoT use cases require support for rapid mobility of Things, for example, sensors on a speeding vehicle communicating with road-side infrastructure or a passenger with a smart-phone commuting on a train [7]. Due to rapid mobility, network conditions may vary frequently, and the location-based data streaming must be set at higher transfer rate, such as 10 Hz.

The second scenario is location-based data distribution. For instance, the above RTCM correction data in the formats of observation-space representations (OSR) or state-space representation (SSR) should also be distributed to real time kinematic (RTK) or precise point positioning (PPP) users. The real-time data streams from a network of CORS receivers are collected by a central server. The computation center generates RTCM data in real time at a real or virtual reference station at or near the initial user position. The generated RTCM data is then sent to the user through a wireless connection, often using the Networked Transport of RTCM via Internet Protocol (NTRIP). The update data rate is usually 1 Hz. Currently RTK mainly serves high-end users for professional GNSS applications, such as surveying and machine guidance. The number of distributions is limited. However, potentially automated and connected vehicles on the roads and smartphones need precise positioning services with the RTCM-OSR or RTCM-SSR corrections. The number of the users can be massive under this scenario in the future. The data transfer rate can be up to 10 Hz. For multiple GNSS constellations, the data rate can be up to 10 Kbps per station. The timeliness and data rate requirements cannot be easily met by satellite broadcasting channels, but it is possible through the cellular networks.

The third scenario is the device-to-device location-based data exchanges, for instance, vehicle-to-everything (V2X), including vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) and vehicle-to-pedestrian (V2P) location data or safety message exchanges. Vehicles spatially located in the proximity exchange their Basic Safety Message (BSM) at the rate of 1–10 Hz to enable cooperative awareness such as emergency electronic brake light and slow/stop vehicle warnings. The BSM Part 1 messages include vehicle states information such as position, velocity and heading and their accuracy parameters as defined in SAE J2735 standards. It is required that the latency of BSM and other safety messages is less than 100 ms for the stringent cooperative safety applications. BSM Part 2 message can also include GNSS data in a RTCM format, so that the vehicles can compute the relative position states every 10th second. V2V and V2I data exchanges more generally refer to sharing arbitrary vehicle sensing data such as images, videos and radar signals. The rate of data exchanges can range from 100 ms to several seconds, depending on latency requirements of the V2V and V2I applications. Currently V2V and V2I communications refer to dedicated short-range communications (DSRC) based on IEEE 802.11p and IEEE1609 standards. Cellular based V2X are still under its early development stage. However, the deployment of these technologies requires new hardware components and roadside infrastructure. The device-to-device (D2D) based LBS is still an internet-based solution, deploying new applications in the application layers based on the current and future mobile broadband services offered by wireless location area or wireless wide area networks.

Overall, the target applications mainly include real time device-to-server (D2S) data streaming, server-to-device (S2D) data distribution and D2D data exchanges. The most essential requirement is the timeliness of 1–10 Hz, or the latency is within 1 s–100 ms. In addition, the following requirements have been identified.

- Supporting a many-to-many communication
- Supporting one-way or unidirectional communication
- Making more efficient use of network and computing hardware resources
- Supporting high scalability and high reliability.

Referring to these requirements, the LBS concept and services are extended in the next section to at least partially cover the above target application scenarios.

4 Extended LBS Architecture and Services

4.1 *Extended LBS Definition and Service*

To support time critical data distribution and exchanges, the definition of LBS given by Ferraro and Aktihanoglu in [2] should be extended. The extended definition of LBS is as follows:

- *The user is able determine their location **and exchange location-related data instantaneously***
- *The **data shared**, and information provided is spatially related **and timely critical** to the user's location*
- *The user is offered **both two-way interaction and unidirectional connections** with the location **data**, information or contents.*

Having just added several words with respect to the definition given in [2], the new definition extends the scope of LBS technical capabilities dramatically. The new definition brings the precise positioning services, that is GNSS data correction services, into the scope of extended LBSs. In other words, most GNSS professional applications can be uniformly served under the extended LBS framework. The new LBS also enables time critical and location sensitive professional services in the connected and automated vehicles domains and IoT domains. V2X and IoT are the two technological megatrends. The extended LBSs do not necessarily replace the current or future V2X technologies, but offer the supplementary Internet-based solutions for the low-to-medium demanding V2X applications. There are clear intersections among the LBS, V2X and IoT concepts in terms of components and interaction requirements for networking, positioning and computing. In fact, the traditional LBS may be classified into the Internet of People (IOP), while part of V2X can to large extent be described as the Internet of Vehicles (IoV). As such, the extended LBSs supplement and enhance the IoV and IoT solutions, also would have good interoperability with IoT framework and V2X standards. One can readily recognize the new opportunities for merging the LBS into the IoV and IoT domains.

4.2 Extended LBS Architecture Based on the Publish/Subscribe Paradigm

The extended LBSs demands the adoption of the publish/subscribe communication paradigm. Publish/subscribe paradigm enables unidirectional communication from a publisher to one or more subscribers. Figure 3 shows the concept of the publish/subscribe communication paradigm with the LBS publishers and subscribers. The IoV subscribers declare their interest in both RTCM data and vehicle BSM messages. When the RTCM and BSM publishers have new data available from that classes, they push their data or messages to the interested vehicle subscribers. Similarly, a LBS mobile user may subscribe the RTCM data and POI contents and a Real Time Kinematic (RTK) GNSS user only subscribes the RTCM data. This paradigm is optimal for the traditional LBS track and monitor services that require one-way communications to enable high-timeliness data streaming. More generally, the publish/subscribe model is a decentralised data model and is suitable for the extended LBS deployments with loose coupling between communication endpoints. This is because the model can leverage parallelism and multicast capabilities of the underlying transport network. It supports point-to-multi-point and multi-point-to- multi-point communications.

Based on the publish/subscribe paradigm, the LBS reference architecture shown in Fig. 1 can then be re-explained as noted before. In the new architecture, two-way communications are replaced by two one-way data flows. One flow represents “publish” and the other represents “subscribe”. All clients or applications can be either publisher or subscriber or both. Therefore, the extended LBSs adds new S2D, D2S and D2D functions. The extended LBS architecture not only extends the scope of LBS for more time critical professional grade services, but also significantly scale up the communication connections or simultaneous mobile users in the services.

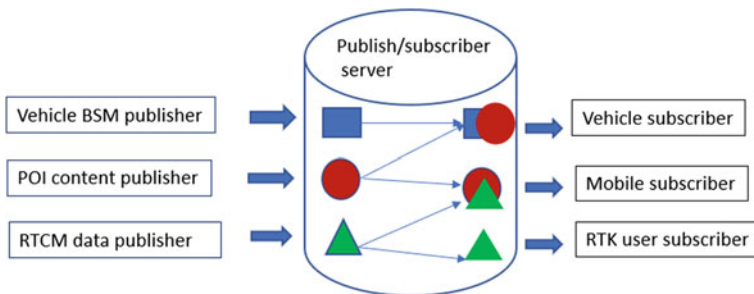


Fig. 3 Diagram of publish/subscribe communication paradigm

4.3 *Publish/Subscribe Application Protocols*

There are several application protocols that can support publish/subscribe communications. Message Queue Telemetry Transport (MQTT) is a light-weight publish/subscribe messaging protocol that was originally developed by IBM for enterprise telemetry and is a proposed OASIS standard [4]. MQTT allows devices to asynchronously and efficiently communicate across constrained networks to remote systems. MQTT is now becoming one of the standard protocols for IoT. Adopting the IoT standard protocol in the proposed LBS extension is one of the potential choices. MQTT follows a client-server architecture where clients connect to a fog or cloud server called broker with IP-addresses. A broker is responsible for registering data sources using the MQTT protocol and assigning unique identifiers for them. Messages or data are published to an address, referred to as a “topic”. Clients subscribe to one or more topics and receive the updates when available. Clients may also publish their data and messages to the broker in the same way. This enables data exchanges between clients through the broker. For instance, in Fig. 3, the vehicle and mobile users exchange their data. In MQTT, topics are hierarchical, and subscriptions may use wildcards. The hierarchical data structure allows the publishers to specify exactly in which topic their published data should reside. Consequently, any subscriber who has subscribed to that specific topic or pool of topics will receive the data when it is received by the broker entity. MQTT offers three levels of Quality of Services (QoS) for data distribution using TCP and UDP as the underlying transport: QoS 0, QoS 1 and QoS 2. They are referred to three message transfer modes respectively: the message is delivered at most once, at least once and exactly once. The default mode is QoS 1. With QoS 1 mode, the sender sends the same message again until an acknowledgment is received. As a result, the receiver can be sent the same message multiple times and may process it multiple times. This transfer mode is more preferable for the proposed LBS applications where the safety messages must be received with the shortest possible latency. There have been several MQTT server products in use. According to the benchmark evaluation about MQTT scalability, JoramMQ and Mosquitto have shown desirable performance in terms of basic measurements (latency, CPU and message rate) [8]. In our evaluation for extended LBS applications in the next section, we choose the widely-used open source Mosquitto [9].

5 **Performance Evaluation of MQTT Protocols for Extended LBS Services**

To prove the concept of unidirectional communications for the time-critical LBS use cases, we perform three experiments to verify the potential of MQTT for D2D data exchanges. The first is to find the CPU usage through settings of multiple connections for base stations and mobile devices in the testing application protocols

installed at both server and client ends. The second is to find the latency and packet loss performance in server-to-device and device-to-device RTCM data distribution or exchanges at 1 Hz. The third is to find the latency performance of D2D data exchanges at higher update rates and at higher QoS levels. The experimental results are presented in the following three subsections.

5.1 Experiment I: CPU Usages in RTCM Data Distribution with MQTT and Ntrip

Networked Transfer of RTCM via Internet Protocol (Ntrip) has been a RTCM recommended standard for real-time dissemination of GNSS observation and correction since 2004. It was developed based on the HTTP/1.1 standard to support hundreds of reference stations and up to thousand users [10]. It has become a vital component for enabling precise positioning (RTK and PPP), CORS network operations and IGS Real Time Pilot Project (RTPP). The purpose of this experiment is to investigate how well MQTT protocol performs in real-time GNSS RTCM data delivery and distributions. MQTT is suitable for low-bandwidth, low-power and requires limited hardware resources and has been part of IoT standards. Therefore, theoretically MQTT is a good choice for GNSS RTCM data distribution, but with unidirectional connections. The setup of the first scenario is shown in Fig. 4, involving four Nectar Virtual Machines (VM) S1 to S4. Each VM has 2.3 GHz RAM. The purpose of the test is to compare the CPU usage of Nectar 2 and 3, where the MQTT broker and NTRIP Caster are hosted, respectively. RTCM v3 data stream from a Geoscience Australia's GNSS base station is about 300 bytes each message and is replicated at Nectar 1 to multiple numbers of reference station RTCM streams, which are then written to Nectar S2 and Nectar 3, respectively as incoming data streams. Nectar 2 and Nectar 3 push the data streams to Nectar 4 as outgoing streams, but based on the requests or subscriptions from a number of connections. Table 1 gives the specifications and settings of Ntrip and MQTT middleware used in the experiments.

Figure 5 plots the CPU usages at Ntrip caster server and as the MQTT broker, against the number of incoming station data streams, where x-axis label (50, 0) denotes 50 incoming data streams and 0 outgoing streams. Figure 6 shows the CPU usages when there are 300 incoming streams and the outgoing streams change from 10 to 1000. In both cases, MQTT requires much lower computer resources to run RTCM data distributions than the use of Ntrip Caster. One Cloud virtual machine can host tens of thousands of RTCM connections if MQTT protocol is used instead of Ntrip caster. In general, MQTT at least is a promising alternative to Ntrip for GNSS data transitions.

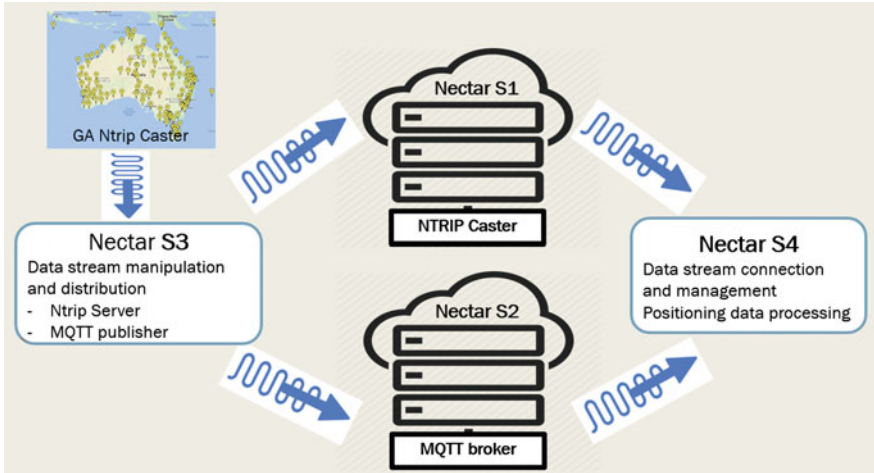


Fig. 4 Diagram of the experiment setup for RTCM data distribution to compare between Ntrip and MQTT protocols

Table 1 Summary of Ntrip and MQTT specifications

	Ntrip	MQTT
Protocol	HTTP 1.1	Publish/subscribe
Middleware	Str2str—NtripCaster and NtripClient	Mosquitto Broker and client
Programing language	Python-based Ntrip client	Paho MQTT python libraries
Underlying transport	TCP/IP	UDP
Quality of service Setting	N/A	QoS 0 (Pub/Sub)

5.2 Experiment II: Latency and Packet Loss Performance in RTCM Data Exchanges with Ntrip and MQTT

This experiment scenario is to examine the latency and packet loss performance when using Ntrip and MQTT with 4G wireless connections. We use the same computer to send to and receive 1 Hz RTCM data from the Nectar cloud virtual machine, so that we can compute the total trip latency. Two Nectar VMs are involved to host Ntrip and MQTT protocols, as we wish to compare the performance. With the 4G wireless connection to the network, the latency results for Ntrip and MQTT are plotted in Fig. 7, comparing the latency readings obtained with Ntrip caster and MQTT broker. The average latency parameters obtained from MQTT and Ntrip tests are 113 and 103 ms, respectively. However, we do observe large fluctuation in the latency from the minimum of ~ 50 ms to several hundreds

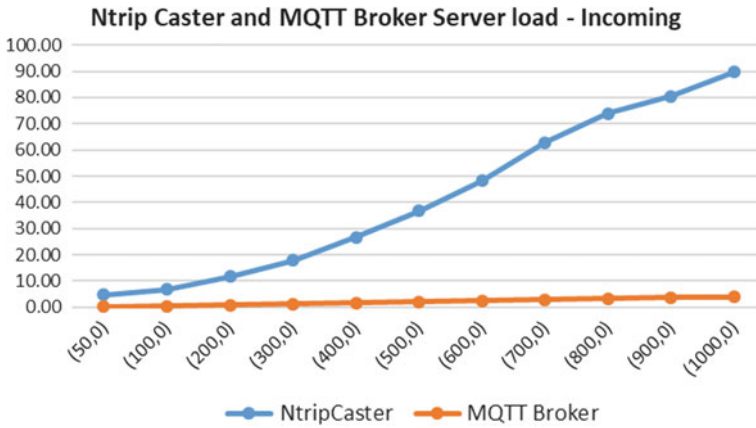


Fig. 5 CUP usage (%) comparison of Ntrip and MQTT for RTCM data transitions against the number of incoming RTCM data streams

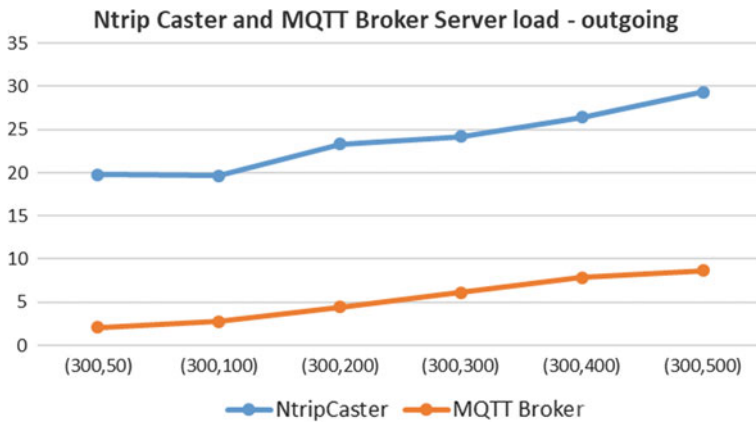


Fig. 6 CUP usage (%) comparison of Ntrip and MQTT for RTCM data transitions against the number of outgoing RTCM data streams

of milliseconds and several seconds in some cases. The packet losses (defined as the latency >2 s) with Ntrip and MQTT are 30 and 21, respectively. Overall, MQTT and Ntrip show similar performance in terms of latency and packet losses under the 4G wireless connections. It is noted that the latency results presented comprised of 4G uplink and downlink directions (i.e. round-trip), the actual latency values could be about half, if the base station has a wired connection to the Internet.

We can further find out to what degree the MQTT protocols can meet the D2D data exchange requirements in terms of transfer rate in Hz or milliseconds. Table 2 outlines the mean latency and reliability of data exchanges versus various transfer latency thresholds. Results demonstrate that the mean latencies are 113 ms when

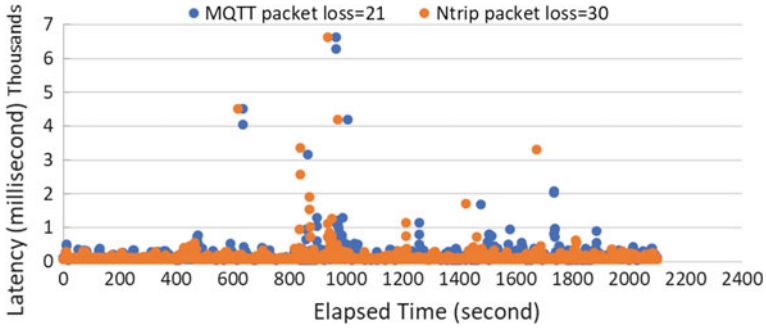


Fig. 7 Illustration of MQTT and Ntrip RTCM data delivery latency with 4G cellular internet connections

meeting the 1 Hz or 1 s latency requirement. The reliability of 99% is achieved for D2D data exchange under the 4G connections in this experiment. For D2D data exchange at the interval of 100 ms, the reliability still reaches 73%. This finding is preliminary, showing the good potential of MQTT for D2D RTCM data exchanges. Referring to the benchmark results by Scalagent [8], lower latency could be offered by alternative MQTT server products.

5.3 Experiment III: Latency Performance of D2D Data Exchanges with QoS Modes 1 and 2

In this experiment, we test the latency performance at the data transfer frequencies of 5 and 10 Hz, with two QoS levels. Table 3 summarizes the latency data and the number of messages that have delayed when the latency exceeds the data transfer interval. The experimental results have also shown that the QoS 1 offers the low latency and high reliability in the 5 and 10 Hz data transfer. The number of delayed messages is defined as the number of the latencies outside the data interval (200 and 100 ms for 5 and 10 Hz, respectively). The results show the good potentials of using MQTT to support V2 V and V2P safety applications. Figure 8 illustrates the

Table 2 Latency and reliability performance of D2D data exchanges with MQTT

Transfer rate (Hz)		Mean latency (ms)	Reliability (%)
10	100	83.67	73.13
5	200	84.75	86.57
2	500	105.40	97.62
1	1000	113.47	99.00
0.5	2000	120.30	99.62
0.2	5000	129.45	99.90

Table 3 Summary of MQTT latency results in 5 and 10 Hz data transfer and two QoS modes

QoS mode	Total record	Size (B)	Frequency (Hz)	Latency (ms)		No. of delayed messages	Reliability (%)
				Maximum	Mean		
1	3000	500	5	187	55	0	100
2	3000	500	5	437	328	3000	0
1	6000	500	10	205	52	8	99.87
2	6000	500	10	500	267	5999	0.02
1	3000	1000	5	1623	59	9	99.7
2	3000	1000	5	593	306	3000	0
1	6000	1000	10	109	54	1	99.98
2	6000	1000	10	452	347	6000	0

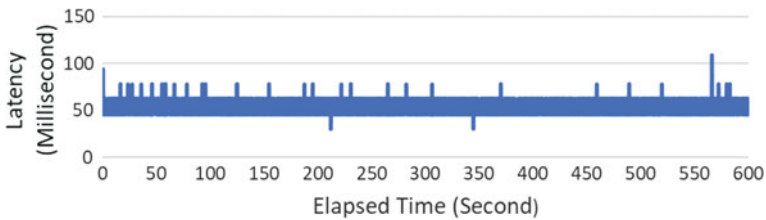


Fig. 8 Illustration of MQTT latency at 10 Hz data transfer and QoS 1 mode with 4G connections

latency over the test period of 10 min at 10 Hz for the message size of 1000 Bytes. It is noted that the QoS 2 option does not meet required latency performance at the 5 and 10 Hz data transfer.

6 Conclusions

The paper has reviewed the existing LBS concept, components and centralised architecture. Most existing Location-based Service (LBS) deployments follow the client-server architecture that enables interactive communication between server and mobile user endpoints. There are more LBS applications that are truly location-based but have not been paid much research and development attentions. We have identified three time-critical professional LBS application scenarios that require one-way communication from mobile devices to a LBS server (the application) or a server to devices. First is location-based data streaming from static and dynamic data sources, such as from the GNSS reference stations, smartphones and IoT sensors. Second is location-based data distributions, such as RTCM correction data and traffic message distributions to road users. Third is the device-to-device location-based high-timeliness data exchanges, such as vehicle-to-vehicle (V2V),

vehicle to infrastructure (V2I) and vehicle to pedestrian (V2P) safety message exchanges. The data sources are spatially widely distributed, and the data is big and is changing fast. The numbers of users can be massive. To implement these new applications, the LBS systems are required to support point-to-multi-point communications and multipoint-to-multipoint connections and device-to-device (D2D) communications. As a result, the LBS has been redefined and the existing architecture has been added with server-to-device and D2D data flow connections. The publish/subscribe communication paradigm for unidirectional connections has been introduced to support decentralised data service with low-latency and high scalability. To prove the extended LBS concept for D2D data exchange, we introduce the Message Queue Telemetry Transport (MQTT) server, which is a lightweight publish-subscribe application protocol widely used in the IoT solutions.

Three types of experiments have been performed. The first is to demonstrate how well the MQTT support GNSS RTCM real time data distributions, which are based on user location updates. Results have shown that the MQTT overwhelmingly outperforms the NTRIP Caster in delivering RTCM data in terms of CPU usage of the network computer server. In other words, with the same hardware and network bandwidth, the MQTT broker can support many time more connections than the Ntrip Caster. MQTT is also tested for V2V data exchanges with cellular connections. Results show two mobile devices can exchange RTCM data at the reliability of 99% with the latency of 1 s under 4G network connectivity. The results demonstrate the benefit of the new LBS architecture for time and location critical professional applications and benefits of publish/subscribe unidirectional communications to extended LBSs for massive number of applications. Further experiments with MQTT have shown that the QoS 1 offers the required performance in data transfer at the 5 and 10 Hz frequencies in terms of latency and reliability, showing the potentials and feasibility of using MQTT to support V2X LBS applications.

References

1. Steiniger S, Neun M, Edwardes A (2006) Foundations of location based services. http://www.e-cartouche.ch/content_reg/cartouche/LBSbasics/en/text/LBSbasics.pdf
2. Ferraro R, Aktihanoglu M (2011) Location-aware applications. Manning Publications, Connecticut, USA
3. Buczkowski A (2012) <http://geoawesomeness.com/knowledge-base/location-based-services/>
4. OASIS (2017) MQTT 3.1.1 specification, 10 Dec 2015. Retrieved 25 April 2017
5. Michael MG, Fusco SJ, Michael K (2008) A research note on ethics in the emerging age of Uberveillance. *Comput Commun* 31(6):1191–1199
6. Michael K, Michael MG (2011) The social and behavioural implications of location-based services. *J Locat Based Serv* 5(3–4):121–137. <https://doi.org/10.1080/17489725.2011.642820>
7. Ammar R, Samer S (2016) Internet of things from hype to reality, the road to digitization, Springer. <https://doi.org/10.1007/978-3-319-44860-2>
8. Scalagent (2015) Benchmark of MQTT servers. http://www.scalagent.com/IMG/pdf/Benchmark_MQTT_servers-v1-1.pdf

9. Light RA (2017) Mosquitto: server and client implementation of the MQTT protocol. *The J Open Source Softw* 2(13) May 2017. <https://doi.org/10.21105/joss.00265>
10. RTCM Paper 234-2004/SC104-PR. <https://igs.bkg.bund.de/rootftp/NTRIP/documentation/NtripPressRelease.pdf>

Virtual Reference Station (VRS) Coordinate's Pattern of QianXun Ground-Based Augmentation System



Xinyu Chen, Chenhao Ouyang and Junbo Shi

Abstract Since May 2016, QianXun Spatial Intelligence Inc. has launched its Beidou ground-based augmentation system (QX-GBAS) which supports high-precision GNSS positioning service within the national coverage. The user first sends its approximate coordinates to QX-GBAS in the NMEA \$GPGGA format and subsequently receives 3D Cartesian coordinates of virtual reference station (VRS) and corresponding GNSS code/phase observations encoded in the format of RTCM. Unlike some operational province- and city-level GBASs whose VRS coordinates received by the user equal to those NMEA \$GPGGA coordinates sent by the user, QX-GBAS does not always return VRS coordinate exactly equal to the user-sent NMEA \$GPGGA coordinates. This study then investigated the pattern of QX-GBAS VRS coordinates in the RTCM message type 1005 and 1006, respectively. Numerical results showed that (1) the QX-GBAS RTCM1006 coordinates were equal to the user-sent NMEA \$GPGGA coordinates within the range of 0.9 m (0.03"); (2) the QX-GBAS RTCM1005 coordinates were different from the user NMEA \$GPGGA coordinates, but followed a $3' \times 3'$ (latitude/longitude) grid pattern. The outcome of this study can be beneficial for both QX-GBAS, or other GBASs, server and users by providing more operational flexibilities and also, protect the homeland security from possible threats.

Keywords QX-GBAS · Network RTK · VRS · NMEA · RTCM
Homeland security

X. Chen · C. Ouyang · J. Shi (✉)
School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei, China
e-mail: jbsshi@sgg.whu.edu.cn

J. Shi
Key Laboratory of Precise Engineering and Industry Surveying of National
Administration of Surveying, Wuhan 430079, Hubei, China

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018
Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_24

1 Introduction

Since May 2016, QianXun Spatial Intelligence Inc. has launched its Beidou ground-based augmentation system (named QX-GBAS hereinafter) within the national coverage. QianXun Spatial Intelligence Inc. takes charge of the Global Navigation Satellite System (GNSS) reference station network construction across the country and high-precision GNSS positioning service provision. On the one hand, QX-GBAS has completed the construction of 1200 satellite navigation reference stations till 2016, becoming the largest ground-based augmentation system in the world [1]. On the other hand, QX-GBAS provides authorized users with three high-precision GNSS positioning services including *Find m*, *Find cm* and *Find mm*. What's more, QX-GBAS concurrently supports three satellite systems including China's BDS, US's GPS and Russia's GLONASS.

As for QX-GBAS high-precision GNSS positioning services, the essence is to provide differential messages via Internet by using the principle of network RTD/RTK positioning. The user first sends its approximate coordinates to QX-GBAS in the National Marine Electronics Association (NMEA) \$GPGGA format and subsequently receives 3D Cartesian coordinates of virtual reference station (VRS) and corresponding GNSS code/phase observations encoded in the format defined by Radio Technical Commission for Maritime Services (RTCM) [2–4].

Unlike some operational province- and city-level GBASs whose VRS coordinates received by the user equal to those NMEA \$GPGGA coordinates sent by the user [3, 5–9], QX-GBAS does not always return VRS coordinates exactly equal to the user-sent NMEA \$GPGGA coordinates [10, 11]. This study then investigated the pattern of QX-GBAS VRS coordinates in the RTCM message type 1005 and 1006, respectively, by first making use of two groups of $1' \times 1'$ (latitude/longitude) grid data and further verified the pattern by using one group of $1^\circ \times 1^\circ$ (latitude/longitude) grid data.

Numerical results showed that (1) the QX-GBAS RTCM1006 coordinates were equal to the user-sent NMEA \$GPGGA coordinates within the range of 0.9 m (0.03"); (2) the QX-GBAS RTCM1005 coordinates were different from the user-sent NMEA \$GPGGA coordinates, but followed a $3' \times 3'$ (latitude/longitude) grid pattern. Such virtual point and grid modes brought more flexibilities for both QX-GBAS server and users. Moreover, through hiding the information related to real reference stations the virtual grid mode can protect the homeland security from possible threats.

2 QX-GBAS High-Precision GNSS Positioning Service

2.1 Flowchart

The flowchart of QX-GBAS high-precision GNSS positioning service is presented in Fig. 1, which contains the following five steps:

- Step 1: The user receives code/phase observations and satellite ephemeris from visible satellites, carries out code-based single point positioning to obtain approximate coordinates.
- Step 2: The user sends its approximate coordinates obtained in Step 1 to QX-GBAS in the format of NMEA \$GPGGA.
- Step 3: After receiving the approximate coordinates in Step 2, QX-GBAS generates 3D Cartesian coordinates of virtual reference station and corresponding GNSS code/phase observations.
- Step 4: QX-GBAS sends the virtual reference station coordinates and corresponding GNSS code/phase observations in Step 3 to users as RTCM message types 1005/1006 and 1074/1084/1124 (GPS/GLONASS/BDS), respectively.
- Step 5: Once getting the code/phase observations and satellite ephemeris in Step 1, VRS coordinates and VRS code/phase observations in Step 4, the user can carry out differential positioning algorithm to obtain high-precision user location.

2.2 RTCM Message Type

Up till 31st August 2017, QX-GBAS provides three mount points including RTCM23 [12], RTCM30 [13] and RTCM32 [2] with specific information presented in Table 1. The required information of using the RTCM32 mount points to conduct high-precision RTD/RTK positioning is shown in Table 2.

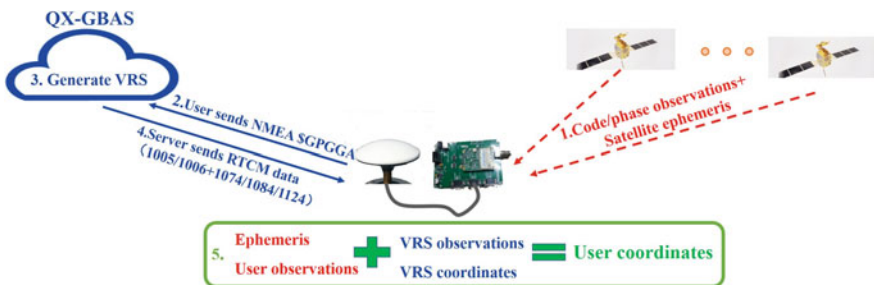


Fig. 1 Flowchart of QX-GBAS high-precision GNSS positioning service

Table 1 Mount points provided by QX-GBAS

Format version	Mount point	Message type (interval in unit of second)	Satellite system
RTCM23	RTCM23_GPS	1(1), 31(1), 41(1), 3(10), 32(30)	GPS
RTCM30	RTCM30_GG	1005(10), 1004-1012(1), 1033(10)	GPS/GLONASS
RTCM32	RTCM32_GGB	1005(10), 1074-1084-1124(1)	GPS/GLONASS/ BDS

Table 2 Required RTCM32 message types for high-precision RTD/RTK positioning

Satellite system	Base station coordinates	Base station observations
GPS	1005/1006	1074
GLONASS		1084
BDS		1124

In the definition of RTCM V3.2, the message type of VRS coordinates is 1005/1006, with structure depicted in Table 3. On the one hand, Message Type 1005 provides the earth-centered, earth-fixed (ECEF) coordinates of the antenna reference point (ARP) for a stationary reference station. On the other hand, Message Type 1006 provides the same information as Message Type 1005, but additionally provides the antenna height above a survey monument. DFXXX is the message number of data field. The argument in parentheses is data type and data length. Data type unit is bit, “unit” represents *unsigned int*; the ellipsis represents that the content of this data field has nothing to do with this study [2, 14].

3 Experiment and Analysis

The user sends its approximate coordinates to QX-GBAS as NMEA \$GPGGA and subsequently receives virtual reference station coordinates returned from QX-GBAS encoded in the format of RTCM. The coordinate message type is RTCM1005/1006 transmitted via three ports (8001, 8002 and 8003) corresponding to various reference coordinate systems. This study concerns the port 8003, i.e. China Geodetic Coordinate System 2000 (CGCS2000). Three experiments are

Table 3 Structure of RTCM message type 1005/1006

Content	Message type	Reference station ID	...	ECEF-X	...	ECEF-Y	...	ECEF-Z	Antenna height (1006)
DF# (size)	002 (unit12)	003 (12 bit)	10 bit	025 (int38)	2 bit	026 (int38)	2 bit	027 (int38)	028 (unit16)

Table 4 Experiment description

Test #	Grid range	Sample interval	# of samples
1	(30°32'00", 114°12'00") →(30°32'50", 114°12'50")	10" × 10"	6
2	(30°28'00", 112°50'00") →(30°28'50", 112°50'50")	10" × 10"	6
3	(30°00'00", 114°00'00") →(31°00'00", 115°00'00")	(3' + random_value < 1") ×(3' + random_value < 1")	2041

designed to investigate the pattern of QX-GBAS VRS coordinates as shown in Table 4.

Test #1.

The test zone is a grid whose bottom left corner is located at (30°32'00", 114°12'00") and the top right corner at (30°32'50", 114°12'50"). In the diagonal of the test zone, choose and send coordinates of six points separated at an interval of 10" in latitude/longitude to QX-GBAS. The specific point distribution is depicted in Fig. 2.

Test #2.

The test zone is a grid whose bottom left corner is located at (30°28'00", 112°50'00") and the top right corner at (30°28'50", 112°50'50"). In the diagonal of the test zone, choose and send coordinates of six points separated at an interval of 10" in latitude/longitude to QX-GBAS. The specific point distribution is depicted in Fig. 3.

Test #3.

The test zone is a grid whose bottom left corner is located at (30°00'00", 114°00'00") and the top right corner at (31°00'00", 115°00'00"). The test zone is divided into 400 grid units at an interval of 3' × 3' (latitude/longitude), as can be seen in Fig. 4 (left). Each grid unit is further divided into four areas at an interval of 1.5' × 1.5' (latitude/longitude) as shown in Fig. 4 (right). A random value within the range of 1" is added to the central point in both latitude and longitude, which

Fig. 2 Simulated user location distribution of Test #1

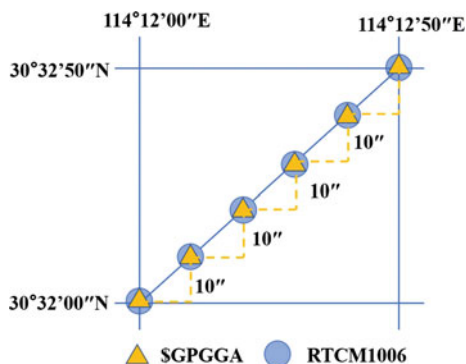
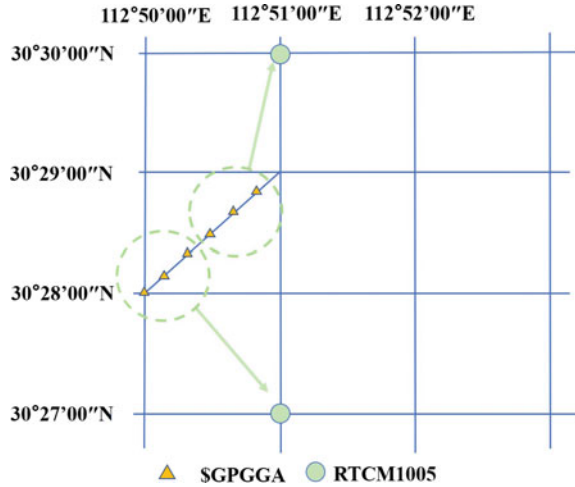


Fig. 3 Simulated user location distribution of Test #2



results in four independent points within each quarter marked as ABCD. In Test #3, there are a total of 2041 points, 441 points on the grid and 1600 points within the grid.

Test #1 As shown in Fig. 2, six simulated user locations represented as yellow triangles were sent to QX-GBAS, and the corresponding VRS coordinates as RTCM1006 were received. A coordinate transformation process was carried out to obtain the VRS coordinates as BLH being consistent with the user-sent NMEA \$GPGGA coordinate form. The results listed in Table 5 told that QX-GBAS RTCM1006 coordinates were equal to the user-sent NMEA \$GPGGA coordinates within the horizontal range of 0.9 m (0.03"), whereas the height difference was within the range of 0.003 m.

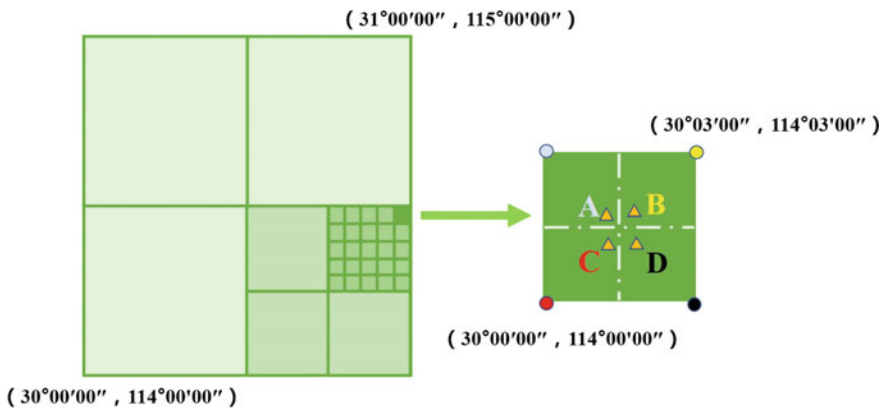


Fig. 4 Simulated user location distribution of Test #3

Table 5 Coordinate analysis of QX-GBAS RTCM1006 (height unit: m)

Sample #	1	2	3	4	5	6
User-sent \$GPGGA coordinates (BLH)	(30°32' 00.00" 114° 12' 00.00" 30)	(30°32' 10.00" 114° 12' 10.00" 30)	(30°32' 20.00" 114° 12' 20.00" 30)	(30°32' 30.00" 114° 12' 30.00" 30)	(30°32' 40.00" 114° 12' 40.00" 30)	(30°32' 50.00" 114° 12' 50.00" 30)
QX-GBAS-returned RTCM1006 coordinates (BLH)	(30°32' 00.00" 114° 11' 59.98" 30.002)	(30°32' 10.00" 114° 12' 9.98" 30.002)	(30°32' 20.00" 114° 12' 19.97" 30.002)	(30°32' 30.00" 114° 12' 29.98" 30.002)	(30°32' 40.00" 114.12' 39.98" 30.002)	(30°32' 50.00" 114° 12' 49.97" 30.002)
Coordinate differences (ΔB, ΔL, ΔH)	0.00" 0.02" -0.002	0.00" 0.02" -0.002	0.00" 0.03" -0.002	0.00" 0.02" -0.002	0.00" 0.02" -0.002	0.00" 0.03" -0.002

Test #2.

As shown in Fig. 3, six simulated user locations represented as yellow triangles were sent to QX-GBAS, and the corresponding VRS coordinates as RTCM1005 were received. A coordinate transformation process was carried out to obtain the VRS coordinates as BLH being consistent with the sent NMEA \$GPGGA coordinates. The results listed in Table 6 told that QX-GBAS RTCM1005 coordinates were not equal to the user NMEA \$GPGGA coordinates in neither horizontal nor vertical directions.

On the one hand, the samples #1, 2, 3 corresponded to the same horizontal coordinate (30°27'00.00", 112°50'59.98") whereas #4, 5, 6 to the other horizontal

Table 6 Coordinate analysis of QX-GBAS RTCM1005 (height unit: m)

Sample #	1	2	3	4	5	6
User-sent \$GPGGA coordinates (BLH)	(30°28' 00.00" 112°50' 00.00" 30)	(30°28' 10.00" 112°50' 10.00" 30)	(30°28' 20.00" 112°50' 20.00" 30)	(30°28' 30.00" 112°50' 30.00" 30)	(30°28' 40.00" 112°50' 40.00" 30)	(30° 28' 50.00" 112° 50' 50.00" 30)
QX-GBAS-returned RTCM1005 coordinates (BLH)	(30°27'00.00", 112°50'59.98", 35.303)			(30°30'00.00", 112°50'59.98", 36.562)		
Coordinate differences (ΔB, ΔL, ΔH)	1' 00.00" -59.98" -5.303	1' 10.00" -49.98" -5.303	1' 20.00" -39.98" -5.303	-1' 30.00" -29.98" -6.562	-1' 40.00" -19.98" -6.562	-1' 50.00" -9.98" -6.562

coordinate (30°30'00.00", 112°50'59.98"). On the other hand, the vertical differences were separated by ~1.2 m. Note that the QX-GBAS-based user coordinates should be further dealt with by the horizontal coordinate transformation and height fitting process, respectively, to derive the local plane and height coordinates. Therefore, the constant height discrepancies (-5.303 and -6.562 m) shown in Table 6 were not further analyzed in this study.

Test #3.

Simulate 2041 points as user locations and send them as \$GPGGA to QX-GBAS, a total of 3466 returned RTCM coordinates were received among which 569 were received as RTCM1006, 37 as RTCM1005, 1430 as both RTCM1005 and RTCM1006.

Firstly, Fig. 5 (left) illustrated the point distribution of 1999 RTCM1006 coordinates where the blue circle represented user-sent \$GPGGA coordinates and the red pentagrams represented returned RTCM1006 coordinates. Figure 5 (right) gave an amplified area. It was quite clear that RTCM1006 coordinates were very close to the user-sent NMEA \$GPGGA coordinates, i.e. with differences no more than 0.9 m (0.03"). Similar to the analysis in Test #2, the vertical coordinate case was not analyzed in this study.

Secondly, Fig. 6 presented the point distribution of 1467 RTCM1005 coordinates where the blue circle represented user-sent \$GPGGA coordinates and the red pentagrams represented returned RTCM1005 coordinates. All 1467 points were located on grid points by a resolution of 3' × 3' in the partition as divided in Fig. 4. These 1467 simulated user locations could be further divided into two categories, i.e. 356 on the grid and 1111 within the grid.

On the one hand, as shown in Fig. 7 (left), blue circles represented 356 user-sent on-the-grid \$GPGGA coordinates and the red pentagrams represented 356 returned RTCM1005 coordinates. Amplify an arbitrary grid unit, as was shown in Fig. 7 (right). It was identified that RTCM1005 coordinates were very close to the

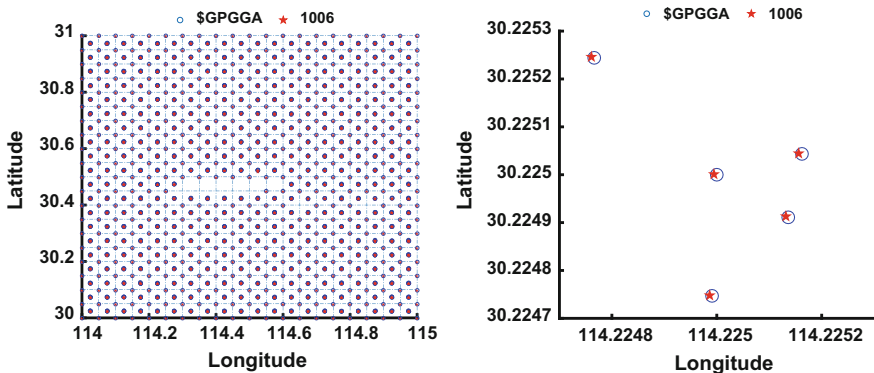


Fig. 5 Point distribution of 1999 (left) and an zoom-in (right) on-the-grid \$GPGGA and RTCM1006 coordinates

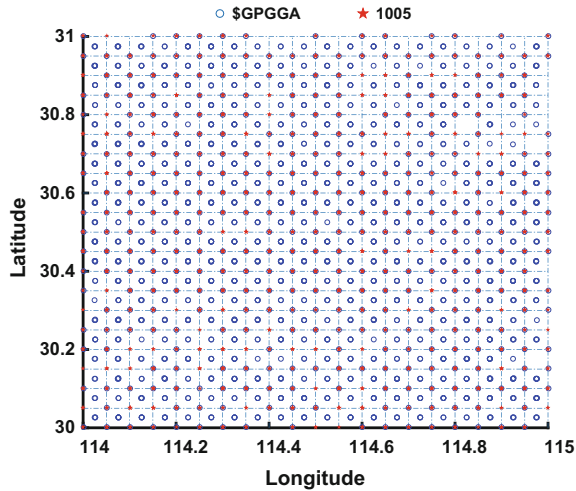


Fig. 6 Point distribution of 1467 RTCM1005 coordinates

user-sent \$GPGGA coordinates by no more than 0.9 m (0.03") in the horizontal direction.

On the other hand, as can be seen from Fig. 8 (left), the blue circle represented 1111 user-sent within-the-grid \$GPGGA coordinates and the red pentagrams represented 1111 returned RTCM1005 coordinates. By a comparison between these two types of coordinates, it could be found that all coordinate differences fell within the range of 1'30" in both latitude and longitude as shown in Fig. 8 (right). Taking both 356 on-the-grid points shown in Fig. 7 and 1111 within-the-grid points in Fig. 8 into consideration, all 1467 returned RTCM1005 coordinates were located at the nearest grid point by a resolution of 3' × 3' as divided in Fig. 4.

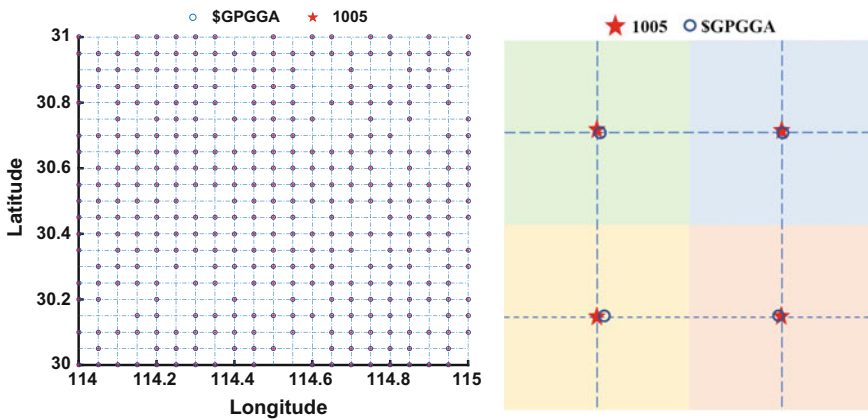


Fig. 7 Point distribution of 356 (left) and an zoom-in (right) on-the-grid \$GPGGA and RTCM1005 coordinates

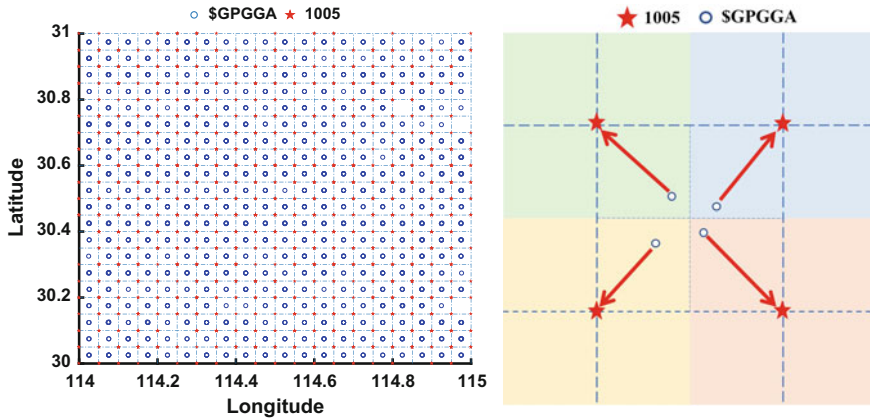


Fig. 8 Point distribution of 1111 (left) and an zoom-in (right) within-the-grid \$GPGGA and RTCM1005 coordinates

4 Conclusion

This paper first presents the flowchart of QX-GBAS based high-precision GNSS positioning. The user first sends its approximate coordinates to QX-GBAS in the NMEA \$GPGGA format and subsequently receives 3D Cartesian coordinates of virtual reference station and corresponding GNSS code/phase observations encoded in the format of RTCM. This study investigated the pattern of QX-GBAS VRS coordinates in the RTCM message type of 1005 and 1006, respectively, by first making use of two groups of $1' \times 1'$ (latitude/longitude) grid data and further verified the pattern by using one group of $1^\circ \times 1^\circ$ (latitude/longitude) grid data.

Some conclusions have been summarized based on three tests:

- (1) The QX-GBAS RTCM1006 coordinates were equal to the user-sent NMEA \$GPGGA coordinates within the range of 0.9 m (0.03");
- (2) The QX-GBAS RTCM1005 coordinates were different from the user-sent NMEA \$GPGGA coordinates, but were located at the nearest grid point by a resolution of $3' \times 3'$ as divided in Fig. 4.

The outcome of this study can be beneficial for both QX-GBAS server and users. First, the grid pattern applied for the virtual reference stations can keep coordinates and GNSS observations of real reference stations confidential, and thus protect the homeland security from possible threats. What's more, the grid pattern could reduce the high computational burden once more users concurrently connect to the server. Second, whenever QX-GBAS broadcasts VRS coordinates in either 1005 (virtual grid) or 1006 (virtual station) mode, there is no need of operational changes for the user's device.

Acknowledgements This work has been supported by CRSRI Open Research Program (Grant #. CKWV2017516/KY), Hubei Nature Science Foundation (Grant #. 2017CFB520) and National Undergraduate Training Program for Innovation and Entrepreneurship (Grant #. 201710486090).

References

1. Zhang F, Hao JM, Chao LI, Wang AB (2007) Algorithm researching of visual reference station based on the RTCM data format. In: *Surveying & mapping of Sichuan, 2007*
2. RTCM Special Committee No. 104 (2013) RTCM standard 10403.2—differential GNSS (global navigation satellite systems) services Version 3 [EB/OL]. 06 Sept 2016. <http://www.rtcn.org/Pub-DGNSS.php2013>
3. Hailing WU, Zuohu LI, Liu H (2014) The overall study on BDS RTCM standardization. In: *GNSS World of China, 2014*
4. Yongshuai H, Junbo S, Chenhao O, Xingning L (2017) Real-time observation decoding and positioning analysis based on Qianxun BeiDou ground based augmentation system. *Bull Surv Mapp* 9:11–14
5. Tang W, Wang J, Jin Y (2004) The design and building of virtual reference station system. In: *Modern surveying & mapping, 2004*
6. Li S (2008) Introduction and decoding arithmetic & realization of RTCM data format of DGPS. In: *Electronic measurement technology, 2008*
7. Xiaoyu S, Pengfei C, Yanhui C, Ziqiao L (2012) DGPS data communication format RTCM 3.1 and its decoding arithmetic realization. In: *Bulletin of surveying and mapping, 2012*, pp 4–6
8. Liu ZM, Lin WJ, Tang WM (2004) Analysis and application on RTCM format for differential GPS. *J Guilin Inst Technol* 24(2):98–102
9. Jingran Z, Shuguo P, Lizhen K (2013) Introduction of RTCM3.1 data format and coding decoding method. In: *China satellite navigation conference 2013*
10. Ling P, Wang Q, Wang HQ (2005) Mode analyses and solution design of real time DGPS data communication. *J Electr Meas Instrum* 19(6):188–191
11. Shou-Zhou GU, Bei JZ, Dang YM (2009) New generation of standard data format: RINEX3.00 and its application. *GNSS World of China* 57:13367–13370
12. Sun LY, Liu C, Hang-Bin WU (2008) The study of RTCM SC-104 decoding. In: *GNSS World of China, 2008*
13. Guo HT, Ren C (2010) DGPS data communication format RTCM3.0 and its application development. In: *Gnss World of China, 2010*
14. Xiaodong YU, Zhiwei L, Wang B, Heli YU, Yan J (2015) The introduction of DGNSS data communication format RTCM3.2 and the decoding research on it. In: *GNSS World of China, 2015*

A Novel Fingerprinting Method of WiFi Indoor Positioning Based on Weibull Signal Model



Zheng Li, Jingbin Liu, Zemin Wang and Ruizhi Chen

Abstract A number of indoor positioning systems based on WiFi fingerprinting were reported thanks to advantages of this method, such as low cost and extensive availability. The Bayesian fingerprinting method needs learn the radio map of probability distribution of WiFi signal strengths over the space of interest through a training phase. Traditionally, the histogram method was used for calculating probability distribution, and it required an adequate number of WiFi samples, which caused a long time taken in the training phase. This study first analyzes the temporal variation of WiFi received signal strength indication (RSSI) at a specific location, and proposes the Weibull signal model for representing the probability density of temporal variation of WiFi RSSI observables. Then, in the positioning phase, the Weibull-based probability density is utilized for Bayesian estimation to resolve the positioning solution. This method is proposed to reduce the required number of RSSI samples for learning probability distribution, and hence improve the efficiency of fingerprinting database training. This method is implemented on Android commodity smartphone, and is evaluated in office building environments. Experiment results show that this method reduces the work loading of fingerprinting

Z. Li · J. Liu (✉) · R. Chen

State Key Laboratory of Information Engineering in Surveying,
Mapping and Remote Sensing, Wuhan University, Wuhan, China
e-mail: jingbin.liu@whu.edu.cn

Z. Li

e-mail: 2016206440005@whu.edu.cn

R. Chen

e-mail: ruizhi.chen@whu.edu.cn

Z. Li · Z. Wang

Chinese Antarctic Center of Surveying and Mapping,
Wuhan University, Wuhan, China
e-mail: zmwang@whu.edu.cn

J. Liu · R. Chen

Collaborative Innovation Center of Geospatial Technology,
Wuhan University, Wuhan, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_25

training due to less samples required, and the positioning accuracy is enhanced by 21–35% up to different building environments, compared to the histogram based method even in which more samples are used.

Keywords WiFi · Indoor positioning · Smartphone positioning
Fingerprint matching · Received signal strength indication (RSSI)
Weibull signal model

1 Introduction

With the increasing demand for location-based services (LBS) especially Indoor Positioning [1], and the rapid development of WiFi communication technology, the WiFi based indoor positioning technology has become a research hotspot [2]. Outdoors people can use the current more mature GNSS and other technologies to locate, but in a complex indoor environment, the positioning accuracy of these technologies is not high enough to meet the needs of indoor positioning [3]. At present, indoor positioning methods mainly include wireless networks, infrared, Bluetooth, ultra-wideband, RFID and pedestrian trajectory estimation PDR, etc. [4]. Among them, the WiFi Based Indoor Positioning technology has strong anti-jamming capability, fast communications, easy access to RSSI measurements and other characteristics, as well as the widespread use of WiFi communication module on smart phone makes wireless indoor location using WiFi into a low cost and easy to implement technology [5, 6]. Therefore, in-depth study on the WiFi based indoor positioning technology has very important scientific significance and engineering value [4].

Domestic and Foreign Scholars have proposed many WiFi based indoor positioning approaches. Such as, Time Difference of Arrival (TDOA), Direction of Arrival (DOA), Phase of Arrival (POA), Time of Arrival (TOA) and so on are based on WiFi signals [7, 8]. However, these approaches require additional hardware and cannot know the true coordinate information of the wireless access point (AP), and there are problems such as poor scalability and long positioning time [9, 10]. In contrast to these approaches, WiFi Based Indoor Positioning Technologies that using RSSI measurements of WiFi only rely on the most common smartphone and wireless network resources available today, are easy to implement and cost-effective. So now the WiFi based indoor positioning technologies is mostly using RSSI measurements. The RSSI based indoor positioning algorithm is divided into triangulation algorithm and fingerprinting algorithm [11]. The principle of triangulation algorithm is to determine the coordinate information of at least three APs first, and then use the signal propagation model to convert the RSSI measurements received by the mobile terminal into the distance from the mobile terminal to the APs. The distances from the mobile terminal to the APs can compute the position of the mobile terminal [12]. However, this algorithm is seriously obstructed by the indoor environment, but also extremely dependent on the signal

propagation model which cannot be defined exactly, so the accuracy and stability are limited and cannot be used widely.

Superior to triangulation algorithm, fingerprinting is characterized by identifying the location of the signal features so that it has less interference with the indoor environment and does not depend on the signal transmission model, does not need to know the AP location [4], etc. Fingerprinting algorithm works in two phases: The first phase is fingerprint training phase [13], which is mainly used to establish the RSSI fingerprint database of the targeted area to be located in advance, but this phase will cost a large amount of computing resources and human resources; The second phase is matching positioning phase [13], the location of the mobile terminal estimated by comparing the real-time RSSI measurements received by the mobile terminal with the RSSI fingerprints in the fingerprint database. Therefore, the positioning accuracy depends largely on the quality of the RSSI fingerprint database [14]. It can be seen that the efficiency and quality of the fingerprint database training are the key to the algorithm.

The algorithm in this paper optimizes these two phases respectively: During fingerprint training phase, the algorithm in this paper proposes the Weibull signal model to approximate the RSSI probability distribution. This model can fit the probability density function very close to the true distribution of RSSI under a limited sample size [15]. Moreover, it only needs to store the parameters of the Weibull signal model when establishing the fingerprint database, and the collection efficiency is also improved under the condition of ensuring the data quality of the fingerprint database, which can greatly save computation and human resources. During matching positioning phase, the algorithm in this paper is improved on Bayesian theorem and Histogram Maximum Likelihood algorithm. According to the RSSI measurements of each AP received by the smart phone in real time, the distribution of each AP is modeled using the Weibull signal model to calculate the probability of occurrence of the received RSSI measurements in real time. Then the maximum likelihood estimation to estimate positioning results. The experiments on smartphones show that the proposed algorithm outperforms the traditional Bayesian position estimation algorithm in accuracy. Because the algorithm in this paper is not integrated with PDR or other localization sources [16], only WiFi fingerprinting, thus the positioning accuracy is not the best. Future fusion with other localization sources such as PDR can obtain better positioning results [16, 17].

2 The Fingerprinting Method Based on RSSI

Location fingerprinting is mainly based on the database of target features to identify, the process of positioning can be divided into fingerprint training and matching positioning [13, 18, 19].

2.1 Fingerprint Training Phase

The first phase is offline fingerprint training phase, the main work is to set a number of reference points in the targeted area, and use the smartphone to collect the received signal characteristic parameter data (such as RSSI measurement, etc.) from multiple APs at each reference point. And these parameters together with the position coordinates of the point as a set of data stored in the database. This set of data is a location fingerprint, which is also referred to as a location fingerprint point [3].

2.2 Matching Positioning Phase

The second phase is online matching positioning phase. Using the smartphone received the same signal parameters stored at the positioning point, and then use the corresponding matching algorithm to determine which fingerprint data in the fingerprint database is the most similar. And then use the most similar set or groups of fingerprint data corresponding to the correlation calculation, which can be used to estimate the actual user location [3].

At present, there are two types of deterministic algorithms and probabilistic algorithms [20]. The deterministic algorithm is relatively simple and the current applications are more extensive, such as the Nearest Neighbor algorithm (NN), K-Nearest Neighbor algorithm (KNN), weighted K-Nearest Neighbor algorithm (WKNN) [21], etc. These algorithms use the Euclidean Metric definition in signal space to understand the minimum distance mapping in RSSI signal space as the minimum distance in physical location. When online positioning, the smartphone receives a set of RSSI measurements and calculates the shortest distance from the RSSI measurements of the same AP in the fingerprint database [22]. The Nearest Neighbor algorithm (NN) outputs the position coordinates of the fingerprint point with the shortest distance solution as a positioning result; K-Nearest Neighbor algorithm (KNN) is finding the nearest neighbor fingerprint point of K ($K > 2$), which is the position coordinates of the smallest K fingerprint points of the shortest distance and obtains the positioning result; Weighted K-Nearest Neighbor algorithm (WKNN) optimization based on K-Nearest Neighbor algorithm, which is to find the nearest neighbor fingerprint point of K ($K > 2$) and multiply the coordinates of each fingerprint point by a weighting coefficient obtains positioning result [21, 22]. Since WiFi signals that are susceptible to interference indoors, the signal space of the RSSI measurement of WiFi is not a one-to-one mapping relationship with the physical location [23], therefore, the accuracy of the deterministic algorithm is limited, while the probabilistic algorithm is better in positioning accuracy. In general, the Bayesian position estimation algorithm is superior to the Weighted K-Nearest Neighbor algorithm (WKNN) [23]. The algorithm proposed in this paper is based on the Bayesian fingerprinting algorithm in probabilistic algorithm.

3 Bayesian Position Estimation Based on Weibull Signal Model

The algorithm proposed in this paper is based on the Bayesian Position Estimation method in probabilistic algorithm and proposes some improvements to the existing problems in the algorithm, which improves the algorithm's efficiency and positioning accuracy. In this paper, we optimize the fingerprint training phase and matching positioning phase of Bayesian probabilistic algorithm.

3.1 Fingerprint Database

In the conventional Bayesian fingerprinting algorithm's fingerprint database, the RSSI probability of all APs received at each fingerprint point are stored [24, 25]. In the conventional Bayesian algorithm, the probability of a RSSI measurement X_n between the reference point R_i and the AP A_m can be expressed as:

$$P(A_m X_n | R_i) = \frac{Y_{x_n}}{N_i} \quad (3.1)$$

where Y_{x_n} is the number of occurrences that the RSSI measurement X_n appeared in the training data set of the i th fingerprint point. Here N_i is the total number of training samples collected at the i th fingerprint point. The entire fingerprint database is expressed as:

$$D = [R_1, R_2, \dots, R_w] \quad (3.2)$$

where W is the total number of the fingerprint points in the target area. In addition, in order to improve the computation process and weaken the RSSI measurements of signal strength singularity, a bin-based solution is adopted. When the conventional algorithm established the fingerprint database to 10 dB for a range of signal strength measurements RSSI by 30–100 divided into 7 ranges, plus 0–30 dB for a single range, a total of 8 ranges, each range can be regarded as a bin. In this way, the probability distribution of each AP is recorded by bin as a unit of the fingerprint database. When positioning, according to the received RSSI measurements to find the corresponding bin acquisition probability of each received AP [15]. In the conventional Bayesian algorithm, at the i th fingerprint point, the probability of the RSSI measurements within the bin B_n for AP A_m can be expressed as:

$$P(A_m B_n | R_i) = \sum_{j > E_{n-1}}^{j \leq E_n} \frac{Y_{x_j}}{N_i} \quad (3.3)$$

where E_{n-1} and E_n are the left and right edges of bin B_n respectively. Y_{x_j} stands for the number of occurrences that the value of the RSSI measurement appeared within the range of $(E_{n-1}, E_n]$. All the RSSI measurements in the bin B_n are cumulated for counting the probability.

3.2 Modelling Fingerprints with the Weibull Signal Model

In this paper, we introduce the Weibull Signal Model to approximate the RSSI probability distribution of all APs received at each finger point. The Weibull function is a traditional method for modelling the signal strength of radio propagation [26]. The probability density function can be expressed as:

$$\text{pdf}(x; \lambda, k, \theta) = \frac{k}{\lambda} \left(\frac{x - \theta}{\lambda}\right)^{k-1} e^{-\left(\frac{x-\theta}{\lambda}\right)^k} \tag{3.4}$$

While the cumulative distribution function can be expressed as:

$$F(x) = 1 - e^{-\left(\frac{x-\theta}{\lambda}\right)^k} \tag{3.5}$$

where X is the variable of the function, K is the shape parameter, λ is the scale parameter, and θ is the shift parameter [15].

The parameters of the Weibull Signal Model can be estimated with a limited number of RSSI sample measurements. The Model parameters (λ, K, θ) can be calculated with [27, 28]:

$$K = \frac{STD}{\ln(2)}, \quad 1.5 \leq K \leq 2.5 \tag{3.6}$$

$$\lambda = \begin{cases} 2 * (K + 0.15) & STD < 2 \\ STD * (K + 0.15) & 2 \leq STD \leq 3.5 \\ 3.5 * (K + 0.15) & STD > 3.5 \end{cases} \tag{3.7}$$

$$\theta = \bar{O} - \lambda * \Gamma\left(1 + \frac{1}{k}\right) \tag{3.8}$$

$$\bar{O} = \frac{1}{n} \sum_{i=0}^n O_i \tag{3.9}$$

$$STD = \sqrt{\frac{1}{n} \sum_{i=0}^n (O_i - \bar{O})^2} \tag{3.10}$$

where \bar{O} is the mean value of the RSSI measurements set O_i , STD the standard deviation. Γ is the gamma function. The term $(K + 0.15)$ is an approximation of the expression $1/\sqrt{\Gamma(1 + \frac{2}{k}) - \Gamma^2(1 + \frac{1}{k})}$ when $1.5 \leq K \leq 2.5$ [15].

Therefore, the distribution probability of each possible RSSI measurement in the fingerprint database can be expressed as:

$$P(x) = F(x + 0.5) - F(x - 0.5) \tag{3.11}$$

Fingerprint database with the Weibull signal model, we can calculate the probability of occurrence of any RSSI measurement. Considering the computational efficiency, the RSSI measurements are still rounded to an integer. The probability for each bin in the fingerprint database can be generated as:

$$P = \int_x^{x+w} f(x)dx = F(x + w) - F(x) \tag{3.12}$$

where w is the width of the bin, x is the RSSI value at the left edge of bin. In theory, the target area can be represented by a set of Weibull signal model that simulate the distribution of RSSI measurements. Each Weibull signal model has three parameters representing the probability distribution of the RSSI measurements between an AP A_m and a smartphone at a fingerprint point R_i . The size of the target area can be reduced in this case because it just requires storing three parameters for each vector between an AP and a fingerprint point.

3.3 Fingerprinting Method with the Weibull Signal Model

Fingerprinting method with the Weibull Signal Model just requires storing three parameters for each vector between an AP and a fingerprint point when creating the fingerprint database, and the data amount can be greatly reduced in this method.

The matching positioning dynamically divides the bin based on the Weibull signal model according to the RSSI measurement of each AP received in real time. Therefore, this paper compares the following three algorithms: The conventional Bayesian fingerprinting algorithm with static bin (Histogram); The fingerprinting algorithm based on Weibull signal model with static bin (Weibull-bin); The fingerprinting algorithm based on Weibull signal model with dynamic calculation probability (Weibull-pdf).

The fingerprinting algorithm in this paper is the use of Bayesian theory and Histogram Maximum Likelihood algorithm [29]. The principle is to use the conditional probability model for location fingerprinting and Bayesian inference mechanism to estimate the position of the smartphone [30], Also called Bayesian probability algorithm, the basic principle can be expressed as:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\sum_{x' \in X} p(y|x')p(x')} \quad (3.13)$$

where x is a fingerprint point in the fingerprint database, y is the RSSI measurements of the AP received by the smartphone at the anchor point, $p(x|y)$ is the probability that the anchor point is the fingerprint point x when the RSSI measurement is y , $p(y|x)$ is the probability that the RSSI measurement is y at the fingerprint point x , $p(x)$ is the probability of a fingerprint point, usually does not consider the difference between the fingerprint points, the default is the fingerprint point equal probability, $p(y)$ is the RSSI measurement occurs the full probability, usually the default AP is mutually independent. From Eq. (3.13), we know that when the value of $p(x|y)$ is the maximum means that the probability of the fingerprint point x occurring when the RSSI value received at the anchor point is y is the maximum. That is the best match with the anchor point can be used as positioning results output. Therefore, the Bayesian probability algorithm is to find the maximum $p(x|y)$, x is the positioning results at this time, the formula can be expressed as:

$$\hat{x} = \arg \max_x p(x|y) \quad (3.14)$$

In order to obtain the maximum value of $p(x|y)$, we know that $p(x)$ and $p(y)$ are the same at each finger point by the Bayesian theory formula. The maximum value of $p(x|y)$ can be transformed into solving maximum value of $p(y|x)$, the meaning of which is the probability of the RSSI measurements of each AP received at the fingerprint point x . Since each AP is independent, it just figuring out the probability product maximum value of the RSSI measurements of each AP, the formula can be expressed as:

$$\hat{x} = \arg \max_x \left(\prod_{j=1}^{n_{MAC}} p(y_j|x) \right) \quad (3.15)$$

where x is a fingerprint point in the fingerprint database; y_j is the RSSI measurements of the j th AP received by the smartphone at the anchor point. Therefore, the conditional probability product of all the APs at each fingerprint point is calculated. The maximum probability is found according to the Histogram Maximum Likelihood algorithm, and the corresponding fingerprint point is the positioning result.

Different from the conventional fingerprinting method, the dynamic calculation probability method takes into account the uncertainty of RSSI measurements and may change within a certain range, is to +5 dB and -5 dB to the RSSI measurements of the AP received in the positioning phase. And then we can dynamically obtain the possible range of RSSI measurements and then calculate the probability, which can more accurately obtain the probability of occurrence of an AP. The

probability of RSSI measurements of all APs received at the anchor point can be expressed as:

$$P(x) = \int_{x-5}^{x+5} f(x)dx = F(x+5) - F(x-5) \tag{3.16}$$

where x is the RSSI value of the AP A_m received by the smartphone at the fingerprint point R_i . The three parameters required by the Weibull signal model have been stored in the fingerprint database to model the probability distribution of the RSSI measurements between an AP A_m and a smartphone at a fingerprint point R_i and then calculates the probability value of the range in real time and dynamically. This is the fingerprinting algorithm based on Weibull signal model with dynamic calculation probability method proposed in this paper.

4 Experiments and Results

In this experiment, the average error and root-mean-square error (RMS) of the position results of the three algorithms are compared respectively in different actual scenarios.

4.1 Experimental Environment and Process

The experiment expressly selected two experiment field, has verified the wide applicability of the proposed algorithm in various environment. The two experiment are conducted respectively on the second floor and the fourth floor of the office building of our laboratory, both of which are 53 m * 73 m in size. The sketch map is shown in Fig. 1

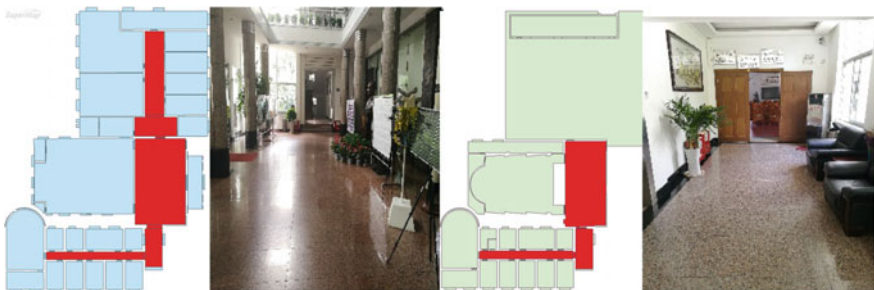


Fig. 1 Second floor and forth floor plan

The red part in the figure is the experimental area that we selected in the public areas such as the lobby and the corridor on the second and fourth floors, 56 and 35 fingerprint points were respectively collected by grid map at the intervals of 2–3 m. The scanning frequency of RSSI measurements of a set of WiFi signals was collected by 2–3 s at each fingerprint point and 30 sets of sample data were collected to establish the fingerprint database. Subsequently, we randomly selected 43 and 20 points in two experimental areas, measured and recorded the true coordinates of these experimental points. At each experimental point collected 5 sets of RSSI measurements with the same sampling frequency at different times. In this way, there are 215 and 100 sets of data for testing in two experimental areas respectively to compare the positioning accuracy of different algorithms (Tables 1 and 2).

4.2 Experimental Results and Analysis

We use three algorithms to calculate the positioning coordinates of all the experimental points, and compared with the true coordinates recorded to calculate the average error and root-mean-square error (RMS) of each algorithm. We also plotted the error distribution graph of three algorithms and the probability Cumulative Distribution Function (CDF) of the positioning error (Figs. 2 and 3).

It can be seen from the positioning error distribution graph and the cumulative probability distribution of error graph that when the collected data of fingerprint database of the conventional Histogram algorithm (red line) and the Weibull-bin algorithm (blue line) are the same, the positioning mean error and RMS of Weibull-bin algorithm has improved. In the same way, the positioning accuracy of the Weibull-pdf algorithm (green line) also significantly better than the Weibull-bin algorithm. As can be seen from the positioning error distribution graph, the green line is obviously lower than the blue line and the red line, that is, the Weibull-pdf algorithm has the best positioning result. By the cumulative probability distribution of error graph analysis, the second and fourth floor of the Weibull-pdf algorithm positioning results were 60 and 70% of the positioning error is less than 2 m, the

Table 1 The error contrast of three algorithms on the second floor

	Histogram	Weibull-bin	Weibull-pdf
RMS	3.27	2.9	2.59
Mean error (m)	2.59	2.25	2.03
Max error (m)	11.15	11.15	9.12

Table 2 The error contrast of the three algorithms on the fourth floor

	Histogram	Weibull-bin	Weibull-pdf
RMS	2.87	2.61	1.86
Mean error (m)	1.96	1.79	1.37
Max error (m)	10.15	10.15	4.32

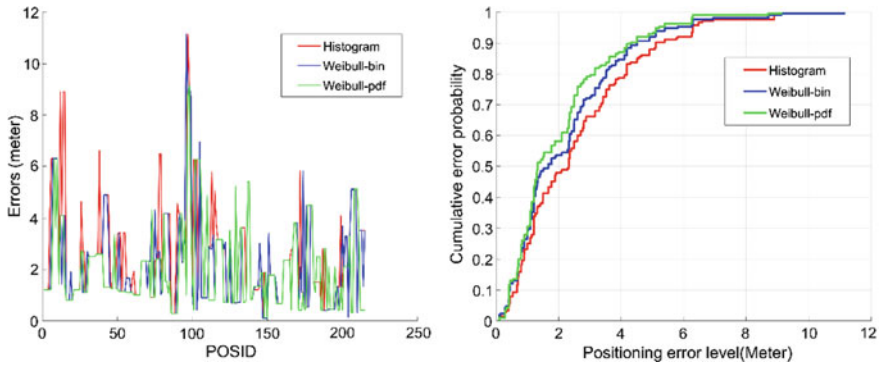


Fig. 2 The positioning error and cumulative probability distribution of the three algorithms on the second floor

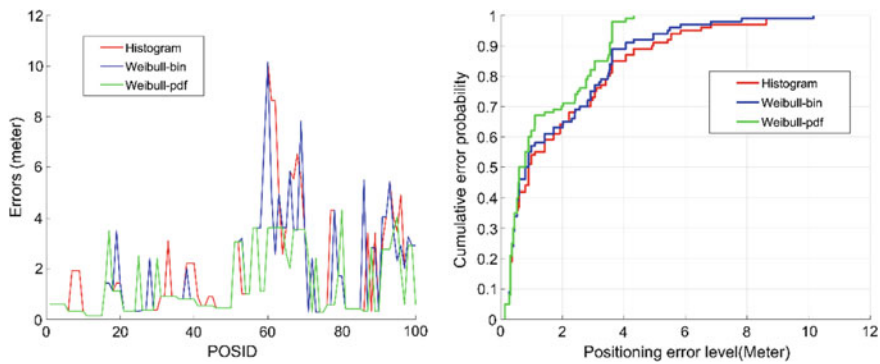


Fig. 3 The positioning error and cumulative probability distribution of the three algorithms on the fourth floor

other two algorithms only about 50 and 60% of the positioning error is less than 2 m. When the cumulative probability of positioning errors is 95%, the positioning errors of Weibull-pdf algorithm are 5.22 and 3.63 m respectively on the second and fourth floors, while the other two algorithms of the positioning errors is only 5.81 and 5.45 m, 6.3 and 6.18 m, obviously the Weibull-pdf algorithm proposed in this paper has the highest accuracy.

In general, the three algorithms in the experiment have some points with larger error, while the third algorithm is obviously better than the first two algorithms. The mean error of this new algorithm is 2.03 and 1.37 m respectively in the two actual scenes on the second and fourth floors, and 90 and 98% of the positioning errors are within 4 m. In addition, the RMS values of the proposed algorithm in this paper and the conventional Histogram algorithm are respectively improved 20 and 30% in

different scenarios. It shows that the new algorithm proposed in this paper not only improve the positioning accuracy but also improve the collection and calculation efficiency of the fingerprint database.

5 Conclusions

Nowadays, the application of WiFi in indoor positioning has been paid more and more attention by researchers. The algorithm proposed in this paper is based on the Bayesian Position Estimation method with RSSI measurements of WiFi, which optimize the fingerprint training phase and matching positioning phase respectively. During fingerprint training phase, the algorithm in this paper proposes the Weibull signal model to approximate the RSSI probability distribution which can greatly save computation and human resources. The contradiction between large sampling statistics and fingerprint collection efficiency is solved well. During matching positioning phase, the method of calculating the probability dynamically is more flexible to obtain the positioning result. The algorithm proposed in this paper not only improve the positioning accuracy but also improve the collection and calculation efficiency of the fingerprint database. Therefore, the new algorithm in this paper has a good prospect of scientific research and application.

Acknowledgements This study is supported in part by the National Key Research Development Program of China with project No. 2016YFB0502204 (High available and high precision indoor intelligent hybrid positioning and indoor GIS technology).

References

1. Chen X, Pang J (2014) Protecting query privacy in location-based services. Kluwer Academic Publishers 18(1):P95–P133
2. Vaughan-Nichols SJ (2009) Will mobile computing's future be location, location, location? Computer 42(2):14–17
3. Luo L (2014) Study on WiFi indoor location techniques based on android. Southwest Jiaotong University, Chengdu
4. Chen B, Liu R, Chen Y, Liu J, Jiang X, Liu D (2015) Wifi fingerprint based self-adaptive indoor localization in the dynamic environment. Chin J Sensors Actuators 28(5):729–738
5. Powell A (2009) Wi-fi as public utility or public park? Metaphors for planning local communication infrastructure. Social Science Electronic Publishing, New York
6. Afanasyev M, Chen T, Voelker GM, Snoeren AC (2008) Analysis of a mixed-use urban wifi network: when metropolitan becomes neapolitan. In: ACM SIGCOMM conference on internet measurement 2008, Vouliagmeni, Greece. DBLP, pp 85–98
7. Günther A, Hoene C (2005) Measuring round trip times to determine the distance between WLAN nodes. In: International conference on research in networking, vol 3462. Springer, Berlin, pp 768–779
8. Shuang H, Luo H, Ying C (2010) The design and implementation of a tdoa-based ultrasonic indoor localizing system. Chin J Sensors Actuators 23(3):347–353

9. Llobart M, Ciurana M, Barcelo-Arroyo F (2008) On the scalability of a novel WLAN positioning system based on time of arrival measurements. In: Workshop on positioning, navigation and communication (WPNC 2008). IEEE, pp 15–21
10. Ciurana M, Barcelo-Arroyo F, Llobart M (2009) Improving the performance of TOA over wireless systems to track mobile targets. IEEE international conference on communications workshops. ICC Workshops, IEEE, pp 1–6
11. Davidson P, Piche R (2016) A survey of selected indoor positioning methods for smartphones. IEEE Commun Surveys Tutorials PP(99):1
12. Li Q, Li W, Sun W, Wang J, Li J (2016) Wi-fi indoor localization algorithm based on rssi and assistant nodes collaboration. J Electron Measur Instrum 30(5):794–802
13. Dortz NL, Gain F, Zetterberg P (2012) WiFi fingerprint indoor positioning system using probability distribution comparison. In: IEEE international conference on acoustics, speech and signal processing, vol 22. IEEE, pp 2301–2304
14. Tang J, Chen Y, Chen L, Liu J, Hyyppä J, Kukko A et al (2015) Fast fingerprint database maintenance for indoor positioning based on ugv slam. Sensors 15(3):5311–5330
15. Pei L, Chen R, Liu J, Kuusniemi H, Tenhunen T, Chen Y (2010) Using inquiry-based bluetooth rssi probability distributions for indoor positioning. 9(2):122–130
16. Li Y (2015) Integration of MEMS sensors, WiFi, and magnetic features for indoor pedestrian navigation with consumer portable devices. Wuhan University, Whuhan
17. Zheng X, Fu J (2015) Study on pdr and rssi based indoor localization algorithm. Chin J Sci Instrum 36(5):1177–1185
18. Li B, Wang Y, Lee HK, Dempster A (2008) Method for yielding a database of location fingerprints in wlan. IEE Proc Commun 152(5):580–586
19. Badawy OM, Hasan, MAB (2007) Decision tree approach to estimate user location in WLAN based on location fingerprinting. In: Radio science conference (NRSC 2007). IEEE, pp 1–10
20. Sun Y, Xu Y, Ma L, Deng Z (2010) KNN-FCM hybrid algorithm for indoor location in WLAN. In: International conference on power electronics and intelligent transportation system, vol 2. IEEE, pp 251–254
21. Davidson P, Piche R (2016) A survey of selected indoor positioning methods for smartphones. IEEE Commun Surveys Tutorials PP(99):1
22. Wu D, Xu Y, Ma L (2009) Research on RSS based indoor location method. In: Pacific–Asia conference on knowledge engineering and software engineering (KESE '09). IEEE, pp 205–208
23. Zhang W, Hua X, Qiu W, Xue W, Zhou D (2017) A new combinatorial optimization algorithm for wifi positioning. Eng Survey Mapp 26(3):14–18
24. Ling P, Liu J, Robert G, Chen Y, Heidi K, Chen R (2012) Using LS-SVM based motion recognition for smartphone indoor wireless positioning. Sensors 12(5):6155–6175
25. Liu J, Chen R, Chen Y, Pei L, Chen L (2012) Iparking: an intelligent indoor location-based smartphone parking service. Sensors 12(11):14612–14629
26. Sagias NC, Karagiannidis GK (2005) Gaussian class multivariate weibull distributions: theory and applications in fading channels. IEEE Trans Inf Theory 51(10):3608–3619
27. Liu J, Chen R, Pei L, Chen W, Tenhunen T, Kuusniemi H et al (2010) Accelerometer assisted robust wireless signal positioning based on a hidden Markov model. In: Position location and navigation symposium, vol 298. IEEE, , pp 488–497
28. Papoulis A (2001) Random variables and stochastic processes. Phys Today 20(1):135
29. Youssef MA, Agrawala A, Shankar AU (2003) WLAN location determination via clustering and probability distributions. In: IEEE international conference on pervasive computing and communications. IEEE, pp 143–150
30. Roos T, Myllymäki P, Tirri H, Misikangas P, Sievänen J (2002) A probabilistic approach to wlan user location estimation. Int J Wirel Inf Networks 9(3):155–164

Development and Assessment of GNSS Online Differential Data Processing System Based on HNCORS



Yaozong Zhou, Cuilin Kuang, Shaohua Dou and Ziping Liu

Abstract This article first introduced the architecture, data processing flow, data processing strategies and data processing results of HNCORS GNSS online differential data processing system, and then analyzed the precision, repeatability and stability of the system by processing the GNSS data of datum stations and measured stations. Finally, based on the simulation deformation monitoring experiment, the feasibility of system application in deformation monitoring was discussed in this article. The experimental results showed that the repeatability and stability of the system is high, which can provide cm or mm level positioning services for Hunan Province GNSS users. What's more, the system has broad application prospects in deformation monitoring.

Keywords HNCORS · Online differential data processing system
Positioning services accuracy · Deformation monitoring

1 Introduction

The GNSS online data processing system is the product of Internet and GNSS high-precision data processing technology, and its appearance greatly simplifies the way of high-precision positioning service, which makes high-precision positioning service more convenient and flexible. According to the GNSS data processing modes, the GNSS online data processing systems can be divided into two types,

Y. Zhou (✉) · C. Kuang (✉) · S. Dou
School of Geosciences and Info-Physics, Central South University,
Changsha 410083, China
e-mail: zhouyaozong@csu.edu.cn

C. Kuang
e-mail: cuilinkuang@163.com

Z. Liu
Hunan Province Mapping and Science and Technology
Investigation Institute, Changsha 410007, China

which are GNSS online PPP data processing system and GNSS online differential data processing system [1, 2]. Compared with the GNSS online PPP data processing system, the GNSS online differential data processing system relies on CORS network, which can provide more accurate and reliable positioning services. Nowadays, there are several mature GNSS online differential data processing systems abroad, such as SCOUT and OPUS in the United States, AUSPOS in Australia, ASG-EUPOS in Poland, and PositionNZ-PP in New Zealand. All these systems can provide cm or mm-level positioning service to GNSS users [3–8]. Many domestic scholars also tried the development work of GNSS online differential data processing system. Back in 2007, Zhang et al. used Internet technology, database technology and Bernese5.0 software to develop the internet-based automatic GPS data processing system (Auto-Bernese) [9]. In 2011, for the complexity of GPS high-precision data processing software, Han et al. developed the GPS network data automatic processing system based on Bernese GPS (AUTO-BGPS) [10]. In recent years, Ye, Hong and Huang also made relevant researches [11–14].

China is a large country of GNSS promotion and application, and there are many CORS networks in China, such as the Crustal Movement Observation Network of China (CMONOC) and city CORS networks, but the mature and open GNSS online differential data processing system based on these networks cannot be found. Therefore, based on Hunan CORS (HNCORS) GNSS data and Bernese5.2 software [15], this study used Perl programming language and WEB programming language to develop an open regional GNSS online differential data processing system for GNSS users in Hunan Province. Then through rigorous data processing experiments, the service performance of the system was tested and analysed from time and space dimensions. Finally, the feasibility of the system application in deformation monitoring was verified through simulation deformation monitoring experiments.

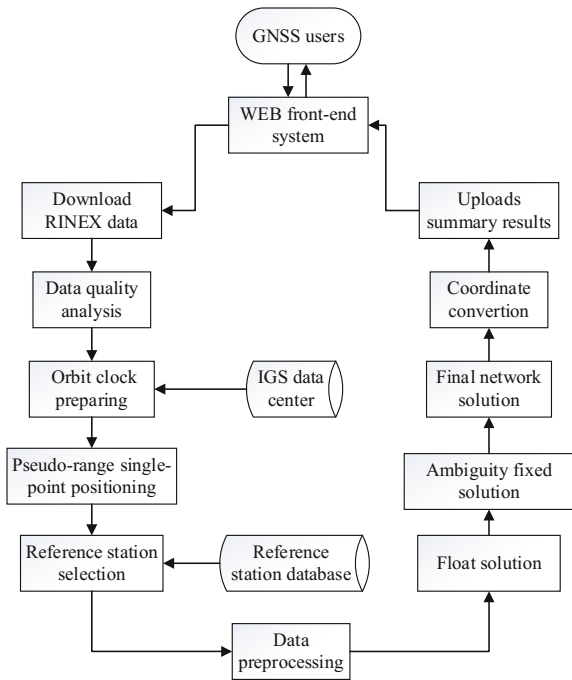
2 System Architecture and Data Processing Flow

This system is deployed on DELL PowerEdge R930 server. The operating system platform is Debian 8.8. The deployment work is mainly divided into three parts: WEB front-end system setup, back-end high-precision GNSS data automatic processing system setup and back-end front-end system docking. The WEB front-end system mainly realizes the functions of GNSS user registration, login and management, also provides a platform for GNSS users to upload GNSS RINEX data and download data processing results. Figure 1 is the schematic diagram of WEB platform. The main function of the back-end system is to automatically process GNSS data with high precision, also including the function of archiving data results, updating antenna information and periodically clearing files. The data processing flow of the system is shown in Fig. 2. The specific data processing procedures are as follows:

Fig. 1 Schematic diagram of WEB platform



Fig. 2 System data processing flow



1. After the GNSS users submit the RINEX data to WEB front-end server through WEB front-end system, the back-end system downloads the users RINEX data from WEB front-end server according to the task lists, analyzes the data quality of the users RINEX data by using anubis software, and generates the data quality analysis report.
2. The back-end system calls the Bernese data processing engine (secondary development of the Bernese 5.2 software), then the Bernese engine downloads

the SP3, CLK and ERP files from IGS data center to prepare orbit and clock. After that, the Bernese engine calculates the approximate coordinates of the user stations by pseudo-range single-point positioning.

3. According to distance, network structure and reference stations data quality, the Bernese engine selects 3 reference stations from reference station database based on the approximate coordinates of user stations. The reference station database of this online system is HNCORS datum station database, which includes more than 110 evenly distributed stations, and the average distance between these stations is about 50 km. Figure 3 is the site distribution of these stations.
4. According to the GNSS data of user and reference stations, the Bernese engine calculates the user station ITRF2008 coordinates through data preprocessing, float solution, ambiguity fixed solution and final network solution, then converts the ITRF2008 coordinates into CGCS2000 coordinates.
5. The back-end system summarizes and archives the data quality analysis and data processing results, uploads the summary results to the WEB front-end server, and notifies the users that the processing task has been completed by email. Afterwards, the GNSS users logs into WEB front-end system to download and view detailed results.

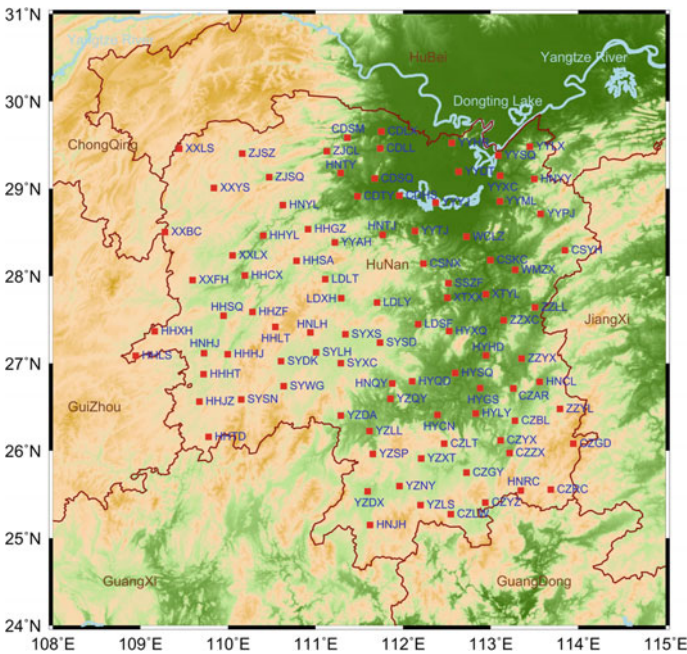


Fig. 3 Site distribution of reference station database

3 System Data Processing Strategies and Results

At present, the system can only process dual frequency data files. It is recommended that the GNSS users submit standard RINEX2 or RINEX3 data files to the system for processing. The system data processing strategies is shown in Table 1, and the system will select the appropriate IGS data product type based on the observation time of user station GNSS data.

The system results mainly include two modules, one is the data quality analysis report and the other is the data processing result report. According to satellite space time distribution, cycle slip, signal-to-noise ratio, multipath and other indicators, the system generates the data quality analysis report by using the anubis software to analyse user station GNSS data, and this report can fully reflect the quality of user data. The data processing result report is relatively rich, including IGS data product type, reference station selection information, user station CGCS2000 coordinates, coordinate accuracy information and so on. Figure 4 is the diagram of reference station selection.

4 System Positioning Service Performance Assessment

4.1 CORS Datum Station Data Testing

4.1.1 Hours Solution Results

The daily GNSS observation data of station CDHS (Changde) and CSKC (Changsha) was selected from HNCORS datum station database, and the selected observation period was day 291 of 2017. The *gfzrxn* software was used to split the

Table 1 GNSS data processing strategies

Processing parameters	Processing parameter values
Satellite orbit, clock and earth rotation parameters	IGS ultra-rapid/rapid/final
Nutation model	IAU2000R06
Subdaily pole model	IERS2010XY
Planetary ephemeris	DE405
Elevation cutoff angle	10°
Sample rate of solutions	30 s
Phase center variations	PCV.I14
Ionospheric delay	Dual-frequency correction
Tropospheric delay	GMF model correction
Ambiguity solution	Quasi-ionosphere-free (QIF) ambiguity resolution

Fig. 4 Schematic diagram of reference station selection



2 station daily data, and the time length of segment results was 1, 2, 3 ... 24 h. Then we submitted the 48 segment result files to the online system for solution, and got the series of three-dimensional coordinates. Finally, the coordinate differences between three-dimensional coordinates series and HNCORS reference values were calculated, and the differences were converted to N, E and U direction. Figure 5 gives the bias of Hours solution results. The figure shows that the N, E and U direction positioning biases of station CDHS and CSKC decrease obviously with the increase of time length. When the time length is one hour, the positioning deviations of station CDHS and CSKC reach mm level.

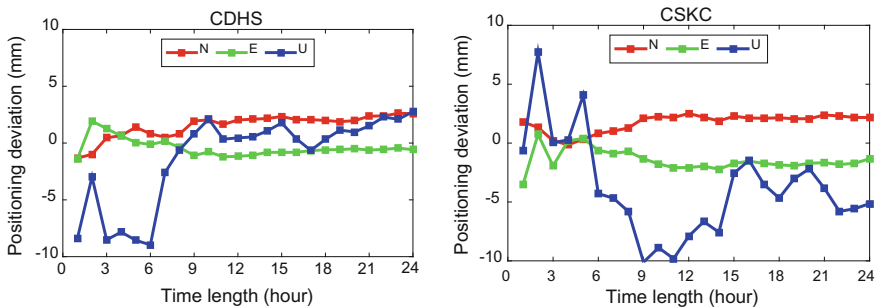


Fig. 5 Hours solution results of station CDHS and CSKC

4.1.2 Daily Solution Results

The daily GNSS data of station CDHS and CSKC was selected from HNCORS database, and the data period was day 291 to 310 of 2017. Then we submitted the selected data to the system for solution to obtain the coordinates time series of station CDHS and CSKC. Finally, we compared the coordinates time series with reference values, and the compared results are shown in Fig. 6. The IGS Rapid (IGR) products are used after day 308 in Fig. 6. From Fig. 6, we can see that the N and E direction positioning biases of daily solution can reach mm level, and the U direction positioning deviation is slightly larger, reaching 1 cm. The single-day time series of station CDHS and CSKC is stable, and has good repeatability.

4.1.3 Multi-station Solution Results

The daily GNSS data of 20 evenly distributed stations was selected from HNCORS database. The selected observation periods are named period 1 (day 291 of 2017) and period 2 (day 300 of 2017). The site distribution of the selected station can be found in Fig. 7. The selected data were submitted to the system for solution, and the differences between single day solution results and reference values were calculated. The calculated results of period 1 and 2 are shown in Fig. 8a, b. In Fig. 8, the colour of the point indicates the U direction positioning deviation. The length of the arrow represents N and E direction positioning bias. The three arrows in the lower-left corner respectively indicate 6, 4 and 2 mm N and E directions positioning deviation. According to this, we can determine the specific value of N and E direction positioning deviation. Figure 8 shows that the N and E direction positioning biases of 20 selected stations are in mm level, and the U direction positioning biases of several stations is slightly larger, up to 2 cm.

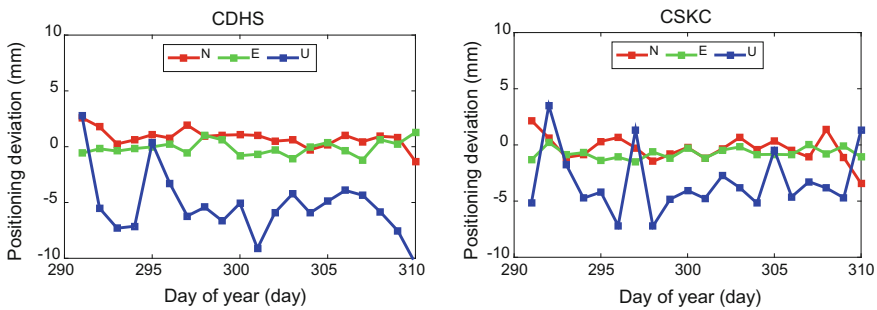


Fig. 6 Daily positioning results of station CDHS and CSKC

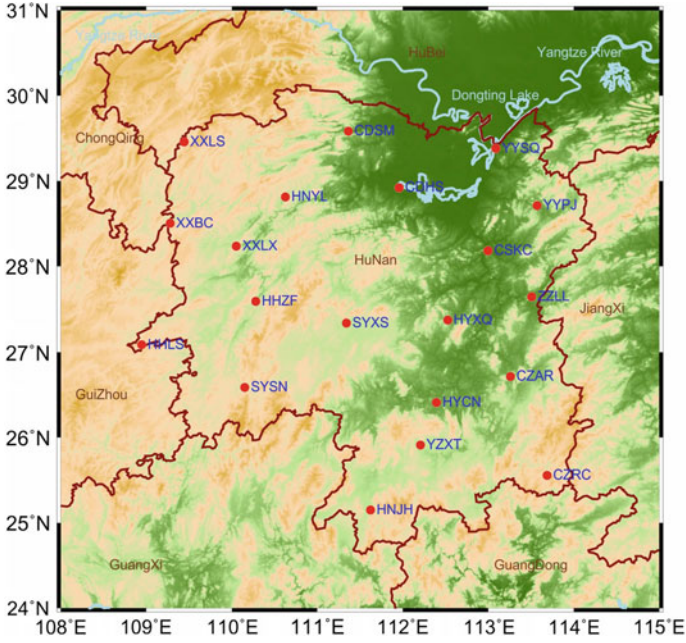


Fig. 7 Site distribution of experimental data

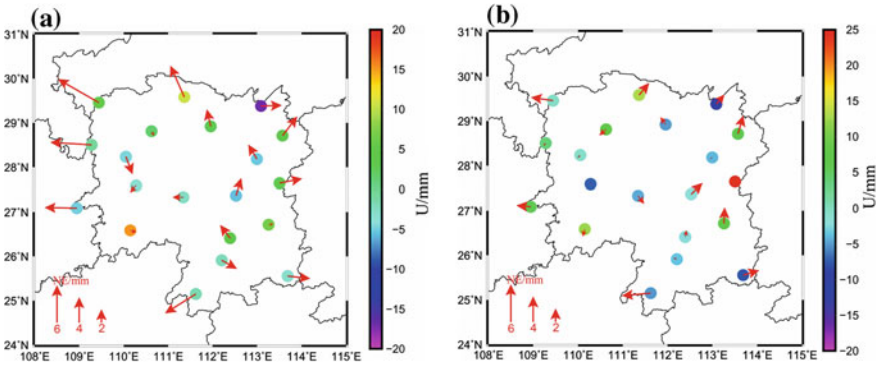


Fig. 8 Spatial distribution of the daily positioning result

4.2 Measured Data Testing

4.2.1 Data Acquisition

On the rooftop of Grammar Building at Central South University, two sets GNSS data acquisition equipment (GNSS receiver and antenna) were installed for GNSS data collection. The two station names were set as CSUB and CSUR. The receiver type is “Trimble NetR9”. The antenna types are “TRM55971.00 NONE” and “TRM55971.00 DOME”, respectively. The data sample interval is 30 s. The elevation cut-off angle is 0° . The data collection time is from October 21 to November 9 in 2017. Figure 9 shows the diagram of experimental field.

4.2.2 Hours Solution Results

We selected the GNSS data of CSUB and CSUR, and the selected observation date was day 294 (October 21) of 2017. The daily GNSS data of station CSUB and CSUR was split into segment files, and the time length was 1, 2, 3 ... 24 h. To get the series of hour solution coordinates, the 48 segment data files were submitted to the online system for solution. The coordinate differences between hour solution coordinates and reference values (average of 20-day solution coordinates) were calculated, and the deviations were converted to N, E and U direction. Figure 10 gives the time series of N, E and U direction positioning bias. The reason for the

Fig. 9 Experimental field of station CSUB and CSUR



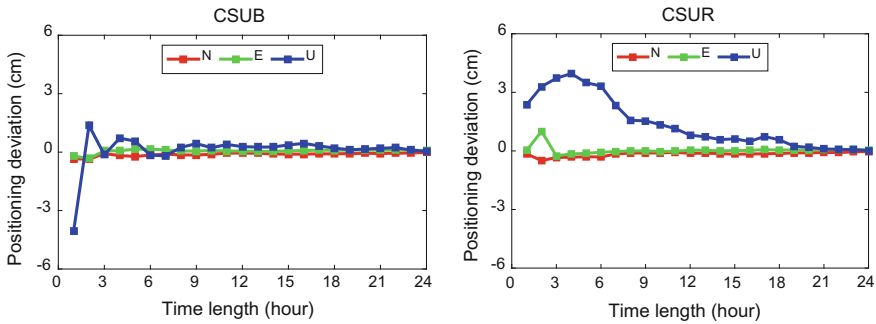


Fig. 10 h solution results of station CSUB and CSUR

gradual convergence of the CSUR station time series in Fig. 10 is that the multipath effect of the CSUR station is more serious. The results of Fig. 10 show that with the increase of time length, the U direction positioning biases of station CSUB and CSUR decrease obviously first, and then tend to be stable within 1 cm. Compared with the U direction results, the N and E direction positioning bias are much smaller, which are stable within 1 cm after the time length reaching 3 h.

4.2.3 Daily Solution Results

We selected 20 days GNSS data of station CSUB and CSUR, and its observation period was from day 294 of 2017 to day 313 of 2017. To get the daily solution coordinates, all selected daily data was processed by the system. Compared the daily solution coordinates with reference values (average of 20-day solution coordinates), we got the positioning biases, which were plotted in Fig. 11. The IGR products were used after day 308 in Fig. 11. Figure 11 shows that the repeatability of station CSUB and CSUR daily solution results is well, and the N, E and U direction positioning deviations are within 6 mm.

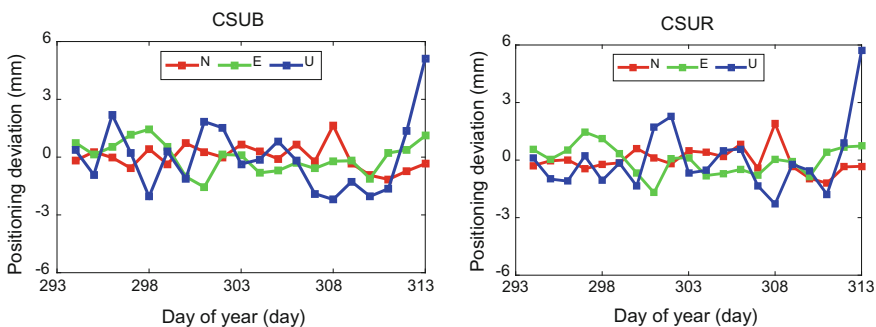


Fig. 11 Daily positioning result of station CSUB and CSUR

5 Simulation Deformation Monitoring

5.1 Data Collection

On the rooftop of Grammar Building at Central South University, another set GNSS data collection equipment (receiver + mobile base + antenna) was installed to collect the GNSS data. The station name was set as CSUM. The receiver type is “Trimble NetR9”. The antenna type is “TRM55971.00 NONE”. The data sample interval is 30 s. The elevation cut-off angle is 0°. The data collection time is from October 21 to November 9 in 2017. Figure 12 is the schematic diagram of experimental field. The position of GNSS antenna was shifted according to Table 2.

Fig. 12 Experimental field of station CSUM



Table 2 Antenna position of station CSUM

Date	Antenna position (mm)			Date	Antenna position (mm)		
	N	E	U		N	E	U
Oct 21	0	0	0	Oct 28	0	0	0
Oct 22	5	5	5	Oct 29	-10	10	10
Oct 23	10	10	10	Oct 30	-20	20	20
Oct 24	15	15	15	Oct 31	-10	10	10
Oct 25	10	10	10	Nov 01	0	0	0
Oct 26	5	5	5	Nov 02	0	0	0
Oct 27	0	0	0	Others	0	0	0

5.2 Data Processing and Results Analysis

We selected 13 days GNSS data of station CSUM. The start and end observation dates are October 21 (day 294 of 2017) and November 02 (day 306 of 2017) in 2017. The daily selected GNSS data was split into 24, 12, 6 and 3 h by using *gfzrnrx* software, and then we submitted all segment data files to the online system for solution. We got the solution coordinates of all segment data files first. Then we selected 7 days GNSS data of station CSUM, and the observation date is from November 03 to November 09. we calculated the reference coordinate (average of 7-day solution coordinates) by submitting the 7 day GNSS data to the online

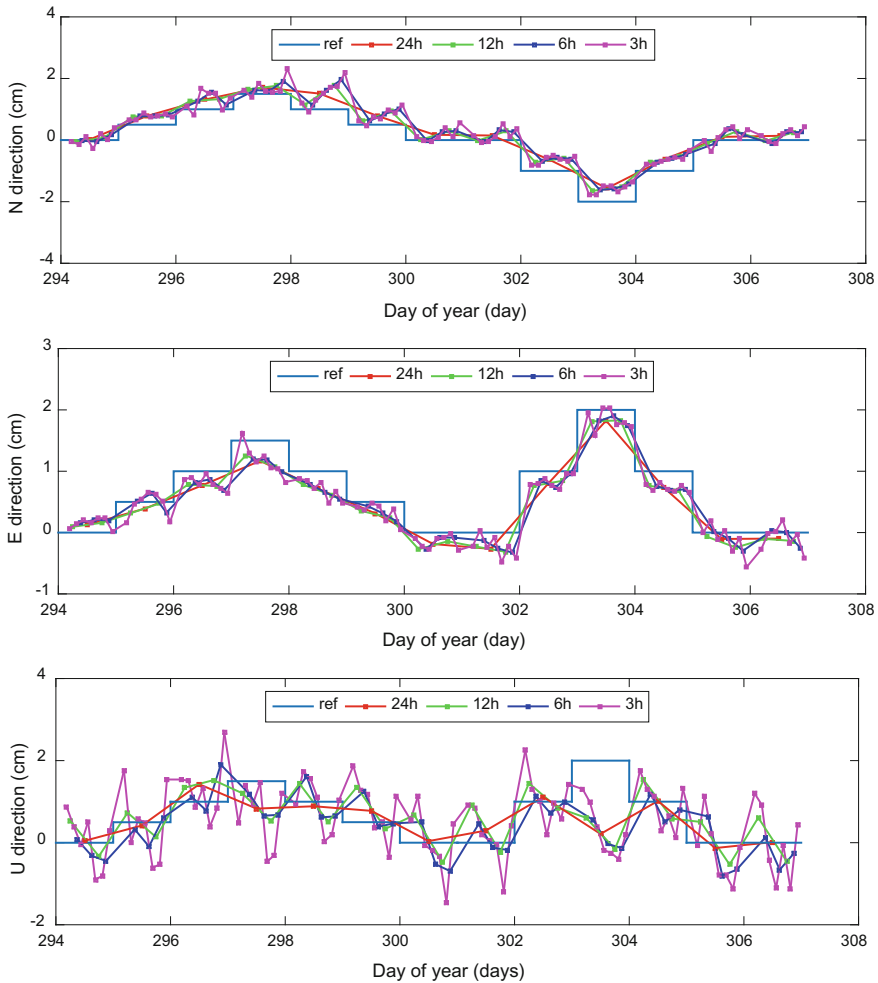


Fig. 13 Simulation results of deformation monitoring

system for processing. Finally, the biases between the solution coordinates and reference coordinates were calculated, also the biases were converted to N, E and U direction. The final calculated results were plotted in Fig. 13.

Figure 13 shows that the N and E direction positioning results of station CSUM have been well when the time length is 3 h, which can well trace N and E direction deformation signals. The U direction positioning results of CSUM station are slightly worse, but it basically reflects the trend of deformation signals.

6 Conclusions

In this paper, the system architecture, data processing flow, data processing strategies and data processing results of HNCORS GNSS online differential data processing system were introduced firstly. Then through CORS datum station data testing and measured data testing, the positioning accuracy, repeatability and reliability of the system were analyzed in detail. Finally, based on simulation deformation monitoring experiment, the feasibility of system application in deformation monitoring was discussed. In the last, we got the following conclusions:

1. The CORS datum station solution results have good repeatability and high precision. The N, E and U direction positioning deviations of station CDHS and CSKC are all within 1 cm when the time length is one hour. The daily solution positioning biases of all tested stations are within 1 cm in E and U directions, and the U direction positioning deviations of all tested stations are less than 2.5 cm.
2. The measured data solution results also have good repeatability and high precision. The N and E direction positioning deviations are about 1.5 cm when the session length is one hour, and the U direction positioning deviations are slightly larger, about 5 cm. The N, E and U direction positioning deviations of single day solution results are in mm level.
3. The N and E direction positioning results of the system have been well when the time length is 3 h, which can well trace N and E direction deformation signals. The U direction positioning results of the system are slightly worse, but it basically reflects the trend of deformation signals.

Based on the above conclusions, authors believe that the system can provide cm or mm level positioning services for Hunan Province GNSS users, and has broad application prospect in deformation monitoring. Of course, there are some deficiencies in the system. For example, the U direction positioning accuracy of the system is relatively poor, which can be improved by refining the tropospheric delay correction model.

Acknowledgements Foundation item: National Natural Science Foundation of China (41774040).

References

1. Ghoddousi-Fard R, Dare P (2006) Online GPS processing services: an initial study. *GPS Solutions* 10(1):12–20
2. Guo Q (2015) Precision comparison and analysis of four online free PPP services in static positioning and tropospheric delay estimation. *GPS Solutions* 19(4):537–544
3. Dawson J, Govind R, Manning J (2001) Application of the AUSLIG online GPS processing system (AUSPOS) to Antarctica. In: *Proceeding of Satnav*
4. El-Mowafy A (2011) Analysis of web-based GNSS post-processing services for static and kinematic positioning using short data spans. *Survey Rev* 43(323):535–549
5. Rapiński J, Cellmer S (2011) Tests of selected automatic positioning systems in post-processing mode. *Technical Sciences/University of Warmia and Mazury in Olsztyn*, pp 45–56
6. Ocalan T, Erdogan B, Tunalioglu N (2013) Analysis of web-based online services for GPS relative and precise point positioning techniques. *Boletim de Ciências Geodésicas* 19(2):191–207
7. Pearson C, Crook C, Jordan A et al (2015) *PositionNZ-PP: an online GPS processing application for New Zealand*. IAG 150 Years. Springer, Cham, pp 549–555
8. Wang G, Turco M, Soler T et al (2017) Comparisons of OPUS and PPP solutions for subsidence monitoring in the greater Houston area. *J Survey Eng* 143(4):05017005
9. Zhang C, Nie G, Xiong X et al (2007) Research on system for internet-based GPS data auto-processing. *Bull Survey Mapp* 08:26–28
10. Han F, Cheng CL, Wang X (2011) A study of automatic GPS data processing system based on Bernese GPS software. *Bull Survey Mapp* 11:8–10
11. Ye HC, Nie GG, Li N (2013) Data processing and re-development with Bernese software. *J Geomatics* 38(02):20–22
12. Hong W (2013) CORS-based GPS data processing web platform design and its application. *Geospatial Inf* 11(04):25–27
13. Ye HC, Nie GG, Yang JH et al (2013) Research on high-precision GPS data online processing system. *Bull Survey Mapp* (05):16–19
14. Huang C, Yang Y, Xiong WD et al (2015) Regional CORS data auto-processing system based on Bernese 5.0. *J Shandong Un Sci Technol (Nat Sci)* 34(05):75–81
15. Dach R, Lutz S, Walser P et al (2015) Bernese GNSS software version 5.2

A Robust Turn Detection Algorithm Based on Periodic Signal Identification



Yu Chen, Haiyong Luo, Fang Zhao, Wenhua Shao and Qu Wang

Abstract Turn detection can be widely used in location navigation, user behavior identification, and scene awareness. Accurate real-time identification of pedestrian turns contributes to improving the accuracy of positioning navigation accuracy and scene awareness. Considering the problem of the insufficient accuracy and robustness of the existing turn detection algorithm, this paper proposes a turn detection algorithm based on periodic signal recognition, which effectively solves the misjudgment caused by the periodic swing of user walking. The algorithm collects the acceleration sensor and the gyro sensor data in real time and calculates the angular velocity of the rotation in the vertical direction by multiplying the vertical gravitational acceleration unit vector by the gyro sensor data. In this paper, fast Fourier transform (FFT) is used to identify and eliminate periodic interfering signals generated by user walking, only keeping the non-periodic signal generated by user walking (such as turning). By integrating the vertical angular velocity data in the sliding window to obtain the angle change value of the user walking within a certain period of time, and the threshold value is compared to realize the accurate identification of the turning behavior. In order to adapt to the user's turn size and

Y. Chen (✉) · F. Zhao · W. Shao

School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100083, China

e-mail: scaldingblood@gmail.com

F. Zhao

e-mail: zfsse@bupt.edu.cn

W. Shao

e-mail: shaowenhua@ict.ac.cn

H. Luo

Research Centre of Pervasive Computing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

e-mail: yhluo@ict.ac.cn

Q. Wang

School of Information and Communication Engineering,

Beijing University of Posts and Telecommunications, Beijing 100083, China

e-mail: wangqu@ict.ac.cn

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_27

the turn speed, the algorithm also proposes a corner detection method based on multi-sliding window. The experiment results show that the proposed algorithm has higher accuracy and lower power consumption than other algorithms which based on GPS and electronic compass. The algorithm can realize the accurate identification of the turning behavior of the mobile phone in various positions of the user, and has good robustness. The accuracy of turning recognition can reach 93% and the average power consumption is about 60 mW.

Keywords Turn detection · Scene awareness · Signal recognition
Multi-sliding window · High accuracy · Low power consumption

1 Introduction

Turn detection can be widely applied for positioning and navigation, scene awareness, user behavior recognition [1] and other areas. Accurately identify turns often can optimize the local details, significantly improve the overall performance. For example, in the process of constructing a map by the SLAM (Simultaneous localization and mapping) algorithms, an accurate turning position is crucial [2]. When positioning, the turn detected could assist calibration using the map [3]. Besides, the step size can be estimated by measuring and counting the distance between two turns detected and the number of steps. Robust and accurate algorithm of turning detection has wide application prospect and research value in traffic pattern recognition, indoor topology mining [4] and other fields. However, the existing algorithms of turn detection cannot meet the demand.

With the application of GPS (Global Position System) [5], high judging accuracy of turn detection could be easily achieved. But it has been proved ineffective in indoor situations. The other usual way is detecting turning using lightweight sensors which have acceptable precision and wide application scenes. But when people put the phone in the pants pocket or take it in the hands of swinging, this cycle will produce a significant influence on the angle, leading to the misjudgments.

In this paper, a turn detection algorithm based on periodic signal identification is proposed. By calculating the vertical gravitational acceleration unit vector, the angular velocity in the direction of the gravity could be obtained to achieve the turn detection in various postures. In order to adapt to the user's different turn size and the turn speed, the algorithm also proposes a turn detection method based on multiple sliding windows. In addition, the algorithm uses FFT (fast Fourier transform) [6] to identify and remove human walking wobble that produces periodic signals, remaining aperiodic turning signals. Thereby algorithm is able to improve the accuracy and robustness and realize the accurate identification of turn when putting the terminal in the pants pocket, in the hands swinging back and forth, and in other complex dynamic scenes.

The paper is organized as follows: In Sect. 2, the related research is introduced. Section 3 describes the system framework. In Sect. 4, details are described. Section 5 uses some experiment to verify the performance and reliability of the algorithm. Finally, Sect. 6 is conclusions and future work.

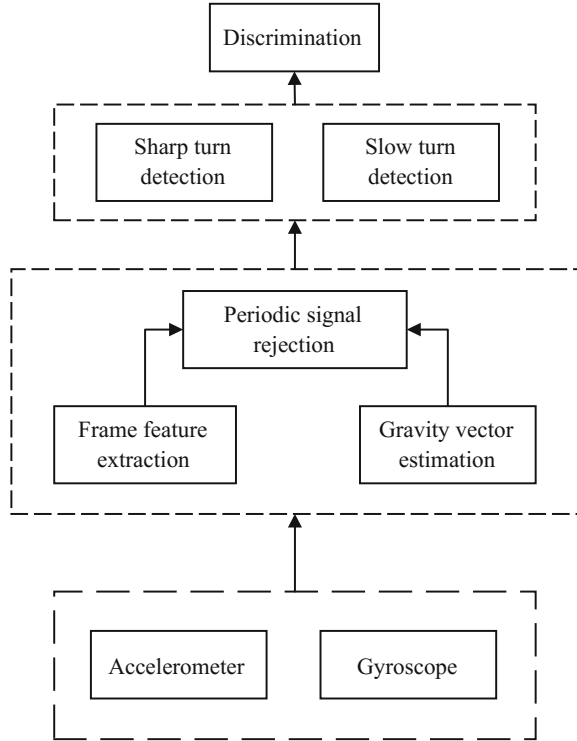
2 Related Work

The existing turn detection algorithms mainly use GPS, electronic compass or inertial sensors. The GPS-based algorithm [7] uses continuously acquired coordinates to detect the turn by observing relative changes in position. This kind of technology does not apply to indoor scenes, having the larger delay and higher power consumption. The use of electronic compass [8, 9] has the advantage of low power consumption, but susceptible to interference caused by ferromagnetic material misjudgment in the indoor environment [10].

In [11], the acceleration sensor and gyroscope are used to determine whether or not turn occurs by calculating the integral of the angular velocity of the device z-axis in the flat end attitude. This way needs to maintain a fixed attitude to the device so that practicality is subject to great restrictions. In order to realize the turn detection of the device in various gestures, the angular velocity calibration value is obtained from the accumulated data in [12]. Then the difference is calculated with the collected angular velocity and calibration value, and compared with the threshold to judge the turning. In [13, 14], the gyroscope and the acceleration sensor are used to estimate the unit gravitational acceleration vector through the sliding window. The angular velocity of turn is calculated by projecting original angular velocity value to the direction of the unit gravitational acceleration vector, and the turn detection is realized by computing the integration of it. This algorithm can adapt to the various gestures of the device, but when the user walks with the device in the pants pocket or in the hands of swing back and forth, such algorithms tend to misjudge the angle change of cyclical swing movement into a turn. Machine learning methods can also be used to determine the turn [15], such as the hidden Markov model [16].

Different from the above algorithms, this paper uses FFT-based periodic signal recognition technology, identifying and eliminating user periodic motion signals and retaining aperiodic signals during the process of turning. So, the action of turning could be identified accurately and robustly when the device in a variety of complex postures.

Fig. 1 The architecture of robust turn detection system



3 System Overview

In this section, the framework of the turn detection system is introduced. This system can be deployed and used on smartphones. The whole system applies the hierarchical architecture. The data is transported between adjacent layers during processing. The system architecture is shown in Fig. 1.

The system encapsulates the captured data into a frame format for the parent module processing through mobile phone's own sensor. Referring to [17], the gravitational acceleration calculation module uses an algorithm that dynamically adjusts the value of gravitational acceleration by analyzing variance. The periodic signal rejection removes the angular deflection caused by walking and preserves the aperiodic signal (such as turn signal) to ensure that the calculated angle is the angle of the turn. The fast and slow turn detection module will process the data and transform the result to the upper layer. The judgment module will combine the results of two modules and make a decision. It always accepts a quick turn and judges whether the slow turn is caused by a quick turn. Eventually, the result will be added up to the counter module.

Turn detection algorithm based on periodic signal identification is shown in Fig. 2.

```

Algorithm Detect Turns
Input: Accswindow, Gyrswindow
Output: Judgment Result
1: function DetectTurn( Accswindow, Gyrswindow )
2:   for all Accsframe, Gyrsframe ∈ Accswindow, Gyrswindow do
3:     Angle ← Angle + GetAngle( Accsframe, Gyrsframe )
4:   end for
5:   if Angle < Anglethreshold then
6:     return True
7:   else
8:     return False
9:   end if
10: end function
11:
12: function GetAngle( Accs , Gyrs )
13:   Accmean ← Mean( Accs )
14:   Accvariance ← Variance( Accs )
15:   Gravity ← GetGravity( Accmean, Accvar )
16:   for all Gyr ∈ Gyrs do
17:     Gyrgravity ← Gyrx * Gravityx + Gyry * Gravityy + Gyrz * Gravityz
18:   end for
19:   Datafreq ← FFT( Gyrs )
20:   len ← Length( Datafreq )
21:   while i < len / 2 do
22:     if i > threshold then
23:       Datafreq[i] ← 0
24:       Datafreq[len - i] ← 0
25:     end if
26:   end while
27:   Gyrs ← iFFT( Datafreq )
28:   result ← interval * Sum( Gyrs )
29:   return result
30: end function

```

Fig. 2 Turn detection algorithm based on periodic signal identification

4 Turn Detection

4.1 Gravitational Unit Vector Estimation Based on Sliding Average

It's necessary to determine the data used in algorithm refer to vertical gravity direction of the local geographical coordinate system while turning as the user carries the phone in the various gestures (flat, vertical, inclined). To this end, gyro data collected in phone coordinate system (carrier coordinate system) should be projected to the direction of gravity to calculate the user's the rotation angular velocity relative to the vertical direction.

The gravity is estimated by calculating the variance of the angular velocity within a frame. If the variance is large, the current state of the phone is unstable and the mean value of the acceleration will be used to approximate the acceleration of gravity. While the variance is small means that the current state of the phone is relatively stable, the acceleration of this frame will be used to approximate gravity acceleration. The gravitational acceleration obtained in this way is more sensitive and reliable.

After gaining the gravitational acceleration, the acceleration can be decomposed into vertical and horizontal directions. Vertical component of the angular velocity can be obtained by dividing the dot product of angular velocity and gravitational acceleration by the modulus of gravitational acceleration, as shown in Eq. (1) where \vec{g} is the gravitational acceleration obtained in the current frame, $\vec{\omega}_i$ is the angular velocity in the i -th sampling period of the current frame, and $\omega_{i\text{-vertical}}$ is the vertical component.

$$\omega_{i\text{-vertical}} = \frac{\vec{\omega}_i \cdot \vec{g}}{|\vec{g}|} = \frac{\omega_{ix} * g_x + \omega_{iy} * g_y + \omega_{iz} * g_z}{\sqrt{g_x^2 + g_y^2 + g_z^2}} \quad (1)$$

4.2 Turn Judgment

The system can get the angular velocity around the direction of gravity by calculating the gravity, that is, the angular velocity of the user's rotation. Integrate this angular velocity to obtain an angle of rotation within a frame. The cumulative multi-frame angle value is the sum of the angles of rotation for a period of time. As shown in Eq. (2), where θ_{sum} is the angle within a window (n frame), θ_{frame} is an intra-frame angle and $\omega_{i\text{-vertical}}$ is the angular velocity around gravity in the i -th sampling period of the current frame, T is a frame length.

$$\theta_{sum} = \sum_{i=1}^n \theta_{frame} = \sum_{i=1}^n \int_0^T \omega_{i\text{-vertical}} dt \quad (2)$$

Algorithm set an angle threshold for turning. If the sum of the calculated angles reaches or exceeds the threshold, it may be considered that a turn has occurred at this time. But the turn is not an instantaneous action, may be over a period of time the angle exceeds the threshold. At this time, a turn could be counted several times. For this reason, the algorithm sets a flag that is initially set to a non-turn state. When the system detects the angle and exceeds the set threshold, it is taken into the corner and set to be marked as turning. Thereafter angle exceeds the threshold value will not be considered to occur a bend until the angle is less than the threshold value. After that, the flag will be reset and waits for a new bend to occur. In this way, the system will get the exact number of turns. The threshold is obtained by taking the mean of the minimum value of the angle of the turn and the maximum value of the angle of the error.

In fact, the calculated angle may fluctuate near the threshold, leading to misjudgment of the turn. The algorithm sets a fixed interval time and it is considered the end of the turn and reset the mark when the angle less than threshold reached this length.

4.3 *Fast Turn and Slow Turn*

Fast turn corresponds to the turn that finishes in a short period while the slow one corresponds to the turn which costs longer or has not obvious range. The algorithm applies multiple sliding windows to calculate the sum of the angles over consecutive time periods to detect the turns with different speed.

The algorithm builds a fixed sliding window. This window converses the angle of several consecutive frames. The oldest frame of data in the window will be removed before adding a new frame to the window. Meanwhile the window slides and the data updates. Each time window slides, the sum of angles in window will be calculated, used to judge occurrence of turn.

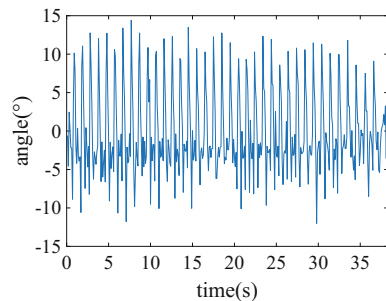
The system sets up several different windows, short window for detecting fast turns, and long window for detecting slow turns. If you only keep a short window, slow turn longer than the short window length will lead to the angle of the window in the cumulative value does not meet the threshold, missing the judgment of turning. Only keeping a long window will also cause the wrong result when several turns occur because the long window will accumulate all the angles within the window, so the system can only detect zero or one corner, resulting in missing quickly turn.

As turns occurred in short window could be captured by the long window redundantly, system will always count turns in short window. If the short window doesn't detect turn while the long window does, this turn will also be recorded.

4.4 *Periodic Signal Identification Based on FFT*

When the user places the device in the pocket or takes it in his hand, the device will swing with the body periodically, resulting in a deviation from the calculated angle and inaccurate results of the turn identification. The system preprocesses the data before calculating the angle to resolve the problem.

Fig. 3 Change of angle cause by swing during walking



Take the sum angle of 0.1 s and draw the graph (as shown in Fig. 3). Observing the picture, the change of angle caused by walking is a periodic signal, whose frequency is about 1 Hz.

The target turn signal is a non-periodic signal. The algorithm uses the fast Fourier transform to transform the signal into the frequency domain to eliminate the wobble signal caused by walking and retain the aperiodic turn signal, see Eq. (3). By plotting the walking swing frequency signal (shown in Fig. 4) and the turning frequency signal (as shown in Fig. 5), it is confirmed that the periodic signal generated from walking is indeed a component of about 1 Hz. β_k is the angle in the k th 0.1 s, representing frequency domain component of the n th point after the discrete Fourier transform.

$$\alpha_n = \sum_{k=0}^{N-1} \beta_k e^{i\frac{2\pi}{N}kn} \quad n = 0, \dots, N - 1 \tag{3}$$

Comparing the frequency domain characteristics of the two signals, it is found that setting a window function can filter the angle effect caused by the swing. The algorithm multiplies the signal by the window function to preserve the aperiodic signal and eliminating the wobble signal generating from walking, as shown in Eqs. (4) and (5). $w(n)$ is the rectangular window, α_n represents the frequency domain component of the n th point after the discrete Fourier transform, and λ_n represents the frequency domain component after the periodic signal is removed.

$$w(n) = R_M(n) = \begin{cases} 1, & 1 \leq n \leq M - 1 \\ 0, & \text{other} \end{cases} \tag{4}$$

$$\lambda_n = \alpha_n \cdot w(n) \tag{5}$$

And then the frequency domain signal will be transformed back to the time domain using the inverse Fourier transform, so as to remove the human walking swing caused by the angle changes and get more accurate identification. The simulation results of the angle that eliminates periodic signal in 0.1 s are shown in Fig. 6.

Fig. 4 Walking signal in the frequency domain

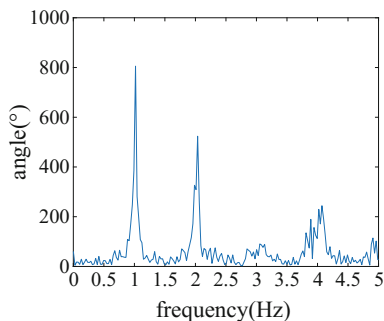


Fig. 5 Turning signal in the frequency domain

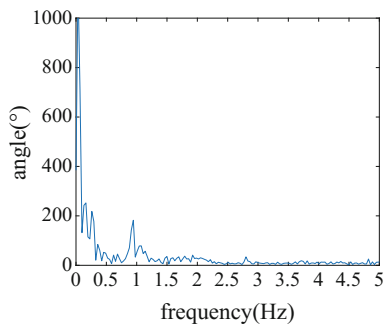
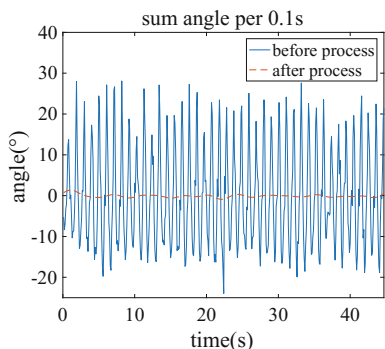


Fig. 6 Change of angle after eliminating periodic signals



5 Experiment Results and Analysis

The experiments evaluated the performance of the entire system and verified the accuracy of the algorithm. There were a number of representative scenes, including indoor laboratories and outdoor roads, and a variety of commonly used transports, such as bicycles, buses, and subways to ensure the universality of the experiments. The devices used in the experiments are Huawei Mate 8 and Nextbit Robin. In addition, the experiments compared the performance of the traditional turn detection algorithm proposed in [18] which could recognize indoor and outdoor scenes and in this paper.

5.1 Evaluate Turn Detection Algorithm Performance When Walking

The experiment shows the performance of the periodic signal elimination algorithm by contrast. The location of the walking experiment is in the laboratory of the Chinese Academy of Sciences. There are two walking routes, one is a straight

corridor and the other is a square aisle, respectively, corresponding to the walking postures of going straight and turning.

The purpose of the straight-line experiment is to show the performance of the algorithm by comparing the magnitude of the angular offset caused by the swing. We experimented with different walking speeds and always kept the phone in a swinging state. In the normal walking speed (about 1.2 m/s), it takes about 90 s to finish the whole process. And then walk with a slow walking speed (about 0.8 m/s), it takes about 135 s to complete. And finally, it takes about 72 s in the faster speed (about 1.5 m/s). In the experiment, we set the length of the short window to 3.84 s. The window slides every 0.1 s and counts the angle values in the window at this time.

Experiments compare the results of the algorithm proposed in this paper and the one used in [18] of indoor and outdoor recognition which doesn't eliminate the periodic oscillation signal rotation detection algorithm experiment results. The resulting image is shown in Fig. 7.

Observing the simulation results obtained in Fig. 7, we find that the data after eliminating the error still have negative deviation. This deviation is generated from the human body that produces a small twist, rather than the inherent deviation of the sensor. In this experiment, the phone is placed in the right pocket. If placed in the left one, there would be a positive deviation. The experimental data is shown in Tables 1 and 2.

The experiment's result shows that the wobble angle is smaller and the variance is significantly reduced. The angular deflection gets more stable.

The test of the turning path also takes a different walking speed, similar to the straight-line experiment. We have completed 10 trajectories of the walk, each one includes four corners.

Experiment is simulated using the collected data, including both clockwise and counterclockwise situations. The simulation results are shown in Figs. 8 and 9.

In the experiment, the test included 40 turn trajectories and 40 straight trajectories. The state of walking was considered, including detours, walking slashes, and so on. Tables 3 and 4 are the results.

Fig. 7 Sum of angles in 3.84 s when walking straight

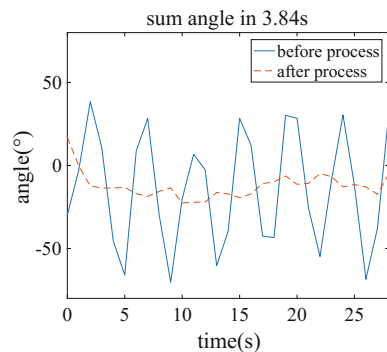


Table 1 Result of turn detection algorithm in [18] when walking on straight path

Walk speed	Average angle (°)	Max angle (°)	Probability of sum angle less than 20° (%)
Slow	-10.60	-35.26	74.1
Normal	-14.42	-34.33	73.7
Fast	-18.08	-42.45	52.0

Table 2 Results of turn detection algorithm based on periodic signal identification when walking on straight path

Walk speed	Average angle (°)	Max angle (°)	Probability of sum angle less than 20° (%)
Slow	-10.19	-33.17	87.6
Normal	-8.28	-33.51	91.3
Fast	-5.31	-25.69	93.7

Fig. 8 Sum of angles in the sliding window when turning clockwise

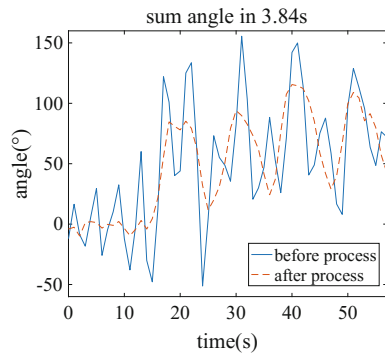


Fig. 9 Sum of angles in the sliding window when turning anticlockwise

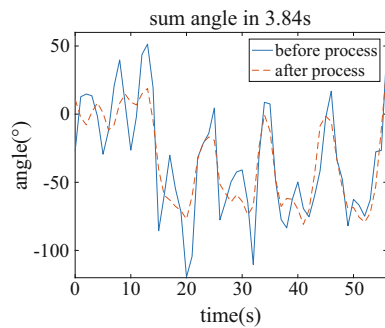


Table 3 Result of turn detection algorithm in [18] when walking on ring path

Walk speed	Detection number	Right number	Accuracy (%)
Slow	80	73	91.3
Normal	80	72	90.0
Fast	80	71	88.8

Table 4 Results of turn detection algorithm based on periodic signal identification when walking on ring path

Walk speed	Detection number	Right number	Accuracy (%)
Slow	80	75	93.8
Normal	80	78	97.5
Fast	80	78	97.5

The results show that the algorithm could reduce the angle deviation caused by the wobble signal while preserving the turning characteristics, improving the accuracy of the turn detection.

5.2 Evaluate Turn Detection Algorithm Performance on Transports

The experiments are carried out on different transports such as bicycles, buses, and subways, verifying the accuracy and robustness of the algorithm.

The procedure is as follows: Open the device's turn detection application and place it in the pocket or hold it. Record the detected turns and actual turns each time. Finally, organize the data. The experimental data is shown in Tables 5 and 6.

The result shows that the accuracy of the recognition is acceptable and the sliding window works in the turning recognition of various vehicles.

5.3 Algorithm Complexity Analysis

Step-by-step analysis of the overall complexity of the algorithm will be illustrated in this part. The time complexity of estimating the gravity acceleration algorithm is

Table 5 Huawei Mate 8's experiment data of different transports

Transport	Detection number	Correct number	Accuracy (%)
Bicycle	32	31	96.9
Bus	12	12	100
Metro	12	12	100

Table 6 Nextbit Robin's experiment data of different transports

Transport	Detection number	Correct number	Accuracy (%)
Bicycle	32	32	100
Bus	12	12	100
Metro	12	12	100

Table 7 Power consumption

Device	Power consumption (mW)
Nextbit Robin	57.6
Huawei Mate 8	59.2

$O(n)$. The complexity of calculating the angular velocity and integrating the angle is $O(1)$. The time complexity of using the fast Fourier transform to filter the periodic signal is $O(n \log(n))$. As a result, the time complexity of the whole system is $O(n \log(n))$.

The system needs to be initialized before running. The first 5 s is used to estimate the gravity. Then collect the original data, sampling cycle set to 10 ms. Data of total 12.8 s is needed to eliminate periodic oscillations of human walking. When the system is running, the data collected in 1.28 s will be encapsulated into a frame. Then sum angle of 3 frames and 20 frames is calculated, judging the occurrence of turning.

Record the running time after each sampling through debugging tool. The average time required to run the algorithm is about 57.5 ms.

5.4 Power Consumption

Power monitor [19], which is an energy consumption tracking devices, was used to detect the use of various resources when the application is turned on. In addition, users can also install the PowerTutor on Android device to get an idea of the power consumption. We recorded the power consumption of two devices in different scenarios and calculated their mean values. The results are shown in Table 7. Application's power consumption is less than 60mW during the test according to the data.

6 Conclusion

In this paper, a robust turn detection algorithm based on periodic signal rejection is proposed because the existing algorithm will be influenced by the swing when device put in the pocket or held in the hand. The algorithm could detect the turn of a

variety of transports and the fast turn of person. Fast Fourier transform is applied to identify and eliminate wobble signals caused by walking and multiple sliding windows are applied to adapt to user's turn size and turn speed. Several experiments proved the feasibility of the algorithm and accuracy.

There are some details to continue to optimize later. At present, the algorithm only applies to the fast turn when detecting the turn of person, unable to detect the slow one. In this case, the turn signal is periodic and rather similar to the walking signal. If using the current algorithm, the turn signal will be recognized as the swing caused by walking. The algorithm still has space for improvement.

Acknowledgements This work was supported in part by the National Key Research and Development Program (2016YFB0502004), the National Natural Science Foundation of China (61374214) and the Open Project of the Beijing Key Laboratory of Mobile Computing and Pervasive Device.

References

1. Yiyang L, Fang Z, Wenhua S, Haiyong L (2016) An hidden Markov model based complex walking pattern recognition algorithm. In: 2016 fourth international conference on ubiquitous positioning, indoor navigation and location based services (UPINLBS). IEEE, Nov 2016, pp 223–229
2. Williams B, Kein G, Reid I (2007) Real-time SLAM relocalisation. In: IEEE 11th international conference on computer vision, 2007. ICCV 2007. IEEE
3. Phuyal BP (2002) Turn detection algorithm for vehicle positioning. US Patent 6,502,033[P], 31 Dec 2002
4. Luo H, Zhao F, Jiang M, Ma H, Zhang Y (2017) Constructing an indoor floor plan using crowdsourcing based on magnetic fingerprinting. *Sensors* 17:2678. <https://doi.org/10.3390/s17112678>
5. Misra P, Enge P (2006) Global positioning system: signals, measurements and performance, 2nd edn. Ganga-Jamuna Press, Lincoln, MA, USA
6. Nasir A, Rao KR (1975) Fast fourier transform. In: Orthogonal transforms for digital signal processing. Springer, Berlin, Heidelberg, pp 54–84
7. Phuyal BP (2002) Turn detection algorithm for vehicle positioning. US Patent No. 6,502,033, 31 Dec 2002
8. Giebeler C, Adelerhof DJ, Kuiper AET, Van Zon JBA, Oelgeschläger D, Schulz G (2001) Robust GMR sensors for angle detection and rotation speed sensing. *Sens Actuators A Phys* 91(1):16–20
9. Ozawa H (1992) Angle detection sensor with setting of ratio of magnetic forces of rotating magnet and bias magnet. US Patent No. 5,148,106, 15 Sep 1992
10. Köping L, Ebner F, Grzegorzec M, Deinzer F (2014) Indoor localization using step and turn detection together with floor map information. *FHWS Sci J*, 9
11. Chung J, Donahoe M, Schmandt C et al (2011) Indoor location sensing using geo-magnetism. In: Proceedings of the 9th international conference on Mobile systems, applications, and services. ACM, pp 141–154
12. Marron JJ et al (2016) Multi sensor system for pedestrian tracking and activity recognition in indoor environments. *Int J Ad Hoc Ubiquitous Comput* 23(1–2):3–23
13. Flores GH, Manduchi R, Zenteno ED (2014) Ariadne's thread: Robust turn detection for path back-tracing using the iPhone. In: Ubiquitous positioning indoor navigation and location based service (UPINLBS). IEEE, Nov 2014, pp 133–140

14. Favre J, Jolles BM, Siegrist O, Aminian K (2006) Quaternion-based fusion of gyroscopes and accelerometers to improve 3D angle measurement. *Electron Lett* 42(11):612–614
15. Eren H, Makinist S, Akin E, Yilmaz A (2012) Estimating driving behavior by a smartphone. In: *Intelligent vehicles symposium (IV)*. IEEE, June 2012, pp 234–239
16. Eddy SR (1996) Hidden markov models. *Curr Opin Struct Biol* 6(3):361–365. MLA
17. Hemminki S, Nurmi P, Tarkoma S (2013) Accelerometer-based transportation mode detection on smartphones. In: *Proceedings of the 11th ACM conference on embedded networked sensor systems*. ACM, Nov 2013, p 13
18. Li F, Zhao C, Ding G, Gong J, Liu C, Zhao F (2012) A reliable and accurate indoor localization method using phone inertial sensors. In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, Sept 2012, pp 421–430
19. Power Monitor [EB/OL] (2015) 08 May 2015. <http://www.msoon.com/>. Accessed 18 Apr 2007

The Indoor Localization Algorithm for Combination of Signal Strength and Anti-disturbance



Wanqing Liu, Juqing Zhang, Guanwen Huang, Ge Wang
and Zhen Zhang

Abstract Signal strength of indoor location is easily influenced by environment, multipath refraction and human disturbance. Considering the characters that long distance propagation of signals is more likely to be influenced by more reflection and multipath effects than short distance propagation, an anti-interference indoor localization algorithm is proposed in this paper. Firstly, the RSSI perturbation of long distance is filtered by setting threshold. Then, based on the principle of trilateration locating and selecting weight iteration (IGG), the coordinates are estimated. Finally, a practical example of indoor localization is given. The comparison test calculation is carried out by the positioning software developed independently on the Android platform. The results show that the method presented in this paper improves the location accuracy and has a certain anti-disturbance ability.

Keywords Indoor location · Human disturbance · Anti-disturbance
IGG

W. Liu · J. Zhang (✉) · G. Huang (✉)
College of Geology Engineering and Geomatics, Chang'an University,
Xi'an, Shanxi 710054, China
e-mail: zhangjq@chd.edu.com

G. Huang
e-mail: huang830928@163.com

G. Wang
School of Land Science and Technology, China University
of Geosciences (Beijing), Beijing 100083, China

Z. Zhang
School of Computer Science, Northwestern Polytechnical University,
Xi'an, Shanxi 710129, China

1 Introduction

With the progress of society, a variety of intelligent systems have become the goal and direction of people's pursuit, such as smart home, intelligent buildings, intelligent parking. Positioning technology is the basis of various intelligent systems. At present, including GPS, GLONASS, BEIDOU and GALILEO, GNSS positioning technology has been quite mature, and the application is also very common [1, 2]. However, due to the characteristics of the signal, GNSS is only suitable for outdoor non occlusion areas [3]. Therefore, Because of the weak signal in the room, it is difficult to use method of GNSS to locate, indoor positioning technology has become the focus of research in recent years.

Since people began to study indoor positioning technology, there have been many indoor positioning techniques and algorithms, such as pseudolite, infrared, ultrasound, Bluetooth radio frequency identification, ultra wideband, ZigBee, Wifi and so on [4]. According to the need to measure the distance between nodes in the process of positioning, it can be divided into two categories: ranging method and range-free method [5, 6]. Indoor positioning method based on RSSI is a kind of ranging positioning method, it mainly records the signal strength value from the anchor node. Use signal propagation model to obtain the distance between anchor nodes and blind nodes. On this basis, the measurement edge positioning method is used to calculate the location information of blind nodes [7, 8]. This method has been widely used because of its simple positioning algorithm, low cost and high accuracy. However, the indoor environment is relatively closed, the positioning equipment is easily affected by temperature, humidity, pressure and so on, In addition, signals are susceptible to furniture, room obstructions and personnel walking during the propagation process, which makes the signal multiple reflections and signal unstable, and influences point positioning accuracy. Scholars at home and abroad have put forward many improved algorithms. For example, Shi gives a spatial compensation model, in order to compensate the error caused by the anchor node and the blind node which is not in the same plane [9]. The localization algorithm of indoor mobile targets in RSSI wireless sensor networks is prone to interference and fluctuation, in document 10, the method proposed in this paper is to improve the accuracy of indoor positioning by iteratively iterating the measured data [10]. And Shi established the phase transition matrix, he proposed an improved RSSI multidimensional scaling indoor positioning algorithm [6]. In terms of parameter determination of signal propagation model, Hu and Ni, they fit parameters by BP neural network method. On this basis, six point centroid positioning algorithm is proposed to replace the algorithm of three point location, in order to continue to improve the positioning accuracy [11]. Considering signal propagation characteristics, and long range propagation is more likely to be affected by more reflection and multipath effects than near distance propagation, firstly, the long distance RSSI value is filtered by setting threshold to ensure the quality of distance acquisition, on this basis, the algorithm in this paper is based on the principle of three side positioning, disturbances such as obstacles and personnel walking are

considered as outliers, it is expected to improve the positioning accuracy by using IGG method with strong robust ability [12, 13].

2 RSSI Positioning Principle

2.1 Distance Calculation

There is a Bluetooth signal propagation model [9, 11]:

$$RSSI = A - 10n \cdot \lg(d) \tag{1}$$

where $RSSI$ is the signal strength, A is the signal strength at 1 m from Bluetooth n is a signal attenuation factor, it refers to the attenuation rate of the signal in a particular environment, its value is related to a particular environment, d is the geometric distance between beacon point and Bluetooth point. In the ideal environment, the attenuation of signal intensity follows the log normal distribution about distance [14]. But in the actual environment, due to the influence of multipath effect, there is a large deviation between the estimated distance from the measured value of the measured signal and the real distance.

In order to determine A and signal attenuation factor n , this usually requires signal acquisition in the experimental environment, then the least square principle is used to solve the parameters. Suppose there are m observations:

$$\begin{cases} RSSI_1 = A - 10n \lg(d_1) \\ RSSI_2 = A - 10n \lg(d_2) \\ \vdots \\ RSSI_m = A - 10n \lg(d_m) \end{cases} \tag{2}$$

The corresponding error equation is:

$$V = B\hat{x} - l \tag{3}$$

where:

$$l = [RSSI_1 \quad RSSI_2 \quad \cdots \quad RSSI_m]^T, \quad \hat{x} = [A, n]^T,$$

$$V = [v_1 \quad v_2 \quad \cdots \quad v_m]^T, \quad B = \begin{bmatrix} 1 & -10 \lg(d_1) \\ 1 & -10 \lg(d_2) \\ \vdots & \vdots \\ 1 & -10 \lg(d_m) \end{bmatrix}$$

Parameters A and n can be obtained by the principle of least squares

$$\hat{x} = (B^T PB)^{-1} B^T P l \tag{4}$$

After determining the parameters of propagation model, the distance between two nodes can be obtained by formula (1):

$$d = 10^{\frac{A-RSSI}{10n}} \tag{5}$$

The differential method is used to deal with the above formula, and the use of error propagation can be obtained as follows:

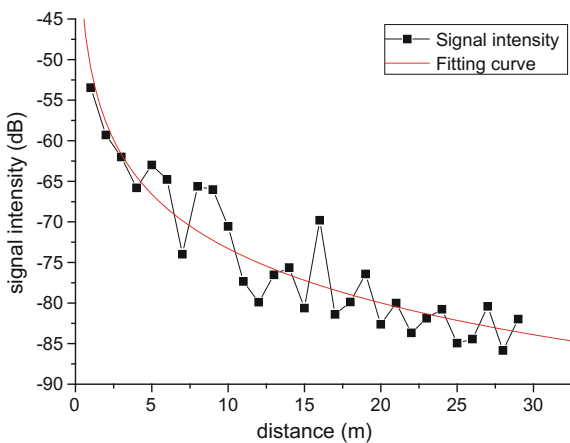
$$m_d^2 = \left(\frac{d \ln(10)}{10n} \right)^2 (m_A^2 + (10n \lg d)^2 m_n^2 + m_{RSSI}^2) \tag{6}$$

m_d represents the distance calculation precision in the formula, m_A , m_n , m_{RSSI} represent the precision of parameters A , n , and signal intensity $RSSI$ respectively.

It can be seen that, the accuracy of m_d^2 is proportional to the distance d between two points, the longer the distance is, the lower the distance precision is obtained by the signal intensity $RSSI$. In order to explain further, the signal acquisition of 30 s was carried out every meter in an experimental environment, the signal intensity value is obtained by filtering the data by Gauss filter. According to formula (2)–(4), the strength and distance relation can be fitted.

As it can be seen from Fig. 1, when A and signal attenuation factors n are determined, the signal intensity gradually slows with the increase of distance. When the distance reaches a certain degree, $RSSI$ will approximate to a certain fixed value, therefore, in order to ensure the precision of distance calculation, a threshold can be set to remove the distance beyond the threshold.

Fig. 1 Related between signal strength and distance



2.2 Positioning Principle

(x_s, y_s) , it is assumed to be blind nodes, (x_i, y_i) , $i = 1, 2, \dots, m$, they are anchor nodes around them, l_i is the plane distance between anchor nodes, (x_i, y_i) , and blind nodes. According to the principle of measurement edge positioning:

$$\begin{aligned}
 (x_s - x_1)^2 + (y_s - y_1)^2 - l_1^2 &= 0 \\
 (x_s - x_2)^2 + (y_s - y_2)^2 - l_2^2 &= 0 \\
 &\vdots \\
 (x_s - x_m)^2 + (y_s - y_m)^2 - l_m^2 &= 0
 \end{aligned}
 \tag{7}$$

In order to eliminate the two term, it is easy to solve the problem, then these $m - 1$ equations are subtracted from the first, these formulas can be properly arranged [15]:

$$\begin{aligned}
 (x_2 - x_1)x_s + (y_2 - y_1)y_s &= [l_1^2 - l_2^2 - (x_1^2 + y_1^2) + (x_2^2 + y_2^2)]/2 \\
 (x_3 - x_1)x_s + (y_3 - y_1)y_s &= [l_1^2 - l_3^2 - (x_1^2 + y_1^2) + (x_3^2 + y_3^2)]/2 \\
 \dots &\dots \\
 (x_m - x_1)x_s + (y_m - y_1)y_s &= [l_1^2 - l_m^2 - (x_1^2 + y_1^2) + (x_m^2 + y_m^2)]/2
 \end{aligned}
 \tag{8}$$

$$\text{Order } G = \begin{bmatrix} (x_2 - x_1) & (y_2 - y_1) \\ (x_3 - x_1) & (y_3 - y_1) \\ \vdots & \vdots \\ (x_m - x_1) & (y_m - y_1) \end{bmatrix} \quad h = \frac{1}{2} \begin{bmatrix} l_1^2 - l_2^2 - (x_1^2 + y_1^2) + (x_2^2 + y_2^2) \\ l_1^2 - l_3^2 - (x_1^2 + y_1^2) + (x_3^2 + y_3^2) \\ \dots \\ l_1^2 - l_m^2 - (x_1^2 + y_1^2) + (x_m^2 + y_m^2) \end{bmatrix}$$

$$\hat{Z} = [\hat{x}_s, \hat{y}_s]^T$$

So the least square principle can be used to solve the problem

$$\hat{Z} = (G^T P G)^{-1} G^T P h \tag{9}$$

P is a weighted matrix corresponding to the compound observation vector h .

3 Trilateration Location Method Based on Robust Estimation

In indoor environment, the signal will be affected by people walking, furniture blocking and so on, these make the signal produce multiple refraction, these affect distance and positioning accuracy. considering the uncertainty of this disturbance,

here we can regard it as an anomaly error, IGG method is used to calculate the coordinate of positioning points, and the effect of signal disturbance on positioning results is expected to be weaken.

Weight selection iteration is a commonly used and effective robust estimation method, it is a method to resist and weaken the influence of outliers on positioning accuracy by reducing the weight of the abnormal observations [12, 16–19]. In this paper, the IGG robust estimation method is adopted, and the specific process is as follows:

- (1) The unknown parameters are solved by using the (9) formula, calculate observation correction.

$$V = G\hat{Z} - h \tag{10}$$

- (2) Calculate equivalent weight matrix, $\bar{P}^k = \text{diag}\{\bar{p}_1^k, \bar{p}_1^k \dots \bar{p}_{m-1}^k\}$

$$\bar{p}_i^k = \begin{cases} p_{ii} & |\tilde{v}_i^k| < 1.5\sigma \\ \frac{p_{ii}}{|\tilde{v}_i^k| + C} & 1.5\sigma \leq |\tilde{v}_i^k| < 2.5\sigma \\ 0 & |\tilde{v}_i^k| \geq 2.5\sigma \end{cases} \tag{11}$$

In the formula [16–19], C is a small amount relative to $|\tilde{v}_i^k|$, \tilde{v}_i^k is the corrected number after standardized treatment, and $\tilde{v}_i^k = \frac{v_i^k}{\sigma\sqrt{Q_{v_i v_i}}}$, $\sigma = \text{med}|v_i|/0.6745$.

- (3) Solving unknown parameter \hat{Z}^k by iteration, until $|\hat{Z}^k - \hat{Z}^{k-1}| < \varepsilon$,

$$\hat{Z}^k = (G^T \bar{P}^k G)^{-1} G^T \bar{P}^k h \tag{12}$$

4 Experimental Test and Result Analysis

In order to facilitate the experiment, we use java to write the positioning algorithm on the Android platform, and install the software on the OnePlus 3T phone, the interface is shown in Fig. 2. In order to verify the effectiveness of the localization algorithm given in this paper, the test calculation was carried out, the details are as follows (Fig. 3).

Select a building corridor as the experimental site, the scene is shown in Fig. 4. Place 10 Bluetooth devices in the corridor, the numbers are 1–10 respectively, the spatial distribution is shown in Fig. 5. S1, S2, S3 and S4 are the blind nodes to be estimated, the corresponding point coordinates of 60 times are calculated by collecting the Bluetooth signal data. In order to verify the effectiveness of the algorithm, the experiment is carried out in the following two schemes.

Fig. 2 Algorithm test software interface



Scheme 1: Normal data acquisition, no human disturbance.

Scheme 2: In the process of data collection, people walk around and interfere with the signal collection.

Before obtaining the coordinates of the unknown point, firstly, we need to determine the parameters of the propagation model in this experimental environment. In this experiment, the 30 s signal acquisition scheme is carried out every one meter. then the propagation parameters are obtained by using the least square estimation, $A = -54.04$, $n = 2.23$. It can be seen from the diagram, when $d > 20$ m, the curve slows down, the threshold value is 20, In the process of location calculation,

Fig. 3 Bluetooth device

the distance value exceeding the value is filtered firstly. On this basis, the least square estimation and IGG two methods are used to solve the two schemes above, the sequence diagram of the calculated positioning error is shown below. The positioning error Δd_i refers to the geometric distance between the true value coordinates of the test points and the calculated coordinates (Figs. 6, 7, 8 and 9).

The statistical results of specific positioning error are shown in Table 1, the mean value $\Delta \bar{d} = \frac{1}{n} \sum_{n=1}^{60} \Delta d_i$, $\sigma_0 = \sqrt{\frac{1}{n} \sum_{n=1}^{60} \Delta d_i^2}$, $\sigma_1 = \sqrt{\frac{1}{n} \sum_{n=1}^{60} (\Delta d_i - \Delta \bar{d})^2}$



Fig. 4 Experimental scene

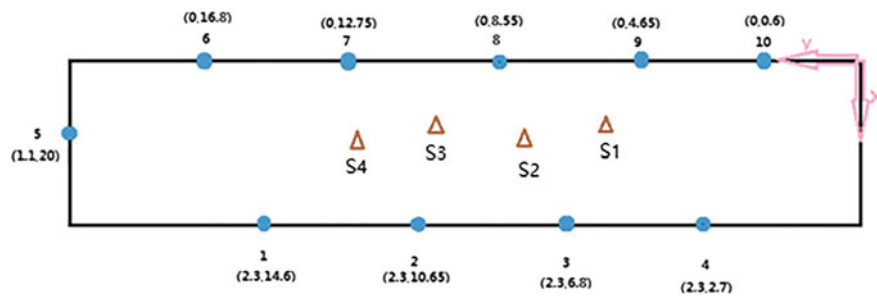


Fig. 5 Bluetooth distribution

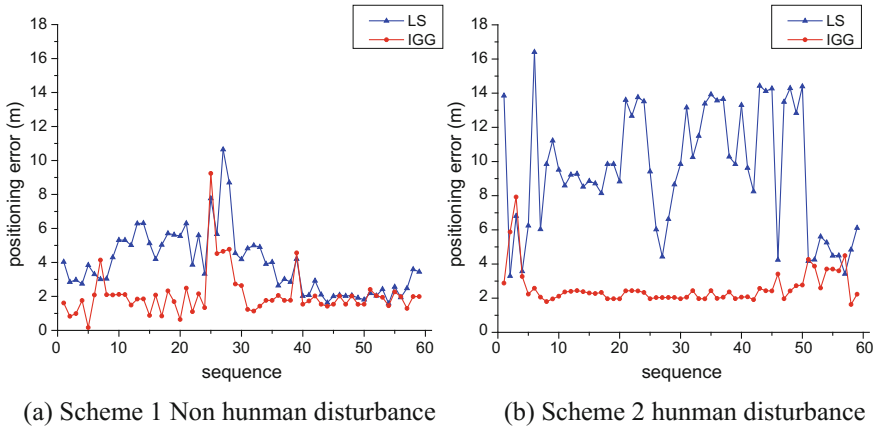


Fig. 6 S1 point positioning error sequence

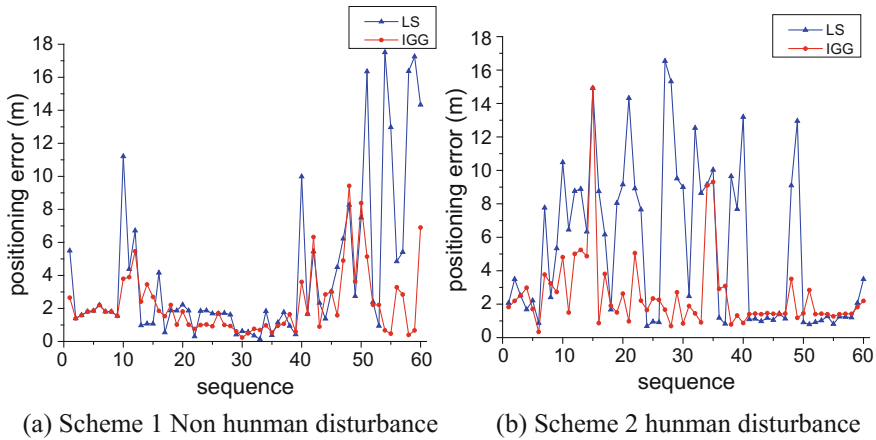


Fig. 7 S2 point positioning error sequence

Analysis of the above figures and tables, we can get the following conclusions:

- (1) Whether there is human disturbance or not, compared to the least square estimation, the positioning accuracy and stability are obviously improved by the IGG estimation method, its positioning precision can be basically controlled within the range of 3 m, standard deviation in the positioning error is basically controlled within 2 m.
- (2) Use LS method to calculate, even without the impact of a human disturbance, the positioning accuracy also has a large jump, the overall positioning accuracy is low, the maximum positioning error of point S3 is nearly 20 m.

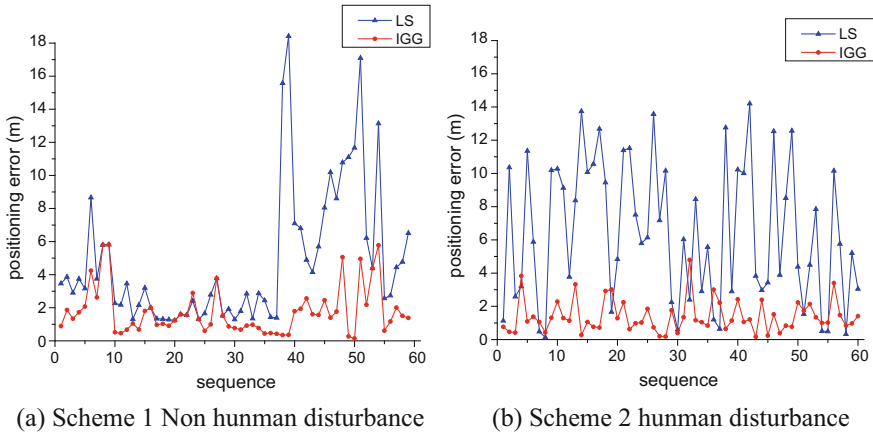


Fig. 8 S3 point positioning error sequence

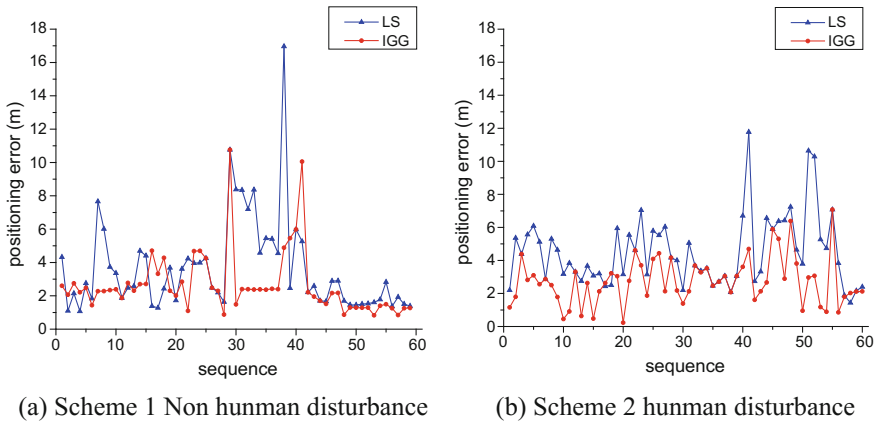


Fig. 9 S4 point positioning error sequence

- (3) Before calculating coordinates by three side location method, the distance of over 20 m is filtered out, so if the perturbation value is over 20 m, this disturbance does not affect the location result, such as point S4, which leads to the illusion that the maximum value of the location error is reduced by the LS method.
- (4) Despite the use of robust method estimation, it improves the positioning accuracy, and the result has relative stability. But there are occasional jumps, for example, the maximum value of S2 point error also reaches 14.92 m (the maximum value of other points is basically below 10 m), which has a certain relationship with the instability of the Bluetooth device itself.

Table 1 The result of different scheme for coordinates calculation

Point number	Scheme category	Method category	Minimum value	Maximum value	Average value	σ_0	σ_1
S1	①	LS	1.59	10.65	3.91	4.31	1.83
		IGG	0.17	9.24	2.08	2.47	1.32
	②	LS	3.30	16.40	9.50	10.17	3.61
		IGG	1.63	7.92	2.59	2.78	1.03
S2	①	LS	0.14	17.51	3.97	6.08	4.57
		IGG	0.24	9.42	2.25	2.96	1.91
	②	LS	0.69	16.54	5.37	7.10	4.61
		IGG	0.34	14.92	2.55	3.48	2.34
S3	①	LS	1.24	18.42	4.78	6.29	4.09
		IGG	0.15	5.82	1.76	2.28	1.45
	②	LS	0.13	14.2	6.41	7.67	4.25
		IGG	0.14	4.79	1.39	1.69	0.98
S4	①	LS	1.07	16.97	3.64	4.58	2.79
		IGG	0.83	10.76	2.651	3.23	1.86
	②	LS	1.45	11.78	4.51	4.98	2.11
		IGG	0.23	7.08	2.73	3.08	1.42

Unit meter

5 Conclusions

Indoor positioning method based on RSSI does not need other auxiliary facilities, and its cost is low, and its algorithm is simple, these advantages make it a major indoor positioning method. However, due to the complex indoor environment, the value of RSSI is easily affected by temperature, humidity, multipath reflection, furniture blocking and personnel walking, and so on, the positioning accuracy is unstable. In this paper, these disturbances are considered as anomalies, On the basis of trilateration location, in this paper, we use a commonly used a robust estimation method—IGG algorithm, an indoor location method for anti- disturbance is presented. Experimental results prove something. An anti-disturbance indoor positioning algorithm based on RSSI is presented in this paper, whether or not there's a human disturbance, the positioning accuracy has been improved, and it has obvious effect in disturbance resistance and stability. Because of the influence of the Bluetooth device's own accuracy and instability, synergetic positioning with other sources is the next step of research.

References

1. Yang Y (2010) Progress, contribution and challenges of Beidou satellite navigation system. *Acta Geodaetica Cartogr Sin* 39(1):1–6
2. Yang Y, Li J et al (2014) Preliminary assessment of the navigation and positioning performance of BeiDou regional navigation satellite system. *Sci China Earth Sci* 57(1):144–152
3. Yang Y (2016) Concepts of comprehensive PNT and related key technologies. *Acta Geodaetica Cartogr Sin* 45(5):505–510
4. Zhao R, Zhong B et al (2014) Overview of indoor localization techniques and applications. *Electron Sci Tech* 27(3):154–157
5. He T, Huang C, Blum BM et al (2003) Range-free localization schemes in large scale sensor networks. In: *Proceedings of the 9th annual international conference on mobile computing and networking (MobiCom)*, San Diego, CA, USA, pp 81–95
6. Xin Shi, Aimin Yin, Xi Chen (2014) RSSI and multidimensional scaling based indoor localization algorithm. *Chin J Sci Instrum* 35(2):261–267
7. Rudafshani M, Datta S (2007) Localization in wireless sensor networks. In: *Information processing in sensor networks (IPSN)*, pp 51–60
8. Yedavalli K, Krishnamachari B, Ravula S et al (2005) Ecolocation: a technique for RF based localization in wireless sensor networks. In: *Proceedings of information processing in sensor networks (IPSN)*, Los Angeles, CA
9. Shi W, Xiong Z, Xu L (2010) In-building RSSI-based user localization algorithm. *Comput Eng Appl* 46(17):232–235
10. Liu D, Xiong Y, Ma J (2011) Exploit kalman filter to improve fingerprint-based indoor localization. In: *2011 international conference on computer science and network technology*, vol 4. IEEE, pp 2290–2293
11. Hu B, Ni X (2017) Improvement of centroid indoor localization algorithm with RSSI. *Comput Technol Dev* 1(7):1–6
12. Yang Y (1994) The principle of equivalent weights—the robust least squares solution for the parameter mode. *Bull Surv Mapp* 1(6):33–35
13. Yang Y, Song L, Xu T (2002) Robust parameter estimation for geodetic correlated observations. *Acta Geodaetica Cartogr Sin* 31(2):95–99
14. Whitehouse K (2002) *The design of Calamari: an ad hoc localization system for sensor networks*. University of California, Berkeley
15. Zhang C (2014) *A Optimization algorithm for indoor positioning based on RSSI*. Northwest University, Xi'an
16. Yang Y (1995) Robust least squares estimation for correlated observations. *Bull Surv Mapp* 1(3):36–38
17. Yang Y (1996) Adaptively robust least squares estimation. *Acta Geodaetica Cartogr Sin* 25(3):206–211
18. Zhang Q, Zhang J, Yue D (2010) *Advanced theory and application of surveying data*. Surveying and Mapping Press, Beijing
19. Zhou J (1989) *Classical error theory and robust estimation*. *Acta Geodaetica Cartogr Sin* 1(18):115–118

A Bluetooth Localization Algorithm Based on Map Path Calibration and Time Series Filtering



Shaomeng Chen, Haiyong Luo, Qu Wang, Wenhua Shao and Fang Zhao

Abstract With the gradual progress of the urbanization process of modern society and the continuous improvement of residents' living standard, people's demand for IPS (Indoor Positioning System) is more urgent than that of the past, higher requirements for positioning accuracy, positioning robustness and positioning energy consumption are put forward. Low-power Bluetooth technology has the characteristics of low-cost, easy deployment and low power consumption, supporting all mainstream smartphone operating systems, and attracting extensive attention from academic and industry field. In this paper, aiming at the situation that current Bluetooth low energy positioning accuracy has large fluctuation under the circumstance of multipath propagation, external disturbance and other factors, and has poor user experience, a Bluetooth localization algorithm based on map path calibration and time series filtering is proposed (MCTF). This algorithm comprehensively utilizes a post-processing method based on the average sliding window, time-series based spatial restraint post-processing method and path information based location calibration method, effectively eliminating the large fluctuation of positioning results caused by multipath propagation and other external wireless

S. Chen (✉) · W. Shao · F. Zhao

School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100083, China
e-mail: chenshaomeng@ict.ac.cn

W. Shao

e-mail: shaowenhua@ict.ac.cn

F. Zhao

e-mail: zfsse@bupt.edu.cn

H. Luo

Research Centre of Pervasive Computing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
e-mail: yhluo@ict.ac.cn

Q. Wang

School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100083, China
e-mail: wangqu@ict.ac.cn

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_29

355

signal interference in complex indoor environment. By using map information, based on the fact that position cannot be transient, utilizing feedback filtering, positioning results can get path correction and direction correction respectively, solving the problem that the positioning coordinates jump between adjacent narrow aisles and jump back and forth on the same path during the positioning process. Extensive experiments performed in shopping mall, building office, exhibition center and other typical indoor environments demonstrate that the average positioning error of the proposed algorithm is within 1.5 m, which has good robustness and ensures the smoothness, accuracy and stability of the positioning result.

Keywords Indoor positioning · Bluetooth location · Map path information
Time series filtering · BLE · Path correction · Direction correction

1 Introduction

In recent years, with the continuous development of science and technology and the gradual improvement of people's living standards, location-based service LBS (Location Based Service) has become the focus of attention and research hot spot, location-aware technology (that is, the positioning technology) as the key technology of LBS, has become the focus of attention from domestic and foreign large commercial companies, research institutes and universities, as well as an important part of national defense, economic construction and social life.

Nowadays, the Global Positioning System (GPS) [1], as a representative of the outdoor positioning technologies, is becoming more and more perfect. As the most mature wireless positioning technology, satellite positioning technology has successfully implemented the popularization and application of commercialization, and has remarkable ability on outdoor positioning. However, due to the high-rise buildings in the city, satellite signal is blocked and difficult to penetrate the reinforced concrete building to capture the indoor people's location information [2, 3], so in the indoor environment, due to various environmental constraints, the traditional positioning technology is difficult to carry out in indoor positioning.

Existing indoor positioning technologies are mainly based on ZigBee [4], Wi-Fi [5], Bluetooth [6], radio frequency identification (RFID) [7], UWB [8], ultrasonic [9], visible light [10, 11] and laser technology [12], among them, technologies which based on ZigBee, RFID, UWB, ultrasonic, and laser need to customize the infrastructure, with the characteristics of high production cost and high system deployment cost which limit the widespread use of the system. Wi-Fi based positioning technology is currently the most popular indoor positioning technology, it's usually based on fingerprint positioning method, according to different sample points intensity and Wi-Fi routing node distribution density, it can provide 2–3 m of positioning accuracy [13, 14], but the Wi-Fi signal in some typical indoor

scenarios such as underground parking lot is not available, the mainstream intelligent operating systems such as the iOS system, has not yet provided API (Application Programming Interface) for scanning indoor Wi-Fi list information [15].

With the continuous development of mobile intelligent terminal technology, smartphone has mature production process and has been successfully popular in people's lives, the arrival of the mobile Internet era gradually changes the way of people's work, life and learning. Intelligent terminal itself has a very powerful information processing capacity, at the same time, a variety of smart sensors built in the terminal have the capability of capturing user's acceleration, walking direction information and environment geomagnetic, air pressure, temperature and humidity, Wi-Fi, Bluetooth and other environmental information. By comprehensively utilizing such kind of information, the mapping relationship between indoor user and the physical space can be established, therefore, based on variety of users, environmental information captured by intelligent terminal, using intelligent terminal equipped with indoor positioning system to get real-time understanding of the user's indoor location is possible.

Compared with other indoor positioning technologies, the positioning technology based on the Bluetooth signal strength has the characteristics of low cost, easy to deploy and low power consumption. The Bluetooth standard works in the 2.4 GHz ISM band, it can avoid crosstalk with the other same frequency band signals by using adaptive frequency modulation technology [16]. This technology can be widely used in various scenes, including underground parking lot and regions where other wireless signals are difficult to form an effective coverage. Android, iOS, Windows phone and other mainstream smartphone operating systems provide an open Bluetooth interface, so this kind of technology can fit with intelligent terminal to achieve accurate positioning.

The main contributions in this paper are that:

- (1) In this paper, a Bluetooth localization algorithm based on map path calibration and time series filtering is proposed. This algorithm comprehensively utilizes a post-processing method based on the average sliding window, time-series based spatial restraint post-processing method and path information based location calibration method, effectively eliminating the large fluctuation of positioning results caused by multipath propagation and other external wireless signal interference in complex indoor environment.
- (2) By using map information, based on the fact that position cannot be transient, utilizing feedback filtering, positioning results can get path correction and direction correction respectively, solving the problem that the positioning coordinates jump between adjacent narrow aisles and jump back and forth on the same path during the positioning process, effectively improving the user experience, positioning accuracy and robustness. A large number of tests show that the proposed algorithm can meet the demand of indoor location with the positioning error within 1.5 m in indoor complex environment.

2 Related Work

Radio frequency identification method is one of the methods first applied to indoor positioning. However, RFID technology can only be applied to short-distance identification communication, so this kind of technology is limited in the application of indoor positioning. Indoor technologies based on ultra-wideband, ultrasonic and laser signals usually use AOA, TOA, TDOA positioning methods [17], but these positioning technologies require specific equipment, not fit with the intelligent terminal, besides, these positioning methods are difficult to use in indoor positioning because of indoor multipath propagation, inaccurate reference clock under short distance and other environmental interference. Wi-Fi positioning technology as a mature wireless positioning technology, has been widely deployed in office buildings, shopping malls, living quarters and other environments, mainly using a fingerprint library to achieve indoor positioning [18], but Wi-Fi positioning technology has relatively low accuracy, the construction of fingerprints is a time-consuming process, due to signal variability and indoor environment layout changes, fingerprints need to be updated regularly which need to invest a lot of manpower and resources.

The positioning methods used in low-power Bluetooth technology are mainly based on the path loss model and the fingerprint-based approaches. Martin et al. [19] based on the path loss model, can achieve 0.53 m positioning accuracy in the indoor open space; Subedi et al. [20] proposed the WCL (Weighted Centroid Localization) algorithm based on the path loss model, and obtained the positioning accuracy of 1.58–2.45 m by using the weight factor associated with the distance between the Bluetooth beacon and the unknown position. Both of the above methods have been simulated in the open space, without taking into account the attenuation of the signals under the multi-path interference and other interfering factors in the complex indoor environment. Peng et al. [21] proposed the Iterative Weighted KNN (IW-KNN) method based on fingerprint to estimate the similarity of signal strength vector by using Euclidean distance and cosine similarity, and obtain an average positioning error of 1.5–2.7 m. Ma et al. [22] proposed BLE signal strength sorting algorithm based on the fingerprint, by using Kendall Tau Correction Coefficient (KTCC), the relationship between position information and multiple low power beacon signal strength ordering, achieving the positioning error of 0.87 m. Similar to the characteristics of Wi-Fi fingerprints, fingerprint construction is a time-consuming process, changes in environmental layout often lead to fingerprint library needs to be rebuilt.

Different from the above-mentioned localization methods, this paper proposes a Bluetooth localization algorithm based on map path calibration and time series filtering. The algorithm comprehensively utilizes a post-processing method based on the average sliding window, time-series based spatial restraint post-processing method and path information based location calibration method, effectively eliminating the large fluctuation of positioning results caused by multipath propagation and other external wireless signal interference in complex indoor environment.

By using map information, based on the fact that position cannot be transient, utilizing feedback filtering, positioning results can get path correction and direction correction respectively, solving the problem that the positioning coordinates jump between adjacent narrow aisles and jump back and forth on the same path during the positioning process, effectively improving the user experience, positioning accuracy and robustness.

3 System Design

The Bluetooth localization algorithm based on map path information calibration and time series filtering proposed in this paper is shown in Fig. 1. The positioning algorithm consists of three modules: data acquisition and preprocessing module, positioning coordinate calculation module and map information based location correction module. The data acquisition and preprocessing module is responsible for Bluetooth beacon data collection in the environment and signal intensity preprocessing of multiple observations of same beacon in the fixed time window, the positioning coordinate calculation module is responsible for completing the initial positioning coordinate calculation and post-processing of positioning coordinates, and the map information based location correction module is responsible for path correction and direction correction of the positioning coordinates respectively.

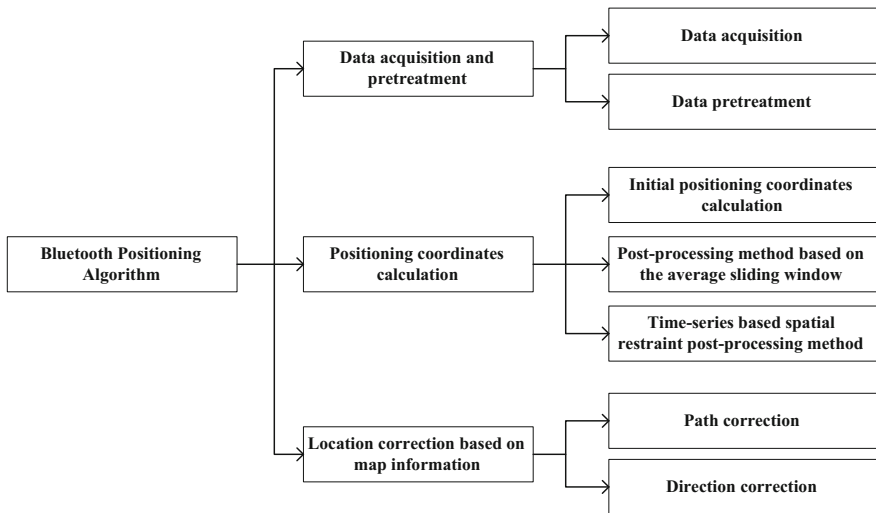


Fig. 1 Positioning algorithm composition module diagram

3.1 Bluetooth Beacon

The Bluetooth beacon specification parameters used in the experiment are shown in Table 1. Bluetooth beacons can broadcast Bluetooth packets at a constant power within a certain range at a fixed frequency, the message contains the Bluetooth beacon's UUID, Major, Minor, MAC address, signal strength and other information, Bluetooth beacon follows the BLE 4.0 standard with very low power consumption, in the case of a button battery, according to different transmitting power and transmission frequency, working hours can reach six months to two years, in the indoor complex environment, such kind Bluetooth beacons can form 25 to 100 m² of effective coverage.

3.2 Intelligent Terminals

The intelligent terminal is used to receive the broadcast message of the Bluetooth beacon, parsing the identifier and signal strength of the Bluetooth beacons, and uploading to the cloud server. Then, the smart terminal obtains and presents the real-time positioning result of the cloud server. The location map is obtained from the cloud server based on identifier matching between the Bluetooth beacon identifier collected in the environment and the beacon identification information on the cloud server.

3.3 Cloud Server

The cloud server is used to communicate with the intelligent terminal, by receiving the locating message information sent by the intelligent terminal, using the Bluetooth location algorithm to calculate the location coordinates, and sending the location result to the locating terminal. The cloud server also stores the location map, the location map description file and the Beacon deployment file. The location map is used to display the topological structure of the location area.

Table 1 Beacon parameter information

Parameters	Value
Core chip	CC2541
Protocol standards	BLE 4.0 standard
Button battery	CR2477H
Transmitting power	-23-4 dBm (configurable)
Launch cycle	100-2000 ms
System support	Android4.3+/iOS7.0+
Transmission distance	Free space < 80 m

Table 2 Path information description

Path number	Start point	End point
1	(x = 1, y = 1)	(x = 2, y = 2)
2	(x = 2, y = 2)	(x = 3, y = 3)
3	(x = 3, y = 3)	(x = 4, y = 4)
...

The location map description file is used to store the path information, the unreachable area information and the item classification information in the location map, the path information description is shown in Table 2. The Beacon deployment file records the coordinates of each beacon in the physical space and the unique identification of the beacon.

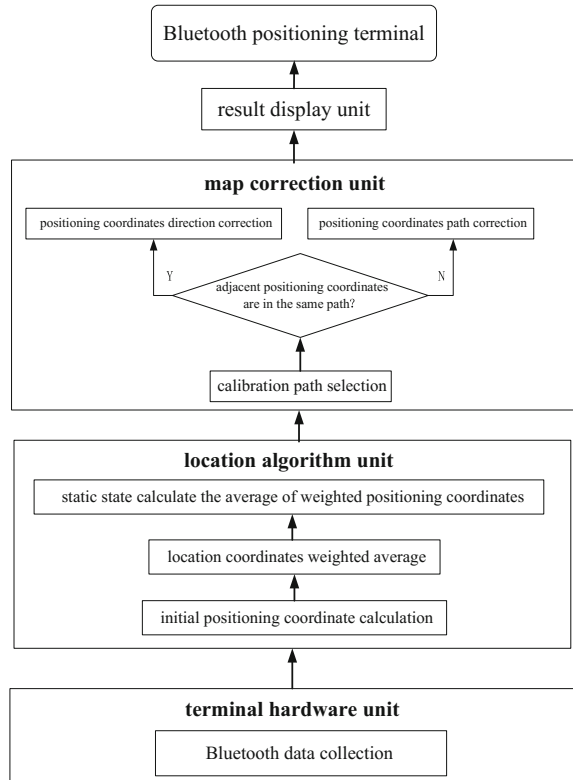
4 Bluetooth Localization Algorithm

Figure 2 shows the system architecture of Bluetooth positioning algorithm based on map path information and time-series filtering, the system consists of four units, which are the Bluetooth beacon data acquisition unit, the location algorithm unit, the map information correction unit and the result display unit. The system operation is divided into two phases: the deployment phase and the positioning stage, in the deployment phase, the Bluetooth beacon is deployed in the positioning scene in advance to generate a Bluetooth beacon deployment file; In the positioning phase, the system first reads the Bluetooth beacon deployment file and the location map description information file, using the terminal Bluetooth scanning function to scan the Bluetooth beacons' broadcast messages in the environment space so as to resolve the beacons' correspondence identifier and the signal strength RSSI information, by using the weighted centroid localization algorithm to obtain the user real-time initial location coordinates, in order to reduce the position coordinate drift, enhance the positioning accuracy and the user experience, the algorithm uses the map path information to modify the location coordinates, the location coordinates correction includes the path correction and the direction correction, the path correction is used to solve the problem of positioning coordinates jump between adjacent narrow aisles, and the direction correction is used to solve the problem of positioning coordinates jump back and forth on the same path during the positioning process.

4.1 Data Acquisition and Pretreatment

The Bluetooth beacon data acquisition unit is responsible for data acquisition. In the experiment, the transmitting frequency of Bluetooth beacon is 10 Hz, for multiple observations of the same Bluetooth beacon in the ϵ time window, the average signal

Fig. 2 System architecture diagram



strength of multiple observations is used as its RSSI value. A queue is used to store the beacon data which consists of beacon identifier and RSSI value, and the queue data is uploaded to the location algorithm unit after ε time, then the storage queue is emptied and the next data storage is performed. In the data acquisition process, the uploading time interval ε can be dynamically adjusted based on the real-time location needs. The Bluetooth beacon data acquisition unit provides raw data for the location algorithm, which is the data source of this system.

4.2 Initial Positioning Coordinates Calculation

The location algorithm unit obtains the data which is preprocessed by the terminal hardware unit, and then performs initial positioning coordinate calculation by using the weighted centroid algorithm. Formula (1) lists the weighted centroid algorithm, for each Bluetooth beacon, using the received Bluetooth beacon signal strength as the weight factor, in order to make the difference between different Bluetooth beacon weights more obvious, the weight factor is set to $1/RSSI^2$.

$$result = \frac{\sum_{i=1}^n \left(\frac{1}{RSSI_i^2} * Coord_i \right)}{\sum_{i=1}^n \frac{1}{RSSI_i^2}} \quad (1)$$

In the above formula, n represents the number of Bluetooth beacons collected in ε time, $RSSI_i$ represents the signal strength of each Bluetooth beacon, and $Coord_i$ represents the position coordinate of the Bluetooth beacon in the physical space.

4.3 Post-processing Method Based on the Average Sliding Window

As shown in Fig. 3, for initial positioning coordinates, the algorithm first conducts post-processing method based on the average sliding window by using two sliding window of length ρ to store the position coordinates respectively, the window length ρ can be adjusted dynamically according to the actual needs, using the initial positioning coordinates window named `initialWindow` to store the initial positioning results, using the weighted positioning coordinates window named `weightedWindow` to store the weighted positioning results. The weights can be set to 0.1, 0.2, 0.2 and 0.5 in the initial positioning coordinate window with length 4. The sliding window is used to save multiple positioning results which can effectively record and use historical location information. By giving different weights to multiple results, that is, giving greater weight to observations closer to the current observation, giving less weight to observations that are father, utilizing the current observation to a greater extent, and using the previous observations to suppress the potential for larger fluctuation so that the positioning results will not be frequent beating, and effectively avoid high positioning error risk caused by large deviation between positioning result and the actual position.

Post-processing method based on the average sliding window

- 1: initialize a sliding window named `initialWindow` with length ρ , used to store the initial positioning results
 - 2: initialize a sliding window named `weightedWindow` with length ρ , used to store the weighted positioning results
 - 3: if `initialWindow.size < ρ` then
 - 4: use the average value of all the positioning coordinates as a weighted positioning result and put it into `weightedWindow`
 - 5: else
 - 6: calculate the weighted positioning result using weighted averages as a weighted positioning result and put it into `weightedWindow`
 - 7: end if
-

Fig. 3 Post-processing method based on the average sliding window

4.4 Time-Series Based Spatial Restraint Post-processing Method

Affected by multipath transmission and other interference factors, the Bluetooth radio frequency signals will produce great volatility, often resulting in the phenomenon that current positioning result relative to the previous positioning result has a large jump, in order to reduce the Bluetooth RF (Radio Frequency) signal large fluctuation which results in positioning results jumping, using the location results of the spatial constraints of the time series to deal with the current positioning result, the processing flow is shown in Fig. 4. For the weighted positioning results in weightedWindow, after each new positioning result is generated, the two adjacent coordinates are compared, and the processing strategy of locating coordinates is as follows: if two position coordinates distance d is greater than the specified threshold θ , the positioning coordinates are moved between two positioning results by one third, the current coordinate is used when the coordinate distance d is less than θ . Then the pedometer output is used to determine whether the user is in a static state or motion state, if the user is in a static state, then data in weightedWindow are averaged. By using this way, the offset phenomenon caused by the difference of two weighted positioning coordinates is effectively prevented, the result shows that the positioning results are smooth and excessive, and the user experience is improved effectively, besides, when the user is in a stationary state, the positioning accuracy can be improved effectively by using larger window to average the location results.

4.5 Location Correction Based on Map Path Information

The map correction algorithm is one of the core units of the whole system which is used to modify the locating results generated by the location algorithm unit. Algorithm execution flow is shown in Fig. 5, the algorithm first generates the path expression according to the starting point and the endpoint coordinates of each path.

Fig. 4 Time-series based spatial restraint post-processing method

Time-series based spatial restraint post-processing method

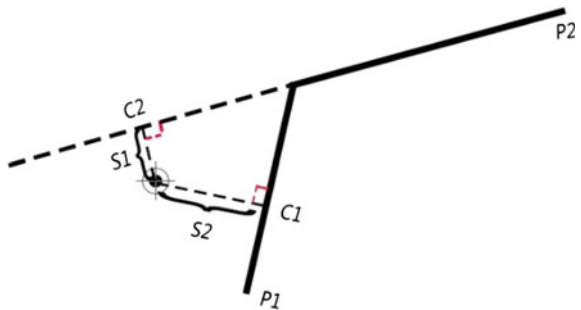
- 1: compare two latest positioning results named curPC and lastPC in weightedWindow
 - 2: get the distance d between curPC and lastPC
 - 3: use finalPC to represent the final coordinate
 - 4: if $d > \theta$ then
 - 5: finalPC = lastPC + (curPC - lastPC) / 3
 - 6: end if
 - 7: if user is in stationary state then
 - 8: finalPC = mean of positioning results in weightedWindow
 - 9: end if
-

Location correction based on map path information

- 1: use the corrPC to represent the coordinate after map path information correction and directionList to store the user's direction on one path
 - 2: calculate the path expression $y = ax + b$ according to the starting and ending points of each path
 - 3: calculate the distance d_i between finalPC and each path in the current map
 - 4: sort d_i in ascending order
 - 5: corrPC = the anchor point in the first nearest path and the anchor point isn't in the extension line of the path at the same time
 - 6: calculate current direction c_i and put it into directionList
 - 7: for two adjacent positioning coordinates p_1, p_2
 - 8: if p_1, p_2 isn't on the same path then
 - 9: if p_1, p_2 isn't on the same path for τ times then
 - 10: change to another path
 - 11: clear the directionList
 - 12: else
 - 13: use last positioning result
 - 14: end if
 - 15: end if
 - 16: if p_1, p_2 is on the same path then
 - 17: use c_1, c_2 to represent two adjacent direction of p_1, p_2
 - 18: if c_1, c_2 isn't on the same direction for κ times then
 - 19: change to another direction
 - 20: else
 - 21: use last positioning result
 - 22: end if
 - 23: end if
-

Fig. 5 Location correction based on map path information

Fig. 6 Path correction diagram



The distance between the weighted positioning coordinate and each path is calculated, then the position coordinate is corrected to the path with nearest distance and the corrected coordinate should not be located on the path's extension line, as shown in Fig. 6, the distances between positioning result and two paths P1, P2 are S1 and S2 respectively, although the logical relationship between the distances is $S1 < S2$, the corrected coordinate C2 is on the extension of path P2, the location coordinate need to be fixed to the C1 position.

After the weighted positioning coordinate is calibrated to the corresponding path, a direction queue is created for storing the forward or backward directions of the adjacent two positioning results, if adjacent two positioning coordinates are located on different paths, path correction for current positioning coordinate is performed, if successive τ times positioning coordinates are not on the same path, then fix to the other path, and clear the directionList, otherwise use the last location result; If the two positioning results are on the same path, perform direction correction to determine whether the two positioning results have the same direction, if yes, return to the current positioning result, otherwise, the direction is corrected if successive κ times direction is different, if less than the κ times, previous positioning result is used.

4.6 Location Map Path Information

Since the positioning coordinates output by the positioning algorithm unit are corrected by the map information in the Bluetooth positioning system, the path information needs to be added to the original map in advance. As shown in Fig. 7, which is a plan view of the seventh floor of the Institute of Computing Technology, Chinese Academy of Sciences, the initial floor plan only shows the layout, there is no corresponding path information, without the amendment of the path information, the positioning results will appear to jump, go through the wall and into the unreachable area and so on. Adding path information on the basis of the original map, which can be used both for map correction in positioning algorithms and for indoor navigation as well.

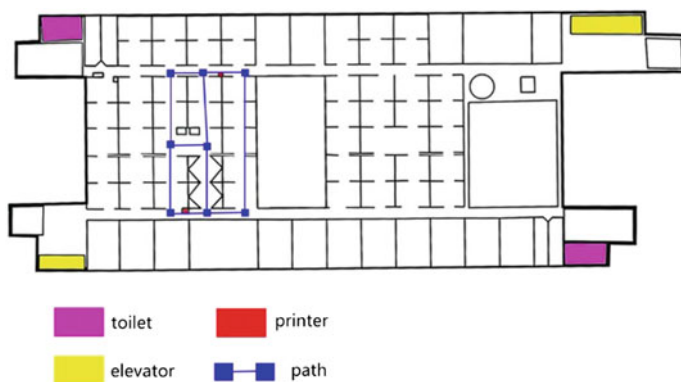


Fig. 7 Location map diagram

5 Experiment and Result Analysis

5.1 Experimental Conditions

The experimental environment is a work area of the seven-storied building of the Institute of Computing Technology, Chinese Academy of Sciences. The environment layout is complex and has many adjacent narrow channels with intensive personnel distribution which is a typical complex indoor scenario. The beacon deployment diagram is shown in Fig. 8. All Bluetooth beacons' transmission frequency is set to 10 Hz, with deployment height of 1.2 m, evenly distributed in $13\text{ m} \times 17\text{ m}$ office area, the distance between adjacent beacons is about 4 m.

The experiment uses the intelligent terminal named HUAWEI Mate8. The test points are distributed as shown in Fig. 9, and 18 test points are evenly distributed in the office environment. The test points are spaced at about 3 m.

5.2 Result Analysis

During each test, several test points are randomly selected, and no less than 20 error samples are taken for each test point. The positioning error between the actual

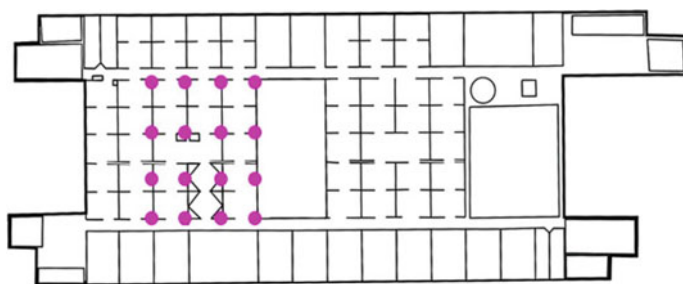


Fig. 8 Beacon deployment diagram

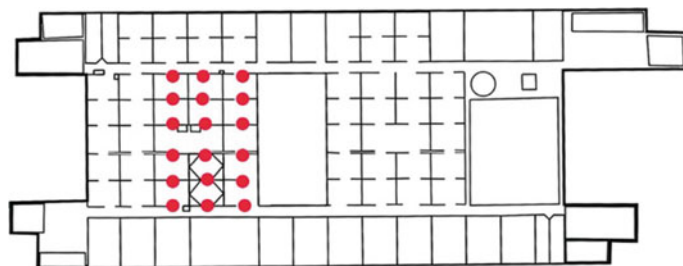
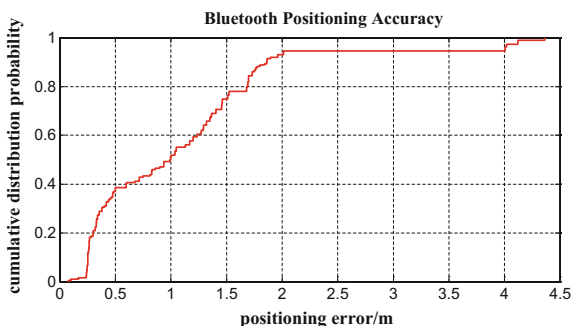


Fig. 9 Sample points diagram

Fig. 10 Positioning error diagram



positioning coordinates and the physical space coordinates is calculated and multiple sampling tests are repeated. The CDF (Cumulative Distribution Function) curve of test results is shown in Fig. 10.

The experimental result shows that in the complex office environment, the algorithm based on map path information and time-series filtering is 80th percentile errors within 1.68 m, and the average positioning error is within 1.5 m, which has good positioning accuracy and robustness.

5.3 Positioning Algorithm Comparison

In order to further verify the effectiveness of the proposed algorithm, we compared with the latest several domestic and international Bluetooth positioning algorithms from the positioning accuracy, the need to build fingerprints and the experimental scenario three dimensions. The comparison result is shown in Table 3.

From Table 3, we can see that compared with the four newly proposed positioning algorithms, the Bluetooth Localization Algorithm (MCTF) based on map path calibration and time series filtering has the advantages of no need to build fingerprints and is suitable for the complex indoor environment. The error is within 1.5 m, which can provide equivalent positioning accuracy as the fingerprint positioning algorithms and meet the positioning needs of complex indoor space environment.

Table 3 Comparison of different positioning algorithms

Algorithm	Accuracy (m)	Fingerprint need	Experimental scenario
MCTF	<1.5	No	Indoor office environment
IW-KNN 0	2.52	Yes	Indoor office environment
WCL [20]	1.58–2.45	No	Free space
ECKP [23]	1.58	Yes	Indoor office environment
KTCC [22]	0.87	Yes	Retail type space

6 Conclusion

This paper proposes a Bluetooth localization algorithm based on map path calibration and time series filtering (MCTF). This algorithm comprehensively utilizes a post-processing method based on the average sliding window, time-series based spatial restraint post-processing method and path information based location calibration method, effectively eliminating the large fluctuation of positioning results caused by multipath propagation and other external wireless signal interference in complex indoor environment. By using map information, based on the fact that position cannot be transient, by using feedback filtering, positioning results can get path correction and direction correction respectively, solving the problem that the positioning coordinates jump between adjacent narrow aisles and jump back and forth on the same path during the positioning process. The average positioning error of algorithm is within 1.5 m, having good robustness. The positioning result has the characteristics of smoothness, accuracy and stability, and compared with the fingerprint based positioning algorithms, it can provide a considerable positioning precision.

It is found that different intelligent terminals have a small fluctuation due to different sensitivity to Bluetooth broadcast message. The next step of our work is to optimize the difference of signal reception between different intelligent terminals, to further deal with the signal intensity of receiving beacon and to try to fuse with other environmental information [24–28].

Acknowledgements This work was supported in part by the National Key Research and Development Program (2016YFB0502004), the National Natural Science Foundation of China (61374214) and the Open Project of the Beijing Key Laboratory of Mobile Computing and Pervasive Device.

References

1. Misra P, Enge P (2006) Global positioning system: signals, measurements and performance, 2nd edn. Ganga-Jamuna Press, Lincoln, MA, USA
2. Al-Ammar MA, Alhadhrami S, Al-Salman A, Alarifi A, Al-Khalifa HS, Alnafessah A, Alsaleh M (2014) Comparative survey of indoor positioning technologies, techniques, and algorithms. In: 2014 international conference on cyberworlds (CW), 6–8 Oct 2014, pp 245, 252
3. Al Nuaimi K, Kamel H (2011) A survey of indoor positioning systems and algorithms. In: 2011 international conference on innovations in information technology (IIT), 25–27 Apr 2011, pp 185, 190
4. Ou CW, Chao CJ, Chang FS et al (2017) A ZigBee position technique for indoor localization based on proximity learning. In: IEEE international conference on mechatronics and automation. IEEE, pp 875–880
5. Kaemarungsi K, Krishnamurthy P (2004) Modeling of indoor positioning systems based on location fingerprinting. Proc IEEE INFOCOM 2(2):1012–1022

6. Rida ME, Liu F, Jadi Y et al (2015) Indoor location position based on bluetooth signal strength. In: International conference on information science and control engineering. IEEE, pp 769–773
7. Xiong R, Waasen SV, Schelten J et al (2016) Development of a quasi-time stretch technology for indoor positioning system based on pulse modulated ultra-high frequency radio. *Sensors*. IEEE, pp 1–4
8. Tiemann J, Schweikowski F, Wietfeld C (2015) Design of an UWB indoor-positioning system for UAV navigation in GNSS-denied environments. In: International conference on indoor positioning and indoor navigation. IEEE, pp 1–7
9. Díaz E, Pérez MC, Gualda D et al (2017) Ultrasonic indoor positioning for smart environments: a mobile application. In: Experiment@ international conference. IEEE, pp 280–285
10. Li L, Hu P, Peng C et al (2014) Epsilon: a visible light based positioning system. In: Usenix conference on networked systems design and implementation. USENIX Association, pp 331–343
11. Wang Q, Luo H, Gao X et al (2017) An indoor three dimensional positioning algorithm based on attitude identification and visible light propagation model
12. Jämsä J, Luimula M, Pieska S et al (2010) Indoor positioning with laser scanned models in the metal industry. In: Ubiquitous positioning indoor navigation & location based service. IEEE, pp 1–9
13. Zhao F, Luo H, Hao G et al (2014) An RSSI gradient-based AP localization algorithm. *China Commun [中国通信(英文版)]* 11(2):100–108
14. Gansemer S, Pueschel S, Frackowiak R et al (2010) Improved RSSI-based Euclidean distance positioning algorithm for large and dynamic WLAN environments. *Computing* 9(1):37–44
15. Thuong NT, Phong HT, Do DT et al (2016) Android application for Wi-Fi based indoor position: system design and performance analysis. In: International conference on information networking. IEEE, pp 416–419
16. “Bluetooth Smart”. <https://www.bluetooth.com/what-is-bluetooth-technology/bluetooth-technology-basics/low-energy>
17. Kumar S, Kumar S, Katabi D (2016) Decimeter-level localization with a single Wi-Fi access point. In: Usenix conference on networked systems design and implementation. USENIX Association, pp 165–178
18. Xu J, Luo H, Zhao F et al (2011) Dynamic indoor localization techniques based on RSSI in WLAN environment. In: International conference on pervasive computing and applications. IEEE, pp 417–421
19. Martin P, Ho BJ, Grupen N et al (2014) An iBeacon primer for indoor localization: demo abstract, pp. 190–191
20. Subedi S, Kwon GR, Shin S et al (2016) Beacon based indoor positioning system using weighted centroid localization approach. In: Eighth international conference on ubiquitous and future networks. IEEE, pp 1016–1019
21. Peng Y, Fan W, Dong X et al (2017) An iterative weighted KNN (IW-KNN) based indoor localization method in bluetooth low energy (BLE) environment. In: Ubiquitous intelligence & computing, advanced and trusted computing, scalable computing and communications, cloud and big data computing, internet of people, and smart world congress. IEEE, pp 794–800
22. Ma Z, Poslad S, Bigham J et al (2017) A BLE RSSI ranking based indoor positioning system for generic smartphones. In: Wireless telecommunications symposium. IEEE, pp 1–8
23. Wang Y, Yang Q, Zhang G et al (2016) Indoor positioning system using Euclidean distance correction algorithm with Bluetooth low energy beacon. In: International conference on internet of things and applications. IEEE, pp 243–247
24. Guo X, Shao W, Zhao F et al (2016) WiMag: multimode fusion localization system based on magnetic/Wi-Fi/PDR. In: International conference on indoor positioning and indoor navigation. IEEE, pp 1–8

25. Wang Q, Luo H, Zhao F et al (2016) An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks. In: International conference on indoor positioning and indoor navigation. IEEE, pp 1–8
26. Luo H, Zhao F, Jiang M, Ma H, Zhang Y (2017) Constructing an indoor floorplan using crowdsourcing based on magnetic fingerprinting. *Sensors* 17:2678
27. Shao W, Zhao F, Wang C et al (2016) Location fingerprint extraction for magnetic field magnitude based indoor positioning. *J Sens* 2016:1–16
28. Zhao F, Luo H, Zhao X et al (2017) HYFI: hybrid floor identification based on wireless fingerprinting and barometric pressure. *IEEE Trans Ind Inform* 99:1

Synthetization of Fingerprint Recognition and Trilateration for Wi-Fi Indoor Localization Through Linear Kalman Filtering



Junxi Tian, Zebo Zhou, Jin Wu, Shuang Du, Changgan Xiang and Changhong Kuang

Abstract In traditional indoor localization, fingerprint localization algorithm fully considers the influences of multipath signals and static obstacles but degrades in case of the changes of observation environment. Trilateration localization has the better robustness to the signal variations but its performance degrades for achieving a certain level of accuracy. In this paper, for the problems above, firstly, a weighted fusion algorithm based on fingerprint recognition algorithm and trilateration algorithm was proposed. Then, adaptive fusion is added to filter the error localization point. Finally, linear Kalman filtering which based on the constant velocity state model assumption is introduced to smooth Wi-Fi localization error. By using the algorithms above, the experimental platform is set up to carry out the localization test. The test result justify that our proposed algorithm has performed better and achieved a better level of accuracy.

Keywords Indoor localization · Trilateration · Fingerprint · Kalman filtering

Z. Zhou (✉)

School of Aeronautics and Astronautics, University of Electronic Science Technology of China (UESTC), Chengdu, People's Republic of China
e-mail: klinsmann.zhou@gmail.com

Z. Zhou

Laboratory of BeiDou Integrated Applications, UESTC, Chengdu, People's Republic of China

J. Tian · J. Wu · S. Du

School of Aeronautics and Astronautics, UESTC, Chengdu, People's Republic of China

C. Xiang · C. Kuang

Sichuan Jiuzhou Beidou Navigation and Location Service Co., Ltd., Chengdu, People's Republic of China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_30

1 Introduction

With the rapid increase of data service and multimedia service, the demand of location-based service increases as well, especially in complex indoor environment. Unfortunately, GNSS fails in indoor environments due to signal attenuations and building blockages etc. [1, 2]. Along with the development of mobile Internet technology, the localization accuracy has become one of the most demanded features. For traditional indoor localization algorithm, the stability is not so satisfied to be applied to the complex environments. For instance, weighted K-nearest neighbour (WKNN) algorithm effectively weakens the influence of multipath propagate and achieves relatively good accuracy [3, 4], but unfortunately, its performance degrades in case of the change of observation environment. On the other hand, with regard to the characteristics of the signal propagation model, trilateration (TRI) method is usually inferior to fingerprint method but has the merit of robustness to the RSSI variations. Therefore, taking complementarity of the both two methods and their error characteristics into account, the combination of two localization methods can be expected to obtain more accurate and reliable localization results [5]. In this paper, a synthetization of fingerprint recognition algorithm and trilateration algorithm is proposed to improve the stability and accuracy of system [6]. According to the experiment, the result from trilateration algorithm has great fluctuation which leads to some error localization points. Therefore, the adaptive fusion is added to the synthetization to filter the error point [7]. Finally, to ensure the accuracy and reliability, the constant velocity state model assumption, different from the information of signal, is introduced to smooth Wi-Fi positioning error [8–10].

The paper is organized as follows: an overview of traditional indoor localization algorithm is presented in Sect. 2. In Sect. 3, a synthetization of fingerprint and trilateration algorithm is proposed and further improved by the adaptive filtering algorithm. The constant velocity state model is additionally established to smooth the Wi-Fi positioning error through a linear Kalman filter (LKF) in Sect. 4. The experiment is carried out to demonstrate the validity of our proposed algorithm compared with the traditional indoor localization algorithm in Sect. 5. Conclusions are drawn in last section.

2 Traditional Indoor Localization Algorithms

The most two popular indoor localization algorithms are weighted K-nearest neighbour (WKNN) and trilateration (TRI). WKNN contains two stages: at the first stage, the Wi-Fi offline localization fingerprint database is created. In the second stage implements, the position matching where the real-time received signal strength indication (RSSI) information is served in the localization algorithm. On the other hand, the RSSI also can be applied to trilateration localization in the sense of least squares (LS) principle. We will discuss that in Sect. 2.2.

2.1 Overview of Fingerprint-Recognition Localization Algorithm

Wi-Fi indoor localization based on the fingerprint recognition mainly involves two stages: one is offline training stage to build a database and the other is online matched positioning stage. The most important part for fingerprint localization is to acquire fingerprint training database. Thus we created a training database and store the measured RSSI values from all the access points (APs) along with their physical locations in this database. The details of database about fingerprint as follows.

Fig. 1 Fingerprint database profile

ID	RSSI	SSID	Var	X	Y
1	(rss1,rss 2...)	(ssid1,ss id2...)	(var1,va r2...)	X1	Y1
2	(rss1,rss 2...)	(ssid1,ss id2...)	(var1,va r2...)	X2	Y2
...
N	(rss1,rss 2...)	(ssid1,ss id2...)	(var1,va r2...)

where ID indicates the number of sampled reference points in database. RSSI information includes the RSSI values from APs with their corresponding SSID. To ensure the accuracy and stability of the signal of referenced points, we measure the wireless signal for 10 times with the sampling interval of 5 s. The horizontal coordinate $\mathbf{r}^F = [x, y]^T$ denotes the ‘true’ location of each fingerprint reference point and N is the number of all fingerprint point.

At online stage, the NN algorithm first calculates Euclidean distance between the user received signal strength (RSS) and the referenced RSS in fingerprint database with

$$d_i^E = \sqrt{\sum_{j=1}^M (RSS_j - rss_{ij})^2} \tag{1}$$

where M is the number of APs. The received signal strength from APs are $(RSS_1 \dots RSS_2 \dots RSS_m)$. rss_{ij} denotes the data from fingerprint database which contains the value of i th fingerprint point receiving signal strength of the j th AP, d_i^E is the Euclidean distance between j th AP and i th fingerprint point. Then the point with the minimum distance in database is selected. Accordingly, its position is coarsely obtained.

Compared to the NN algorithm, KNN algorithm which does not select the minimum Euclidean distance but selects K nearby fingerprint points to generate the mean of physical coordinates included in fingerprint database. Essentially, it makes improvements on NN via the average position.

$$\hat{\mathbf{r}}^{\text{KNN}} = \frac{1}{K} \sum_{k=1}^K \mathbf{r}_k^{\text{F}} \quad (2)$$

where \mathbf{r}_k^{F} is the coordinate of nearby fingerprint point from fingerprint database according to the Fig. 1, and its subscript k represents the index of neighbour fingerprint point. K is the number of neighbourhood fingerprint points. $\hat{\mathbf{r}}^{\text{KNN}}$ is the KNN solution. In Eq. (2), there is a basic presumption that the weighting scalars of the KNN points are the same. However, in most applications, it is much better to increase the weighting value of the nearer points among those K points. Then, the weighting of each point can be calculated by

$$h_k = \frac{\left(\frac{1}{d_k^{\text{E}}}\right)^2}{\sum_{k=1}^K \left(\frac{1}{d_k^{\text{E}}}\right)^2} \quad (3)$$

in which h_k is a weighting scalar. Thus we adjust the Eq. (2) as

$$\hat{\mathbf{r}}^{\text{W}} = \sum_{k=1}^K h_k \cdot \mathbf{r}_k^{\text{F}} \quad (4)$$

where $\hat{\mathbf{r}}^{\text{W}} = [\hat{x}^{\text{W}}, \hat{y}^{\text{W}}]^{\text{T}}$ represents the WKNN estimate vector.

2.2 Trilateration Localization Algorithm

Differing from fingerprint positioning, alternatively, we can also give a distance estimate by Keenan-Motley model according to (5) [11].

$$L = L(d_0) + 10n \lg\left(\frac{d}{d_0}\right) + \sum_{z=1}^Z K_z L_z \quad (5)$$

where $L(d_0)$ is a reference RSS value with the distance d_0 between user and AP. n is the factor of path loss. Z ($1 \leq z \leq Z$) is the number of walls. K_z is the number of walls that signal passes through and L_z is the attenuation factor. L is the observed RSS value with the distance d between user and AP. The empirical model has been widely used in indoor distance measuring. Therefore, according to the RSS, a set of geometrical formulae are generated as follows.

$$d_j = \sqrt{\left(\mathbf{r}_j^{\text{AP}} - \hat{\mathbf{r}}^{\text{TR}}\right)^2} + e_j \quad (6)$$

where d_j ($1 \leq j \leq M$) is the distance between user and the AP point which is obtained from Eq. (5); e_j is the random error subjecting to the Gaussian white noise; \mathbf{r}_j^{AP} is the position vector of the j th AP; $\hat{\mathbf{r}}^{\text{TR}}$ is defined as follows

$$\hat{\mathbf{r}}^{\text{TR}} = [\hat{x}^{\text{TR}}, \hat{y}^{\text{TR}}]^T \tag{7}$$

where $\hat{\mathbf{r}}^{\text{TR}}$ is the estimate vector of TRI; e_j is the random error subjecting to the Gaussian white noise. We solve the nonlinear problem and derive the LS estimate by following equations

$$\hat{\mathbf{r}}_i^{\text{TR}} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{b}$$

$$\mathbf{F} = \begin{bmatrix} 2(\mathbf{r}_M^{\text{AP}} - \mathbf{r}_1^{\text{AP}}) \\ \vdots \\ 2(\mathbf{r}_M^{\text{AP}} - \mathbf{r}_i^{\text{AP}}) \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} (\mathbf{r}_M^{\text{AP}})^2 - (\mathbf{r}_1^{\text{AP}})^2 + d_M^2 - d_1^2 \\ \vdots \\ (\mathbf{r}_M^{\text{AP}})^2 - (\mathbf{r}_{M-1}^{\text{AP}})^2 + d_M^2 - d_{M-1}^2 \end{bmatrix} \tag{8}$$

3 Adaptive Trilateration/Fingerprint Fusion Algorithm

For traditional localization indoors, the single localization algorithm e.g. WKNN and trilateration is usually adopted. This will not achieve high accuracy since each of them has drawbacks in the complex environment. For instance, WKNN will be degraded in the time-variant environment where multipath strongly changes the RSSI of users. On the contrary, the trilateration way is more suffered from the multipath. Therefore, fusion of two different localization algorithm will potentially benefit the localization accuracy and reliability. Here, the architecture of proposed localization system is exhibited in the Fig. 2.

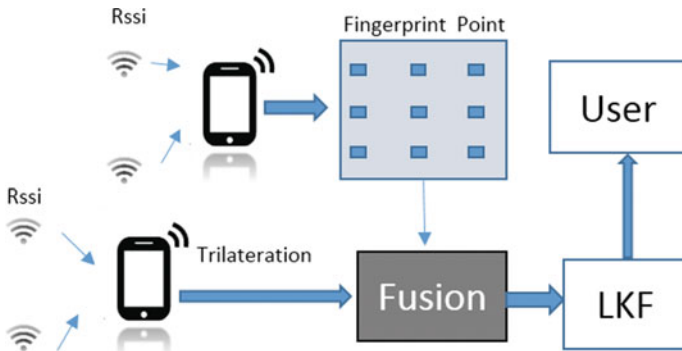


Fig. 2 The architecture of proposed indoor localization system

3.1 Integrated Trilateration/Fingerprint Algorithm

In Sect. 2, two types of localization solutions are derived. However, these two algorithms still have their own drawbacks due to jamming, multipath and the variations of indoor environment. Thus neither of them is always effective for real indoor applications. In this section, with the fingerprint point database information and real-time RSS value, a new indoor localization algorithm is proposed by fusing two indoor localization algorithms. Firstly, the difference value, $\tilde{\mathbf{r}}_i^{\text{TR}}$ and $\tilde{\mathbf{r}}_i^{\text{W}}$ between ‘true’ value and two calculated value, is introduced. The variable i ($1 \leq i \leq N$) is the index number of fingerprint point in database.

$$\begin{aligned}\tilde{\mathbf{r}}_i^{\text{TR}} &= \hat{\mathbf{r}}_i^{\text{TR}} - \mathbf{r}_i^{\text{F}} \\ \tilde{\mathbf{r}}_i^{\text{W}} &= \hat{\mathbf{r}}_i^{\text{W}} - \mathbf{r}_i^{\text{F}}\end{aligned}\quad (9)$$

where \mathbf{r}^{F} is the ‘true’ vector for the i th fingerprint point in fingerprint database. $\hat{\mathbf{r}}^{\text{W}}$ is the estimated vector by WKNN. $\hat{\mathbf{r}}^{\text{TR}}$ is the estimated vector calculated by TRI. $\tilde{\mathbf{r}}^{\text{TR}}$ and $\tilde{\mathbf{r}}^{\text{W}}$ are the difference vector between ‘true’ vector and estimated vector calculated by corresponding algorithm above. Then, the vector $\tilde{\mathbf{r}}^{\text{TR}} = [\hat{x}^{\text{TR}}, \hat{y}^{\text{TR}}]^{\text{T}}$ and $\tilde{\mathbf{r}}^{\text{W}} = [\hat{x}^{\text{W}}, \hat{y}^{\text{W}}]^{\text{T}}$ are used to acquire the weighting:

$$\begin{cases} \mathbf{w}^{\text{TR}} = \frac{\sum_{i=1}^N \frac{((\text{diag}(\hat{\mathbf{r}}_i^{\text{TR}}))^{-1})^2}{((\text{diag}(\hat{\mathbf{r}}_i^{\text{TR}}))^{-1})^2 + ((\text{diag}(\hat{\mathbf{r}}_i^{\text{W}}))^{-1})^2}}{N} \\ \mathbf{w}^{\text{W}} = \frac{\sum_{i=1}^N \frac{((\text{diag}(\hat{\mathbf{r}}_i^{\text{W}}))^{-1})^2}{((\text{diag}(\hat{\mathbf{r}}_i^{\text{TR}}))^{-1})^2 + ((\text{diag}(\hat{\mathbf{r}}_i^{\text{W}}))^{-1})^2}}{N} \end{cases} \quad (10)$$

where:

$$\begin{cases} \mathbf{w}^{\text{TR}} = \begin{bmatrix} w_x^{\text{TR}} & 0 \\ 0 & w_y^{\text{TR}} \end{bmatrix} \\ \mathbf{w}^{\text{W}} = \begin{bmatrix} w_x^{\text{W}} & 0 \\ 0 & w_y^{\text{W}} \end{bmatrix} \end{cases} \quad (11)$$

In Eq. (11), \mathbf{w}^{TR} is the weighting matrix of TRI, while \mathbf{w}^{W} is the weighting matrix of WKNN. Then we can set it as the prior weight matrix for WKNN and TRI fusion. Thereby the fused location solution calculated by weighting trilateration and fingerprint (WTF) becomes:

$$\hat{\mathbf{r}}^S = (\mathbf{w}^W) \cdot \hat{\mathbf{r}}^W + (\mathbf{w}^{TR}) \cdot \hat{\mathbf{r}}^{TR} \quad (12)$$

where

$$\hat{\mathbf{r}}^S = [\hat{x}^S, \hat{y}^S]^T \quad (13)$$

3.2 Adaptive Fusion

The adaptive WKNN/TRI fusion algorithm (AWTF) is introduced for weakening sharp errors caused by the change of the signal power or the environment. Firstly, the fingerprint database information is used to calculating $\tilde{\mathbf{r}}^A$ to filter the sharp errors.

$$\tilde{\mathbf{r}}^A = \mathbf{r}^{TR} - \mathbf{r}^W \quad (14)$$

where $\tilde{\mathbf{r}}^A = (\tilde{x}^A, \tilde{y}^A)^T$ indicates the difference value between the estimate value \mathbf{r}^{TR} and \mathbf{r}^W .

Furthermore, $\tilde{\mathbf{r}}_i^A = (\tilde{x}_i^A, \tilde{y}_i^A)^T$ is defined as the result of AWTF algorithm at i th fingerprint point. Finally, we can set the variance matrix $\tilde{\mathbf{D}}$ as follows.

$$\tilde{\mathbf{D}} = \frac{\sum_{i=1}^N \left(\text{diag}(\tilde{\mathbf{r}}_i^A) - \frac{\sum_{i=1}^N \text{diag}(\tilde{\mathbf{r}}_i^A)}{N} \right)^2}{N} \quad (15)$$

where

$$\tilde{\mathbf{D}} = \begin{bmatrix} D_x & 0 \\ 0 & D_y \end{bmatrix} \quad (16)$$

Defining \mathbf{D}_V as

$$\mathbf{D}_V = \begin{bmatrix} D_x \\ D_y \end{bmatrix} \quad (17)$$

where D_x and D_y are the variance of $\tilde{\mathbf{r}}^A$. Therefore, the estimate of AWTF algorithm as follows.

$$\begin{cases} \hat{\mathbf{r}}^A = \mathbf{w}^W \cdot \hat{\mathbf{r}}^W + \mathbf{w}^{TR} \cdot \hat{\mathbf{r}}^{TR}, & \|\tilde{\mathbf{r}}^A\| < \alpha \|\mathbf{D}_V\| \\ \hat{\mathbf{r}}^A = \mathbf{w}^W \cdot \hat{\mathbf{r}}^W + \mathbf{w}^{TR} \cdot \tilde{\mathbf{D}}^{-1} \cdot \hat{\mathbf{r}}^{TR}, & \|\tilde{\mathbf{r}}^A\| > \alpha \|\mathbf{D}_V\| \end{cases} \quad (18)$$

where $\hat{\mathbf{r}}^A = (\hat{x}^A, \hat{y}^A)^T$ is the estimate vector of AWTF and the parameter α is usually chosen as 1–3.

4 Linear Kalman Filtering

To deal with the problems above and smooth the position error, it is essential to introduce Linear Kalman filtering (LKF) based on dynamic model assumption which could potentially benefit the continuity and accuracy of indoor localization [12, 13].

LKF is an effective algorithm for real-time data processing and deals with noise in short time. The basic concepts of LKF are described by Eqs. (19) and (20):

$$\mathbf{X}_{k+1} = \Phi_{k+1,k} \mathbf{X}_k + \mathbf{W}_k \quad (19)$$

$$\mathbf{Z}_{k+1} = \mathbf{H}_{k+1} \mathbf{X}_{k+1} + \mathbf{V}_{k+1} \quad (20)$$

where the subscript k denotes the epoch index; \mathbf{X} is the state vector. Φ is the state transition matrix. \mathbf{W}_k is the process noise of the k th epoch with $\mathbf{W}_k \sim N(0, \mathbf{Q}_k)$. \mathbf{Q}_k denotes the variance of process noise which be set as a constant. \mathbf{H} is the measurement matrix. \mathbf{Z}_{k+1} is the measurement vector, while \mathbf{V}_{k+1} is the measurement error with $\mathbf{V}_{k+1} \sim N(0, \mathbf{R}_k)$. \mathbf{R}_k is the variance of measurement noise which obtained from the variance matrix of $\hat{\mathbf{r}}^A$. In this paper, the user device equipped on a pedestrian is assumed as constant-velocity motion behaviour. The state vector is defined as $\mathbf{X}_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]$. Therefore, the state transition matrix as follows.

$$\Phi_{k+1,k} = \begin{bmatrix} \mathbf{I}_2 & \Delta t \mathbf{I}_2 \\ \mathbf{0}_{2 \times 2} & \mathbf{I}_2 \end{bmatrix} \quad (21)$$

where \mathbf{I}_2 denotes the two dimensional identity matrix and t is the difference time; Based on the Eq. (18), the position solution of AWTF algorithm can be obtained and further used as the observation information:

$$\mathbf{H} = [\mathbf{I}_2 \quad \mathbf{0}_{2 \times 2}] \quad (22)$$

The initial covariance matrix \mathbf{P}_0 is set as:

$$\mathbf{P}_0 = \mathbf{I}_4 \quad (23)$$

Then the prediction step:

$$\mathbf{X}_{k+1,k} = \Phi_{k+1,k} \mathbf{X}_k \quad (24)$$

$$\mathbf{P}_{k+1,k} = \Phi_{k+1,k} \mathbf{P}_k \Phi_{k+1,k}^T + \mathbf{Q}_k \quad (25)$$

Kalman gain:

$$K_k = P_{k+1,k} H_{k+1}^T R_{k+1}^{-1} \tag{26}$$

The measurement update:

$$X_{k+1} = X_{k+1,k} + K_k [Z_{k+1} - H_{k+1} X_{k+1,k}] \tag{27}$$

$$P_{k+1} = [I - K_k H_{k+1}] P_{k+1,k} \tag{28}$$

5 Experiment and Analysis

The test field is located in the laboratory of the research building in University of Electronic Science and Technology of China (UESTC) and the version of test cellphone is Meizu 3S with Android system. All APs are around the corner of the laboratory regarding the environment condition. The APs deployment detail is depicted in Fig. 3. The antenna of cellphone faces to the up direction for a good signal capture ability. Moreover, to avoid interferences, we closed other electronics devices during the test. The true values of tested points are obtained by designing a planned special trajectory plotted in Fig. 3. The coordinates are precisely

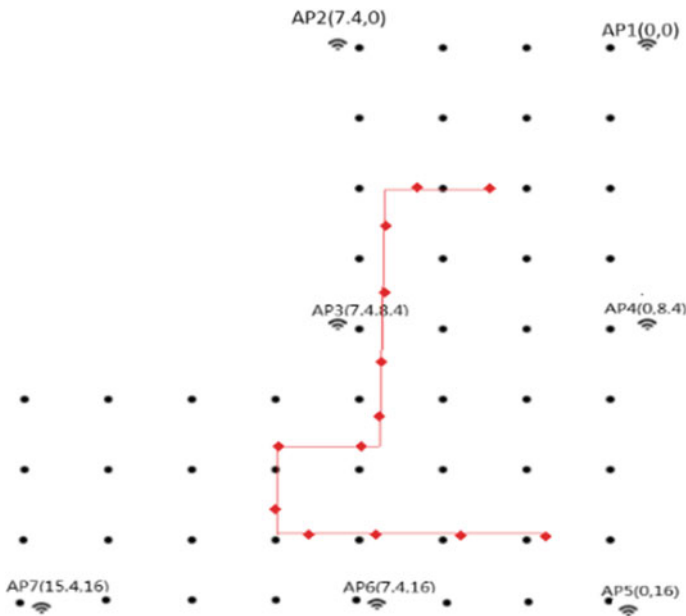


Fig. 3 The true motion trajectory

measured with a tapeline. The grid cell size is set as $2\text{ m} \times 2\text{ m}$. During the test, the RSS is collected in real time with the sampling rate of 2 s . The field test is carried out to illustrate the performances of our proposed algorithm. To ensure the reliability of the RSSI in database, the data is repeatedly sampled for seven times per fingerprint point. The weighting matrix is $\mathbf{w}^W = \text{diag}([0.55, 0.65]^T)$ and $\mathbf{w}^{\text{Tr}} = \text{diag}([0.45, 0.35]^T)$.

Four schemes, i.e. Weighted K-nearest neighbour (WKNN), trilateration (TRI), weighted trilateration and fingerprint (WTF), adaptive weighted trilateration and fingerprint (AWTF) are examined here for comparison purpose. The computed X and Y component errors of all algorithms are shown in Figs. 4 and 5.

It can be seen from that TRI is inferior to WKNN in most of time, whereas shows its merit on some points (see Figs. 4 and 5). The WTF generates the best

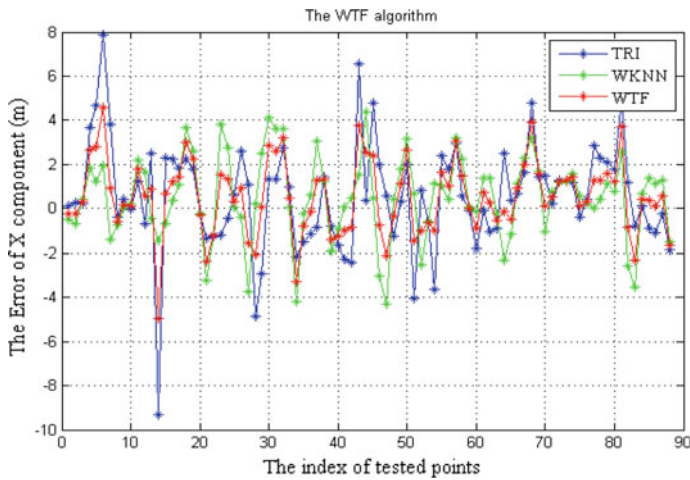


Fig. 4 X component errors of WKNN, TRI and WTF

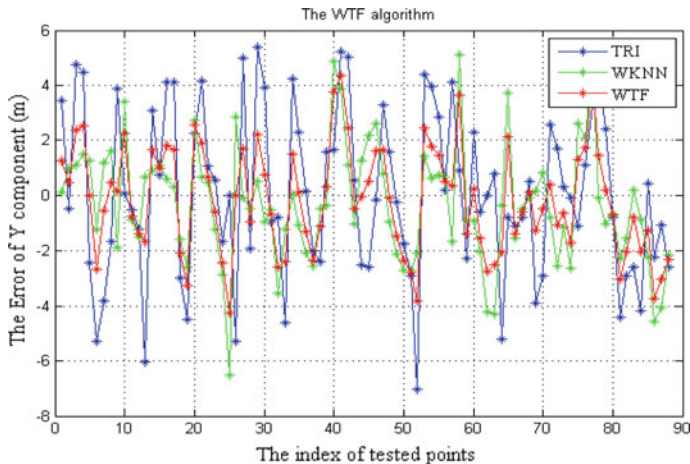


Fig. 5 Y component errors of WKNN, TRI and WTF

result by fusing these two algorithms (see Table 1). If the indoor environment varies, the better result can be expected.

Then, we focus on testing the efficiency of AWTF. The computed results and comparisons are presented in Tables 1 and Figs. 6, 7.

Table 1 The error of WKNN, TRI, WTF and AWTF

	X component				Y component			
	WKNN	TRI	WTF	AWTF	WKNN	TRI	WTF	AWTF
Max	4.36	9.3	4.19	3.68	6.53	7.01	4.56	4.24
Min	0.21	0.11	0.08	0.01	0.12	0.13	0.15	0.11
Mean	1.53	1.81	1.36	1.32	1.71	2.47	1.61	1.61

Unit m

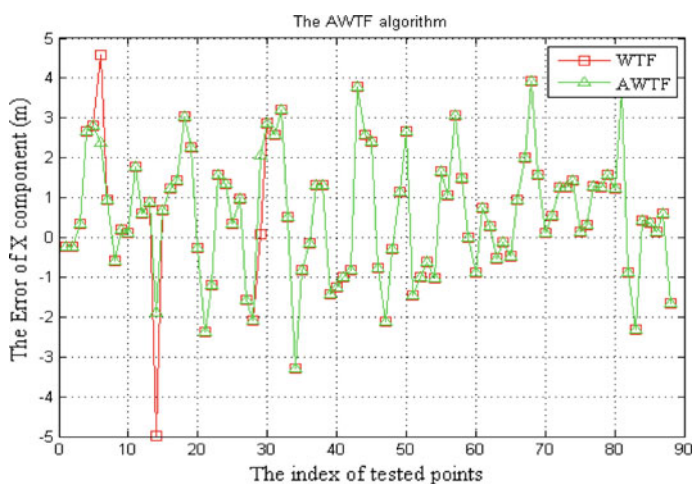


Fig. 6 The X component error of WTF and AWTF

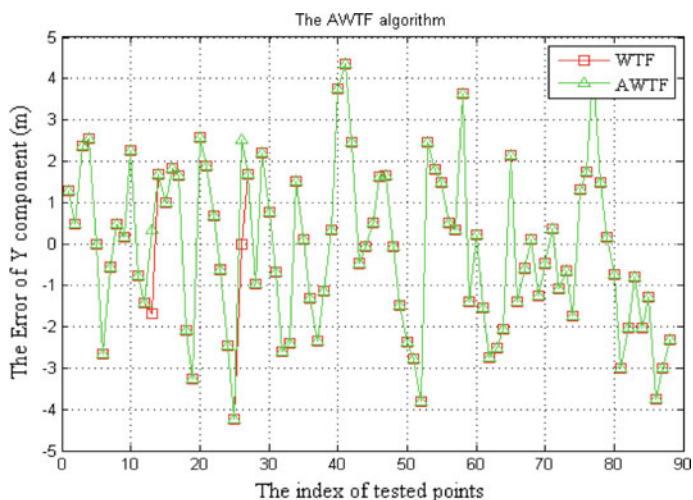


Fig. 7 The Y component error of WTF and AWTF

In Figs. 6 and 7 above, the sharp errors are successfully minimized. The overall accuracy is be better (see Table 1). This proves the effectiveness of AWTF algorithm.

Finally, to verify the contribution of dynamic model, we walked following the designed route and the cellphone located once every two seconds. The constant velocity is about 1 m/s and the whole walking path includes 13 collected points. Seen from Table 2 and Figs. 8, 9, LKF effectively smoothes the positioning result of AWTF. The results above three algorithms are very close. The LKF achieves the best result compared with WTF and AWTF. Furthermore, for the aspect of stability, LKF greatly smoothes the amplitude of noise thus generates a more reliable estimate compared with other two algorithms (see Figs. 8 and 9).

Table 2 The error of WTF, AWTF and LKF

	X component			Y component		
	WTF	AWTF	LKF	WTF	AWTF	LKF
Max	4.19	3.68	3.61	4.56	4.24	3.62
Min	0.18	0.11	0.04	0.15	0.11	0.02
Mean	1.36	1.32	1.31	1.61	1.61	1.20

Unit m

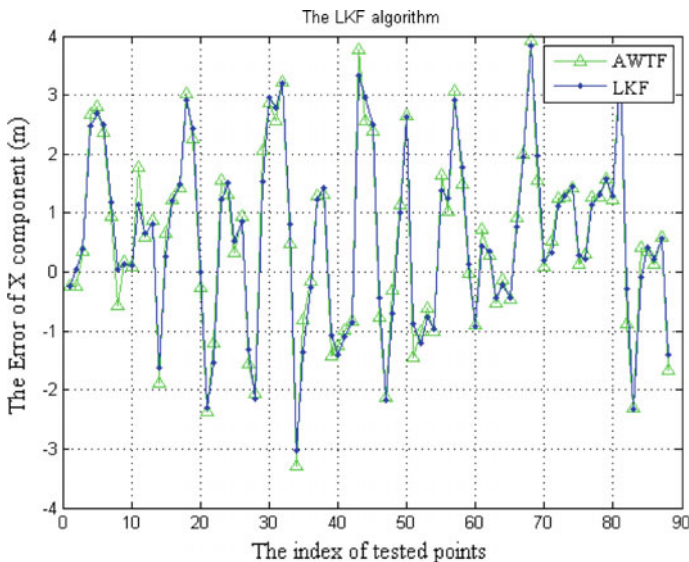


Fig. 8 The X component error of AWTF and LKF

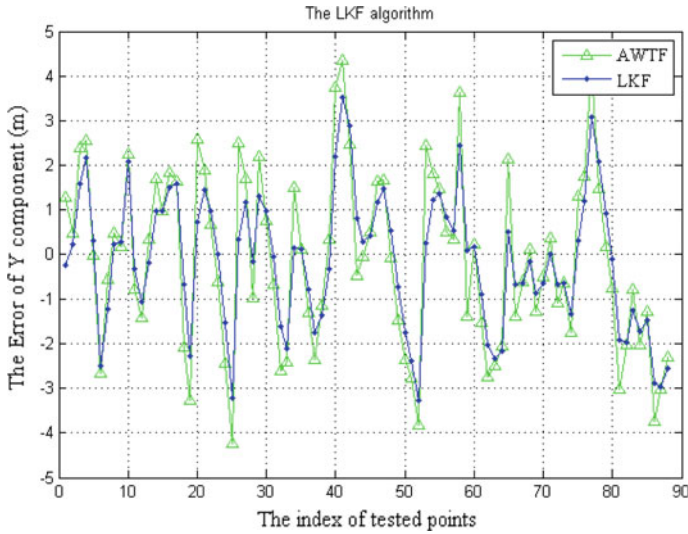


Fig. 9 The Y component error of AWTF and LKF

6 Concluding Remark

In summary, the fusion of weighted K-nearest neighbor (WKNN) and trilateration (TRI) algorithm is proposed for indoor localization algorithm and is further improved through a linear Kalman filter (LKF). Two main contributions are proposed: (1) the positioning based on the adaptive fusion of trilateration and fingerprint localization algorithms. (2) The dynamic model is also established and subsequently used in a LKF to smooth the noise and error peaks. LKF is used as dynamic model for estimating the position, while the measurement value is given by the result of Adaptive filtering and the synthetization of two algorithms above. The result show that: the fusion of TRI and WKNN effectively improves the accuracy compare with each one. By constructing the dynamic model and observation model through the LKF, the system can achieve more accurate and reliable results.

Acknowledgements This work is supported by the National Natural Science Funds of China (No. 41604025 and 41704029), Sichuan Province Science and Technology Project (No. 2016GZ0062), the State Key Laboratory of Geodesy and Earth’s Dynamics (Institute of Geodesy and Geophysics, CAS) Grant No. SKLGED2018-3-2-E and the Key Laboratory of Precise Engineering and Industry Surveying of National Administration of Surveying, Mapping and Geoinformation (PF2015-11).

References

1. Farivar R, Wiczer D, Gutierrez A, Roy H (2009) A statistical study on the impact of wireless signals' behavior on location estimation accuracy in 802.11 fingerprinting systems. In: IEEE international symposium on parallel & distributed processing, Rome, Italy, pp 1–8, 23–29 May
2. Zhou Z, Shen Y, Li B (2011) Moving time-window based real-time estimation algorithm for the stochastic model of GPS/Doppler navigation. *Acta Geodaetica Cartogr Sin* 40(2):220–225
3. Roshanaei M, Maleki M (2009) Dynamic-KNN: a novel locating method in WLAN based on angle of arrival. In: IEEE symposium on industrial electronics and applications, Kuala Lumpur, Malaysia, pp 722–726, 4–6 Oct
4. Liang P, Krause P (2016) Smartphone-based real-time indoor location tracking with 1-m Precision. *IEEE J Biomed Health Inf* 20(3):756–762
5. Zhang M, Shen W, Yao Z, Zhu J (2016) Multiple information fusion indoor location algorithm based on WIFI and improved PDR. In: Chinese control conference, Chengdu, China, pp 5086–5092, 27–29 July
6. Biswa J, Veloso M (2010) WiFi localization and navigation for autonomous indoor mobile robots. In: IEEE international conference on robotics and automation anchorage convention district, Anchorage, Alaska, USA, pp 4379–4384, 3–7 May
7. Yang Y, Gao W (2006) An optimal adaptive Kalman filter. *J Geodesy* 80(4):177–183
8. Chen B, Liu R, Chen Y, Liu J, Jiang X, Liu D (2015) WIFI fingerprint based self-adaptive indoor location in the dynamic environment. *Chin J Sensors Actuators* 28(5):729–738
9. Joydeep B, Manuela V (2010) WiFi localization and navigation for autonomous indoor mobile robots. IEEE international conference on robotics and automation anchorage convention district, vol 58(8), pp 4379–4384
10. Koch KR, Yang Y (1998) Robust Kalman filter for rank deficient observation models. *J Geodesy* 72(7):436–444
11. Zhou Z, Li Y, Rizos C, Shen Y (2009) A robust integration of GPS and MEMS-INS through trajectory-constrained adaptive Kalman filtering. In: Proceedings of international technical meeting of the satellite division of the Institute of Navigation, vol 5538(1), pp 995–1003
12. Xu J, Song E, Luo Y, Zhu Y (2012) Optimal distributed Kalman filtering fusion algorithm without invertibility of estimation error and sensor noise covariances. *IEEE Signal Process Lett* 19(1):55–58
13. Zhou Z, Shen Y, Li B (2010) A windowing-recursive approach for GPS real-time kinematic positioning. *GPS Solutions* 14(4):365–373
14. Wu J, Zhou Z, Chen J, Fourati H, Li R (2016) Fast complementary filter for attitude estimation using low-cost MARG sensors. *IEEE Sens J* 16(18):6997–7007

Positioning Performance Analysis of Combined BDS/GPS-RTK Based on iGMAS Products



Subo Wan, Qianxin Wang, Wenwen She and Mingbin Zhang

Abstract With the rapid development of the self-determination navigation system in many countries, multi-band multi-mode GNSS integrated positioning has become the current research hotspot, which can overcome the limitations that the navigation based on single satellite system would be not continuous and reliable while GNSS signal is sheltered seriously. According to the development status of GNSS navigation positioning system in China in recent years, the paper intended to study on the performance of combined BDS/GPS-RTK with the application of products of iGMAS. The result of static experiment showed that the positioning accuracy of single GPS was better than single BDS and combined BDS/GPS in open area, combined BDS/GPS was better than single BDS and single GPS in enclosed area. The result of dynamic vehicle experiment showed that combined BDS/GPS could still provide continuous and stable positioning service although the positioning reliability was greatly affected when the observation environment was bad. The result of antenna movement experiment showed that the positioning accuracy of single GPS and combined BDS/GPS was almost the same in either vertical or horizontal direction in open area and single GPS was better than combined BDS/GPS in enclosed area.

Keywords iGMAS · Combined BDS/GPS · RTK · Positioning performance

1 Introduction

iGMAS is the abbreviation of International GNSS Monitoring & Assessment Service, whose basic functions are the acquisition, transmission, storage, analysis of data and the release of information. Its main purpose is to provide users with secure

S. Wan · Q. Wang (✉) · M. Zhang
School of Environment Science and Spatial Informatics,
China University of Mining and Technology, Xuzhou 221116, China
e-mail: wqx@cumt.edu.cn

W. She
Jiangsu Surveying and Mapping Product Quality Supervision and Inspection station,
Nanjing 210013, China

and reliable service of GNSS. The website of iGMAS is <http://124.205.50.178>. Users can download the products of orbit and clock correction with high precision or other products freely from it.

Along with the development of autonomous satellite navigation systems, the number of GPS satellites in orbit has been 32 by December 2017. The number of China's BDS satellites in orbit has reached 19 and there are 5 GEO satellites, 7 MEO satellites and 7 IGSO satellites among them. The GALILEO navigation system of European Union has 8 orbiting satellites in addition to the experimental satellites. The number of GLONASS satellites in orbit has been 24 up to now after the rescue plan which is aimed at the supplement of satellites. It can be seen that the combination and cooperation of multi satellite navigation systems will be the trend in the development of global navigation satellite system (GNSS).

At present, the research of combined BDS/GPS-RTK positioning is mainly focused on two aspects. On one hand, the research on mathematical model, such as literature [4, 6] studied on the observation equation which is combined by BDS and GPS in RTK positioning, literature [9] aimed at improving the Kalman filter algorithm of BDS/GPS integrated positioning system. On the other hand, the analysis of position performance in combined BDS/GPS-RTK [7, 8]. Although there is plenty of research on positioning algorithm, the experimental part mostly concentrates on the analysis of static data instead of dynamic data. In addition, most of the data and products are downloaded from IGS, there is no significance for us to verify the precision and reliability of the products provided by iGMAS. Therefore, the position performance of combined BDS/GPS-RTK will be analyzed in this essay according to the products of iGMAS.

2 Mathematical Model

2.1 GNSS Positioning Observation Equation

RTK mathematical model is based on the observation equation of pseudo range and carrier phase [2], as follows:

$$\widetilde{\rho}_i^p(t) = \rho_i^p(t) + c\delta t_i(t) - c\delta t^p(t) + \delta\rho_{i,trop}^p(t) + \delta\rho_{i,iono}^p(t) + \varepsilon_\rho \quad (1)$$

$$\phi_i^p(t) = \frac{f}{c}\rho_i^p(t) + f\delta t_i(t) - f\delta t^p(t) + \frac{f}{c}\delta\rho_{i,trop}^p(t) - \frac{f}{c}\delta\rho_{i,iono}^p(t) - N_0^p + \varepsilon_\phi \quad (2)$$

The double difference observation equation of pseudorange and carrier phase can be expressed as follows:

$$DD_{ij,\rho}^{pq}(t) = \nabla\Delta\rho_{ij}^{pq}(t) + \nabla\Delta\rho_{ij,trop}^{pq}(t) + \nabla\Delta\rho_{ij,iono}^{pq}(t) + \varepsilon_{ij,\rho}^{pq} \quad (3)$$

$$DD_{ij,\varphi}^{pq}(t) = \frac{f}{c}\nabla\Delta\rho_{ij}^{pq}(t) + \frac{f}{c}\nabla\Delta\rho_{ij,trop}^{pq}(t) - \frac{f}{c}\nabla\Delta\rho_{ij,iono}^{pq}(t) - \nabla\Delta N_{ij}^{pq} + \varepsilon_{ij,\varphi}^{pq} \quad (4)$$

where $\widetilde{\rho}_i^p(t)$ and $\varphi_i^p(t)$ represent the pseudorange observation and carrier phase observation, $\rho_i^p(t)$ is the geometric distance from satellite p to station i , $\delta t_i(t)$ is the receiver clock error of station i , $\delta t^p(t)$ is the clock error of satellite p , $\delta\rho_{i,trop}^p(t)$ and $\delta\rho_{i,iono}^p(t)$ are the delay errors of troposphere and ionosphere, ε_ρ and ε_φ denote the residual errors of pseudorange and carrier phase, $DD_{ij,\rho}^{pq}(t)$ and $DD_{ij,\varphi}^{pq}(t)$ represent double difference observations of pseudorange and carrier phase, $\nabla\Delta$ is the difference operator. Among them, delay errors of troposphere and ionosphere can be well eliminated in short-baseline RTK.

The matrix form of observation equation in BDS/GPS combinations can be expressed as:

$$\begin{bmatrix} \mathbf{A}_B & 0 & 0 \\ \mathbf{A}_G & 0 & 0 \\ \mathbf{A}_B & \mathbf{B}_B & 0 \\ \mathbf{A}_G & 0 & \mathbf{B}_G \end{bmatrix} \begin{bmatrix} d\mathbf{X} \\ \nabla\Delta\mathbf{N}_B \\ \nabla\Delta\mathbf{N}_G \end{bmatrix} = \begin{bmatrix} \nabla\Delta\mathbf{l}_{\rho B} \\ \nabla\Delta\mathbf{l}_{\rho G} \\ \nabla\Delta\mathbf{l}_{\varphi B} \\ \nabla\Delta\mathbf{l}_{\varphi G} \end{bmatrix} \quad (5)$$

$$\mathbf{A} = \begin{bmatrix} \frac{X_0 - X^1(t)}{\rho_0^1(t)} & \frac{Y_0 - Y^1(t)}{\rho_0^1(t)} & \frac{Z_0 - Z^1(t)}{\rho_0^1(t)} \\ \vdots & \vdots & \vdots \\ \frac{X_0 - X^n(t)}{\rho_0^n(t)} & \frac{Y_0 - Y^n(t)}{\rho_0^n(t)} & \frac{Z_0 - Z^n(t)}{\rho_0^n(t)} \end{bmatrix} \quad (6)$$

$$\mathbf{B} = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix} \quad (7)$$

where $d\mathbf{X}$ is the correction of coordinate, $\nabla\Delta\mathbf{N}$ is the double difference ambiguity, the subscript B represents BDS, the subscript G represents GPS, (X_0, Y_0, Z_0) is the approximate coordinates of receiver, $(X^n(t), Y^n(t), Z^n(t))$ is the coordinates of satellite, $\rho_0^n(t)$ is the geometric distance from satellite to station, λ represents the wavelength, t represents the epoch [5, 10].

2.2 Adaptive EKF Algorithm

The state equation and observation equation of Extended Kalman Filter (EKF) are written as:

$$\mathbf{X}_k = \Phi_{k,k-1}\mathbf{X}_{k-1} + \Gamma_{k,k-1}\mathbf{u}_{k-1} + \mathbf{W}_{k-1} \quad (8)$$

$$\mathbf{Z}_k = h(\mathbf{X}_k) + \mathbf{V}_k \quad (9)$$

The iterative algorithm of EKF can be shown as follows:

$$\widehat{\mathbf{X}}_{k,k-1} = \Phi_{k,k-1}\widehat{\mathbf{X}}_{k-1} + \Gamma_{k,k-1}\mathbf{u}_{k-1} \quad (10)$$

$$\mathbf{P}_{k,k-1} = \Phi_k\mathbf{P}_{k-1}\Phi_k^T + \mathbf{Q}_{k-1} \quad (11)$$

$$\widehat{\mathbf{V}}_k = \mathbf{Z}_k - h(\widehat{\mathbf{X}}_{k,k-1}) \quad (12)$$

$$\mathbf{P}_{\widehat{\mathbf{V}}_k} = \mathbf{H}_k\mathbf{P}_{k,k-1}\mathbf{H}_k^T + \mathbf{R}_k \quad (13)$$

$$\mathbf{K}_k = \mathbf{P}_{k,k-1}\mathbf{H}_k^T \left[\mathbf{P}_{\widehat{\mathbf{V}}_k} \right]^{-1} \quad (14)$$

$$\mathbf{X}_k = \widehat{\mathbf{K}}_{k,k-1} + \mathbf{K}_k\widehat{\mathbf{V}}_k \quad (15)$$

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\mathbf{P}_{k,k-1} \quad (16)$$

where \mathbf{X}_k is the state vector, \mathbf{Z}_k is the observation vector, $\Phi_{k,k-1}$ is the state transition matrix, $\Gamma_{k,k-1}$ is the dynamic noise transition matrix, \mathbf{u}_{k-1} is the control vector, $h(\cdot)$ is the nonlinear function, \mathbf{W}_{k-1} and \mathbf{V}_k represent the dynamic noise and observation noise separately, they conform to the normal distribution $p(w) \sim (\mathbf{0}, \mathbf{Q}_{k-1})$ and $p(v) \sim (0, \mathbf{R}_k)$.

The state vector can be expressed as:

$$\mathbf{X}_k = [x_k \quad y_k \quad z_k \quad \dot{x}_k \quad \dot{y}_k \quad \dot{z}_k \quad \ddot{x}_k \quad \ddot{y}_k \quad \ddot{z}_k \quad \nabla\Delta N_1 \quad \dots \quad \nabla\Delta N_n] \quad (17)$$

The state transition matrix can be expressed as:

$$\Phi_{k,k-1} = \begin{bmatrix} \mathbf{I}_3 & dt\mathbf{I}_3 & \frac{dt^2}{2}\mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_3 & dt\mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_n \end{bmatrix} \quad (18)$$

where \mathbf{I}_3 is 3×3 identity matrix.

The coefficient matrix can be shown as follows:

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{A}_{\rho B} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{\rho G} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{\varphi B} & \mathbf{B}_B & \mathbf{0} \\ \mathbf{A}_{\varphi G} & \mathbf{0} & \mathbf{B}_G \end{bmatrix} \quad (19)$$

$$\mathbf{A}_{\rho B} = \mathbf{A}_{\varphi B} = \mathbf{A}_{\rho G} = \mathbf{A}_{\varphi G} = \begin{bmatrix} \frac{X_0 - X^1(t)}{\rho_0^1(t)} & \frac{Y_0 - Y^1(t)}{\rho_0^1(t)} & \frac{Z_0 - Z^1(t)}{\rho_0^1(t)} & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{X_0 - X^n(t)}{\rho_0^n(t)} & \frac{Y_0 - Y^n(t)}{\rho_0^n(t)} & \frac{Z_0 - Z^n(t)}{\rho_0^n(t)} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{20}$$

$$\mathbf{B}_B = \mathbf{B}_G = \begin{bmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda \end{bmatrix} \tag{21}$$

The \mathbf{R}_k matrix of measurement equation uses the altitude angle model, that is $\sigma^2 = \sigma_0^2 / \sin^2(E)$, the state equation \mathbf{Q}_k can be determined by IAE and RAE algorithm.

3 Positioning Results and Accuracy Analysis of Combined BDS/GPS-RTK

In order to test the performance of combined BDS/GPS-RTK, The static experiment, dynamic vehicle experiment and antenna movement experiment have been carried out. The data of reference station in these three experiments are from the Continuously Operating Reference Station (CORS) of China Mining University (CUMT). The model of CORS is ThumbT300 and the model of moving station is Trimble R10. By the test, the positioning performance of receiver is reliable and accurate. The products required for experiments are all from iGMAS, and the accuracy of the products has been analyzed in detail in literature [1, 3] which will not be referred to again in this paper.

3.1 Static Experiment

3.1.1 Data Collection

The experiment used observation of multi-GNSS from moving station on the roof and balcony of School of Environment Science and Spatial Informatics, CUMT. As Fig. 1a showed, the roof belonged to the open area where the observation condition was good and the signal would not be interfered. The data of open area was collected from 5:45 pm to 8.31 pm in January 16, 2016 and the data sampling rate was 5 s, the elevation angle cutoff was set as 10°. Figure 1b was the balcony called

enclosed area where the observation condition was bad and the signal would be interfered to a certain extent. The data of enclosed area was collected from 1:22 pm to 6:01 pm in January 17, 2016 and the data sampling rate was 5 s, the elevation angle cutoff was set as 5° .

Figure 2a–b showed the projection of satellite orbit during data reception and Fig. 3a–b showed the number of BDS and GPS satellites in open area and enclosed area respectively.

3.1.2 Result Analysis

Figure 4 shows the result of open area. The GPS positioning result is selected as the reference value of the precision statistics. Table 1 shows the statistical result of precision. As we can see from Fig. 4 and Table 1:

- (1) The RMS of GPS positioning in North, East and Up components are 0.35, 0.22 and 0.52 cm respectively, those of BDS positioning are 10.87, 5.28 and 13.82 cm, those of BDS/GPS positioning are 1.78, 1.84 and 4.58 cm.
- (2) The positioning result of BDS is relatively worse than GPS in stability and reliability. The possible reason lies in that the orbit error of GEO satellites in BDS may affect the positioning or BDS data is of poor quality which might contain too many cycle slips and gross errors.
- (3) Positioning result of combined BDS/GPS is greatly affected due to the poor stability of BDS positioning, therefore GPS-RTK is better than combined BDS/GPS-RTK and much better than BDS-RTK.

Figure 5 shows the result of enclosed area. The combined BDS/GPS positioning result is selected as the reference value of the precision statistics. Table 2 shows the precision statistics. As we can see from Fig. 5 and Table 2:

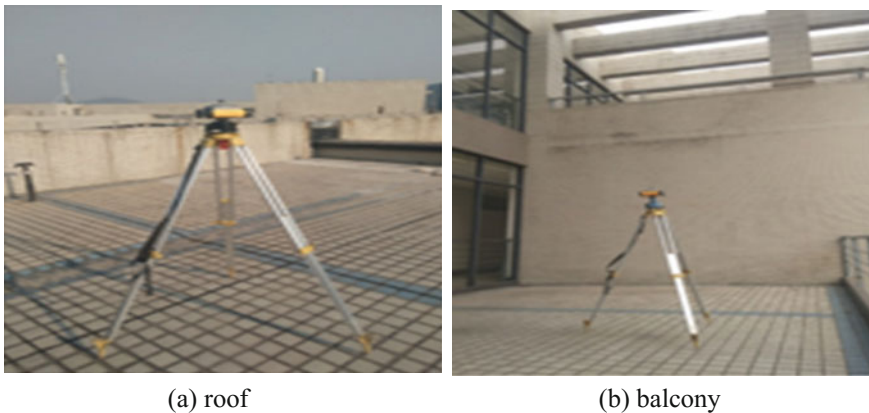


Fig. 1 The field of experimental data acquisition

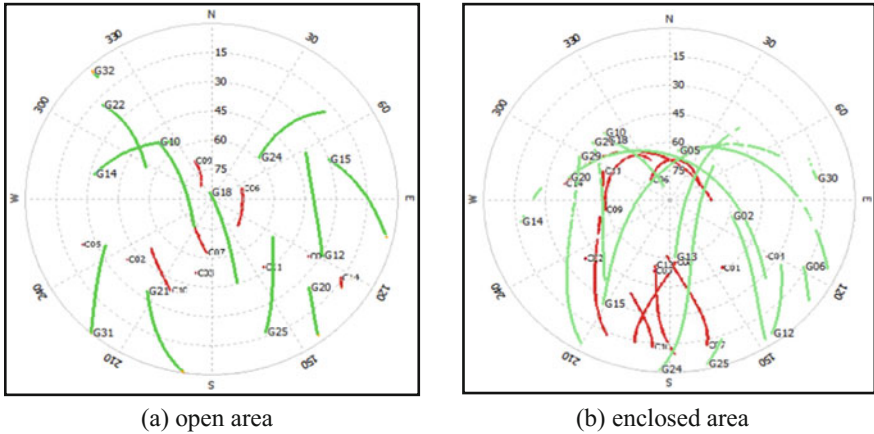


Fig. 2 The projection of satellite orbit

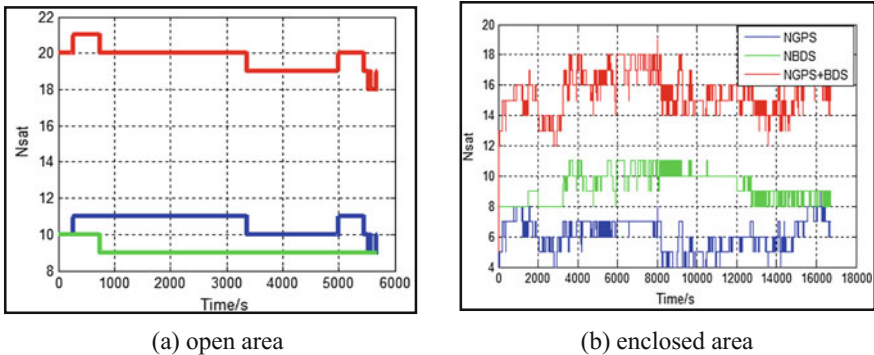


Fig. 3 The number of satellites

- (1) The RMS of GPS positioning in North, East and Up components are within 2.5 m. In BDS positioning, the North and East components are within 0.2 m, the Up component is within 0.4 m. In combined BDS/GPS positioning, the North and East components are within 0.1 m, the Up component is within 0.2 m.
- (2) In enclosed area, the result of BDS-RTK is much better than that of GPS-RTK and the combined BDS/GPS-RTK has the highest accuracy.
- (3) The main reason that the positioning result of GPS-RTK is poor lies in the fewer GPS satellites. From Fig. 3b, we can see that the number of GPS satellites is about 5 in most cases and the phenomenon of signal lock-lose occurs frequently. The number of BDS satellites is generally more than 7 and there is few interruption of signal. The number of combined BDS/GPS is more than 15 and the stability and reliability of positioning is greatly increased.

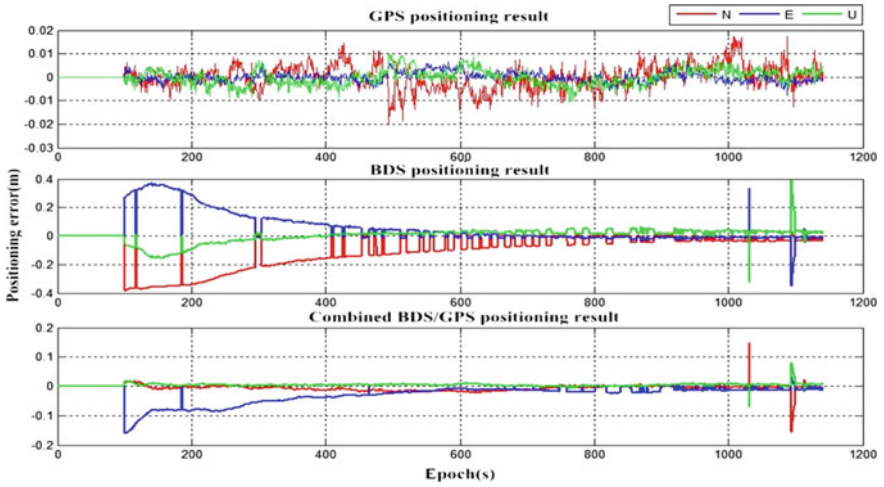


Fig. 4 The result of open area

Table 1 The statistical result of precision

Satellite system	RMS (cm)		
	N	E	U
BDS	10.87	5.28	13.82
GPS	0.35	0.22	0.52
BDS/GPS	1.78	1.84	4.58

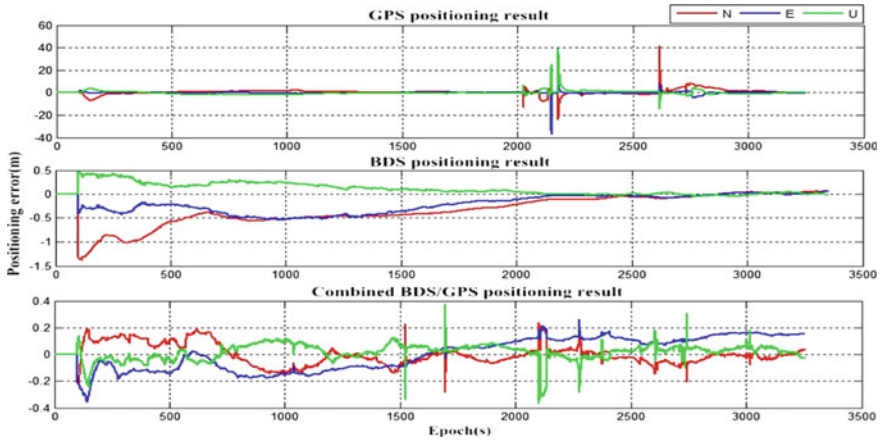


Fig. 5 The result of enclosed area

Table 2 The statistical result of precision

Satellite system	RMS (m)		
	N	E	U
BDS	0.12	0.18	0.31
GPS	1.10	2.43	2.20
BDS/GPS	0.08	0.06	0.18

3.2 Dynamic Vehicle Experiment

3.2.1 Data Collection

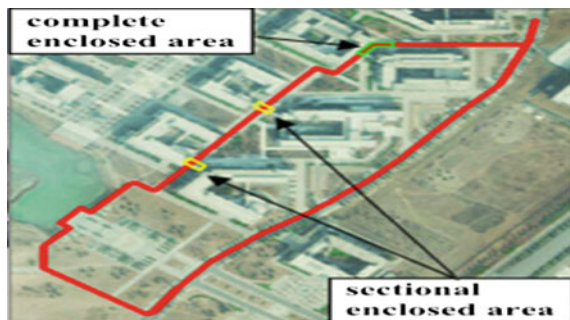
The data of dynamic vehicle experiment was collected in the campus of China University of Mining and Technology. In Fig. 6, the yellow part was complete enclosed area of signal, the green part was sectional enclosed area and the red part was open area. The data was collected from 6:36 pm to 7:17 pm in January 19, 2016 and the data sampling rate was 5 s, the elevation angle cutoff was set as 5° .

3.2.2 Result Analysis

The result of dynamic experiment is shown in the Fig. 7. It's difficult for us to select the reference value for accuracy evaluation due to the frequent phenomenon of signal lock-lose. Therefore, we just give a qualitative description for the reliability and stability of single system positioning and combination positioning. As we can see from Fig. 7:

- (1) In complete enclosed area (a) and (b), single BDS and single GPS can't provide navigation and positioning services, but combined BDS/GPS can.
- (2) In sectional enclosed area (c), the reliability and stability of single GPS is poor from the view of dispersion degree although it can provide positioning. Combined BDS/GPS is better than GPS, but the result is not very ideal in reliability.

Fig. 6 The field of experimental data acquisition



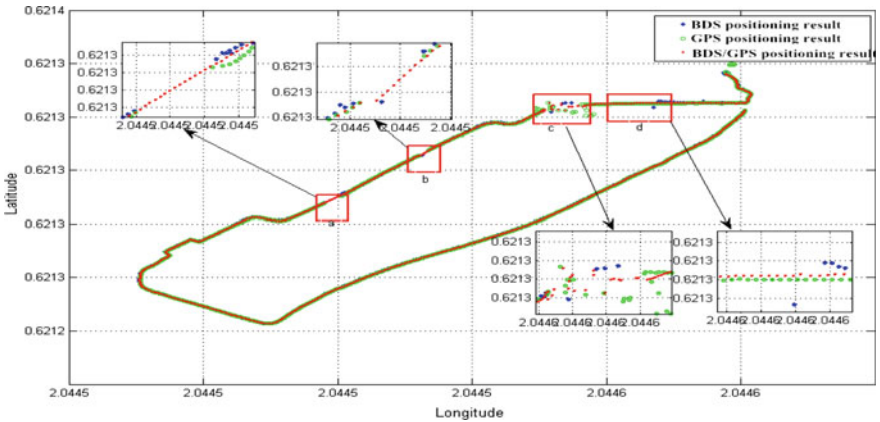


Fig. 7 The result of dynamic experiment

- (3) In sectional enclosed area (d), single GPS and combined BDS/GPS can provide stable positioning, but single BDS can't.
- (4) In open area, the results of single BDS, single GPS and combined BDS/GPS are all good and stable.

3.3 Antenna Movement Experiment

3.3.1 Data Collection

This experiment in open area included two parts, vertical movement and horizontal movement. Vertical movement was conducted on the roof of School of Environment Science and Spatial Informatics, CUMT. The data was collected from 10:07 am to 2:00 pm in May 13, 2016. In Fig. 8, the height was adjusted from 1.5 to 1.7 m at 11.33 am. It ensured that there was only change in vertical direction. Horizontal movement was conducted on the square which was in front of the gymnasium. The data was collected from 16:28 pm to 19:35 pm in May 13, 2016. In Fig. 9, the receiver was moved from 9 to 29 cm along the ruler at 18.10 pm. There was only change in horizontal direction. The data sampling rate was 1 s and the elevation angle cutoff was set as 5°.

3.3.2 Result Analysis

The vertical movement is 20 cm in theory. The processing results are shown in Fig. 10 and Table 3. We can see that:

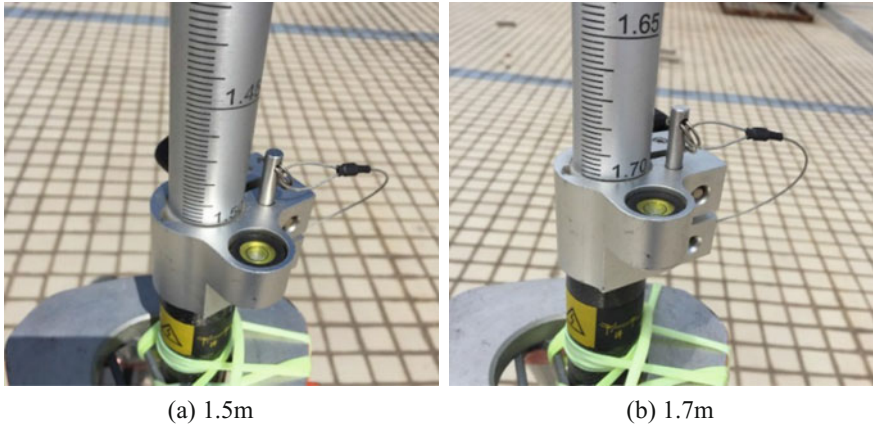


Fig. 8 Vertical movement

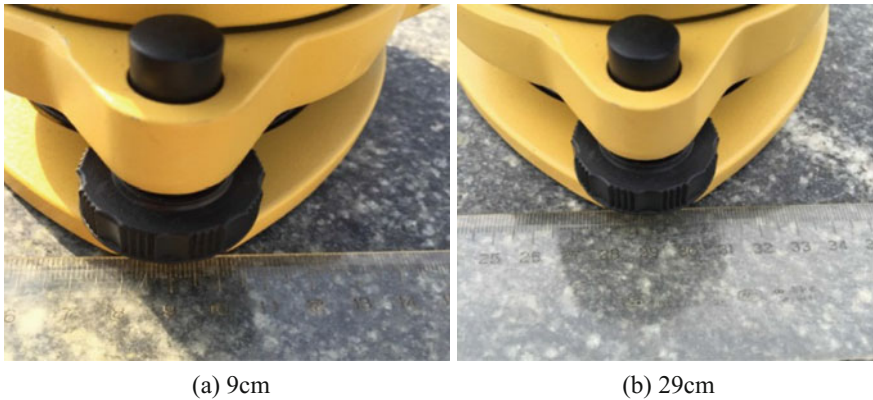


Fig. 9 Horizontal movement

- (1) Single BDS, Single GPS and combined BDS/GPS all can monitor the deformation accurately in a good observation environment, the position error is within 0.5 cm.
- (2) GPS-RTK is slightly better than BDS-RTK and combined BDS/GPS-RTK. The result of combined BDS/GPS is similar to that of BDS for there may be some problems in determining the weight of BDS and GPS.

The vertical movement experiment of enclosed area is simulated by removing some satellites artificially. Figure 11 shows the reception of satellites. We can see from it that the number of BDS satellites is generally 8–10, GPS satellites is generally 7–9. Next, the following 10 satellites G03, G08, G09, G14, G16, G23,

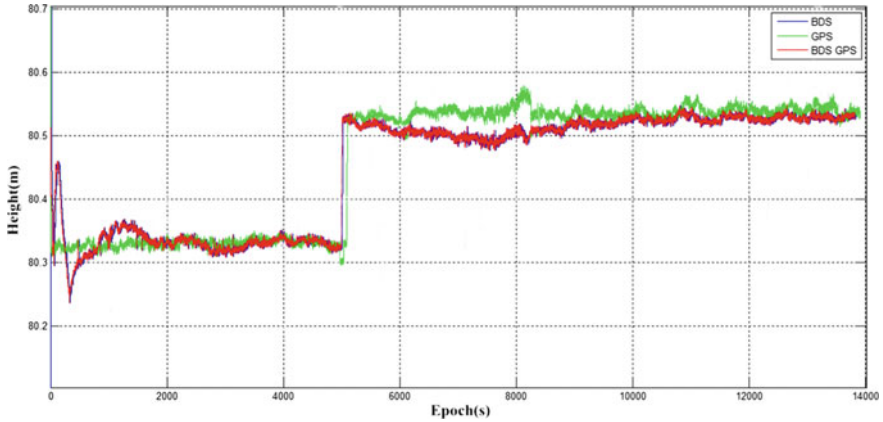


Fig. 10 The result of vertical movement in open area

Table 3 The statistical result of precision

Satellite system	Open area		
	Before adjustment (m)	After adjustment (m)	Change (m)
BDS	80.327	80.524	19.7
GPS	80.332	80.533	20.1
BDS/GPS	80.327	80.524	19.7

C07, C08, C09, C10 are removed and not involved in calculation. The result can be seen in Fig. 12 and Table 4:

- (1) Single GPS is accurate and the position error is small although the GPS satellites are few.
- (2) The position errors of single BDS and combined BDS/GPS are large, about 6 cm. The reason lies in the irrational structure of BDS satellites that IGSO

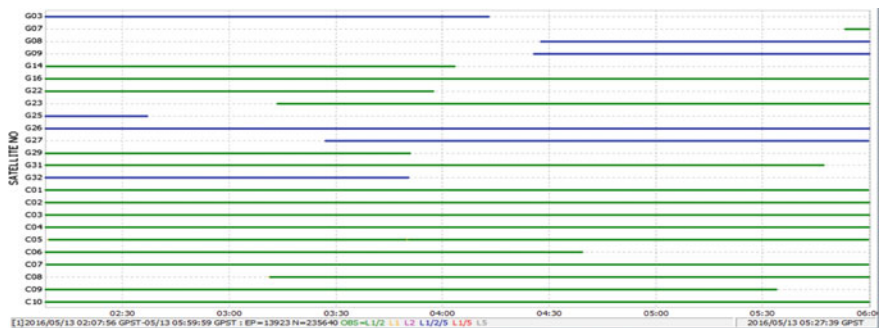


Fig. 11 The reception of satellites

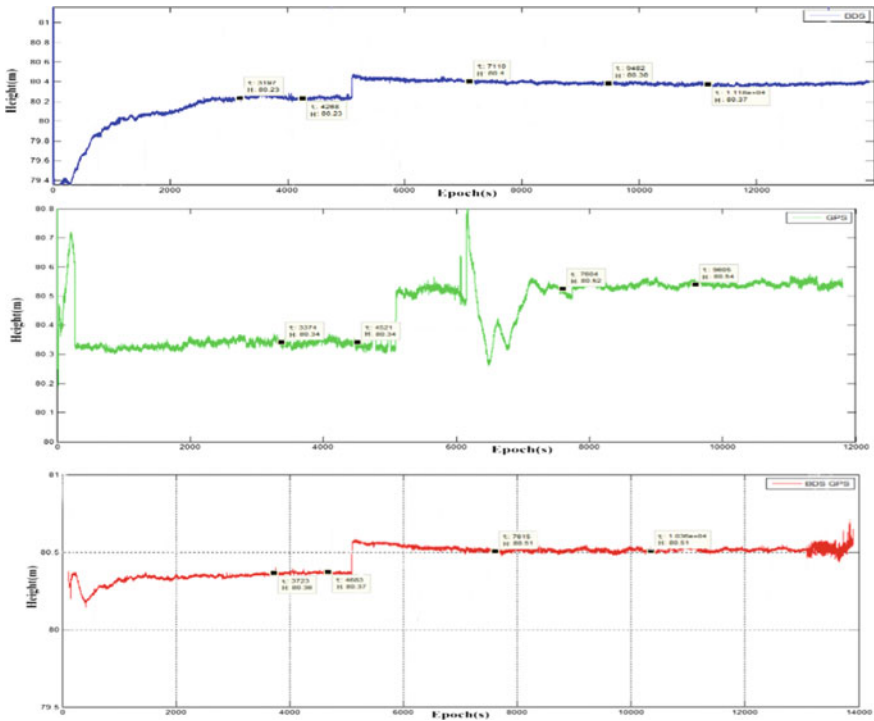


Fig. 12 The result of vertical movement in enclosed area

Table 4 The statistical result of precision

Satellite system	Enclosed area		
	Before adjustment (m)	After adjustment (m)	Change (m)
BDS	80.234	80.375	14.1
GPS	80.336	80.538	20.2
BDS/GPS	80.364	80.512	14.8

satellites are removed excessively and most of the remaining satellites are MEO satellites.

The horizontal movement is 20 cm in theory. The processing result is shown in Fig. 13 and Table 5. We can see that:

- (1) Both single GPS and combined BDS/GPS can well measure the variation of 20 cm. However, the accuracy of single GPS is slightly higher than combined BDS/GPS.
- (2) The result of single BDS is not stable, so we don't take it into account.

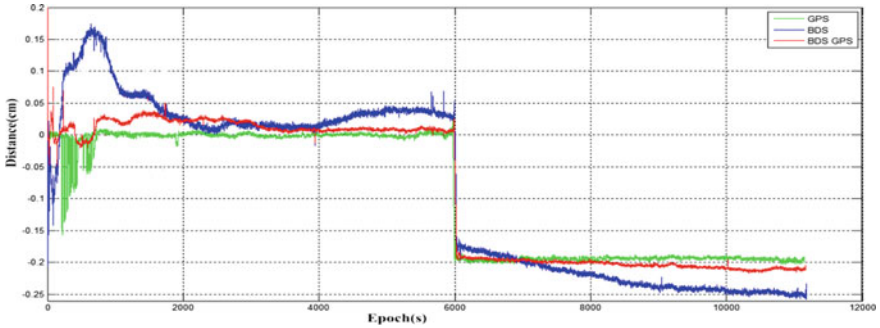


Fig. 13 The result of horizontal movement in open area

Table 5 The statistical result of precision

Satellite system	Open area		
	Before adjustment (cm)	After adjustment (cm)	Change (cm)
GPS	0	-19.9	19.9
BDS/GPS	0.1	-20.1	20.2

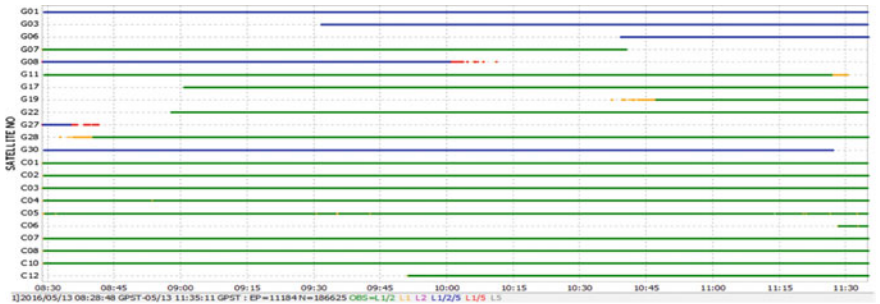


Fig. 14 The reception of satellites

Figure 14 represents the situation of satellite reception. In order to simulate the bad observation condition, we get rid of 7 satellites, including G01, G03, G08, G17, C01, C02, C03. The remaining number of GPS satellites is 4–5, which of BDS satellites is 5–6. From Fig. 15 and Table 6, we can see that:

- (1) The result of single GPS is the best and there is no change in accuracy.
- (2) Single BDS is still not stable and it is similar to the previous result. The result of combined BDS/GPS is not ideal, the measurement error becomes large.

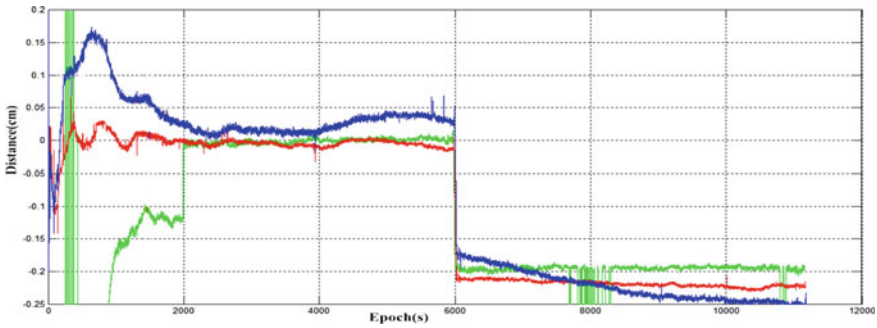


Fig. 15 The result of horizontal movement in enclosed area

Table 6 The statistical result of precision

Satellite system	Enclosed area		
	Before adjustment (cm)	After adjustment (cm)	Change (cm)
GPS	0	-19.9	19.9
BDS/GPS	-0.5	-22.2	21.7

4 Conclusion

In this paper, the positioning performance of combined BDS/GPS-RTK under the different conditions was analyzed. In static experiment, it was shown that the positioning accuracy of single GPS was better than single BDS and combined BDS/GPS in open area, combined BDS/GPS was better than single BDS and single GPS in enclosed area. In dynamic vehicle experiment, it was shown that combined BDS/GPS could still provide continuous and stable positioning service, but the positioning reliability was greatly affected when the observation environment was bad. In antenna movement experiment, the positioning accuracy of single GPS and combined BDS/GPS was almost the same in either vertical or horizontal direction in open area. And when it was in enclosed area, single GPS was better than combined BDS/GPS.

The paper briefly analyzes the positioning performance of combined BDS/GPS based on the iGMAS products, which can provide some useful reference for subsequent scholars. With the limitation of experimental conditions, the conclusions obtained in this paper need to be further test. Next step, we can consider the influence of occlusion in different directions on positioning accuracy and how to determine the weight of BDS and GPS according to the quality of measured data.

Acknowledgements We would like to thank the International GNSS Monitoring Assessment System (iGMAS) for their support for this paper. We would like to acknowledge the support of the Jiangsu dual creative talents and Jiangsu dual creative teams programme projects awarded in 2017.

This paper was funded by The Natural Science Foundation of China (41404033), Supported State Key Laboratory Lake Science and Environment (SKLGIE2014-Z-1-1) and the Fundamental Research Funds for the Central Universities (2015QNA31).

References

1. Chen K, Xu T, Chen G et al (2015) The orbit and clock combination of iGMAS analysis centers and the analysis of their precision. In: China satellite navigation conference (CSNC) 2015 proceedings, vol II. Springer, Berlin, pp 421–438
2. Gao X, Dai W, Pan L (2014) Robust EKF and its application to combined GPS/BDS kinematic relative positioning. *J Geodesy Geodyn* 04(34):140–144
3. Jiao W, Ding Q, Li J et al (2011) Monitoring and assessment of GNSS open services. *Sci Sin Phys Mech Astron* 41(5):521–527
4. Lou Y, Gong X, Gu S, Zheng F (2016) An algorithm and results analysis for GPS + BDS inter-system mix double-difference RTK. *J Geodesy Geodyn* 01(36):1–5
5. Mohamed AH, Schwarz KP (1999) Adaptive Kalman filtering for INS/GPS. *J Geodesy* 73(4):193–203
6. Odolinski R, Teunissen PJG, Odijk D (2015) Combined GPS + BDS for short to long baseline RTK positioning. *Measur Sci Technol* 26(4)
7. Odolinski R, Teunissen PJG, Odijk D (2014) Combined BDS, Galileo, QZSS and GPS single-frequency RTK. *GPS Solutions* 19(1):151–163
8. Sihao Z, Xiaowei C, Feng G et al (2014) A Kalman filter-based short baseline RTK algorithm for single-frequency combination of GPS and BDS. *Sensors* 14(8):15415–15433
9. Wang S, Bei J, Li D, Zhu H (2014) Real-time kinematic positioning algorithm of GPS/BDS. *Geomatics Inf Sci Wuhan Univ* 05(39):621–625
10. Yang Y, He H, Xu T (2001) Adaptive robust filtering for kinematic GPS positioning. *Acta Geod Cartogr Sin* 30(4):294–298

Environment Recognition Based on Temporal Filtering SVM



Yuze Wang, Peilin Liu, Xiaoguang Zhu, Xiaoxi Jin, Qiang Liu
and Jiuchao Qian

Abstract Since the signal quality of global navigation satellite system (GNSS) is extremely vulnerable to the surrounding environment, the environment-aware adaptive positioning algorithm has drawn wide attention. In order to select the suitable positioning method in different types of environment, the receiver need to recognize the type of surrounding environment in real-time. Targeting on the vehicle positioning applications in the city, this paper divides the urban environment into six categories: canyon, downtown, suburb, viaduct-up, viaduct-down and boulevard, and proposes a novel environment recognition algorithm based on the navigation signal characteristics. Firstly, a five dimension signal feature vector is proposed to describe the quality of navigation signal. The vector elements are signal power attenuation mean, power attenuation standard deviation, signal blocking coefficient, DOP value expansion ratio and power fluctuation coefficient. Then, taking this vector as environmental attribute, this paper proposes an environment recognition algorithm based on the temporal filtering support vector machine (SVM). In the experiment, the raw navigation signal data are collected for more than 100 thousand epochs in six types of urban environment, with no less than 10 thousand epochs for each type. In order to verify the validity of the proposed recognition algorithm, the five cross validation method is used to train and test all the collecting data. The testing results show that the recognition accuracy of the algorithm are higher than 90% for all types of environment.

Keywords GNSS · Signal feature vector · Environment recognition

Y. Wang (✉) · P. Liu · X. Zhu · Q. Liu · J. Qian
Shanghai Jiao Tong University, Shanghai, China
e-mail: wangyuze90@163.com

X. Jin
Beidou Satellite Navigation Center, Beijing, China

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_32

1 Introduction

The quality of satellite navigation signal is very easy to be affected by the surrounding environment, which leads to the decrease of the positioning performance of GNSS receiver. The interference factors in the environment mainly include multipath, signal occlusion and ambient thermal noise [1]. In order to improve the robustness of positioning, researchers have proposed a large number of anti-interference algorithms [2], such as multipath suppression algorithm [3], high sensitivity capture tracing, filtering algorithm, multi-source fusion algorithm, etc. However, most performance improvements of anti-interference algorithms are based on specific application scenarios, which needs to consume more computational resources. Therefore, the positioning performance of the same receiver is usually very different in different scenarios.

Considering that the inconsistencies of statistical characteristics between different interference source and disturbance intensity in different environment, it is necessary to adopt adaptive algorithm strategy. For example, in the multi-source positioning system, the error weight distribution of satellite measurements and other sensor measurements directly affects the final positioning accuracy. Therefore, it is necessary to design the weight coefficient to match different characteristic types, so as to improve receiver's environmental adaptability. In addition, the algorithm based on the environment recognition can reduce the unnecessary consumption of computing resources, thus saving energy efficiency. For example, MEDLL, the classic multi-path suppression algorithm, needs to increase a large number of correlators in the tracking loop to detect multipath signals [4]. Actual test results show that MEDLL in urban canyon environments can be significant to suppress the multipath signal, but not significant in the suburban environment. Therefore, the receiver can choose to close the algorithm in suburban environment to save hardware energy consumption. Although the environment recognition algorithm can effectively promote receiver's comprehensive performance in all kinds of scenes, but first step of the method is that receiver how automatically identifies the type of scene by its own perception characteristics. So we need to build the environment recognition algorithm based on navigation signal characteristics.

At present, the environment recognition algorithm has been widely applied. This algorithm is first applied to the visual field, which is mainly based on image feature to identify each type. Tang and Breckon [5] used the color, texture and edge features of the image to identify whether the motor vehicle is on the highway, in the city, suburban. Bosch et al. [6] proposed a dimension reduction method to identify whether the environment is forest, mountain range, highway, street, beach, etc. Although the visual classification method can identify better, but most satellite navigation terminals cannot perceive visual characteristics of the surrounding environment. Therefore, it is necessary to develop a environment classification algorithm based on satellite navigation signal characteristics. Hsu et al. [7] used satellites visibility and pseudorange measurements to identify the receiver above or below the highway. However, the algorithm can only identify two kinds of

scenarios, and which needs to know the height of viaduct as a prior condition. Drawil et al. [8] used the mean and variance characteristics of satellite signal strength to identify environments, which were divided into three types: high, medium and low precision in real time. Groves et al. [9] used the signal strength of satellite and the intensity of WIFI signal to judge the environments, which were divided into five categories: open space, suburb, urban, urban canyon and tunnel. Lighari et al. [10] analyzed the intensity and variance of satellite signals in each type of scenes, which were divided into open ground, forest shade, unilateral occlusion and bilateral occlusion. Bancroft et al. [11] used the sensory information and satellite signals to identify the user's location, whether is indoor or outdoor, and the type of motion.

Above all, although there are many kinds of environment recognition algorithms based on navigation signals, the main disadvantages are significant, which include two aspects. One is that some algorithms need to use other information more than the navigation signals, and the second is that the types of recognition environments are not comprehensive. Therefore, based on vehicle-mounted receiver, this paper proposes a environment recognition algorithm based on navigation signal characteristics for different environments, which are divided into urban canyons, urban, suburban, above highway, below highway and shade.

2 Signal Feature Vector

To recognize each type of environment based on GNSS signal characteristics, an important task is to build a reliable signal feature vector. Most of the previous studies use signal power, the number of visible satellites and the GDOP value as feature elements, but it is only suitable for distinguishing environment with fewer categories. In order to recognize more environmental categories, the signal feature space need to be expanded.

This paper proposes a five dimensional signal feature vector to characterize the environmental interference. The vector parameters are signal power attenuation mean and standard deviation, blockage coefficient, GDOP expansion ratio and power fluctuation coefficient. The detailed calculation of each parameter are given as follows.

- Mean μ and standard deviation σ of signal power

The satellite signal power is influenced by many factors such as atmospheric attenuation, antenna gain and equipment noise. In order to calculate the power attenuation only caused by surrounding environment, the standard signal power with no attenuation need to be measured in the open sky environment. The signal power attenuation of each satellite can be calculated by using standard power minus measured power. Thus, the power attenuation of satellite k at time t can be calculated as follows:

$$A_k(t) = P_k^{std}(t) - P_k(t) \quad (1)$$

where $P_k(t)$ is measured power of satellite k at time t . $P_k^{std}(t)$ is the standard power.

Then, $\mu(t)$ and $\sigma(t)$ can be calculated as follows:

$$\mu(t) = \frac{1}{N} \sum_{k=1}^N A_k(t) \quad (2)$$

$$\sigma(t) = \sqrt{\frac{1}{N} \sum_{k=1}^N (A_k(t) - \mu(t))^2} \quad (3)$$

where N is the number of visible satellites.

- Blockage coefficient α

Blockage coefficient can be calculated as follows:

$$\alpha(t) = 1 - \frac{N_{visible}(t)}{N_{all}(t)} \quad (4)$$

where $N_{visible}$ is the number of visible satellite, and N_{all} that can be calculated by almanac is the number of all satellite in the sky.

- GDOP expansion ratio λ

The GDOP value can be calculated as follows:

$$\tilde{G} = \begin{bmatrix} -\cos \theta^{(1)} \sin \alpha^{(1)} & -\cos \theta^{(1)} \cos \alpha^{(1)} & -\sin \theta^{(1)} & 1 \\ -\cos \theta^{(2)} \sin \alpha^{(2)} & -\cos \theta^{(2)} \cos \alpha^{(2)} & -\sin \theta^{(2)} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ -\cos \theta^{(N)} \sin \alpha^{(N)} & -\cos \theta^{(N)} \cos \alpha^{(N)} & -\sin \theta^{(N)} & 1 \end{bmatrix} \quad (5)$$

$$\tilde{H} = (\tilde{G}^T \cdot \tilde{G})^{-1} \quad (6)$$

$$GDOP = \sqrt{\text{tr}(\tilde{H})} \quad (7)$$

where $\theta^{(k)}$ and $\alpha^{(k)}$ is the elevation and azimuth respectively. Then, the λ can be calculated:

$$\lambda(t) = \frac{GDOP(t)}{GDOP^{std}(t)} \quad (8)$$

where $GDOP(t)$ is the measured value of the receiver, and $GDOP^{std}(t)$ is value in the open sky environment.

- power fluctuation coefficient β

The signal power fluctuation coefficient is defined as the standard deviation of the signal power attenuation in one second. It can be calculated as follows:

$$t_0 = t_1 - fs \quad (9)$$

$$D_k(t_1) = \frac{1}{t_1 - t_0} \sum_{t=t_0}^{t_1} A_k(t) \quad (10)$$

$$\beta_k(t_1) = \sqrt{\frac{1}{t_1 - t_0} \sum_{t=t_0}^{t_1} (A_k(t) - D_k(t_1))^2} \quad (11)$$

$$\beta(t_1) = \frac{1}{N} \sum_{k=1}^N \beta_k(t_1) \quad (12)$$

where fs is the data recording frequency of receiver. N is the satellite number.

According to the above description, the five dimension signal feature vector can be express as:

$$\gamma = [\mu, \sigma, \alpha, \lambda, \beta] \quad (13)$$

This vector will be used in environment recognition algorithm in the next part.

3 Data Collection

This section will illustrate the experimental equipment of data collection and characteristics of each environmental type. First, the collection route and the subjective description for each type of environment is given. Then, the statistical distributions of signal feature vector is studied in different types of environment.

In order to obtain the raw data of GNSS signal, a commercial receiver is used to log the message five times every second. The antenna of the receiver is mounted on a van, and the motion state of the van is medium and low speed. Figure 1 shows the experimental platform and the collecting route in Shanghai, China. More than 100 thousand seconds data have been collected and the collecting area is almost 250 km². It can be seen that the collecting route covers most of the typical environment in the city.

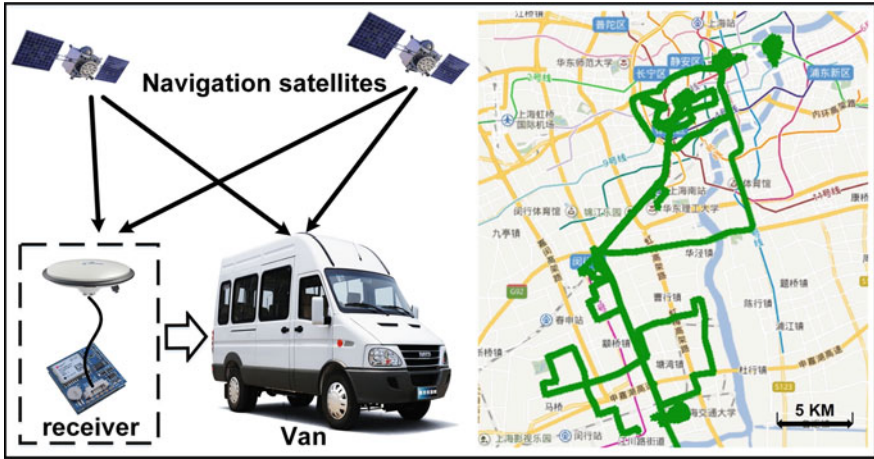


Fig. 1 Signal collecting platform and route

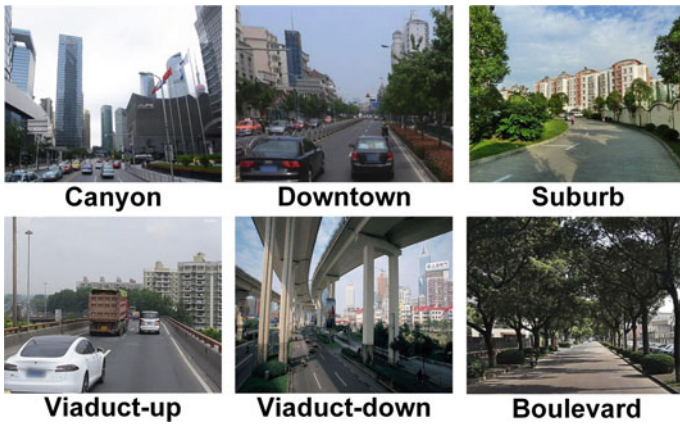


Fig. 2 Live photo of each type environment

This paper divides urban environment into six categories: urban canyon, downtown, suburb, viaduct-up, viaduct-down and boulevard. The photos of each type are shown in Fig. 2.

As shown in Fig. 2, the characteristics of the surrounding environment are distinctly different in each type, and the description are as follows:

- Canyon: In this environment, there are a large number of skyscrapers on both sides of the road which is taller than 100 m and widely use glass walls. The GNSS signal will be easily blocked and multipath interference is serious.
- Downtown: In downtown environment, the height of most buildings are around 50 m, and the multipath interference will be lower.

- Suburb: In general, the buildings are no higher than 20 m in suburb. The positioning accuracy of receiver is usually good.
- Viaduct-up: In such environment, the interference is very weak, so the signal quality is good and hardly to be interfered.
- Viaduct-down: In viaduct-down environment, the signal is very easily to be blocked, especially for the satellite with high elevation.
- Boulevard: Almost all satellite signals will be attenuated by trees under this environment, and the short delay multipath is serious.

4 Environment Recognition Algorithm

Taking the feature vectors as the training data, this section proposes a temporal filtering SVM algorithm to recognize the environmental type. The feature vectors need to be standardized before training, and the equation is as follows.

$$\gamma' = \frac{\gamma - \mu(\gamma)}{\sigma(\gamma)} \quad (14)$$

where γ is the raw feature parameter. $\mu(\gamma)$ is the mean of γ . $\sigma(\gamma)$ is the standard deviation. In the following section, γ' will be used to replace γ in the recognition algorithm.

4.1 SVM Classifier

The basic theory of support vector machine (SVM) is to find a partition hyperplane in the sample space and separate the different categories of samples, which is based on the training set D . The sample set is defined as follows.

$$D = \{(x_1, y_1), \dots, (x_m, y_m)\}, \quad y_i \in \{-1, +1\} \quad (15)$$

In the sample space, the division of the hyperplane can be described by the following linear equation:

$$\omega^T x + b = 0 \quad (16)$$

where $\omega = (w_1; w_2; \dots; w_d)$ is the normal vector which determines the direction of hyperplane. b is the displacement which determines the distance between origin and hyperplane. Then, The distance between any point x and the hyperplane (ω, b) in the sample space can be calculated as:

$$r = \frac{|\omega^T x + b|}{\|\omega\|} \tag{17}$$

If hyperplane (ω, b) could classify the training data correctly, the following function can be get for $(x_i, y_i) \in D$.

$$\begin{cases} \omega^T x_i + b \geq +1, & y_i = +1 \\ \omega^T x_i + b \leq -1, & y_i = -1 \end{cases} \tag{18}$$

Then

$$r = \frac{2}{\|\omega\|} \tag{19}$$

In order to maximum the space, we need to calculate;

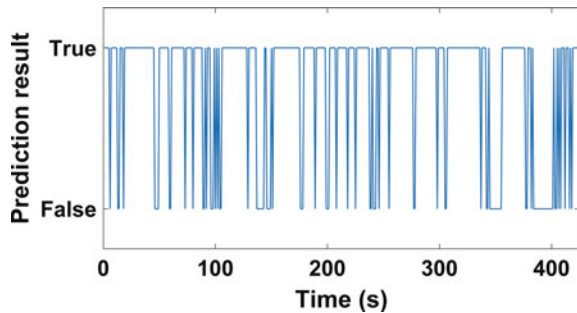
$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \tag{20}$$

$$s.t. \quad y_i(\omega^T x_i + b) \geq +1, \quad i = 1, 2, \dots, m \tag{21}$$

Finally, the hyperplane coordinates can be calculated by solving the above formula. Then, the hyperplane can be used to classify the categories of observation data in real-time.

By training the signal feature vector obtained from the collecting data, the SVM algorithm can be used to identify the environmental type. Figure 3 shows the classification results of SVM algorithm for downtown environment in a short period of time. The peak value represents the correct prediction, and the trough represents the wrong prediction. It can be seen that the recognition result of SVM algorithm is correct in most epochs. However, the recognition result will fluctuate frequently in some period, because the signal feature parameters are affected by random noise.

Fig. 3 Recognition result of SVM in downtown environment



4.2 Prediction Updating Based on Temporal Filtering

According to the previous analysis in Fig. 3, the random noises of measured feature vector will cause the fluctuation of recognition result and decrease the recognition accuracy. However, the users usually stay in the same type of environment for a period of time in practical applications. Thus, the filtering algorithm can be considered to improve the environment recognition performance.

An important work of filtering algorithm is to design the filtering window parameters. In this research, the type conversion of environment is determined by the spatial coordinate of the receiver. Therefore, the moving mileage of the receiver will be an important reference to design the length of filtering window. In addition, it is necessary to consider the running time to ensure that the window length will not increase infinitely in static cases. Thus, taking both moving mileage and the running time into consideration, a SVM classification algorithm based on the temporal filtering (SVM-FL) is proposed in this paper.

In this algorithm, the recognition result with maximum probability is chosen to represent the environmental type for all epochs in one filtering window. The detailed calculating procedures are as follows.

1. By training the signal feature vector, the basic SVM algorithm is used to get a preliminary prediction $\hat{s}(t)$ at every epoch t .

$$\hat{s}(t) = SVM(\gamma(t)) \quad (22)$$

where $\gamma(t)$ is the signal feature vector.

2. The time $T_s(t)$ it takes the receiver to move s meters is calculated. The length of filtering window k is chosen as the minimum between $T_s(t)$ and fixed value T_0 .

$$s = f(t_0) = \sum_{t_i=t_0}^t v_i \cdot (t_i - t_{i-1}) \quad (23)$$

$$T_s(t) = t - t_0 = t - f^{-1}(s) \quad (24)$$

$$k = \min(T_s(t), T_0) \quad (25)$$

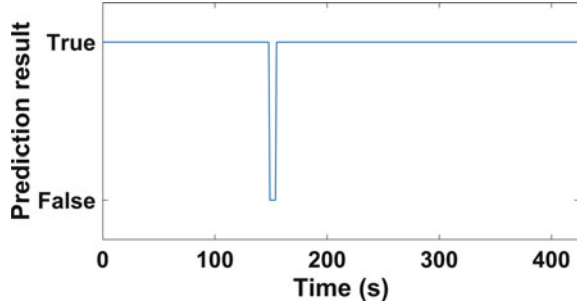
where v_i is the mean speed in the period of $t_i - t_{i-1}$.

3. The maximum probability of k preliminary predictions is selected as the final recognition result $s(t)$, and it can be calculated as follows:

$$s(t) = \text{MAX}(\{\hat{s}(t - k + 1), \hat{s}(t - k + 2), \dots, \hat{s}(t)\}) \quad (26)$$

Figure 4 shows the environment recognition results for the test example in previous section by using the SVM-FL algorithm, and it can be seen that the recognition accuracy has been significantly improved.

Fig. 4 Recognition of SVM-FL in urban environment



5 Performance Test

5.1 Testing Method

In order to test the performance of proposed recognition algorithm, the five cross validation method is used to train and recognize the data which collected in Sect. 3. Firstly, the data set is randomly divided into 5 mutually exclusive subsets equally. Four subsets are mixed as the training set, and the rest set is taken as the test set. Every subset will be test once, so totally five training and tests need to be done. Finally, the recognition performance of all collecting data can be obtained.

In this experiment, three methods are used to recognize the environmental types for the collecting data set. The first method uses the traditional signal feature vectors \bar{h} and SVM algorithm to recognize the environment. Its vector parameters are signal power mean, signal power variance, DOP value and number of visible satellite. The second method uses the SVM algorithm to recognize the environment based on the signal feature vectors γ that proposed in this paper. The third one uses the SVM-FL algorithm to recognize the environment based on feature vectors γ .

5.2 Recognition Result

Figure 5 shows the recognition performance of three methods in different types of environment. The blue dotted line $SVM(\bar{h})$ is the performance of first method. The yellow dashed line $SVM(\gamma)$ is the performance of second methods, and the red solid line $SVM-FL(\gamma)$ represents the performance of third methods.

According to the figure, it can be seen that the traditional recognition algorithm has a good performance in three types of environment which are suburb, viaduct-up and viaduct-down. But its recognition accuracy is not good in other three environmental types. When using the signal feature vector proposed in this paper, the accuracy of recognition algorithms can be improved by 10–20%. The recognition performance can be further improved by using SVM-FL algorithm, and the recognition accuracy is higher than 90% for all types of environment.

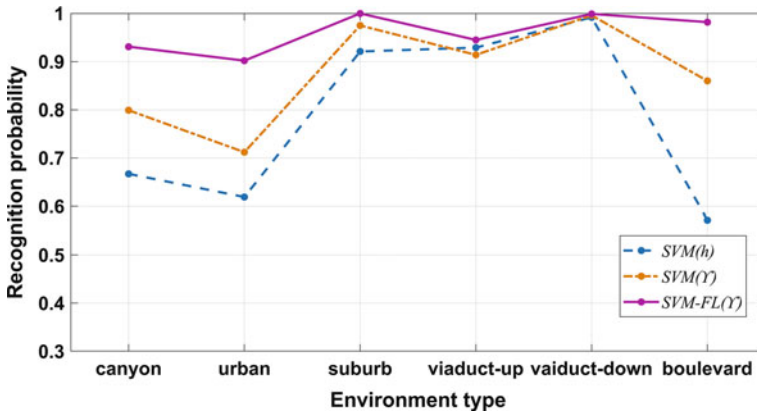


Fig. 5 performance comparison among three method

6 Conclusion

Based on the influence of environmental interference on GNSS signal characteristics, this paper divides urban environment into six types which are canyon, downtown, suburb, viaduct-up, viaduct-down and boulevard. Then an environment recognition algorithm is proposed based on the GNSS signal feature vector. It contains two main innovations:

- A feature vector is proposed for describing the influence of environmental interference on GNSS signal, and the parameters contain the power attenuation mean, power attenuation standard deviation, blockage coefficient, GDOP expansion ratio and power fluctuation coefficient. The statistical distribution of each parameter is analyzed in different types of environment.
- An environmental type recognition algorithm based on the feature vector is proposed. The recognition accuracy are more than 90% for all types of environment.

The research results in this paper provide a reliable guidance for designing high-accuracy positioning algorithms under different environment, which is beneficial to improve the comprehensive performance of receivers.

References

1. Misra P, Enge P (2006) Global positioning system: signals, measurements and performance second edition. Ganga-Jamuna Press, Massachusetts
2. Gustafsson F, Gunnarsson F, Bergman N et al (2002) Particle filters for positioning, navigation, and tracking. IEEE Trans Signal Process 50(2):425–437

3. Van Nee RDJ, Siereveld J, Fenton PC et al (1994) The multipath estimating delay lock loop: approaching theoretical accuracy limits. In: Position location and navigation symposium, IEEE. IEEE, pp 246–251
4. Alban S, Akos DM, Rock SM et al (2003) Performance analysis and architectures for INS-aided GPS tracking loops. In: Proceedings of the Institute of Navigation National Technical Meeting. Institute of Navigation, pp 611–622
5. Tang I, Breckon TP (2011) Automatic road environment classification. *IEEE Trans Intell Transp Syst* 12(2):476–484
6. Bosch A, Zisserman A, Muñoz X (2008) Scene classification using a hybrid generative/discriminative approach. *IEEE Trans Pattern Anal Mach Intell* 30(4):712–727
7. Hsu LT, Gu Y, Kamijo S (2017) Intelligent viaduct recognition and driving altitude determination using GPS data. *IEEE Trans Intell Veh* 2(3):175–184
8. Drawil NM, Amar HM, Basir OA (2013) GPS localization accuracy classification: a context-based approach. *IEEE Trans Intell Transp Syst* 14(1):262–273
9. Groves PD, Martin HFS, Voutsis K et al (2013) Context detection, categorization and connectivity for advanced adaptive integrated navigation. The Institute of Navigation
10. Lighari RUR, Berg M, Salonen ET et al (2017) Classification of GNSS SNR data for different environments and satellite orbital information. In: 2017 11th European conference on antennas and propagation (EUCAP). IEEE, pp 2088–2092
11. Bancroft JB, Garrett D, Lachapelle G (2012) Activity and environment classification using foot mounted navigation sensors. In: 2012 international conference on indoor positioning and indoor navigation (IPIN). IEEE, pp 1–10

Pseudorange Double Difference Algorithm Based on Duty-Cycled Carrier Phase Smoothing on Low-Power Smart Devices



Qiang Liu, Rendong Ying, Yuze Wang, Jiuchao Qian and Peilin Liu

Abstract In recent years, due to the open access to raw GNSS measurements, it is possible to study other PVT algorithms to improve the positioning accuracy on smartphones. In this paper, we first give the analysis of the characteristics of the smartphone observations and measurement error sources, then focus on the continuous cycle slip of carrier phase with duty cycle mode. Then, we propose an improved Hatch filtering algorithm in order to use the carrier phase observations which have continuous loss of cycles to reduce the measurement noise of pseudoranges. Finally, we give a comparison of several pseudorange positioning algorithms, the model of pseudorange double difference algorithm and the parameter selection principle of Kalman filter. Experiments based on a real smartphone in the open sky environment and the post-processing of the collected data is discussed in detail later. The results show that the proposed carrier smoothing algorithm can effectively reduce the noise of pseudorange observations and the positioning accuracy of the pseudorange double difference algorithm with carrier smoothing technique is less than 5 m, which shows obvious improvement compared with the positioning accuracy of the traditional positioning method on smartphones.

Keywords Smartphone GNSS measurements · Duty cycle · Pseudorange double difference · Carrier smoothing

1 Introduction

Since Nokia N95 first built a built-in GPS module on the mobile phone, GNSS has become one of the standard parts of smartphones during the decade. Today, however, the positioning accuracy of GNSS module on smartphone is still not very

Q. Liu (✉) · R. Ying · Y. Wang · J. Qian · P. Liu
School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University, Shanghai, China
e-mail: lugialiu@sjtu.edu.cn

high, about 5–10 m [1]. On the one hand, the performance of GNSS hardware on smartphone is generally poorer than the survey-grade receiver's hardware due to low cost and limited power consumption, these components have led to large errors and noises in the smartphone GNSS observations. On the other hand, most of the smartphone's operating system provide developers application programming interface (API) which only supports the output of the positioning results, it cannot improve the GNSS positioning accuracy essentially just by filtering these information. These two factors contribute to a simple positioning method and low positioning accuracy on smartphones. But so far there has been a breakthrough in May 2016, Google first announced the possibility to retrieve GNSS raw data from smartphones equipped with Android Nougat or later [7], so it is possible for researchers to study on other methods such as difference, carrier phase to achieve high precision GNSS positioning with low-cost platform. K. M. Pesyna et al. were the first group to explore the feasibility of high precision GNSS positioning on smartphones, they analysed the smartphone-quality GNSS antenna [8] and duty cycle mode of GNSS chipset [9], and came to the conclusion that centimetre accurate positioning is possible based on smartphone's hardware [5] and carrier phase difference algorithm before 2016. However, their work was not based on real smartphone GNSS measurements but used the external RF front-end and processing software, so it is impractical. Since the open access to GNSS raw data, more and more scholars are able to analyse the data which is recorded by a real phone, such as Realini [10], Banville [2] and Garcia [3] and others, they analysed the raw observations with different approaches and obtained some preliminary conclusions, some of them even obtained the positioning results by using carrier phase observations. But their work also has limitations, they adopted Nexus 9 which is currently the only one tablet supports to disable the duty cycle mode, that is, it can track the carrier phase continuously. Duty cycle mode can greatly reduce the power consumption on smartphone, and is very likely to be retained in the GNSS chipset of future.

In this paper, we firstly analyse the characteristics of the raw observations with duty cycle mode and the relationship between their error sources and hardware. Then we propose an improved carrier smoothing algorithm and pseudorange double difference (DD) model. Finally, the performance of the proposed algorithm is verified by experiments and the conclusion and prospect are given.

2 Error Analysis of Smartphone Measurements

2.1 Performance Limits of Smartphone's GNSS Module

The GNSS hardware on smartphones is shown in Fig. 1, which has no difference with the hardware in common GNSS receiver, but the performance of smartphone's hardware is poorer than the latter's. The smartphone's antenna loses between 5 and

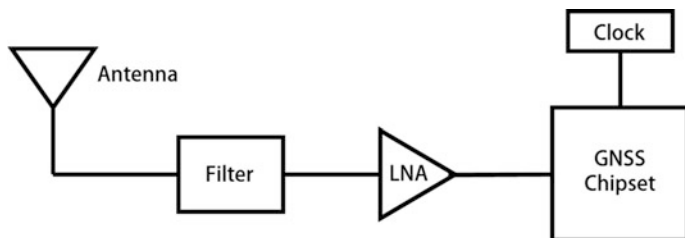


Fig. 1 GNSS hardware component in smartphone

15 dB in sensitivity as compared to the survey-grade antenna [8]. In addition, the antenna's ability of multipath suppression is poor by its linear polarization. The internal clock of smartphone adopts the low quality TCXO which clock drift is large, and exhibits significant low-frequency phase noise, making cycle-slip-free carrier phase recovery more challenging [5]. In order to reduce power consumption, GNSS chipset usually adopts duty cycle mode, namely in each signal processing cycle just only a small part of the receiving signal will be processed, and the rest time is idle, this mode can effectively reduce the power consumption of the chip, and typical value is 13 mW [9], but it has also led to non-continuous GNSS carrier phase tracking, thus this characteristic becomes an important obstacle to realize high precision GNSS positioning on smartphones. Finally, due to the technology for circuit designing, the power of thermal noise in smartphone is also larger than that of the survey-grade receiver. These characteristics of hardware will be reflected in the measurement errors of GNSS observations, and these errors will also have an impact on the positioning results, so it is necessary to analyze in-depth in the post.

Although we can access to raw observations by Android APIs, but it does not support directly give us those observations such as pseudorange which can be used to calculate position, so we need to convert recorded data into RINEX files for post-processing. In addition, the APIs also provide other information including the accuracy of measurement, cycle slip flag and multipath indicator and so on [4], and RINEX files don't have these information, it is beneficial for designing algorithm.

2.2 Experimental Setup

Difference GNSS experiments take place in Fig. 2 (a). The position of the reference station's antenna (c) is prior known (see Table 1). First, the position of the two rover stations in the figure (b) need to be calibrated. The reference station is NovAtel's PP6 receiver (upper right of Fig. 2), while the rover station is NovAtel's PP6D receiver, and the baseline length is about 400 m. We use the Waypoint Inertial Explorer software to process the data, the calibration positions of A and B are shown in Table 1, which is accurate to a centimeter level.

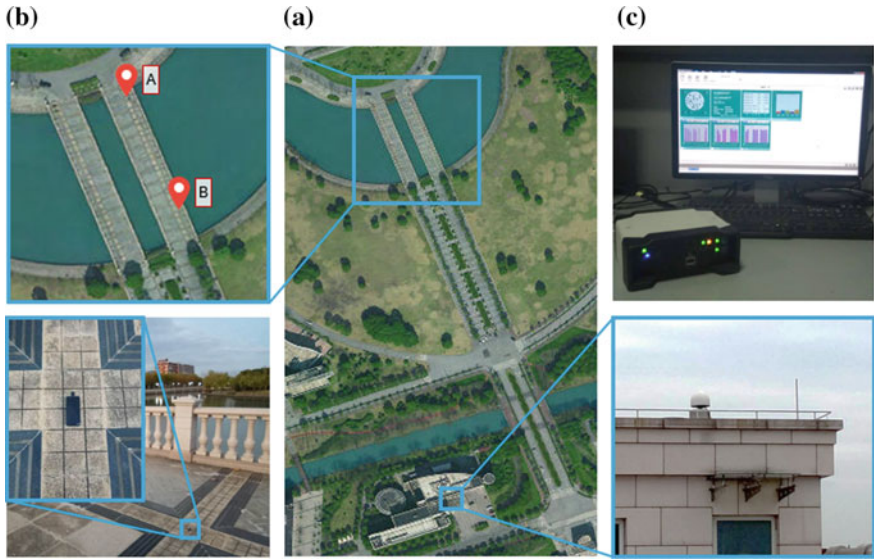


Fig. 2 Environment of data logging with smartphone

Table 1 Ground truth of reference and rover stations

	ECEF X [m]	ECEF Y [m]	ECEF Z [m]
Reference Station	-2,853,445.34	4,667,464.96	3,268,291.03
Rover Station A	-2,853,315.49	4,667,286.38	3,268,615.41
Rover Station B	-2,853,348.66	4,667,300.27	3,268,567.25

During the following smartphone experiments, reference station still uses PP6 receiver, while using Huawei P10 smartphone as rover station (lower left of Fig. 2), its Android version is 7.0 and adopts Broadcom GNSS chipset internally with duty cycle mode. The data of these two test stations are collected in real time and the observations are processed in RINEX format.

2.3 Error Analysis of Pseudorange Observations

The pseudorange observation equation of reference and rover station is:

$$\rho_r^k = r_r^k + c(\delta t_r - \delta t^k) + I_r^k + T_r^k + \varepsilon_r^k \tag{1}$$

$$\rho_u^k = r_u^k + c(\delta t_u - \delta t^k) + I_u^k + T_u^k + \varepsilon_u^k \tag{2}$$

In the above equations, the superscript k represents the satellite, the subscript r and u represent the reference and rover station, ρ is pseudorange observation, r is the real distance between satellite and receiver, c is the speed of light, δt_r and δt_u is clock bias of two stations, δt^k is clock bias of satellite k , I and T represent the propagation delay of ionosphere and troposphere, and ε is the sum of all the rest errors.

In this paper, the following strategy is used to analyze the error of pseudoranges. First, by observing the pseudorange observation equations, we can see that the equations consist of three parts: one is real distance; second one is the error associated with the receiver, such as the clock errors, thermal noise and multipath errors; third one is the errors unassociated with the receiver, such as the satellite clock errors, ephemeris errors, ionosphere and tropospheric delay. Error analysis is as follows:

Condition 1: because of the coordinates of reference and rover stations are known, and the space position of satellites also can be known by ephemeris at a particular period of time, so we can get the real distance r between stations and satellite k .

Condition 2: for the reference station, because the output observations have been corrected, so the clock bias in RINEX file for each epoch is zero, and due to its high quality antenna and hardware components, and the pseudoranges also have been smoothed by carrier phase, so the errors associated with the receiver can be considered as zero.

Based on the above two conditions, we can get the errors unassociated with the receiver of the reference station. Because it is a short baseline case in this experiment, so the ionosphere and troposphere delay of two stations about satellite k is almost equal, so this part can be regarded as common-mode errors.

For the rover station, the real distance and common-mode errors are known, so the errors associated with the receiver can be obtained. These errors will be discussed in detail in the following. In Fig. 3 (a), the test results of the errors associated with the smartphone in pseudorange observations of satellite G18 are given.

As we can see from figure (a), not only satellite G18 but also other satellites have the same decline pattern, which is the result of smartphone's clock drift. In order to separate the smartphone's clock errors from thermal noises, the following least squares method are adopted:

$$\mathbf{y} = \mathbf{C}\delta t_u + \boldsymbol{\varepsilon}_u \tag{3}$$

$$\mathbf{y} = \begin{bmatrix} (\rho_u^1 - r_u^1) - (\rho_r^1 - r_r^1) \\ (\rho_u^2 - r_u^2) - (\rho_r^2 - r_r^2) \\ \vdots \\ (\rho_u^N - r_u^N) - (\rho_r^N - r_r^N) \end{bmatrix}_{N \times 1} \quad \mathbf{C} = \begin{bmatrix} c \\ c \\ \vdots \\ c \end{bmatrix}_{N \times 1} \quad \boldsymbol{\varepsilon}_u = \begin{bmatrix} \varepsilon_u^1 \\ \varepsilon_u^2 \\ \vdots \\ \varepsilon_u^N \end{bmatrix}_{N \times 1}$$

$$\widehat{\delta t_u} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{y} \tag{4}$$

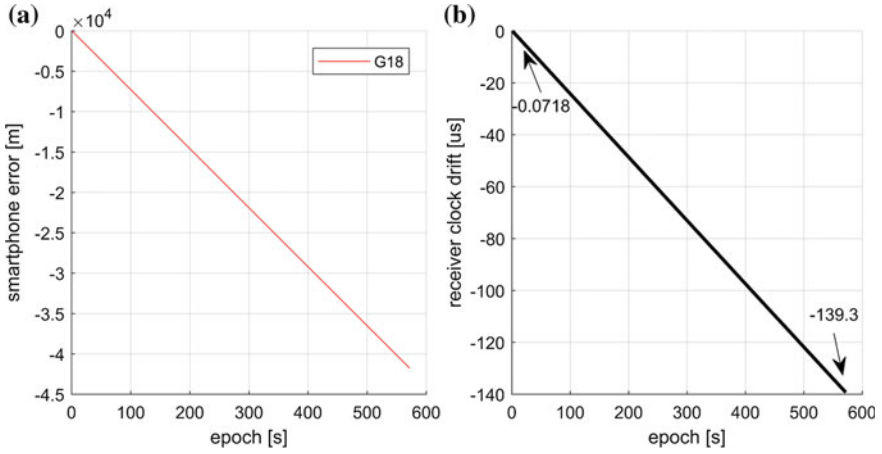
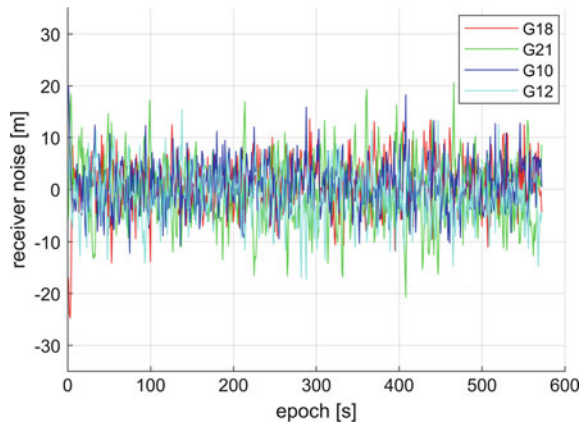


Fig. 3 The smartphone’s own error and the clock drift estimate

where y is known in the equation, and formula (4) gives an estimate of the clock error in each epoch. It is important to note that ϵ_u is not a zero mean Gaussian noise due to other errors such as multipath, but we also can use this method to give an enough good estimation of clock error. Finally the clock error estimation is shown in Fig. 3 (b), thus it can be seen that clock drift is -0.243×10^{-6} s/s on average.

When the clock error is separated from the errors associated with the receiver, the result shown in Fig. 4 is obtained. The result mainly consists of thermal noises and multipath errors, and all other errors that are ignored during the whole analysis. Among them, the standard deviation of satellite G18 is 5.02 m, satellite G21 is 6.36 m, satellite G10 is 4.84 m, and satellite G12 is 5.22 m. This error is also very serious and needs to be eliminated.

Fig. 4 Thermal noises and multipath errors



2.4 Error Analysis of Carrier Phase Observations

Similar to the analysis of the pseudoranges, the analysis method of carrier phase observations is as follows:

Condition 1: because the positions of reference and rover stations are known, so the theoretical number of total cycles from satellite k to station is also known. For reference station, the change in the number of cycles introduced by the errors associated with the receiver such as clock error is almost zero.

Condition 2: the RINEX file of the reference station gives the number of total carrier cycles between the satellite and the station without integer ambiguity, so we can obtain the change in the number of cycles introduced by common-mode errors such as ionosphere and troposphere delay.

Based on above two conditions and the results of pseudorange error analysis, we can estimate the effects in carrier phase measurements caused by clock error, but the thermal noises and multipath effects on pseudorange and carrier phase observations are different, so the impact of these errors on carrier cycles is unpredictable. However, it is still available for us to estimate the carrier phase cycles.

$$\widehat{\phi}_u = \lambda^{-1}r_u + f\widehat{\delta t}_u + \phi_{cm} \quad (5)$$

$\widehat{\phi}_u$ is the estimated value of the complete carrier cycles between satellite and smartphone, λ and f is the carrier wavelength and frequency, r_u is real distance between satellite and smartphone, the $\widehat{\delta t}_u$ is the clock error estimation which is presented in the previous section, and ϕ_{cm} is common-mode errors. This estimation does not include the impact of thermal noise, multipath effects, as well as the estimation error of the clock error.

Accumulated delta range (ADR) observations from smartphone do not resolve the integer ambiguity, and the initial value of the observations at the same moment are also different among each satellite. Therefore, we need to subtract the initial values of the respective observations, and then we can compare the trend of actual observations and estimations. In Fig. 5 (a), the estimated total carrier cycles of four satellites are given, and (b) shows the difference between the estimated value and the observed value (both have removed initial value).

The explanation of above figure (b) is as follows: First, it is observed that all four satellites have the common-mode error, this error is the estimated error of the clock error mentioned before. Second, the difference between the actual measured value and the estimated value of each satellite accumulates over time, about 2 cycle/s change trend, that is to say, carrier phase observation of smartphone has continuous loss of cycles. Since there is very little public information about GNSS chipset now, this phenomenon cannot be explained from the bottom to up. The author argues that the reason for this phenomenon lies in duty cycle mode of GNSS chipset, the APIs return cycle slip flag of ADR_STATE_CYCLE_SLIP specified that there are cycle

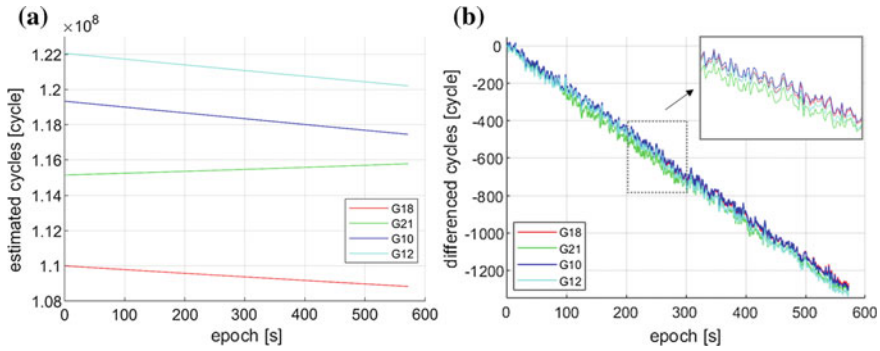


Fig. 5 Estimation and measurement error of smartphone's ADR

slips in the observations, but the specific effect of duty cycle on ADR observations is not clear in academia until today. Following will give a reasonable explanation for the phenomenon of continuous loss of cycles.

In the case of satellite G18, the doppler shift from 2135.661 to 1948.476 Hz during this observation time, and the number of carrier cycles from Fig. 5 (a) is also reduced. The diagram is as follows:

Consider the situation in Fig. 6 (a), at moment t_0 , the satellite is far from the receiver and the doppler in vector direction is large, when at moment t_1 , the satellite is close to the receiver and the doppler is small. Because the satellite gets close to the receiver, so the total number of carrier cycles is reduced. In the continuous tracking mode, the waveform is shown in the below of figure (b). Although the doppler changes, the number of cycles reduced from t_0 to t_1 can still be recorded. However in the duty cycle mode, the receiver only tracks the signal from t_0 to t_d , and it will use the number of reduced cycles during this period to estimate the entire number of reduced cycles from t_0 to t_1 .

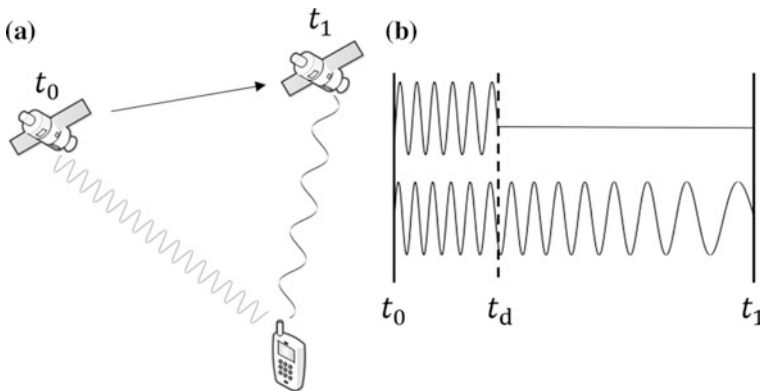


Fig. 6 The explanation for the continuous loss of cycles

$$\widehat{\Delta\phi} = \frac{t_1 - t_0}{t_d - t_0} \cdot \Delta\phi_{dc} \quad (6)$$

So the above estimation (observations in duty cycle mode) is greater than the real value (observations in continuous mode), that is to say for the same observation time, the cycle loss in duty cycle mode is larger than that in continuous mode, this is consistent with the situation of satellite G18. Because this is the case for each measurement period, the overall result is the continuous loss of cycles. Besides, it is the same reason for the case that the satellite is away from the receiver.

3 Carrier Smoothing and Pseudorange DD Algorithm

According to the conclusions of the previous section, the carrier phase observations of the smartphone within duty cycle mode have continuous loss of cycles, so the carrier phase cannot be directly used for high precision positioning now. However, it is still possible to use the carrier phase observations to smooth the pseudoranges, which will be discussed in this section. After that, the model of pseudorange DD and the selection principle of filter parameters will be introduced in this section.

3.1 Carrier Smoothing with Continuous Loss of Cycles

Unlike most survey-grade receivers, which output the smoothed pseudoranges by carrier, but the pseudorange observations of the smartphone are calculated directly by the signal transmission time, so the pseudoranges are not smoothed by carrier phase. However, it is known from previous analysis that the thermal noise power of smartphone is very large, if not eliminated, the noises will directly affect the positioning performance. It can be seen from the analysis of carrier phase that the carrier phase observations have continuous loss of cycles cannot be used directly in the classical Hatch recursion filtering algorithm. Therefore, the method of using carrier phase with continuous loss of cycles is discussed.

The formula of the classical Hatch recursive filter algorithm is as follows [6]:

$$\begin{aligned} \rho_{s,k} &= \frac{1}{M} \rho_k + \frac{M-1}{M} [\rho_{s,k-1} + \lambda(\phi_k - \phi_{k-1})] \\ \rho_{s,1} &= \rho_1 \end{aligned} \quad (7)$$

ρ_k and $\rho_{s,k}$ is the pseudoranges before smoothing and after smoothing of epoch k , the ϕ_k is the ADR observations of epoch k , λ is the carrier wavelength, M is the

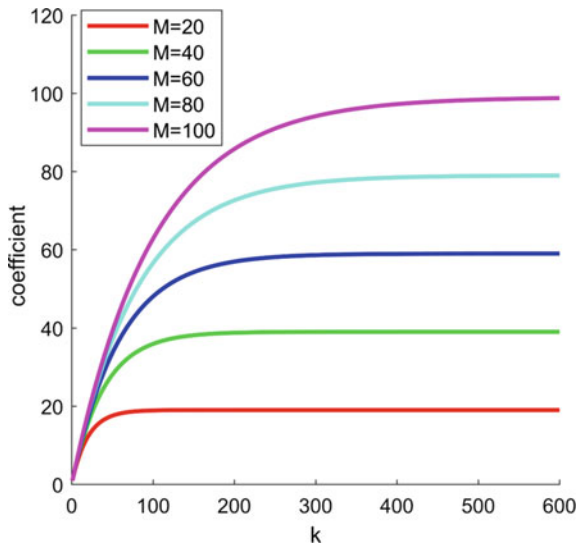
smoothing time constant. The main purpose of carrier phase smoothing is to reduce the thermal noise of the pseudorange observations, which is the term ε in Eq. (2), because this term in the carrier observation is far less than the pseudorange. Using the Hatch filter, after a period of iteration, the noise term is almost completely eliminated, and the filtered result will be a very smooth curve.

Normally, if the carrier phase has cycle slip, the Hatch filter needs to be reset when use it to smooth pseudoranges. But when just some individual cycle slips are observed in the ADR observations, the Hatch filter still can be used directly, because it will converge to the correct result over a period of time, the disadvantage is that the smoothed pseudoranges will make the position results fly away when the filter is in pull-in state. In that case, a small cycle slip has little effect on smoothing results, while a huge cycle slip will has a significant impact. When the ADR observations have continuous loss of cycles, assume that the difference of ADR between two adjacent epochs is ΔN and we can know that this value can be thought as a constant based on previous analysis. Therefore, after k iterations, this accumulated error is as following:

$$\varepsilon_{\Delta N,k} = \sum_{i=1}^k \left(\frac{M-1}{M}\right)^i \cdot \lambda \cdot \Delta N \tag{8}$$

The sum in this equation is called the coefficient, and the relationship between it and M and k is shown in the Fig. 7. Thus, this coefficient converges to a fixed value, and the larger the M , the longer the convergence time, the final convergence value is

Fig. 7 The relation between error coefficient and M



$$\varepsilon_{\Delta N, \infty} = (M - 1) \cdot \lambda \cdot \Delta N \quad (9)$$

When there are continuous loss of cycles in observations, even a very small number of cycle slips can be amplified by parameter M . However, this feature can also be used to modify the Hatch filter, so the author proposes Continuous Loss of Cycles (CLOC) Hatch filter algorithm as follows:

- (1) First, the default ΔN is zero, and use pseudorange observations minus the smoothing result, and the smoothing error is obtained.
- (2) Different from the smooth curve in Fig. 7, the smooth curve superposition of the random noise in real conditions, so we need to average the smoothing error values for each P epochs, and reduce the update rate of the data processing. The selection of P is generally $1/3$ to $1/2$ of M .
- (3) In each update cycle, the above average is summed over history, then will be averaged. This step is used to further reduce the impact of noise.
- (4) Make a difference to the historical average of each update cycle except the first value, i.e., the last moment minus the previous moment, and then we take the absolute value.
- (5) When the absolute value is smaller than a certain threshold (usually 1), the current historical average as convergence values will be substituted into Eq. (9) to obtain ΔN , then use ΔN to compensate term $\phi_k - \phi_{k-1}$ in formula (7).

3.2 Pseudorange DD and Filter Design

The positioning algorithm by using pseudorange includes absolute positioning, such as single point positioning and D-GNSS, or relative positioning, such as single difference, double difference, etc. Compared with the former, the latter can reduce the dependence on the transmission model, and can output the positioning results continuously, and can build the motion model of the receiver with Kalman filter. In addition, double difference even can eliminate the receiver clock error compared with single difference, so it is widely used in the high precision relative positioning.

The double difference observation formula of pseudorange is as follows [11]:

$$\rho_{ur}^{kj} = (\rho_u^k - \rho_r^k) - (\rho_u^j - \rho_r^j) = r_{ur}^{kj} + \varepsilon_{ur}^{kj} \quad (10)$$

In the equation, the satellite j which has the highest elevation, is the reference satellite. As we can see by the equation, the error and noise in double difference observations are almost eliminated, only the real distance (in double difference form) and term ε are remained. In addition, the pseudoranges have been smoothed by carrier phase, so the thermal noise power in the term ε_{ur}^{kj} is very small and mainly is the multipath error.

We can get $M - 1$ double difference observation equations from M satellites, and the relationship between baseline vector and DD measurements can be expressed as follows in the form of matrix:

$$\begin{bmatrix} \rho_{ur}^{21} \\ \rho_{ur}^{31} \\ \vdots \\ \rho_{ur}^{M1} \end{bmatrix} = \begin{bmatrix} -(\mathbf{1}_r^2 - \mathbf{1}_r^1)^T \\ -(\mathbf{1}_r^3 - \mathbf{1}_r^1)^T \\ \vdots \\ -(\mathbf{1}_r^M - \mathbf{1}_r^1)^T \end{bmatrix} \mathbf{b}_{ur} + \boldsymbol{\varepsilon}_{ur} \quad (11)$$

In the above equation, satellite 1 is the reference satellite. \mathbf{b}_{ur} is the baseline vector, $\boldsymbol{\varepsilon}_{ur}$ is residual error, $\mathbf{1}_r^k$ is the normalized vector from reference station to satellite k . We can use two methods to solve this equations. One is to find the least squares solution for each epoch, the result is not smooth, but we can use this method to observe the remaining errors. Another solution is Kalman filter. For static situation, the state is \mathbf{b}_{ur} , the measurements are pseudorange DD observations, the state transition matrix is a three order identity matrix, and the observation matrix is the design matrix in Eq. (11). We can use least squares to initialize, then its mean value can be used as the initial state of the Kalman filter, and the process noise covariance matrix \mathbf{Q} and measure noise covariance matrix \mathbf{R} and the initial posteriori estimation mean square error matrix \mathbf{P} can be set to the following values:

$$\begin{aligned} \mathbf{P} &= \text{diag}(d, d, d) \\ \mathbf{Q} &= \text{diag}(0.1, 0.1, 0.1) \\ \mathbf{R} &= \text{diag}(\sigma^2, \sigma^2, \sigma^2) \end{aligned} \quad (12)$$

The diagonal element of \mathbf{P} is mainly the initial error, which is introduced by multipath and can be set to 2–3. The \mathbf{Q} is not set to a zero matrix, because the error in Eq. (11) is not zero mean, so there is a slight jitter between the states. The σ^2 in \mathbf{R} can be set as error variance in Fig. 4, such as 25–100.

4 Experimental Results

4.1 The Results of Carrier Smoothing Methods

In order to evaluate the performance of carrier smoothing, two results are given. One is the differential value of pseudoranges in time, which can reflect the smoothness of pseudoranges. Another one is the smooth error between the raw pseudorange and the smoothed pseudorange, only need to remove the random noise

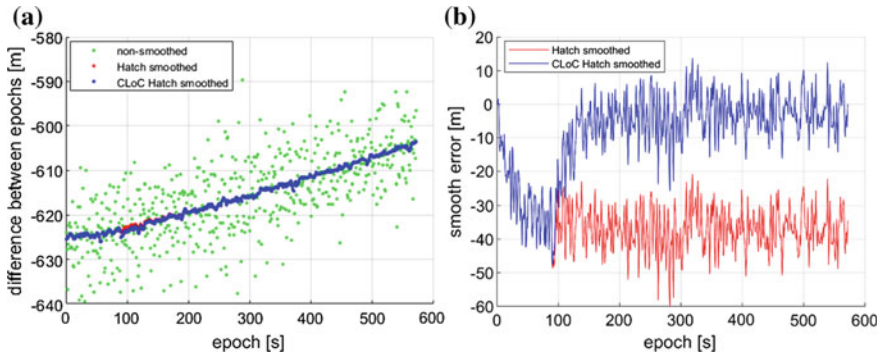


Fig. 8 The performance of CLoC Hatch filter

from the raw pseudoranges, and should not introduce any other errors. In Fig. 8 (a) we compared three cases of pseudorange differential value, we can find that the Hatch filter and CLoC Hatch filter compared with no smoothed pseudoranges, greatly reduce the influence of noise. In addition, the two smoothing methods overlap most of the time in figure (a), but figure (b) reflects the difference between the two algorithms. Figure (b) shows that Hatch filter will eventually produce more than 30 m smooth error, and we will not be able to positioning by using those smoothed pseudoranges, and CLoC Hatch Filter begins to adjust after a period of time and eventually corrected smooth error caused by continuous loss of cycles.

4.2 Positioning Methods Comparison and Multipath Effects

This section compares three methods: single point solution (SP) without carrier smoothing, single point solution with carrier smoothing (S-SP), and carrier smoothed pseudorange DD (S-PDD-KF) based on Kalman filter. Figure 9 shows the 2D positioning results of the three methods, where the origin of coordinates is the calibrated position of rover station B. By comparing SP and S-SP, the effect of random noise on positioning solution is reduced effectively by the carrier smoothing. The S-PDD-KF further reduces the influence of noise and the results are smoother. In general, the positioning accuracy of the pseudorange DD is less than 5 m. Figure 10 shows the height error of the three methods, we zoom in figure (a) to figure (b), which can show details. In order to observe the results of two methods, it can be seen that error of the pseudorange DD is smaller than the other two methods, and the error is also less than 5 m.

It can be seen from Eq. (10) that most errors have been eliminated when adopting pseudorange DD. The remaining errors in the equation are mainly thermal

Fig. 9 The 2D positioning accuracy of the three algorithms

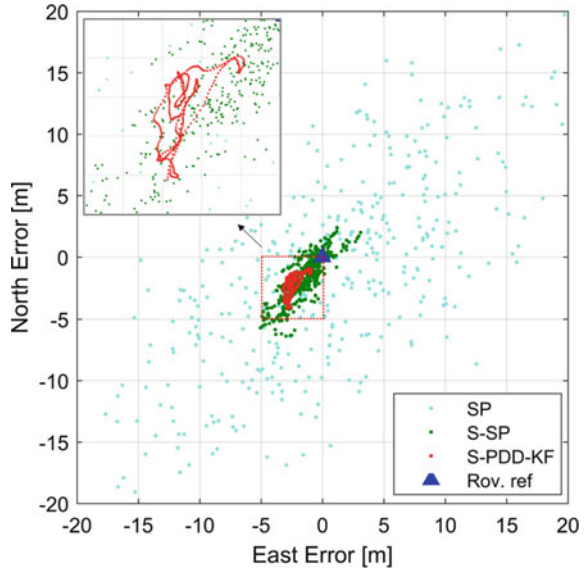
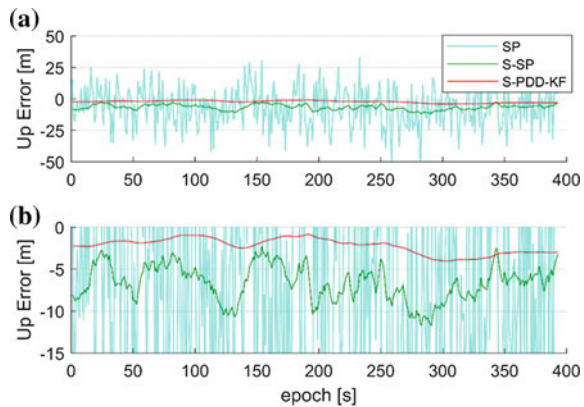
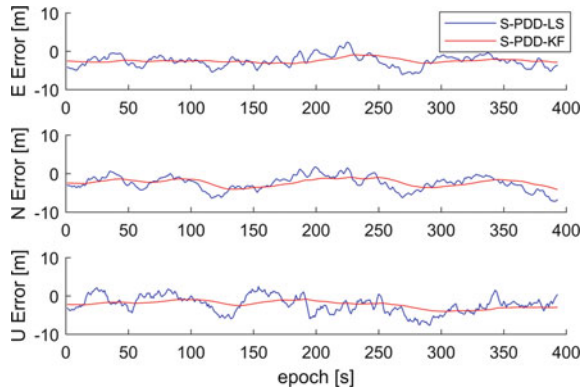


Fig. 10 The height accuracy of the three algorithms



noise and multipath error. The influence of thermal noise is greatly weakened when the carrier phase is used to smooth the pseudorange. Therefore, this rest error is mainly the influence of the multipath, because of multipath effect, resulting in Fig. 9, positioning results of pseudorange DD still have a bias from the rover reference, about 2–3 m. In Fig. 11, compared S-PDD-LS (using least squares to instead of Kalman filter) and S-PDD-KF, the ENU errors of the former are not zero mean Gaussian noises, but have bias errors which are changing over time. The Kalman filter is used to further filter the noise in three directions, but it still cannot eliminate the bias errors caused by multipath effect.

Fig. 11 Errors caused by multipath effect



5 Conclusion

In this paper, a method to analyze the measurement errors of smartphone is presented, and the characteristics of carrier phase observation in duty cycle mode are analyzed, and the causes of continuous loss of cycles are explained. Then the author proposed an improved Hatch Filter algorithm based on the carrier phase with continuous loss of cycles. Finally, pseudorange DD algorithm based on Kalman filter is used to positioning, and the positioning accuracy is improved effectively compared with other pseudorange positioning algorithms. Based on the research of this paper, the future work for realizing high precision with low cost device such as smartphone still has two challenges. First, multipath mitigation algorithm on smartphone should be studied. Second, we need to study the carrier phase reconstruction algorithm, to realize the high precision positioning based on carrier phase in duty cycle mode.

References

1. Bi J, Zhen J, Guo Y (2016) Positioning accuracy analysis of GPS and A-GPS for Android mobile phone. *Bull Surv Mapp* 7:10–13
2. Banville S, Diggelen FV (2016) Precise GNSS for everyone: precise positioning using raw GPS measurements from android smartphones. *GPS World* 27(11):43–48
3. GNSS Carrier Phase from Nexus 9. Available online: <http://rokubun.cat/2017/07/24/gnss-carrier-phasenexus-9/> (accessed on 11 Aug 2017)
4. GnsMeasurement. Available online: <https://developer.android.com/reference/android/location/GnsMeasurement.html>. Accessed on 11 Nov 2017
5. Humphreys TE, Murrian M, Diggelen FV et al (2016) On the feasibility of cm-accurate positioning via a smartphone’s antenna and GNSS chip. In: *Position, location and navigation symposium*. IEEE, pp 232–242
6. Misra P (2011) *Global positioning system: signals, measurements, and performance*. Ganga-Jamuna Press, Massachusetts

7. Malkos S (2016) User location takes center stage in new android OS: Google to provide raw GNSS measurements. *GPS World* 27:36
8. Pesyna KMJ, Heath RWJ, Humphreys TE (2014) Centimeter positioning with a smartphone-quality GNSS antenna
9. Pesyna KM, Kassas ZM, Heath RW et al (2015) A phase-reconstruction technique for low-power centimeter-accurate mobile positioning. *IEEE Trans Signal Process* 62(10):2595–2610
10. Realini E, Caldera S, Pertusini L et al (2017) Precise GNSS positioning using smart devices. *Sensors* 17(10):2434
11. Xie G (2009) *GPS principle and receiver design*. Publishing House of Electronics Industry, Beijing

Communication Tower Based Experiment and Analysis of Differential Augmentation for Auto-Steering Guidance of Agricultural Machinery



Bingbing Hu, Caicong Wu, Jinjiang Li, Xiaolong Li and Xingtao Liu

Abstract UHF based single station RTK is widely used for auto-steering guidance of agricultural machinery in China currently. Too many reference stations in a region will cause serious interference of radio frequency. Moreover, the reliability is very low, since most of the stations were built on the residential buildings. Considering the above problems, we propose to construct the reference station on the communication tower because of the advantages including distribution density, infrastructure guarantee, high quality communication, and relative height compared with ordinary building. We select three communication towers for experiment in Beijing. The average baseline is 43.3 km. We place the GNSS antennas on the roof of communication equipment room, and put the reference receivers (PD318) in the room. An agri-CORS is constructed by using PowerNetwork software. We use 4G of China Mobile to transfer observation data and ephemeris data in real time. We put the antennas of UHF radio on the tower, which is nearly 50 m high. Results show that the data integrity of three reference stations are better than 99%. The signal-to-noise ratio of L1, L2, B1, B2, and B3 are greater than 46, 35, 44, 46, and 40% respectively. The multipath of all the bands are less than 0.50. The average accuracy of baseline after adjustment is better than 0.001 m. Both average internal accord accuracy of CORS and single station RTK are better than 0.01 m, and both average external accord accuracy of CORS and single station RTK are better than

C. Wu (✉)

506 Xindian Building, 17 Qinghua East Road, Beijing 100083, China

e-mail: wucc@cau.edu.cn

B. Hu · C. Wu

College of Information and Electrical Engineering,

China Agricultural University, Beijing, China

J. Li

The Eighth & Shihezi Agriculture Mechanization Technology
Extension Station, Xinjiang Production and Construction Corps,
Shihezi, Xinjiang, China

X. Li · X. Liu

Beijing Agricultural Machinery Experiment Appraisal Popularize Station,
Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_34

0.025 m. Therefore, we get the basic conclusion that the selected communication towers are suitable for construction of GNSS reference station and the CORS and single station RTK meet the application requirement of auto-steering guidance of agricultural machinery.

Keywords CORS · RTK · Communication tower · Differential augmentation
Auto-steering guidance · Agricultural machinery

1 Introduction

Auto-steering guidance of agricultural machinery, as the important precision agriculture technology nowadays, developed quickly in recent years all over the world [1–3]. The technology can reduce labor costs, extend operating hours, and improve operation quality and land utilization [4]. The technology requires GNSS differential signal of high accuracy, high reliability, and high concurrency, in order to achieve path-to-path accuracy of ± 2.5 cm [5]. It is a worthwhile exploration to build low-cost, high precision, and reliable differential signal sources for agricultural application in China.

At present, four kinds of differential augmentation technology can provide centimeter level accuracy for agricultural machinery. These technologies include real-time kinematic (RTK), Continuous Operational Reference System (CORS), Satellite Based Augmentation Systems (SBAS), and precise point positioning (PPP) [6–8]. Among them, single station RTK and CORS have higher positioning accuracy and shorter convergence time, and depend on digital radio and mobile communication networks. PPP's optimum positioning accuracy is about 3.8 cm, and has been used in precision agriculture in recent years [9]. However, PPP needs long convergence time about 30 min. 2015 precision agricultural services dealership survey report [10] showed that 11.7% of respondents in the United States built single station RTK, 25.7% of respondents purchased virtual reference services (VRS), and 27.2% of respondents purchased OminSTAR XP or HP, StarFire 2. According to the author's survey, auto-steering guidance of agricultural machinery in China almost make use of the fixed or mobile single station RTK.

To improve the reliability of differential augmentation, Trimble, John Deere, and related agricultural companies provide more than two differential resources for agricultural machinery. For example, Trimble RTX technology can maintain RTK solutions in 20 min after losing RTK or VRS services [11, 12]. RTX uses single station RTK or CORS as the main differential source, and utilizes satellite-based PPP as the backup signal source. The two kinds of differential augmentation signal can switch between each other instantaneously without any delay.

Auto-steering guidance of agricultural machinery in China developed quickly in recent years [4, 13]. The suppliers generally construct single station RTK and broadcast differential data through UHF radio. However, too many UHF radios cause heavily problems, such as high redundancy, low reliability, and frequency

interference. In addition, coverage uncertainty of UHF radio exists due to irregular terrain and other environmental impact. In Shihezi, the stations even built by the same company encounters frequency interference. When rover switches from one station to another without network adjustment, auto-steering guidance system would change navigation line [14], resulting in disorder of farm operations. CORS covers a much larger area compared with single station RTK and it can maintain a consistent reference for coordinative operation and repetitive work. It is meaningful to make use of both CORS and single station RTK. In practice, most of the agricultural GNSS receivers are compatible with CORS and UHF radio based RTK.

During the past several decades, three mobile communication operators in China, i.e., China Mobile, China Unicom, and China Telecom, have built a total of about 1.5 million communication towers, accounting for 65% of the world [15]. In 2014, Chinese government created state owned company, China Tower, to manage all the communication towers. The towers have the obvious advantages of large density, reliable infrastructure, and good communication access, which ordinary residential houses do not have. Therefore, the authors infer that the communication towers might be suitable for construction of agricultural GNSS reference stations. However, we also need to consider about the disadvantages of the tower, such as swaying, signal blockage, multipath, and frequency interference. Can the communication towers meet the requirements of GNSS station construction? We think it is necessary to do some further researches and explorations. Therefore, this paper proposes to make use of the communication tower and its infrastructure to construct a CORS for experiments, in order to explore the new differential augmentation mode for agricultural application.

2 Material and Method

In this section, we introduce the construction process of CORS and validating method of the network and single station RTK.

2.1 Reference Station Construction

Three communication towers were selected for CORS construction from three suburban districts of Beijing, i.e., Shunyi, Pinggu, and Tongzhou. The station's name are NDSY, NDPG, and NDTZ. The average baseline is about 40 km, considering the effective radio transmission radius usually not more than 20 km. To avoid the shake of the tower, we decided to place the GNSS receiver antenna on the roof of the communication equipment room, and to fix the UHF antenna on the top of the tower, which cable is less than 30 m.

We choose PD318 as the reference receiver, which can track BeiDou B1, B2, B3, GPS L1, L2, L5, and GLONASS L1, L2. The receiver supports CMR, CMR+, RTCM2.1, RTCM2.2, RTCM2.3, RTCM3.1, RTCM3.2, and related differential formats. The rover uses S9 receiver, which can track the same band of reference receiver. Each reference receiver uses China Mobile SIM card to transmit original observation data to the data center in real time with the transmission frequency of 1 Hz. Meanwhile, we use HX-U202 UHF digital radio to broadcast differential data. The frequencies of UHF radio of NDSY, NDPG, and NDTZ are set up as 441.125, 440.125, and 442.125 MHz.

We rent a server for CORS from Alibaba Cloud (Aliyun), which uses Intel Xeon CPU E5-2630@2.30 GHz, memory 8G, and Windows Server 2008 R2 Standard 64-bit operating system. PowerNetwork software developed by Wuhan University is used for data processing and VRS service.

2.2 Evaluation Method

Three reference stations synchronously recorded 4-day observation data with a sampling interval of one second. We use DataQC software to evaluate data integrity, signal-to-noise ratio, and multipath effect [16] of the observation data. The data integrity reflects the availability of the data in the range and is an important indicator of data quality. The signal-to-noise ratio is the ratio of the carrier signal strength received by the GNSS receiver to the observed noise strength. The greater the signal-to-noise ratio, the weaker the observed noise intensity and the better the quality of the received carrier observations. The multipath effect is a delay phenomenon of the composite signal generated by the transmitting satellite, the receiver antenna, and the relative spatial relationship between objects in the vicinity of the antenna with respect to the direct receiver antenna signal. Suppose GPS L1 and L2 signal propagation path is the same, while assuming BDS B1, B2 and B3 signal propagation path is exactly the same, and then calculate the signal corresponding MP1, MP2, and MP3 multipath effect. The smaller the multipath effect value is, the stronger the anti-multipath effect capability is.

We arranged experiments to validate both internal and external accuracy of CORS and single station RTK. Five cites were selected for validations, and four of them lie in the coverage of CORS network. 5# is very near to the NDPG station. On September 15, 2017, to validate the availability, we connected two vehicle-mounted RTK receivers to the CORS and single reference stations respectively. We also placed the rover at five cites to validate the performance of CORS and single station RTK with the duration of 8 h.

3 Results and Discussion

In this section, we give the main results and discuss the cause or phenomenon.

3.1 Data Quality Evaluation

The evaluation results of data integrity are shown in Table 1. We can find that the data integrity of all three reference stations are more than 98%, which means that all the receivers have very good data integrity.

From the above table, we can find current total of 31 GPS operational satellites are all visible, and total of 13 BDS MEO satellites are visible, which indicate that the reference stations have a good observation vision and the roofs of equipment room of communication tower are suitable for placing GNSS antenna. The result is quite important for the further research of this paper.

The evaluation results of signal-to-noise ratio are shown in Table 2. We can find that the signal-to-noise ratios of GPS L1 and L2 of all the reference stations are greater than 46 and 35% respectively, and the signal-to-noise ratios of BDS B1, B2, and B3 are greater than 44, 46, and 40% respectively. Therefore, we can conclude that all the three reference stations meet the construction requirements of reference station.

The evaluation results of multipath are shown in Table 3. The multipath index of NDPG station is relatively large, but it is less than 0.50, which indicate that the multipath effects of all the reference stations meet the construction requirements. The multipath effect error of the GPS L1 is always less than that of L2, whereas BDS B2 is always smaller than that of B1 and B3.

3.2 CORS Network Adjustment

The results of network adjustment as shown in Table 4. We can find that the internal accord accuracy of each reference station is better than 0.0003 m.

Based on the above experiments, we construct the agri-CORS with three reference stations, as shown in Fig. 1. The average baseline is 43.3 km.

In further, NDPG and NDSY have very good transmission reliability. During 24 h, they only missed 5 epochs and 6 epochs respectively. While NDTZ missed 848 epochs, its availability is still over 99%. Therefore, to some extent, we think that 4G wireless network is suitable for observation data transmission. However, we also need to improve the transmission reliability by using backup transmission method, since auto-steering guidance of agricultural machinery need extremely high reliability.

Table 1 Evaluation of data integrity of three reference stations in Beijing

	NDPG		NDSY		NDTZ	
	GPS	BDS	GPS	BDS	GPS	BDS
Total number of satellites observed	31	13	31	13	31	13
Total number of observable ephemeris	86,400	86,400	86,400	86,400	86,400	86,400
Actual intact observation epoch	86,400	86,400	86,400	86,400	86,400	86,400
Estimated total number of observations > 0.00°	940,864	867,678	940,930	868,044	941,415	869,003
Estimated total number of observations > 10.00°	747,447	807,866	747,052	807,913	747,954	808,767
Total number of pseudoranges/phase observations received	740,850	797,563	739,138	801,275	745,656	801,988
Delete the total number of observations (for any reason)	6494	3372	4288	307	3586	301
The number of intact observations > 10.00°	734,356	794,191	734,850	800,968	742,070	801,687
Delete the total number of observations > 10.00°	6142	3095	3942	37	3214	40
Intact observation ratio	98	98	98	99	99	99

Table 2 Evaluation results of signal-to-noise ratio of three reference stations in Beijing

GNSS system	Band	NDPG	NDSY	NDTZ
GPS	L1	46.22	46.32	46.08
	L2	36.63	36.67	36.65
BDS	B1	44.51	44.80	44.76
	B2	46.76	47.52	47.70
	B3	42.91	42.54	43.41

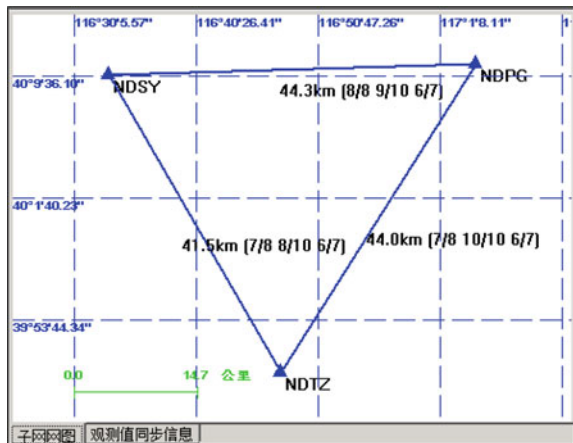
Table 3 Evaluation of multipath analysis of three reference stations in Beijing

	NDPG		NDSY		NDTZ	
	GPS	BDS	GPS	BDS	GPS	BDS
MP1RMS average value (m)	0.3582	0.3096	0.2547	0.2488	0.2866	0.2432
MP2RMS average value (m)	0.4966	0.2162	0.3808	0.1658	0.3947	0.1576
MP3RMS average value (m)	/	0.2427	/	0.2207	/	0.1908

Table 4 Internal accord accuracy of reference station after network adjustment

Station	X (m)	Y (m)	Z (m)
NDPG	0.0003	0.0003	0.0003
NDSY	0.0002	0.0003	0.0002
NDTZ	0.0002	0.0003	0.0002

Fig. 1 Agri-CORS network with three reference stations



3.3 Validating of Positioning Accuracy

As Fig. 2 shows, the average internal accord accuracy of CORS achieve 0.01 horizontal accuracy and 0.006 m vertical accuracy. The worst horizontal and vertical internal accord accuracy are 0.01 and 0.018 m. The average external accord

Fig. 2 Internal and external accord accuracy of CORS. Internal_H and Internal_V mean internal accord accuracy in horizontal and vertical direction. External_H and External_V mean external accord accuracy in horizontal and vertical direction

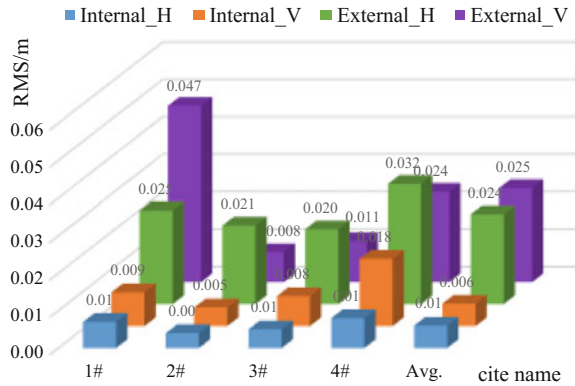
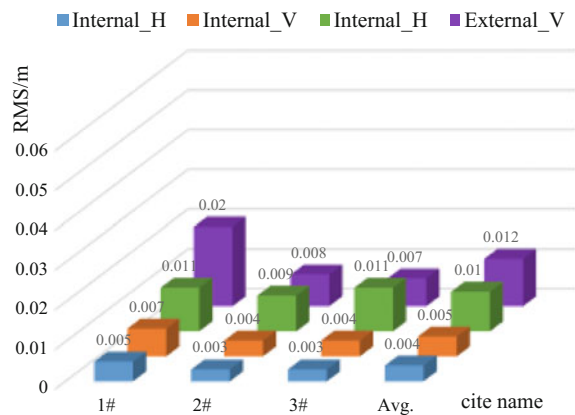


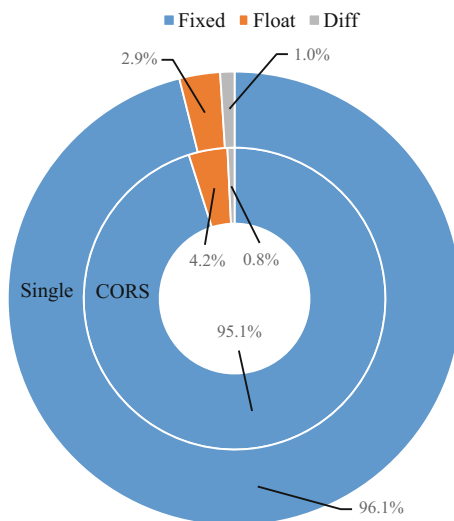
Fig. 3 Internal and external accord accuracy of single station RTK



accuracy of CORS achieve 0.024 m horizontal accuracy and 0.025 m vertical accuracy. The worst horizontal and vertical external accord accuracy are 0.032 and 0.047 m respectively. The results show that the agri-CORS can meet the requirements of auto-steering guidance of agricultural machinery.

As Fig. 3 shows, the average internal accord accuracy of single station RTK achieve 0.004 m horizontal accuracy and 0.005 m vertical accuracy. The worst horizontal and vertical internal accord accuracy is 0.005 and 0.007 m respectively. The average external accord accuracy of single station RTK achieve 0.011 m horizontal accuracy and 0.012 m vertical accuracy. The worst horizontal and vertical external accord accuracy achieve 0.011 and 0.02 m respectively. The results show that the single station can also meet the requirements of auto-steering guidance of agricultural machinery.

Fig. 4 Availability of CORS and single reference station RTK



3.4 Availability Evaluation

The ratio of fixed solution, float solution, and differential solution of static positioning are shown in Fig. 4. The proportions of fixed solutions of CORS and single stations reach 95.1 and 96.1% respectively. The experiments show that it is impossible to achieve a 100% complete fixed solution whether by CORS or by single station RTK. Therefore, it is necessary to make use of the above two differential resources to improve the success rate of positioning. In theory, the fixed solution can be increased to 99.8%, while fixed solution or float solution can be increased to 99.99% by using both CORS and single reference station RTK.

The spatial availability of CORS was tested both inside and outside the network. The main reason for the non-fixed solution is that there are shelterbelt along the test road and the vehicle encountered serious satellite occlusion. However, the fixed solution can be obtained in different regions, so the spatial availability of CORS satisfies the system design.

4 Conclusions

We placed GNSS antennas on the roof of equipment room of communication tower, and find that the data integrity, multipath effect, and signal-to-noise ratio meet the construction requirement of reference stations. The data integrity is higher than 99%. The signal-to-noise ratios of GPS and BDS are all greater than 44%. Therefore, we get the basic conclusion that the selected communication towers are suitable for the construction of GNSS reference station. In further, we validate the

positioning accuracy of CORS and single station RTK. Both average internal accord accuracy of CORS and single station RTK are better than 0.01 m, and both average external accord accuracy of CORS and single station RTK are better than 0.025 m. Therefore, we infer that the CORS and single station RTK meet the application requirement of auto-steering guidance of agricultural machinery.

With the above conclusion, we plan to move all single stations build on the residential houses to the communication towers in Shihezi, to explore new construction mode of CORS or single station RTK for agricultural application.

Acknowledgements Thanks for the support of National Key Research and Development Program of China (No. 2016YFB0501805). Thanks for Wuhan Panda Space and Time Technology Ltd and Beijing Zhongguangtiangong Technology Ltd. who provided GNSS reference receivers and communication tower for experiments.

References

1. Bechar A, Vigneault C (2017) Agricultural robots for field operations. Part 2: operations and systems. *Biosys Eng* 153:110–128. <https://doi.org/10.1016/j.biosystemseng.2016.11.004>
2. Han X, Kim H-J, Jeon CW, Moon HC, Kim JH (2017) Development of a low-cost GPS/INS integrated system for tractor automatic navigation. *Int J Agric Biol Eng* 10(2):123–131. <https://doi.org/10.3965/j.ijabe.20171002.3070>
3. Bechar A, Vigneault C (2016) Agricultural robots for field operations: concepts and components. *Biosys Eng* 149:94–111. <https://doi.org/10.1016/j.biosystemseng.2016.06.014>
4. Hu J, Gao L, Bai X, Li T, Liu X (2015) Review of research on automatic guidance of agricultural vehicles. *Nongye Gongcheng Xuebao/Trans Chin Soc Agric Eng* 31(10):1–10. <https://doi.org/10.11975/j.issn.1002-6819.2015.10.001>
5. Rovira-Más F, Chatterjee I, Sáiz-Rubio V (2015) The role of GNSS in the navigation strategies of cost-effective agricultural robots. *Comput Electron Agric* 112:172–183. <https://doi.org/10.1016/j.compag.2014.12.017>
6. Zhang B-C, Odijk D (2015) A method for processing GNSS data from regional reference networks to enable single-frequency PPP-RTK. *Acta Geophys Sin* 58(7):2306–2319. <https://doi.org/10.6038/cjg20150709>
7. Zhang X, Li X, Li P (2017) Review of GNSS PPP and its application. *Cehui Xuebao/Acta Geod et Cartographica Sin* 46(10):1399–1407. <https://doi.org/10.11947/j.AGCS.2017.20170327>
8. Odijk D, Zhang B, Teunissen PJG (2015) Multi-GNSS PPP and PPP-RTK: some GPS + BDS results in Australia. Paper presented at the 6th China satellite navigation conference, CSNC 2015, May 13, 2015–May 15, 2015, Xian, China
9. Siejka Z (2014) Verification of the usefulness of the trimble RTX extended satellite technology with the xFill function in the local network implementing RTK measurements. *Artif Satell* 49(4):191–209. <https://doi.org/10.2478/arsa-2014-0015>
10. Erickson B, Widmar DA (2015) 2015 precision agricultural services dealership survey results
11. Carballido J, Perez-Ruiz M, Emmi L, Aguera J (2014) Comparison of positional accuracy between RTK and RTX GNSS based on the autonomous agricultural vehicles under field conditions. *Appl Eng Agric* 30(3):361–366. <https://doi.org/10.13031/aea.30.10342>
12. Landau H, Brandi M, Chen X, Drescher R, Glocker M, Nardo A et al (2013) Towards the inclusion of Galileo and Beidou/compass satellites in trimble centerpoint RTX. Paper presented at the 26th international technical meeting of the satellite division of the institute of

- navigation, ION GNSS 2013, September 16, 2013–September 20, 2013, Nashville, TN, United states
13. Ji C, Zhou J (2014) Current situation of navigation technologies for agricultural machinery. *Nongye Jixie Xuebao/Trans Chin Soc Agric Mach* 45(9):44–54. <https://doi.org/10.6041/j.issn.1000-1298.2014.09.008>
 14. Hameed IA, Bochtis DD, Sorensen CG, Noremark M (2010) Automated generation of guidance lines for operational field planning. *Biosys Eng* 107(4):294–306. <https://doi.org/10.1016/j.biosystemseng.2010.09.001>
 15. Shi Q (2016) Research on the development strategy and operation strategy of China tower. NanJing University of Posts and Telecommunications, Nanjing
 16. Strode PRR, Groves PD (2016) GNSS multipath detection using three-frequency signal-to-noise measurements. *GPS Solutions* 20(3):399–412. <https://doi.org/10.1007/s10291-015-0449-1>

Low-Cost Receiver Robust-RTD Algorithm Based on the Constraints of Velocity



Jian Zhang, Guorong Yu, Shuguo Pan, Guoliang Liu and Fei Shen

Abstract The improvement of low-cost receiver performance and the development of Multi-GNSS will make high-precision dynamic positioning, which is for the general public, become the main growth point of GNSS applications gradually. However, under the complex road conditions of the city, there are apparent facts that seriously affect the continuity and reliability of GNSS positioning, like satellite occlusion, the multi-path effects and so on, and restrict the rapid development of GNSS mass navigation. In this paper, based on the static and dynamic data of dual system (GPS + BDS) acquired by ublox M8T module and ordinary ceramic patch antenna, an algorithm of phase smoothed pseudorange is used to improve the accuracy of pseudorange observations, the speed and direction of the receiver are constrained by doppler observations. The joint pseudorange and velocity pseudo-observations are used to establish the robust kalman filter for real-time dynamic positioning. The experimental results show that the proposed algorithm can guarantee the fast and reliable positioning of the low-cost receiver, under the open environment, the sub-meter level positioning accuracy can be basically achieved, and the reliable positioning within 2 m can be guaranteed effectively; in the complex environment, the positioning accuracy within 5 m can also be basically guaranteed.

Keywords Low-cost · Robust kalman · RTD · Velocity determination

J. Zhang · G. Yu (✉)

School of Transportation, Southeast University, Nanjing 210096, China

e-mail: 476310930@qq.com

S. Pan (✉)

School of Instrument Science and Engineering, Southeast University,

Nanjing 210096, China

e-mail: psg@seu.edu.cn

G. Liu

Nanjing Compass Navigation Technology Co., Ltd, Nanjing 210096, China

F. Shen

JiangSu BDS Application Industry Institute, Nanjing 210032, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_35

1 Instruction

At present, the level of urbanization has reached 60% in our country. With the economic development and the expansion of urban scale, mass navigation has become the main growth point of GNSS application. For the traditional surveying and mapping industry, the survey-grade receivers were expensive and bulky, these defects greatly restricted the development of GNSS application in most civilian areas. Therefore, the low-cost and small GNSS navigation devices have become the key to the development of mass navigation.

Zuo et al. [1] proposed a combined BDS/GPS RTK (Real-time kinematic) algorithm to obtain centimeter level positioning using a low-cost single-frequency helix antenna. Mongredien et al. [2] achieved centimeter level positioning accuracy using ublox NEO-M8P receivers. The above research greatly satisfied the demands of many industries, like unmanned aerial vehicles (UAV), precision agriculture and so on. However, in complex environment of the city, because the satellite occlusion is serious, there are few satellites and the visible satellites are not continuous in serial time, which lead lack of carrier phase observations and the occurrence of cycle-slip, centimeter RTK positioning is difficult. In fact, for many civilian areas, users may not need centimeter level or millimeter level positioning accuracy, such as LBS (Location Based Service), but instead focus on the continuity, reliability and environmental adaptability of GNSS positioning. Peng et al. [3] established kalman filter based on pseudorange and doppler observations using ublox LEA-6T module and ordinary navigation antenna, it improve the continuity of positioning under urban road conditions, however, this method still has some shortcomings in the reliability of positioning.

From the practical value, taking into account problems of the carrier phase observations in the complex environment, this paper uses RTD (Real Time Differential) positioning mode to ensure the continuity of positioning. However, in view of the impacts of the pseudorange observation noise and multi-path effects, this paper uses the phase smoothed pseudorange algorithm [4–8] to improve the accuracy of pseudorange observations under the condition of keeping carrier phase observation quality. And the use of robust M-LS filter [9] eliminates the impact of gross errors on the filtering solution, it ensures the reliability of positioning. In fact, the robust filtering algorithm obtains the optimal estimation by adjusting gain matrix, it overcomes the influence of the observed model error and the dynamic model error [10].

In this paper, the formula of velocity determination first is deduced using doppler observations, and the obtained velocity, which is used as pseudo-observations, combined with pseudorange observations establish the robust kalman filter to constrain positioning solution. Then, the proposed algorithm is verified by the collected data, the positioning performance of the algorithm is analyzed by the statistical analysis of horizontal positioning. A summary and future research is finally given.

2 RTD Algorithm

2.1 Velocity Determination Based on Doppler Observations

Doppler shift observations represent the magnitude of the doppler effect caused by the relative motion of the satellite and GPS receiver antennas, that is, the instantaneous observations of the change rate of carrier phase observations [11].

GPS pseudorange observation equation is as follows:

$$P_m^s = \rho_m^s + d_{t_m} - d_{t^s} + I_m^s + T_m^s + \varepsilon \tag{2.1}$$

The superscript s denotes satellites; and the subscript m denotes receivers; P is pseudorange observation; ρ represents true geometric distance between satellite antenna phase center and receiver antenna phase center (m); d_{t_m} , d_{t^s} are the satellite and receiver clock errors (m), respectively; I is the ionosphere delay (m); T is the slant troposphere delay (m); ε is the unmodeled error.

From the convenience of discussion, the troposphere delay and ionosphere delay are combined into Δ . The distance of satellite and receiver ρ in Eq. (2.1) is linearized, and the above equation get the total differential:

$$\delta \dot{P}_m^s = \begin{bmatrix} \frac{X_m - X^s}{\rho_m^s} & \frac{Y_m - Y^s}{\rho_m^s} & \frac{Z_m - Z^s}{\rho_m^s} \end{bmatrix} \begin{bmatrix} \delta X_m - \delta X^s \\ \delta Y_m - \delta Y^s \\ \delta Z_m - \delta Z^s \end{bmatrix} + \delta d_{t_m} - \delta d_{t^s} + \delta \Delta_m^s \tag{2.2}$$

where (X^s, Y^s, Z^s) is satellite coordinate in ECEF frame; (X_m, Y_m, Z_m) is receivers antenna coordinate.

Divide both sides of Eq. (2.2) by time and take zero limit:

$$\dot{P}_m^s = \begin{bmatrix} \frac{X_m - X^s}{\rho_m^s} & \frac{Y_m - Y^s}{\rho_m^s} & \frac{Z_m - Z^s}{\rho_m^s} \end{bmatrix} \begin{bmatrix} \dot{X}_m - \dot{X}^s \\ \dot{Y}_m - \dot{Y}^s \\ \dot{Z}_m - \dot{Z}^s \end{bmatrix} + \dot{d}_{t_m} - \dot{d}_{t^s} + \dot{\Delta}_m^s + \dot{\varepsilon} \tag{2.3}$$

where \dot{P} is pseudorange rate (m/sec); $(\dot{X}_m, \dot{Y}_m, \dot{Z}_m)$ is receivers velocity; $(\dot{X}^s, \dot{Y}^s, \dot{Z}^s)$ is satellite velocity; \dot{d}_{t_m} , \dot{d}_{t^s} are receiver and satellite clock drift, respectively; $\dot{\Delta}$ is the change rate of the atmosphere delay; $\dot{\varepsilon}$ is the unmodeled error.

The above only deduce velocity determination equation of pseudorange rate, pseudorange rate and doppler has the following relationship:

$$\dot{P} = -\lambda D \tag{2.4}$$

In Eq. (2.3), the receivers coordinate can be obtained by single point positioning. Assuming that the receivers coordinate error are 10 m, the effect on velocity accuracy is about 2 mm [12]; the satellite coordinate can be calculated by the numerical differencing method based on position series of the satellite [13]. The

satellite clock drift can be obtained by the change of the satellite clock error over a period of time. The effect of the change rate of atmosphere delay is ignored in this paper.

The above velocity deduction is mainly used to obtain the receiver’s velocity, and it provides pseudo-observations for pseudorange differential positioning to constrain the position solution.

2.2 RTD Algorithm Based on the Constraints of Velocity

The pseudorange differential equation of j satellite can be expressed as follows:

$$\Delta\nabla P_{m,n}^{i,j} = \Delta\nabla\rho_{m,n}^{i,j} + \Delta\nabla T_{m,n}^{i,j} + \Delta\nabla I_{m,n}^{i,j} + \varepsilon \tag{2.5}$$

where $\nabla\Delta$ denote double difference operator; the superscript i and j denote the reference satellite and non-reference satellite, respectively; and the subscript m and n denote the base-station and rover-station, respectively; P is smoothed pseudorange observation (m); ρ represents true geometric distance between satellite antenna phase center and receiver antenna phase center (m); I is the ionosphere delay (m); T is the slant troposphere delay (m); ε is the unmodeled error.

Single epoch GPS + BDS observation equation can be expressed as follows:

$$V = BX - L \tag{2.6}$$

where:

$$V = \left[V_{P_{i,G}^j} \quad \cdots \quad V_{P_{i,C}^k} \quad V_{\dot{X}} \quad V_{\dot{Y}} \quad V_{\dot{Z}} \right]^T \tag{2.7}$$

$$X = \left[\Delta X \quad \Delta Y \quad \Delta Z \quad \Delta\dot{X} \quad \Delta\dot{Y} \quad \Delta\dot{Z} \right]^T \tag{2.8}$$

$$B = \begin{bmatrix} l_{i,G}^j & m_{i,G}^j & n_{i,G}^j & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & 0 & 0 & 0 \\ l_{i,C}^k & m_{i,C}^k & n_{i,C}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & & \\ 0 & 0 & 0 & & 1 & \\ 0 & 0 & 0 & & & 1 \end{bmatrix} \tag{2.9}$$

$$L = \left[L_{P_{i,G}^j} \quad \cdots \quad L_{P_{i,C}^k} \quad L_{\dot{X}} \quad L_{\dot{Y}} \quad L_{\dot{Z}} \right]^T \tag{2.10}$$

where G denotes GPS system; C denotes BDS system; l, m, n represent the first order Taylor series expansion; X denotes the estimated parameter vector; L denotes the observation vector.

In above model, the paper uses velocity pseudo-observations to constrain the estimated parameters, the equations of velocity pseudo-observation are as follows:

$$\tilde{X} = \dot{X} \tag{2.11}$$

where \dot{X} is the estimated parameters; \tilde{X} is a priori velocity pseudo-observations. The value can be obtained by Sect. 2.1. The above equation also applies to Y, Z directions.

3 Experiment and Analysis

The experiment uses the ordinary ceramic patch antenna, the module is ublox M8T, the sample interval of the data is 1 s, the elevation cut-off angle is 12° . RTK data is obtained synchronously by ZHD H32 survey-grade receiver, as reference value of the coordinate. The relationship between the antenna of the ublox M8T and the ZHD H32 receiver is shown in Fig. 1. The GPS + BDS data (static and pedestrian dynamic in open environment, pedestrian dynamic in complex environment and vehicle dynamic in the urban road) were analyzed from 4 different time periods in October 2017.

Fig. 1 The relationship between the antenna of the ublox M8T and the ZHD H32 receiver



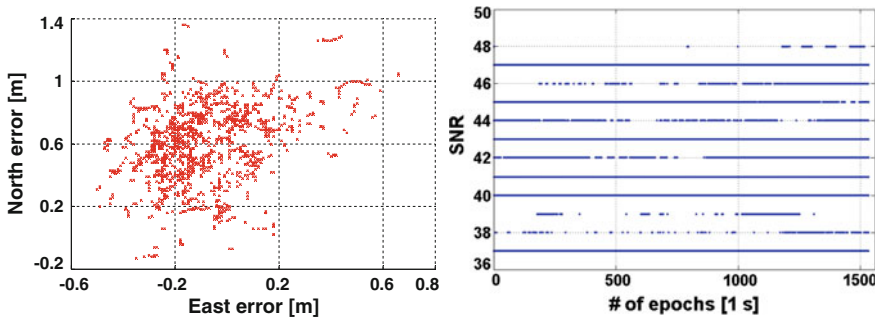


Fig. 2 The horizontal (N, E) position scatter (left) and SNR (right) of static test on open environment

Table 1 The statistics analysis of static test on open environment

Content	Precision (m)				
	RMS	≤ 0.5 (%)	≤ 1 (%)	≤ 3 (%)	≥ 3
N	0.66	30.5	96.5	100	0
E	0.21	99.0	100	100	0
Horizont	0.69	25.1	95.4	100	0

3.1 Static Test on Open Environment

Figure 2 shows the horizontal positioning performance and SNR under the static condition of the open environment. The reference value of static coordinate takes RTK fixed solution of ZHD H32 receiver. It can be seen from Fig. 2 (right) that the SNR of satellites is better than 30 in an open environment, and the signal quality is better, which also guarantees the horizontal positioning performance. Table 1 presents the statistical analysis of the horizontal positioning performance in open environment.

As Table 1 is shown, RMS of the horizontal positioning is 0.69 m and the rate of the points which are better than 1 m is 95.4%. And it is seen from Fig. 2, the whole points are better than 2 m in open environment.

3.2 Dynamics Test on Open Environment

The test was conducted on the roof of Southeast University ZhongXin Building. And Fig. 3 shows the track of dynamic test on open environment. In Fig. 3, the white points are fixed resolution of RTK which were regarded as reference value, and the red points are test results of the self-made program. Figures 4 shows the horizontal positioning performance and SNR under the dynamic condition of the open environment. The number of points in Fig. 4 (left) is not equal to the number



Fig. 3 The track of dynamical test on open environment

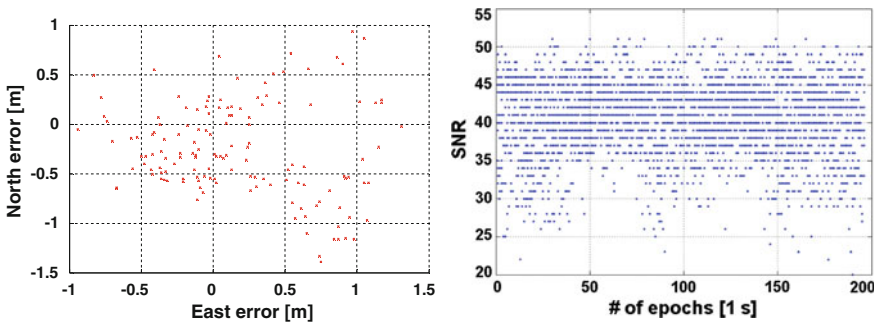


Fig. 4 The horizontal (N, E) position scatter (left) and SNR (right) of dynamical test on open environment

Table 2 The statistics analysis of dynamical test on open environment

Content	Precision (m)				
	RMS	≤ 0.5 (%)	≤ 1 (%)	≤ 3 (%)	≥ 3
N	0.53	58.0	94.4	100	0
E	0.55	61.5	92.3	100	0
Horizont	0.77	37.7	76.9	100	0

of epochs in Fig. 4 (right). It results from the loss of lock when the receiver pass through a house in the measured environment (Fig. 3). Table 2 presents the statistical analysis of the horizontal positioning performance according to Fig. 4 (left).

It can be seen from Fig. 4 (right) that under dynamic conditions, the signal quality received by the low-cost receiver is inferior to the static condition, which is also confirmed by the horizontal positioning performance (Fig. 4 (left)). However, as can be seen from Table 2, under dynamic conditions, the rate of the points which are better than 0.5 m is superior to the static condition, which may be related to the length of the collected data. As Table 2 is shown, under dynamic conditions, RMS of horizontal positioning is 0.77 m. And it is seen from Fig. 4 (left), the whole points are better than 2 m in open environment.

3.3 Dynamics Test on Complex Environment

The test was conducted on playground of Southeast University on October 18, 2017. The playground is surrounded by lush trees and tall buildings, the observation environment is terrible. And Fig. 5 shows the track of dynamic test on the complex environment. In Fig. 5, the white points are fixed resolution of RTK which were regarded as the reference value, and the red points are test results of self-made program. Figure 6 shows the horizontal positioning performance and SNR of dynamic test on complex environment.

Because of the loss of lock, the number of points in Fig. 6 (left) is not equal to the number of epochs in Fig. 6 (right). Compared with Fig. 4 (right), signal quality

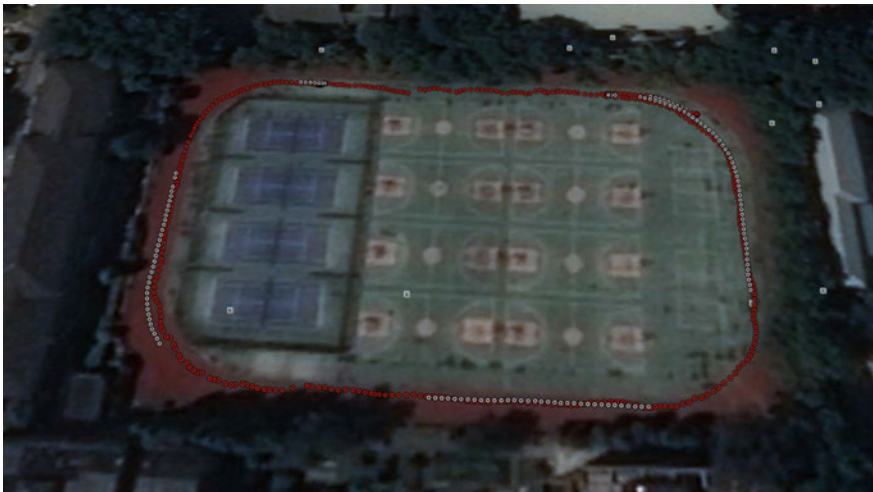


Fig. 5 The track of dynamic test on complex environment

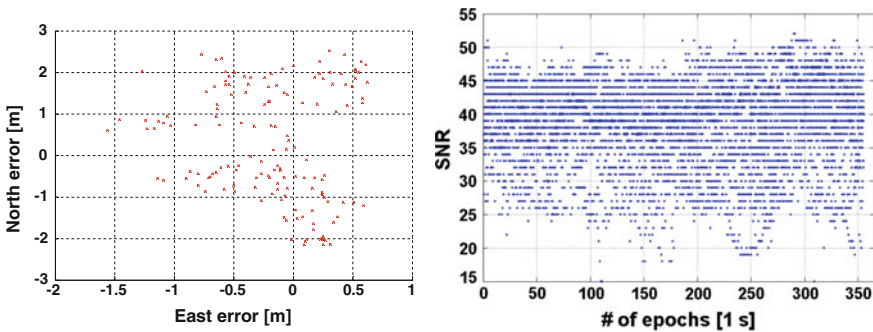


Fig. 6 The horizontal (N, E) position scatter (left) and SNR (right) of dynamic test on complex environment

Table 3 The statistics analysis of dynamic test on complex environment

Content	Precision (m)				
	RMS	≤ 0.5 (%)	≤ 1 (%)	≤ 3 (%)	≥ 3
N	1.39	15.4	44.3	100	0
E	0.52	65.1	92.0	100	0
Horizont	1.48	8.1	33.6	100	0

of low-cost receiver is worse on complex environment. Table 3 shown the statistics analysis of horizontal positioning performance on complex environment. As it is shown, RMS of horizontal positioning is 1.48 m. However, the rate of the points which are better than 1 m is 33.6%, and the whole points are better than 3 m.

3.4 Dynamic Car-Borne Test

Figure 7 shows the track of dynamic car-borne test on urban road on October 23, 2017, the red points are test results of self-made program, the blue points are reference values of the coordinate. The reference values are obtained by the post-resolution of integrated GNSS/INS BDS terminal dynamic detection system of Nanjing Institute of Measurement and Testing Technology, the plane precision of post-resolution is better than 1 cm.

As can be seen from SNR of dynamic car-borne test shown in Fig. 8 (left), signal quality of low-cost receiver is poor on most epochs, combining with the number change of satellites in Fig. 8 (right), when traveling to city canyon and other complex environment, the road is lined with lush trees and tall buildings, resulting in serious blockage of satellite signals and obvious multipath effects. Especially around the 1900th epoch, when traveling to the campus of Southeast University, there are tall trees on both sides of the road, as can be shown from the sky plot in Fig. 9 (left), the satellite structure is terrible, resulting in the plane positioning performance of some epochs have a large offset. Combined with the horizontal positioning

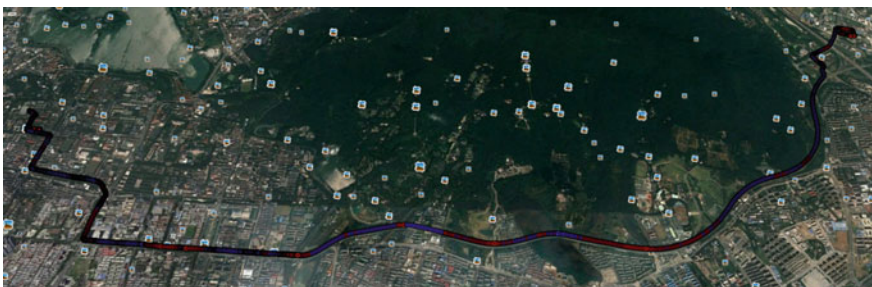


Fig. 7 The track of dynamic car-borne test

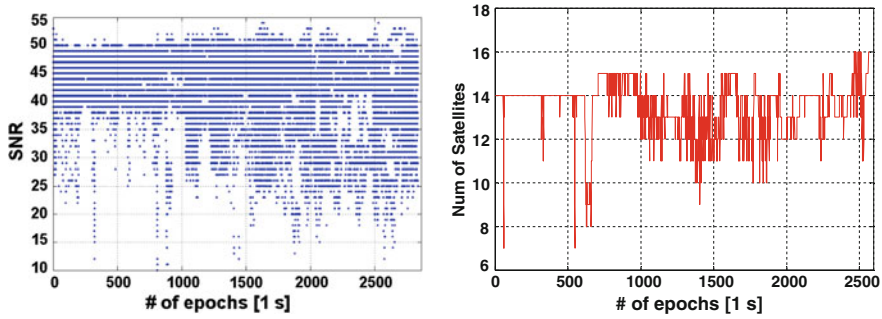


Fig. 8 SNR (left) and the number change of satellites (right) of dynamical car-borne test

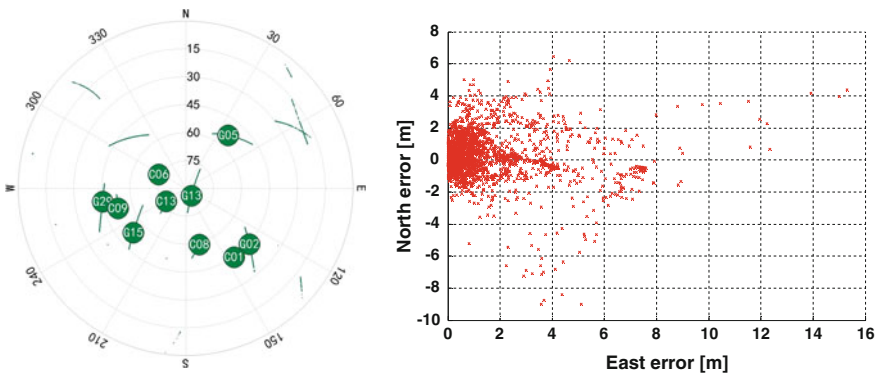


Fig. 9 The sky plot (left) and the horizontal (N, E) position scatter (right) of dynamical car-borne test

performance of dynamic car-borne test on urban road shown in Fig. 9 (right), in some epoch, which is serious occlusion and poor satellite structure, the offset of E direction are more than 10 m. However, it can be found from most epochs, the reliable positioning within 5 m can be guaranteed on dynamic car-borne test.

Table 4 shows the statistics analysis of dynamic car-borne test. As it is shown, RMS of horizontal positioning is 2.63 m, resulted from poor environment. But the rate of points which are better than 5 m is 93.0%, satisfied the needs of many civilian areas.

Table 4 The statistics analysis of dynamic car-borne test

Content	Precision (m)					
	RMS	≤ 0.5 (%)	≤ 1 (%)	≤ 3 (%)	≤ 5 (%)	>5 (%)
N	1.43	37.9	63.6	95.4	99.0	1.0
E	2.21	36.4	62.2	86.5	94.8	5.2
Horizont	2.63	14.7	37.5	81.4	93.0	7.0

4 Conclusions and Future Research

In this paper, RTD positioning algorithm for low-cost receiver is studied carefully. The position solution is constrained by velocity pseudo-observation, and the effects of gross error was eliminated by robust kalman filter, the continuity and reliability of GNSS positioning have been effectively guaranteed, it has a good engineering value. Experiments also verify the positioning performance of the algorithm in different situations, basically meet the needs of location based service. But it can be seen from above experiments, GNSS positioning performance is still poor in extremely harsh environments. So, the author plan to get better positioning performance by using low-cost INS (Inertial Navigation System) devices, which is used to assist GNSS.

Acknowledgements We appreciate anonymous reviewers for their valuable comments and improvements to this manuscript. Thanks also go to the National Natural Science Foundation of China (No: 41574026, 41774027), Primary Research & Development Plan of Jiangsu province (Grant Number BE2016176), National Key Technologies R&D Program (Grant Number 2016YFB0502101) and Six Talent Peaks Project in Jiangsu Province (Grant Number 2015-WLW-002).

References

1. Zuo X, Chen Y, Shi X et al (2017) Centimeter level precise positioning with a low-cost GNSS antenna. In: China satellite navigation conference (CSNC) 2017 proceedings, vol III
2. Mongredien C, Doyen JP, Strom M, Ammann D (2016) Centimeter-level positioning for UAVs and other mass-market applications. In: Proceedings of the ION GNSS 2016, Institute of navigation, Portland, Oregon, September 12–16, pp 1441–1454
3. Zhou P, Liu H, Qian C (2015) Research and implementation of real-time GPS dynamic position with single frequency observations. *Bull Surv* 5:13–16
4. Peng XY, Guo JM, Xue SQ (2011) Accuracy analysis on GPS single frequency phase-smoothed pseudorange and it's point positioning. *Sci Surv Mapp* 36(5):39–41
5. Tang W, Liu Z (2005) Precision evaluating and applications of GPS code pseudorange smoothing algorithm based on carrier observables. *Wtsum Bull Sci Technol* 30(3):37–39
6. Zhong LN, Liu JY, Rong-Bing LI et al (2014) Adaptive robust filtering algorithm for tightly-coupled integrated navigation system based on carrier phase smoothing pseudorange. *J Chin Inert Technol* 22(2):205–210
7. Park B, Sohn K, Kee C (2008) Optimal hatch filter with an adaptive smoothing window width. *J Navig* 61(3):435–454
8. Hwang PA, Mcgraw GA, Bader JR (2001) Enhanced differential GNSS carrier-smoothed code processing using dual frequency measurements. *US* 6198430 B1
9. Yang Y (2006) Adaptive navigation and kinematic positioning. *Surveying and Mapping Press*, Beijing
10. Gui Q, Xu A, Han S (2009) A stepped robust and adaptive filtering and its applications in GPS kinematic navigation. *Geomat Inf Sci Wuhan Univ* 34(6):719–723

11. Sun H (2004) DGPS/INS intergrated position and attitude determination and its application in MMS. Wuhan University
12. Wang T (2012) Preliminary accuracy analysis on velocity determination using compass doppler measurements. *J Geomat* 37(2):16–19
13. Li X, Wu M, Zhang K et al (2012) Navigation satellites velocity and acceleration computation: methods and accuracy analysis. *Acta Geod Cartogr Sin* 41(6):816–824

The Research on Single Satellite Positioning System of Aircraft Based on Communication Signals of Beijing Inmarsat Ground Earth Station



Yuanliang Liu and Hongxia Kang

Abstract The traditional satellite navigation and positioning technology uses a number of satellites to locate, for example, GPS and BDS use four satellites to locate. However, in some special cases, when GPS or BDS positioning device is invalid, it becomes very urgent to realize the positioning. In order to solve this problem, a single satellite positioning method and system based on the communication signals of the International Maritime Satellite (Inmarsat) Ground Earth Station (GES) is proposed in this paper. The standard rectangular space coordinate system is established with the center of the geocentric as the starting point, the flight path model of the aircraft is constructed, and all the possible flight paths of the aircraft are found by using the method of iterative analysis. Method to find out all the possible trajectories of the aircraft and visualize them by using the relevant data such as the Burst Timing Offset (BTO) provided by the GES to simulate, and determine the possible location of aircraft. Finally, the cubic spline interpolation is used to estimate the velocity components of the aircraft, combined with the Burst Frequency Offset (BFO) data formula to estimate the BFO value of each route, and compared with the standard BFO value to calculate the probability of each possible flight route, simulate the possible flight path. Taking the Beijing GES and the Asia-Pacific Satellite as an example, this paper analyzes. The research results of this paper will help to shorten the positioning time of the aircraft crash, find the aircraft as soon as possible.

Keywords Inmarsat · GES · Communication signals · Aircraft BTO · BFO · Single satellite positioning

Y. Liu (✉)

China Transport Telecommunication & Information Center, Beijing, China
e-mail: ddx303@126.com

Y. Liu

Transport Information & Communication Technology R&D Center, Beijing, China

H. Kang

China Transport Infojet Technologies Co., Ltd., Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_36

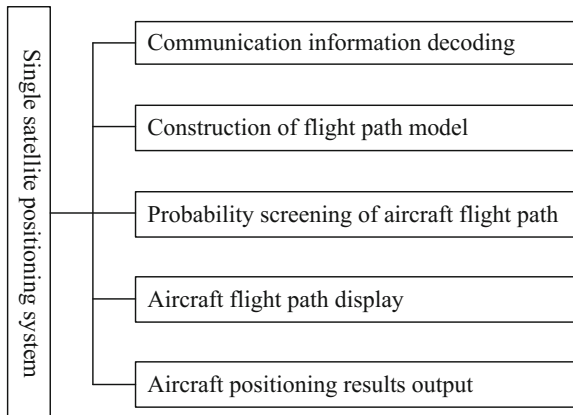
1 Introduction

The traditional satellite navigation and positioning technology uses a number of satellites to locate, for example, GPS and BeiDou Navigation Satellite System (BDS) use four satellites to locate. The Real-time kinematic (RTK) and other technologies can be used to achieve more accurate positioning [1]. However, in some special cases, when GPS or BDS positioning device is invalid or out of work, it becomes very urgent to realize the positioning. For example, in Malaysia Airlines MH370 lost contact event, the aircraft did not provide GPS location information, only returned to seven handshake information of the aircraft and the International Maritime Satellite (Inmarsat). The traditional four-satellites positioning technology can not meet the requirements of the flight safety positioning. To solve this problem, this paper proposes a single satellites positioning method and system based on Inmarsat Ground Earth Station (GES) communication signals, and takes Inmarsat Beijing GES and the Asia-Pacific satellite as an example.

2 System Structure

The single satellite positioning system based on Beijing GES communication signals is mainly composed of six parts: the decoding of communication information of Beijing GES, the construction of flight path model, the probability of aircraft flight path selection, the flight path of aircraft and the output of aircraft positioning results. System function shown in Fig. 1.

Fig. 1 Functional composition of aircraft single satellite positioning system



3 System Workflow

Input various initial conditions, calculate the circle center and radius of the relevant circle, calculate the coordinate value of aircraft on the relevant circle, perform cubic spline difference fitting on the coordinate value of the aircraft, calculate the BFO value of the intermediate path of the aircraft, and screen the intermediate path by using the probability screening method, analyze and calculate the aircraft flight path and display, calculate the location longitude and latitude and positioning deviation of key nodes, then output these data. System workflow shown in Fig. 2.

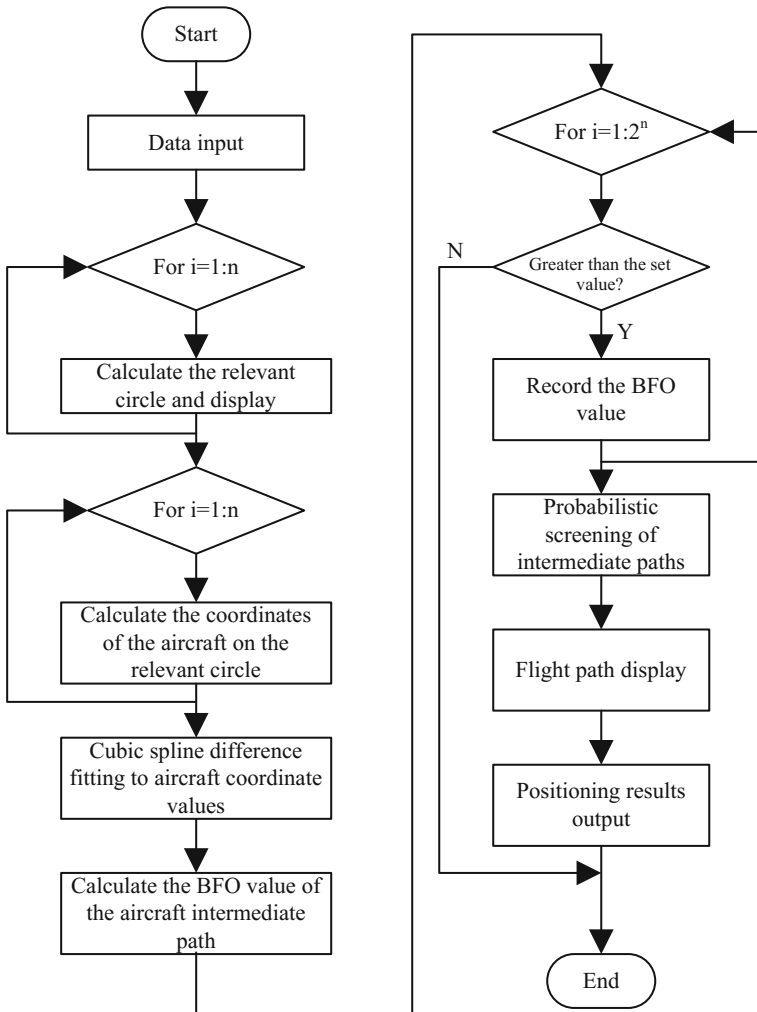


Fig. 2 Workflow of aircraft single satellite positioning system

4 Communication Information Decoding

In order to capture the interactive communication information between the aircraft and Beijing GES, the Wireshark network capture tool which is widely used is selected as a foundation, and the characteristics of the Wireshark protocol such as expansibility, open architecture and convenient and flexible script of Lua scripting language design function, design protocol of communication package for private protocol, realize the decoding analysis of Beijing GES communication signal, improve analysis efficiency and reduce human error of data packet analysis [2].

Through the real-time analysis of Beijing GES communication information, we can get the Burst Timing Offset (BTO), Burst Frequency Offset (BFO), system delay compensation, system frequency compensation, airborne terminal IMSI number and GPS positioning device location and other information.

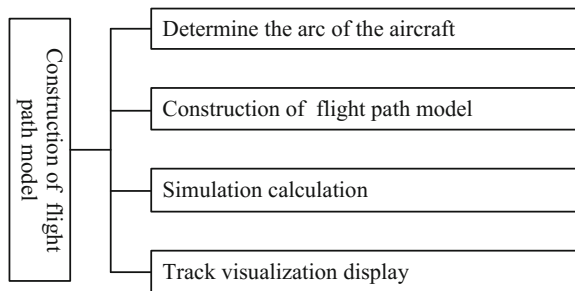
5 Flight Trajectory Model Construction [3–5]

The flight trajectory model construction includes the following steps: using BTO to determine the arc of aircraft; using the iterative method to build the flight trajectory model; simulating calculation; displaying the trajectory. Method steps shown in Fig. 3.

5.1 Use BTO to Determine the Arc of Aircraft

BTO consists of two parts, one is the system transmission delay, that is, the signal from GES to the satellite and then to the aircraft round trip time; the second is the system processing delay (bias), that is, system processing time. Therefore, the signal propagation distance can be expressed as follows:

Fig. 3 Model construction of flight trajectory



$$L_c = \frac{BTO - bias}{2} c \tag{1}$$

where c is the speed of light.

Thus the distance from the satellite to aircraft can be expressed as:

$$L_{S-A} = L_t - L_{S-G} \tag{2}$$

where L_{S-G} is the distance from the satellite to GES.

As shown in Fig. 4, with the geocentric as the origin, a standard space cartesian coordinate system is established. Suppose the location of aircraft is (x, y, z) and the location of satellite is (a, b, c) . From Eq. (2) we can see the following distance relationship:

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = \left(\frac{BTO - bias}{2} c - L_{S-G} \right)^2 \tag{3}$$

In addition, aircraft flies in the air at a height of h from the ground, so the spherical coordinate equation is:

$$x^2 + y^2 + z^2 = (R + h)^2 \tag{4}$$

where R is the average radius of the earth, h is the height of aircraft from the ground, and the simultaneous Eqs. (3) and (4) can get the trajectory arc equation of aircraft.

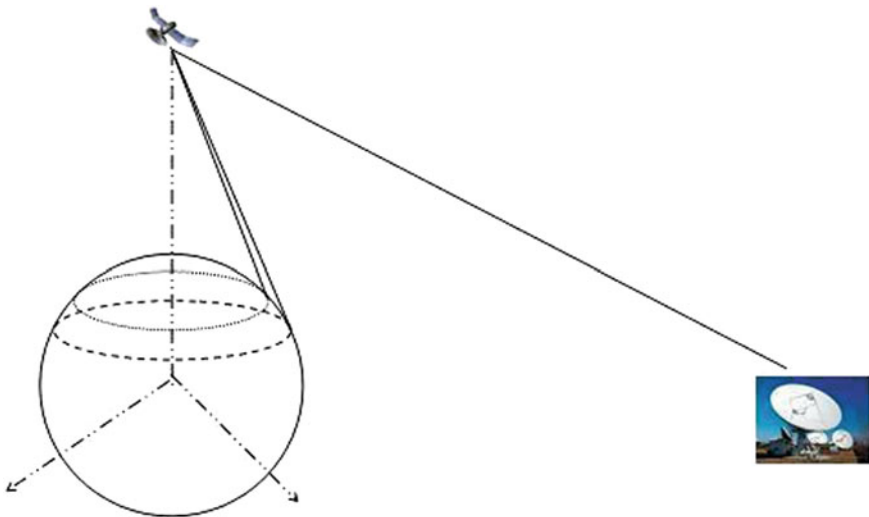


Fig. 4 Schematic diagram of the trajectory arc based on BTO

5.2 Using Iterative Method to Build Aircraft Trajectory Model

Assuming GES receives the ping signal of aircraft for n times and the corresponding time point as t_n , selects a previous time point of receiving the first ping signal as the initial position point of aircraft. Because the signal propagation speed is very fast, propagation time is very short, and the moving speed of the aircraft is slower than that of the signal. Therefore, the moving distance of aircraft during the signal propagation is negligible, and the speed of aircraft is constant.

From the time point t_0 to the time point t_1 , aircraft is on the first arc of the trace, which is obtained from the arc equation:

$$(x - a_1)^2 + (y - b_1)^2 + (z - c_1)^2 = \left(\frac{BTO_1 - bias}{2}c - L_{(S-G)_1} \right)^2 \tag{5}$$

where (x, y, z) is the possible position that aircraft may be located on the arc of the first trace, (a_1, b_1, c_1) is the position of the satellite at time t_1 , $L_{(S-G)_1}$ is the distance of satellite and GES at time t_1 . According to the spherical coordinate equation can be obtained:

$$x^2 + y^2 + z^2 = (R + h)^2 \tag{6}$$

Finally, according to the distance that aircraft moves from time point t_0 to time point t_1 , it can be obtained that:

$$\cos \theta_1 = \frac{x_0x + y_0y + z_0z}{(R + h)^2} \tag{7}$$

$$S_1 = (R + h)\theta_1 \tag{8}$$

$$S_1 = v(t_1 - t_0) \tag{9}$$

where (x_0, y_0, z_0) is the position of aircraft at the time point t_0 , S_1 is the distance that aircraft moves from the time point t_0 to the time point t_1 , and θ_1 is the central angle corresponding to S_1 .

The simultaneous Eqs. (4)–(8) can obtain two coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2) that aircraft may be located on the first trajectory arc, then (x_1, y_1, z_1) and (x_2, y_2, z_2) , respectively, as the initial point, it is also possible to obtain four points (x_3, y_3, z_3) , (x_4, y_4, z_4) , (x_5, y_5, z_5) , (x_6, y_6, z_6) . And so on, 2^n points on the n th trajectory arc can be obtained.

That is, from the time point t_i to t_{i+1} , taking a certain point of the i -th trajectory arc as an initial point, it may be denoted as (x_a, y_a, z_a) , to the $i + 1$ trajectory arc, there is the following equation:

$$(x - a_{i+1})^2 + (y - b_{i+1})^2 + (z - c_{i+1})^2 = \left(\frac{BTO_{i+1} - bias}{2} c - L_{(S-G)_{i+1}} \right)^2 \tag{10}$$

$$x^2 + y^2 + z^2 = (R + h)^2 \tag{11}$$

$$\cos \theta_{i+1} = \frac{x_a x + y_a y + z_a z}{(R + h)^2} \tag{12}$$

$$S_{i+1} = (R + h)\theta_{i+1} \tag{13}$$

$$S_{i+1} = v(t_{i+1} - t_i) \tag{14}$$

where (x, y, z) is the point at which the aircraft may be located on the (i + 1) trajectory arc, (a_{i+1}, b_{i+1}, c_{i+1}) is the position of the satellite at time point t_{i+1}, L_{(S-G) i+1} is the distance between the satellite and GES at time point t_{i+1}, S_{i+1} is the distance that the aircraft moves from the time point t_i to the time point t_{i+1}, and θ_{i+1} is the corresponding central angle of S_{i+1}.

5.3 Simulation Calculation

Extract 1 group of data packets that the aircraft communicates with Beijing GES and extract 12 location information points from the data packet for analysis. Design related iterative algorithm flow according to Eqs. (10)–(14), develop software by using Matlab.

5.4 Track Visualization

Using Matlab GUI and Google Earth to visualize the above process, the trajectory visualization diagram is obtained. The trajectory visualization display method is as follows:

(1) Method of Realizing the Trajectory Arc

By setting coordinates of Beijing GES, coordinates of initial time point of aircraft, BTO, space coordinates of satellite, times of information exchange between aircraft and satellite and time difference as receiving parameters, calculate the expression of the trajectory arc by using Eqs. (1)–(4), connect Matlab with Google Earth and finally draw the arc of the trajectory.

(2) Method of Visualization of Aircraft Trajectory

First, set the coordinates of Beijing GES, the coordinates of the initial time point of aircraft, BTO, bias, satellite coordinates, the initial speed of aircraft and the number of exchange of information between the aircraft and the satellite, and the time difference as the receiving parameters. Calculate the coordinates of each possible point using Eqs. (5)–(14). The process from the initial point to the first arc needs to be calculated separately, using Eqs. (5)–(9), and then iteratively using the iterative method, using Eqs. (10)–(14). Through the Matlab xlsxwrite function, the output of the aircraft coordinates stored in the Excel file. Then through the Matlab importdata function, input the previously calculated aircraft coordinates, the establishment of the path lookup table, so that you can get the coordinates of the corresponding point directly according to the route. Finally, the Cartesian coordinates are converted to the geodetic coordinates, and the trajectory map is drawn on the Google Earth.

(3) Calculation Results

Through simulation calculation, the possible trajectory arc of aircraft is obtained, and the Google Earth software is used to display the predicted arc trajectory, as shown in Fig. 5.

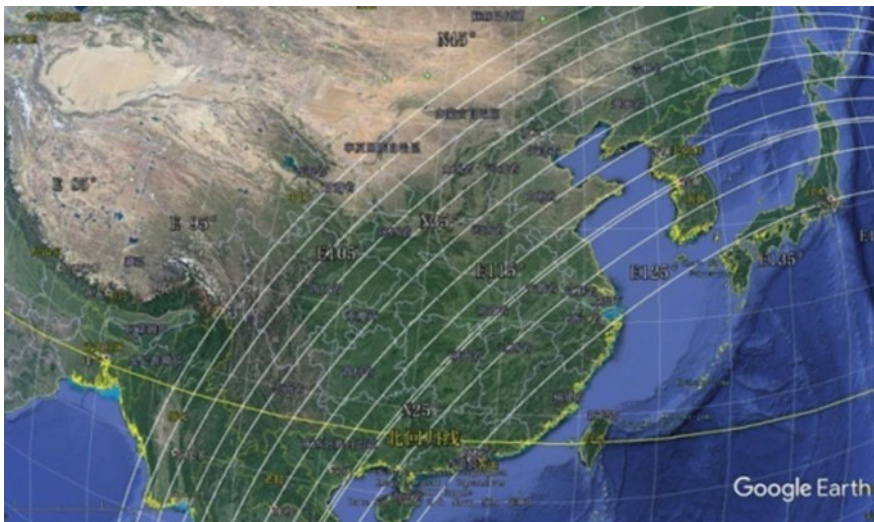


Fig. 5 Aircraft prediction arc trajectory

6 Probability Screening of Flight Path [6]

6.1 Probability Screening Model

First of all, we use the x as the parameter, find the cubic spline interpolation of z with respect to x , and the cubic spline interpolation of y with respect to x . Then find the derivative of z about x and the derivative of y about x . Furthermore, the number of tangential directions can be obtained. Then normalized, and multiplied by the velocity scalar size, thereby calculating the velocity component estimates in the x , y , z directions.

Because BFO calculation equation is:

$$BFO = \Delta F_{up} + \Delta F_{down} + \Delta F_{comp} + \Delta F_{bias} + \Delta F_{AS} \tag{15}$$

Among them, ΔF_{bias} and ΔF_{AS} are fixed values, obtained from the GES Decoding.

So first calculate ΔF_{up} :

$$\Delta F_{up} = \frac{F_L \times \Delta V \times (Satellite - Aircraft)}{c \times \sqrt{Satellite^2 - Aircraft^2}} \tag{16}$$

where c is the speed of light, ΔV is the aircraft speed component, $Satellite$ is the satellite position, and $Aircraft$ is the aircraft position.

$$\Delta F_{down} = \frac{F \times \Delta V \times (Aircraft_0 - Satellite)}{c \times \sqrt{Satellite^2 - Aircraft_0^2}} \tag{17}$$

where c is the speed of light, ΔV is the satellite speed component, $Satellite$ is the satellite position, and $Aircraft_0$ is the initial position of the aircraft.

$$\Delta F_{comp} = \frac{F_L \times (Aircraft_0 - Aircraft)}{c \times \sqrt{Aircraft_0^2 - Aircraft^2}} \tag{18}$$

where c is the speed of light, ΔV is the satellite speed component, $Aircraft_0$ is the initial position of the aircraft, and $Aircraft$ is the aircraft position.

By calculating the difference between the BFO predicted value and the standard value of each node in each route, absolute values are taken and summed to find the maximum value. You can calculate the probability of each route.

$$P_i = 1 - \frac{d_i}{\max(d_i)} \quad (i = 1, 2, 3, \dots, 2^n) \tag{19}$$

Wherein, d_i is the sum of the absolute value of the difference between the BFO predicted value and the standard value for each node of each route, $\max(d_i)$ is the maximum value of the route of 2^n , n is the number of selected nodes.

6.2 Calculation Results and Diagrams

After selecting 12 key nodes in the flight path, the BFO estimates of each route are obtained through simulation and compared with the standard BFO values. The probability of each intermediate route may be calculated according to formula (19), displayed with different colors, the color from dark to light represents the probability from low to high, the calculation results are displayed in Google Earth, as shown in Fig. 6.

7 Flight Trajectory Display

According to the probability calculation results of each path, the path with the highest probability is selected to form the prediction of the final flight positioning trajectory of the aircraft. The white lines and key points are the location and key points of the GPS positioning device of the airborne terminal. The red lines and key points are the paths and key points of the system positioning in this paper, as shown in Fig. 7.

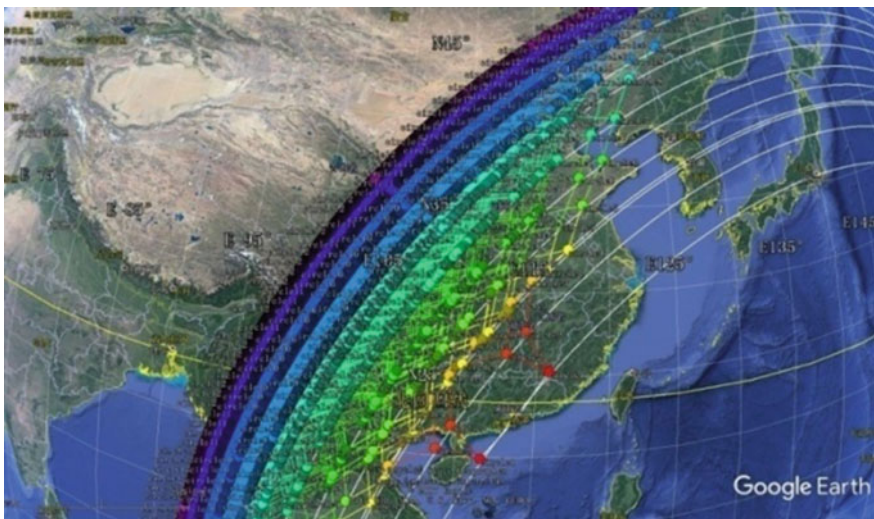


Fig. 6 Probability calculation results of intermediate path

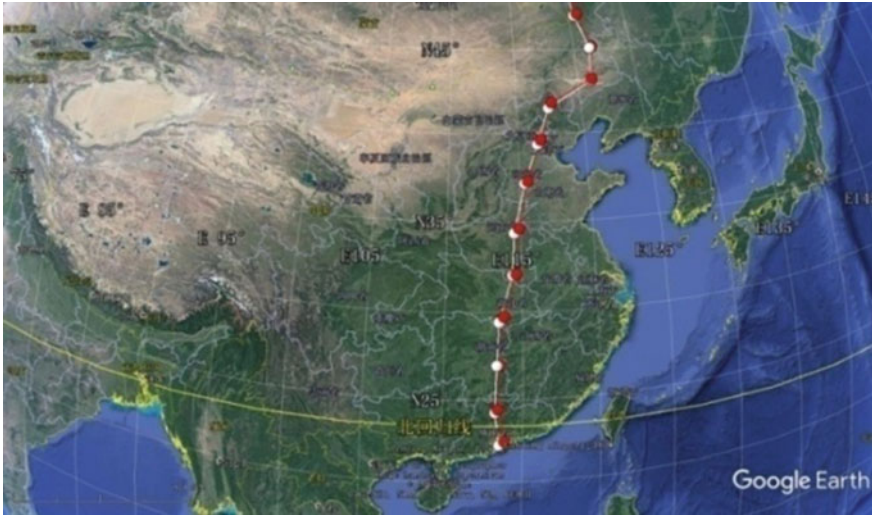


Fig. 7 Prediction path and GPS location path

8 Aircraft Positioning Results Output

12 key positions are selected on the predicted track to predict the latitude and longitude of the position, and then compared with the longitude and latitude of the positioning position of the airborne terminal GPS positioning device to obtain the positioning error of the single satellite positioning result, as shown in Table 1.

As can be seen from Table 1, the single satellite positioning system based on Beijing GES communication signals can be used to locate aircraft in an emergency situation. the positioning accuracy is about 100 km, and meet the development and usage requirements, provides a new kind of technical means for aviation safety.

Table 1 Single satellite positioning results

Serial number	Positioning results		GPS positioning results		Positioning deviation (km)
	North latitude (°)	East longitude (°)	North latitude (°)	East longitude (°)	
1	22.55563889	113.8584722	22.33203	113.51305	129.9878519
2	22.55444444	113.9571667	22.3276	113.57258	140.6583227
3	22.71169444	113.8750833	22.42421	113.51903	145.5666113
4	22.89875	113.7606389	22.53555	113.45383	153.2997717
5	23.08322222	113.63325	23.04596	113.37597	80.05621315
6	23.29377778	113.5168056	23.16976	113.31005	75.68184991
7	23.50438889	113.4048889	23.29758	113.24176	84.99236271
8	23.73341667	113.3993889	23.44003	113.23578	109.6007987
9	23.97697222	113.4060278	23.58371	113.24217	139.9434195
10	24.21827778	113.4243333	24.13058	113.24876	60.94084169
11	24.47133333	113.4243889	24.28168	113.25278	81.85106123
12	24.72266667	113.4363889	24.43216	113.2571	110.8481242

Acknowledgements This study is supported by the National Science and Technology Support Program Project of China (No. 2014BAB12B01).

References

1. Liu Y (2017) The study on ship auxiliary navigation system of intelligent lockage based on BDS. The 8th China satellite navigation conference, Shanghai, p 5
2. Tang H (2016) Wireshark based plugin implementation on railway signal system application data's dissection. *Traffic Transp* 7:84–88
3. Australian Transport Safety Bureau. MH370 flight path analysis update. ATSB transport safety report, External Aviation Investigation AE-2014-054, Oct 2014
4. Australian Transport Safety Bureau. MH370 search area definition update. ATSB transport safety report, External Aviation Investigation AE-2014-054, Dec 2015
5. Australian Transport Safety Bureau. MH370 Search and debris examination update. ATSB transport safety report, External Aviation Investigation AE-2014-054, Nov 2016
6. Davey S, Gordon N, Holland I, Rutten M, Williams J (2016) Bayesian methods in the search for MH370. Springer Open, Aug 2016

Assessment of the Contribution of QZSS Combined GPS/BeiDou Positioning in Asia-Pacific Areas



Yize Zhang, Nobuaki Kubo, Junping Chen, Hu Wang
and Jiexian Wang

Abstract Three QZSS satellites are launched in 2017, which implies that a four satellites regional system is to be established in 2018. There is no doubt that QZSS will play a more important role in the future global GNSS constellations. So it is quite necessary to investigate the importance of current QZSS constellation in positioning. In this paper, the number of visible satellite and PDOP (Position Dilution of Precision) value improvement by combining QZSS with the existing GPS and BeiDou system is analyzed among Asia-Pacific areas. 9 IGS stations are selected to evaluate the performance of SPP (Single Point Positioning) and PPP (Precise Point Positioning) using GPS, BeiDou and GPS + QZSS, BeiDou + QZSS system. Analysis results show that QZSS improves SPP performance for both GPS and BeiDou at different level. Especially when the satellite number is reduced, such as in urban areas or when the elevation cutoff is high, the positioning error will reduce after adding QZSS satellite and the availability of other GNSS systems will also improve. For kinematic PPP users, QZSS could also reduce the convergence period. Meanwhile, the dual frequency and single frequency RTK (Real Time Kinematic) positioning performance is compared after adding QZSS satellite into GPS and BeiDou. Kinematic car test in urban environment shows that when combining QZSS satellite with GPS and BeiDou, the rate of instantaneous ambiguity resolution will increase for both single and dual-frequency users.

Y. Zhang (✉) · N. Kubo
Tokyo University of Marine Science and Technology, Tokyo, Japan
e-mail: zhyize@163.com

Y. Zhang · J. Chen
Shanghai Key Laboratory of Space Navigation and Positioning Techniques, Shanghai, China

Y. Zhang · J. Chen
Shanghai Astronomical Observatory, Chinese Academy of Science, Shanghai, China

H. Wang
Chinese Academy of Surveying and Mapping, Beijing, China

J. Wang
College of Surveying and Geo-Informatics, Tongji University, Shanghai, China

Keywords PDOP · SPP · PPP · RTK · Ambiguity resolution

1 Introduction

QZSS satellite system, known as Quazi-Zenith Satellite System, is a regional satellite navigation system developed by Japan. This first QZSS satellite, also named as Michibiki, was launched by JAXA (Japan Aerospace Exploration Agency) in September, 2010. From February, 2017 [1]. JAXA discontinues the operation of QZS-1 and the control of QZSS is transferred to the Japan Cabinet Office [2]. After that, QZSS satellite speed up its launch plan. The 2nd to 4th QZSS satellites were launched from June to October in 2017 [3, 4]. So far, QZSS has preliminary finished its first step of QZSS system including 4 satellites. According to Japan Cabinet Office, QZSS system will start official service from 2018 and a system consists of seven satellites is planned as the second step of QZSS, which will be finished by 2024 [5].

Starts from the launch of first QZSS satellite, many researches have been done from QZSS signal transmission to POD (Precise Orbit Determination) and positioning. The satellite attitude and orbit determination was analysis by Hauschild and Montenbruck [6–8]. QZSS also releases the official satellite information and control mode in 2017, which would be great help for QZSS orbit determination [9]. The SISRE(Signal in Space Range Error) was analyzed by Montenbruck [10], which indicates that it can reach below 0.6 m and is comparable with GPS. Quan et al. assessed the signal quality of QZSS together with other GNSS systems using real-data [11] and Odolinski et al. analyzed the performance of long baseline and single frequency QZSS aided RTK [12, 13]. The QZSS ISB is also estimated with other GNSS systems [14].

However, almost all these researches only focus on QZS-01. With the available of other satellites of QZSS, it is necessary to investigate the contribution of QZSS constellation under multi-system.

Due to the constellation character of QZSS, the service area of QZSS is limited within Asia-Pacific region. In this paper, we mainly focus on its contribution in Asia-Pacific area.

2 QZSS System

Till the end of November, 2017, the satellites number of QZSS in orbit is 4. Among which QZS-3(PRN J07) is the GEO satellite above the longitude of 127°E, the other three satellites are QZO (Quazi-Zenith Satellite Orbit) satellites [5]. Different from the IGSO satellite of BeiDou, trajectory of QZO satellite is an asymmetric eight to keep the satellite stay longer in Japan.

Table 1 Information of QZSS constellation

Sat_Name	QZS-1	QZS-2	QZS-3	QZS-4
SCN	J001	J002	J003	J004
PRN	J01	J02	J07	J03
BLOCK	IQ	IIQ	IIG	IIQ
Sat_Type	GSO	GSO	GEO	GEO
Longitude	130–140°E	130–140°E	126.9–127.1°E	130–140°E
Launch date	2010.09.11	2017.06.01	2017.08.19	2017.10.09

Table 1 summarizes the basic information of current QZSS constellation.

According to QZSS ICD and the official information [9, 15], the coordinate system that QZSS adopts is JGS (Japan satellite navigation Geodetic System), which is similar with ITRS. The newest JGS2010 is established by more than 40 stations of SLR, GPS and QZSS [16]. JGS uses the Geodetic Reference System 1980 (GRS80) ellipsoid, while WGS84 uses the WGS84 ellipsoid. The difference between these two ellipsoids is only the flattening parameter, which can be neglected in most situations. QZSS is the time system applied in QZSS. Its definition is the same as GPST. QZSS is maintained by JAXA, its difference with UTC(NICT) maintained by NICT (National Institute of Information and Communications Technology) is less than 50 ns [15].

According to the pamphlet provided by QZSS [17], the SIS (Signal In Space) of QZSS broadcast ephemeris is under 2.6 m (95%); the URE (User Range Error) of ionosphere is below 7 m (95%); the time difference with UTC is 40 ns (95%) or less; the satellite service availability for QZO is 0.95 or more and 0.80 or more for GEO; the constellation service availability is 0.99 or more. Besides basic PNT (Positioning, Navigation, Timing) service, QZSS also provides code-based augmentation service of SLAS (Sub-meter Level Augmentation System) and DC Report (Satellite Report for Disaster and Crisis Management) on L1S band, together with PPP-RTK like precise service of CLAS (Centimeter Level Augmentation System) on L6 band.

Currently, the IGS ACs (Analysis Center) that provide QZSS precise orbit and clock include GFZ, CODE, WHU, TUM, et al. From the end of September of 2017, QZS-2 is included in POD processing [18]. Montenbruck assessed the SIS of broadcast ephemeris by comparing with POD result and concluded that the SIS of QZS-1 is less than 0.6 m. Richard assess the clock stability of QZS-1 and QZS-2 in 2017 and results show that the 100 s Allan Deviation is better than 3×10^{-13} and the long stability of QZS-1 and QZS-2 reaches the same level of GPS Block IIF satellite.

With the available of more QZSS satellites, users may enjoy a better positioning performance.

3 Data Description

The number of QZSS satellite is still too few to provide a stand-alone positioning service. To assess the contribution of QZSS, we combine QZSS with GPS and BeiDou in Asia-Pacific areas. The comparison of positioning results is used for the assessment.

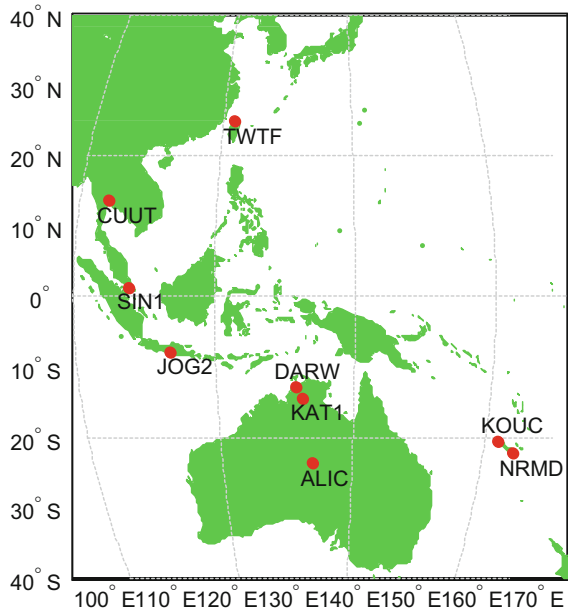
Although J02 (QZS-2) started signal transmission from the end June in 2017, many IGS station haven't upgrade their receivers firmware and can't track the signal of J02 or J03. After carefully distinguish, 9 of IGS stations distributed in Asia-Pacific areas are selected, which shows in Fig. 1. All these stations can track signals of J01, J02 and GPS. BeiDou signals can also be tracked except stations of TWTF and JOG2. 10 days of data from Nov. 6th to 15th in 2017 are chose for the assessment in this paper.

4 Contribution of QZSS

4.1 Contribution of QZSS in DOP Value

Before the assessment of positioning performance, it is necessary to evaluate the contribution of QZSS in satellite number and DOP value.

Fig. 1 Tested IGS stations



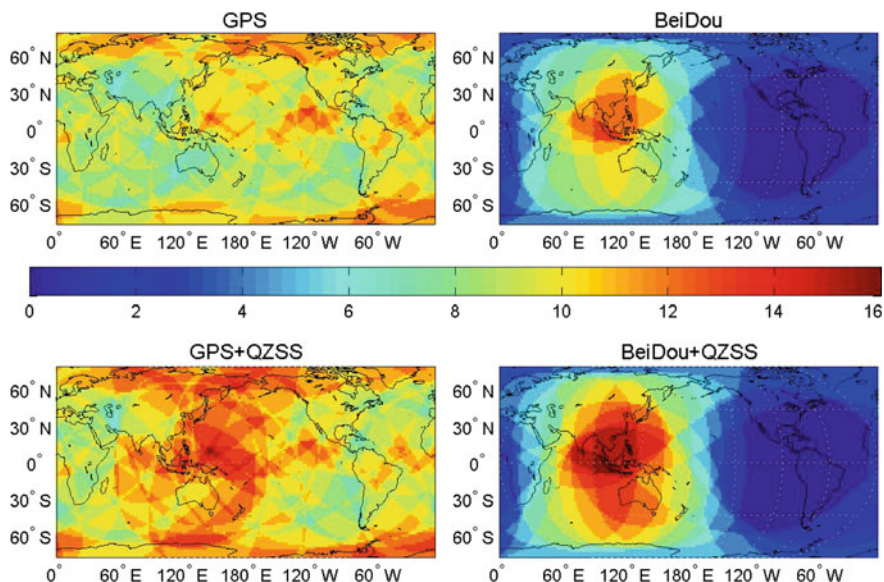


Fig. 2 Satellite number comparison for difference GNSS system (2017/11/15 23:00:00)

Figure 2 explains the visible satellite number change after combining QZSS with GPS and BeiDou at the epoch of November 15th 23:00(UTC) in 2017. The satellite elevation cutoff is 10° and the left two subplots compare the satellite number for GPS and GPS + QZSS, while the right two subplots compare that for BeiDou and BeiDou + QZSS. The corresponding Fig. 3 illustrates the improvement of PDOP for GPS and BeiDou after combining with QZSS.

As can be seen from Figs. 2 and 3, QZSS satellites shows a better improvement in Southeast Asia and West Pacific, where the mean improved satellite number can reach up to 3. With the help of QZSS, this area is the hottest GNSS area in the world and shows a best PDOP value. At the same time, for the current regional satellite system of BeiDou, QZSS would help to enlarge its service area.

4.2 Contribution of QZSS in SPP

From analysis in the section above, we can find the improvement of QZSS in Asia Pacific areas in satellite number and DOP value, which means that it would be great help in user positioning, theoretically. In this part, we will analysis the SPP performance after adding QZSS to GPS and BeiDou.

Based on the data introduced in 1.3, we perform GPS and BeiDou based SPP. After that one and two QZSS satellites are added into the solution

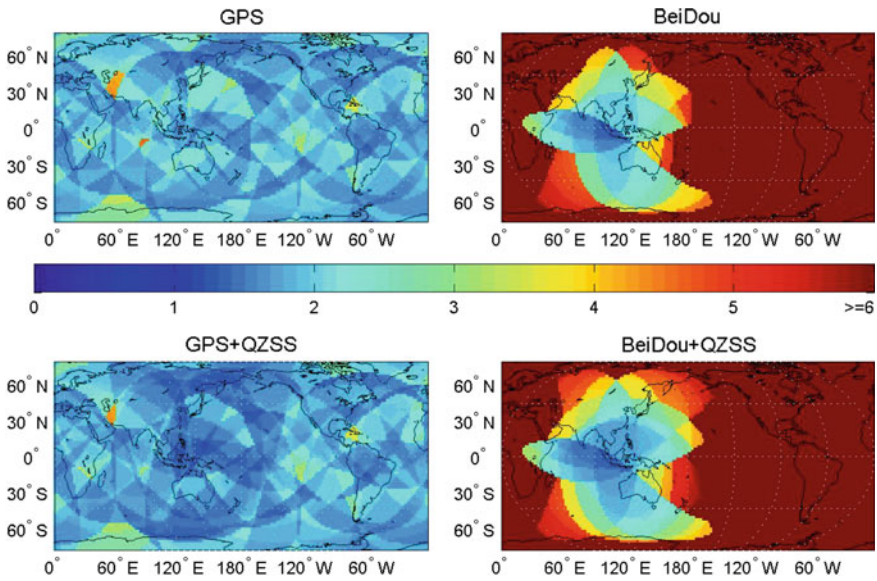


Fig. 3 PDOP comparison for difference GNSS system (2017/11/15 23:00:00)

successively. Due to the high elevation of QZSS, we set the elevation cutoff every 5° from 10° to 25°.

Take BeiDou as example, Fig. 4 compares the SPP error in NEU(North, East, Up) direction of BeiDou alone and BeiDou + QZSS combined SPP at station KAT1 on November, 6th, 2017. The elevation cutoff in the left subplots is 10° and the right one is 25°. Tables 2 and 3 summarize the GPS and BeiDou SPP performance after combining with 0 to 2 QZSS satellites.

From Fig. 4 and Tables 2, 3, one can clearly see that the positioning performance will improve for both GPS and BeiDou at the elevation cutoff from 10° to 25°. The more QZSS satellites, the better positioning RMS. And we can find that with the increase of elevation cutoff, the positioning accuracy decreases in general. However, when combined with QZSS, this trend of decreasing seems to be not so dramatically, which shows the advantage of high elevation for QZSS.

Besides, compared with GPS, this elevation depended decreasing of precision is more slight for BeiDou. This may due to the reason that most BeiDou GEO and IGSO satellites are at a high elevation in this areas.

From Fig. 2 we know that the service area of BeiDou and QZSS is limited due to the limit of satellite number, which influence the service availability. We assume that the service is unavailable when PDOP value is less than 6 [15]. Table 4 is a statistical result of time percentage when PDOP is less than 6 for different system at different elevation cutoff. From the table we can see that the availability of GPS is better than BeiDou even in Asia-Pacific areas. However, when adds QZSS, the

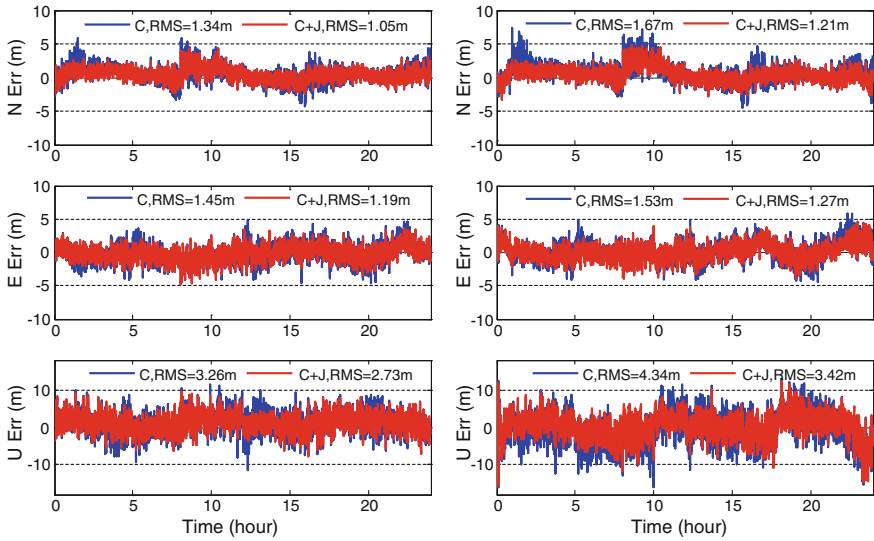


Fig. 4 SPP Performance for BDS after Adding QZSS at An Elevation Cutoff of 10° (left) and 25° (right) at Station KAT1

Table 2 GPS and GPS + QZSS SPP Result

Elevation (°)		10	10	15	20	25
G	N(m)	0.94	1.03	1.35	1.64	
	E(m)	0.00	1.05	1.18	1.44	
	U(m)	2.49	2.90	4.20	5.56	
G + J01	N(m)	0.92	0.99	1.24	1.57	
	E(m)	0.98	1.02	1.14	1.40	
	U(m)	2.42	2.75	3.77	5.16	
G + J01 + J02	N(m)	0.90	0.96	1.14	1.51	
	E(m)	0.96	1.01	1.10	1.33	
	U(m)	2.36	2.65	3.45	4.70	

Table 3 BeiDou and BeiDou + QZSS SPP Result

Elevation (°)		10	10	15	20	25
C	N(m)	1.43	1.50	1.55	1.89	
	E(m)	2.38	2.50	2.54	2.60	
	U(m)	4.44	4.77	4.98	6.05	
C + J01	N(m)	1.27	1.32	1.40	1.51	
	E(m)	2.06	2.16	2.30	2.44	
	U(m)	3.92	4.14	4.45	5.04	
C + J01 + J02	N(m)	1.18	1.21	1.27	1.36	
	E(m)	1.94	2.03	2.19	2.37	
	U(m)	3.56	3.77	4.07	2.55	

Table 4 Percentage of time when PDOP < 6 at different elevation cutoff

System	Elevation (°)	0 QZSS (%)	1 QZSS (%)	2 QZSS (%)
GPS+	10	99.8	99.8	99.8
	15	99.3	99.6	99.7
	20	95.3	96.4	97.9
	25	85.2	88.9	92.8
BeiDou+	10	75.4	87.2	95.8
	15	71.8	84.4	92.7
	20	68.8	78.9	85.3
	25	65.0	73.6	77.7

availability will greatly improve. For GPS, the availability would also improve at the situation of high elevation cutoff. This proves that QZSS would benefit the continuity and availability of other GNSS systems.

4.3 Contribution of QZSS in PPP

From September of 2017, some IGS ACs starts to provide precise orbit and of QZS-2. In this part we try to assess the contribution of current two QZSS satellite in PPP.

Again, we choose GPS and BeiDou as the base system, respectively. A kinematic PPP solution is carried here. To make full of the satellite observations, the satellite elevation cutoff is set as 10°. The precise multi-GNSS satellite orbit and clock are provided by GFZ. The estimated kinematic PPP positions are compared with true coordinates provided by IGS.

Figure 5 shows the 24 h kinematic PPP error in North, East and Up direction at station KAT1 on Nov. 6th, 2017. The RMS is calculated after convergence of half an hour. From the result we can see that the RMS of GPS will improve from (1, 2, 6 cm) to (1, 2, 5 cm) after adding QZSS satellite. As for BeiDou, the value is from (2, 2, 8 cm) to (1, 2, 5 cm). The better improvement of BeiDou is due to the better performance of convergence period after adding QZSS satellite. In general, more QZSS seems not much improvement on GPS, but really do a good help for BeiDou.

To have a more reliable conclusion, we compute the convergence performance for all stations every day during the first one hour of PPP. For GPS, the result is defined as convergence when 3D positioning error is less than 0.2 m; for BeiDou, the value is 0.4 m. Percentage the convergence stations are calculated every 5 min. Figure 6 shows the statistical result, from which we can see that for BeiDou, the convergence speed improves after combining QZSS satellite. However, the current PPP performance of BeiDou is still not better than GPS, one reason is that the orbit and clock accuracy of BeiDou not as good as GPS, other may due to that the change of DOP value is much slower than GPS.

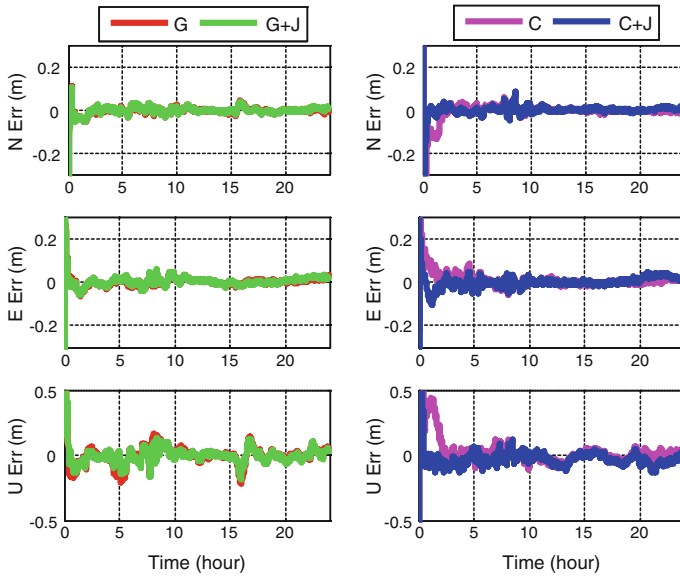


Fig. 5 PPP performance for GPS and BDS combined with QZSS at station of KATI

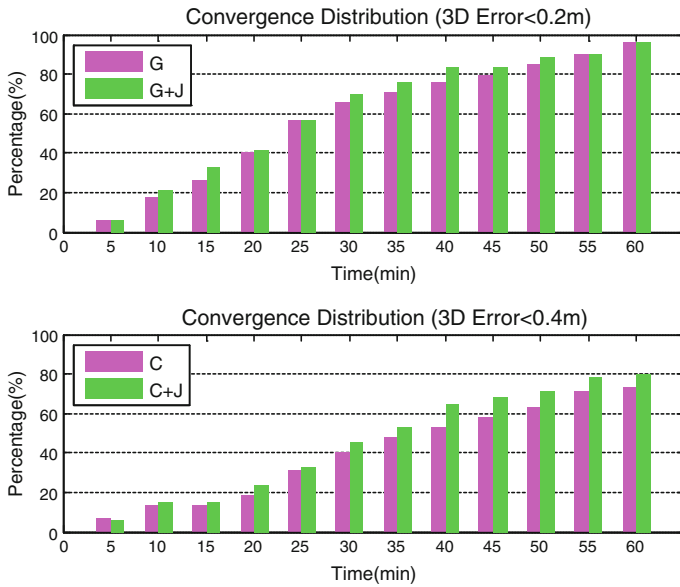


Fig. 6 Percentage of positioning error during convergence period

4.4 Contribution of QZSS in RTK

Thanks to the same frequency of QZSS with GPS, an additional QZSS satellite is almost equivalent with a high elevation GPS satellite for Asia-Pacific users during the process of RTK. However, in urban environments, blocking by the high buildings and serious multipath effect, it is not easy to fix the ambiguity even for short baselines. In this part, we attempt to investigate the contribution of QZSS in RTK.

The experiment is carried out using a Trimble Net_R9 geodetic receiver in the rooftop of a car. One hour of kinematic multi-GNSS system data was collected on July 19th, 2017. The trajectory of the kinematic RTK test is showed in Fig. 7.

An instantaneous AR(Ambiguity Resolution) is applied in processing for both single and dual frequency data. For dual frequency data, the fixing rate of QZSS combined GPS and GPS alone AR is compared. For single frequency data, the fixing rate of QZSS combined GPS + BeiDou and GPS + BeiDou AR is compared. Corresponding results are showed in Fig. 8, from which we can see that the fixing rate will improve from 62.9 to 77.0% for QZSS aided GPS dual frequency AR and from 60.0 to 68.8% for QZSS aided GPS + BeiDou single frequency AR. Compare with the map in Fig. 7, we can clearly see that when the receiver is between high buildings or under high-way, the AR fixed rate would dramatically



Fig. 7 Trajectory of Kinematic RTK Test

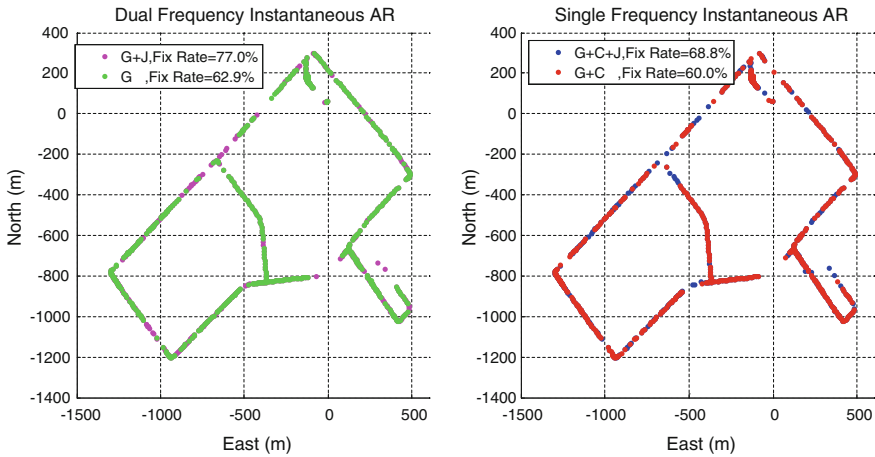


Fig. 8 Dual and single frequency instantaneous AR

decrease, this is due to the frequently satellite signal break off. The result indicates that the AR fixing rate of RTK would improve when high elevation QZSS satellites are introduced, especially in urban environments.

5 Conclusions

In this paper, we introduce the status and development of QZSS. The contribution of QZSS in Asia-Pacific areas is analyzed in DOP, SPP, PPP and RTK. Results show that:

- (1) QZSS improves the visible satellite number and DOP value in Asia-Pacific areas. At the meanwhile, it can help to enlarge the service region of BeiDou.
- (2) When QZSS is combined with GPS or BeiDou, the SPP accuracy would improve at different level, especially for the situation of high elevation cutoff. QZSS could also speed up the convergence period in PPP.
- (3) The instantaneous AR fixing rate can improve with the help of QZSS in RTK.

With the development of QZSS and the ongoing official service of QZSS, it would be more and more important as part of GNSS.

Acknowledgements This work is supported by the National Natural Science Foundation of China (No. 11673050) and the Opening Project of Shanghai Key Laboratory of Space Navigation and Positioning Techniques (No. KFKT_201706).

References

1. <http://www.insidegnss.com/node/5435>. Accessed 30 Nov 2017
2. <http://global.jaxa.jp/projects/sat/qzss/>. Accessed 30 Nov 2017
3. <http://gpsworld.com/qzs-2-signal-analysis-qzs-3-launched/>. Accessed 30 Nov 2017
4. http://mgex.igs.org/IGS_MGEX_Status_QZSS.php. Accessed 30 Nov 2017
5. http://qzss.go.jp/en/overview/services/sv01_what.html. Accessed 30 Nov 2017
6. Hauschild A, Steigenberger P, Rodriguez-Solano C (2012) QZS-1 Yaw attitude estimation based on measurements from the CONGO network. *Navigation* 59(3):237–248
7. Hauschild A, Steigenberger P, Rodriguez-Solano C (2012) Signal, orbit and attitude analysis of Japan's first QZSS satellite Michibiki. *GPS Solutions* 16(1):127–133
8. Montenbruck O, Steigenberger P, Prange L et al (2017) The multi-GNSS experiment (MGEX) of the international GNSS service (IGS)—achievements, prospects and challenges. *Adv Space Res*
9. <http://qzss.go.jp/en/>. Accessed 30 Nov 2017
10. Montenbruck O, Steigenberger P, Hauschild A (2015) Broadcast versus precise ephemerides: a multi-GNSS perspective. *GPS Solutions* 19(2):321–333
11. Quan Y, Lawrence L, Roberts GW et al (2016) Measurement signal quality assessment on all available and new signals of multi-GNSS (GPS, GLONASS, Galileo, BDS, and QZSS) with real data. *J Navig* 69(2):313–334
12. Odolinski R, Teunissen PJG, Odijk D (2015) Combined BDS, Galileo, QZSS and GPS single-frequency RTK. *GPS Solutions* 19(1):151–163
13. Odolinski R, Teunissen PJG, Odijk D (2014) Combined GPS + BDS + Galileo + QZSS for long baseline RTK positioning. In: ION GNSS + 2014, Tampa, Florida, 8–12 Sept 2014
14. Odijk D, Nadarajah N, Zaminpardaz S et al (2017) GPS, Galileo, QZSS and IRNSS differential ISBs: estimation and application. *GPS Solutions* 21(2):439–450
15. Japan Aerospace Exploration Agency (2016) Interface specification for QZSS (IS-QZSS)
16. <http://www.unoosa.org/pdf/icg/2012/template/QZSSupdated.pdf>. Accessed 30 Nov 2017
17. <http://qzss.go.jp/en/overview/downloads/pamphlet.html>. Accessed 30 Nov 2017
18. <ftp://cddis.gsfc.nasa.gov/pub/gps/products/mgex/>

The Service Improvement of BDS Positioning Based on Advanced Equivalent Satellite Clock Calculation



Yangfei Hou, Junping Chen, Bin Wang and Jiexian Wang

Abstract Positioning precision of navigation satellite system can be measured by two indicators: the dilution of precision (DOP) and the user equivalent range error (UERE). As the DOP values are only related to the spatial distribution of navigation satellites, the reduction of UERE is the main approach to improve the positioning precision. Equivalent satellite clock (ESC) has been used by Beidou satellite system (BDS) to reduce the UERE and to improve the user's positioning accuracy. In this contribution, both the pseudo-range and carrier-phase measurements of BDS are used to compute the ESC respectively, and the corresponding navigation positioning performance are also compared. It is shown that the UERE improvement based on phase observables is 50.1%, while 32.1% based on pseudo-range observables. Kinematic positioning experiments of 4 MEGX stations are performed respectively under the standard PNT service and wide area differential service (WADS). It is shown that horizontal, vertical and three-dimensional positioning results of WADS are better than that of the standard PNT service.

Keywords Equivalent satellite clock · User equivalent range error
Navigation performance improvement

Y. Hou · J. Chen (✉) · B. Wang
Shanghai Astronomical Observatory Chinese Academy of Sciences,
Shanghai 200030, China
e-mail: junping.chen@shao.ac.cn

Y. Hou · J. Wang
College of Surveying and Geo-Informatics, Tongji University,
No. 1239, Siping Rd., Shanghai 200092, China

J. Chen
School of Astronomy and Space Science, University of Chinese
Academy of Sciences, Beijing 100049, China

1 Introduction

Users' positioning precision is mainly related to the precision of navigation signals and the spatial distribution of navigation satellites, which can be described as $Pos_{error} = UERE \times DOP$. The DOP is the indicator of navigation satellites distribution, and the UERE is related to the users' range errors and users' equipment errors [1–3]. Beidou satellite system (BDS) is designed to provide two kinds of navigation service, including the standard navigation service (to all users) and the wide area differential service (only to authorized users). For authorized users, navigation positioning precision can be improved by using the equivalent satellite clock (ESC) parameter to reduce the UERE. Currently, the model used to calculate the ESC values is based on pseudo-range observables [4, 5]. In this contribution, the advanced algorithm based on phase observables is proposed and experiments to calculate more precise ESC values are performed in order to improve the UERE.

2 Equivalent Satellite Clock Computation Model

2.1 Equivalent Satellite Clock

As the radial orbit error of BDS satellites present small projection differences in China area, it is difficult to fully separate the radial orbit errors with satellite clock errors. Therefore, it is better to merge them, called the equivalent satellite clock. Ionosphere-free observables for one station and one satellite can be expressed as:

$$P_i = \rho(x^{sat}) + c \cdot (dt_{rec} - dt^{sat}) + (b_{ifb} - b^{igd}) + m \cdot ZTD + \varepsilon \quad (1)$$

P_i is the range observables, ρ is the geometric range between the satellite and receiver, dt_{rec} , dt^{sat} are respectively the receiver and satellite clock offset, b_{ifb} , b^{igd} are respectively the receiver and satellite code hardware delays, m and ZTD are the troposphere mapping function and corresponding zenith troposphere delays, ε represents many kinds of errors including multipath effects. The equivalent satellite clock can be expressed as:

$$ESC = \delta_{clk} + \delta_{orb} \quad (2)$$

δ_{clk} is the satellite clock error, while δ_{orb} is the projection error of satellite orbit. By taking all the other corrections out of the observation equation, the previous equation can be rewritten as:

$$p_i^j = \delta_{rec,i} - \delta_{ESC}^j + \varepsilon_i^j \tag{3}$$

p_i^j is the observation residuals between satellite j and station i , $\delta_{rec,j}$ is the receiver clock error, δ_{ESC}^j is the equivalent satellite clock. Through aforementioned equations, we can construct the following equation, which can be solved using the least squares adjustment:

$$\begin{bmatrix} p_1^1 \\ p_1^2 \\ \vdots \\ p_n^1 \\ p_n^2 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & -1 & 0 & \cdots \\ 1 & 0 & \cdots & 0 & -1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & 0 & 1 & -1 & 0 & \cdots \\ \cdots & 0 & 1 & 0 & -1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \delta_{rec,1} \\ \delta_{rec,2} \\ \vdots \\ \delta_{ESC}^1 \\ \delta_{ESC}^2 \\ \vdots \\ \vdots \end{bmatrix} \tag{4}$$

Due to lack of the time reference, the constructed normal equation is rank defect, which means the normal equation is singular. Therefore, we should firstly define one station clock as the reference, then iteratively estimate all the other clock errors using the least square adjustment, finally we can obtain a set of equivalent satellite clock.

2.2 Variation of the ESC

In Eq. (4), the ESC values are estimated only using pseudo-range observables, which are heavily affected by measurement noise. To improve the estimation precision of the ESC, a two-step smoothing method can be applied to suppress the pseudo-range measurement noise. The first step is to calculate the variation of the epoch-wise ESC values using phase-differenced observables, and the second step is to smooth the pseudo-range observables using the variation of epoch-wise ESC values [6]. The ionosphere-free phase measurement equation is as follows:

$$L = \rho(x^{sat}) + c \cdot (dt_{rec} - dt^{sat}) + (b_{ifb} - b^{tgd}) + N + m \cdot ZTD + \varepsilon \tag{5}$$

ε is the error of phase observable, b_{ifb} , b^{tgd} are receiver and satellite hardware delays respectively; other parameters have the same meaning as the Eq. (1) except for the phase ambiguity parameter N . The phase ambiguity can't be fixed in real

time, which hinder the ESC calculation when using phase observables, however the epoch-wise variation of the ESC can be obtained simply:

$$\Delta L(t_{i-1}, t_i) = \Delta \rho(x_{i-1}^{sat}, x_i^{sat}) + c \cdot (\Delta dt_{rec} - \Delta dt^{sat}) + \Delta m \cdot ZTD + \Delta \varepsilon \quad (6)$$

From Eq. (6), It can be seen that the phase ambiguity problem donot exist any more, and hardware delays are canceled, and the troposphere errors are only related to the mapping function differences. Until now, the high-precision phase observables can be processed using similar least squares adjustment as pseudo-range observables, and the epoch-wise ESC values can be acquired.

2.3 Smoothing

Estimation of the epoch-wise ESC variation in Sect. 2.2 can be used to smooth the estimated ESC absolute values in Sect. 2.1. With the absolute ESC values calculated by pseudo-range observables, we can form a new set of measurement equations as:

$$\hat{X}_i - X_{l,i} = V_{l,i} \quad (7)$$

On the other hand, the estimation of the epoch-wise ESC variation can be represented as:

$$(\hat{X}_i - \hat{X}_{i-1}) - (X_{\varphi,i} - X_{\varphi,i-1}) = V_{\Delta\varphi,i} \quad (8)$$

Through Eqs. (7) and (8), we can get the normal equation as:

$$E^T P_l E \hat{X} = E^T P_l E X_l \quad (9)$$

$$C^T \cdot P_\varphi \cdot C \cdot \hat{X} = C^T \cdot P_\varphi \cdot \Delta x_\varphi \quad (10)$$

C is the corresponding coefficients matrix:

$$C = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}_{n \times n} \quad (11)$$

P_l and P_φ are weight matrix for pseudo-range and phase observables respectively. Through aforementioned smoothing procedure, we can acquire a set of better-quality ESC values by using both pseudo-range and phase observables.

2.4 Improved Equivalent Satellite Clock Estimation Algorithm

Although the two-step algorithm is quite effective in ESC estimation, the absolute ESC values derived in above sections can still be improved. When there are no cycle slips in the carrier phase, the integer ambiguity keeps unique. With this property in mind, we can directly calculate the ESC using phase observables, and it is much more convenient than the above two-step procedure. The improved ESC estimation algorithm based on phase observables can be described using the following steps. It is assumed that the reference epoch of ESC is t_i .

1. Using long-time phase observables, we can fix the phase ambiguity N_i^j in the period of $[t_i - \Delta_t, t_i]$;
2. At epoch t_i , firstly, we perform carrier phase cycle-slip detection. Then we perform polynomial fitting of the phase measurements $\tilde{\varphi}_i$ in the period of $[t_i - \Delta_t, t_i]$ [7]:

$$\tilde{\varphi}_i = a_0 + a_1(t_i - t_0) + a_2(t_i - t_0)^2 + \dots + a_n(t_i - t_0)^n \tag{12}$$

$(i = 1, 2, \dots, m; m > n + 1)$

Through fitting, coefficients a_0, a_1, \dots, a_n , and corresponding root mean squares error $\sigma = \sqrt{\frac{\sum V_i^2}{m - (n + 1)}}$ can be obtained, where V_i are the fitting residuals. With the coefficients estimation, we can extrapolate the phase measurements in the following epoch, and make a comparison between the extrapolated value and real measurement. If the difference is less than 3σ , the total period can be regard as continuous without any cycle slips. In experience, three order polynomials should be used.

3. If there is no cycle slip in the period $[t_i - \Delta_t, t_i]$, the carrier phase ambiguity can be estimated, and then the measurement equations can be simplified as

$$\begin{bmatrix} l_1^1 \\ l_1^2 \\ \vdots \\ l_n^1 \\ l_n^2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & -1 & 0 & \dots \\ 1 & 0 & \dots & 0 & -1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & 0 & 1 & -1 & 0 & \dots \\ \dots & 0 & 1 & 0 & -1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \delta_{rec,1} \\ \delta_{rec,2} \\ \vdots \\ \delta_{ESC}^1 \\ \delta_{ESC}^2 \\ \vdots \end{bmatrix} \tag{13}$$

Using Eq. (13), all the ESC can be estimated directly.

4. Excluding the first measurement in aforementioned period, we add the next carrier phase measurement into the measurement value series, and reprocess

using the above steps. When the difference between extrapolated value and real measurement is larger than 3σ , the ESC estimation procedure should be stopped till the next carrier phase ambiguity solved.

3 User Equivalent Range Error

The UERE are the errors related to the satellite orbits errors, satellite clock errors, atmosphere model errors, receiver clock errors, and so on. It includes the navigation signal range errors (URE) and the users' equipment errors (UEE) [8].

The UERE can be expressed as:

$$UERE^j = \bar{P}^j - \rho^j - \delta_{sat}^j - \delta_{model}^j \quad (14)$$

\bar{P} is the range observation, ρ is the geometric range between receiver and satellite, δ_{model} includes troposphere, ionosphere, relativity, earth displacement and receiver clock correction. By subtracting the ESC, we can get a new variable: the user difference range error (UDRE) [9]:

$$UDRE^j = \bar{P}^j - \rho^j - \delta_{sat}^j - ESC - \delta_{model}^j \quad (15)$$

4 Experiments

The experiment period is from 23/10/2017 to 25/10/2017. In the experiment period, ESCs are calculated using the two-step algorithm and the improved estimation algorithm. Three types of the UERE/UDRE corrections are compared with each other. Mode 1 is pseudo-range corrections based on two-step algorithm, mode 2 is pseudo-range corrections based on improved ESC estimation algorithm, and mode 3 is phase corrections based on improved ESC estimation algorithm.

4.1 User Ranging Error Statistics

4.1.1 Statistical Results of the UERE, UDRE

In Fig. 1, we present the RMS of UERE and UDRE of C14 (MEO) satellite for 3 kinds of ESC correction on 24/10.

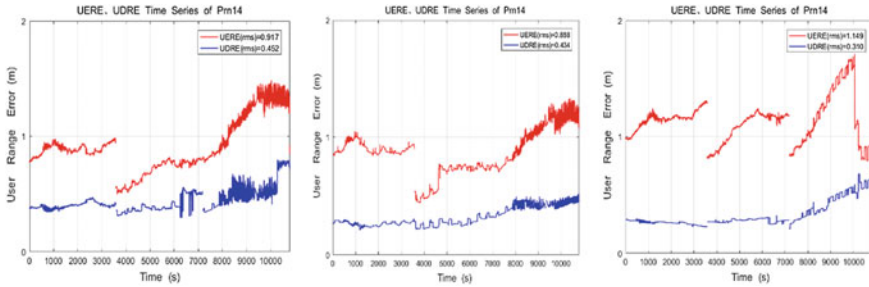


Fig. 1 UERE and UDRE of three kinds of ESC correction for C14 satellite

In the above figures, the red dots represent the UERE of the BDS standard service users, the blue dots represent the UDRE of the BDS authorized service users. From those figures, we can find that the values of UDRE are less than corresponding UERE corrections, which imply the effective use of ESC corrections. Meanwhile, the UDRE of mode 3 is much smoother than that of mode 1 and mode 2, which indicates that the ESC estimation calculated by improved ESC estimation algorithm is much more stable than the two-step algorithm; the differences between UERE and UDRE are similar in mode 1 and mode 2, which indicates that the ESC estimated by improved algorithm may be not compatible with pseudo-range observables.

Table 1 RMS statistics of all UDRE satellites in 3 days (unit: m)

PRN	October 23 UDRE (RMS)			October 24 UDRE (RMS)			October 25 UDRE (RMS)		
	1	2	3	1	2	3	1	2	3
1	0.334	0.307	0.215	0.419	0.407	0.172	0.374	0.393	0.154
2	0.338	0.354	0.236	0.442	0.472	0.256	0.406	0.380	0.158
3	0.375	0.387	0.232	0.420	0.472	0.180	0.343	0.361	0.212
4	0.473	0.455	0.215	0.785	0.472	0.287	0.346	0.359	0.264
5	0.431	0.382	0.288	0.436	0.482	0.294	0.529	0.520	0.294
6	0.447	0.391	0.319	0.479	0.511	0.256	0.922	0.857	0.368
7	0.631	0.608	0.381	0.455	0.505	0.293	0.587	0.653	0.264
8	0.447	0.440	0.331	0.501	0.436	0.371	0.521	0.473	0.235
9	0.415	0.383	0.315	0.559	0.614	0.356	0.487	0.414	0.158
10	0.373	0.432	0.240	0.581	0.577	0.399	0.362	0.350	0.141
11	0.322	0.281	0.208	0.712	0.727	0.575	0.649	0.697	0.367
12	0.461	0.478	0.258	0.495	0.469	0.301	0.584	0.575	0.265
13	0.525	0.519	0.295	0.528	0.554	0.293	0.529	0.574	0.264
14	0.497	0.494	0.370	0.660	0.658	0.378	0.514	0.546	0.328

4.1.2 RMS of the UDRE

The RMS of UDRE is shown in Table 1 and Fig. 2 for three days and three modes.

From Table 1 and Fig. 2, we can find that mode 3 is superior to mode 1 and 2 with much smaller RMS of UDRE. The RMS of UDRE for mode 1 and mode 2 are similar, which indicates that phase observables are superior to pseudo-range observables in UDRE calculation.

4.1.3 Comparison

To quantify the precision improved in mode 3, we present the proportion reduction for mode 1 and mode 3 in Fig. 3:

Form Fig. 3, we can find that the average of proportion reduction is 32.1% for mode 1, and 50.1% for mode 3, which verifies the high-precision property of the improved ESC estimation.

4.2 Positioning Results Statistics

In order to study the effect of ESC on the positioning results, four MEGX stations DUND, DAE2, GMSD and JFNG were selected. And the results of dual-frequency positioning under standard service and enhanced service are compared.

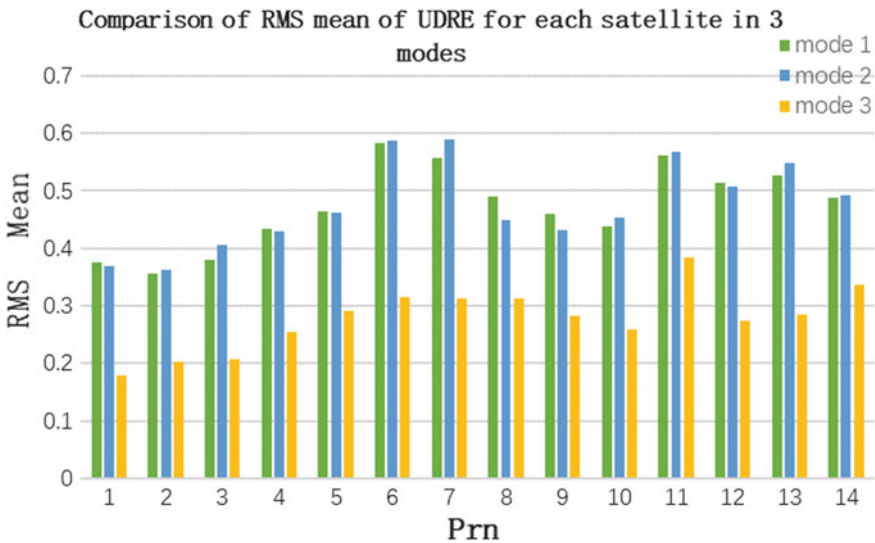


Fig. 2 RMS of UDRE for three kinds of mode for all BDS satellites

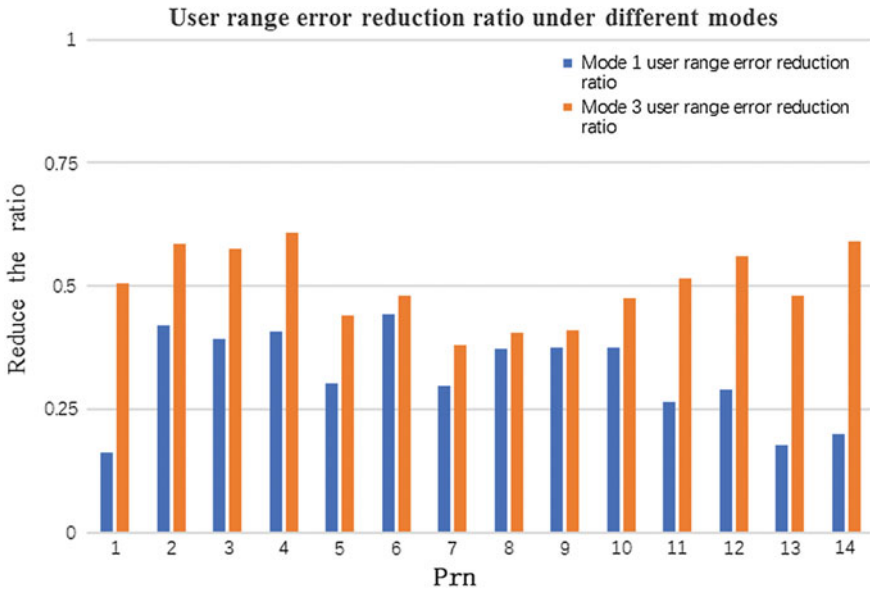


Fig. 3 Proportion reduction of satellite user range error for Mode 1, Mode 3

4.2.1 Positioning Results of Standard Service and Enhanced Service

Taking the positioning results of GMSD station on October 24 as an example, statistics are made on the kinematic positioning results of the users under the standard and enhanced service, respectively. The results are as shown in Fig. 4.

The left two pictures in Fig. 4 are the positioning results of pseudo-range and carrier phase measurements under the standard service of GMSD station respectively. The two pictures on the right are the positioning results of pseudo-range and carrier phase measurements under enhanced service of GMSD station. In Fig. 4, RMS of north, east and up positioning error shows that after the ESC correction, enhanced positioning results are significantly better than that of standard positioning results.

4.2.2 Positioning to Enhance the Proportion of Statistical Results

Horizontal, vertical and three-dimension positioning results of four MEGX stations are obtained. And the proportion improvement of enhanced navigation relative to the standard navigation is shown in Table 2.

As can be seen from Table 2, the horizontal, vertical and three-dimension positioning results of the four MEGX stations have been improved after the ESC correction adding to the measurements. This proves that the ESC can be used to improve the positioning accuracy of users required better navigation services. It is

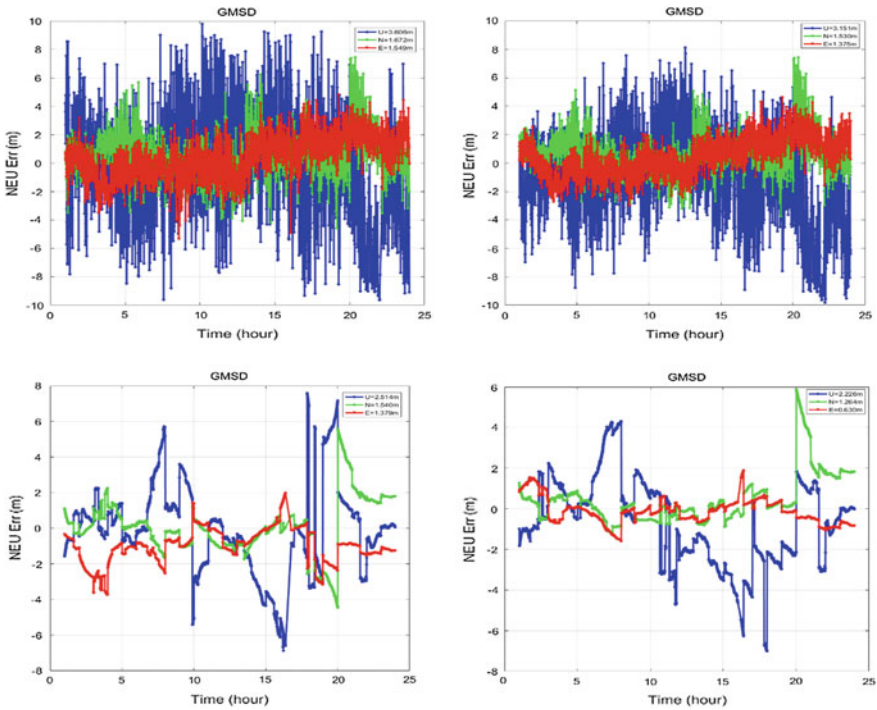


Fig. 4 Positioning results of GMSD under standard service and enhanced service respectively

Table 2 Proportion improvement of enhanced navigation relative to the standard navigation

Station	Pseudo-range enhancement navigation boost ratio (%)			Carrier phase enhancement navigation boost ratio		
	Horizontal	Vertical	3D	Horizontal	Vertical	3D
DUND	10.54	3.32	5.83	12.79	7.86	10.62
DAE2	13.43	7.98	6.13	4.58	23.0	18.46
GMSD	9.75	12.62	11.79	31.68	11.46	19.01
JFNG	8.45	8.80	8.84	23.53	22.72	21.09

noted that the average improvement percentage of the positioning result is smaller than the proportion reduction of the user’s ranging error. This is because the BDS reference station is selected to calculate the user ranging error, and the quality of the observations of the MEGX stations used in the positioning is relatively poor, and the DOP of satellites for MEGX station is not too good.

5 Conclusion

The broadcast equivalent satellite clock corrections are very useful to improve navigational positioning precision in the BDS. In this study, we present the old-fashion two-step pseudo-range-based algorithm to calculate the ESC corrections, and propose a new advanced phase-based algorithm to improve the calculation of ESC corrections. Through 3 days experiment Beidou observables, the results show that the RMS of the UDRE corrections gets significant reduction, with reduction ratio reaching to 50.1%; The results of dynamic positioning of four MEGX observatories under basic service and enhanced service are calculated, it is shown that horizontal, vertical and three-dimensional positioning results of WADS are better than that of the standard PNT service.

Acknowledgements This work is support by the NSFC (No. 11673050) and the 863 (No. 2014AA123102) projects.

References

1. Misra P (2011) Global positioning system: signals, measurements, and performance. Ganga-Jamuna Press, Lincoln
2. Wei Z, Wang G (2000) Wide area differential GPS. Bull Survey Mapp 2000(06)
3. Tsai Y-J (1999) Wide area differential operation of the global positioning system: Ephemeris and clock algorithms. Ph. D. thesis, Stanford University
4. Zhou SS, Hu X, Wu B et al (2011) Orbit determination and time synchronization for a GEO/IGSO satellite navigation constellation with regional tephemeris and clock algorithm tracking network. Sci China Phys Mech Astron 54:1089–1097
5. Xing N (2013) Research on the application of wide area differential technique to compass system. University of Chinese Academy of Science, Beijing
6. Chen J, Yang S, Zhou J et al (2017) A pseudo-range and phase combined SBAS differential correction model. Acta Geodaet Cartogr Sinica 46(5):537–546
7. Yang S (2017) Research on BDS decimeter level SBAS and its performance assesment. University of Chinese Academy of Science, Beijing
8. Jiao W, Ding Q, Li J, Lu X et al (2011) Monitoring and assessment of GNSS open service. Sci China Phys Mech Astron
9. Cao Y (2014) Studies on the wide-area differential correction and integrity monitoring for the regional satellite navigation system of BeiDou. University of Chinese Academy of Science, Beijing

Rapid Ambiguity Resolution Algorithm for Multi-constellation Between Reference Stations Based on Ambiguity Tight Constraint



Ruicheng Zhang, Chengfa Gao, Shuguo Pan, Zhiyue Yan
and Qing Zhao

Abstract The development of multi-GNSS remarkably increased the number of available satellites, but how to solve the multi-dimensional ambiguity parameters quickly and accurately in network RTK technology is still an issue full of perplexity and significance. To ensure the virtual station have more available satellites under the occlusion environment, a fast ambiguity resolution method for base station based on ambiguity tight constraint was proposed in this paper. Firstly, the optimal subset of ambiguity is selected by partial ambiguity resolution (PAR) strategy, and then impose strong constraints on the ambiguities of these satellites. Finally, update the filter equation and assist in fixing the ambiguity of other satellites. The real measured baseline data which contain GPS, BDS and GLONASS from Tianjin CORS and Curtin University was used in the experiments, and the results illustrated that this method could significantly shorten the initialization time of ambiguity between base stations, accelerate the convergence speed of newly-arisen satellites, and increase the number of available satellites of RTK virtual observations (especially low-elevation angle satellites), that providing a reliable guarantee for high-precision positioning in the occlusion environment, such as the roads in cities.

Keywords Rapid ambiguity resolution · Multi-GNSS · Reference station
Tight constraint · Partial ambiguity resolution (PAR)

R. Zhang · C. Gao (✉) · Q. Zhao
School of Transportation, Southeast University, Nanjing 210096, China
e-mail: gaochfa@163.com

S. Pan (✉)
School of Instrument Science and Engineering, Southeast University,
Nanjing 210096, China
e-mail: psg@seu.edu.cn

Z. Yan
Nanjing Compass Navigation Technology Co., Ltd., Nanjing 210096, China

1 Instruction

With the continuous development of autonomous vehicle and digitized cities, all walks of life have put forward higher requirements for the continuity, real-time and precision positioning. What's more, combined with the completed GPS, GLONASS and rapidly idealizing BDS and Galileo, the number of orbiting satellites in the global satellite navigation system has also rose up considerably, which laying the foundation for high-precision positioning. At the same time, network RTK technology attracts much attention in positioning applications owing to its uniform and high precision positioning results over a large space [1, 2], among which, how to fix the ambiguity quickly and accurately has always been the research hotspot and the kernel contents of GNSS data processing.

For the sake of improving the reliability, success rate and initial time of ambiguity resolution in different length, some scholars took advantages of tri-frequency that can be composed of a series of long-wavelength, low-noise levels and reduced ionospheric effects observations [3]. However, the currently used CORS system and RTK equipment are generally dual-frequency receivers, this method has not been put into practical application.

In addition, in order to further reduce the search space of ambiguities for improving the search efficiency, many scholars have proposed the ambiguity resolution method that only fix partial ambiguities that are easy to fix. Yang and Wang et al. took the variance of ambiguity as the first choice principle, gave priority to fix ambiguities with small variance [4, 5]. Li et al. proposed a partial ambiguity resolution method of increasing the cut-off angle of the satellite to satisfy the fixed success rate and the Ratio threshold [6]. What's more, there are some researchers declared a partial ambiguity fixed scheme, by selecting the optimal subset to ameliorate the ambiguity fixed efficiency [7].

But most of the existing methods tend to fix the ambiguities that are easy to fix, in particular, the partial ambiguity resolution abandon some of the available satellites and fail to make full use of all the satellite observations. For network RTK (NRTK) data processing centres, how to increase the number of fixed satellites as much as possible and use them to generate virtual observations is the premise to promote usability and reliability in occlusion areas. Based on this reason, this paper presented a method of rapid ambiguity resolution for GPS/BDS/GLONASS (G/C/R) triple-system base station based on the tight constraint of ambiguity. Firstly, the PAR strategy was used to select optimal ambiguity subset, in which the ambiguities are considered as accurately fixed. Then, the fixed consequence of these ambiguities were used to update the filter equation as "pseudo" observations, and this process was called "ambiguity tight constraint". Finally, the ambiguities of all the satellites are solved by Lambda. What's more, the real measured baseline data of Tianjin CORS was used to validate this method.

2 Optimal Subset Selection Scheme

When the tight constraint algorithm is used, incorrect strong constraints will have an irreversible effect on the ambiguities. Therefore, how to make sure the correctness of these ambiguities, that will be used to constraint other ambiguities, become the most essential part of this algorithm. In this paper, the optimal subset of ambiguities is selected by the idea of PAR, and the ambiguities corresponding to the selected optimal subsets are considered to be exact and fixed, preparing for the ambiguity constraint in the next step. Specific algorithm is as follows:

- (1) According to the deviation of wide-lane ambiguity float solution, elevation angle and the number of continuous filtering (threshold values of 0.25 cycle/30°/20 times), the ambiguities are simply divided into two parts, easy to fix and hard to fix. The corresponding variance-covariance matrix can be written as:

$$Q = \begin{bmatrix} Q_{\hat{N}_s} & Q_{\hat{N}_s \hat{N}_b} \\ Q_{\hat{N}_b \hat{N}_s} & Q_{\hat{N}_b} \end{bmatrix} \tag{1}$$

- (2) Sorting the diagonal elements of $Q_{\hat{N}_s}$ ascending, then a new sequence of satellite ambiguities is obtained: $N = \{N_1, N_2, \dots, N_s | Q_{n_1} < Q_{n_2} < \dots < Q_{n_s}\}$, where Q_{n_i} represents the ambiguity variance of i th un-reference satellite.
- (3) Lambda search is applied to the initial subset of ambiguities, and the ambiguity subset would be considered as successfully selected, while the ambiguity bootstrapping rounding success rate P (hereinafter referred to as ambiguity fixed success rate) is higher than 99.9% and the ratio value exceed 4.0.
- (4) If the ambiguity search results do not satisfy the conditions in step (3), remove the satellites with the largest variance in the ambiguity subset and repeat the steps in (3) after obtaining a new subset of ambiguities until all the conditions in (3) are satisfied, then the selected subset can be contemplated as the optimal subset. If the number of remained satellites less than 10, the selection process will be terminated too, and the “tight constraint” would not be applied to current epoch.

3 Ambiguity Tight Constraint Algorithm

During the process of fixing the ambiguities between network RTK base stations, the ambiguities of newly-arisen satellites are very arduous to fix on account of their large atmospheric error (especially the tropospheric delay), and the ratio value of the overall ambiguity resolution will be affect when they are added to the filter equation. In this paper, a strong constraint is imposed on the ambiguities belong to

the optimal subset that are selected by the partial ambiguity resolution strategy in Sect. 1, then utilizing the correlation between ambiguities (a relative zenith tropospheric delay is estimated), assist low-altitude satellites and other hard-to-fix satellites to be fixed. The specific algorithm is as follows.

3.1 Construct “Pseudo” Observations

Construct “pseudo” observations for the ambiguities in the optimal subset, including the value and precision of the observations.

The value of the “pseudo” observation is the base ambiguity that are fixed, acquiring from the previous epoch lambda search. In view of the strong constraint of ambiguity will be imposed on the optimal subset, the accuracy of the “pseudo” observations was set at 0.001 cycle, which is considered to be fixed correctly. “Pseudo” observation L and its degree of accuracy R are as follows:

$$L = \begin{bmatrix} N_{fixed}^1 \\ \vdots \\ N_{fixed}^i \\ \vdots \\ N_{fixed}^H \end{bmatrix}, \quad R = \begin{bmatrix} \delta_c^2 & & & \\ & \ddots & & \\ & & \delta_c^2 & \\ & & & \underbrace{\delta_c^2}_s \end{bmatrix} \quad (2)$$

where N_{fixed}^i represents the double-difference ambiguity integer solution of the i th satellite in the optimal subset, and H represents the size of the optimal subset. $\delta_c = 0.001$ cycle.

3.2 Update the Filter Equations

Adding the “pseudo” observations to the Kalman filter (KF) equations and update them, then the Lambda algorithm will be used to search all ambiguities. The wide-lane observation and ionospheric-free model are utilized to fix the ambiguities in this paper, specific algorithm can consult the reference literature [8].

It can be seen from the above introduction that the “pseudo” observation constructed in this section is the state estimation value of the filter equations, combined with the “pseudo” observations in (1), matrix B and V can be written as:

$$\mathbf{B} = \begin{bmatrix} 1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & \ddots & \vdots \\ & & \underbrace{\hspace{2cm}}_H & 1 & & \\ & & & & \underbrace{\hspace{2cm}}_{m-H+1} & \\ & & & & & 0 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} N_{float}^1 - N_{fixed}^1 \\ \vdots \\ N_{float}^i - N_{fixed}^i \\ \vdots \\ N_{float}^H - N_{fixed}^H \end{bmatrix} \quad (3)$$

where H denotes the size of the optimal subset, and m presents the number of all satellites participating in the KF equations; N_{float}^i represents the double difference float ambiguity of the i th satellite in the optimal subset.

Then, substitute the observation noise, design matrix, observation residuals matrix mentioned above in the KF equations and update the equations.

4 Experiment and Analysis

The triangle network contributed with BD, DHP and DWG stations in Tianjin CORS is selected for experimental verification, calculating effect of ambiguity resolution of three baselines. The distribution of base stations and the length of baselines are shown in Fig. 1. In addition, two stations of Curtin, CUTB and CUTC, are selected to form a short baseline, which is used to verify the positioning accuracy promoted by newly-arisen satellites in the occlusion environment, the distribution and baseline length of the base stations are shown in Fig. 1 too. What’s more, the details of two sets of data are shown in Table 1.

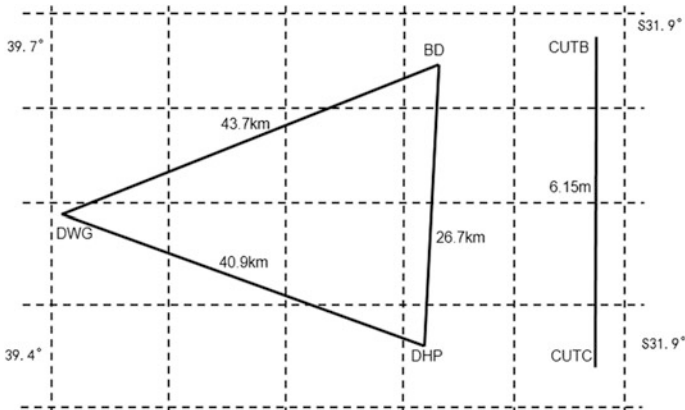


Fig. 1 The base station distribution of Tianjin CORS and curtin

Table 1 The detailed information of observations

	Time	Receivers	Interval (s)	Ele threshold (°)	Remark
Tianjin CORS	2016:267 03:00–07:00	TRIMBLE BD970	1	10	The filter interval if 5 s
Curtin	2017:113 01:00–02:00	TRIMBLE NETR9	1	10	The existential time of new rising satellite is less than 1 h

4.1 *The Function of Tight Constraint on Ambiguity Resolution*

Figures 2 and 3 show the ADOP and Ratio values of the three baselines (DWG-DHP, DWG-BD and BD-DHP respectively) before and after the use of the ambiguity tight constraint algorithm. The ADOP values can be utilized for assessing the success rate of ambiguity resolution. It is generally accepted that the ambiguity fix success rate exceed 99.9% if ADOP value is less than 0.12 cycles [9]. Moreover, it should be noted that the number of fixed satellites in all three baselines during the entire observation period exceeded 21, and the maximum number reached 26. Due to space limitations, this article will not detail the states of fixed satellites.

As illustrated in Fig. 2, the ADOP values of the three baselines are markedly reduced with the tight constraint strategy, and the ambiguity fixed success rate of 99.9% can be accomplished within only 1–2 epochs, and the two bumps in the figure are due to the initialization or reinitialization. After partial magnification, it could be found that the ADOP value that after using the tight constraint strategy can reach the threshold quite faster, and correspondingly, the initialization time is shorter.

As seen from Fig. 3, the Ratio value had been significantly ameliorated with the use of the ambiguity constraint, and Ratio could recover to a threshold of 2.0 over several epochs even in the presence of newly-arisen satellites. However, the traditional method requires hundreds of epochs to reach the threshold. What's more, the statistics show that the percentage of epochs where Ratio exceeded the threshold under the constrained strategy is more than 80%, which is obviously better than the traditional filtering method. The mutation case of Ratio was caused by the change of the number of satellites. It can be seen from the results that the tight constrained strategies, due to the strong constraints imposed on the ambiguities in the optimal subset, expedites the ambiguity resolution of newly-arisen satellites and makes better use of the information of all satellites, especially the low elevation satellites.

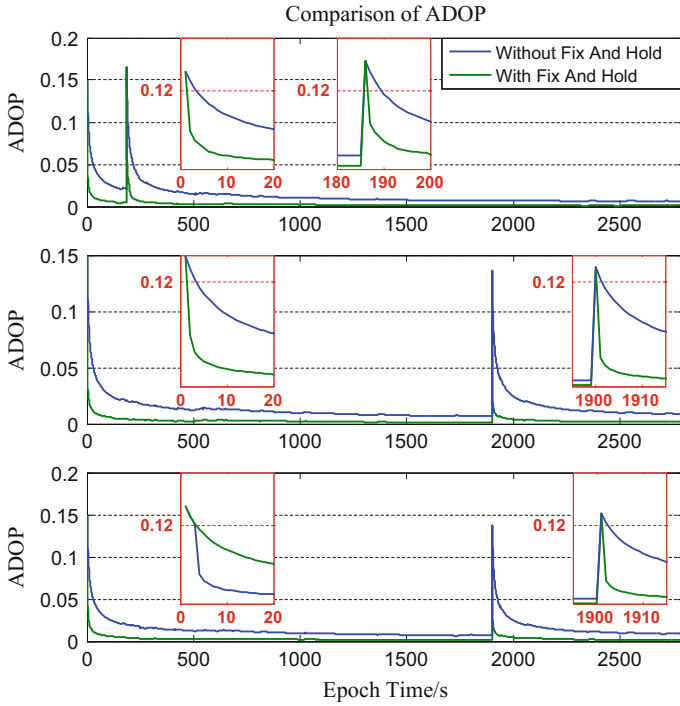


Fig. 2 The comparison of ADOP of two methods

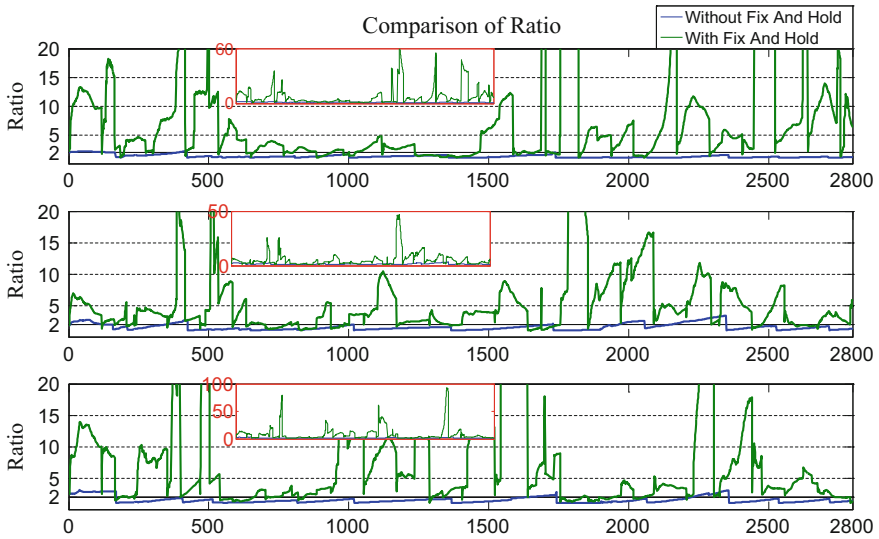


Fig. 3 The ratio of three baselines with two methods

4.2 The Effect of Newly-Arisen Satellite on Positioning Accuracy

In this paper, some satellites on specific azimuth were removed to simulate the RTK positioning in the occluded environment. In the comparative experiment, newly-arisen satellite R21 was added to verify the improvement of the positioning accuracy of low-elevation satellites in the occluded environment. The sky plot of two experiments are shown in Fig. 4.

Calculating the PDOP values (Position Dilution of Precision) of two sets of data, it was found that the satellite structure has been outstandingly enhance after adding the newly-arisen satellite R21. At the same time, the space distribution of the satellites in the east-west direction was more reasonable (since most of the satellites were on the left side and the R21 satellite was on the right side). The comparison of PDOP is shown in Fig. 5.

Further analysis indicated that after the addition of newly-arisen satellite R21, the positioning accuracy in E direction and U direction was greatly improved, and the accuracy in N direction remains unchanged (but the accuracy is better than E direction), which was also consistent with the satellite spatial distribution analysis results above. It can also be seen from Fig. 6 that there are two mutations in RTK positioning results (marked by the red dashed box), which may be due to the overall poor satellite geometry distribution. The analysis of the mutation results evidenced that the outlier after adding newly-arisen satellite R21 was distinctly smaller than the outlier without adding newly-arisen satellite, and the outlier time range was also less than the former, which certified the addition of new-arisen satellite has a certain degree of resistance to the outliers, thereby increasing the reliability of positioning.

The RTK positioning results of two groups were evaluated. The statistical results are shown in Table 2.

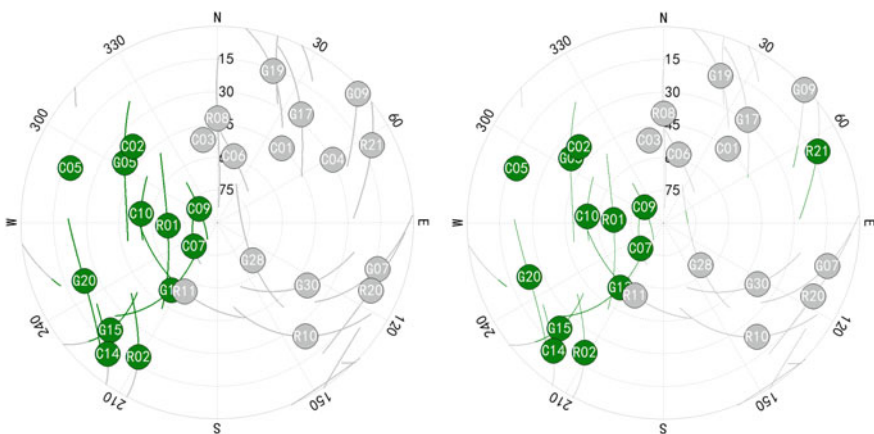


Fig. 4 The sky plot of two experiments

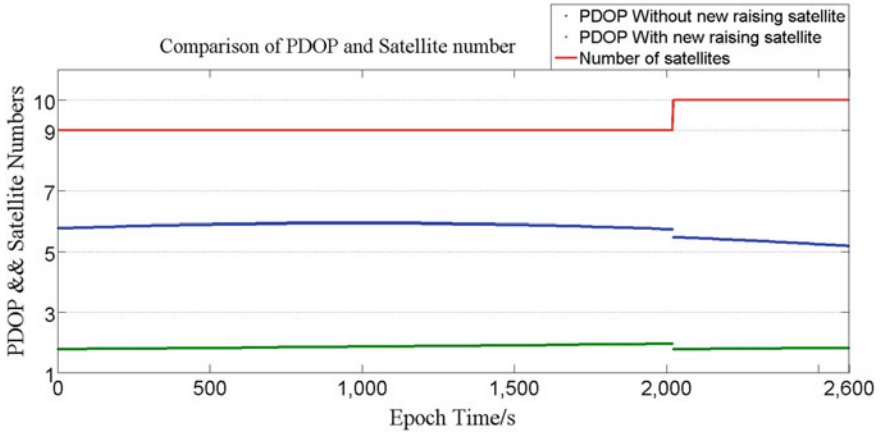


Fig. 5 The PDOP and number of satellites of two experiments

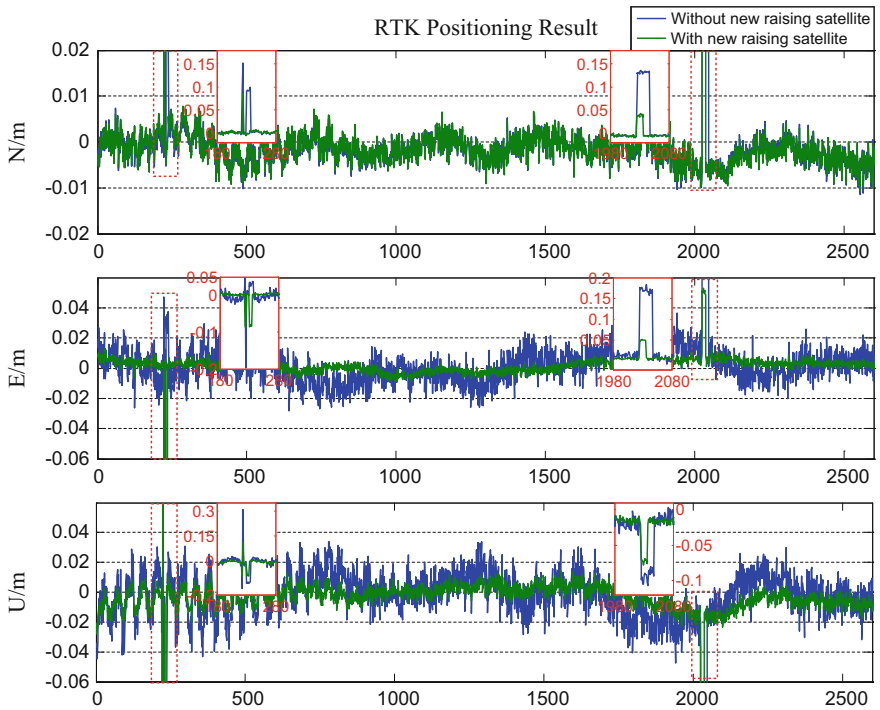


Fig. 6 The positioning result of two experiments

Table 2 The statistical results of two experiments

Experiment		Content			
		N direction (m)	E direction (m)	Plane (m)	Elevation (m)
1st experiments	MEAN	-0.0001	0.0033	0.0033	-0.0026
	RMSE	0.0143	0.0197	0.0243	0.0189
2nd experiments	MEAN	-0.0014	0.0018	0.0023	-0.0035
	RMSE	0.0045	0.0068	0.0082	0.0100

It can be observed from the above table that the positioning accuracy after adding satellite R21 has been appreciably improved both in plane and vertical direction, and in N/E/U directions, has a rise of 68, 65 and 47% respectively.

5 Conclusion

Aiming at the difficulty of ambiguity resolution when newly-arisen satellites appears during the process of network RTK baselines, a fast ambiguity resolution algorithm based on constraint ambiguity is proposed derived from the traditional ambiguity resolution model method. The method sets floating deviation, continuous locking times and elevation threshold as preconditions, and determines the optimal subsets selection order by the variance of state estimator. Then the ratio and the ambiguity fixed success rate are selected as the optimal subset selection conditions, then strong constraints are imposed on the ambiguities among the selected optimal subsets to speed up the fixed rate of newly-arisen satellites. The experimental results show that the ratio and success rate of the ambiguity are noticeably improved after using the tight constraint strategy, and the initialization time is correspondingly shortened. Furthermore, in the case of newly-arisen satellites, using the tight constraint strategy can make the Ratio value recover to above the threshold in several epochs, and compared with the traditional PAR method, the new algorithm can heighten the availability of satellites to RTK users. At the same time, the experiments verify that the newly-arisen satellites in the occlusion area can improve the RTK positioning accuracy by 65 and 47%, respectively in the plane and elevation directions, which also confirms that the proposed method of tight constraint ambiguity makes a significant contribution to the positioning accuracy under complex environment.

Acknowledgements Thanks for the generous help from my teachers and schoolmates, who provided some advice for this paper. Thanks also go to the National Natural Science Foundation of China (No: 41574026, 41774027), Primary Research & Development Plan of Jiangsu province (Grant Number BE2016176), National Key Technologies R&D Program (Grant Number 2016YFB0502101) and Six Talent Peaks Project in Jiangsu Province (Grant Number 2015-WLW-002).

References

1. Landau L, Vollath U, Chen XM (2002) Virtual reference station system. *J Global Position Syst* 1(2):137–143
2. Liu J, Liu H (2003) Continuous operational reference system: infrastructure of urban spatial data. *Geomatics Inf Sci Wuhan Univ* 28(3):259–264
3. Gao W, Gao C, Pan S (2015) Single-epoch positioning method in network RTK with BDS triple-frequency widelane combinations. *Acta Geodaetica Cartogr Sin* 44(6):641–648
4. Wang J (2013) Reliability of partial ambiguity fixing with multiple GNSS constellations. *J Geodesy* 87(1):1–14
5. Yang R, Ou J, Yuan Y (2007) Facilitating efficiency and success rate of resolving GPS phase ambiguity with parts search method. *Geomatics Inf Sci Wuhan Univ* 32(2):160–163
6. Li B, Shen Y, Feng Y (2014) GNSS ambiguity resolution with controllable failure rate for long baseline network RTK. *J Geodesy* 88(2):99–112
7. Gao W, Gao C, Pan S (2017) Fast ambiguity resolution between GPS/GLONASS/BDS combined long-range base stations based on partial-fixing strategy. *Geomatics Inf Sci Wuhan Univ* 42(4):558–562
8. Allison T (1991) Multi-observable processing techniques for precise relative positioning. In: *Proceedings ION GPS-91*. Albuquerque, New Mexico
9. Teunissen PJG (1997) A canonical theory for short GPS baselines. Part IV: precision versus reliability. *J Geodesy* 71(9):513–525

Part III
Time-Space Datum and Temporal
Frequency Technologies

Using IGS RTS Products for Real-Time Subnanosecond Level Time Transfer



Guojun Li, Yuting Lin, Fengfeng Shi, Jialin Liu,
Yuting Yang and Junwen Shi

Abstract Time transfer by precise point positioning has the defect of long latency resulting from IGR products. GPS common-view can be updated once every 16 min, but with a precision of about 3–5 ns. A real-time precise point positioning time transfer algorithm using IGS RTS (Real-time Service) products was proposed. It was proved to be practical through the time transfer experiments among 4 time laboratories in Western Europe. The time transfer results show that the accuracy of the new algorithm can be reach to 0.30 ns for RMS and 0.25 ns for STD. Moreover, the stability of the time transfer results is up to $2E-15$ at 1 day averaging.

Keywords IGS RTS · Time transfer · Real-time precise point positioning
Subnanosecond

1 Introduction

Presently, the time transfer methods for time laboratories participating in TAI (International Atom Time) consist of GPS AV (GPS All-in-view) [1], GLONASS AV (GLONASS All-in-view) [2], GPS PPP (GPS Precise Point Positioning) [3], TWSTFT (Two-way Satellite Time and Frequency Transfer) and et al. As an important time transfer method, GPS PPP accounts for 48% time links with increasing proportion in recent years [4]. It's widely used in worldwide time laboratories.

GPS PPP has great advantages in high-precision time and frequency transfer over a long distance. It can provide a frequency stability of $1E-15$ to $1E-16$ over an averaging period of 1 day, moreover, as type B uncertainty is less than 0.3 ns [5, 6]. However, GPS PPP requires precise ephemeris products. The ephemeris

G. Li (✉) · Y. Lin · F. Shi · J. Liu ·
Y. Yang · J. Shi
Beijing Satellite Navigation Center, Beijing, China
e-mail: 1010551750@qq.com

products used in BIPM for GPS PPP time transfer is the IGR (IGS Rapid), whose latency is 17–41 h [7]. The IGU is a real-time precise ephemeris product for real-time cm-level positioning. The accuracy of satellite clocks is about 0.15 ns for the IGU observed half, and is 3 ns for the IGU predicted half. Thus, the IGU is unusable for real-time subnanosecond level time transfer.

There are many researches on GPS PPP, which mainly focus on post-processing time transfer [8, 9]. Many time laboratories can hardly monitor the real-time time and frequency signals. The methods for nanosecond and subnanosecond level time transfer are far from practical applications. In this paper, firstly, we briefly introduced the organization of IGS RTS and the usages of IGS RTS products. Secondly, the accuracy and availability of IGC01 were analyzed. Thirdly, the experiment, GPS PPP time transfer with IGC01 among four time laboratories in Western Europe, proved that it was feasible to realize subnanosecond level time transfer with IGS RTS products.

2 The IGS RTS Products

2.1 *The Organization of the IGS RTS*

Since its inception in 1994, IGS has produced high-quality GNSS data products from a cooperative global infrastructure. The IGS products enable access to the definitive global reference frame for scientific, educational, infrastructure, and people's livelihood. With the growing developments of GNSS application, IGS users have expressed a desire for real-time products. In 2001, the IGS Real-Time Working Group (RTWG) was established [10]. During the IGS 2002 Workshop, held under the theme "Towards Real-Time", a framework for development of a real-time service (RTS) was defined [11]. In 2007, the IGS Real-time Pilot Project started to be constructed, and the RTWG declared that the pilot project had reached the initial operating capability.

The IGS RTS is overseen by the RTWG. Parts of importation organizations are as follow [12]:

- (1) Analysis centers, including BKG, CNES, DLR, GFZ, ESA/ESOC, GMV, Geo++, NRCAN, TUW and WUHAN, has responsibility for the generation of precise ephemeris products with observations provided by global GNSS tracking networks. In addition, NRCAN, ESA/ESOC, and BKG are in charge of supervision, coordination and operation.
- (2) Combination centers, including ESA/ESOC, BKG and NRCAN, produce the official combination products by realigning, detection and elimination of outliers and averaging.
- (3) The products distribution centers, including two primary products distribution centers and a number of secondary centers, use Networked Transport of RTCM via Internet Protocol (NTRIP) for streaming GNSS and differential correction

Table 1 The information of IGS RTS products

RTS products	Ref point	RTCM messages
IGS01	APC	1059 (5), 1060 (5)
IGC01	CoM	1059 (5), 1060 (5)
IGS02	APC	1057 (60), 1058 (10), 1059 (10)
IGS03	APC	1057 (60), 1058 (10), 1059 (10) 1063 (60), 1064 (10), 1065 (10)

data over the Internet. Users must complete the online subscriber registration. After authentication, users will be in contact with login and be access to the RTS streams.

The official products currently include IGS01/IGC01, IGS02 and IGS03 [12] (Table 1). IGS01/IGC01 is a single-epoch GPS combination solution. The solution epochs in this product are completely independent of each other. IGS02 is a Kalman filter combination product with precise ephemeris products estimated by individual analysis centers. IGS03 is similar to IGS02, except that IGS03 contains the GPS and GLONASS corrections which are offered as an experimental product.

2.2 Real-Time Satellite Precise Ephemeris and Clocks

(1) Real-time precise satellite ephemeris

RTS products are the corrections refer to broadcast navigation data. The orbit corrections are defined in radial, along-track and cross-track in satellite body-fixed coordinate system. If the orbit correction vector is $[\delta O_r \quad \delta O_a \quad \delta O_c]^T$ at epoch t_0 , and the velocity vector is $[\delta \dot{O}_r \quad \delta \dot{O}_a \quad \delta \dot{O}_c]^T$. Three steps are required to compute the real-time satellite position at current epoch. Firstly, compute the real-time correction in satellite body-fixed coordinate system at current epoch t [13].

$$\begin{bmatrix} \delta_r \\ \delta_a \\ \delta_c \end{bmatrix} = \begin{bmatrix} \delta O_r \\ \delta O_a \\ \delta O_c \end{bmatrix} + \begin{bmatrix} \delta \dot{O}_r \\ \delta \dot{O}_a \\ \delta \dot{O}_c \end{bmatrix} (t - t_0) \quad (1)$$

where δ_r , δ_a , and δ_c are the radial, along-track and cross-track corrections. The orbit corrections must be transfer to geocentric corrections. Secondly, compute the transformation matrix:

$$\mathbf{R} = \begin{bmatrix} \frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|} \times \frac{\mathbf{r} \times \dot{\mathbf{r}}}{|\mathbf{r} \times \dot{\mathbf{r}}|} & \frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|} & \frac{\mathbf{r} \times \dot{\mathbf{r}}}{|\mathbf{r} \times \dot{\mathbf{r}}|} \end{bmatrix}^T \quad (2)$$

where \mathbf{r} is the satellite broadcast position vector and $\dot{\mathbf{r}}$ is the satellite broadcast velocity vector. Thirdly, apply the real-time corrections to the broadcast coordinates.

$$\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \begin{bmatrix} X_{\text{brdc}} \\ Y_{\text{brdc}} \\ Z_{\text{brdc}} \end{bmatrix} - \mathbf{R} \begin{bmatrix} \delta_r \\ \delta_a \\ \delta_c \end{bmatrix} \quad (3)$$

where X_s, Y_s, Z_s are the satellite coordinate in the ECEF coordinate system. $X_{\text{brdc}}, Y_{\text{brdc}}, Z_{\text{brdc}}$ are the satellite position computed from broadcast ephemerides.

The RTS products refer to the satellite CoM (Center of Mass) or APC (Antenna Phase Center). There is no need to apply APC correction if the products refer to the APC, such as IGS01, IGS02 and IGS03. However, the APC correction must be applied if the products refer to the CoM.

(2) Real-time precise satellite clocks

If the polynomial coefficients in RTS products are C_0, C_1 and C_2 at epoch t_0 . The precise satellite clock correction at current epoch t can be calculated as follows:

$$\delta C = C_0 + C_1(t - t_0) + C_2(t - t_2)^2 \quad (4)$$

Apply the corrections to the broadcast clock:

$$dt^{\text{sat}} = dt_{\text{brdc}}^{\text{sat}} + \frac{\delta C}{c_{\text{light}}} \quad (5)$$

where $dt_{\text{brdc}}^{\text{sat}}$ is the broadcast clock corrected with relativistic effect at current epoch t . c_{light} is the speed of light in vacuum.

2.3 The Accuracy and Availability of the IGS RTS Products

To monitor the accuracy of the RTS products, the daily comparison reports between the real-time products of each analysis center and the IGR are provided by EAS/ESOC. The precision of precise point positioning with the RTS products is monitored. According to the statistical results in recent years [14], the satellite orbit accuracy is about 5 cm, as same as IGU. While the satellite clock accuracy is about 0.3 ns, which is much better than the accuracy of the IGU predicted half.

For the convenient to analysis, we take IGC01 as an example to analyze its accuracy and availability owing that its reference is same with the IGR's. The software BNC [15] recommended by the official was used to store the IGC01 streams from WUHAN (<http://gnsslab.ntrip.cn>) in RINEX at intervals of 30 s. The totaled data span is 16 day, from 2017-09-10 to 2017-09-25.

It's remarkable that the real-time satellite clocks of individual analysis center are possible in different time scale. The difference of the time references must be taken into consideration when analyze the accuracy of the satellite clocks. In this paper we take the average value of all the satellite clocks in each epoch as the time

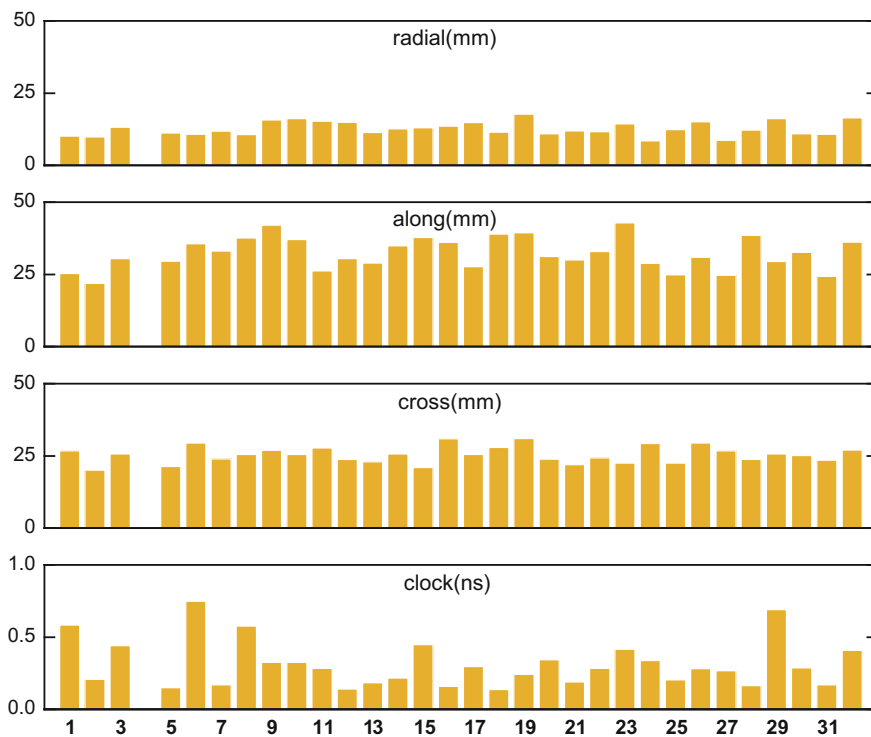


Fig. 1 The RMS of IGC01 ephemeris and clock

reference. In addition, the clock correction resulting from the orbit radial correction is applied. Figure 1 shows the RMS of the satellite orbit and clock difference of the IGC01 against the IGR. Table 2 presents the statistic of the IGC01 products.

As shown in Fig. 1 and Table 2, The RMS of the satellite orbits in IGC01 is 33 mm for the radial component, which is 1–2 times better than 32 mm for the along-track component and 26 mm for the cross-track component. The STD and the RMS of the satellite 3D-orbit is 20 and 44 mm. The STD and the RMS of the satellite 3D-orbit are respectively 20 and 44 mm. The STD and the RMS of the satellite clocks are respectively 0.12 and 0.30 ns. However, the accuracy of a small number of the satellite clocks is slightly low, such as G06 and G29, the RMS is up to 0.75 ns. This may be related to the data span or the performance of the satellite clocks.

The availability of the real-time products directly determines the reliability of the time transfer method. The availability is 99.99% for the broadcast ephemeris, 99% for the IGS final, 95% for the IGU and the IGR. Figure 2 shows the distribution of data interruptions in IGC01. Figure 3 shows the average availability of each satellites.

Table 2 The statistic of IGC01 ephemeris and clock

IGC01	Statistic type	Max/PRN	Min/PRN	Avg.
Orbit (mm)	STD	27.6/G19	13.6/G02	19.9
	RMS	53.1/G19	31.3/G02	43.5
Clock (ns)	STD	0.31/G06	0.08/G27	0.12
	RMS	0.75/G06	0.14/G18	0.31

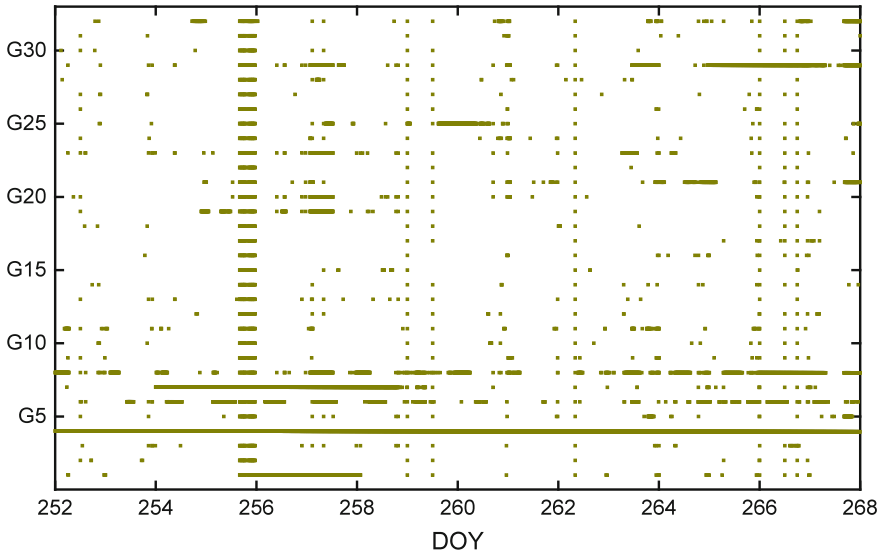


Fig. 2 The distribution of data interruptions in IGC01

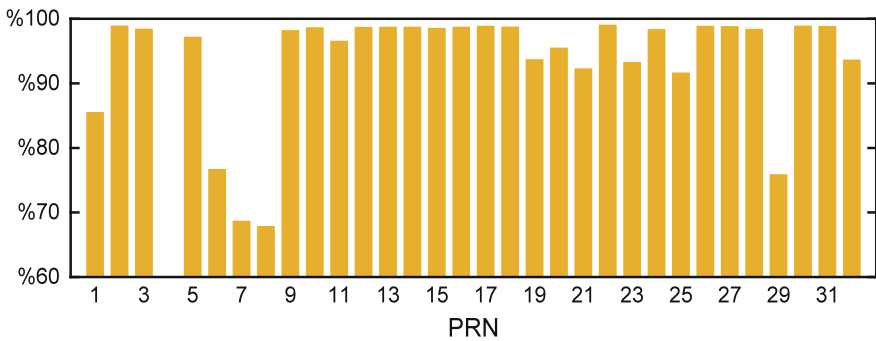


Fig. 3 The availability of IGC01

As shown in Figs. 2 and 3, the availability of the IGC01 for most of the satellites is up to 98%. Except for G04, the average availability of the other satellites is about 93.8%, which is slightly lower than the availability of about 95% obtained from the long term statistics.

3 GPS PPP Time Transfer Based on the RTS Products

3.1 GPS PPP Time Transfer

The traditional ionosphere-free model of GPS PPP time transfer is as follows [14]:

$$P_3 = \rho + cdt^r - cdt^s + T + b_{P_3}^r - b_{P_3}^s + \varepsilon_{P_3} \quad (6)$$

$$\Phi_3 = \rho + cdt^r - cdt^s + T + b_{\Phi_3}^r - b_{\Phi_3}^s + \lambda_3 N_3 + \varepsilon_{\Phi_3} \quad (7)$$

where P_3 and Φ_3 are respectively the ionosphere-free combinations of the pseudorange measurements and the carrier phase measurements. ρ is the topocentric distance between satellite and receiver. c is the speed of light in vacuum. dt^r and dt^s are the satellite and receiver clock errors. T is the slant tropospheric delay. $b_{P_3}^r$ and $b_{P_3}^s$ are the satellite and receiver code hardware delays. $b_{\Phi_3}^r$ and $b_{\Phi_3}^s$ are the frequency-dependent carrier phase hardware delays for satellite and receiver. λ_3 is the narrow-lane wavelength. N_3 is the narrow-lane ambiguity. ε_{P_3} and ε_{Φ_3} are unmodeled errors.

A reparameterized version of Eqs. (6) and (7) which eliminates linear dependencies can be written as:

$$P_3 = \rho + cdt_{P_3}^r - cdt_{P_3}^s + \varepsilon_{P_3} \quad (8)$$

$$\Phi_3 = \rho + cdt_{P_3}^r - cdt_{P_3}^s + T + \lambda_3 N_3' + \varepsilon_{\Phi_3} \quad (9)$$

where $dt_{P_3}^r = dt^r + b_{P_3}^r/c$, $dt_{P_3}^s = dt^s + b_{P_3}^s/c$, $N_3' = (-b_{P_3}^r + b_{P_3}^s + b_{\Phi_3}^r - b_{\Phi_3}^s)/\lambda_3 + N_3$.

The receiver clock can be calculated with Eqs. (8) and (9) by Kalman filtering. The time reference of the receiver clock is depending on the IGS products used. For example, the time reference of the GPS PPP time transfer results in BIPM is IGRT, while it changes to be IGST if the IGS Final products are used.

3.2 The Time Reference of the RTS Products

Compared with the IGS Final products, the RTS products have the advantages of low latency of about 30 s [16] and high sampling rates (Table 1). The GPS combined products in IGS01/IGS01 are updated once every 5 s. However, owing to the time reference difference of the RTS products produced from individual analysis centers, the time reference for the combined products will be discontinuous due to the absence of the analysis centers' RTS products, the outliers or the network interruption. The GPS PPP time transfer results will occur a jump resulting from the discontinuity of the time reference, which can be avoided when the IGS post-processed products are used. Figure 4 shows the GPS PPP time transfer results using the RTS products between PTBB and IENG from 2017-09-17 to 2017-09-19.

As shown in Fig. 4, there often occurs clock jumps when carrying out time transfer at PTBB and IENG, which is the same at each epochs. When looking at the difference between the two stations, these jumps cancel. Therefore, the time transfer

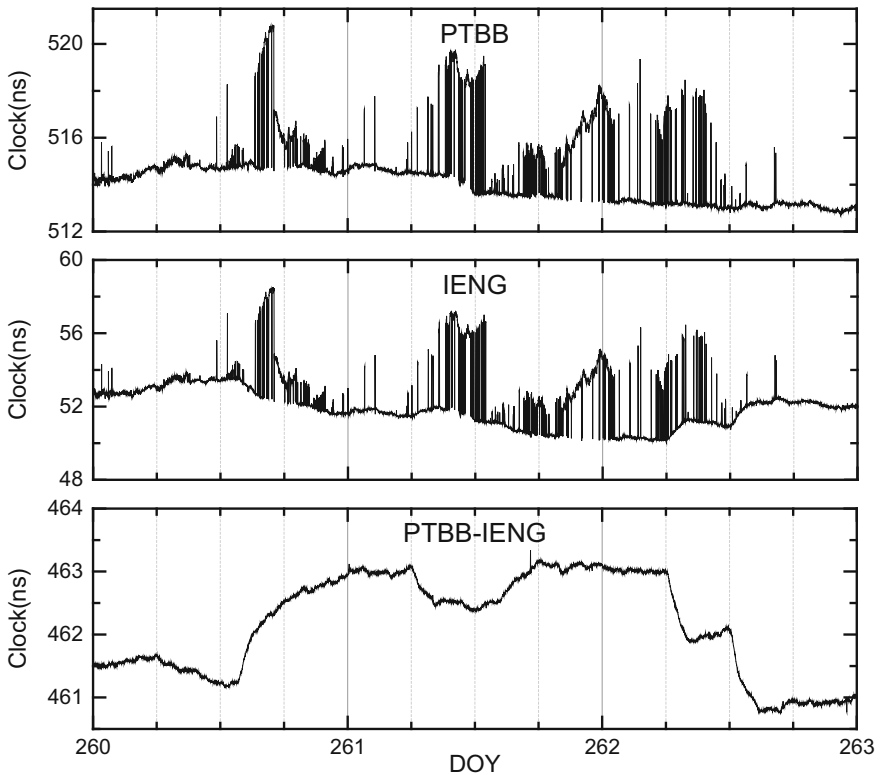


Fig. 4 PPP time transfer solutions between PTBB and IENG

results at a single station by GPS PPP will be damaged resulting from the discontinuity of the RTS products. Fortunately, this effect could be canceled for the time link [17].

3.3 The Experiments

To further analyze and confirm the practicability of the RTS products for subnanosecond level real-time time transfer, we carried out a simulated real-time time transfer experiment between PTBB and BRUX, IENG or OPMT. The observation data which has a total of 16 days from 2017-09-10 to 2017-09-25 is downloaded. The GPS PPP time transfer algorithm is implemented on the modified version of RTKLIB-2.4.3b29 [18]. The precise ephemeris and clocks products are the IGR and the IGC01. The IGC01 is obtained by the same way as Sect. 1.2.2. There are three reasons why the IGS Final products are not used. Firstly, the latency of the IGS Final products is too long, about 12–18 days. Secondly, there hardly exists any differences of the time transfer results no matter which one is used, IGR or IGS Final [9]. Thirdly, it's more comparable when using the IGR for time transfer so as to be in accordance with BIPM. Figure 5 shows the converged time transfer results (150 min later) between PTBB and BRUX, IENG or OPMT. The black curve denotes the IGR, and has been shifted by 0.5 ns. The green curve denotes the PPP (IGC01). The red curve denotes the PPP(IGR), and has been shift by -0.5 ns.

In Fig. 5, we observed that the IGR is better than the PPP(IGR) and the PPP (IGC01). The PPP(IGC01) is almost the same as the PPP(IGR). The clocks at IENG occur a jump about 2 ns. There is no time transfer results in DOY253, DOY254 or DOY260 due to the absences of the observation data at OPMT.

Next, we calculate the accuracies of the PPP time transfer results using different IGS products. Here we take the IGR station clocks as the reference owing to that the accuracy is up to 0.075 ns [7]. Figure 6 shows the clock difference between the PPP (IGC01) or the PPP(IGR) and the IGR. Table 3 summarizes the statistic of difference for the three time links.

As shown in Fig. 6 and Table 3, the time transfer accuracies for the three links are almost the same. The accuracy for the link PTBB-BRUX is the same with that for the link PTBB-OPMT, which is 0.15 ns for the RMS, and 0.12 ns for the STD. The accuracies for the link PTBB-IENG are 0.30 ns for the RMS and 0.24 ns for the STD, they are slightly lower than that for the other two links. This may be related to the jump occurred within DOY261–DOY264.

The time transfer stability is also crucial. Apart from the link PTBB-OPMT for the incompleteness of the observation data or the IGR station clocks, we only calculate the stabilities of the links PTBB-BRUX and PTBB-IENG. The overlapping Allan variance is as follows [18]:

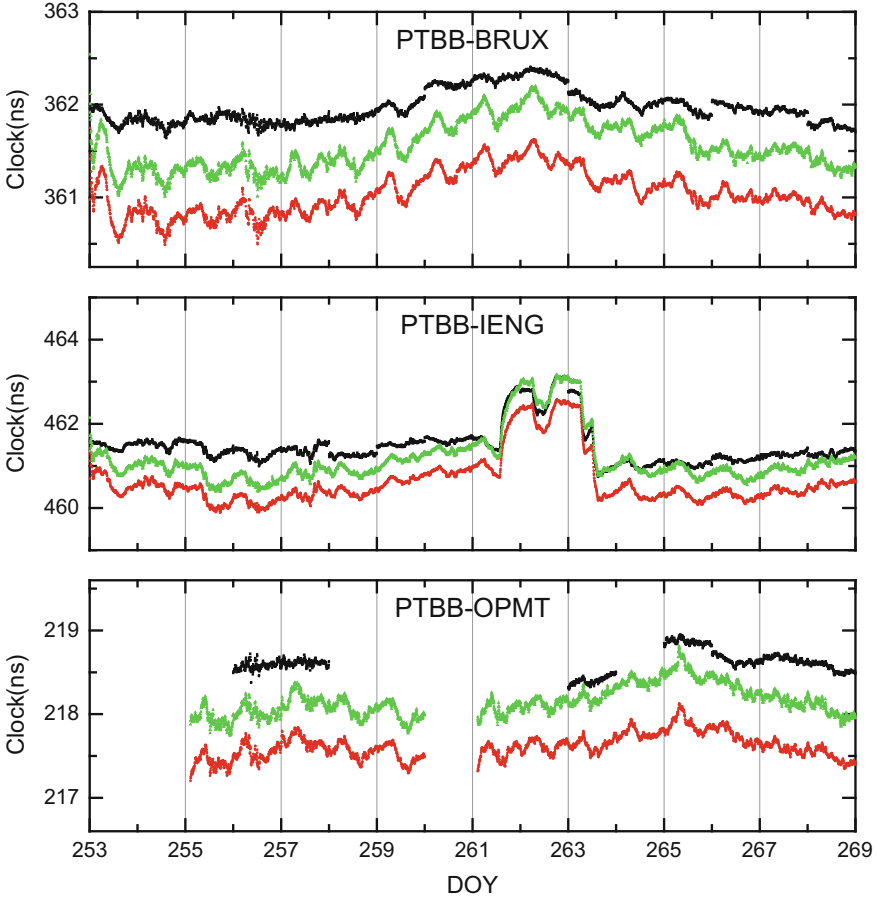


Fig. 5 Time transfer solutions obtained using IGR, PPP(IGC01) and PPP(IGR) (Black: IGR, +0.5 ns. Green: PPP(IGC01). Red: PPP(IGR), -0.5 ns)

$$\sigma_y^2(\tau) = \frac{1}{2(N-2m)\tau^2} \sum_{i=1}^{N-2m} (x_{i+2m} - 2x_{i+m} + x_i)^2 \quad (10)$$

where x_i is the clocks. N is the number of the clocks. $\tau = m\tau_0$, τ_0 is sampling interval of 5 min. Figures 7 and 8 respectively shows the time transfer stabilities of the two links using different IGS products.

As shown in Figs. 7 and 8, the time transfer stability of the IGR is much better than the PPP(IGR), which is almost the same with the PP(IGC01). The stabilities of the PPP(IGC01) for the links PTBB-BRUX and PTBB-IENG are up to $2.0E-15$ and $7.5E-15$, which not only contain the stabilities of the reference time, but also the stabilities resulting from the time transfer methods, the length of the links and

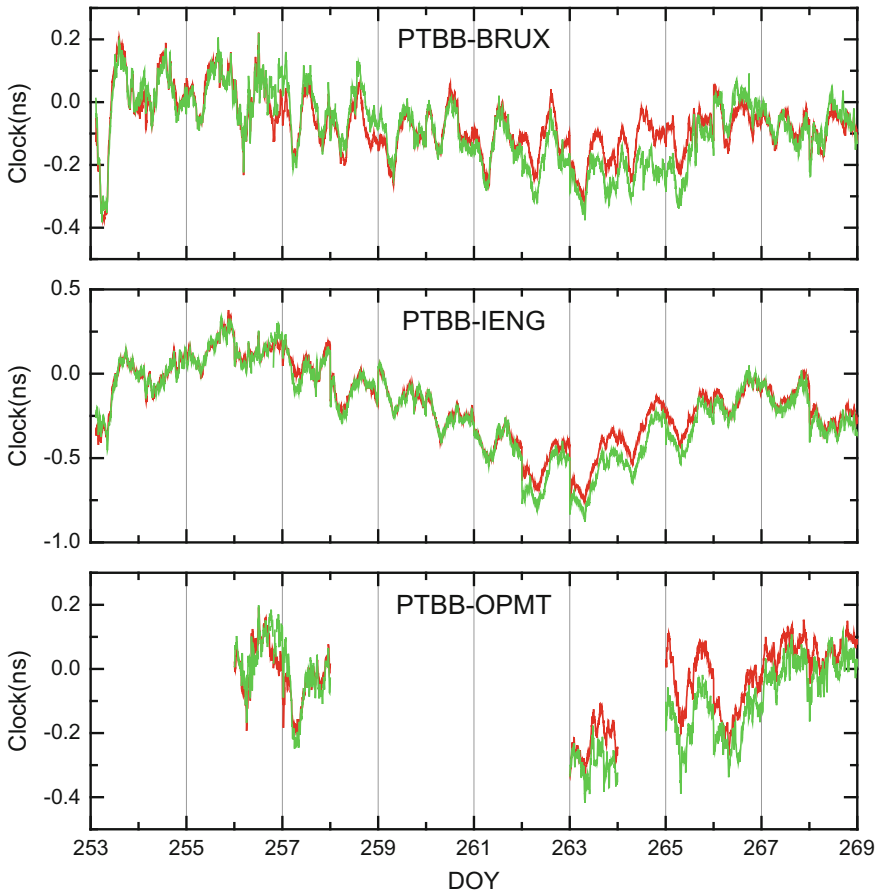


Fig. 6 Difference between the time transfer solutions obtained with PPP(IGC01) or PPP(IGR) with respect to IGR (Green: PPP(IGC01). Red: PPP(IGR))

Table 3 The statistic of difference between the time transfer solutions obtained with PPP(IGC01) or PPP(IGR) with respect to IGR (Unit: ns)

Link	IGS products	Min	Max	Avg.	RMS	STD
PTBB-BRUX	IGR	-0.38	0.22	-0.07	0.12	0.09
	IGC01	-0.38	0.22	-0.08	0.14	0.11
PTBB-IENG	IGR	-0.76	0.38	-0.17	0.28	0.22
	IGC01	-0.88	0.33	-0.21	0.32	0.25
PTBB-OPMT	IGR	-0.36	0.19	-0.04	0.12	0.11
	IGC01	-0.42	0.20	-0.09	0.16	0.13

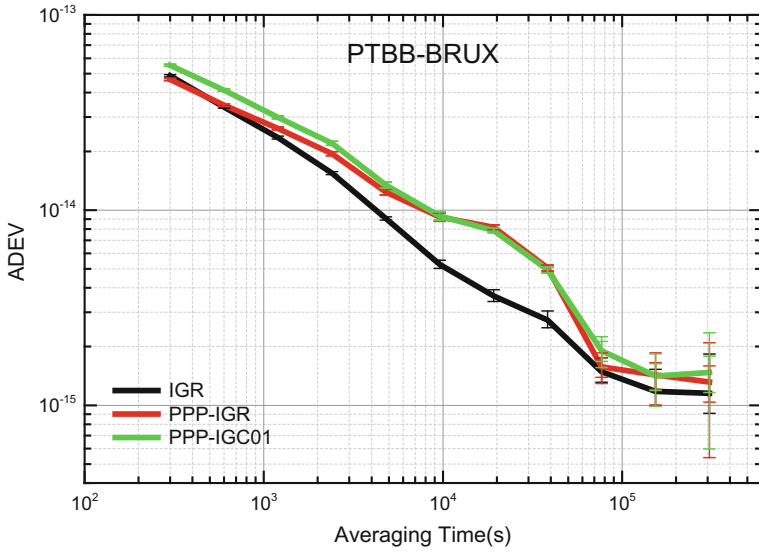


Fig. 7 The time transfer stability between PTBB and BURX

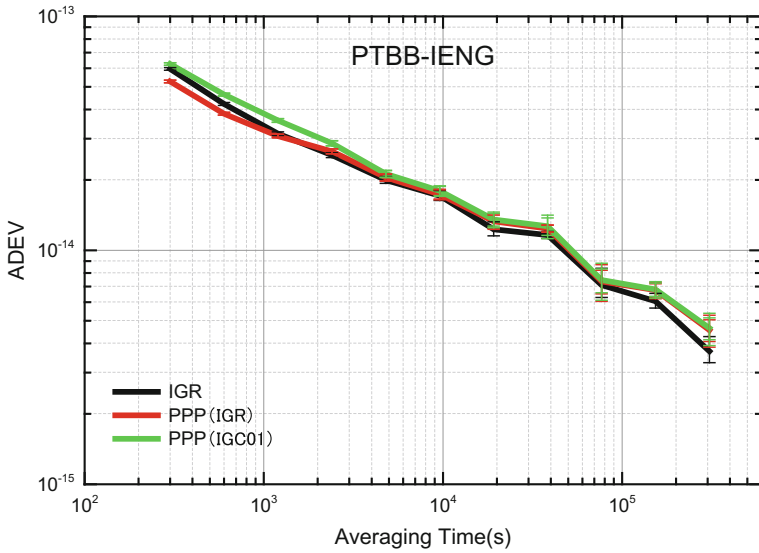


Fig. 8 The time transfer stability between PTBB and IENG

the environments. With the growing of the averaging time, the stability for the link PTBB-BRUX becomes stable, while the stability for the link PTBB-IENG is still decreasing. Combined with Figs. 5 and 6, we can conclude that the low stability for the link PTBB-IENG mainly results from the fluctuation of the reference time at IENG.

4 Conclusion

- (1) The accuracy and availability of the RTS products are analyzed. The RMS of the satellite orbit in IGC01 is 33 mm for the radial component, 32 mm for the along-track component and 26 mm for the cross-track component. The STD and the RMS of the satellite 3D-orbit is 20 and 44 mm. The STD and the RMS of the satellite 3D-orbit are respectively 20 and 44 mm. The average availability is about 93.8%.
- (2) Focus on the long latency of the GPS PPP time transfer and the low accuracy of the GPS CV time transfer, a real-time subnanosecond level time transfer algorithm using the RTS products was proposed. The time transfer experiment among four time laboratories in Western Europe shows that our algorithm is feasible. The time transfer results show that the accuracy of our algorithm can be reach to 0.30 ns for RMS and 0.25 ns for STD. Moreover, the stability of the time transfer results is up to $2E-15$ at 1 day averaging.
- (3) The robustness of the real-time time transfer using the RTS products heavily relies on the availability of the RTS products. Interruption of observation data, precise ephemeris missing or network interruption may lead to the failure of time transfer. How to judge and weigh the real-time requirement and how to improve the robustness of the algorithm will be researched next.

Acknowledgements Thanks to the real-time precise ephemeris provided by the IGS RTS analysis center of WUHAN for our experiments. This work is funded by National Natural Science Foundation of China (No. 11403112).

References

1. Petit G, Jiang Z (2008) GPS all in view time transfer for TAI computation. *Metrologia* 45 (1):35
2. Lewandowski W, Jiang Z (2009) Use of GLONASS at the BIPM. 41st annual Precise Time and Time Interval (PTTI) meeting, pp 5–13
3. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. *GPS Solutions* 5(2):12–28
4. Arias EF (2017) BIPM annual report on time activities. BIPM, Sevres Cedex
5. Ray J, Senior K (2005) Geodetic techniques for time and frequency comparisons using GPS phase and code measurements. *Metrologia* 42(4):215–232

6. Petit G, Kanj A, Loyer S et al (2015) 1×10^{-16} frequency transfer by GPS PPP with integer ambiguity resolution. *Metrologia* 52(2):1–4
7. IGS (2007) GPS satellite ephemerides/satellite and station clocks. <http://www.igs.org/products>. Accessed 12 Dec 2017
8. Xiandong C (2008) Precision time transfer methods based on geodetic time and frequency transfer receivers. *J Wuhan Univ (Inf Sci Ed)* 33(3):245–248
9. Zhang X, Cai S, Li X et al (2010) Accuracy analysis of time and frequency transfer based on precise point positioning. *J Wuhan Univ (Inf Sci Ed)* 35(3):274–278
10. Caissy M, Agrotis L (2011) Real-time working group and real-time pilot project. *Int GNSS Serv Tech Rep*, pp 183–190
11. IGS (2017) Why is IGS involved in real-time GNSS?. ftp://igs.org/pub/resource/pubs/IGS_why_in_RT.pdf. Accessed 12 Dec 2017
12. IGS (2017) IGS real-time service. <http://www.igs.org/rtts>. Accessed Dec 2017
13. Hadas T, Bosy J (2015) IGS RTS precise orbits and clocks verification and quality degradation over time. *GPS Solutions* 19(1):93–105
14. Elsobeiey M, Al-Harbi S (2016) Performance of real-time precise point positioning using IGS real-time service. *GPS Solutions* 20(3):565–571
15. BKG (2017) BKG Ntrip client. <https://igs.bkg.bund.de/ntrip/download>. Accessed 12 Sept 2017
16. IGS (2017) IGS technical report. ftp://igs.org/pub/resource/pubs/2016_techreport.pdf. Accessed Oct 2017
17. Defraigne P, Aerts W, Pottiaux E (2015) Monitoring of UTC (k)'s using PPP and IGS real-time products. *GPS Solutions* 19(1):165–172
18. RTKLIB (2017) RTKLIB: an open source program package for GNSS positioning. <http://www.rtklib.com/>. Accessed Sept 2017
19. Riley WJ (2008) Handbook of frequency stability analysis. NIST Special Publication 1065

Time Transfer Analysis of GPS- and BDS-Precise Point Positioning Based on iGMAS Products



Yulong Ge, Xuhai Yang, Weijin Qin, Hang Su, Meifang Wu,
Yuanxi Wang and Shengli Wang

Abstract International time transfer experiment based on GPS- and BDS-precise point positioning (PPP) technique is carried out using final (ISC) and rapid (ISR) precise satellite clock products provided by iGMAS (international GNSS continuous Monitoring and Assessment System). As a comparison, the same processing is conducted by using COD products provided by the Center for Orbit Determination in Europe (CODE). Experimental data come from observation of seven stations from Multi-GNSS-Experiment (MGEX) and two stations including PT11 (PTB) and BRUX (ROB) from time keeping laboratory in 70 days from 5 August to 15 October 2017. With PTB as the center node, the solutions of 8 time-links are formed. In order to verify the type A uncertainty (u_A) of time transfer based on GPS- and BDS-PPP by using iGMAS products, this paper compares the differences between GPS- and BDS-PPP results and GPS PPP results using IGR products. The experimental results demonstrate that u_A of time transfer based on GPS- and BDS-PPP using iGMAS products are roughly equal to CODE products. For the results of time-links based on GPS- and BDS-PPP, the STD of one-day arc solutions can reach better than 0.1 and 0.8 ns for all of processing, while 0.1 and 1 ns were achieved for ten-day arc solutions, respectively. Since u_A of GPS- and BDS-PPP time transfer using iGMAS products was up to sub nanosecond, there-

Y. Ge · X. Yang (✉) · W. Qin · H. Su · M. Wu · Y. Wang
National Time Service Center, Chinese Academy of Sciences, Xi'an, China
e-mail: yyang@ntsc.ac.cn

Y. Ge · X. Yang · W. Qin · H. Su · M. Wu · Y. Wang
Key Laboratory of Precise Positioning and Timing Technology, Chinese Academy of Sciences, Xi'an 710600, China

X. Yang
School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China

Y. Ge · W. Qin · H. Su · M. Wu · Y. Wang
University of Chinese Academy of Sciences, Beijing 100049, China

S. Wang
Institute of Ocean Engineering, Shandong University of Science and Technology, Qingdao 266590, China

fore, the experiments can provide a reference for BIPM to use BDS PPP based on iGMAS products for International UTC/TAI comparison. Moreover, this also illustrates iGMAS products is reliable.

Keywords Precise point positioning · Time transfer · iGMAS
GPS · BDS

1 Introduction

Currently, thanks to precise orbit and clock products [5] provided by the International GNSS Service (IGS), precise point positioning technique has been used to compute time and frequency links by using GPS pseudorange and carrier-phase observations in the time community [3, 14]. At BIPM, the technique has been operationally used to compute time links for International Atomic Time TAI since September 2009 and now concerns over 50% of the more than 70 laboratories contributing to TAI and UTC [12, 13]. Compared to GNSS code-only techniques, PPP provides much better short-term stability due to the low noise of phase measurements. Present typical uncertainty of PPP frequency comparisons is about 1×10^{-15} at 1 d averaging and about 1×10^{-16} at 30 d averaging, corresponding to the use of ‘type A’ uncertainty of 0.3 ns for PPP links in the BIPM Circular T [14].

With the development of multi-constellation GNSS, multi GNSS-based PPP time transfer has become a hot research topic. Many scholars have used BDS Common-View (CV) to compute time and frequency links based on BDS satellites. And corresponding ‘type A’ uncertainty could reach ns level. [9, 19]. P. Defraigne and G. Petit presented the Version 2E of Common GPS GLONASS Time Transfer Standard (CGGTTS), i.e. Common GNSS Generic Time Transfer Standard, the extended version includes all the constellations presently available, i.e. GPS, GLONASS, BDS, Galileo, and QZSS [4]. The standard BDS CV has been investigated by some scholars as well, and the results indicated that the uncertainty of BDS CV can achieve ns magnitude [10]. But the disadvantage of the CV model is obvious, for example, it only applies to shorter distance time-links comparison. For long distance time-links, relay stations should be selected for user in the middle of time-links, otherwise it cannot be guaranteed that those time-links can observe the same satellite [17]. Guang et al. used BDS carrier phase (BDS CP) method for time transfer by using Broadcast ephemeris. The results demonstrate that the uncertainty of BDS CP is better than BDS CV [8]. Zhang et al. indicated that the performance of time transfer based on BDS PPP and GPS PPP showed similar characteristics [16]. But above studies are not for the international UTC and TAI comparison, and the amount of data tested is also very small. In this contribution, based on precise products provided by iGMAS (international GNSS continuous Monitoring and Assessment System), with PT11 provided by PTB as the center

node, the solutions of 8 time-links are presented using GPS- and BDS-PPP technique. The dataset is obtained from 5 August to 15 October 2017.

The remaining paper is organized as follows: PPP time transfer methods are firstly presented in Sect. 2. Experiment and data processing strategies are then described in Sect. 3. In Sect. 4, the solutions of GPS- and BDS-PPP time transfer are conducted. And the solutions of GPS- and BDS-PPP time transfer using different precise products (ISC, ISR, and COD) are compared. In Sect. 5, we end with the conclusions.

2 CP Time Transfer Model

The CP time transfer algorithm uses PPP model. Dual-frequency pseudorange and carrier phase observations of Ionosphere-free positioning model are applied to mitigate the effect of first-order ionosphere. It can be expressed as follows [6, 20]:

$$P_{IF} = \frac{f_1^2 \cdot P_1 - f_2^2 \cdot P_2}{f_1^2 - f_2^2} = \rho + cdt - cdt^s + d_{trop} + \varepsilon(P_{IF}) \quad (1)$$

$$\Phi_{IF} = \frac{f_1^2 \cdot \Phi_1 - f_2^2 \cdot \Phi_2}{f_1^2 - f_2^2} = \rho + cdt - cdt^s + d_{trop} + \lambda_{IF} N_{IF} + \varepsilon(\Phi_{IF}) \quad (2)$$

where, P_{IF} and Φ_{IF} refer to ionosphere-free pseudorange and carrier-phase observations, respectively. δt and δt^s are the receiver and satellite clock offsets in reference to system time, respectively. c is the speed of light in vacuum. f_i is the signal frequency, and subscript i is the frequency number. λ_{IF} is the wavelength of carrier phase, N_{IF} is the float ambiguity. $\varepsilon(\Phi_{IF})$ and $\varepsilon(P_{IF})$ include ionospheric-free measurement noise and other measurement errors. d_{trop} denotes tropospheric delay in meters and can be written as:

$$d_{trop} = mf_h \cdot ZHD + mf_{nh} \cdot ZWD \quad (3)$$

where ZHD and ZWD denote the zenith hydrostatic and non-hydrostatic/wet delays, respectively, mf_h and mf_{nh} are the hydrostatic and non-hydrostatic mapping functions [1]. A priori ZHD is calculated through the empirical models (Saastamoinen) based on the provided meteorological information for GPT model [11] at a given location. The ZWD is estimated as an unknown parameter in the adjustment together with other parameters, such as the clock. The least squares are employed for GPS- and BDS-PPP processing.

3 Experimental Data and Processing Strategy

3.1 Dataset

Figure 1 displays the geographical distribution of 9 GNSS tracking stations, which includes 7 globally-distributed stations of the Multi-GNSS-Experiment (MGEX) network and PT11 (PTB) and BRUX (ROB) from time keeping laboratory, used in this study and also their supported GPS + BDS constellations, which can be tracked by all stations. Note that all GNSS tracking stations are equipped with atomic clocks. The summaries of all stations are listed in Table 1 in detail. The PT11 are regarded as the center node, and 8 time-links were designed in this work. Observations of a 70 days period from 5 August to 15 October 2017 are processed as well. Final (ISC), rapid (ISR) precise orbit and clock products at intervals of 15 and 5 min, respectively, which are provided by iGMAS are employed. In order to verify the feasibility of GPS- and BDS-PPP time transfer toward International UTC and TAI comparison by using iGMAS precise products, the solutions of GPS PPP using IGR products provided by IGS [7] are regarded as external reference according to the strategies of BIPM [13]. Moreover, using the same processing strategies, the solutions of GPS- and BDS-PPP time transfer using precise orbit and clock products at intervals of 5 min and 30 s provided by the Centre for Orbit Determination in Europe (CODE) are presented for verifying the reliability of

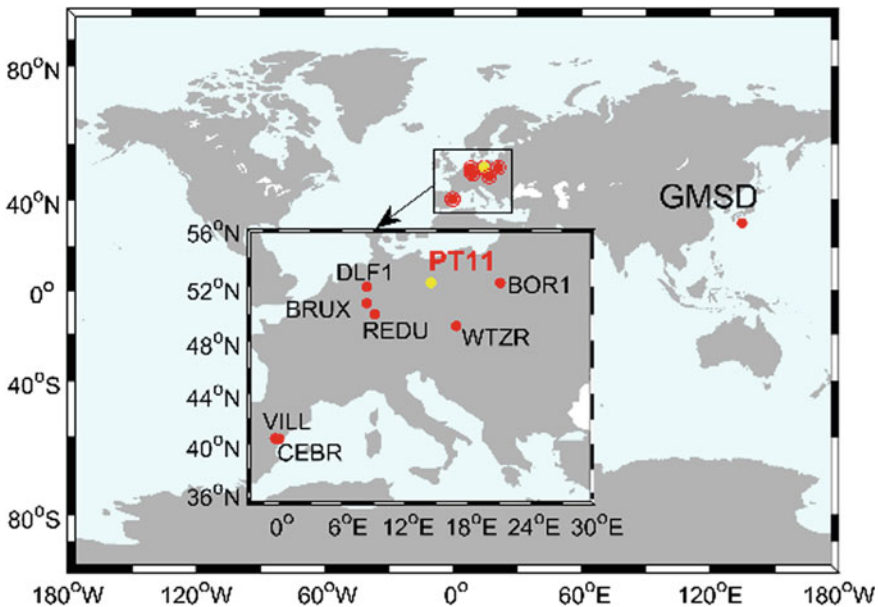


Fig. 1 Geographical distribution of GNSS tracking stations and their supported navigation satellite constellations

Table 1 Station involved in the time transfer

Station	Receiver	Antenna	Clock	The distance from PT11 (km)
PT11	SEPT POLARX4TR	LEIAR25.R4 LEIT	H-MASER	
GMSD	TRIMBLE NETR9	TRM59800.00	Cesium	8401.46
BOR1	TRIMBLE NETR9	TRM59800.00	H-MASER	450.78
BRUX	SEPT POLARX4TR	JAVRINGANT_DM	Cesium	454.88
CEBR	SEPT POLARX4	SEPCHOKE_MC	H-MASER	1730.11
DLF1	TRIMBLE NETR9	LEIAR25.R3	Cesium	417.24
REDU	SEPT POLARX4	SEPCHOKE_MC	Cesium	451.08
VILL	SEPT POLARX4	SEPCHOKE_MC	Cesium	1710.73
WTZR	LEICA GR25	LEIAR25.R3	H-MASER	389.84

iGMAS products. In order to meet requirements of data processing in this work, an improved Bernese 5.2 software is used to deal with it. And Table 2 summarizes the detailed processing strategy for GPS- and BDS-PPP.

3.2 The Methods of Uncertainty Assessment

Currently, type A uncertainty (u_A), which is the statistical uncertainty, of 0.3 ns are published for PPP links in the BIPM Circular T without the calibration. In order to evaluate type A uncertainty of GPS- and BDS-PPP time transfer using iGMAS precise products, the solutions of GPS PPP using IGR precise products are regarded as external reference values. The Standard deviation (STD) of the difference between the results using iGMAS products and IGR products are regarded as its evaluation criteria [12]. Type A uncertainty is by definition obtained by statistical methods, so the standard deviation of the clock differences is a valid technique to estimate the type A uncertainty of BDS PPP and GPS PPP.

4 Results and Analysis

In this section, the uncertainty of GPS- and BDS-PPP time transfer toward international UTC and TAI comparison using iGMAS precise products are presented. Moreover, in order to verify the reliability of iGMAS products, the solutions using iGMAS products and CODE products are compared.

Table 2 Summary of GPS- and BDS-PPP processing strategies

Items	Descriptions
First order ionospheric delay	Ionosphere-free combination
Estimator	Least squares (LSQ) estimator in batch mode
Tropospheric delay	ZHD: corrected with global pressure and temperature (GPT) model using the formulas of Saastamoinen; ZWD: estimated as a continuous piece-wise linear function (2 h parameter spacing), GMF mapping function
Satellite antenna PCO and PCV (phase center variations)	IGS ANTEX (GPS)
Observables	Undifferenced ionosphere-free combined observables from smoothed code and phase observations
Sampling rage	300 s
Signal selection	GPS: L1/L2; BDS: B1/B2
Observation weighting	$\text{Cos}(e)^2$ [2]
Elevation cut off	10°
Relativistic effect	Corrected
Sagnac effect	Corrected
Solid earth tides	IERS2010 [15]
Ocean loading	IERS2010 [15]
Polar tides	IERS2010 [15]
Phase wind-up effect	Corrected [18]
Receiver clock errors	Estimated as parameters
Station coordinates	Estimated as static
Phase ambiguities	Estimated, constant for each continuous arc; float value

We tested two different parts.

1. The results of 8 time-links were obtained from GPS PPP technique by using ISR, ISC and COD precise products, respectively. The clock difference between GPS PPP using three different products and GPS PPP using IGR products are presented.
2. The BDS PPP technique is applied to 8 time-links by using ISR, ISC and COD precise products, respectively. The clock difference between BDS PPP using three different products and GPS PPP using IGR products are displayed.

The clock differences series using GPS PPP technique on BRUX-PT11 and DLF1-PT11 time-links from part 1 are displayed in Fig. 2. The results illustrate that the clock differences between GPS PPP by using ISR, ISC, and COD products and GPS PPP by using IGR products are almost a straight line. Other time-links show same characteristics, we won't display in detail. Hence, we can conclude that the comparable uncertainty of GPS PPP time transfer based on IGR products can be achieved by using iGMAS precise products. The statistics of ten-day arc solutions (a–c) and one-day arc solutions (d–f) using GPS PPP technique on 8 time-links

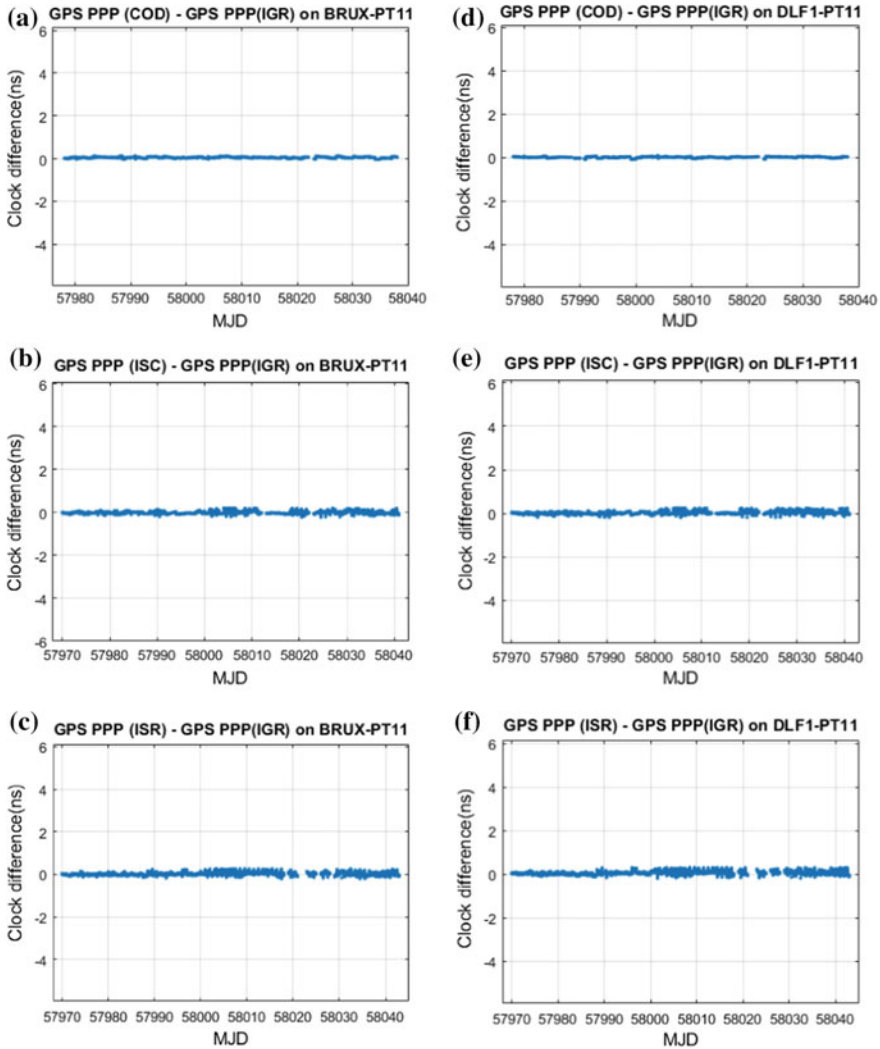


Fig. 2 The clock differences series using GPS PPP technique on BRUX-PT11 and DLF1-PT11 time-links from part 1

from part one is shown in Fig. 3. As shown in Fig. 3, the STD of ten-day arc solutions based on ISR and ISC precise products can reach 0.1 ns. Hence, we can further conclude the reliability of GPS PPP time transfer based on iGMAS products. For the results of GPS PPP using CODE products, the performance is relatively better because it is one of IGS’s analytical centres and adopts 30 s precise clock products. From another perspective, compared with the STD of ten-days arc solutions, the STD of one-day arc solutions shows better performance. Hence, the

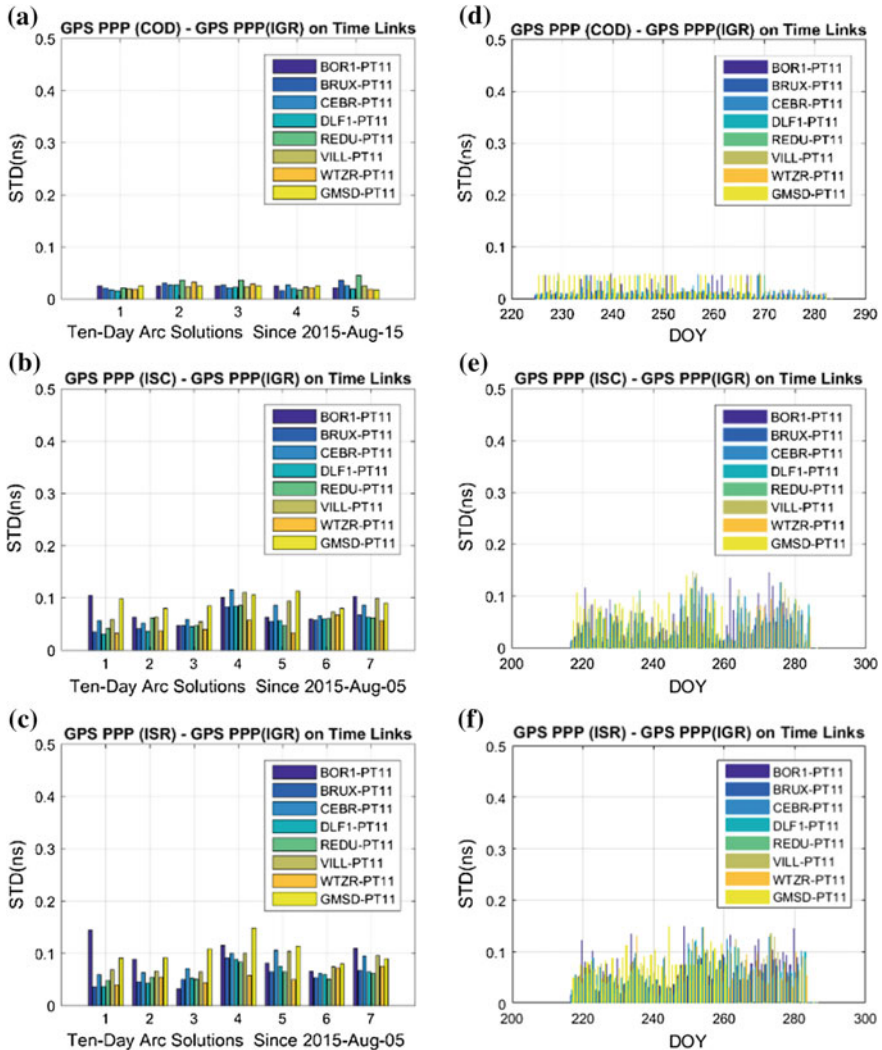


Fig. 3 The statistics of ten-day arc solutions (a–c) and one-day arc solutions (d–f) using GPS PPP technique on 8 time-links form part one

day-boundary jumps are also a problem to be solved, for that it would influence the uncertainty of time transfer.

To verify the feasibility of BDS PPP based on iGMAS products toward international UTC and TAI comparison at the current stage. Figure 4 presents the clock differences series using BDS PPP technique on BRUX-PT11 and DLF1-PT11 time-links from part 2. The results demonstrate that the clock differences between BDS PPP by using ISR, ISC, and COD products and GPS PPP by using IGR

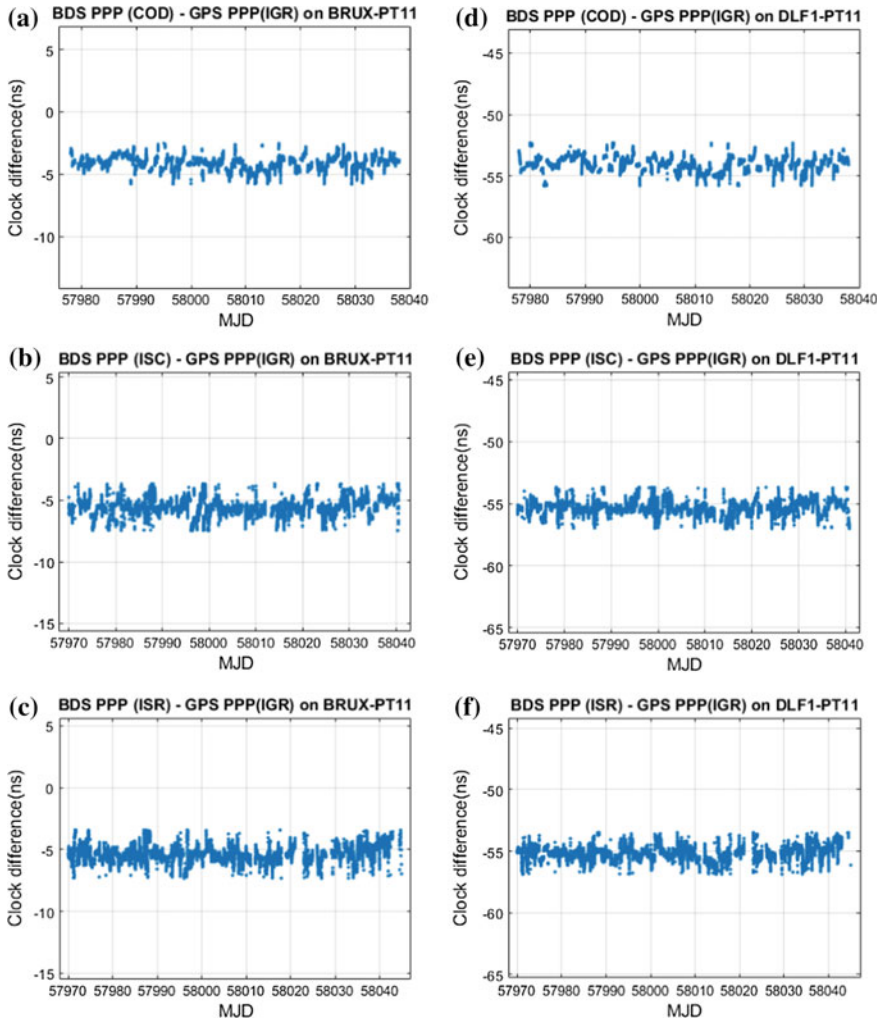


Fig. 4 The clock differences series using BDS PPP technique on BRUX-PT11 and DLF1-PT11 time-links from part 2

products have less fluctuation and better stability. Other time-links show same characteristics, thus we won't display in detail. Hence, we obtain that the performance of BDS PPP by using iGMAS precise products is better. Therefore, it also provides a good reference for international UTC and TAI comparison based on BDS PPP. The statistics of ten-days arc solutions (a–c) and one-day arc solutions (d–f) using BDS PPP technique on 8 time-links from part 2 is shown in Fig. 5. As shown in Fig. 5, the STD of ten-days arc solutions using ISR and ISC precise

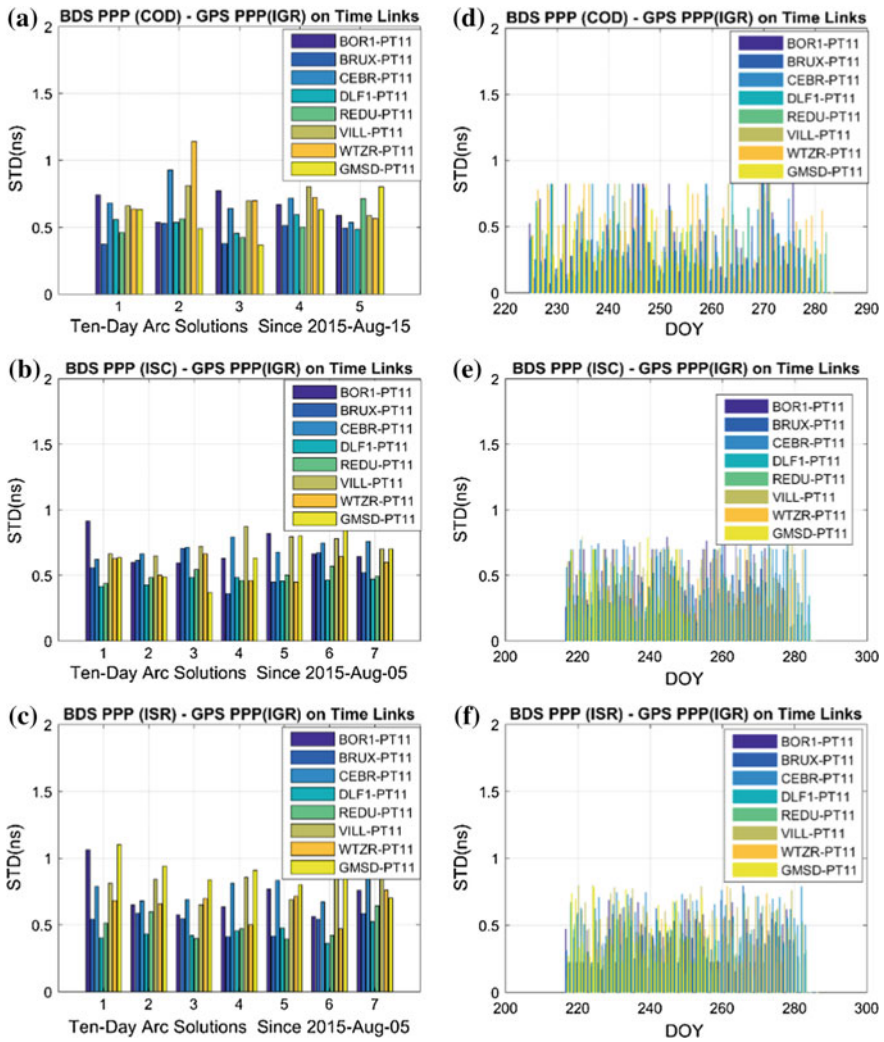


Fig. 5 The statistics of ten-days arc solutions (a–c) and one-day arc solutions (d–f) using BDS PPP technique on 8 time-links from part 2

products can reach better than 1 ns. Hence, we can further conclude the feasibility of BDS PPP time transfer based on iGMAS products. For the results of BDS PPP using CODE products, the comparable uncertainty can be reached by using iGMAS products. Compared with Fig. 5a–c, the day-boundary jumps are also a problem to be solved on BDS PPP time transfer from Fig. 5d–f.

5 Summary

In this contribution, based on final (ISC) and rapid (ISR) precise products provided by iGMAS, GPS- and BDS-PPP techniques are employed for time transfer. In order to verify the feasibility of GPS- and BDS-PPP time transfer toward International UTC and TAI comparison by using iGMAS precise products and the reliability of iGMAS products. With PT11 provided by PTB as the center node, the solutions of 8 time-links are presented using 70 days observations. The solutions of GPS PPP using IGR products are regarded as external reference values. Compared with GPS PPP using IGR products, the results show that STD of GPS PPP using ISR and ISC products can reach 0.1 ns and can remain on the same magnitude for a long time. The STD of BDS PPP can be better than 1 ns. Therefore, it can provide a good reference for international UTC and TAI comparison based on BDS PPP. Moreover, the comparable uncertainty of BDS PPP using iGMAS and CODE products are presented. Hence, we can further conclude the reliability of iGMAS products.

Acknowledgements This work is supported by National Natural Science Foundation of China (No.41104021; No.11173026; No.11703033). The authors gratefully acknowledge iGMAS for providing multi-GNSS ground tracking data and funding. We would like to thank The National R&D Infrastructure and Facility Development Program of China, “Fundamental Science Data Sharing Platform” (DKA2017-12-02-24) who funded the project. Many thanks go to the IGS MGEX for providing precise orbit and clock products MGEX, PTB and ROB (Royal Observatory of Belgium) for providing data.

References

1. Boehm J, Niell A, Tregoning P, Schuh H (2006) Global mapping function (GMF): a new empirical mapping function based on numerical weather model data. *Geophys Res Lett.* <https://doi.org/10.1029/2005gl025546>
2. Dach R, Lutz S, Walser P, Fridez P (2015) Bernese GNSS software version 5.2. <https://doi.org/10.7892/boris.72297>
3. Defraigne P, Carine B, Nicolas G (2007) PPP and phase-only GPS time and frequency transfer. In: 2007 IEEE international frequency control symposium joint with the 21st European frequency and time forum. IEEE, 2007
4. Defraigne P, Petit G (2015) CGGTTS-version 2E: an extended standard for GNSS time transfer. *Metrologia* 52(6):G1–G1. <https://doi.org/10.1088/0026-1394/52/6/G1>
5. Dow JM, Neilan RE, Gendt G (2005) The International GPS service: celebrating the 10th anniversary and looking to the next decade. *Adv Space Res* 36(3):320–326. <https://doi.org/10.1016/j.asr.2005.05.125>
6. Ge Y, Sun B, Wang S, Shen P, Liu J (2016) Convergence time analysis of multi-constellation precise point positioning based on iGMAS products. In: China satellite navigation conference (CSNC) 2016 proceedings, vol III, 390, pp 297–306
7. Griffiths J, Ray JR (2009) On the precision and accuracy of IGS orbits. *J Geodesy* 83(3–4):277–287. <https://doi.org/10.1007/s00190-008-0237-6>

8. Guang W, Zhang P, Yuan H, Li W, Zhang Q, Huang G, Zhang R (2014) The Research on carrier Phase time transfer of BeiDou navigation satellite system. In: European frequency and time forum (EFTF), 2014. IEEE, 2014
9. Guo MJ, Lu H, Xiao Y, Mo YJ, Liu XG (2015) Analysis of BeiDou signals on the accuracy of common-view time transfer. In: China satellite navigation conference (CSNC) 2015 Proceedings, vol I, 340, pp 425–436. https://doi.org/10.1007/978-3-662-46638-4_39
10. Huang W, Defraigne P (2016) BeiDou Time Transfer With the Standard CCGTTS. *IEEE Trans Ultrason Ferroelectr Freq Control* 63(7):1005–1012. <https://doi.org/10.1109/TUFFC.2016.2517818>
11. Kouba J (2009) Testing of global pressure/temperature (GPT) model and global mapping function (GMF) in GPS analyses. *J Geodesy* 83(3–4):199–208. <https://doi.org/10.1007/s00190-008-0229-6>
12. Petit G (2009) The TAIPPP pilot experiment. In: Frequency control symposium, 2009 joint with the 22nd European frequency and time forum. *IEEE international*, 1–2, pp 116–119
13. Petit G, Jiang Z (2008) Precise point positioning for TAI computation. *Int J Navig Obs* 2008:1–8. <https://doi.org/10.1155/2008/562878>
14. Petit G, Kanj A, Loyer S, Delporte J, Mercier F, Perosanz F (2015) 1×10^{-16} frequency transfer by GPS PPP with integer ambiguity resolution. *Metrologia* 52(2):301–309. <https://doi.org/10.1088/0026-1394/52/2/301>
15. Petit G, Luzum B (2010) IERS conventions (2010) Bureau international des poids et mesures Sevres (France). (No. IERS-TN-36)
16. Ruifei Z, Hui L, Bao S, Chuang Q (2017) Research on time transfer: based on BDS precise point positioning and accuracy comparison. *J Geodesy Geodyn* 37(10)
17. Sheng G, Hongqing C, Liang Z (2012) Review of the precise trace-ability of time and frequency based on Beidou/GNSS Transfer. *J Astronaut Metrol Measur* 32(1):41–44
18. Wu JT, Wu SC, Hajj GA, Bertiger WI, Lichten SM (1992) Effects of antenna orientation on GPS carrier phase. In *Astrodynamic*
19. Yuan H, Guang W (2012) Common view time transfer by a GEO navigation satellite. In: 2012 IEEE international frequency control symposium proceedings
20. Zhang X, Cai S, Li XX, Guo F (2010) Accuracy analysis of time and frequency transfer based on precise point positioning. *Geomatics and Information Science of Wuhan University* 135(3):274–278

Simulation and Experiment of Multi-section Coil for C Field Used in Passive Hydrogen Maser



Zhibing Pan, Yonghui Xie, Tao Shuai, Pengfei Chen, Yuxian Pei, Chao Shen, Xiaoyan Pan, Yang Zhao and Chuanfu Lin

Abstract Passive hydrogen maser (PHM) has the advantages of high frequency stability and low frequency drift. It has been widely used in the navigation system and frequency calibration. Atomic transition signal amplitude of PHM directly determines the system signal to noise ratio, and thus affect system performance, which makes it one of the most important parameters of PHM. Straight solenoid is widely used in PHM to produce the C field for atomic transition. Due to restriction of structural of the physical package, there is improvement potential for homogeneity of magnetic field. This paper discusses the feasibility of multi-section coil for the C field of PHM. Simulation and optimization of the parameters of the multi-section coil is carried out, including length, number of sections, turns, spacing and diameter. Design of multi-section coil with better homogeneity of the C field can be given by the simulation software. Inhomogeneity of about 2% of the C field is realized comparing to about 10% of the straight solenoid. According to the simulation results, several multi-section coils are manufactured. Application and test of these multi-section coils in physical package of hydrogen maser are carried out. Experimental results show that atomic transition signal gain can be increased by about 10%, which is beneficial to the performance of PHM.

Keywords Passive hydrogen maser · C field · Multi-section coil
Homogeneous magnetic field

Z. Pan · Y. Xie (✉) · T. Shuai · P. Chen · Y. Pei · C. Shen · X. Pan · Y. Zhao · C. Lin
Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030,
People's Republic of China
e-mail: xyh@shao.ac.cn

Z. Pan
University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018
Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_42

1 Introduction

Passive hydrogen maser (PHM) relies on hyperfine transition signal of hydrogen atoms of ground state for time keeping. It has advantages of high frequency stability and low frequency drift [1], and has been widely used in navigation system and frequency calibration. Performance improving is an important direction of PHM development. Amplitude of atomic transition signal of PHM directly determines system Signal to Noise Ratio (SNR), and thus affects output performance, and it is one of the most important parameters of PHM.

Hydrogen atoms (0–0) transition ($F = 1, mF = 0$ to $F = 0, mF = 0$) in PHM is stimulated by microwave field in the resonant cavity, and effected by axial static magnetic field (C field). The coupling between the axial magnetic induction of the microwave field and the C field has directly effect on the hydrogen atom (0–0) transition. If the C field in the storage bulb area is not homogeneous, some of the hydrogen atoms will experience a deviating magnetic field, and take transition with deviating frequency. This would reduce amplitude of the (0–0) transition signal, and the SNR and gain of spectral line would deteriorate [2]. Therefore, homogeneity of C field is an important factor for quality of transition.

Through calculation and computer simulation, we have designed a multi-section coil with better magnetic field homogeneity. It can be used for C field instead of conventional straight solenoid. Experimental results on PHM show it can effectively enhance amplitude of the hydrogen atom (0–0) transition signal, and improve system performance of PHM.

2 Calculation and Simulation of Homogeneity of Straight Solenoid

Equation (1) can be used to describe magnetic induction in the axis of straight solenoid [3]:

$$B = \frac{\mu_0 n I}{2} \left[\frac{l_1}{(l_1^2 + R^2)^{1/2}} + \frac{l_2}{(l_2^2 + R^2)^{1/2}} \right] \quad (1)$$

μ_0 is vacuum permeability, n is turn per unit length, I is current through the straight solenoid, R is radius of the straight solenoid, and l_1, l_2 is distance from the point to two endpoints of straight solenoid.

If length of straight solenoid is much longer than its radius ($l \gg R$), magnetic induction is about $B \approx \mu_0 n I$ in its central region. In other words, magnetic field is almost homogeneous in central region of the straight solenoid. However, Magnetic induction can be reduced to $B \approx \mu_0 n I / 2$ near each endpoint of straight solenoid,

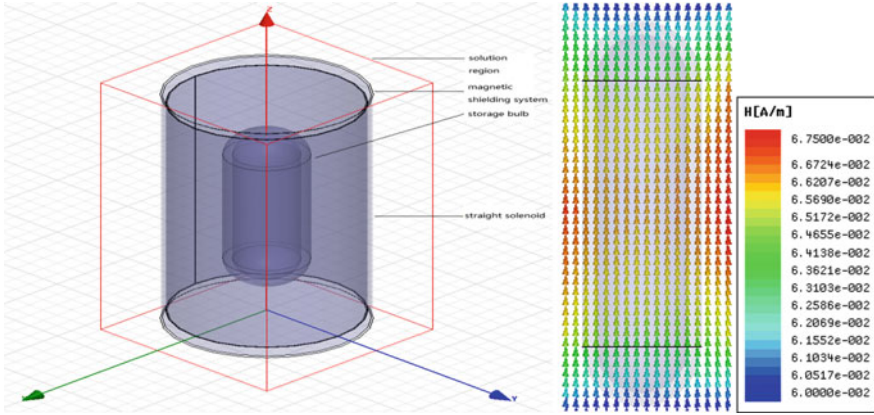


Fig. 1 Simulation model and simulation results of straight solenoid in the area of storage bulb

only half of the central region. Actually end surface of the storage bulb are close to the end surface of the C field in practical application of PHM. Homogeneity of C field in the area of the storage bulb is inevitably limited.

According to actual size of the straight solenoid, we can simulate static magnetic field generated by straight solenoids in the area of the storage bulb by using electromagnetic field simulation software. Simulation model and results are shown separately in Fig. 1.

As in Fig. 1, the maximum of magnetic field strength in the area of storage bulb is $6.75E-2$ A/m, and the minimum is $6.00E-2$ A/m. Inhomogeneity is about 11.11%. This shows that homogeneity of the magnetic field strength of the straight solenoid has great potential of improvement.

3 Calculation and Simulation of Homogeneity of Multi-section Coil

Helmholtz Coil has two identical coils which are coaxially placed. The distance between two centers of Helmholtz coil is equal to radius of the coils. If the two coils have currents with same amplitude and same direction, homogeneous magnetic field could be established in its central area [4].

If the distance between two centers of Helmholtz coil is d , the radius is R , setting the middle point of the line connecting centres of the two coils as coordinate origin, magnetic induction at the central axis is [5]:

$$B(x) = \frac{\mu_0 I R^2}{2 \left[R^2 + \left(\frac{d}{2} + x \right)^2 \right]^{\frac{3}{2}}} + \frac{\mu_0 I R^2}{2 \left[R^2 + \left(\frac{d}{2} - x \right)^2 \right]^{\frac{3}{2}}} \quad (2)$$

When $d = R$, magnetic induction in the region between the centres of the two coils is homogeneous [6].

Two-section Helmholtz coil can produce a homogeneous magnetic field in its central region. However, the axial length of the storage bulb is about 100 mm. If a two-section Helmholtz coil is used to generate C field containing storage bulb, radius of coils is about 100 mm, which result in unacceptable volume for PHM. Multi-section coil is an assembly of multiple coils coaxially with same current. Parameters of multi-section coil include radius, total length, length of each section, number of sections, spacing between sections, turns per unit length and so on. Considering all above parameters and adjusting them correspondingly, we can design a multi-section coil with high homogeneity of C field throughout the storage bulb area.

We have carried out computer simulation of seven-section coil. The model is shown in Fig. 2. Structures including the holder of C field, storage bulb and magnetic shielding system are taken into account. Input current of seven-section coil is 0.02 mA. After multiple adjustment and simulation, a set of optimal coil parameters are obtained, and the simulation result is also shown in Fig. 2.

The simulation results show that the maximum of magnetic field strength in area of storage bulb is $4.66E-2$ A/m and the minimum value is $4.59E-2$ A/m. Inhomogeneity between the two values is 1.51%. Comparing to the straight solenoid, magnetic field homogeneity has been greatly improved.

Homogeneity of the magnetic field strength generated by the seven-section coil and the straight solenoid in the storage bulb area are shown in the following Fig. 3.

In addition to the seven-section coil, we have also carried out calculation and simulation of three-section coil and five-section coil. Simulation model and results

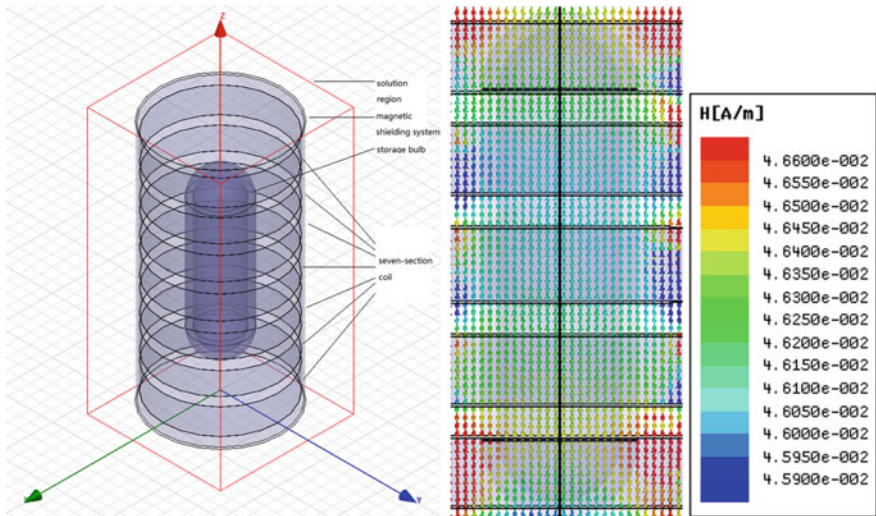


Fig. 2 Simulation model and simulation results of seven-section coil in the area of storage bulb

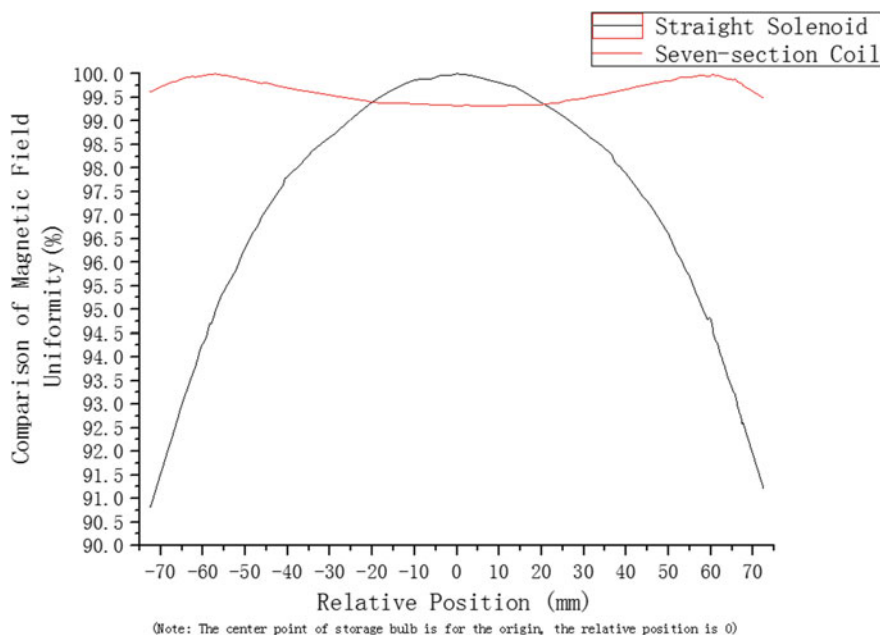


Fig. 3 Comparison of magnetic field homogeneity between seven-section coil and straight solenoid in storage bulb region

are shown separately in Figs. 4 and 5. Results show that with increase of number of sections it would be easier to achieve better magnetic field homogeneity in the region of storage bulb.

The magnetic field homogeneity of various types of coil in the storage bulb area is shown in Table 1.

4 Experiment and Application of Multi-section Coil

According to the design parameters of the seven-section coil given by simulation software, an actual C-field coil is manufactured and applied to PHM. In the cases of keeping the other experimental parameters to be same, signal amplitude of hydrogen atom (0–0) transition is measured. Measurement of the straight solenoid is also carried out to make comparison.

During the experiment, magnetic shielding system was degaussed. Microwave input power of resonator is -90 dBm. Microwave resonant frequency of cavity is 1420.405751 MHz \pm 5 kHz. Hydrogen atom (0–0) transition signal gain is 2.49 dB when straight solenoid coil is used as c field, and hydrogen atom (0–0)

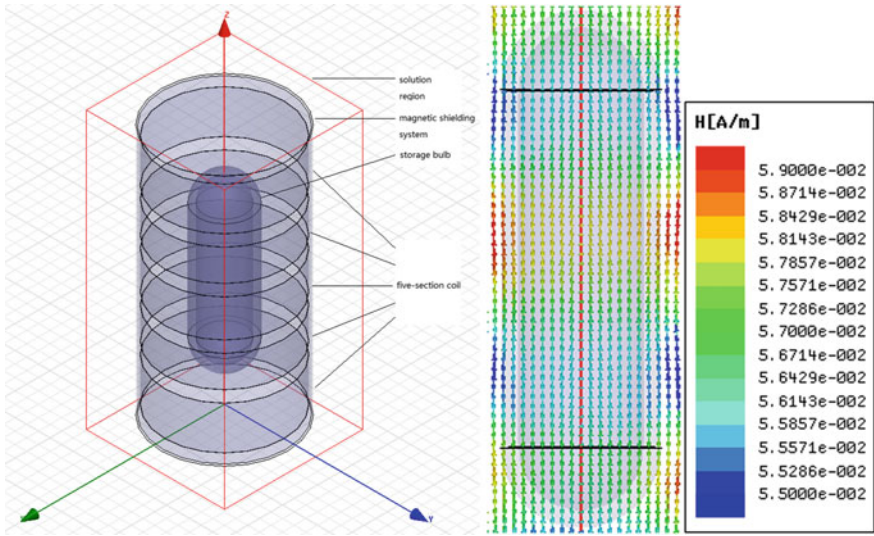


Fig. 4 Simulation model and simulation results of five-section coil in the area of storage bulb

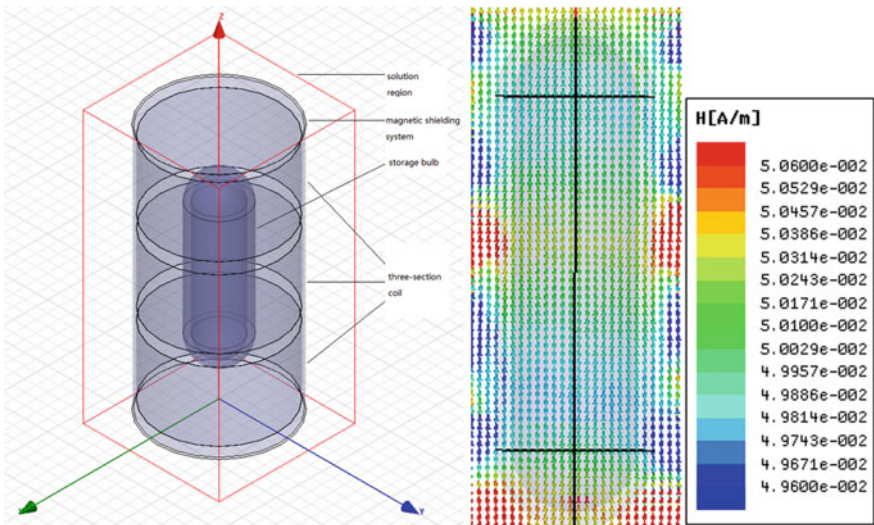


Fig. 5 Simulation model and simulation results of three-section coil in the area of storage bulb

transition signal gain is 2.81 dB when seven-section coil is used as c field. The test results given by vector network analyzer are obtained. An improvement of 0.32 dB or 12.8% is achieved with seven-section coil. Their comparison is shown in Fig. 6.

Table 1 The homogeneity of magnetic field of various types of coil in the storage bulb area

Type of C field	The difference between the maximum and minimum values in the storage bulb area	The maximum values in the storage bulb area	The difference between the maximum and minimum values in the storage bulb area/the maximum values in the storage bulb area (%)
Straight solenoid	7.50E-03	6.75E-02	11.1
Three-section coil	4.00E-03	5.90E-02	6.78
Five-section coil	1.00E-03	5.06E-02	1.98
Seven-section coil	0.70E-03	4.66E-02	1.51

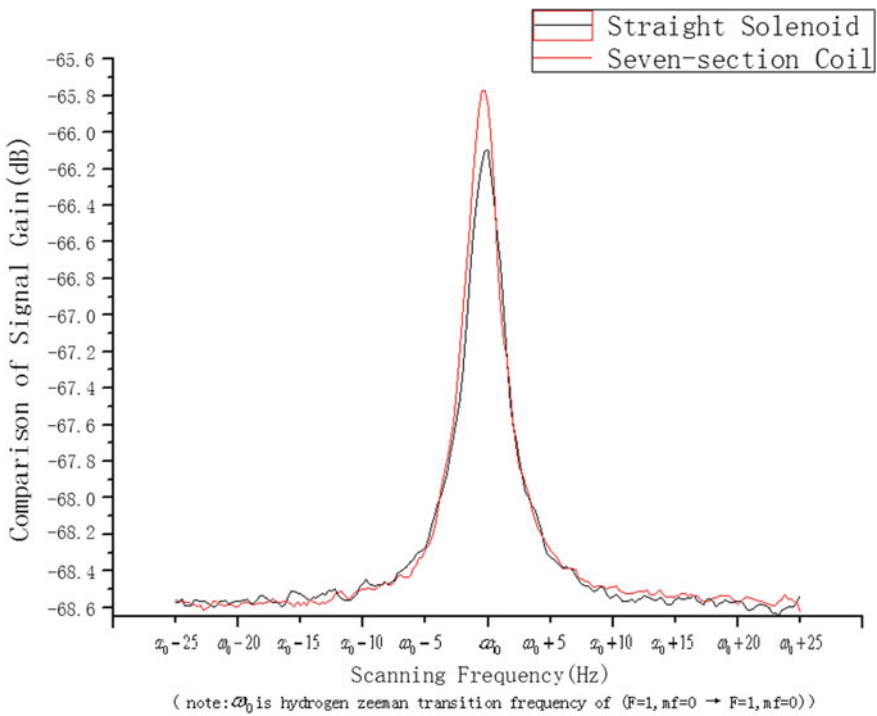


Fig. 6 Comparison of signal gain of hydrogen atom (0-0) transition of straight solenoid and seven-section coil

5 Conclusion

Homogeneity of C field is an important factor of spectral quality of hydrogen (0–0) transitions. Calculation and software simulation results show that Inhomogeneity of the C field generated by straight solenoid is about 10%. However, less than 2% of inhomogeneity of the C field can be realized by seven-section coil. Experimental application shows that the gain of the hydrogen atom (0–0) transition signal can be increased by about 10% with application of seven-section coil.

Application of multi-section coil is benefit to performance of PHM without increasing its volume and weight. We will carry out closed-loop testing with electronics package to evaluate its advantage on system performance such as short-term and long-term frequency stability and frequency drift. At the same time, we will optimize the parameters such as the number of sections and spacing between sections to explore parameters with better magnetic field homogeneity.

References

1. Wang Y (1986) The principle of quantum frequency standards. China Science Publishing
2. Vanier J, Audoin C (1989) The quantum physics of atomic frequency standards. Adam Hilger, Bristol and Philadelphia
3. Xu Y (2008) Electromagnetism. China Science Publishing
4. Song X (2016) Comparison of magnetic field distribution and homogeneity between Helmholtz coil and Maxwell coil. *J Magn Mater Devices* 47(5):16–18, 77
5. Ye B (2014) Electromagnetism. University of Science and Technology of China Press
6. Liu J (1997) Theoretical calculation of magnetic equal region inside Helmholtz coil. *J Hebei Univ* 17(2):87–89

Periodicity Systematic Error Analysis of GPS in Nonlinear Variations of Station Coordinates



Yanbo Fu, Fuping Sun, Xinhui Zhu, Kai Xiao, Hailiang Dai and Longlong Zhang

Abstract Analysing the mechanism of nonlinear variations in GPS stations coordinates time series is an effective way to improve the coordinates accuracy of GPS stations. The constellation design characteristic of GPS may cause the periodic variations of GPS stations coordinates and affect the precision of GPS stations coordinates. In order to analyze the systematic error of the GPS, the spectrum analysis was applied to detect the periodic terms of the GPS and VLBI vertical coordinates time series in the selected 6 amalgamated stations, the amplitude and initial phase of periodic terms were extracted. The annual terms were commonly found in the coordinates time series of GPS and VLBI, the annual terms' changing trend of GPS and VLBI were basically same, and the systematic error was commonly found between GPS and VLBI. Meanwhile, it was also found that the systematic error of the semi-annual terms exist between the two technologies in some amalgamated stations, which may be related to the systematic error of GPS itself and the false periodic terms caused by the data's quality problems.

Keywords Amalgamated station · GPS · VLBI · Periodic system error

1 Introduction

In order to establish and maintain the global millimeter—level terrestrial reference frame, eliminating the nonlinear variation in the coordinates time series of GPS stations is one of the important means. GPS (Global Positioning System), VLBI (very long baseline interferometry), SLR (Satellite Laser Ranging).

GPS (Global Positioning System), VLBI (very long baseline interferometry), SLR (Satellite Laser Ranging) and other spatial geodetic techniques are used to model and analyze the various periodic terms which are universally found in the coordinates time series of crustal surface stations.

Y. Fu (✉) · F. Sun · X. Zhu · K. Xiao · H. Dai · L. Zhang
Information Engineering University, Science Avenue 62, Zhengzhou, Henan, China
e-mail: fybxrlb1993@163.com

Due to the different climate and geophysical tectonic environment of each GPS station in the world, there are differences in the periods and causes of all kinds of periodic terms.

According to existing studies, there is a strong correlation between seasonal temperature changes and the GPS annual changes of coordinates time series [1].

Seasonal temperature changes can explain 11.2% of the vertical amplitude [2], Surface quality correction can explain 40% of the vertical nonlinear changes [3], Environmental loading correction can weaken 70% of the nonlinear changes in the vertical and eastern of most of the GPS stations in China [4], High order ionospheric change is one of the factors of the annual and semi-annual terms of the GPS stations [5]. The above three factors are the main geophysical factors that lead to the nonlinear variation of the coordinates time series of the GPS stations. literature [6] and literature [16] point out that besides geophysical factors, the annual effect of the GPS itself may also be a factor that leads to the annual variations in the coordinates time series of the GPS stations, literature [7] considers that the GPS periodic systematic errors can be separated by the amalgamated stations, literature [8] and literature [9] preliminarily validate that the GPS annual systematic errors exist in the annual terms of GPS vertical coordinates time series.

In this paper, the periodic systematic errors in the vertical coordinates time series of GPS stations are analyzed by using the amalgamated stations of GPS and VLBI through the contrast experiment.

2 The Adopted Data and Analysis Methods

2.1 The Adopted Data

The data used in this article is the GPS and VLBI coordinates residual time series come from ITRF2008 (The International Terrestrial Reference Frame) website, the time span of the selected GPS data and VLBI data is from 2000 to 2009, and the data sampling interval is 7 days. In order to make the separated GPS systematic error more accurate, 6 amalgamated stations with good data quality are selected, as shown in Table 1.

Table 1 Information of collocated stations

Names of amalgamated stations	Coordinates of amalgamated stations	Names of GPS stations	Names of VLBI stations
Noto	14.99°E 36.87°N	NOTO	7547
Urumqi	87.60°E 43.81°N	URUM	7330
Algonquin	78.07°W 45.96°N	ALGO	7282
WETZELL	12.88°E 49.14°N	WTZR	7224
Matera	16.70°E 40.65°N	MATE	7243
Medicina	4.52°E 11.65°N	MEDI	7230

2.2 Data Analysis Method

On the basis of eliminating the long-term trend term in the time series, the three spline interpolation method is used to interpolate the discontinuous points in the original data, then GPS and VLBI coordinates residual time series are analyzed by the method of spectrum analysis, and screen out the obvious periodic terms in each time series by setting the threshold, as shown in Tables 2 and 3.

As we can see from Tables 2 and 3, there are various periodic terms in the coordinates time series of GPS and VLBI stations, the annual and semi-annual terms are universally exist in the time series of each station. Therefore, we can compare the annual and semi-annual terms of GPS and VLBI time series in the same amalgamated station to verify whether there is periodic systematic errors in GPS. In order to extract the amplitude and initial phase of the periodic terms in the time series, the harmonic analysis is used to fit time series:

$$X_t = \sum_{i=1}^N A \cos(\omega_i t + \varphi_i) \tag{1}$$

Expand the formula with coefficient:

$$X_t = \sum_{i=1}^k (a_i \cos 2\pi f_i t + b_i \sin 2\pi f_i t) \tag{2}$$

Table 2 The periodic terms of the vertical residual time series of GPS stations

Names of GPS stations	Periodic terms (spectrum energy from large to small)/d			
NOTO	364	150	270	675
URUM	350	601	168	210
ALGO	362	774	185	1162
WTZR	352	169	538	915
MATE	352	169	570	1460
MEDI	361	521	173	521

Table 3 The periodic terms of the vertical residual time series of VLBI stations

Names of VLBI stations	Periodic terms (spectrum energy from large to small)/d			
7547	355	177	568	86
7330	385	177	577	105
7282	710	368	191	903
7224	370	560	189	764
7243	779	364	181	571
7230	348	185	726	562

The amplitude and phase of each periodic term are extracted by using the least square fitting method:

$$A_i = \sqrt{\left(\frac{2}{N} \sum_{t=1}^N X_t \cos 2\pi f_i t\right)^2 + \left(\frac{2}{N} \sum_{t=1}^N X_t \sin 2\pi f_i t\right)^2} \tag{3}$$

$$\varphi_i = \arctan\left(-\frac{\sum_{t=1}^N X_t \sin 2\pi f_i t}{\sum_{t=1}^N X_t \cos 2\pi f_i t}\right) \tag{4}$$

3 Analysis of Calculation Results

The annual and semi-annual terms in each amalgamated station are separated by using the above methods, the annual and semi-annual terms of GPS and VLBI coordinates time series of each amalgamated station are shown in Figs. 1, 2, 3, 4, 5 and 6 (The accuracy is $\pm 0.2\text{mm}$).

It can be seen from Figs. 1, 2, 3, 4, 5, 6 and Table 4, ① the maximal difference of annual and semi-annual terms' initial phase between GPS and VLBI in a same amalgamated station is 40° (The semi-annual term of the Matera amalgamated station), the minimal difference is 0.39° (The semi-annual term of Noto

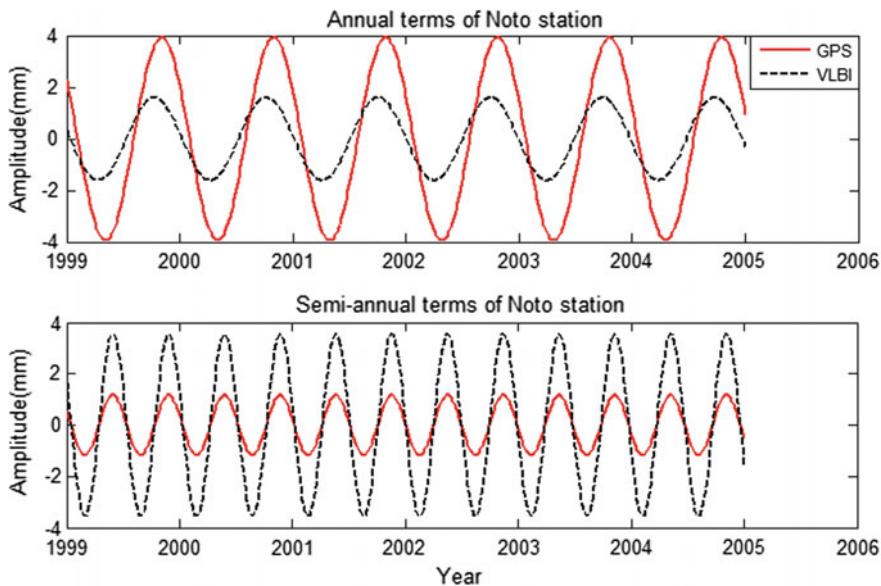


Fig. 1 The contrast between the periodic terms of the Noto collocated station

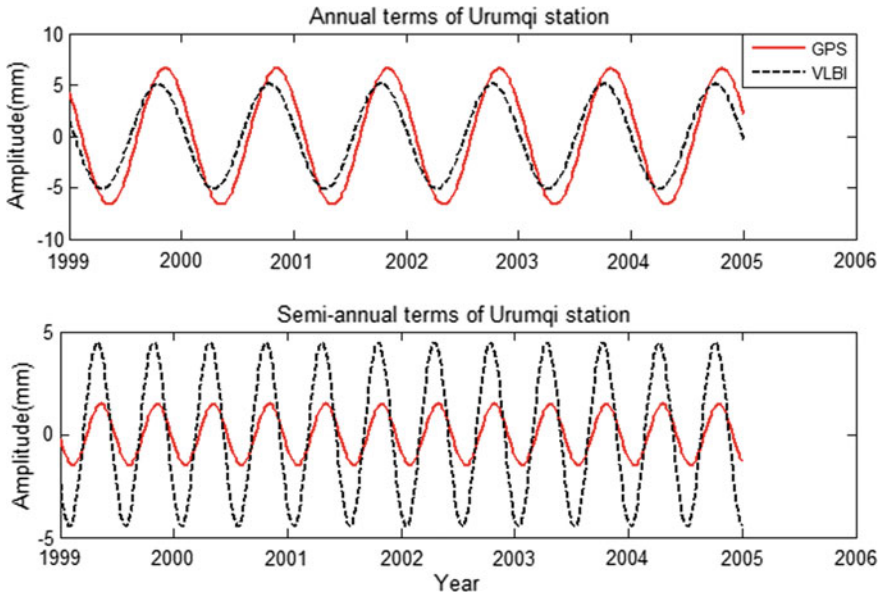


Fig. 2 The contrast between the periodic terms of the Urumqi collocated station

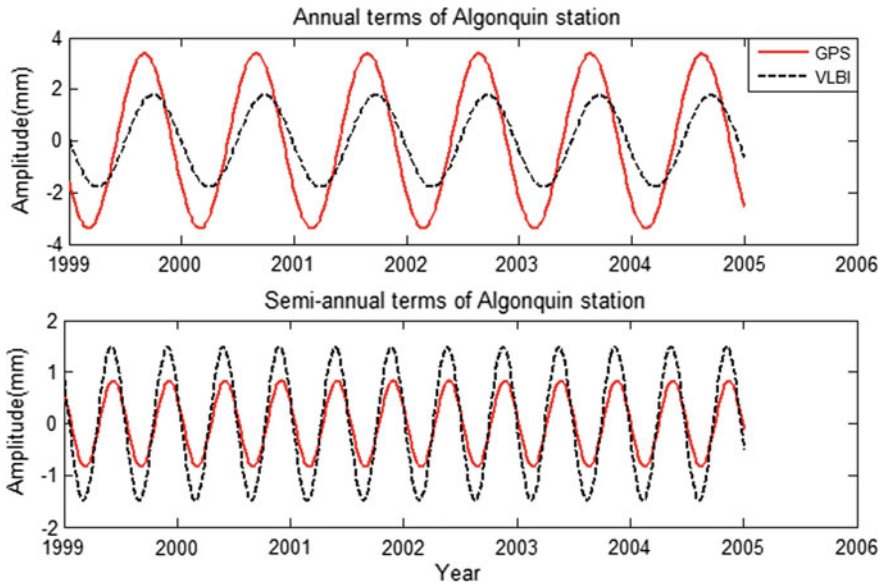


Fig. 3 The contrast between the periodic terms of the Algonquin collocated station

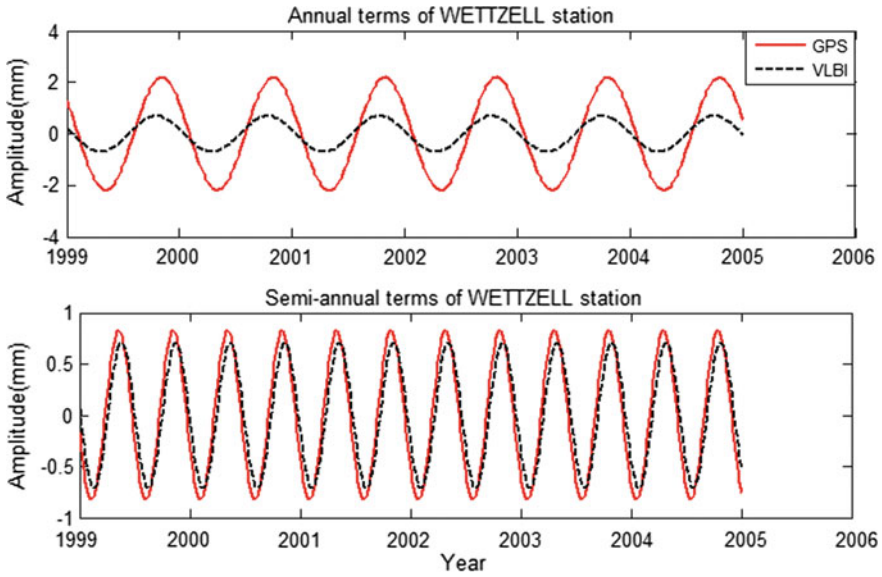


Fig. 4 The contrast between the periodic terms of the WETTZELL collocated station

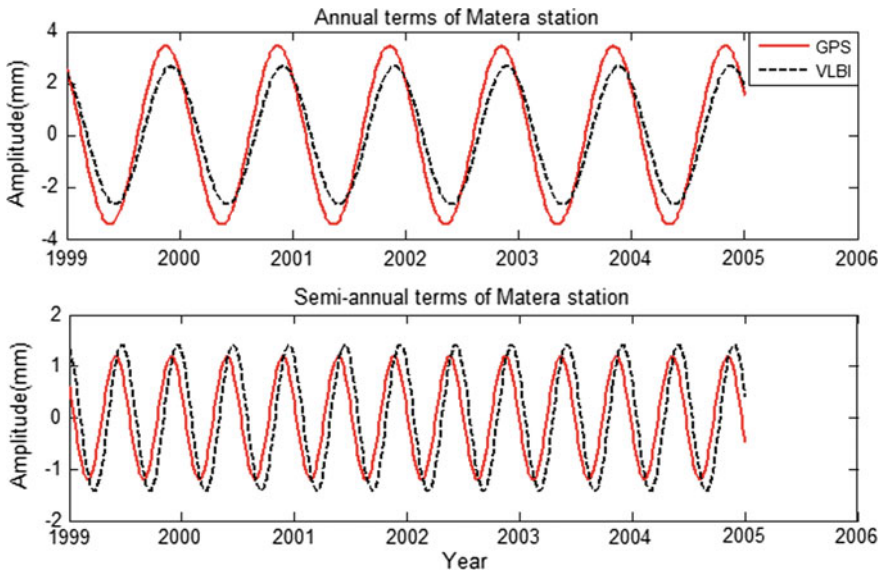


Fig. 5 The contrast between the periodic terms of the Matera collocated station

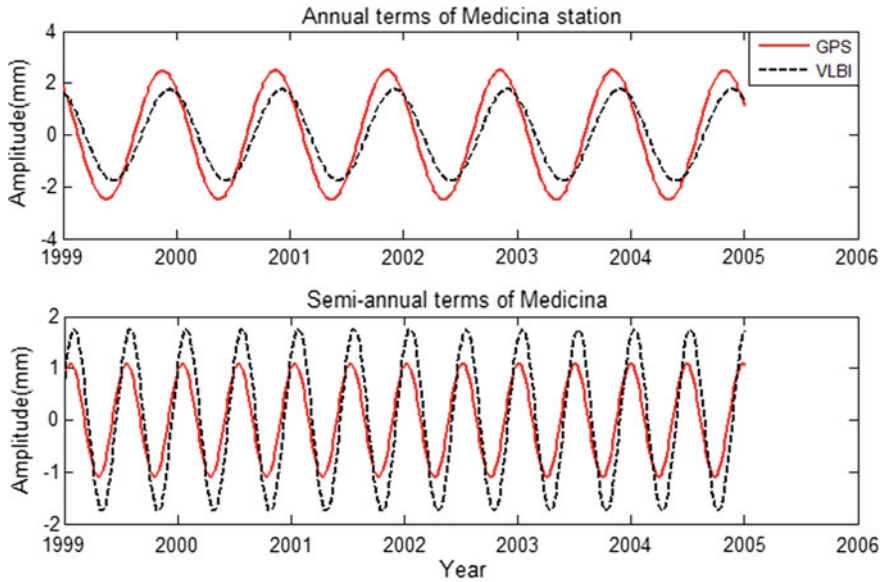


Fig. 6 The contrast between the periodic terms of the Medicina collocated station

amalgamated station), the difference of initial phase is universally small, it is shown that the monitoring of the annual and semi-annual changing laws of the surface deformation by the GPS and VLBI in the same amalgamated station is basically consistent, it is further proved that the GPS and VLBI have good spatiotemporal consistency in the coordinates of amalgamated stations, which is discussed in literature [10]. ② The amplitudes of annual terms of GPS coordinates time series are all larger than the amplitude of VLBI. Except the WETTZELL amalgamated station, the amplitudes of semi-annual terms of GPS are all smaller than the amplitudes of VLBI. The maximal difference of amplitudes of annual terms between GPS and VLBI is 2.33 mm (Noto station), the minimal difference of amplitude is 0.73 mm (Medicina station), the maximal difference of amplitudes of semi-annual terms between GPS and VLBI is 3.01 mm (Urumqi station), the minimal difference of amplitude is 0.12 mm (WETTZELL station). ③ The difference of annual amplitudes of amalgamated stations with similar latitudes is small, the region that Urumqi station is located may be due to the severe local geological structure changes, so the amplitude is obviously larger than the stations located in near latitudes.

Compared with VLBI, systematic errors of annual and semi-annual terms obviously present. Because the geological conditions and meteorological conditions of GPS and VLBI stations in a amalgamated station are basically the same, VLBI measures the propagation delay of two signals from extragalactic radio sources between two receiving antennas to accurately determines the of the signal from the extragalactic radio source between two receiving antennas using radio telescope propagation delay to accurately determine the relative position between the ground

Table 4 The comparison between the GPS's and the VLBI's amplitude and phase of the annual terms and the semi-annual terms

Names of amalgamate stations	GPS			VLBI								
	Annual terms			Semi-annual terms			Annual terms			Semi-annual terms		
	Amplitude (mm)	Phase (°)		Amplitude (mm)	Phase (°)		Amplitude (mm)	Phase (°)		Amplitude (mm)	Phase (°)	
Noto	3.94	48.16		1.16	48.55		1.61	72.23		3.57	48.16	
Urumqi	6.62	42.52		1.49	85.50		5.09	65.47		4.5	108.42	
Algonquin	3.39	110.23		0.84	29.54		1.79	81.58		1.49	40.99	
WETTZELL	2.19	47.49		0.83	87.62		0.69	64.68		0.71	70.44	
Matera	3.46	34.98		1.20	44.98		2.67	17.79		1.42	4.87	
Medicina	2.49	34.97		1.09	305.82		1.76	12.05		1.76	282.90	

Notes The reference epoch for calculating the initial phase is 1999.0

points [11], So the observation object and the observation equipment do not have the possibility to cause the annual systematic errors, therefore, there are two aspects of the reasons of the annual systematic errors: ① The systematic error exists in GPS itself, ② There is a quality problem of the observed data used. And GPS has now been found three kinds of annual effects[12, 5]: (1) The annual effect caused by the solar radiation perturbation to the satellite sails, (2) The annual effect of the configuration between stations and constellation, the same station is 4 min earlier every day to return to the same geometric configuration between the station and the GPS constellation the day before, (3) There is uncertainty in the phase center of the GPS receiver antenna, which is closely linked to the strong seasonal nonlinear changes, such as annual and semi-annual variations. For the quality of the data used, the quality of VLBI global stations data is uneven, because of the limit of equipment and observation condition and the low data sampling rate, the amplitudes of periodic terms are seriously effected.

For the systematic errors between the semi-annual amplitudes of GPS and VLBI, we infer the reasons are from two aspects: ① the semi-annual effect caused by the design characteristics of GPS, ② the false periodic effect caused by the quality of observation data of GPS and VLBI.

4 Conclusions

In the comparison of the periodic terms of the coordinates time series of the two technical stations of GPS and VLBI at the same amalgamated station, it is found that the trends of the annual and semi-annual terms of GPS and VLBI are basically the same, at the same time, there are systematic deviations in the amplitude. We infer the reasons are from two aspects: ① the annual systematic errors caused by the solar radiation perturbation, annual configuration effect between the stations and the GPS constellation, and the semi-annual errors caused by the design characteristics of GPS, ② the false periodic terms caused by the quality of observation data of GPS and VLBI. The semi-annual systematic errors of GPS need to be further studied to determine the source of errors in the GPS.

References

1. Sun F, Tian L, Men B et al (2012) Study on correlation of temperature changes with GPS station's nonlinear movement. *Acta Geodaetica Cartogr Sin* 41(5):723–728
2. Jiang W, Wang K, Deng L et al (2015) Impact on nonlinear vertical of GNSS reference stations caused by thermal expansion. *Acta Geodaetica Cartogr Sin* 44(05):473–480
3. Dong D, Fang P, Bock Y, Cheng MK et al (2002) Anatomy of apparent seasonal variations from GPS-derived site position time series. *J Geophys Res* 107:1–16
4. Jiang W, Xia C, Li Z et al (2014) Analysis of environment loading effects on regional GPS coordinates time series. *Acta Geodaetica Cartogr Sin* 43(12):1217–1223

5. Zhang X, Lv Z (2007) Realization and improvement of ITRF2005. *Bull Surv Mapp* 7(16): 16–18
6. Zhu W, Xiong F, Song S (2008) Notes and commentary on the ITRF2005. *Prog Astron* 26 (1):1–13
7. Beutler G et al (2000) In: Gowey K, Neilan R, Moore A (eds) IGS directory 1999, JPL, California
8. Tian L, Sun F, Zhao Y et al (2013) Detection and analysis of GPS seasonal systematic errors. *Sci Surv Mapp* 03:47–48
9. Tian L, Duan J, Wang Q et al (2013) Analysis of GPS seasonal systematic error based on collocation sites. *J Geodesy Geodyn* 33(1):78–80
10. Volker T, Peter S, Markus R et al (2009) Annual deformation signals from homogeneously reprocessed VLBI and GPS height time series. *J Geodesy* 83:973–988
11. Sun F, Zhao M (1995) The methods for determining plate motion parameters from VLBI and SLR data. *Ann Shanghai Obs Acad Sinica* 16:7–13
12. Yuan L, Ding X, Wu C et al (2008) Characteristics of daily position time series from the Hong Kong GPS Fiducial Network. *Chin J Geophys* 51(05):1372–1384
13. Dong D, Herring TA, King RW (1998) Estimating regional deformation from a combination of space and terrestrial geodetic data. *J Geodesy* 72:200–214
14. Freymueller JT (2008) Seasonal position variations and regional reference frame realization. Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK 99775 USA, pp 191–196
15. Simon D, Williams P (2004) Error Analysis of continuous GPS position time series. *J Geophys Res* 109:1–19

An Accurate and Efficient Detecting and Correcting Algorithm for Abnormal Clock Data



Fan Yang and Ran Zhang

Abstract Because of various factors, there are many abnormal situations in the atomic clock measurement data from time-keeping system, such as data lost, error data, phase jump. The abnormal data can not reflect the real status of atomic clock, it is necessary to correct the abnormal data caused by short-term external disturbances. In this paper, based on analyzing for frequently used algorithm objectively and combining multi-methods, an accurate and efficient detecting and correcting algorithm is put forward. At last, the algorithm is validated using the clock measurement data, the result shows that the algorithm can accurately locate the abnormal data, analyze abnormal type, the correction effect is satisfactory.

Keywords Time-keeping system · Abnormal clock data · Detecting algorithm
Correcting algorithm

1 Introduction

Reliable atomic clock comparison data is the prerequisite and basis for clock performance analysis and time-keeping. In time-keeping system, due to many external factors, such as system power down, clock lockout, circuit thermal noise and device aging, a series of abnormalities inevitably exist in the atomic clock comparison data, common abnormal situations are data lost, error data, phase jump [1, 2]. Abnormal atomic clock comparison data does not truly reflect the atomic clock performance, which can not be applied to time scale calculation. Only by removing the effects of abnormal data, can we obtain accurate and meaningful analysis results. Therefore, it is of great significance to carry out the research on the abnormal detection and correction algorithms for atomic clock data.

The detecting method of abnormal atomic clock comparison data is mainly includes threshold method, 3σ criteria, wavelet analysis, median method and mean

F. Yang (✉) · R. Zhang
Beijing Institute of Radio Metrology and Measurement, Beijing, China
e-mail: yeyingfanfan@163.com

comparison method of adjacent data window. The experience threshold value of threshold method can only reflect the long-term data distribution in the past, cannot adapt to short-term data changes [3]. The 3σ criteria is suitable for data sequences that conform to normal distribution, it can effectively detect the individual abnormal data in the data sequence, however, it cannot detect the phase jump of the clock. Wavelet analysis has the characteristics of multi-resolution analysis, and the ability to characterize the local characteristics of the signal in the time-frequency domain [4], but it is too complex to realize. The median method is proposed by Riley, it is a simple and effective data detecting method for the large data sequence, which can effectively improve the computational efficiency, and often used to detect abnormal data [5–7]. Phase jump is a special abnormal situation, which is the appearance of the atomic clock phase from one value to another, resulting in the discontinuity of phase data. Most algorithms detect phase jump by comparing the mean of two adjacent windows that are moved on relevant data, while others use cumulative summation curves or other similar methods [8].

Based on the above analysis, this paper selects the improved median method to detect abnormal atomic clock comparison data. This algorithm can not only detect the location of abnormal data, but also can judge the abnormal type and select the appropriate correction algorithm. The improved median method is validated using the clock measurement data of time-keeping system in 203 institute, and works well. The sampling interval is 1 h. When the abnormal situation lasted less than 5 h, the abnormality was analyzed and corrected. However, when the continuous abnormal situation lasted more than 5 h, indicating that the state of atomic clock was abnormal, and the clock was unable to participate in time-keeping work, so only alarm for the abnormal but not correct.

2 Detecting and Correcting Algorithm for Abnormal Clock Data

2.1 Detecting Algorithm and Application for Abnormal Data

2.1.1 The Median Detection Method

The expression of the median detection method is as follows:

$$MAD = Median \left(\frac{y_i - m}{0.6745} \right) \quad (1)$$

In the formula, m represents the medium of the sequence $\{y_i, i = 1, 2, \dots, N\}$, when the observation $|y_i| > (m + n \times MAD)$ (n is a positive integer, the value of n is different for different atomic clocks), it is considered to be abnormal data, and set to zero.

As can be seen from the above formula, the median method has a good robustness, since it is not sensitive to the size of the gross errors. It should be noted that the median detection method is generally used for frequency data, because the gross error of the phase data is corresponding to the peak of frequency data, and the peak is easy to be detected.

2.1.2 Application of Improved Median Detection Method in Detecting for Abnormal Data

In atomic clock comparison data, the common abnormal situations include data loss, error data, phase jump. The traditional median detection method sets the abnormal data to zero, but the atomic clock comparison data is large and complex, simply setting to zero will generate new abnormal data. In order to analyze the abnormal situations according to the actual situation, the median detection method is improved. The improved algorithm can locate the abnormal data accurately and determine which kind of abnormal situations it belongs to, so as to select the appropriate correction algorithm.

The atomic clock comparison data is phase data, which is needed to convert to frequency data. The abnormal frequency data that detected by the algorithm in 2.1.1 is marked by “false”. The abnormal data detecting algorithm is shown in Fig. 1. According to the correspondence between frequency data and phase data, the abnormal type of phase data can be judged by the number of “false”, and the abnormal type detecting algorithm is shown in Fig. 2.

2.2 Correcting Algorithm for Abnormal Data

2.2.1 Linear Interpolation

The formula of linear interpolation is as follows:

$$y_i = y_{pre} \times \frac{x_{next} - x_i}{x_{next} - x_{pre}} + y_{next} \times \frac{x_i - x_{pre}}{x_{next} - x_{pre}} \quad (2)$$

In the formula, y_{pre} , y_{next} respectively represents the normal point before and after the abnormal point, x_i , x_{pre} , x_{next} respectively represents the position of the abnormal point and the two normal points before and after the abnormal point.

Through analysis and verification, corresponding to Fig. 2, when $n = \{1, 2\}$, with the condition of “The phase corresponding to $flg[i]$ is abnormal and then returns to normal”, the linear interpolation method has a good correcting effect.

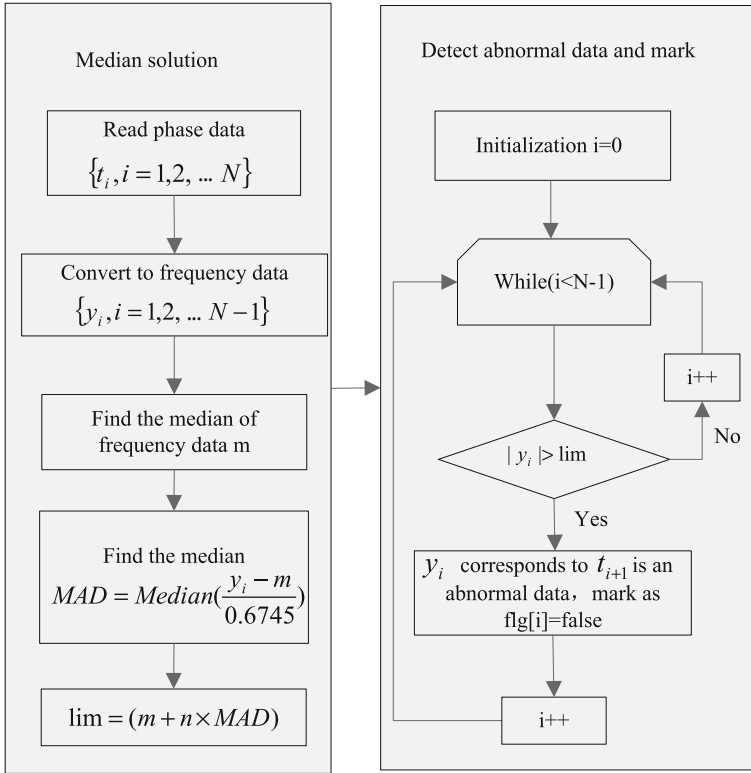


Fig. 1 Abnormal data detecting algorithm

2.2.2 Phase Compensation Method

Assuming that (x_i, y_i) is the starting point of phase jump, (x_{i-1}, y_{i-1}) and (x_{i-2}, y_{i-2}) respectively represent two adjacent normal points before jump point, then:

$$\begin{cases} const = y_{i-1} - y_{i-2} + lim \times |x_{i-1} - x_{i-2}| \\ y_{i,new} = y_i + const \end{cases} \quad (3)$$

In the above formula, “lim” is the frequency threshold between two adjacent phase data obtained by the median method. $y_{i,new}$ is the value of the phase jump point after phase compensation.

Through analysis and verification, corresponding to Fig. 2, when there is a short jump phase, that is, $1 \leq n < 5$, with the conditions of $m < 5$, this method can effectively compensate phase jump value.

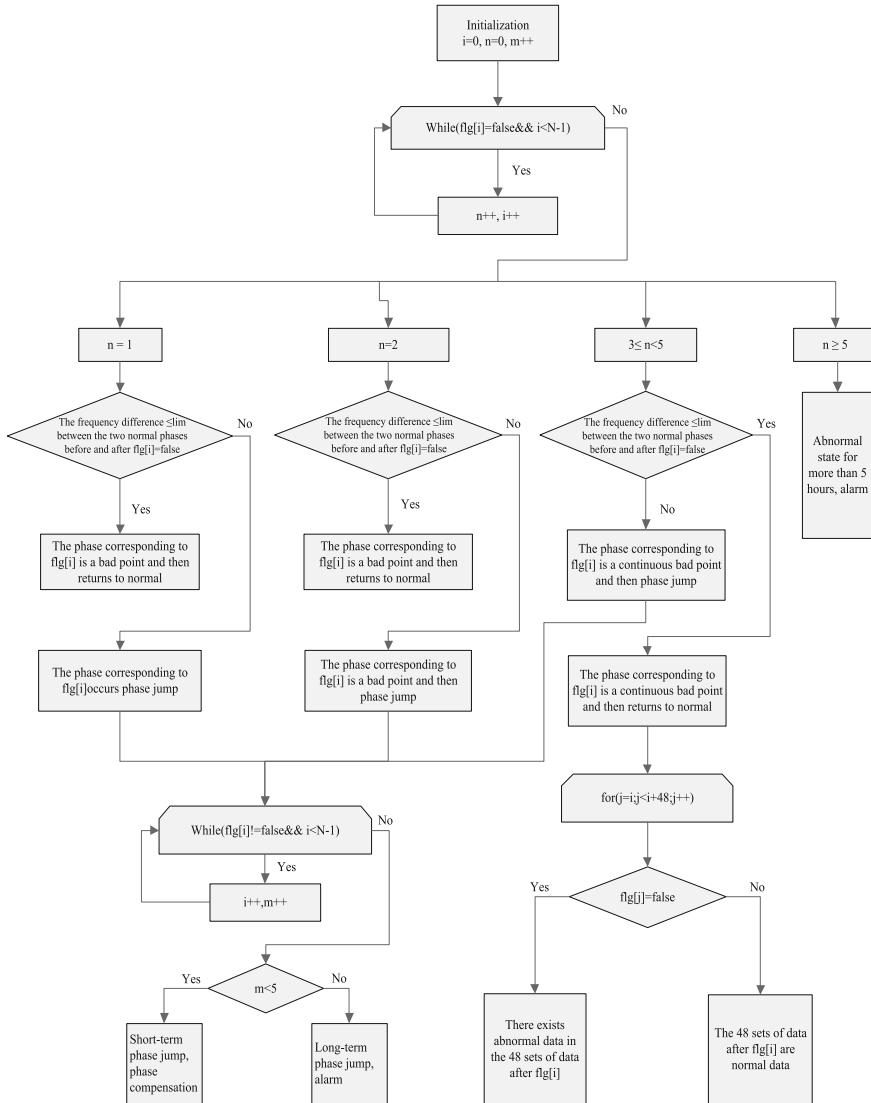


Fig. 2 Abnormal type detecting algorithm

2.2.3 Least Squares Method

The idea of least square algorithm is: for a given set of data $(x_i, y_i) (i = 0, 1, \dots, m)$, it is required to find a function $y = S^*(x)$ in the function set $\varphi = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$ to minimize the quadratic sum of the errors, as [9] proposed.

$$\|\delta\|_2^2 = \sum_{i=0}^m \delta_i^2 = \sum_{i=0}^m [S^*(x_i) - y_i]^2 = \min_{S(x) \in \varphi} \sum_{i=1}^m [S(x_i) - y_i]^2 \tag{4}$$

$$S(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \dots + a_n\varphi_n(x) \quad (n < m) \tag{5}$$

The least square algorithm processes as follows:

- (1) Consider $\|\delta\|_2^2$ as weighted quadratic sum;

$$\|\delta\|_2^2 = \sum_{i=0}^m \omega(x_i) [S(x_i) - f(x_i)]^2 \tag{6}$$

- (2) In order to find a function $y = S^*(x)$ to minimize $\|\delta\|_2^2$, the problem is transformed into finding the minimum of multivariate function;

$$\left. \begin{aligned} I(a_0, a_1, \dots, a_n) &= \sum_{i=0}^m \omega(x_i) \left[\sum_{j=0}^n a_j \varphi_j(x_i) - f(x_i) \right]^2 \\ \frac{\partial I}{\partial a_k} &= \sum_{i=0}^m 2\omega(x_i) \left[\sum_{j=0}^n a_j \varphi_j(x_i) - f(x_i) \right] \varphi_k(x_i) = 0 \quad (k = 0, 1, \dots, n) \end{aligned} \right\} \tag{7}$$

- (3) According to Eq. (7) can obtain the only solution

$$a_k = a_k^* \quad (k = 0, 1, \dots, n) \tag{8}$$

- (4) Thus, the least squares solution of the function $f(x)$ is:

$$S^*(x) = a_0^*\varphi_0(x) + a_1^*\varphi_1(x) + \dots + a_n^*\varphi_n(x) \tag{9}$$

In practical applications, the least square fitting times is determined by the type of atomic clock, e.g. cesium clocks usually choose least squares linear fitting, meanwhile, hydrogen clocks generally use least squares quadratic fitting. The least squares fit sample uses 48 consecutive normal data before the abnormal data.

Through analysis and verification, the abnormal data can be well corrected by the least squares fitting in the following two kinds of abnormal situations: (1) the phase jump after the abnormal data, corresponding to Fig. 2, that is, under the condition of $n = 2$, $n = 3$ and $3 < n < 5$, “The phase corresponding to $\text{flg}[i]$ is (continuous) a bad point and then phase jump”; (2) there exists abnormal point in

the 48 consecutive after abnormal data, corresponding to Fig. 2, that is, under the conditions of $3 < n < 5$, “The phase corresponding to $\text{flg}[i]$ is continuous bad points and then returns to normal, but there exists abnormal point in the 48 consecutive after $\text{flg}[i]$ ”.

2.2.4 Least Squares Combination Model

When the number of continuous abnormal data increases, only choosing the normal data points prior to the abnormal data to do the least squares results in a gradual increase in the deviation between the predicted and actual measurements. This is due to the limitations of the model itself, which takes into account only the trend of data sequence changes before the abnormal data and neglects the trend of data changes after the abnormal data. In view of this situation, this paper presents a least squares combination model, the algorithm process is as follows:

- (1) Assuming that n is the number of continuous abnormal data, k pairs of normal data points before abnormal data are $\{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\}$, k pairs of normal data points after abnormal data are $\{(x_{k+n+1}, y_{k+n+1}), (x_{k+n+2}, y_{k+n+2}), \dots, (x_{2k+n}, y_{2k+n})\}$, and both data points are known;
- (2) Using the least squares fitting to model the forward k pairs of data, and using the model to forecast n pairs of abnormal data and backward k pairs of normal data;
- (3) To solve the residual between the predicted values and the measured values of backward k pairs of data points;
- (4) Using the least squares fitting to model the residuals of the backward k pairs of data, and using the model to predict residuals Δy of n pairs of abnormal data;
- (5) The final correcting values of n pairs of abnormal data are $y = \bar{y} + \Delta y$.

In practical application, the cesium clock uses the least squares linear fitting to model the 48 pairs of data before and after the abnormal data, the hydrogen clock uses the least squares quadratic fitting to model the 48 pairs of data before and after the abnormal data. Through analysis and validation, when the phase does not jump after the continuous abnormal values, corresponding to Fig. 2, that is, under the condition of $3 < n < 5$, “The phase corresponding to $\text{flg}[i]$ is a continuous bad point, then returns to normal, and the 48 pairs of data after $\text{flg}[i]$ are normal”, the least-squares combination model has a good correction effect.

3 Result Analysis of Algorithm Examples

The paper selects a piece of representative data to verify algorithm, due to external factors, from 11:00 on October 31, 2017 to at 11:00 on November 4, 2017, there is an abnormality in the measured clock value between the clock 2 and the main clock

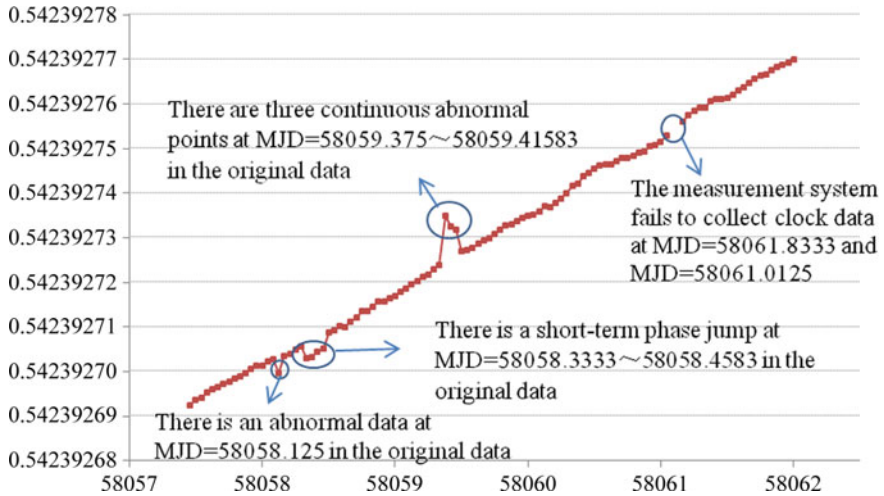


Fig. 3 Original phase of UTC(BIRM)-clock 2

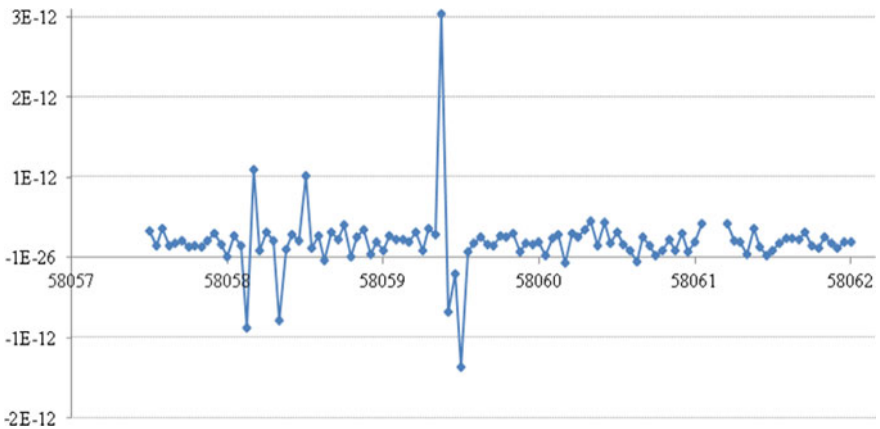


Fig. 4 Original frequency of UTC(BIRM)-clock 2

collected by the laboratory clock measurement system. The original phase and the abnormal situation are shown in Fig. 3, and the original frequency is shown in Fig. 4.

According to the detecting algorithm in 2.1 and the correcting algorithm in 2.2, the location of the abnormal situation and the corresponding correcting algorithm are as follows:

- (1) MJD = 58,058.125: there is an abnormal data in the original data, which can be corrected by linear interpolation;

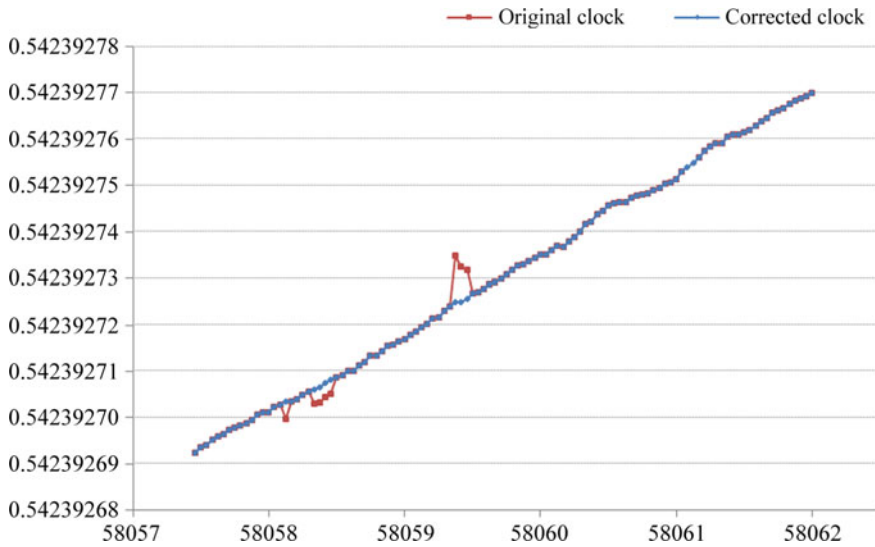


Fig. 5 Correction effect

- (2) $MJD = 58,058.3333-58,058.4583$: there is a short-term phase jump in the original data, which can be corrected by the phase compensation method;
- (3) $MJD = 58,059.375-58,059.4583$: there are three continuous abnormal points, and exists abnormal data in the 48 sets of data after abnormal point, which can be corrected by the least square fitting method using forward 48 pairs of data (the 48 pairs of data have been corrected as available data);
- (4) $MJD = 58,061.8333$ 和 $MJD = 58,061.0125$: the measurement system fails to collect clock data at these two moments, and the frequency anomalies of these two moments can be detected by using the median detection method, which can be corrected by linear interpolation.

The correcting effect is shown in Fig. 5, where the red curve represents the original phase and the blue curve represents the corrected phase. The stability before and after the correction is shown in Fig. 6, due to the limited amount of data, only calculated to 16 h of stability. As can be seen from the figure, compared with the original data, the stability of the corrected data has been significantly improved, the correction effect is fine.

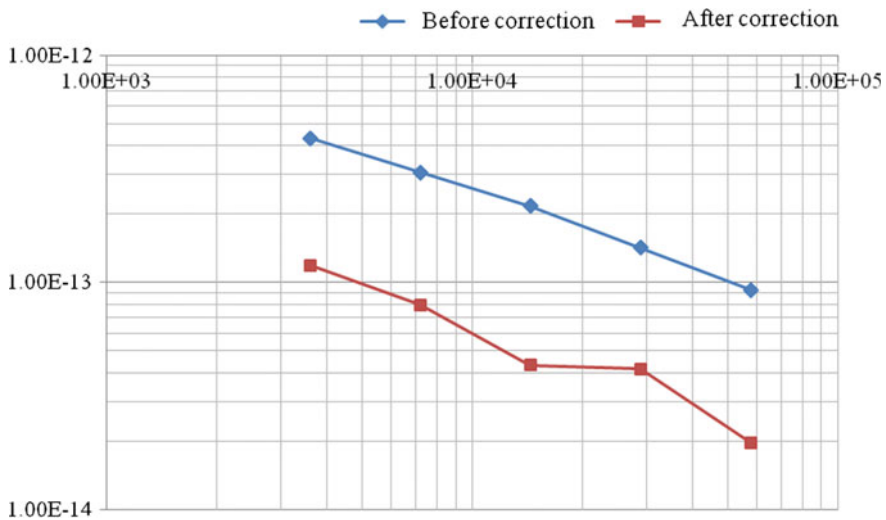


Fig. 6 Stability comparison before and after correction

4 Conclusion

This paper focuses on the detection and correction of common abnormal situations in time-keeping system clock measurement data. It introduces an improved median gross error detection method, which can accurately locate the abnormal data, analyze abnormal type and choose different correcting algorithm for different abnormal type: (1) For individual (less than 2) abnormal values, choose linear interpolation method to correct; (2) For short-term phase jump (jump time less than 5 h), choose phase compensation method to correct; (3) If the phase jump occurs after the abnormal value, choose the least square fitting method to correct the abnormal value; (4) If the 48 pairs of data after the abnormal value did not jump, and there are no other anomalies, then choose the least squares combination model to correct; (5) For long-term (more than 5 h) abnormal situations, indicating abnormal clock state, giving alarm information. It is validated using the clock measurement data of time-keeping system in 203 institute, the results show that the atomic clock comparison abnormal data detecting and correcting algorithm proposed in this paper is accurate and effective, the corrected data can objectively reflect the atomic clock performance, and participate in time-keeping work.

References

1. Yang J, Li SG, Xu YQ (2014) Effect of temperature on hydrogen maser performance. *J Test Meas Technol* 28(2):93–97
2. Li LH, Peng JH, Zhang JY (2013) Influence of environment factors on atomic clock
3. Yuan HB, Li W, Qu LL (2010) Research on error checking and correction algorithm for the measurement data from time-keeping system. In: International conference on bio-inspired systems and signal processing (ICBSSP 2010), pp 253–256
4. Feng SL (2009) Study on the methods of data preprocessing and performance analysis for atomic clocks. In: PLA Information Engineering University, pp 20–21
5. Riley WJ (2007) Handbook of frequency stability analysis. Hamilton Technical Services, pp 86–96
6. Riley WJ (2002) The calculation of time domain frequency stability. Hamilton Technical Services, pp 360–366
7. Riley WJ (2003) Dead time and uneven spacing in frequency stability. Hamilton Technical Services, pp 5–10
8. Riley WJ (2008) Frequency jump detection and analysis. In: The 40th annual PTTI meeting
9. Li QY, Wang CN, Yi DY (2001) Numerical analysis. Huazhong University of Science and Technology Press

Processing Method of Clock Break in VLBI Data Analysis



Han Lei, Lu Cao, Ke Xu and Chong Wang

Abstract Each observation station in the VLBI has an independent hydrogen clock. In particular, on rare occasions the phase lock synthesizer loss phase lock on the signal from the hydrogen maser cavity resulting in a discontinuity in the clock time obtained from the maser, this phenomenon is referred to as a “clock break”. Clock breaks can be several hundred nanoseconds. In the VLBI data analysis and analysis to consider the special operation of the clock breaks. In the least squares solution method, it is necessary to reduce the influence of the clock break on the residuals of the observed data and the accuracy of the solution. The method of positioning the clock break is used to reduce the influence of the clock break by means of piecewise linearization. In the Kalman filter method, the stochastic model can not contain such a rapid clock time variation, but also the method of locating the clock break and then including the clock break in the state transition noise of the Kalman filter. Thus improve the accuracy of VLBI data solution.

Keywords VLBI · Clock break · Piecewise linear · Least squares method
Kalman filter

1 Introduction

Very Long Baseline Interferometry (VLBI) use to accurately record time-tagged signals from extragalactic radio sources. The two radio telescopes are in pair and the two received signals are cross-correlated to obtain the difference (delay) between the two signals arriving at the antenna and the rate of change (delay rate) of

H. Lei (✉) · L. Cao · K. Xu · C. Wang
State Key Laboratory of Astronautic Dynamics, Xi'an, China
e-mail: 583346800@qq.com

H. Lei
Xi'an Jiaotong University, Xi'an, China

H. Lei · L. Cao · K. Xu
State Key Laboratory of Geo-Information Engineering, Xi'an, China

the interferometric phase delay (delay rate). Radio telescopes receiver reference frequency, calibration signal and time-tag information are derived from the accurate clocks at the radio telescopes. Almost all of these clocks are hydrogen maser frequency standards [1]. When observing many different radio sources, a observation of a radio source can last 100–400 s and the entire observation lasts for 24–48 h. The obtained delay and delay rate can be used to determine the relative position of the radio telescopes and the coordinates of the radio sources [2, 3].

VLBI determines the relative position of radio telescopes with a precision of better than 1 part in 10^9 . Measurements with a precision of 2×10^{-9} have been obtained, the general accuracy of measurement results have reached 4×10^{-9} . However, to achieve these accuracies, the effect of clock errors from different radio telescopes and signal delays caused by propagation medium must be taken into account to enable data analysis that attains the accuracy requirements. Because any signal received by a station from a radio source needs to be time stamped by the local clock time, the difference of the time of the independent radio telescope clock will directly affect the measurement of the delay, especially the discontinuity of the clock signal, The accuracy of VLBI data analysis also has a greater impact. Because any signal received by a station from radio sources needs to be time-tag by the local clock time, the difference of the time of the independent radio telescope clocks will directly affect the measurement of the delay. In particular, the discontinuity of the clock signal also has a greater impact on the accuracy of the VLBI data analysis results.

2 VLBI System Principle

2.1 VLBI Geometric Principle

VLBI technology is based on two stations to form a baseline, through the radio signal received from radio sources recording, transmission and correlation processing to obtain the delay and delay rate measurements that reflect the geometric relationship between the tracking target and the position of the two stations.

Due to the different distances between the two antennas and the observation radio source, the time when the same wavefront of the radio signal reaches the two antennas will also be different, and there will be a delay τ_g [4, 5]. As shown in Fig. 1, the basic concept of VLBI is that the two antennas make up the interferometer with a baseline vector \bar{B} between them. The plane wave front propagating along the unit vector \bar{k} of the radio source arrives at the two antennas pointing to this radio source. From the geometric relationship analysis of Fig. 1, the inner product of \bar{B} and \bar{k} is divided by the speed of light c as the time delay τ_g , and the formula is expressed as follows:

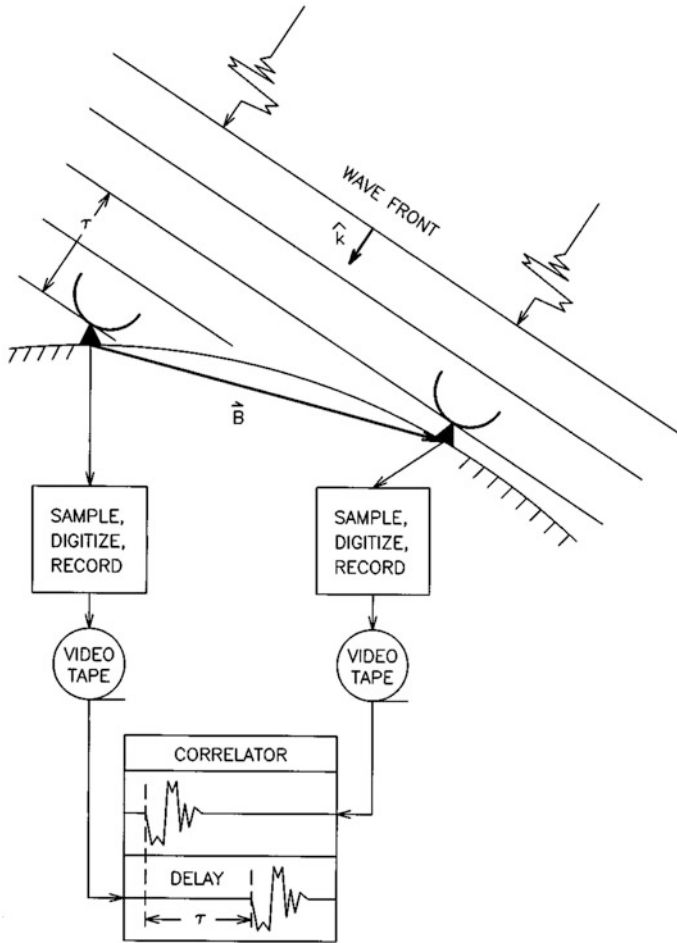


Fig. 1 Basic concept of very long baseline interferometry

$$\tau_g = - \frac{(\vec{B} \cdot \vec{k})}{c} \tag{2.1}$$

The derivation of formula (2.1) is from purely geometric relation, so called τ_g geometric delay. In the actual VLBI observations τ_g can not be measured directly. Because the actual observations are subject to other factors, such as clock errors, equipment delays, delays in the propagation medium, etc. If set τ as the actual delay observation, then:

$$\tau = \tau_g + \tau_{clk} + \tau_{inst} + \tau_{trop} + \tau_{iono} + \dots \quad (2.2)$$

The geometric delay τ_g can only be obtained by deducting the rest of the measured value. From formula (2.2) we can see that the influence of clock parameters such as clock errors on the actual delay, so clock parameters would affect the result of the final VLBI data analysis [6].

2.2 VLBI Data Calculation Principle

The principle of estimating the parameters is to establish the corresponding observation error equation by using the delay obtained from the actual observation and the theoretical calculation delay, and the error equation is solved by the least square method or linear filter principle to obtain the required parameters [7]. VLBI basic observation equation is below

$$\mathbf{O}_t = \mathbf{C}(\mathbf{X}, t) + \mathbf{V}_t \quad (2.3)$$

In the formula, \mathbf{O}_t is the delay obtained at time t ; \mathbf{X} represents the vector of observations related to the parameters; $\mathbf{C}(\mathbf{X}, t)$ is the mathematical model in which parameter \mathbf{X} affect the observation \mathbf{O}_t , also known as the theoretical value of the observation. \mathbf{V}_t is the observed noise residual vector. Assuming that the priori value of \mathbf{X} is \mathbf{X}_0 and the correction value is \mathbf{x} , the linearized observation equation is:

$$\mathbf{O}_t = \mathbf{C}(\mathbf{X}_0, t) + \frac{\partial \mathbf{C}(\mathbf{X}, t)}{\partial \mathbf{X}} \Big|_{\mathbf{X}_0} \cdot \mathbf{x} + \mathbf{V}_t \quad (2.4)$$

The above equation was linearized observation equation:

$$\mathbf{y}_t = \mathbf{A}_t \mathbf{x} + \mathbf{V}_t \quad (2.5)$$

where $\mathbf{y}_t = \mathbf{O}_t - \mathbf{C}(\mathbf{X}_0, t)$ is the difference between the observed value and the theoretical value at time t . \mathbf{A}_t is the partial derivative matrix, which is related to the change of the parameter value with respect to the observed value. All observation equations are expressed in matrix:

$$\mathbf{Y} = \mathbf{A}\mathbf{x} + \mathbf{V} \quad (2.6)$$

Suppose the observation weight matrix is \mathbf{P} , then the least squares theory can find the correction values of the parameters is:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{P} \mathbf{A}^T \mathbf{Y} \quad (2.7)$$

Using Kalman filtering technique to solve the VLBI parameter, linearization observation equation is:

$$\mathbf{y}_t = \mathbf{A}_t \mathbf{x}_t + \mathbf{V}_t \quad (2.8)$$

Kalman filter estimation is based on the observation epoch time sequence, First, the state at time t is used to predict the state at time $t + 1$, then the observation at time $t + 1$ is added to modify the predicted value to obtain the state estimate at time $t + 1$.

Prediction:

$$\mathbf{x}'_{t+1} = \mathbf{S}_t \mathbf{x}'_t \quad (2.9)$$

$$\mathbf{C}'_{t+1} = \mathbf{S}_t \mathbf{C}'_t \mathbf{S}_t^T + \mathbf{W}_t \quad (2.10)$$

Modification:

$$\mathbf{x}^{t+1}_{t+1} = \mathbf{x}'_{t+1} + \mathbf{K}_{t+1} (\mathbf{y}_{t+1} - \mathbf{A}_{t+1} \mathbf{x}'_{t+1}) \quad (2.11)$$

$$\mathbf{C}^{t+1}_{t+1} = \mathbf{C}'_{t+1} - \mathbf{K}_{t+1} \mathbf{A}_{t+1} \mathbf{C}'_{t+1} \quad (2.12)$$

The above t represents the moment corresponding to the last data used to estimate, \mathbf{C}'_t is the covariance matrix, \mathbf{K}_{t+1} is the Kalman gain.

Equations (2.9)–(2.12) are the complete process of Kalman filter. When the calculation at time $t + 1$ is completed, the calculation process will be repeated until all the observations are processed [8].

3 Clock Break Positioning

In particular, on rare occasions the phase lock synthesizer loss phase lock on the signal from the hydrogen maser cavity resulting in a discontinuity in the clock time obtained from the maser, this phenomenon is referred to as a “clock break”. Clock breaks can be several hundred nanoseconds. In many cases, such information is available from the logs of the operations at each of the telescopes. However, often these events are sufficiently small that they are not noticed during the experiment in which case clock breaks must be detected using the VLBI data themselves. In the VLBI data analysis we should consider to do a special treatment for clock break to reduce the impact of clock break on the accuracy of the solution.

We take the example of “09NOV24XA_N004” provided by International VLBI Service (IVS) as an example for the following discussion. We can see from Table 1 that the position of the BADARY site perturbed by clock break is estimated to 52 mm. According to our calculation of normal data without clock break, the standard deviation of the horizontal coordinate correction of stations is between 4

Table 1 Effects of a clock break on estimates of site coordinates

Site	ΔN (mm)	ΔE (mm)	Δh (mm)
WETTZELL	7	-9	0
RICHMOND	-1	7	14
WESTFORD	-2	-3	10
HRAS 085	0	-6	16
BADARY	-4	-8	52

and 10 mm, while the standard deviation of the vertical coordinate correction is between 10 and 20 mm. The vertical correction at BADARY site is significantly larger. Here we recognize that this data is different from the other data, and we further hypothesize that this is due to the clock break in the five stations included in this data. Below we will further verify this hypothesis.

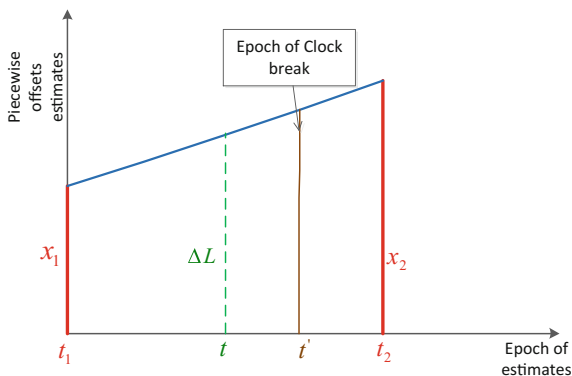
3.1 Piecewise Linearization

The method of piecewise linearization is used to solve the parameter in the process of solving the VLBI data by the least square method.

In Fig. 2, the vertical axis represents the parameter correction value, and the horizontal axis represents the epoch time. x_1 is the parameter correction value at time t_1 , x_2 is the parameter correction value at time t_2 , and ΔL is the correction value at time t of observation epoch. The piecewise linearization means that the observed residuals at observation epoch t are represented by the correction values at estimated moments t_1 and t_2 , and the parameter correction values at the specified moment t_1 and t_2 to be estimated is obtained. Figure 2 shows the following relationship:

$$\Delta L = x_1 + \frac{t - t_1}{t_2 - t_1}(x_2 - x_1) \tag{3.1}$$

Fig. 2 Basic concept of piece-wise linear



Partial derivation of x_1 and x_2 respectively leads to:

$$\frac{\partial \Delta L}{\partial x_1} = 1 - \frac{t - t_1}{t_2 - t_1} \tag{3.2}$$

$$\frac{\partial \Delta L}{\partial x_2} = \frac{t - t_1}{t_2 - t_1} \tag{3.3}$$

Set observation time Epoch observation equation is:

$$\mathbf{y}_t = \mathbf{A}_t \Delta L + \mathbf{V}_t \tag{3.4}$$

From (3.2) and (3.3) can be obtained the observation equation which is expressed with x_1 and x_2 is:

$$\mathbf{y}_t = \mathbf{A}_t \left(1 - \frac{t - t_1}{t_2 - t_1} \right) x_1 + \mathbf{A}_t \left(\frac{t - t_1}{t_2 - t_1} \right) x_2 + \mathbf{V}_t \tag{3.5}$$

The new observation equation established in (3.5) can continue to solve the data according to the least square method described above. We generally use piecewise linearization over a period of one hour. It can be seen from Fig. 2 that if there is a clock break at the station’s hydrogen clock at time t' , this period can not be well represented by linearization. In order to use the method of piecewise linearization, we must consider effective methods to deal with the clock break, thus reducing or eliminating the impact of clock break.

3.2 Clock Break Positioning and Processing

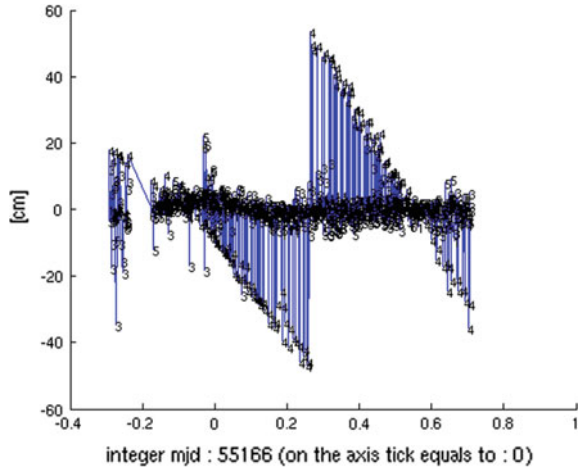
We use the least square method to initially solve the clock parameters, and get the solve residuals of local clock parameters. Solution steps are as follows:

- (1) Calculate the partial derivative matrix of the clock parameters (including clock error, clock speed and clock acceleration).
- (2) Solve the clock parameters using the least squares method introduced in Sect. 2.2.
- (3) Calculate the solution residuals of each station’s solution clock parameters. Calculate the residual calculated as follows:

$$\mathbf{residuals} = \mathbf{y} - \mathbf{Ax} \tag{3.6}$$

Figure 3 is the clock parameters solution residuals obtained from the five stations. We note that the residuals for the baseline corresponding to the

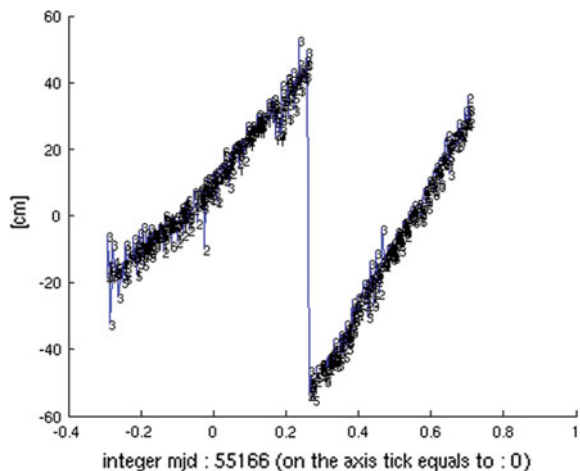
Fig. 3 Residuals for station WETTZELL



“WETTZELL” station and the “BADARY” station numbered “4” are very large. And the “BADARY” station clock parameter solution residuals shown in Fig. 4, the residual values in a short time quickly become from positive to negative, indicating the “BADARY” station hydrogen clock at that epoch appeared clock break.

We will zoom in Fig. 4 positioning clock break epoch time, record the clock station and epoch time. The information is written to the solution operation file. Introduced above is to manually locate the clock break. We designed a simple algorithm to automatically search for clock break by solving residual data at each station. The following pseudo-code algorithm:

Fig. 4 Residuals for station BADARY



for i = 1:ob_num % Observations for a single station

total = 0+residual(i) % Find the sum of a single station solving residuals

end

resdiff = residual(2:end)-residual(1:end-1)

resstd = standard deviation (resdiff)

num = find(| resdiff | > resstd*10)

Calculate the standard deviation of the variation of residuals between each observation. Taking this ten times of the standard deviation as the threshold, if the variation of the solution residual between two observations is greater than this threshold, we assume that a clock break has occurred between the two observations. Once again, when performing piecewise linearization, the epoch time of the clock break is set as a segment time point, the data before the clock break occurs is put into a segment for linearization, and the data after the clock break is placed in another segment for linearization. The final piecewise linearization will eliminate the effect of clock break.

Figure 5 shows the solution residual before the clock break correction. Figure 6 shows the solution residual after clock break correction. It is obvious from the comparisons of the two figures that the distribution of residuals in Fig. 6 is more concentrated, and the residual values have also been significantly improved. In order to quantify the comparison, we calculated the standard deviation of the residuals, and the standard deviation decreased from 2.32 in Fig. 5 to 0.69 in Fig. 6. Therefore, we found the clock break through the above methods, and locate and correct, to a certain extent, improve the use of least squares method to calculate the accuracy of the data.

Fig. 5 Residuals of total observation before clock break correction

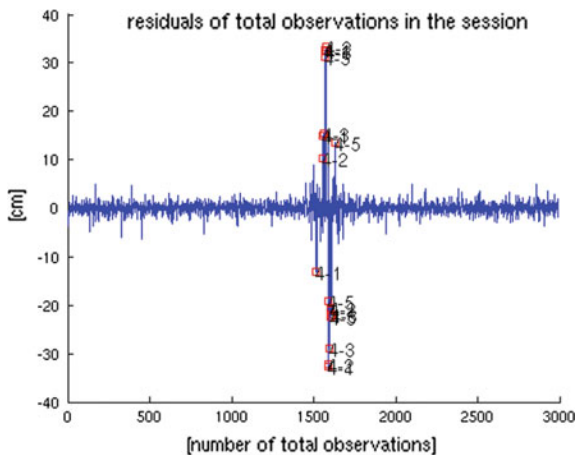
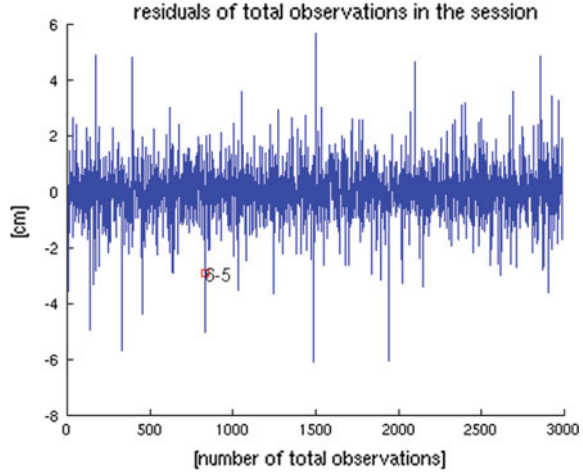


Fig. 6 Residuals of total observation after clock break



4 Clock Break in Kalman Filter

In the Kalman filter method, the stochastic model we use can not completely represent the clock variation of the hydrogen clock of a station. In particular, there is a clock break in the station hydrogen clock, which can reach hundreds of nanoseconds. since the stochastic models do not account for such rapid changes in the clock time, these changes are explicitly accounted for in the kalman filter equations by increasing the elements of \mathbf{W}_t , which are affected by the clock break, to reflect the uncertainty of the new value of the clock time. Compared with the variance of the clock variation between two successful observations, the variance of the elements corresponding to the clock parameters increases by a large amount during the epoch of clock break because of the discontinuity caused by clock break. The algorithm assumes that we know the epoch time before the Kalman filter processing the data. For clock break greater than a few nanoseconds, pre-loaded delay residuals may be used; for smaller clock break, variations in the estimated clock or changes in atmospheric zenith delay may be used to include clock break.

The smoothing effect of the Kalman filter as it tried to filter through the discontinuity, there is a rapid fluctuation of the atmospheric zenith delay estimates around the time of the clock break. Because of the atmospheric delay white noise excitation power spectral density (PSD) is usually much larger than the clock white noise corresponding PSD. Filters therefore tend to attribute rapid changes to atmospheric delays [9]. Biases between the estimates of the clocks and the atmospheric delays, with and without the clock break epoch in the solution, arise because the Kalman filter distributes the anomalous behavior to all of parameters being estimated.

The clock break algorithm nominally only needs the epoch time of clock break, and suppose the clock break is smaller than the prior variance of the clock model.

We usually set the a priori variance of the clock model to be $(10 \text{ ns})^2$ and the clock offset can be estimated with a variance of $<(0.1 \text{ ns})^2$, the estimation of the clock break of 100 ns will be less than 0.01 ns due to the restriction of its constraint. Before data filtering, it is also possible to incorporate very large clock break into the theoretical model, allowing any deviation to be reduced to negligible values.

5 Discussion and Conclusion

This article briefly introduces the principle of VLBI technology and VLBI data solution. The effect of clock break on data resolution is discussed. The residual data is calculated by using the clock parameter to manually or automatically locate the clock break. The piecewise linearization method is used to eliminate the impact of clock break on the data solution. A simple comparison of the results before and after the clock break correction verifies the validity of the clock break correction method. Finally, the clock break correction of Kalman filtering and related problems are briefly discussed. In the future, we will conduct a more in-depth study on the specific technical methods of Kalman filter clock break correction.

Acknowledgements This study is funded by State Key Laboratory of Geo-information Engineering, Xi'an, China(NO. SKLGIE2016-M-1-6). This work is also supported by State Key Laboratory of Astronautic Dynamics and Xi'an Jiaotong University, Xi'an, China.

References

1. Vessot RFC (1980) Hydrogen maser frequency standards. In: Proceedings of conference on radio interferometry, NASA Conference Publication 2115, 1980, pp 203–222
2. Rogers AEE et al (1983) Very-long-baseline interferometry: the mark III system for geodesy, astrometry, and aperture synthesis. *Science* 219:51–54
3. Clark TA et al (1985) Precision geodesy using the mark III very-long-baseline interferometer system. In: Transactions on Geoscience and Remote Sensing, 1985, GE-23, pp 438–449
4. McCarthy DD, Gérard P (2003) IERS conventions 2003, IERS technical note 32, 2004
5. Gérard P, Brian L (2010) IERS conventions, 2010, IERS technical note 36, 2011
6. Sekido M, Fukushima T (2005) Relativistic VLBI delay model for finite distance radio source. *J Geodesy* 128(G02):141–145
7. Sovers OJ, Fanselow JL (1998) Astrometry and geodesy with radio interferometry: experiments, models, results. *Rev Mod Phys* 70(4):1393–1454
8. Nilsson T, Soja B, Karbon M, Heinkelmann R, Schuh H (2015) Application of Kalman filtering in VLBI data analysis. *Earth, Planets Space*
9. Soja B, Nilsson T, Karbon M, Zus F, Dick G, Deng Z, Wickert J (2015) Tropospheric delay determination by Kalman filter VLBI data. *Earth, Planets and Space*

An Improved Abnormity Detection Method for Atomic Clock Frequency Based on Chi-Squared Statistics



Huijie Song, Shaowu Dong, Meng Jiang, Yongnan Rao, Haibo Yuan and Hong Zhang

Abstract An improved chi-squared test method used to detect frequency abnormality of atomic clock is presented. This method is independent of choice of reference value, which is applicable to the measure of data in pairs. The method may speed up the process of analysis and detection for frequency abnormality. The result shows that compared with the classical chi-squared test method, the detection performance of frequency abnormality of atomic clock are improved and the calculation procedure is simplified.

Keywords Atomic clock · Frequency abnormality · Chi-square statistics
Detection probability

1 Introduction

Atomic clock frequency abnormalities detection is of importance in time-keeping. In order to maintain the continuity of phase and frequency of time scale, it is necessary to detect and calculate atomic clock phase and frequency abnormalities. Atomic clock real-time physical signal, especially the phase and frequency abnormalities of the master clock signal need to be monitored.

In global navigation satellite system, the accuracy and stability of on-board clocks can affect the precision of ranging, and then affect the positioning precision [1]. Space clocks are subject to several abnormalities that can be caused by

H. Song (✉) · S. Dong · M. Jiang · H. Yuan · H. Zhang
National Time Service Center, Chinese Academy of Sciences, Xi'an, China
e-mail: songhuijie@ntsc.ac.cn

H. Song · S. Dong · M. Jiang · H. Yuan · H. Zhang
Key Laboratory of Time and Frequency Primary Standards,
Chinese Academy of Sciences, Xi'an, China

Y. Rao
Navigation & Communication Research Lab, Beijing, China
e-mail: raoyongnan@163.com

environment or internal clock malfunction. These abnormalities must be detected and communicated to the user as soon as possible. The detection of atomic clock abnormalities is therefore a critical problem. This algorithm must be fast and it must have a high probability of detection (PD) and a low probability of false alarm (PFA) [2].

Frequency abnormalities are common in atomic clocks, including frequency jump and frequency drift jump. Therefore, several detection algorithms are used to detect frequency abnormalities [2–16]. The literature [2, 3] introduce the Kalman filtering algorithm for the frequency abnormalities detection of atomic clocks. Literature [4–10] discuss the frequency abnormalities detection algorithms of atomic clocks based on frequency spectrum density and dynamic Allan variance. In the paper [11–17], the frequency abnormalities detection algorithms of atomic clocks are studied according to mathematical model and statistical method.

In this paper, the detection algorithm of frequency abnormality of atomic clock based on pair-difference chi-squared test is developed, the pair-difference chi-squared statistics is used to test whether or not two data sets are equivalent with their associated uncertainties [18].

2 Chi-Squared Tests for Atomic Clock Frequency Abnormality

The classical chi-squared statistics starts with the inverse-variance weighted mean as the locator of central tendency used as the frequency reference value. For N participating atomic clocks, with the i th atomic clock reporting frequency value y_i and standard uncertainty u_i , this weighted mean is $y_w = (\sum_{i=1}^N y_i/u_i^2) / (\sum_{i=1}^N 1/u_i^2)$. For variances away from this weighted mean, the classical chi-squared statistic, expressed as a reduced chi-squared with $(N - 1)$ degrees of freedom, is

$$\chi_c^2 = (N - 1)^{-1} \sum_{i=1}^N (y_i - y_w)^2 / u_i^2 \quad (1)$$

If all N frequencies of atomic clocks are drawn from N independent Gaussian distributions, with the same unknown mean, and with standard deviations equal to the reported standard uncertainties, then the number $y_c = \chi_c^2$ computed for the frequency comparison data, $\{y_i, u_i\}$, is expected to be a sample from a positive random variable y distributed as an exact reduced chi-squared with $N - 1$ degrees of freedom, with a probability density function(PDF) that is proportional to $y^{(N-3)/2} \exp(-\frac{1}{2}(N-1)y)$ [18].

For the i th and j th atomic clocks in a laboratory, the correlation coefficient of frequency comparison data are r_{ij} , the standard uncertainties are u_i and u_j , and that

the components of the covariance matrix $\text{cov}_{ij} = r_{ij}u_iu_j$, for $i = 1, \dots, N$ and $j = 1, \dots, N$. The general least squares fit as a frequency reference would be [19]

$$\bar{y}_g = \left(\sum_{i=1}^N y_i \sum_{j=1}^N [\text{cov}^{-1}]_{ij} \right) / \sum_{i=1}^N \sum_{j=1}^N [\text{cov}^{-1}]_{ij} \tag{2}$$

Note that if the covariance matrix is diagonal, then so is its inverse, and Eq. (2) simplifies to give the usual inverse variance weighted mean discussed above.

It is the best fit in the general least squares sense, which is to say that it has the smallest ‘general chi-squared’. This general chi-squared statistic, expressed as a reduced chi-squared with $(N - 1)$ degrees of freedom, is

$$\chi_g^2 = (N - 1)^{-1} \sum_{i=1}^N \sum_{j=1}^N (y_i - \bar{y}_g) [\text{cov}^{-1}]_{ij} (y_j - \bar{y}_g) \tag{3}$$

Equation (3) is expected to draw from a chi-squared distribution with $(N - 1)$ degrees of freedom when all N results are drawn from a multivariate Gaussian distribution where each of the N variates has the same mean and where the covariance matrix is derived from the uncertainty budgets for the N measurements.

3 Pair-Difference Chi-Squared

It is possible to define a more robust variant of a reduced chi-squared that can treat inter-laboratory covariances simply and in their proper metrological context: one that requires the selection of neither a reference value nor a reference value algorithm. In N frequencies comparison, the pair-difference reduced chi-squared for frequency j is [19]

$$\begin{aligned} \chi_{jPD}^2 &= (N - 1)^{-1} \sum_{i=1; i \neq j}^N (y_i - y_j)^2 / (u(y_i - y_j))^2 \\ &= (N - 1)^{-1} \sum_{i=1; i \neq j}^N (y_i - y_j)^2 / (u_i^2 + u_j^2 - 2r_{ij}u_iu_j) \end{aligned} \tag{4}$$

This pair-difference chi-squared has the enormous advantage of detection the frequency abnormalities with the other $(N - 1)$ frequencies, relative to the stated uncertainties, without invoking any particular choice of a reference value. It is the exact form of chi-squared that is required to supported value.

4 Frequency Abnormity Chi-Squared Test

The frequency abnormities test statistic are selected as (1), (3) and (4) respectively. The test statistic can be represented as λ_N . Where the test statistic follow χ^2 distribution of $N - 1$ freedom. That is, $\lambda_N \sim \chi_{N-1}^2(\lambda)$, N represents the dimension of frequency. The criterion of frequency abnormities is written as

If $\lambda_N > T_D$, there is a frequency abnormities,
If $\lambda_N \leq T_D$, there is no frequency abnormities.

Where T_D represents the threshold, it determines the detection performance of frequency abnormities. When the probability of false alarm is α , according to the Neyman-pearson theorem, the threshold T_D is solved by the equation

$$P_{FA} = P[\lambda_k > T_D, H_0] = \alpha \quad (5)$$

Which make the probability of leakage minimum and can be expressed as

$$P_{FA} = \int_{T_D}^{\infty} \chi_{N-1}^2(\lambda) d\lambda = 1 - \int_0^{T_D} \chi_{N-1}^2(\lambda) d\lambda \quad (6)$$

The detection probability is given by $P_D = P[\lambda_{N-1} > T_D, H_1]$ and is written as

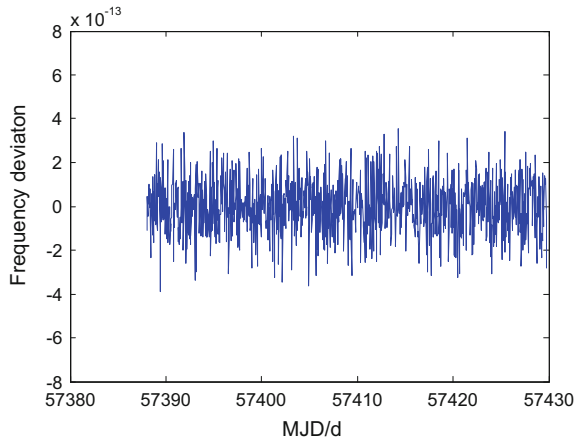
$$P_D = \int_{T_D}^{\infty} \chi_{N-1}^2(\lambda, \mu) d\lambda = 1 - \int_0^{T_D} \chi_{N-1}^2(\lambda, \mu) d\lambda \quad (7)$$

where $\chi_{N-1}^2(\lambda, \mu)$ is the probability density function of non-center chi-squared distribution, its mean is μ .

5 The Results Analysis

In order to verify the detection performance of pair-difference chi-squared statistics algorithm, the classical chi-squared statistics algorithm is used to compare the detection performance. The data is derived from the measurement data of atomic clocks of UTC(NTSC) system. The sampling interval is 3600 s. The sampling period is from January 1, 2016 to February 11, 2016 (MJD: 57388–57429), the total length is 42 days. The differences between cesium atomic clock Cs2959 and other atomic clocks are used to calculate. The frequency data is calculated by phase data of atomic clocks. It is assumed that the atomic clocks are not related to each other. According to formula (1) and formula (4), the classical chi-squared statistics and pair-difference statistics are calculated respectively. In order to verify the

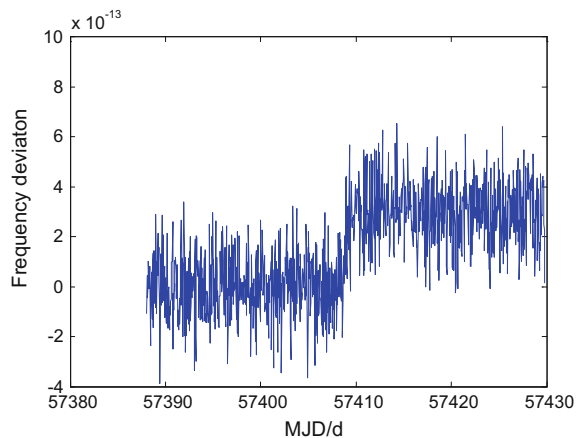
Fig. 1 The normal frequency deviation of Cs2959



effectiveness of the proposed algorithm, the different frequency jumps are added to atomic clock Cs2959. Figures 1 and 2 indicate the normal frequency and the frequency of the jump respectively. In two algorithms, the chi-squared statistic and the detection probability are calculated based on different frequency jump sizes respectively. The results are shown in Fig. 3, Tables 1 and 2. The probability of false alarm is 0.05, the degree of freedom is 9, the threshold is 16.92.

Form Table 1, it can be seen that the statistic based on pair-difference chi-squared test is greater than classical chi-squared test when the frequency jumps are less than 8×10^{-13} . When the frequency jumps are greater than 8×10^{-13} , the results of the two test algorithms are almost same. Form Fig. 3 and Table 2, the detection probability of pair-difference chi-squared algorithm is higher than that of classical test algorithm. When the jump is smaller than 7×10^{-13} , the detection probability is improved obviously. The reason is that the statistic of pair-difference

Fig. 2 The abnormal frequency deviation of Cs2959



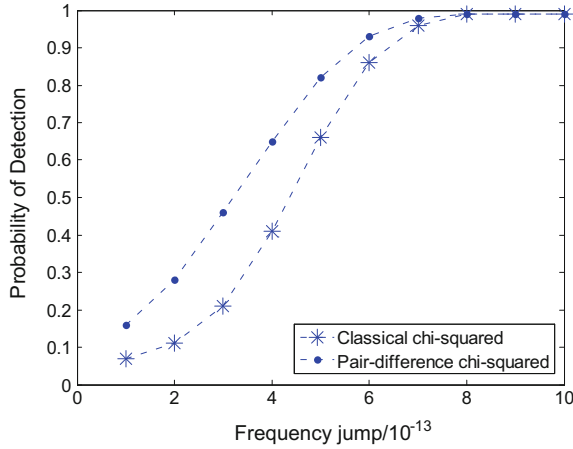


Fig. 3 The detection probability trend of different frequency jumps of two algorithms

Table 1 The chi-squared statistic of different frequency jumps of two algorithms

Frequency jump	1×10^{-13}	2×10^{-13}	3×10^{-13}	4×10^{-13}	5×10^{-13}
Classical	0.67	1.61	3.80	7.42	11.92
Pair-diff	2.67	4.96	8.01	11.82	16.41
Frequency jump	6×10^{-13}	7×10^{-13}	8×10^{-13}	9×10^{-13}	1×10^{-12}
Classical	17.89	25.09	33.55	43.26	54.22
Pair-diff	21.76	27.88	34.76	42.41	54.50

Table 2 The detection probability of different frequency jumps of two algorithms

Frequency jump	1×10^{-13}	2×10^{-13}	3×10^{-13}	4×10^{-13}	5×10^{-13}
Classical	0.07	0.11	0.21	0.41	0.66
Pair-diff	0.16	0.28	0.46	0.65	0.82
Frequency jump	6×10^{-13}	7×10^{-13}	8×10^{-13}	9×10^{-13}	1×10^{-12}
Classical	0.86	0.96	0.99	0.99	0.99
Pair-diff	0.93	0.98	0.99	0.99	0.99

algorithm is based directly on frequency differences of measurement data and the corresponding uncertainties, which can reflect the abnormal frequency of atomic clock. For the classical algorithm, the average frequency of atomic clocks need to be calculated, the uncertainties of atomic clocks need to be estimated, the differences between the frequencies and the average frequency also need to be calculated. So, the procedure is increased and the error is brought in, which lead to bad detection performance in classical method.

6 Conclusions

Compared with the classical test algorithm, the amount of calculation of pair-difference test algorithm is reduced. Because the average frequency of atomic clocks involved is calculated, and the frequency uncertainties are estimated, so the amount of calculation of classical method is large. For pair-difference test algorithm, the measurement uncertainties of the clock differences are only need to be calculated, the amount of calculation is small.

The detection performance is improved by using the proposed algorithm. When the frequency jump is smaller, the detection probability of the pair-difference algorithm is improved obviously and the abnormal frequency is detected easily. When the frequency jump is smaller than 3×10^{-13} , the detection probability is 1 times higher than that of the classical test algorithm. The detection probability increased by 20% when the frequency jump is 4×10^{-13} and 5×10^{-13} . When the frequency jump is larger, which is bigger than 6×10^{-13} , the detection probability of the two algorithms is getting same. Compared with the classical test algorithm, the pair-difference test algorithm can effectively detect the small frequency jump.

References

1. Niu F, Han CH, Zhang YS (2009) Analysis and detection on atomic clock abnormities of navigation satellite. *Geomat Inf Sci Wuhan Univ* 34(5):585–588
2. Galleani L, Tavella P (2012) Detection of atomic clock frequency jumps with Kalman filter. *IEEE Trans Ultrason Ferroelectr Freq Control* 59:504–509
3. Huang XM, Gong H, Zhu XW, Qu G (2014) Detection of weak frequency abnormities for atomic clocks with a Kalman filter. *Metrologia* 51:183–190
4. Riley WJ (2008) Algorithms for frequency jump detection. *Metrologia* 45:S154–S161
5. Galleani L (2008) Detection of changes in clock noise using the time-frequency spectrum. *Metrologia* 45:S143–S153
6. Galleani L, Tavella P (2008) Detection and identification of atomic clock abnormities. *Metrologia* 45:S127–S133
7. Nunzi E, Galleani L, Tavella P, Carbone P (2007) Detection of abnormities in the behavior of atomic clocks. *IEEE Trans Instrum Meas* 56:523–528
8. Galleani L, Tavella P (2009) The dynamic Allan variance. *IEEE Trans Ultrason Ferroelectr Freq Control* 56:450–464
9. Galleani L (2010) The dynamic Allan variance: II. A fast computational algorithm. *IEEE Trans Ultrason Ferroelectr Freq Control* 57:182–188
10. Sesia I, Galleani L, Tavella P (2011) Application of the dynamic Allan variance for the characterization of space clock behaviour. *IEEE Trans Aerosp Electron Syst* 47:884–895
11. Zucca C, Tavella P (2015) A mathematical model for the atomic clock error in case of jumps. *Metrologia* 52:514–521
12. Panfilo G, Tavella P (2008) Atomic clock prediction based on stochastic differential equations. *Metrologia* 45:S108–S116
13. Tavella P (2008) Statistical and mathematical tools for atomic clocks. *Metrologia* 45:S183–S192
14. Galleani L (2008) A tutorial on the two-state model of the atomic clock noise. *Metrologia* 45: S175–S182

15. Galleani L, Sacerdote L, Tavella P, Zucca C (2003) The mathematical model for the atomic clock error: an overview. *Metrologia* 40:S257–S264
16. Nunzi E, Carbone P (2008) Monitoring signal integrity of atomic clocks by means of the GLRT. *Metrologia* 45:S103–S107
17. Yang Y (1993) Robust estimation of variance components and its application in geodesy. IAG Meeting, Beijing
18. Douglas RJ, Steele AG (2006) Pair-difference chi-squared statistics for key comparisons. *Metrologia* 43:89–97
19. Mardia KV, Kent JT, Bibby JM (1979) *Multivariate analysis*. Academic, Longdon, p 521

Research on Robust Timing Receiver with Joint Verification of Location and Clock Offset



Dan Shen, Hang Gong, Zengjun Liu, Xiangwei Zhu and Feixue Wang

Abstract GNSS timing receiver is widely used in various fields. Unfortunately, it has weak capability of anti-jamming and anti-spoofing, so an integrated processing technique based on joint verification of location and clock offset is proposed. The algorithm is verified by simulation with timing receiver data. Simulation results show that 95% probability of the spherical error of location result can be used as the threshold of the location to detect strong intentional or unintentional jamming signal efficiently. When the location check is passed, the clock offset checking is used to monitor the receiver clock offset anomaly, with processing error less than 0.2 ns. The joint verification of location and clock offset method can greatly improve the robust performance of timing receivers.

Keywords Timing receiver · Location assistance · Clock offset assistance
Robust timing

1 Introduction

Due to wide coverage and high precision of GNSS, the GNSS timing receiver is widely used in the national infrastructures that require high-precision time reference, such as electricity, communication and transportation. GNSS signals are vulnerable to jamming and spoofing attacks during transmission, which would affect timing accuracy. While the timing receiver almost has no anti-jamming and anti-spoofing capability at present. Literature [1] proposed two ideas of interference monitoring of timing receiver, one is multi satellite clock difference aided method, the other is communication network assistance method. Literature [2] proposed a robust timing method of multi-layered multi-receiver architecture, and this method only present the anomaly monitoring flow and the verification summary; when

D. Shen · H. Gong (✉) · Z. Liu · X. Zhu · F. Wang
Engineering Research Center for Position, Navigation and Time, College of Electronic Science, National University of Defense Technology, Changsha 410073, Hunan, China
e-mail: gonghang@nudt.edu.cn

anomaly was detected, how to handle the anomaly and obtain time were not given; and the clock offset monitoring was considered as a secondary aid. These anomaly monitoring approaches of multi-level structure proposed above are too complicated to be applied in timing receiver.

2 Principle of Timing Receiver

GNSS timing receiver is comprised of five parts: antenna, RF front-end, baseband digital signal processing, and navigation signal solution and time generation. Wherein the antenna receives each frequency signal; the RF module carries out frequency down conversion and analog-to-digital conversion of the navigation signal; the baseband part finishes signal acquisition and tracking; the navigation signal resolving part uses pseudorange and carrier phase measurement to calculate location, clock offset.

The working mode of time generation module is determined by the result of abnormal clock offset monitoring. If in time keeping mode, it means that clock offset produced by information processing unit is invalid, the receiver's time would be maintained by crystal oscillator. When working in timing mode, it indicates that the clock offset can be used, and the disciplining module controls crystal oscillator to drive the generation of time-frequency signal and time information [2]. The system block diagram is shown in Fig. 1.

3 Principle of Timing Anomaly Monitoring

By analyzing the characteristics of timing receiver, this paper presents a robust timing approach with joint verification of location and clock offset. For most jamming and spoofing attacks, which would directly lead to large position deviation, can be monitored by the known precise location. When location monitoring passed, it does not mean that the receiver's clock offset is credible. Literature [3]

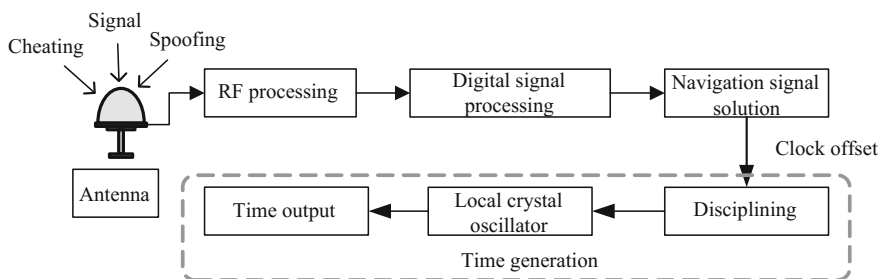


Fig. 1 System block diagram of satellite timing receiver

proposed a cheating and interference method that can affect the receiver’s clock offset without affecting the position. When receiver is influenced by external interference, power supply abnormality, receiver equipment failure, etc., it may lead to phase outlier and phase jump anomalies. These two anomalies of receiver clock offset are very common [4], and that are what this paper mainly deals with in clock offset monitoring. Figure 2 shows the workflow of stable timing method for joint verification of location and clock offset.

3.1 Location Monitoring Principle

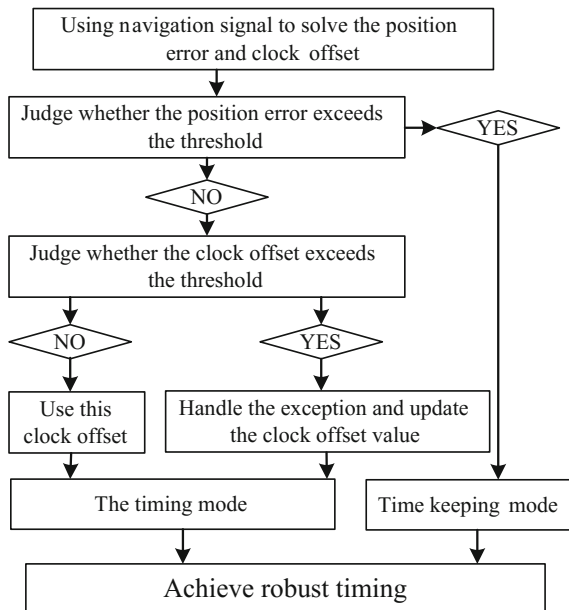
When GNSS signals are received by the receiver, the optimal combination of four satellite signals is selected to calculate the receiver’s position. If

$$\sqrt{(x_u - x_0)^2 + (y_u - y_0)^2 + (z_u - z_0)^2} \leq R_0 \tag{1}$$

Which means the distance error between the receiver’s position (x_u, y_u, z_u) and the exact position (x_0, y_0, z_0) is within the threshold range R_0 , that is to say, the signal is reliable under the location monitoring.

The choice of threshold R_0 is crucial. On the foundation of probability and statistics approach in [5], this paper sets the position threshold to be a spherical error of 95% probability of position error of a large amount of received data in a good environment under actual conditions.

Fig. 2 Workflow of robust timing receiver



3.2 Clock Offset Monitoring Principle

3.2.1 Clock Offset Prediction Model

Abnormal monitoring of receiver clock offset needs to establish a more accurate model of clock offset. According to [6], receiver clock offset can be modeled by using a generalized quadratic polynomial model if frequency source (atomic clock, crystal oscillator, etc.) satisfies the frequency offset is less than 1×10^{-6} . Since the current receiver uses oven-controlled crystal oscillators, whose accuracy can reach 10^{-9} level, short-term stability is less than 10^{-11} , the receiver clock offset model can be described as:

$$x(t) = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 + \varepsilon_x(t) \quad (2)$$

Which a_0 represents clock offset of the receiver at time t_0 ; a_1 represents frequency offset; a_2 represents frequency drift; $\varepsilon_x(t)$ is the random deviations. Equation (2) can be expressed in matrix form

$$X = Ha + e \quad (3)$$

Which X is n-dimensional observation vector, a is 3-dimensional unknown vector of parameters; e is n-dimensional error vector; H is $n \times 3$ -dimensional matrix.

$$H = \begin{bmatrix} h_1^T \\ h_2^T \\ \vdots \\ h_n^T \end{bmatrix} = \begin{bmatrix} 1 & t_1 - t_0 & (t_1 - t_0)^2 \\ 1 & t_2 - t_0 & (t_2 - t_0)^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n - t_0 & (t_n - t_0)^2 \end{bmatrix} \quad (4)$$

Then, we can get the least squares solution

$$\hat{a} = (H^T H)^{-1} H^T X \quad (5)$$

At present, the main methods of clock offset anomaly monitoring include: high-precision frequency or phase compared measurement, dynamic Allan variance and sliding window based least squares fitting algorithm, which are mainly used for satellite clock offset. However, they are hardly to meet the real-time requirements. Literature [7] proposed a recursive forgetting factor least squares (RFFLS) algorithm for real-time monitoring of clock anomaly, which can make full use history information of clock offset and enhance the impact of current data, and whose clock offset forecasting accuracy and time-consuming are better. So this algorithm is applied to anomaly monitoring of receiver clock offset in this paper, its flow is described as follows:

- (1) Initialization. Use m data ($m \geq 3$) to calculate the fitting parameters $\hat{\alpha}_m$ and T_m .

$$\begin{cases} \hat{\alpha}_m = (H^T W H)^{-1} H^T W X \\ T_m = (H^T W H)^{-1} \end{cases} \quad (6)$$

Which W represents the forgetting factor matrix, it can strengthen the contribution of new data and weaken the influence of old data. $W = \text{diag}(\lambda^{n-1}, \lambda^{n-2}, \dots, \lambda^0)$, $0 < \lambda \leq 1$, T_m is the weight vector for iterative updating.

- (2) Iteration updating. According to Eq. (7), calculate and store $\hat{\alpha}_k$ and T_k at the moment t_k iteratively.

$$\begin{cases} T_k = \frac{1}{\lambda} \left(T_{k-1} - \frac{T_{k-1} h_k h_k^T T_{k-1}}{\lambda + h_k^T T_{k-1} h_k} \right) \\ \hat{\alpha}_k = \hat{\alpha}_{k-1} + T_k h_k (\hat{x}_{RFFLS-t_{k-1}} - h_k^T \hat{\alpha}_{k-1}) \end{cases} \quad (7)$$

wherein, $\hat{x}_{RFFLS-t_{k-1}}$ represents clock offset processed by RFFLS algorithm at the moment t_{k-1} , which is determined by the algorithm decision criterion. $k = m + 1, m + 2, \dots, n$.

- (3) Calculate the predicted value of clock offset at the moment t_k :

$$\hat{x}_{t_k} = h_k^T \hat{\alpha}_k \quad (8)$$

3.2.2 Principle of Abnormal Clock Offset Monitoring Algorithm

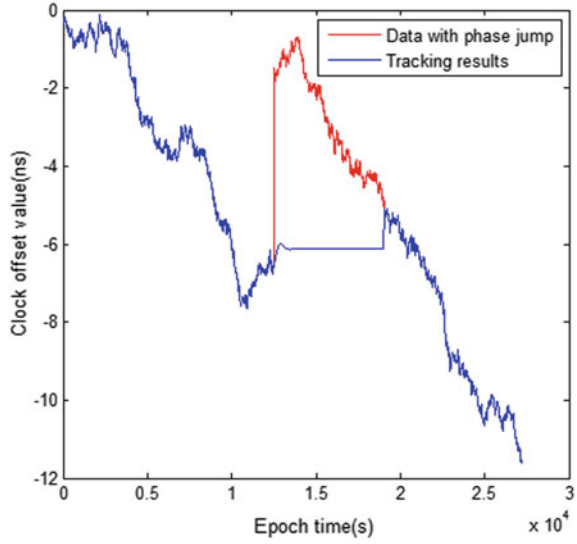
The real-time anomaly detecting algorithm of clock offset can achieve the aim of real-time detection of abnormal data, excluding and updating, by using the single-step predictive value of each input data to monitor clock offset.

The basic processes of real-time anomaly monitoring algorithm for clock offset are as follows:

- (1) Establish receiver clock offset model, and use recursive forgetting factor least squares algorithm to predict receiver's clock offset.
- (2) Calculate the difference between the predicted and the solved, then compare the monitoring threshold to judge whether an exception occurs or not. If real-time solution of clock offset is x_{t_k} , when $|\hat{x}_{t_k} - x_{t_k}| \geq T$, then x_{t_k} is an anomalous value and the processing unit would give an alarm; if not, the clock offset is considered without exception.

The selection of monitoring threshold T is very crucial as well. In this paper, we set the threshold T to be 95% confidence of all the results of single-step prediction error, which means that 95% of the single-step prediction error is less than T .

Fig. 3 Tracking performance of RFFLS algorithm for phase jump under simple criterion



- (3) For the removal and updating of anomalies, as in [7], when the difference is greater than T , the anomaly is removed and replaced by the predicted, then use this predicted to predict the next value. This judgment criterion is too simple that it is suitable for phase outlier, but not for phase jump, the effect is shown in Fig. 3.

Therefore, make the following improvements to step (3):

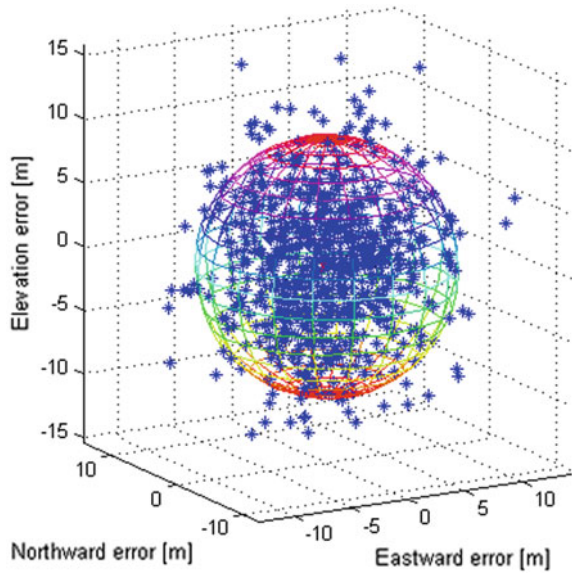
- (a) If $|\hat{x}_{RFFLS-t_{k-1}} - x_{t_{k-1}}| < T$ & $|\hat{x}_{t_k} - x_{t_k}| < T$, then $\hat{x}_{RFFLS-t_k} = x_{t_k}$. It means the clock offset is normal at this moment.
- (b) If $|\hat{x}_{RFFLS-t_{k-1}} - x_{t_{k-1}}| < T$ & $|\hat{x}_{t_k} - x_{t_k}| \geq T$, then $\hat{x}_{RFFLS-t_k} = \hat{x}_{t_k}$. It means the clock offset just encounter abnormal values.
- (c) If $|\hat{x}_{RFFLS-t_{k-1}} - x_{t_{k-1}}| \geq T$ & $|\hat{x}_{t_k} - x_{t_k}| < T$, then $\hat{x}_{RFFLS-t_k} = x_{t_k}$. It means the clock offset return to normal at this time.
- (d) If $|\hat{x}_{RFFLS-t_{k-1}} - x_{t_{k-1}}| \geq T$ & $|\hat{x}_{t_k} - x_{t_k}| \geq T$, then $\hat{x}_{RFFLS-t_k} = x_{t_k}$, this indicates that the clock offset is in the state of phase jump at this moment.

4 Simulation and Analysis

4.1 Location Monitoring Simulation Analysis

For the monitoring of location, the BDS timing receiver is adopted for experiments, it is fixed upon the open roof, and no high-power interference signal around. The position result of 800 data is shown in Fig. 4. The position threshold is adopted by 95% probability of spherical error, so $R_0 = 10.3$ m, the position accuracy is

Fig. 4 Positioning accuracy diagram



basically the same as that of BDS system, so 95% probability of spherical error threshold is credible.

In general, the location solution of receiver under interference of jammer is jammer’s position [1–3]. For strong intentional or unintentional signal jamming and deception, the position solution results deviate far from the exact position, exceeds 10 m range. Therefore, it is reasonable to use the 95% spherical error threshold of location. This method can effectively detect the position solution exceeds threshold caused by strong interference and spoofing signals, and thus the receiver enters time keeping mode to achieve a steady timing.

4.2 Clock Offset Monitoring Simulation Analysis

In order to verify the performance of clock offset detection algorithm adopted in this paper, the normal clock data is collected by BDS timing receiver is analyzed. Figures 5 and 6 show the tracking features and modeling accuracy of RFFLS algorithm for receiver clock offset, the RFFLS algorithm predicts the next moment each time based on the original data, it can be seen:

- (1) The RFFLS algorithm can reasonably model receiver clock offset, and the predicted value curve basically coincides with the real value curve.
- (2) The single-step prediction accuracy of RFFLS algorithm is basically within 0.2 ns, and the standard deviation of single-step prediction error is 0.0632 ns.

Fig. 5 Tracking performance of RFFLS algorithm

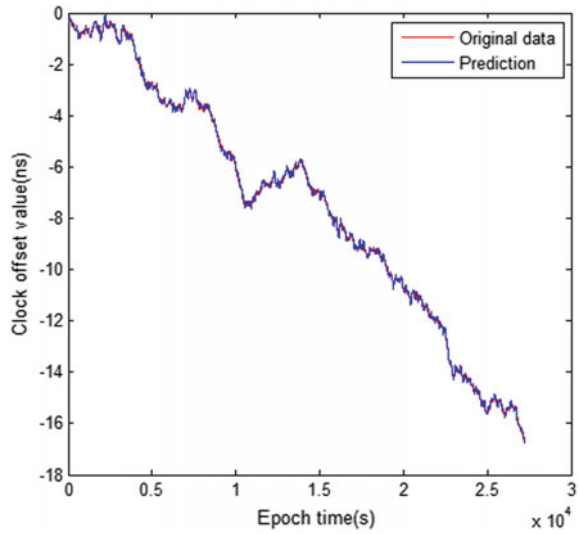
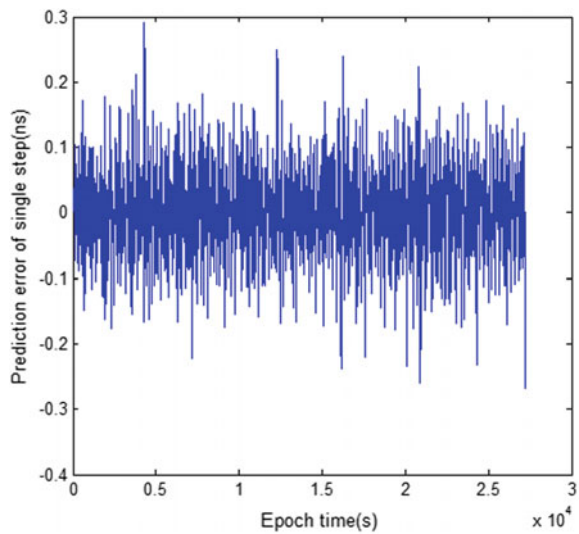


Fig. 6 Single-step prediction error graph of RFFLS algorithm for clock offset



The clock monitoring threshold T takes 0.15 ns in this paper. For the iterative updating of the algorithm, the algorithm would adopt the clock offset processed by RFFLS to predict the next moment. The following will analyze and verify the abnormalities of two common clock offsets. The abnormality is characterized by superimposing abnormal data on normal clock offset [4].

4.2.1 Phase Outlier

For phase outlier, the normal clock offset is the receiver clock offset mentioned earlier, and then add 5 ns, 20 ns, 38 ns, 56 ns, 15 ns, to the 378th, 858th, 1462th, 2780th, 3500th points, simulates clock offset that contains phase outlier. As shown in the red line in Fig. 7.

It can be seen from Fig. 7 that the algorithm processed curve almost matches the normal clock curve and does not keep up with the trend of outliers. Figure 8 is drawn from the results of $\hat{x}_{t_k} - x_{t_k}$, we can see that the algorithm can efficiently monitor phase outlier. Moreover, there is almost no large error value above the positive threshold value +T, which also shows that the predicted value quickly returns to normal value at the next moment of the abnormal.

4.2.2 Phase Jump Abnormalities

Based on the normal receiver clock offset data mentioned earlier, a phase jump of 5 ns is added at 2500th point and following points. For phase jump, the algorithm must be able to detect the change and the processed result must keep up with the trend of phase jump.

As it can be seen from Fig. 9, the RFFLS algorithm can effectively monitoring performance for phase jump and quickly follows the trend of phase jump. Figure 10 is drawn from the results of $\hat{x}_{t_k} - x_{t_k}$, we can see that, RFFLS algorithm can also monitor phase jump well. Above the positive threshold +T, existing individual values exceed the threshold is due to the fact that original data varies greatly during this period and the algorithm has to keep up with stepwise trend at the moment.

Fig. 7 The tracking performance of the RFFLS algorithm eliminating phase outlier anomalies

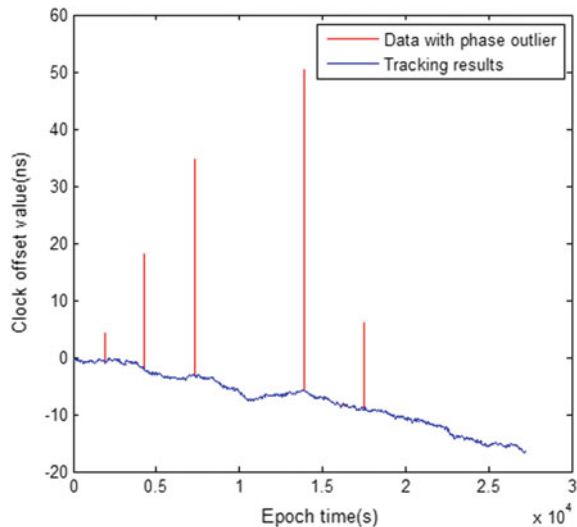


Fig. 8 The difference between RFELS predicted and the original with phase outlier anomalies

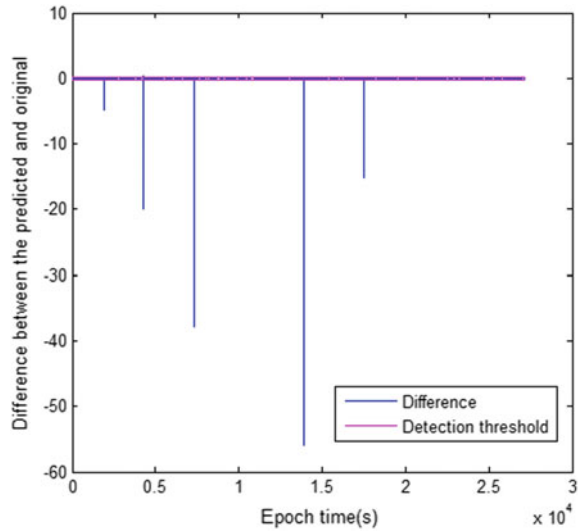
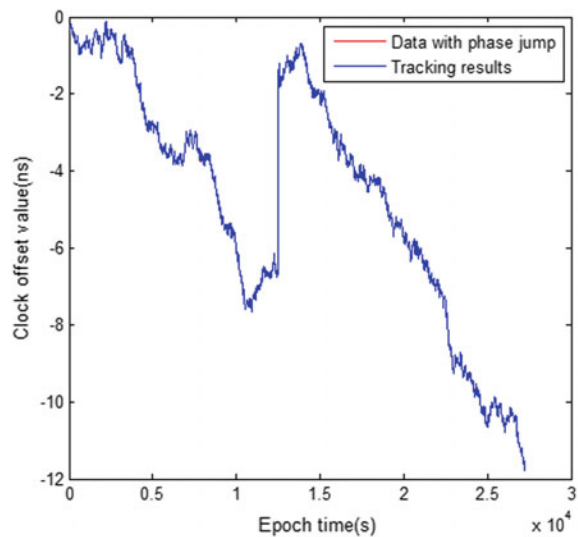


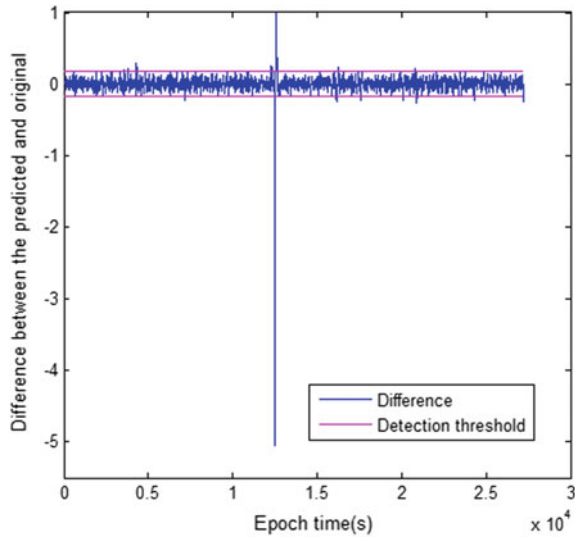
Fig. 9 The tracking performance of the RFELS algorithm eliminating phase jump anomalies



From Figs. 9 and 10, we would know that the algorithm can be more ideal for monitoring phase jump abnormalities of receiver clock offset.

For phase outlier and phase jump, the algorithm can detect abnormal values effectively in real time and eliminate anomalies, quickly returning to the normal trend of data, thus providing valid clock offset data to disciplining the oven-controlled crystal oscillators to ensure the receiver timing stably.

Fig. 10 The difference between RFFLS predicted and the original with phase jump anomalies



5 Conclusion

Since timing receiver almost does not have the ability of anti-jamming and anti-spoofing, this paper presents a robust timing approach with joint verification of location and clock offset. Simulation results show that 95% of receiver’s location results are reliable as location thresholds, and strong intentional or unintentional interference spoofing signals can be detected in location domain that timing receivers are in time keeping mode. When location checking is passed, the recursive forgetting factor least squares algorithm can effectively detect phase outlier and phase jump in clock offset domain checking. The updated clock offset can reasonably disciplining the crystal oscillator for robust timing, and the processing error of RFFLS is less than 0.2 ns. Compared with those complex multi-level structures approaches, this method can greatly improve robust timing performance of timing receivers.

References

1. Zhu X, Wu Y, Gong H et al (2015) GNSS timing receiver toughen technique in complicated jamming environments. *J Natl Univ of Def Technol* 37(3):1–9
2. Heng L, Mekala J, Dominguez-Garcia AD et al (2014) Reliable GPS-based timing for power systems: a multi-layered multi-receiver architecture. In: *Power and energy conference at Illinois, PECEI, 2014*, pp 1–7
3. Jiang X, Zhang J, Harding BJ, Makela JJ et al (2013) Spoofing GPS receiver clock offset of phase measurement units. *IEEE Trans Power Syst* 28(3):3253–3262

4. Gong H (2013) Study on frequency stability estimation techniques for GNSS on-board clock based on single station observation. Graduate School of National University of Defense Technology
5. Kaplan ED, Hegarty CJ (2008) Understanding GPS principles and applications, 2nd edn. Artech House, Boston, London
6. Misra P, Pratt M, Muchnik B et al (1995) A general RAIM algorithm based on receiver clock. In: Proceeding of the ION GPS proceedings of the 1995 8th international technical meeting of the satellite division of the institute of navigation, pp 1941–1947
7. Yu H, Hao J, Liu W et al (2016) A real-time anomaly monitoring algorithm for satellite clock. *Geomat Inf Sci Wuhan Univ* 41(1):106–110

The Long-Term Performance Analysis and Comparison of BDS on-Board Atomic Clock Based on Precise Clock Products of IGS Analysis Centers



Huijie Xue, Tianhe Xu, Qingsong Ai and Shengchao Wang

Abstract As an important load of BDS, the performance of satellite-borne atomic clock directly affects the service level of navigation system. Based on the requirements of navigation system on atomic clocks in orbit, the performance of BDS spaceborne atomic clocks has been studied in recent years. Precise clock products of two IGS analysis centers wum and gbm with sampling interval of 300 s from 2014 to 2017 were adopted. The BDS spaceborne atomic clocks were evaluated and analyzed from accuracy, drift rate and stability respectively. The results indicate that, overall, MEO satellite clock performance is the best, followed by IGSO satellite clock, the worst is GEO satellite clock; the accuracy, the daily drift rate and the 10,000 s stability fluctuate at 7×10^{-12} – 4×10^{-11} , -3×10^{-13} – 3×10^{-13} , 3×10^{-14} – 1.5×10^{-13} respectively. In terms of drift rate and the 10,000 s stability, the results of performance analysis of GEO satellite clock using two different precise clock products have great differences, while those of IGSO and MEO satellite clocks have little difference.

H. Xue · S. Wang

School of Geology Engineering and Surveying, Chang'an University,
No. 126, Yanta Road, Xi'an, Shanxi, China
e-mail: 1357643345@qq.com

S. Wang

e-mail: 364693662@qq.com

T. Xu (✉)

Institute of Space Science, Shandong University, Weihai, Shandong, China
e-mail: thxugfz@163.com

T. Xu

State Key Laboratory of Geo-Information Engineering, Xi'an, Shanxi, China

Q. Ai

State Key Laboratory of Geodesy and Earth's Dynamics,
Institute of Geodesy and Geophysics, Wuhan, China
e-mail: 1305583364@qq.com

Q. Ai

University of Chinese Academy of Sciences, Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_48

Keywords BDS · Satellite atomic clock · Accuracy · Drift · Stability
wum · gbm

1 Introduction

Beidou Navigation System (BDS) is currently under construction. Studying the performance analysis of BDS satellite atomic clock is of great significance to improve the service performance of the system and the next generation system construction and maintenance. Based on the data of satellite clocks, Huang et al analyzed the variation law of phase, frequency, frequency drift and short-term stability of BDS satellite-borne atomic clocks [1]. Wang et al conducted a long-term performance analysis of BDS satellite-borne atomic clocks [2]. Li et al conducted a performance analysis of the BDS satellite atomic clock based on the dynamic Allen variance [3]. At present, the performance analysis of BDS spaceborne atomic clock is still limited and the research on the performance analysis of BDS satellite atomic clock mainly focuses on using one analysis center precise clock products. The data length is generally about one year, and a long-term analysis should be conducted. At present, there is little research on the analysis of long-term performance analysis of BDS spaceborne atomic clocks and comparative analysis of using different precise clock products. Based on this motivation, we use the precise clock bias products (Product names are: wum*****.clk and gbm *****.clk) with the sampling interval of 300 s from 2014.01.01 to 2017.06.27, which are released by Wuhan University and the German Geoscience Research Center, respectively, for BDS satellite clock performance analysis. Based on the variation of frequency stability, frequency-daily drift rate and frequency accuracy, the long-term performance of BDS satellite atomic clock is comprehensively evaluated. In addition, the differences of BDS satellite clock long-term performance analysis between the precise clock products of IGS analysis centers wum and gbm are compared, and some useful conclusions are drawn.

2 Algorithm Principle

2.1 Data Preprocessing Method

Gross errors often occur during the satellite clock bias data acquisition. In this paper, the data preprocessing is performed by a median method (MAD) with good robustness and timeliness. Median MAD can be expressed as [4]:

$$MAD = Median\{|y_i - m|/0.6745\} \quad (2.1)$$

where m is the median of the time series, that is, $m = Median \{y_i\}$; When the $|y_i| > (m + n \times MAD)$ (n is valued as 5 in this article), it is considered an outlier.

In this paper, the clock bias sequence is obtained from the precise clock file per day. When there are some missing data in the whole day, interpolation is conducted to get the missing items. The processed complete clock bias sequence is carried out by the first difference and divided by the sampling interval to transform into frequency data. The MAD is used to detect the gross error, after eliminating these outlier points, linear interpolation is used to fill up [5].

2.2 Frequency Accuracy Calculation Method

The frequency accuracy characterizes the correspondence between the measured frequency or the calculated frequency and the defined value of the frequency, and the formula is:

$$A = \frac{f_x - f_0}{f_0} \quad (2.2)$$

where A is the frequency accuracy, f_0 is the nominal frequency of the measurement scale, and f_x is the actual frequency measurement or calculated value. In satellite clock performance evaluation, we can not get the difference between the nominal frequency and the actual frequency of the satellite clock. The frequency accuracy is usually calculated by the method of time difference comparison. The formula is [6]:

$$A = \frac{-\Delta t}{T} = \left| \frac{\hat{x}(t_2) - \hat{x}(t_1)}{t_2 - t_1} \right| = \left| \frac{\hat{x}(t + \tau) - \hat{x}(t)}{\tau} \right| \quad (2.3)$$

where Δt is the time difference of the adjacent interval observations; T is the length of the observation period, and \hat{x} is the clock bias data. According to the above formula, we can get multiple values of the frequency accuracy, in this paper, the average of the absolute value is used as the accuracy index of the final results.

2.3 Frequency Drift Rate Calculation Method

Atomic frequency standard in operation, due to the aging of internal components, as well as environmental factors and load energy changes, its output frequency often monotonically increase or decrease with the running time. In general, the rate of the frequency values monotonically change with time is called the frequency drift rate. It can be expressed as [7]:

$$D = \frac{\sum_{i=1}^N (y_i - \bar{y})(t_i - \bar{t})}{\sum_{i=1}^N (t_i - \bar{t})^2} \quad (2.4)$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad \bar{t} = \frac{1}{N} \sum_{i=1}^N t_i \quad (2.5)$$

where D is the frequency drift rate, \bar{y} is the mean value of the relative frequency value y_i , t_i is the measurement time, and N is the number of samples.

2.4 Frequency Stability Calculation Method

The time service stability of atomic clock is a key factor to determine the performance level of the user's real-time position [8, 9]. Hadamard variance is a three-sample variance, because it can effectively counteract the impact of linear frequency drift on satellite stability estimates, it is suitable for the long-term stability analysis of the orbital rubidium atomic clock. In this paper, the standard Hadamard variance is used and the formula for the frequency data is [10]:

$$H\sigma_y^2(\tau) = \frac{1}{6(M-2)} \sum_{i=1}^{M-2} [\bar{y}_{i+2} - 2\bar{y}_{i+1} + \bar{y}_i]^2 \quad (2.6)$$

where τ is the smoothing time, M is the number of samples, and \bar{y}_i is the mean value of relative frequency deviation in the corresponding period of time.

3 BDS Satellite Clock Performance Analysis

The long-term changes in the stability of 10,000 s and daily drift rates of Beidou satellite of two clock bias products are shown in Figs. 1 and 2. The blue line corresponds to the wum clock bias product, and the red corresponds to the gbm clock bias product. Daily clock bias data is used to calculate the 1000 s stability (smoothing time is 900 s), the 10,000 s stability (smoothing time is 9900 s), daily drift rate, and accuracy, as shown in Figs. 3 and 4. One month average is used for the calculation of the stability of day, as shown in Figs. 5 and 6. The statistical results of performance analysis are shown in Table 1.

3.1 The 10,000 s Stability Long-Term Changes

It can be seen from Fig. 1 that:

- (1) In general, the 10,000 s stability of 14 satellite atomic clocks maintains between 4×10^{-14} and 4×10^{-13} . Except for the C02, C06 and C08, which have large fluctuations, others are relatively stable.

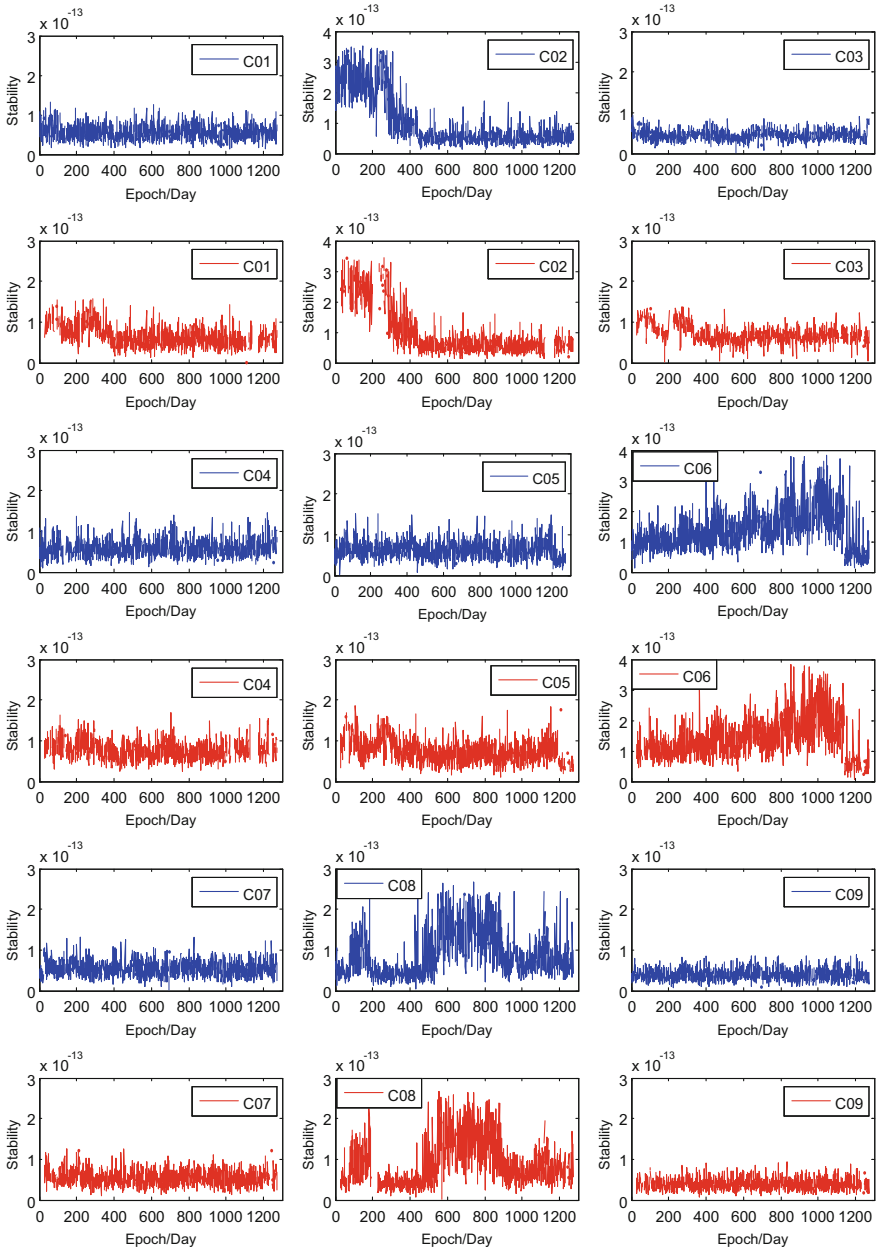


Fig. 1 BDS spaceborne atomic clock 10,000 s stability long-term changes

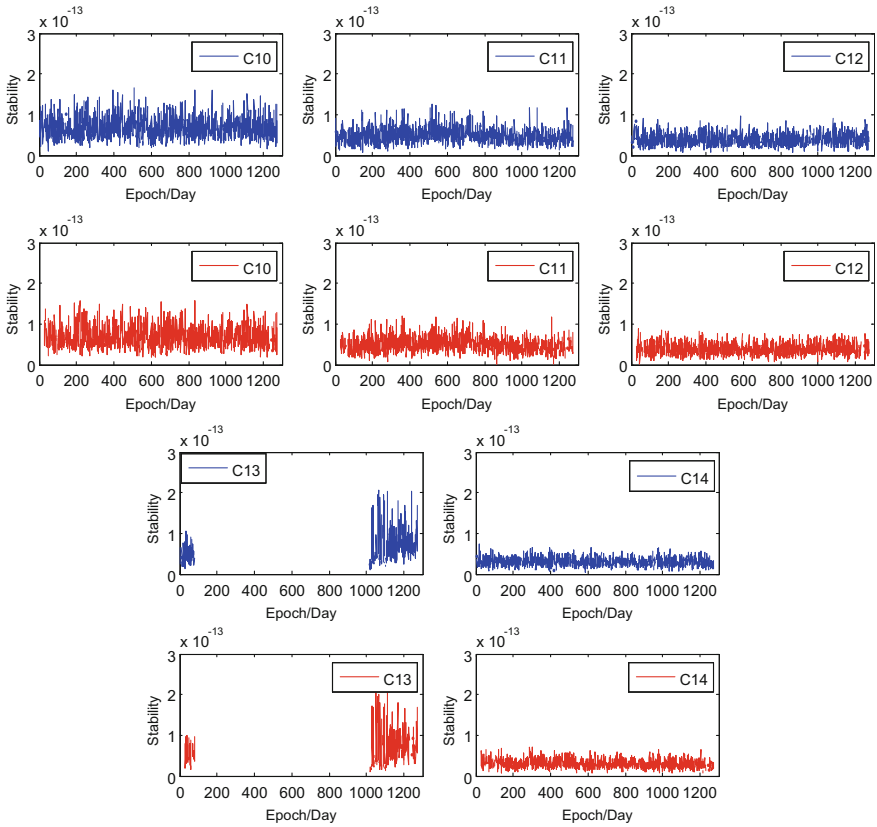


Fig. 1 (continued)

- (2) In early 2014, the 10,000 s stability of C02 kept at 2.5×10^{-13} and gradually decreased to about 5×10^{-14} from the late stage of 2014 and remained stable thereafter. The 10,000 s stability of C06 gradually increases from 1×10^{-13} to about 2.5×10^{-13} , and then there is a big jump in 2017, the 10,000 s stability downs to about 5×10^{-14} . The 10,000 s stability of C08 has a clear periodicity and shows two peaks, the cycle is about three years, the first peak keeps at about 1×10^{-13} , the second peak retains at about 2×10^{-13} , the minimum value maintains at about 5×10^{-14} .
- (3) The 10,000 s stability of C01, C03, C04 and C05 calculated by using the gbm clock bias product in 2014 is significantly higher than those by using the wum clock bias product in 2014. Moreover, the 10,000 s stability of this 4 satellites calculated by using the former product has obvious periodicity in 2014, and then tends to be stable. For other satellites (except C02), the 10,000 s stability calculated by using two clock bias products has better consistency.

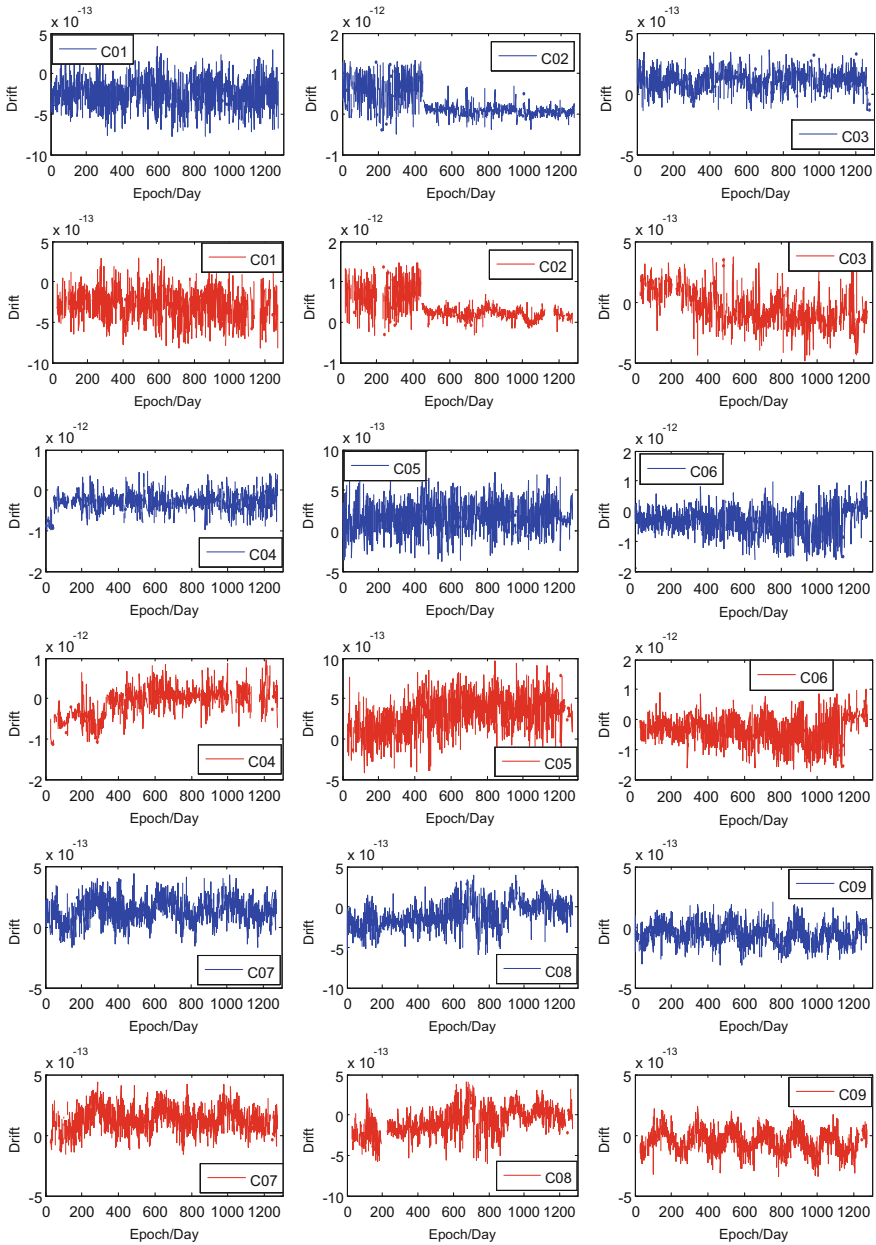


Fig. 2 Long-term changes in daily drift rate of BDS spaceborne atomic clock

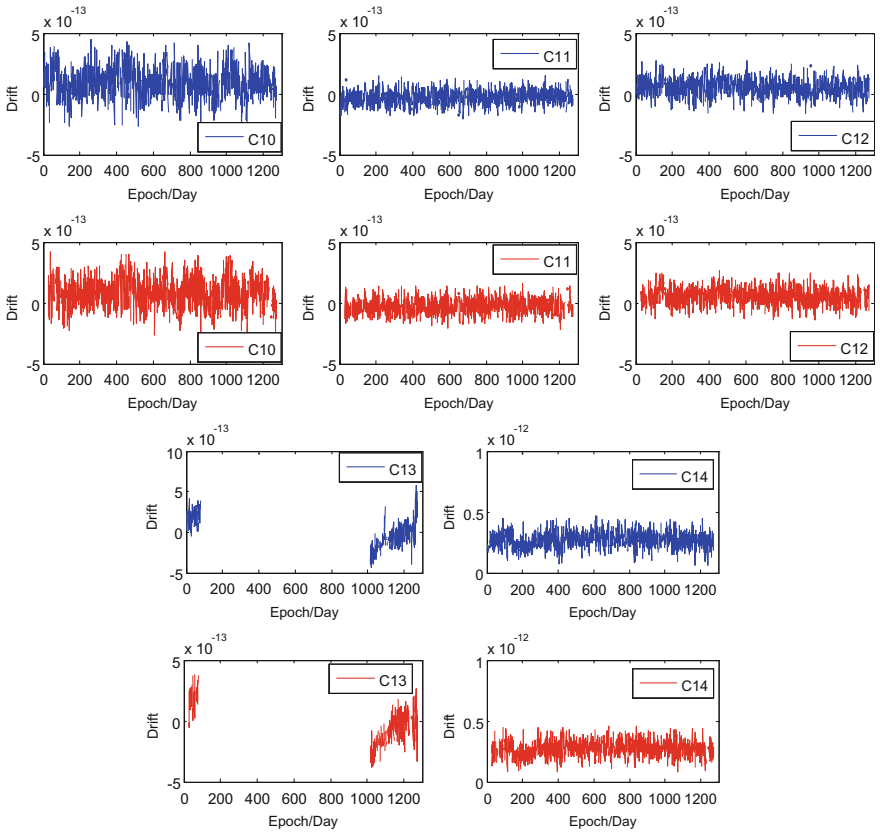


Fig. 2 (continued)

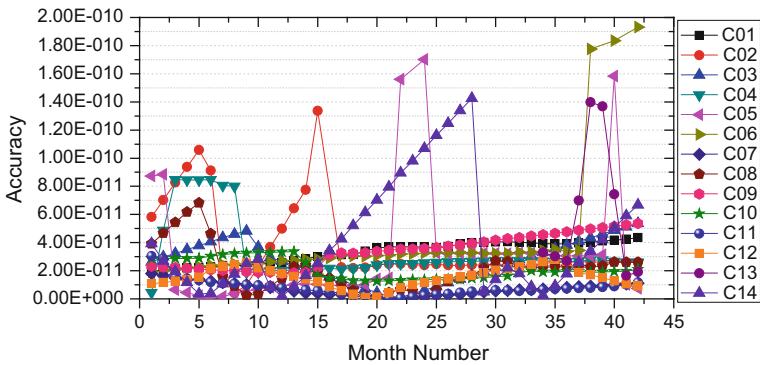


Fig. 3 BDS spaceborne atomic clock accuracy of the long-term changes using wum clock bias product

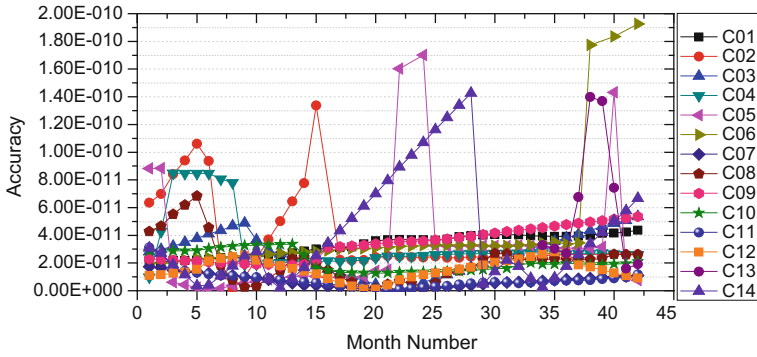


Fig. 4 BDS spaceborne atomic clock accuracy of the long-term changes using gbm clock bias product

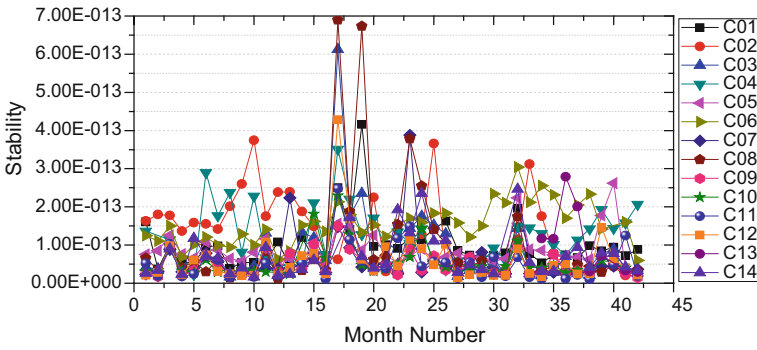


Fig. 5 BDS spaceborne atomic clock day stability long-term changes using wum clock bias product

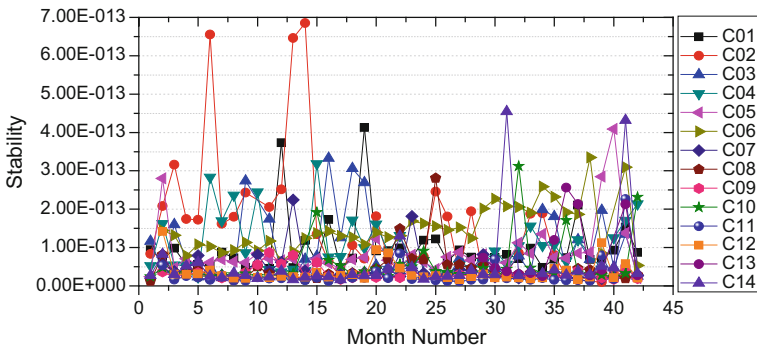


Fig. 6 BDS spaceborne atomic clock day stability long-term changes using gbm clock bias product

Table 1 BDS satellite clock performance indicators results charts

Satellite type	Satellite	Accuracy (E-11)		Daily drift rate (E-13)		1000 s stability (E-13)		10,000 s stability (E-14)		Day stability (E-14)	
		wum	gbm	wum	gbm	wum	gbm	wum	gbm	wum	gbm
GEO	C01	3.317	3.314	-2.402	-2.798	1.641	1.639	5.511	6.567	10.010	9.835
	C02	3.790	3.843	2.570	3.495	1.390	1.387	10.150	9.540	12.660	15.580
	C03	2.680	2.683	1.088	-0.487	0.851	0.837	4.449	6.948	8.141	9.367
	C04	3.299	3.288	-2.834	-0.854	1.324	1.325	5.895	7.673	12.420	10.640
	C05	3.177	3.156	1.802	3.065	1.566	1.569	6.256	7.484	8.869	8.922
	Mean	3.253	3.257	2.139	2.140	1.354	1.351	6.453	7.642	10.420	10.870
IGSO	C06	4.081	4.080	-3.457	-3.628	2.007	2.014	13.780	14.020	15.440	14.940
	C07	0.709	0.702	1.362	1.293	1.511	1.515	5.422	5.398	5.869	4.290
	C08	2.084	2.091	-0.944	-0.859	2.990	3.046	8.716	8.879	9.930	4.420
	C09	3.384	3.381	-0.547	-0.528	1.223	1.225	3.923	3.898	4.820	3.111
	C10	2.104	2.101	0.911	0.811	2.322	2.317	6.865	6.857	5.528	5.496
	Mean	2.472	2.471	1.444	1.424	2.011	2.023	7.741	7.810	6.991	6.451
MEO	C11	0.793	0.789	-0.276	-0.281	2.460	2.481	4.924	4.972	5.182	2.866
	C12	1.536	1.535	0.562	0.563	1.105	1.107	3.877	3.870	6.153	3.479
	C13	4.873	4.869	0.109	-0.361	1.698	1.619	6.711	6.893	9.167	9.478
	C14	4.226	4.200	2.709	2.729	0.744	0.751	3.014	3.067	7.051	5.457
		Mean	2.857	2.938	0.914	0.983	1.502	1.489	4.632	4.701	5.081

3.2 *The Daily Drift Rate Long-Term Changes*

It can be seen from Fig. 2 that:

- (1) As a whole, the daily drift rate of 14 satellites basically maintains between -3×10^{-13} and 3×10^{-13} .
- (2) The daily drift rate of C06, C07 and C09 has obvious periodicity, the period of C06 is about 100 days, the period of C07 and C09 is about 200 days. In 2014, the daily drift rate of C02 kept at about 5×10^{-13} and was very unstable. And then a jump occurred, the daily drift rate significantly reduces and becomes stable, maintaining at about 3×10^{-13} . The daily drift rate of MEO satellites (except C13) is more stable than other satellites. The daily drift rate of C08 appears multiple jumps, overall it is gradually rise, but its absolute value has little change. The fluctuation range of the daily drift rate of C06 became bigger and bigger before 2017.
- (3) The daily drift rate of C03 calculated with wum clock bias product basically maintains at about 1×10^{-13} , while the daily drift rate of C03 satellite calculated with gbm clock bias product maintained at about 1×10^{-13} in early 2014, then gradually reduced to -1×10^{-13} in late 2014. For C04 atomic clock, the daily drift rate calculated with wum clock bias product tends to be stable, however, the daily drift rate calculated with gbm clock bias product rises after a jump in the middle, and the absolute value of the daily drift rate does not obviously increase.
- (4) The daily drift rate of C05 has raised, indicating that the satellite clock is aging with time. The rise of daily drift rate of C05 calculated with gbm clock bias product is more obvious than that of C05 calculated with wum clock bias product.

3.3 *Accuracy Long-Term Changes*

It can be seen from Figs. 3 and 4 that:

- (1) Overall, the accuracy of 14 satellites basically maintain between 1×10^{-11} and 4×10^{-11} . For C02, C04, C06, C13 and C14, the accuracy has a relatively large fluctuation. The accuracy of other satellites is relatively stable.
- (2) The accuracy of C05 appears several large jumps. After two jumps occurred around November 2015, the accuracy became significantly larger. The accuracy of C14 increased roughly linearly between December 2014 and April 2016, and then a jump occurred in April 2016, the accuracy dropped to about 1×10^{-11} . In October 2016, the accuracy of C14 increased again roughly linearly.
- (3) The accuracy of C01, C05, C06 and C09 has a gradual increasing trend, the aging is obvious. Overall the accuracy of C03, C07 and C11 shows a gradual

decline and then a gradual upward trend. The accuracy of C08 and C12 shows obvious periodicity, and the period of accuracy of C12 is about 20 months.

- (4) The accuracy of each satellite calculated by wum and gbm clock bias products has good consistency.

3.4 Day Stability Long-Term Changes

It can be seen from Figs. 5 and 6 that:

- (1) As a whole, the day stability of 14 satellites basically kept between 3×10^{-14} and 4×10^{-13} . The day stability of C02, C04 and C06 is poor. The fluctuation of day stability of C02 and C04 is more frequent than that of other satellites.
- (2) The day stability of C06 satellite shows a gradual upward and then gradual decline trend.
- (3) The fluctuation of day stability of C03 calculated with gbm clock bias product is stronger than the fluctuation of day stability of C03 calculated with wum clock bias product.
- (4) The day stability of each satellite calculated by wum clock bias products shows multiple fluctuations. However, the day stability of each satellite calculated by gbm clock bias product shows more stable.

4 Conclusions

- (1) Among the three types of satellite atomic clocks, the IGSO satellite clock is the most accurate, followed by the MEO satellite clocks, the accuracy of GEO satellite clock is the worst. As far as the short-term stability is concerned, the MEO satellite clock and the GEO satellite clock are not much different and both are superior to the IGSO satellite clock. For 10,000 s stability, the worst is the IGSO satellite clock, followed by the GEO satellite and the MEO satellite clock is the best. As far as the overall performance is concerned, the C06 atomic clock is the worst. For day stability and daily drift rate, the performance of GEO satellite clock is relatively poor, followed by the IGSO satellite clock, the best is the MEO satellite clock, while in MEO satellite clock, the performance of the C13 satellite is obviously worse than other satellites clocks.
- (2) For 1000 s stability and accuracy, the results of each satellite calculated with two kinds of clock bias products have good consistency and the differences are very small. For the GEO satellite clock, daily drift rate and 10,000 s stability of each satellite calculated with two kinds of clock bias products have relatively large differences. The differences of daily drift rate are all above 4×10^{-14} , The differences of 10,000 s stability are greater than 6×10^{-15} . For the IGSO and

MEO satellite clocks, daily drift rate and 10,000 s stability of each satellite calculated with two kinds of clock bias products have small differences. The differences of daily drift rate (except C06, C13) are less than 1×10^{-14} . The differences of 10,000 s stability maintain within 2.4×10^{-15} . Using two clock bias products to calculate the day stability, the difference of C08 clock is the largest, reaching 5.5×10^{-14} , and other satellites' differences are between 1×10^{-14} and 3×10^{-14} . On the whole, day stability of the three kinds of satellite clocks calculated with two kinds of clock bias products have big differences.

- (3) The reason of the above differences may be related to factors such as orbit type, station selection, data processing strategy of analysis center, quality of clock bias products and the length of the smooth time. More BDS satellite clock products of IGS analysis centers as well as iGMAS analysis centers should be included into the performance analysis in the further work.

Acknowledgements The research was supported by National Natural Science Foundation of China (Grant No. 41574013, 41731069) and National Key Research and Development Program of China (2016YFB0501701).

References

1. Huang G, Yu H, Guo H et al (2017) Analysis of the long-term clock bias characteristics of the Beidou satellite clock. *J Wuhan Univ (Inf Sci Edit)* 42(7):982–988
2. Wang Y, Lv Z, Wang N (2017) BDS spaceborne atomic clocks long-term performance analysis. *J Surv Mapp* 46(2):157–169
3. Li Z, Mi K, Zhou P (2016) Analysis of Beidou satellite's atomic clock based on dynamic Allan variance. *Beijing Surv Mapp* 2:36–39
4. Yu H (2015) Research on BDS/Galileo spaceborne atomic clock performance and short-term forecast algorithm. Chang'an University
5. Ai Q, Xu T, Sun D et al (2016) Forecasting Beidou satellite clock bias with consideration of cycle term error and starting point deviation correction. *Acta Surv Mapp* 45(S2):132–138
6. Sun D (2016) GNSS star borne atomic clock performance analysis theory and algorithm research. Chang'an University
7. Jia X, Feng L, Mao Y et al (2010) Performance evaluation of GPS spaceborne atomic clocks. *Time Freq J* 33(2):115–120
8. Liu S, Jia X, Sun D (2017) GNSS spaceborne atomic clocks performance analysis. *J Wuhan Univ (Inf Sci Edit)* 42(2):277–284
9. Chen K, Xu T, Yang Y et al (2016) Synthesis and evaluation of iGMAS GNSS clock bias products. *J Mapp* 45(s2):46–53
10. Wang Y, Lv Z, Li L et al (2017) Analysis of long-term performance of GPS BLOCK IIF spaceborne atomic clock. *J Astron* 58(3):11–21

A Method on Laser Power Stabilization in Optical Detection Cesium Atomic Clock



Ziyu Chen, Chang Liu, Shuhui Wang and Yanhui Wang

Abstract Laser detection avoid the limitation on cesium clocks' lifespan due to electron multipliers. However, the laser inevitably induces light shift. By measuring the frequency shifts under various light powers, we study the light shift. The dependence of the light shift on the light power is $2 \times 10^{-12}/\text{mW}$. The Allan deviation of the output power of the laser diode is 4.5×10^{-3} at 10^5 s under free running. The temperature coefficient of the clock is $1 \times 10^{-12}/^\circ\text{C}$. For a high-performance cesium clock, the light shift is one of the main restricts to accuracy and long-term stability. In order to improve the clock's long-term frequency stability, accuracy and temperature coefficient, we propose a method for laser power stabilization. This method uses a liquid crystal variable retarder to tune the polarization of the light, in order to feedback control the light power with respect to the photoelectric detector. We theoretically analyze the method, and find that the drift of light power mainly results from the temperature drift of electronic elements. The temperature drift is below $1 \times 10^{-5}/^\circ\text{C}$, manifesting the theoretical feasibility. We build an experimental system and measure the light power for 20 days. The stability of light power is 6×10^{-5} at 10 s, and 3×10^{-6} at 10^5 s. Moreover, the temperature coefficient reduces to $2 \times 10^{-13}/^\circ\text{C}$ from $1 \times 10^{-12}/^\circ\text{C}$. Therefore, the method effectively improves the performance of the clock. Besides, it can be easily applied to other experiments that requires laser power stabilization.

Keywords Atomic clock · Light shift · Laser power stabilization
Stability · Liquid crystal variable retarder

Z. Chen · C. Liu · S. Wang · Y. Wang (✉)
School of Electronics Engineering and Computer Science,
Peking University, Beijing, China
e-mail: wangyanhui@pku.edu.cn

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_49

607

1 Introduction

Since the second was defined based on the transition frequency between the two hyperfine levels of the ground state of the ^{133}Cs atom in 1967, International Atomic Time has been widely used as a time scale. It is kept by over 400 atomic clocks, the majority of which are Cs clocks, because Cs clocks have outstanding accuracy and long-term stability. Nowadays, Cs clocks have been applied widely in many fields including time-keeping, communication, electric power, and navigation.

Cs clocks use Ramsey's separated-field method that can effectively narrow the linewidth of the microwave spectral line and improve the short-term stability. On the other hand, Cs clocks use atomic beam that can eliminate the collisions between atoms, hence improving the long-term stability. There are two main types of the Cs clocks, the magnetic-state-selection clocks and the optical pumping clocks. The conventional Cs clocks use the magnetic-state-selection scheme, where magnets select the atomic states and electron multiplier counts the number of the atoms. The advantages of the conventional Cs clocks are the simplicity of the Cs tubes, high reliability and high environmental suitability. Moreover, the velocity distribution of the selected atoms is narrow which also narrows the linewidths of the microwave spectral lines. The disadvantage is the low utilization of the atomic beam, which limits the signal to noise ratio. Besides, the technical barrier of the electron multipliers hinders the researches of the commercially available conventional Cs clocks. On the other hand, optical pumping scheme uses lasers for the state selection and detection, which largely enhance the utilization of the atomic beam. The utilization ratio even reaches 100% in the double-laser scheme. Since the optically pumped Cs clocks were proposed in 1970s, many researches have been made. The optically pumped Cs frequency standard was reported in USA in 2001 [1]. It has a short-term stability of $1.2 \times 10^{-12} \tau^{-1/2}$, and the Allan deviation reaches 3.8×10^{-15} at averaging time of a day, which is the best performance of small Cs clocks. For the consideration of space-borne clocks, many researches focuses on the optically pumped Cs clocks with single laser. One of the best results of this type of Clocks is reported by Lecomte et al. from Switzerland [2]. They acquire a short-term stability of 1.14×10^{-12} and $1.5 \times 10^{-12} \tau^{-1/2}$ respectively, by using the D1 and D2 lines of Cs for the optical pumping and detection. The performances are close to those of the laboratory Cs frequency standards. However, the complicated, low-reliability and costly laser system limits the performance of the clocks. And the light shift is inevitable in this scheme and has become one of the main barriers for the long-term stability of the clocks. As a result, there are no commercial products of the optically pumped Cs clocks at present.

Recently, there have been many researches on small Cs clocks in China. Based on the comparisons of conventional Cs clocks and optically pumped Cs clocks, we proposed a new scheme [3], which uses both the magnetic state selection and the optical detection. This method avoids the limitation on the cesium clocks' lifespans due to the electron multipliers. Meanwhile, it maintains the advantages of the

conventional Cs clocks, i.e. narrow microwave spectral lines and low complexity, although the light shift still exists in this scheme.

Light shift is a phenomenon that the atomic energy levels are shifted by light interaction. According to quantum theory, the light shift is proportional to the intensity of the light field. In our method, the light shift mainly stems from the fluorescence, the scattered light from the detection, and the stray light from the environment. The stray light from the environment can be eliminated by a careful design of light blocking. As a result, the dependence of light shift on the detection light power represents the light shift's influence on the clock's performance. The current coefficient of light shift with respect to the detection light power is $2 \times 10^{-12}/\text{mW}$. For high performance clocks, the light shift has become one of the main limitations on the accuracy and the long-term stability. Moreover, the laser diode is quite sensitive to the environmental temperature. In order to improve the accuracy and long-term frequency stability the clock, we need effective control of the light shift. One of the most effective ways is to control the light power. Based on the variable-phase property of liquid crystal retarders [4], we propose a method that can stabilize the light power. The method uses a liquid crystal retarder for controlling the laser polarization, and feedback controls the light power according to the signal obtained by a photodiode. With this method, the laser power and light shift can be well regulated.

2 Laser Power Stabilization Scheme Based on Liquid Crystal Variable Retarder

Liquid crystal variable retarder uses a liquid crystal cell to function as a variable wave plate. Due to the birefringence of the liquid crystal material, the light polarization can be modulated by the voltage applied to the cell. When placed between a pair of polarizers with perpendicular Optic axes, the output light power can be varied continuously.

The experimental setup of the optical system is shown in Fig. 1, where P and LCVR represent the polarizers and the liquid crystal variable retarder mentioned above, respectively. The LCVR used in the experiment is LCC1111-B from Thorlabs. It has a broad operating wavelength from 650 to 1050 nm. The laser diode (LASER), with a central wavelength of 852 nm and a linewidth of 2 MHz, is a distributed-feedback laser diode from Eagleyard. An optical isolator (ISO) prevents the reflected light and a pair of beam expanders (BE) adjust the laser beam diameter. The power-variation system is placed between the expanders for consideration of the overall volume. A photodiode (PD) acquires a fraction of the light power via a beam splitter (BS). The signal is pre-amplified, and compared with a reference voltage. The error signal is coupled back, via a proportion-integration-differentiation circuit, to the voltage applied to the LCVR.

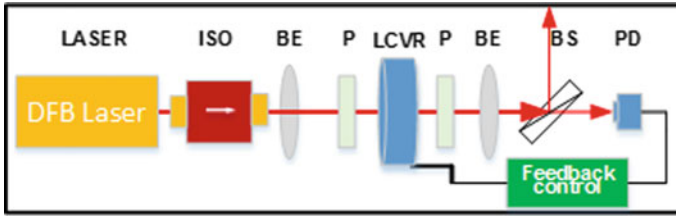


Fig. 1 Optical system of detection for Cs atomic clock

3 Results and Discussion

3.1 Results of Light Power Stability

We measure the relative light power variation without the light power stabilization, as plotted in Fig. 2. The total measuring time is 1.5×10^6 s, with a sampling time of 10 s. It is indicated that the light power varies largely and shows long-term drift without light power stabilization.

The variation, represented by Allan deviation, is shown in Fig. 3. The Allan deviation increases for averaging time larger than 100 s, and reaches 4.5×10^{-3} at 10^5 s.

Figure 4 shows the relative light power when the light power is stabilized. It is seen that the long-term drift of the light power is well eliminated. It is worth noting that the temperature drift of the electronic circuits has been removed.

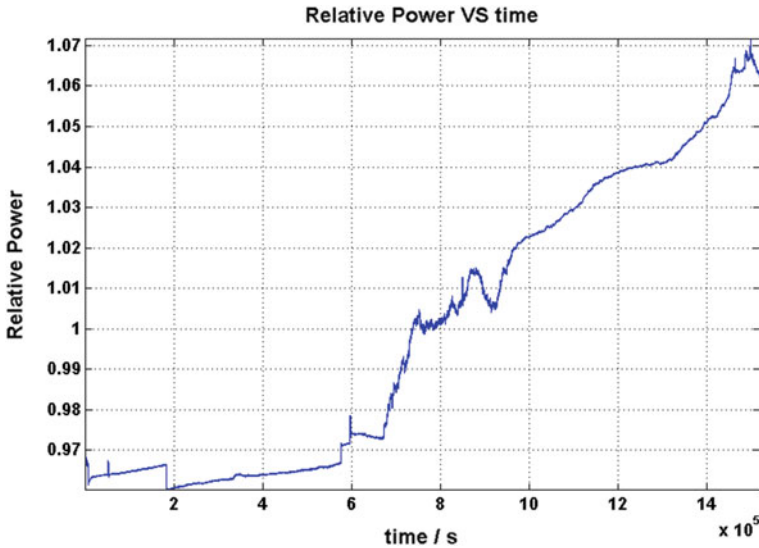


Fig. 2 Relative light power without light power stabilization

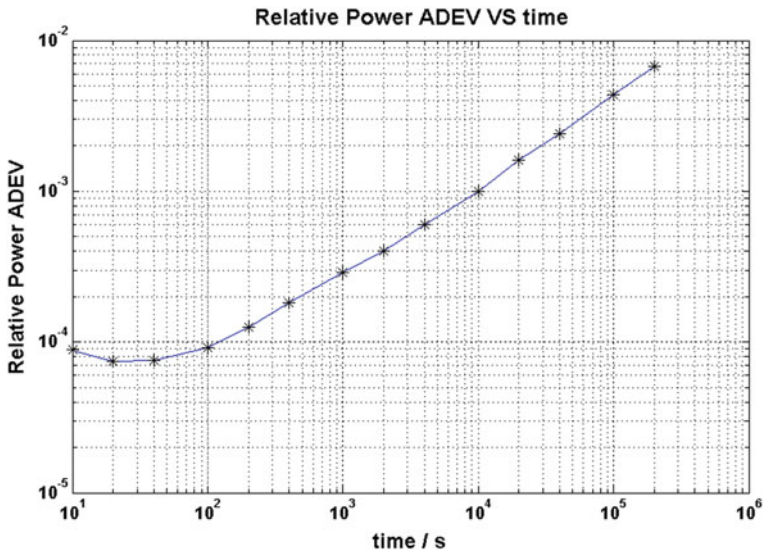


Fig. 3 Allen deviation of relative light power without light power stabilization

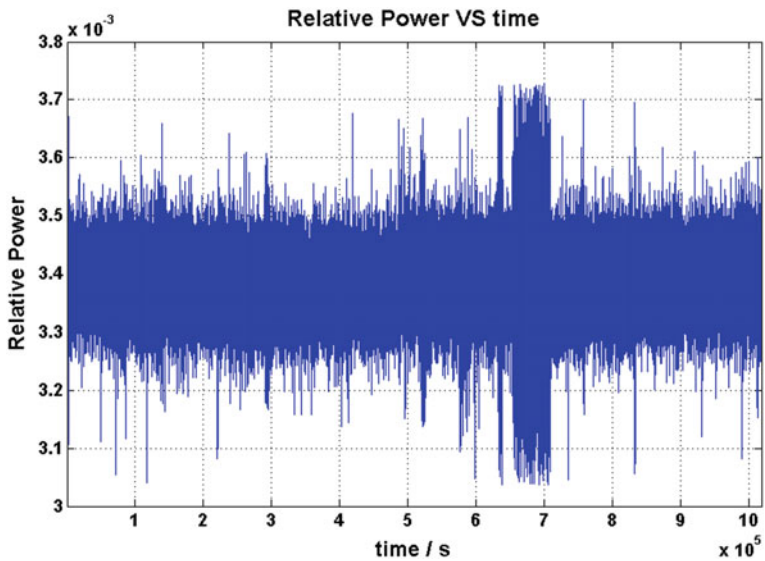


Fig. 4 Relative light power with light power stabilization

The corresponding Allan deviation is shown in Fig. 5. The Allan deviation of light power is 6×10^{-5} at 10 s, and 3×10^{-6} at 10^5 s. The long-term stability of the light power is remarkably improved.

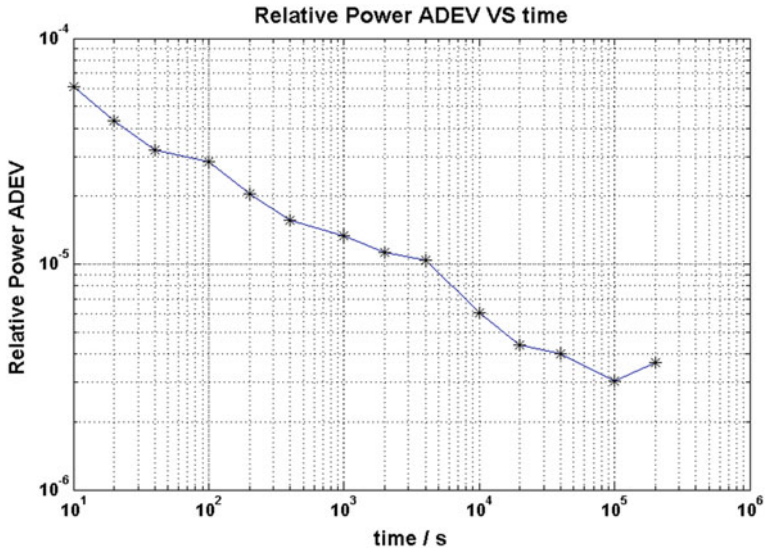
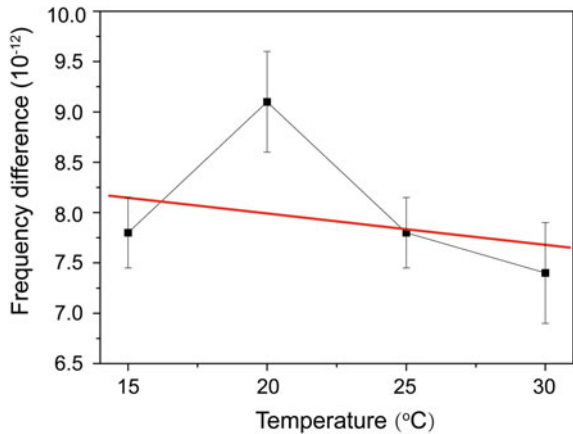


Fig. 5 Allen deviation of relative light power with light power stabilization

3.2 Temperature Coefficient Measurement Results

The temperature coefficient, i.e. the dependence of the output frequency of the clock on the environmental temperature, is a critical parameter of the clock. Since the laser system is usually sensitive to the environments, reducing the temperature coefficient is difficult both in our scheme and in the optical pumping scheme. We have measured the temperature coefficient of the clock without laser power stabilization, which is $1 \times 10^{-12}/^{\circ}\text{C}$. The coefficient is reduced to $-3 \times 10^{-14}/^{\circ}\text{C}$ when

Fig. 6 Temperature coefficient of Cs atomic clock with light power stabilization



the laser power is stabilized, as in Fig. 6. In conclusion, the light power stabilization method effectively controls the light power, stabilizes the light shift and reduces the temperature coefficient.

3.3 Discussion

The results have manifested the effect of the detection light power stabilization. In this scheme, the power stability is limited by the temperature drift of the components in the electronic circuits. Theoretically, the overall temperature drift of the circuits is below $1 \times 10^{-5}/^{\circ}\text{C}$. In order to remove the influence of temperature drift from the circuits, we use the difference of the acquired voltage of the PD and the reference voltage as the result.

The light power used in the clock detection is 5 mW, and the dependence of the clock frequency to light power is $2 \times 10^{-12}/\text{mW}$. Accordingly, the partial Allan deviation of the clock frequency due to the light shift is 4.5×10^{-14} , which is a main limitation to the accuracy and long-term stability of the clock. After the light power stabilization is applied, as the Allan deviation of the light power is 3×10^{-6} , the Allan deviation due to the light shift at 10^5 s is 3×10^{-17} , which is negligible to the overall deviation. Moreover, the decrease of the temperature coefficient from 1×10^{-12} to $3 \times 10^{-14}/^{\circ}\text{C}$ manifests that the method largely improves the environmental suitability. Besides, this method can be easily applied to other experiments that require laser power stabilization.

4 Conclusion

We propose a method for laser power stabilization in a Cs clock. This method uses a liquid crystal variable retarder to tune the polarization of the light, in order to feedback control the light power with respect to the photoelectric detector. We theoretically analyze the method, and find that the drift of light power mainly results from the temperature drift of electronic elements. The temperature drift is below $1 \times 10^{-5}/^{\circ}\text{C}$, manifesting the theoretical feasibility. We build an experimental system and measure the light power for 20 days. The stability of light power is 6×10^{-5} at 10 s, and 3×10^{-6} at 10^5 s, with the temperature drift of electronic elements removed. Moreover, the temperature coefficient reduces to $3 \times 10^{-14}/^{\circ}\text{C}$ from $1 \times 10^{-12}/^{\circ}\text{C}$. Therefore, the method effectively improves the performance of the clock. Besides, it can be easily applied to other experiments that requires laser power stabilization.

References

1. Lutwak R, Emmons D, Garvey RM, Vlitaz P (2001) Optically pumped cesium-beam frequency standard for GPS-III. In: Proceedings of the 33th annual precise time and time interval systems and applications meeting, Long Beach, California, November 2001, pp 19–32
2. Lecomte S, Haldimann M, Ruffieux R, Berthoud P, Thomann P (2007) Performance demonstration of a compact, single optical frequency cesium beam clock for space applications. In: IEEE international frequency control symposium joint with the 21st European frequency and time forum, pp 1127–1131
3. Liu C, Shi R, Wang Y et al (2015) An optically detected cesium beam frequency standard with magnetic state selection. In: European frequency and time forum. IEEE, pp 175–177
4. Liang Q, Chen L, Lei G, Wu W, Zhou B (2015) Laser intensity stabilization with a liquid crystal variable retarder for a nuclear magnetic resonance gyroscope prototype. In: Proceedings, vol 9671, AOPC 2015: advances in laser technology and applications; 96711Q

Analysis of BDS Satellite Clock in Orbit with ODTS and TWTT Satellite Clock Data



Bin Wang and Junping Chen

Abstract Two Way Time Transfer (TWTT) is used in BDS to synchronize the time between satellite and ground system. Therefore, both Orbit Determination and Time Synchronization (ODTS) and TWTT can obtain BDS satellite clocks. For ODTS, satellite clock is estimated with satellite orbit, and the consistency of them is better, but estimation error of them cannot be separated. Compared with ODTS, satellite clock estimated by TWTT is less impacted by satellite orbit, because of the difference between uplink and downlink pseudo-ranges in TWTT processing. Consistency of ODTS satellite clock and TWTT satellite clock is of importance for the performance analysis of BDS satellite clock in orbit. Based on the satellite clocks estimated by ODTS and TWTT respectively, this paper analyzes the systematic, periodical and stochastic characteristics of BDS satellite clock, taking into account the characteristics of satellite clock data. The result can provide reference for BDS service performance optimization. It is shown that the performance of BDS satellite clock is approximately $5 \times 10^{-12} \tau^{-0.5}$ (WHFM, $10^3 \text{ s} < \tau \leq 10^4 \text{ s}$) + 4×10^{-14} (FLFM, $\tau > 10^5 \text{ s}$), while during the 10^4 – 10^5 s period, the orbital-period characteristics is notable.

Keywords Beidou navigation satellite system • Orbit determination and time synchronization • Two way time transfer • Satellite clock in orbit
Noise analysis

B. Wang · J. Chen (✉)

Shanghai Astronomical Observatory, Chinese Academy of Sciences,
Shanghai 200030, China
e-mail: junping@shao.ac.cn

J. Chen

School of Astronomy and Space Science, University of Chinese
Academy of Sciences, Beijing 100049, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_50

1 Introduction

Performance of GNSS is closely related to that of on board atomic clock. Navigation, positioning and timing accuracy is affected by the performance of satellite clock, through the influences of satellite clock modeling and forecasting. Different from the other GNSSs (such as GPS, GLONASS and Galileo) using the orbit determination and time synchronization (ODTS) technique for the time synchronization between satellites and ground system, two-way time transfer (TWTT) is used by BDS [1]. In ODTS algorithm, ionosphere-free linear combinations of L1 and L2 phase and code observations are taken as input. Through the separation of various propagation delays (such as tropospheric delay), estimations of satellite orbit and satellite clock are obtained [2]. Up-link and down-link pseudo-ranges between satellite and ground station is used by TWTT to perform the time comparison between satellite and the station, and only the satellite clock is estimated.

Through the differences between up and down-link pseudo-ranges, TWTT can eliminate the influences of some common errors, such as the tropospheric delay, satellite ephemeris errors and so on, to the satellite clock estimations [3]. Compared with satellite clock estimation obtained by ODTS algorithm, TWTT satellite clock estimation is relatively less affected by the satellite orbit, however ODTS satellite clock is more consistent with the ODTS satellite orbit. Code pseudo-range and carrier phase observations are generally processed by ODTS algorithm. Absolute reference of satellite clock is provided by the pseudo-range, while the precision of satellite clock estimation is usually determined by the precision of carrier phase observations. Precision of ODTS satellite clock is around 33 ps [4], while the precision of TWTT satellite clock is affected by the precision of code pseudo-range. Therefore, ODTS satellite clock and TWTT satellite clock have different but complementary attributes. They can be used to reveal different aspects of BDS satellite clock performance. Comparative analysis of ODTS and TWTT satellite clocks are useful to identify the estimation error of BDS satellite orbit and satellite clock, and also useful to improve the estimation accuracy of them.

Analysis of BDS satellite clock is of great significance for understanding the status of BDS service and discussing the improvement of BDS service performance. At present, many scholars have carried out a large number of work of performance analysis of BDS satellite clocks based on ODTS clock data [5–8]. However, performance analysis of BDS satellite clock based on TWTT clock data is rarely carried out, and comparative analysis of BDS satellite clock performance result based on these two kinds of different clock data has not been carried out yet. In this paper, we compare TWTT and ODTS satellite clock data of BDS, and analyze their performance. The analysis is carried out from three aspects: characteristics of satellite clock data, characteristics of the periodic variation of satellite clock, and random characteristics of satellite clock.

2 Comparison of Satellite Clock Data

2.1 Mutual Agreements of ODTS and TWTT Satellite Clock Data

Due to different observations, processing algorithms and software implementations used by ODTS and TWTT algorithms, satellite clock data obtained by these two algorithms can be regarded as independent from each other, and therefore the mutual agreements of them can be obtained by comparative analysis of these two kinds of satellite clock data.

After removing the influences of different time reference, mutual agreements between TWTT and ODTS satellite data is shown in Fig. 1. Figure 1 shows the time variations of the difference between ODTS and TWTT satellite clock data from September 18 to September 21, 2013 and the functional relationship between the difference and the satellite orbit angle μ . It is can be seen that there is some correlation between the difference of these two kinds of satellite clock data and satellite orbit. Amplitudes of 1 cpr (cycle per revolution) for C01 and C05 satellites are 1 ns and 0.6 ns respectively, while those of C06 and C09 satellites are 0.4 and 0.3 ns, respectively.

In TWTT algorithm, differences of up-link and down-link code pseudo-range are used to estimate the satellite clock, thus the orbit error has little effect on the satellite clock estimates. ODTS algorithm estimates both satellite orbit and satellite clock, and satellite clock estimates would be affected by satellite orbit estimation error. Under the assumption that TWTT satellite clock data is relatively less affected by the orbit estimation error, the above phenomenon indicates that there is an influence of the satellite orbit estimation error in the ODTS satellite clock data.

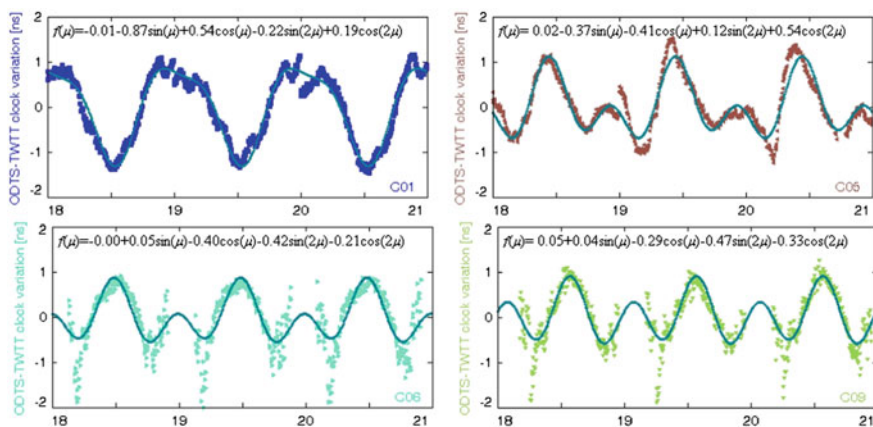


Fig. 1 Periodicity of difference between ODTS and TWTT clock

2.2 Internal Consistency Comparison of OTDS and TWTT Satellite Clock Data

The internal consistency was evaluated by the 95.5 quantile of the RMS error for quadratic polynomial fitting of daily satellite clock data, and the results are shown in Table 1. It can be seen that for GEO satellites, the internal consistency of TWTT satellite clock data is better than that of ODTS satellite clock data, while for IGSO satellites, the internal consistency of them is comparable.

3 Comparison of Periodicity Analysis Results Based on ODTS and TWTT Satellite Clock Data

Periodicity analysis results based on ODTS and TWTT satellite clock data are compared from two aspects: period, amplitude. Analysis results using ODTS satellite clock data can be referred to relevant literature, such as [9].

Analysis results based on the TWTT satellite clock are shown in Fig. 2. It can be seen that 12 h (2 cpr) harmonics of GEO and IGSO satellites clock is noticeable, and IGSO satellite clock has the n cpr harmonics. The amplitude of harmonics decreases with the increase of n . Compared with the analysis results based on ODTS satellite clock, the amplitude of 24 h (1cpr) harmonic of GEO and IGSO satellites using TWTT satellite clock data is smaller.

In summary, periodic variation characteristics of n (n is a positive integer) cpr for the BDS satellite clock can be obtained whether using ODTS satellite clock data or TWTT satellite clock data. Results of harmonic period of BDS satellite clock analyzed by these two kinds of satellite clock data are consistent with each other, while there are some differences in the amplitudes. Amplitude results analyzed by TWTT satellite clock data is smaller than the results analyzed by the ODTS satellite clock data, especially for the 1 cpr harmonic.

Table 1 Internal consistency comparison of ODTS and TWTT clock (second half of the year 2013)

Orbit/PRN		Type	
		ODTS [s]	TWTT [s]
GEO	C01	9.72E-10	6.71E-10
	C02	4.52E-09	3.91E-09
	C03	1.10E-09	6.28E-10
	C04	1.73E-09	1.63E-09
	C05	1.23E-09	8.61E-10
IGSO	C06	1.87E-09	1.49E-09
	C07	8.33E-10	6.87E-10
	C08	4.83E-10	5.72E-10
	C09	6.57E-10	7.81E-10
	C10	9.19E-10	7.18E-10

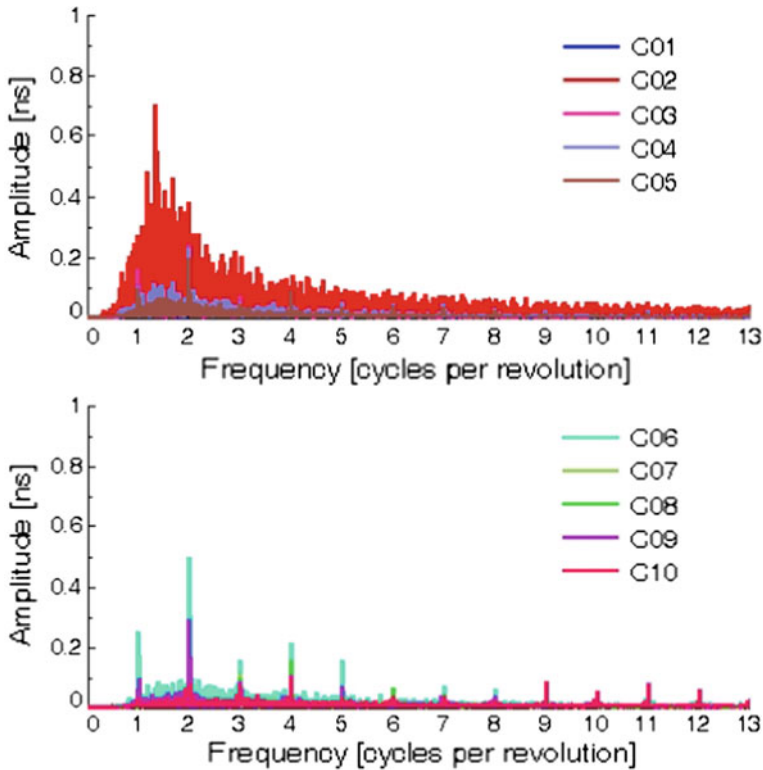


Fig. 2 Periodogram of BDS satellite clock based on TWTT satellite clock data

4 Comparison of Stability Analysis Results Based on ODTS and TWTT Satellite Clock Data

Stability analysis results based on ODTS and TWTT satellite clock data are compared from three aspects: time domain variances of frequency stability, dominant power law noise type and power law spectral densities. Analysis results of C05 (GEO) satellite and C07 (IGSO) satellite clock are described. Allan, Hadamard and Modified Allan deviation calculated by these two kinds of clock data are respectively compared, and the noise types in different time periods are analyzed. Among these methods, the Hadamard deviation is insensitivity to linear frequency drift, and comparison of Hadamard deviation and Allan deviation is useful to validate the frequency stability analysis algorithm; modified Allan deviation is used to distinguish between white and flicker PM noise.

Stability analysis results of C05 satellite clock are shown in Fig. 3. It can be seen that variances calculated using TWTT satellite clock data is larger than that calculated using ODTS satellite clock data in the period of $3 \times 10^2 - 5 \times 10^4$ s, while

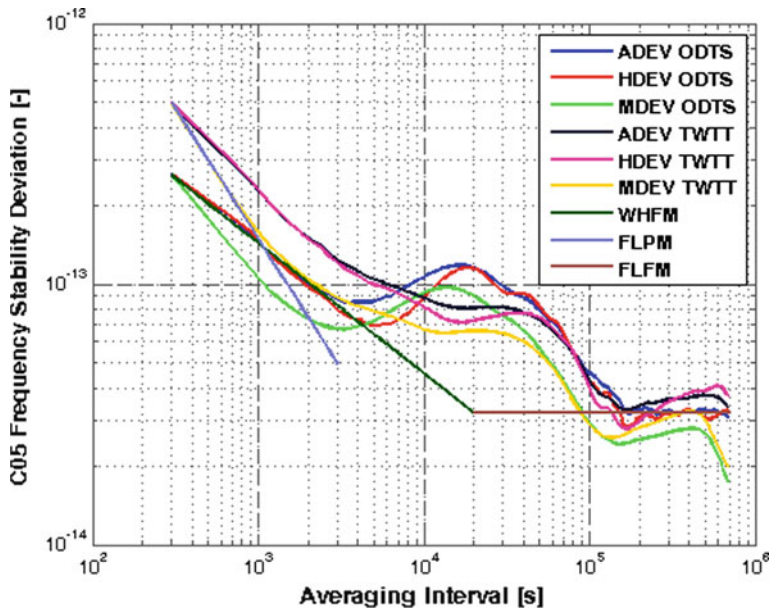


Fig. 3 Result comparison of frequency stability analysis for C05 satellite clock based on TWTT and ODTS satellite clock data. Allan deviation denoted by ADEV, Hadamard deviation denoted by HDEV, and modified Allan deviation denoted by MDEV

the variances calculated by both satellite clock data in the period of larger than 5×10^4 s are consistent with each other.

For TWTT satellite clock data, remarkable effect of flicker phase modulation (FLPM) noise can be seen in the period of $3 \times 10^2 - 1 \times 10^3$ s. While in the period of $1 \times 10^3 - 1 \times 10^4$ s, power law noise type analyzed from TWTT satellite clock data is white frequency modulation (WHFM) noise, which is the same as that analyzed using ODTS satellite clock data. TWTT satellite clock data is affected by the periodic variations of satellite orbit during the period of 10^4 s to one orbit period (satellite orbit period of GEO and IGSO satellites is one day). However, the periodic variations influences on TWTT are much smaller than of ODTS. In the period of longer than one satellite orbit period, power law noise type obtained using both satellite clock data is the same, and both are the flicker frequency modulation (FLFM) noise.

Frequency stability analysis results of C07 satellite clock are shown in Fig. 4. Comparison results of frequency stability analysis using both satellite clock data is consistent with that of C05 satellite clock.

Power law spectral densities analyzed from ODTS and TWTT satellite clock data are the same, except for the FLPM noise in the short period. It is inferred that performance parameters of BDS satellite clock in orbit is $5 \times 10^{-12} \tau^{-0.5}$ (WHFM, $10^3 \text{ s} \leq \tau \leq 10^4 \text{ s}$) + 4×10^{-14} (FLFM, $\tau > 10^5 \text{ s}$).

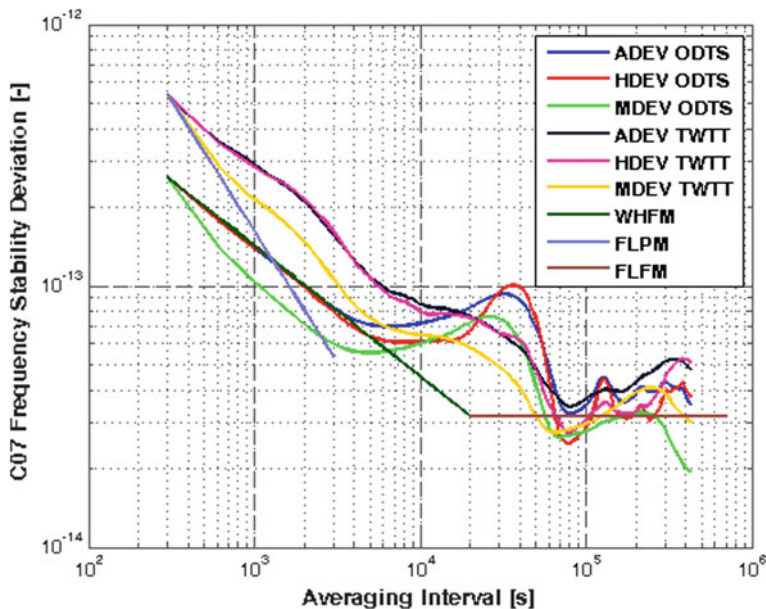


Fig. 4 Result comparison of frequency stability analysis for C07 satellite clock based on TWTT and ODTS satellite clock data. Allan deviation denoted by ADEV, Hadamard deviation denoted by HDEV, and modified Allan deviation denoted by MDEV

5 Conclusions

Analysis of BDS satellite clock based on ODTS and TWTT satellite clock data shows that, (1) for GEO satellites, the internal consistency of TWTT satellite clock data is better than that of ODTS satellite clock data, while for IGSO satellites, the internal consistency of both satellite clock data is comparable. (2) Harmonic period of satellite clock periodic variation obtained from ODTS and TWTT satellite clock data is consistent, and the amplitude of TWTT satellite clock data is less than that of ODTS, especially for the one cpr harmonic. (3) Difference between ODTS and TWTT satellite clock data has relationship with the satellite orbit angle, this phenomenon indicates that one cpr periodic variation of ODTS satellite clock data is related to the satellite orbit estimation error, especially for GEO satellites. (4) Comparative analysis of frequency stability analysis results obtained from ODTS and TWTT satellite clock data shows that, in the period of 10^2 – 10^3 s, remarkable effect of FLPM noise on the analysis results of TWTT satellite clock data can be seen. Power law spectral densities analyzed from ODTS and TWTT satellite clock data are almost the same. It is inferred that performance parameters of BDS satellite clock in orbit is $5 \times 10^{-12} \tau^{-0.5}$ (WHFM, $10^3 \text{ s} \leq \tau \leq 10^4 \text{ s}$) + 4×10^{-14} (FLFM, $\tau > 10^5 \text{ s}$).

Acknowledgements This work is supported by the NSFC (No.11673050), the 863 (No. 2014AA123102) projects and also the open project program of the key lab of space navigation and positioning technology (No. KFKT-201705), Shanghai.

References

1. Han C, Cai Z, Lin Y (2013) Time synchronization and performance of Beidou satellite clocks in orbit. *Int J Navig Obs* 2013:5
2. Zhao Q, J Guo, M Li (2013) Initial results of precise orbit and clock determination for compass navigation satellite system. *J Geod* 87(5):475–486
3. Liu L, Zhu L, Han C (2009) The model of radio two-way time comparison between satellite and station and experimental analysis. *Chin Astron Astrophys* 33(4):431–439
4. Senior KL, Ray JR, Beard RL (2008) Characterization of periodic variations in the GPS satellite clocks. *GPS Solut* 12(3):211–225
5. Gong H, Ni S, Mou W, et al (2012) Estimation of compass on-board clock short-term stability. In: *Proceeding of the 26th European frequency and time forum (EFTF)*, Gothenburg, Sweden
6. Zhou P, Du L, Zhang Z, et al (2014) Preliminary performance evaluation of Beidou spaceborne atomic clocks. In: *Proceedings of China satellite navigation conference (CSNC) 2014*. Springer, Berlin, Heidelberg
7. Montenbruck O, Steigenbrger P, Hugentobler U, et al (2013) Initial assessment of the Compass/Beidou-2 regional navigation satellite system. *GPS Solut* 17(2): 211–222
8. Steigenberger P, Hugentobler U, Hauschild A, et al (2013) Orbit and clock analysis of compass GEO and IGSO satellites. *J Geod* 87(6):515–525
9. Wang B, Lou Y, Liu J, et al (2016) Analysis of BDS satellite clocks in orbit. *GPS Solut* 20(4):783–794

Analytical Analysis Method of Whispering Gallery Mode of Sapphire Dielectric Resonators and the Optimization of the Q Value in Liquid Helium Temperature



Xi Zhu, Haibo Chen, Kai Huang and Lianshan Gao

Abstract A novel approximate analytical method is presented for calculating parameters of Whispering Gallery Mode (WGM) in a sapphire dielectric resonator (SDR). It is based on ray optics and wave theory. The frequency of WGM is computed with less computation time and the influence of sapphire's parameters on the frequency is also analyzed. The Q value of SDR in the liquid helium temperature is optimized by adjusting the coupling coefficient and the surface roughness of the sapphire. According to the computed results, a SDR employing WGM at X-Band is fabricated. The loaded Q_L value reaches 4.0×10^8 and the eigen frequency is 9.205 GHz in the liquid helium temperature.

Keywords Sapphire dielectric resonator · Q value · Oscillator
Whispering gallery mode · Eigen frequency

1 Introduction

Cryogenic sapphire oscillator (CSO) has been studied extensively for its very low phase noise and excellent short-term frequency stability, which can be used in radar, metrology and fundamental physics measurement [1–3]. The phase noise of CSO has reached -160 dBc/Hz at 1-kHz offset from the 11.2 GHz carrier and the frequency stability can reach a level of 1×10^{-15} @1s [1]. The noise and short-term stability of CSO is far better than the frequency multiplication method of quartz crystal oscillator. This is benefited from the extremely high Q value of its SDR. Therefore, as the key part of CSO, SDR has been extensively studied. The Q value of SDR can reach 10^5 at room temperature, and 10^9 at liquid helium [4–8]. John et al. fabricated the CSO, reaching the fractional frequency stability of 10^{-15} with

X. Zhu (✉) · H. Chen · K. Huang · L. Gao
Beijing Institute of Radio Metrology and Measurement, Beijing 100854, China
e-mail: 1989zhuxi@163.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_51

loaded Q value of 1×10^9 . Christophe et al. fabricated three CSOs and reached a stability of 5×10^{-16} in 1s using three-cornered-hat method with three SDRs, which had loaded Q value 1×10^9 , 3.5×10^8 and 4×10^8 , respectively [9].

Sapphire material has been widely used to fabricate the SDR employing WG modes for its very low loss tangent at high frequency. The WGMs can be analyzed by Maxwell equations combining the boundary conditions or numerically computed by Finite element method. The SDR without enclosure of the metal cavity has been analyzed by M. E. Tobar using analytical method with rather high precision [5]. The SDR with the enclosure of metal cavity has been analyzed by E. N. Ivanov, R. C. Taber and Yoshio Kobayashi [6–8] and the results are more precise than that reported in Ref. [5], while these methods are very complicated.

In this paper, a novel analytical method combining the wave theory and the ray optics is proposed to solve the problem of WGM. This method has more intuitional field distribution and requires less computation. The influences of sapphire material's parameters are also analyzed. Furthermore, the results of this analytical method, numerical method and experimental data are compared. The method presented in this paper is consistent with the experiment and can easily get the eigen frequency of the SDR. An SDR is fabricated according to the computed results and the Q value in the liquid helium temperature is optimized by adjusting the coupling coefficient. The experimental resonant frequency is 9.205 GHz, which is in consistent with the analytical results. The loaded Q value of SDR reaches 4.0×10^8 .

2 Theoretical Model

The SDR is composed of a sapphire and a metal enclosure. The schematic section of the sapphire is shown in Fig. 1. When the eigen modes are WGMs, the electro-magnetic field is mainly confined in the sapphire and distributes along the circumference, as shown in Fig. 2. The frequency and the Q value are determined by the size of sapphire. The frequency variation is small regardless of the metal enclosure. The sapphire cylinder can be regarded as a ring waveguide when oscillating in the WGM, as shown in Fig. 3.

Fig. 1 The schematic section of SDR

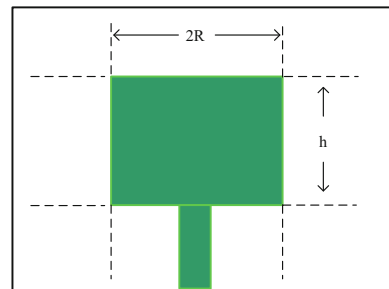


Fig. 2 The schematic distribution of WGM

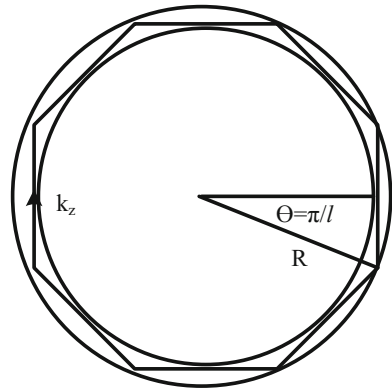
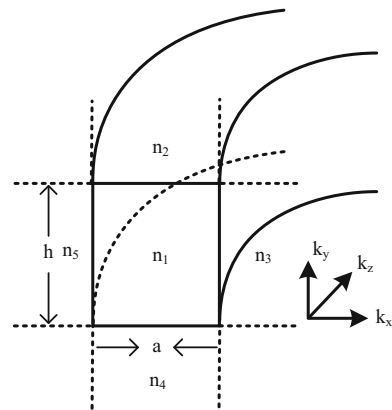


Fig. 3 The equivalent ring waveguide of the sapphire cylinder



There are two types of modes in the ring waveguide, the E mode and the H mode. According to the boundary conditions of waveguide, the equations are derived.

For the E mode

$$k_y h = n\pi - \tan^{-1} \frac{k_y n_2^2}{k'_{2y} n_1^2} - \tan^{-1} \frac{k_y n_4^2}{k'_{4y} n_1^2} \tag{1a}$$

$$k_x a = m\pi - \tan^{-1} \frac{k_x}{k'_{3x}} - \tan^{-1} \frac{k_x}{k'_{5x}} \tag{1b}$$

For the H mode

$$k_y h = n\pi - \tan^{-1} \frac{k_y}{k'_{2y}} - \tan^{-1} \frac{k_y}{k'_{4y}} \tag{2a}$$

$$k_x a = m\pi - \tan^{-1} \frac{n_3^2 k_x}{n_1^2 k'_{3x}} - \tan^{-1} \frac{n_5^2 k_x}{n_1^2 k'_{5x}} \quad (2b)$$

$$k_z 2n_1 R' = l 2\pi \quad (3a)$$

$$R' = R(1 + \cos(\pi/l))/2 \quad (3b)$$

where $k'_{2y} = \sqrt{4\left(\frac{\pi}{\lambda}\right)^2 \left(n_1^2 - n_2^2\right) - k_y^2}$, $k'_{3x} = \sqrt{4\left(\frac{\pi}{\lambda}\right)^2 \left(n_1^2 - n_3^2\right) - k_x^2}$, $n_1^2 k_0^2 = k_x^2 + k_y^2 + k_z^2$, m, n, l are integer, $k_0 = 2\pi/\lambda$. $n_1 = \sqrt{\varepsilon_z}$ for (1a) and (2a), $n_1 = \sqrt{\varepsilon_r}$ for (1b) and (2b). $\sqrt{\varepsilon_z}$, $\sqrt{\varepsilon_r}$ are the permittivity parallel and perpendicular to the c-axis of sapphire, respectively. $f_0 = k_0 c/2\pi$, c is the speed of light in vacuum. When m, n, l, R, h are determined, f_0 will be solved.

3 Analytical Frequency and Optimization of the Q Value

The designed eigen frequency is close to 9.19 GHz, which is the same as that of the local oscillator of Cesium fountain clock. According to Ref. [7], ε_z is about 11.37 and ε_r is about 9.28 near the liquid helium temperature. Figures 4, 5, 6 and 7 show the frequency variations with height h , radius R , ε_z and ε_r , respectively. From the results, we can see the frequency variation with R is larger than that with h . It is because k_y is smaller than k_z , and a small variation in circumference will result in a great change in k_z . Furthermore, with the increase in h, R, ε_z or ε_r , the frequency reduces, which causes negative temperature frequency. When h and R is 29.9 and 24.3 mm respectively, the eigen frequency is near 9.2 GHz. Analytical and numerical methods are used to analyze the different SDR's WGMs, as shown in Fig. 8. Furthermore, the frequency increases with the l number. As shown in Fig. 8, the result computed by this analytical method is lower than the simulated results of numerical method and experimental results, especially for the small azimuthal number l . This is caused by two reasons. One is that the analytical model does not include the effect of k_z 's variation with the bended waveguide. Specifically, the component of k_z will leak from the interface between sapphire and air. That will make the computed frequency lower than the experimental results. Another reason is the (3b) is approximately derived from the ray theory. The error in waveguide width will cause variation of k_x , and then the frequency will be larger, especially when the azimuthal number l is small. The field distribution is verified by HFSS as shown in Fig. 9.

According to the analysis above, sapphire cavity is developed. To obtain how the surface finish of sapphire samples affects their Q value, which is difficult to analyze by theoretical method, two sapphire samples are made with different surface finish, sample A 40/20 and sample B 80/50, as Fig. 10 shows. The sapphires

Fig. 4 The variation of f with h ($m = 1, n = 1, l = 12$)

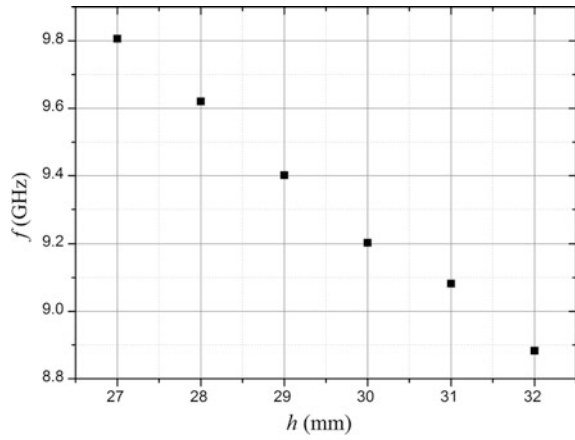


Fig. 5 The variation of f with R ($m = 1, n = 1, l = 12$)

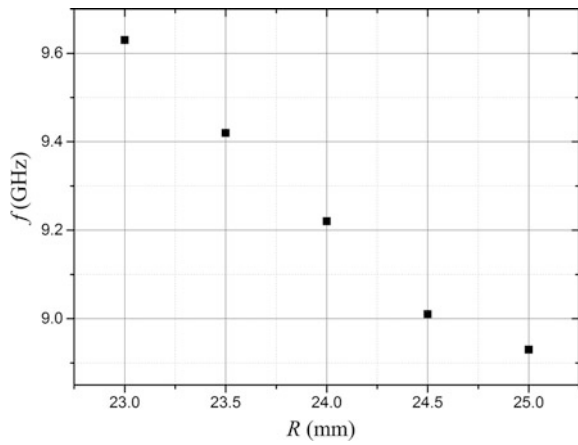


Fig. 6 The variation of f with ϵ_z ($m = 1, n = 1, l = 12$)

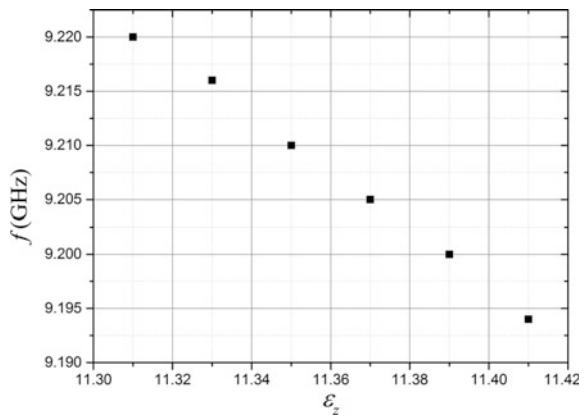


Fig. 7 The variation of f with ϵ_r ($m = 1, n = 1, l = 12$)

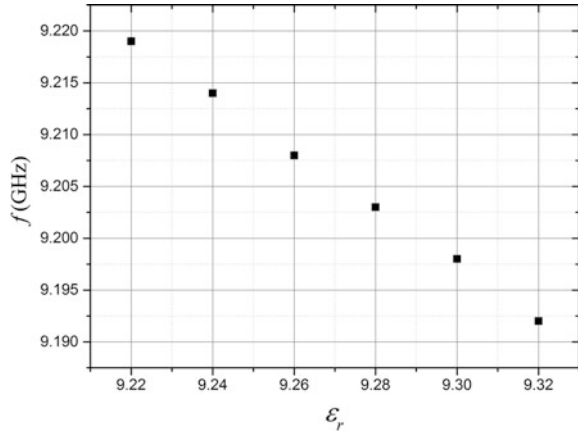
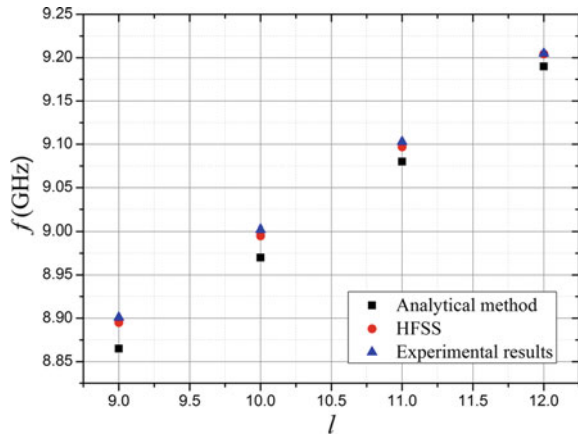


Fig. 8 The variation of f with l



are ensembled with copper cavity and deposited in cryostat as shown in Fig. 11. The Gifford-McMahon cycle cryocooler is used to liquefy helium gas, and then liquid helium flows into liquid chamber. While a copper plate with heating wires is placed under the liquid chamber which is used for stabilizing temperature inside the vacuum dewar. Sapphire cavity is screwed under the copper plate. With vacuum dewar preventing the heat exchange from liquid helium to room environment and makes the temperature more stabilized. The sapphire cavity is excited by two coupling rings. One is set as input port with long length of coupling ring and strong coupling coefficient. Another is set as output port with short length of coupling ring and weak coupling coefficient. The Q value is optimized by adjusting the length of input port. Figure 12 shows the relationship between the length of input port and Q value, and also shows the relationship between the length of input port and the insertion loss of sapphire cavity which include the loss of cables in the dewar. In order to measure these data, every measurement are performed by manual adjusting

Fig. 9 Field distribution of l simulated by HFSS soft

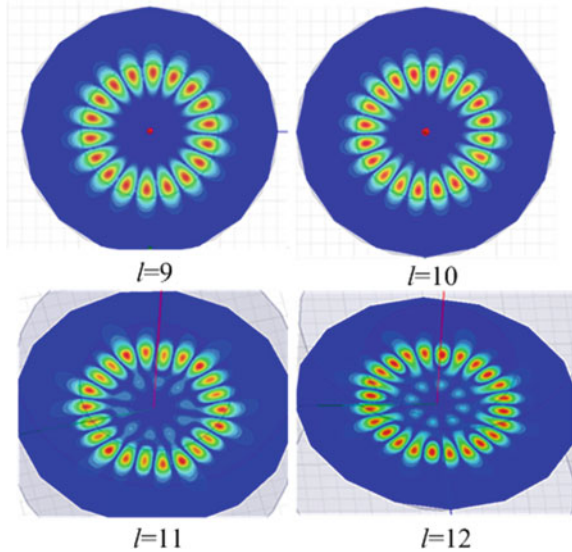
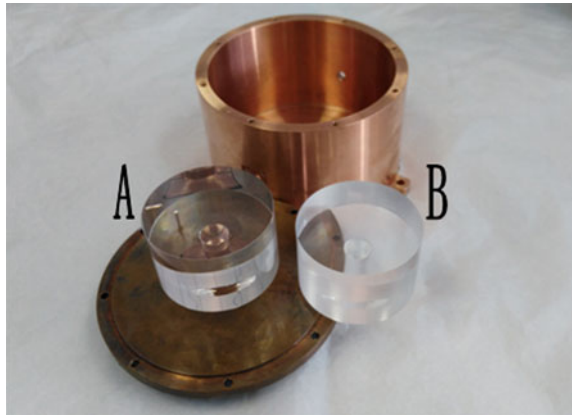


Fig. 10 Two sapphires with different surface (Finish A: 40/20, B: 80/50)



the length of coupling ring in the room temperature and then cooling down the sapphire cavity to 6.5 K. The temperature is controlled by Lakeshore 350.

The Q value varies exponentially with the length of coupling ring because the field distribution is a function of Bessel. When the distance between the coupling ring and sapphire is large, there is less energy flow out of the cavity through the ring, so the storage energy inside the sapphire cavity is larger. When the length is 6.5 mm the Q reaches 4.0×10^8 . It can be inferred that when the length is less than 6.5 mm the Q will be bigger while the insertion loss will also be larger. The variation of insertion loss and coupling ring length is also given in Fig. 12. It is clearly shown that the insertion loss increases quickly with the coupling ring length

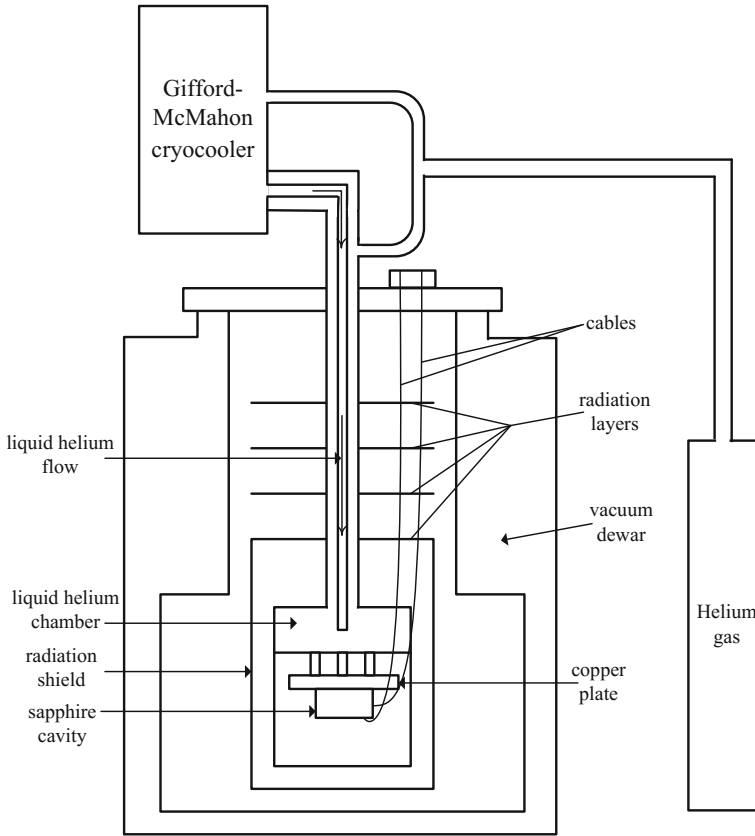


Fig. 11 Schematic of cryocooler cryostat with sapphire cavity

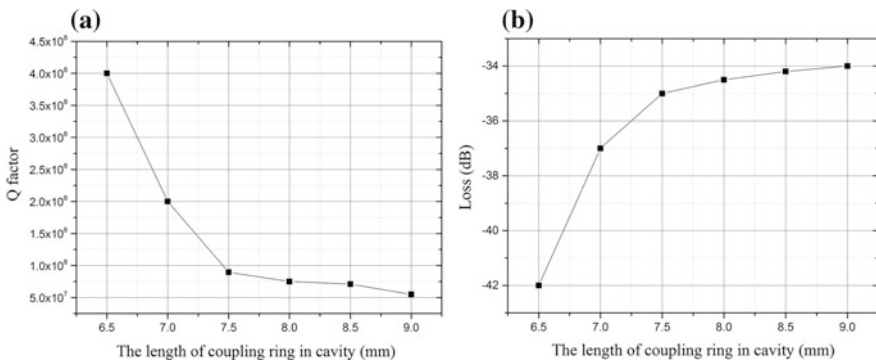
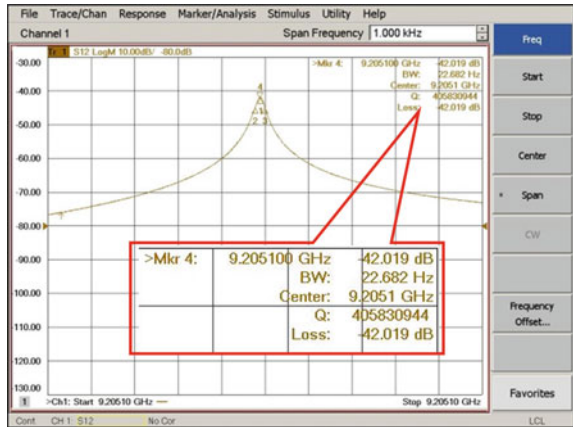


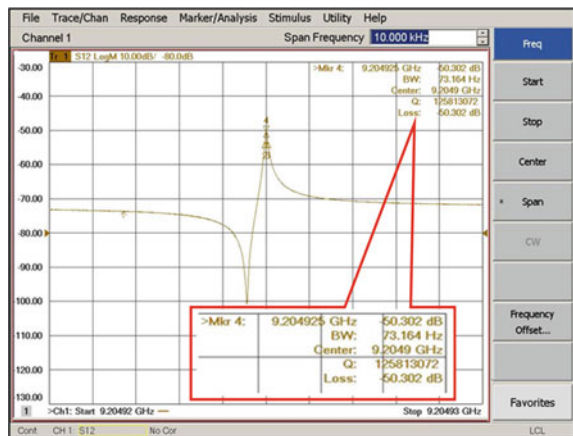
Fig. 12 Q value and insertion loss with the length of coupling ring

Fig. 13 The frequency and Q value of sapphire cavity employing WGH_{12,0,0} mode (A cavity)



reducing. It must be pointed out that the choice of Q value should also take the insertion loss into account. As CSO is a positive feedback loop, that is to say, if the insertion loss is too large, the whole loop oscillating will not be sustained. The maximum gain of the additional low-phase-noise amplifier can reach 56 dB in the typical CSO loop. While the insertion loss is 42 dB which includes the 22 dB loss of sapphire cavity and 20 dB loss of the 0.86 mm diameter cryo-cables in dewar. Besides, the loss of other device and cables outside the dewar together are about 10 dB, so the loop loss will be appropriate for oscillating. If the Q value is further improved, the loss of the cables in the dewar will be decreased and the length of coupling ring will be reduced. Figure 13 shows the frequency and Q value of cavity A (enssembled with A sapphire), the analysis above is based on cavity A. The frequency and the Q value are 9.205 GHz and 4.0×10^8 respectively at 6.5 K. Figure 14 shows the frequency and the Q value of cavity B (enssembled with B sapphire). The frequency and the Q value are 9.204 GHz and 1.2×10^8

Fig. 14 The frequency and Q value of sapphire cavity employing WGH_{12,0,0} mode (B cavity)



respectively at 6.5 K. The sapphire A and B are the same and the input and output coupling ring used are the same ones, except that the surface Finish of A sapphire is 40/20 and surface Finish of B sapphire is 80/50. The Q value of A sapphire cavity is over 3 times larger than B sapphire cavity since the leak of WGH mode energy is less in smooth surface.

4 Conclusion

An analytical method combining the wave theory and the ray optics is proposed to compute the eigen frequency with intrinsic phenomena and less computation. The Q value of a sapphire cavity in 6.5 K is optimized by adjusting the length of coupling ring and the influence of surface Finish is verified by experiment. Considering the loop gain of CSO, the Q value reaches 4.0×10^8 at 9.205 GHz.

References

1. Hartnett JG, Nand NR (2010) Ultra-low vibration pulse-tube cryocooler stabilized cryogenic sapphire oscillator with 10^{-16} fractional frequency stability. *IEEE Trans Microw Theory Tech* 10:1–8
2. Locke CR, Ivanov EN, Hartnett JG, Stanwix PL, Tobar ME (2008) Invited article: design techniques and noise properties of ultrastable cryogenically cooled sapphire-dielectric resonator oscillators. *Rev Sci Instrum* 79:1–12
3. Grop S, Bourgeois PY, Boudot R et al (2010) 10 GHz cryocooled sapphire oscillator with extremely low phase noise. *Electron Lett* 46:420–422
4. Hartnett JG, Tobar ME, Ivanov EN et al (2013) Optimum design of a high-Q room-temperature whispering-gallery-mode X-band sapphire resonator. *IEEE Trans Ultrason Ferroelectr Freq Control* 60:1041–1047
5. Tobar ME, Mann AG (1991) Resonant frequencies of higher order modes in cylindrical anisotropic dielectric resonators. *IEEE Trans Microw Theory Tech* 39:2077–2082
6. Ivanov EN, Blair DG, Kalinichev VI (1993) Approximate approach to the design of shielded dielectric disk resonators with whispering-gallery modes. *IEEE Trans Microw Theory Tech* 41:632–638
7. Taber RC, Flory CA (1995) Microwave oscillators incorporating cryogenic sapphire dielectric resonators. *IEEE Trans Ultrason Ferroelectr Freq Control* 42:111–119
8. Kobayashi Y, Senju T (1993) Resonant modes in shielded uniaxial-anisotropic dielectric rod resonators. *IEEE Trans Microw Theory Tech* 41:2198–2205
9. Fluhr C, Grop S, Dubois B et al (2016) Characterization of the individual short-term frequency stability of cryogenic sapphire oscillators at the 10^{-16} level. *IEEE Trans Ultrason Ferroelectr Freq Control* 63:915–921

Part IV
Satellite Navigation Augmentation
Technology

Application of Gaussian Overbounding on Ionospheric Error Analysis for China SBAS



Dun Liu, Liang Chen, Xiao Yu and Weimin Zhen

Abstract Delay errors will follow a biased distribution under the ionospheric anomaly or storms conditions, making it difficult to overbound effectively the delay errors with the estimated error bounds for SBAS. Gaussian-bounding theorem provides a way to test the effectiveness of ionospheric error bound estimate. With the theorem, various bound estimating methods and its overbounding capability were analyzed with observations during storms for China area where ionospheric anomaly persists. It also shows the strength of different estimations, in which the irregularity detector is an integral component to minimize the size of bound while maintaining the effectiveness, bound estimate with dynamic inflation factor outperforms over that of static inflating in system availability, and the undersampled threat model improves integrity margin in the bound at the cost of system availability, so diverse threat models are proposed to be constructed based on the ionospheric activities.

Keywords Ionosphere · SBAS—space-based augmentation system
Error overbounding · Integrity · Availability

1 Instruction

Aerospace applications have strict requirements for SBAS. SBAS has a top-level safety requirement to protect users against Hazardously Misleading Information (HMI) with 10^{-7} [1]. It is generally assumed that errors in positioning follow zero-mean normal distributions when analyze whether SBAS meet this strict requirement [2, 3].

Ionospheric delay is the largest source of positioning error. To allow the user to correct the error due to ionospheric delay, the grid model is adopted in SBAS to derive from measurements at Ionospheric Pierce Points (IPPs) a vertical delay

D. Liu (✉) · L. Chen · X. Yu · W. Zhen
No. 22nd Research Institute, CETC, Qingdao 266107, Shandong, China
e-mail: Dun.L@163.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_52

estimate at each ionospheric grid point (IGP), namely the grid ionospheric vertical delay (GIVD). In addition, SBAS computes at each IGP an integrity bound called the Grid Ionospheric Vertical Error (GIVE). For each GNSS signal detected by a user's receiver, SBAS guarantees that the user's computed integrity bound on the vertical delay error at the signal IPP should bound, with a sufficiently high probability, the actual slant delay error converted to vertical [1, 2].

The grid model applies well in mid-latitude area when ionosphere is calm. The storm causes spatial and temporal gradient in ionospheric delay, breaking down the shell model assumption underlying the grid model and leading to a biased normal distribution in error. Within approximately $\pm 20^\circ$ on both side of the magnetic equator lies the equatorial anomaly. It is the occurrence of a trough in the ionization in the F2 layer at the equator and crests at larger magnetic latitude depending solar activity. Ionospheric anomaly could also produce large delay gradient, and then larger error and biased error distribution [4, 5]. All these biased errors will affect the integrity risk analysis of SBAS.

Bruce DeCleene derived pseudorange-level requirement from the position domain integrity requirement, and proved that the assumption for a zero-mean, normal error distribution can be replaced by a requirement that the error distribution is symmetric, unimodal, and whose cumulative distribution function (CDF) is bounded by a normal error distribution [6]. Further, Timothy R. Schempp applied the Gaussian-bounding Theorem for the WAAS fault free error analysis, with the emphasis on satellite error and ionospheric error [3, 7]. In this paper, we make analysis on delay error from ionospheric anomaly and storm with the Gaussian-bounding Theorem for SBAS in China.

SBAS uses the methods of irregularity detecting, error inflating, and under-sampled threat model in delay error bound estimation to protect user with enough integrity margin [8–11]. However, a larger bound would reduce the system availability. So we make a further study on the overbounding effectiveness of estimated bounds with these various methods.

The second section of paper gives a brief introduction to the GIVE monitor responsible to compute the integrity bound associated ionospheric error. The third section introduces the Gaussian-bounding Theorem. The fourth section analyzes the effectiveness of the estimated delay error bound with the theorem. Effectiveness of different methods in error bound estimation is compared and analyzed in the fifth section. Finally, the conclusion is given.

2 GIVE Monitor in SBAS

GIVE monitor is designed to handle the threat originating from ionospheric error in SBAS. GIVE monitor is responsible to calculate GIVD, the delay at IGP and its error estimation GIVE, which in essence is the integrity bound information σ_{GIVE} associated with the delay correction. A user interpolates the grid delays and integrity bounds from the IGPs surrounding a signal's IPP to form the user's

ionospheric vertical delay (UIVD), and its error bound σ_{UIVE} which UIVE stands for user ionospheric vertical error [8–11].

In the case of quiet ionosphere in mid-latitude region, the grid model performs well. However, the delay error will increase during a storm, making it hard to be effectively bounded, and leading to the system integrity risk. In addition, the ionospheric anomaly in low latitudes could also seriously affect the accuracy of delay estimate and cause a system integrity risk [4, 5, 7]. To protect user from all these kind of integrity risks, a variety of measures is designed in SBAS to deal with the ionospheric threats.

GIVE monitor consists of two main components, each is designed to mitigate a particular error source: grid model error and undersampling threat [8, 9].

GIVE monitor calculates the delay at a IGP with measurements at surrounding IPPs with certain planer-fitting schemes, with the Kriging method a specific one implementation. The irregularity detector, a goodness-of-fit indicator in essence, is designed to distinguish a possible ionospheric disturbance at the same time. When the detector is triggered, the GIVE is set to the maximum value of 45 m [9, 12, 13].

To protect user from conditions when the irregularity detector is “near triggering”, the uncertainty in the planar fit is multiplied by an additional inflation factor denoted R_{irreg} . There are two different ways for R_{irreg} estimation: the static inflation factor and the dynamic inflation factor [9, 12, 14].

An integrity risk may arise when the ionospheric condition is not sampled sufficiently. Due to the poor coverage of SBAS network, a regional area where ionospheric irregularities present is not sampled by SBAS, but may be sampled by a user. The situation is particularly serious in the case of ionospheric storm. The ionospheric anomaly in low latitudes could make the situation even more severe. To handle this threat, GIVE monitor introduces an undersampled threat model to further inflate the estimated error bound.

The undersampled threat model is constructed using GNSS observations during historical ionospheric storm events. Irregularity detecting is an integral part of the threat model building. Only the data that pass through the detection are used to construct the model. Data depriving strategy is also considered to simulate the scenarios of undersampling measurements. There is a potential integrity risk from the aging delay corrections. These errors are taken into account in the undersampled threat model, and in SBAS correction algorithm as well [9, 15, 16].

3 Gaussian-Bounding Theorem

A SBAS user compute Horizontal Protection Levels (HPL) and Vertical Protection Levels (VPL) to determine whether the system has sufficient accuracy for a particular application. If the horizontal or vertical error exceeds the HPL or VPL, then the user has misleading information. For applications where the vertical positioning

error plays the major role, the mathematical expression for the probability that a user’s vertical positioning error (VPE) exceeds the user’s VPL is [3, 6]:

$$P(|VPE| > VPL) = P\left(\left|\sum_{i=1}^n k_i J_i\right| \geq K \sqrt{\sum_{i=1}^n (k_i \sigma_{B,i})^2}\right) \tag{1}$$

Where, n is the number of satellites used in the position solution, k_i comes from the third column of the use’s weighting matrix, J_i represents the errors on the i th ranging measurement, $\sigma_{B,i}$ is the estimated error bound of the i th ranging source, K is 5.33 here.

Defining a parameter A as:

$$A = \bar{\Phi}^{-1}\left(\frac{P_{HMI}}{2}\right) \tag{2}$$

where, $\bar{\Phi}(u)$ denotes $P(X \geq u)$, X is normal with mean 0 and variance 1.

Suppose there is a choice of γ such that all J_i are (μ_i, σ_i^2) Gaussian-bounded, and

$$(1 + \gamma^2)\sigma_i^2 + n(1 + \gamma^{-2})\left(\frac{\mu_i}{A}\right)^2 \leq \left(\frac{K}{A}\sigma_{B,i}\right)^2 \tag{3}$$

Then for a specific value of probability of HMI (P_{HMI}):

$$P\left(\left|\sum_{i=1}^n k_i J_i\right| \geq K \sqrt{\sum_{i=1}^n (k_i \sigma_{B,i})^2}\right) \leq P_{HMI} \tag{4}$$

Equation (3) can be referred as the “Threshold test”, where the threshold is $\left(\frac{K}{A}\sigma_{B,i}\right)^2$. For any particular distribution J_i , many different values of γ satisfy the threshold test. The theorem provides a criteria for satisfying the HMI problem for a specific value of P_{HMI} [3, 6].

To demonstrate the SBAS ionospheric error distribution meets the criteria of the Gaussian-bounding theorem, one should conduct statistical analysis on the delay error of the grid model and put the statistics into the theorem. Specially, the J_i will represent the ionospheric delay error on the i th ranging source, and $\sigma_{B,i}$ will be the bound σ_{UIVE} .

4 Gaussian-Bounding Test for Ionospheric Error

In this section, we conduct a study on the overbounding effectiveness of the bound estimate on delay error under the ionospheric anomaly and storm for SBAS in China. The way is as follows:

Construct the SBAS grid model with dual-frequency GNSS observation from network in China and surrounding area, and estimate the delays (GIVDs) and its error bounds (GIVEs) at grids.

Taking each station in the network as a virtual user at the same time, compute the accurate slant delays with the leveled carrier measurements for the viewed satellites at each station, and convert it into the vertical delays I_{Truth} at the corresponding IPPs.

Using the established grid model and the interpolating algorithm for user to compute the delays I_{SBAS} and its error bounds (UIVE) at the IPPs for the viewed satellites at each station.

The aggregate user error bounds are divided into several bins. The divisions are denoted UIVEI, and the range of each UIVEI bin is shown in Table 1 [3].

For each bin, the error ε_{UIVE} is computed as the normalized difference between I_{Truth} and I_{SBAS} with σ_{UIVE} (Eq. 5). The values of σ_{UIVE} are taken as the upper limit of the bin (see Table 2).

$$\varepsilon_{UIVE} = \frac{I_{Truth} - I_{SBAS}}{\sigma_{UIVE}} \tag{5}$$

Compute the mean μ and standard deviation σ of a Gaussian distribution that CDF overbounds the empirically derived distribution for each bin. Substitute μ , σ into the left-hand side of Eq. 3. The threshold used in the analysis was based on an allocation of 1.15×10^{-8} probability of HMI [3].

From the Gaussian-bounding theorem, when the ionospheric delay error could be overbounded with the estimated error bounds, there should be different values of γ satisfy the threshold test.

The following shows the results of the Gaussian-bounding threshold test with GPS data in China and the surrounding areas. The period of data is from September 10 to September 13, 2005, during which a strong ionospheric storm happened with a maximum value of 9 for the Kp index. Table 2 lists the mean and variance of

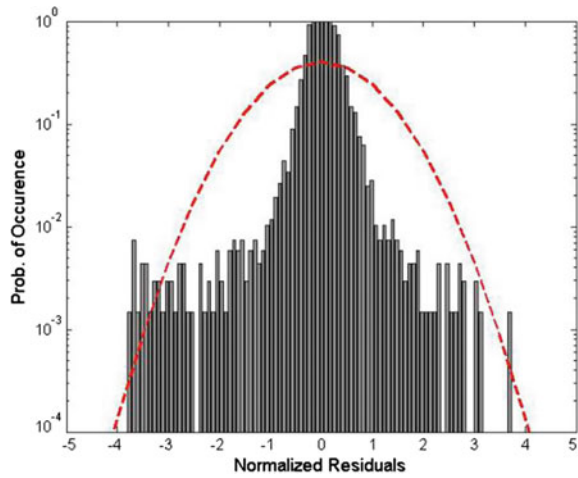
Table 1 UIVEI bins

UIVE index	σ_{UIVE} range (m)
UIVEI = 9	$\sigma_{UIVE} \leq 0.91$
UIVEI = 10	$0.91 < \sigma_{UIVE} \leq 1.09$
UIVEI = 11	$1.09 < \sigma_{UIVE} \leq 1.37$
UIVEI = 12	$1.37 < \sigma_{UIVE} \leq 1.82$
UIVEI = 13	$1.82 < \sigma_{UIVE} \leq 4.56$
UIVEI = 14	$4.56 < \sigma_{UIVE} \leq 13.68$

Table 2 Overbounding statistics for errors of UIVEI bins

σ_{UIVE}	μ	σ	Pts.
0.912	0.026	0.462	8298
1.094	0.005	0.466	17,052
1.368	-0.012	0.435	34,248
1.824	-0.059	0.655	1788
4.559	0.075	0.458	226
13.678	-0.037	0.116	442

Fig. 1 Delay error distribution for bin of UIVEI = 9

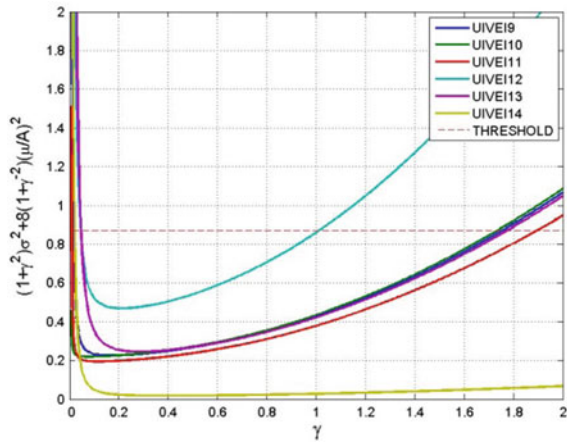


Gaussian distribution that can CDF overbound the empirically derived distribution for each UIVEI bin. As an example, Fig. 1 shows the delay error distribution for UIVEI = 9 bin the standard normal distribution is shown as well (red curve). For simplicity, the aggregate of all of the UIVE slices is plotted in Fig. 2. The threshold value is shown as the red dashed line. The Gaussian bounding curve for each UIVEI bin, which is derived with the corresponding μ and σ substituted into the left-hand side of Eq. 3, is plotted in different color. It can be clearly seen that there is some γ for all curves lies below the threshold, indicating the integrity of the delay error is maintained for SBAS in China during this severe ionospheric storm event.

5 Comparison of Methods in Bound Estimation

GIVE monitor is designed to improve the overbounding effectiveness with various methods, including the irregularity detector, error inflating and undersampled threat model. Larger margin in the bound could overbound the error more effectively.

Fig. 2 Gaussian bounding for data of Sep. 10 to Sep. 13, 2005



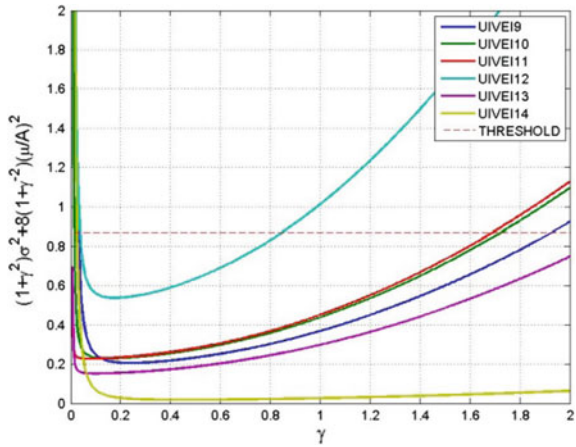
However, an excessively large bound would reduce the system availability. It is necessary to analyze the effectiveness of various methods in the bound estimation after its effectiveness has been test with the Gaussian-bounding theorem.

5.1 Threshold of Irregularity Detector

Increasing the threshold of irregularity detector will allow measurement with larger error passed into the grid model construction. Consequently, the chi-square statistic value χ^2 will increase at the grid point. Since the inflation factor R_{irreg}^2 in bound estimation depends on χ^2 proportionally, larger chi-square value will lead to a larger inflation factor, and a larger error bound as the consequence. At the same time, an increasing threshold for irregularity detector will allow measurement with larger error to be used to construct the undersampled threat model (one should remember the threat model is established only with the observation not trigged the detector). Therefore, a larger undersampled threat model would be resulted. Taking all these factors into account, the increase of the irregularity detector threshold will eventually lead to an increasing error bound estimate.

On the other hand, larger detecting threshold will increase the error in grid model, since measurement with larger error is applied to the grid model establishment. Therefore, the increase of irregularity detector threshold will increase the error at the same time. As a result, the increasing error bound does not assure the effective overbounding of the larger error, or on the contrary even a loss of overbounding capability could happen. Meanwhile, a larger error bound will also reduce the availability of the system. Therefore it is essential to design the irregularity detector with reasonable threshold to maintain the system availability while effectively overbound the error.

Fig. 3 Gaussian bounding with irregularity detector threshold of 3, dynamic inflation factor and simplified spatial threat model



Figures 3, 4 and 5 show the results of Gaussian bounding threshold test with different irregularity detector thresholds of 3, 5 and NaN (no irregularity detection is conducted). The date of the measurements is September 12, 2005 (the same data set were also used in the following analysis). For a consistent result, the dynamic inflation factor algorithm and the undersampled threat model established with the detecting threshold of 3 are adopted.

It can be seen that the estimated ionospheric delay error and its bound all can satisfy the Gaussian bounding test when the irregularity detector thresholds are 3 and 5 respectively. For the result of threshold of 5, the range of γ satisfying the threshold test narrowed for UIVEI = 12, which means a decreased integrity margin. The Gaussian bounding test shows that the system does not meet the 1.15×10^{-8} probability of HMI requirement with the irregularity detector turned off since no value of γ can be found where all curves lie below the threshold.

Fig. 4 Gaussian bounding with irregularity detector threshold of 5, dynamic inflation factor and simplified spatial threat model

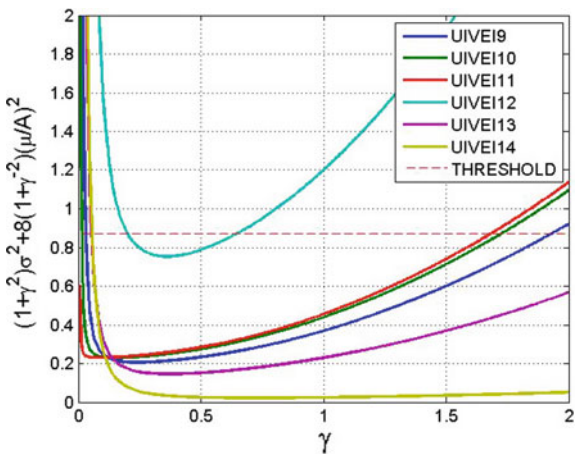
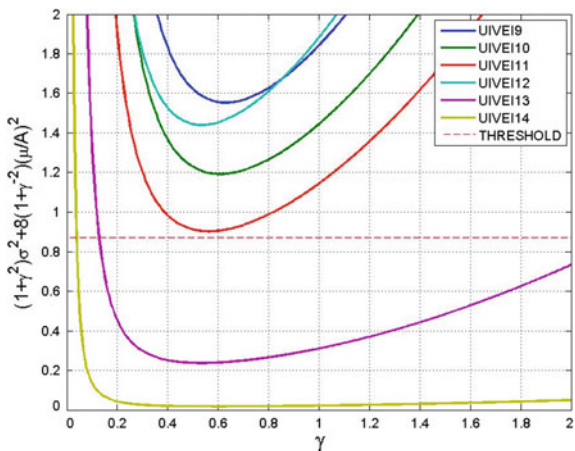


Fig. 5 Gaussian bounding with no irregularity detector, dynamic inflation factor and simplified spatial threat model



5.2 Static Inflation and Dynamic Inflation

There are two algorithms for the inflation factor in the error bound estimation: the static inflation factor and the dynamic inflation factor [9, 12].

The static inflation factor is determined by taking the ratio of the upper threshold to the lower bound of the chi-square statistic for the allocated probability of false alarm P_{fd} and missed detection P_{md} . In the static factor approach, the upper threshold is actually the maximum value in the chi-square statistic with the allowable false alarm rate. In the dynamic factor approach, the upper threshold is replaced by the actual value of the chi-square statistic. Therefore, the static factor approach obtains a larger error bound estimate, and consequently has a wider γ value range below the threshold in the Gaussian bounding test. However, the increase of the bound inevitably leads to the decrease of system availability.

Figure 6 shows the results of the Gaussian bounding test using the static inflation factor. The threshold for the irregularity detector is 3, and the same undersampled threat model as in Sect. 5.1 is used. Compared with the result in Fig. 3, it can be seen that with the static inflation factor, the range of γ satisfying the Gaussian bounding test increased, indicating a larger margin in the estimated error bound.

Figures 7 and 8 show the availability of SBAS in China with the error bound estimate with the dynamic inflation factor and the static inflation factor respectively. It can be seen that the system availability of the static inflation factor method is reduced by about 5.7% compared with that of the dynamic inflation factor method.

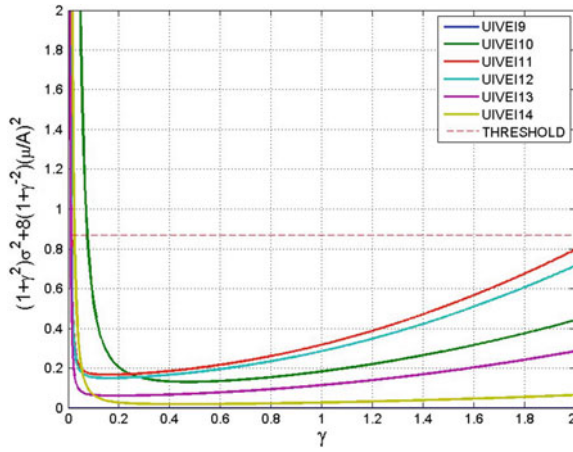


Fig. 1.6 Gaussian bounding with irregularity detector threshold of 3, static inflation factor and simplified spatial threat model

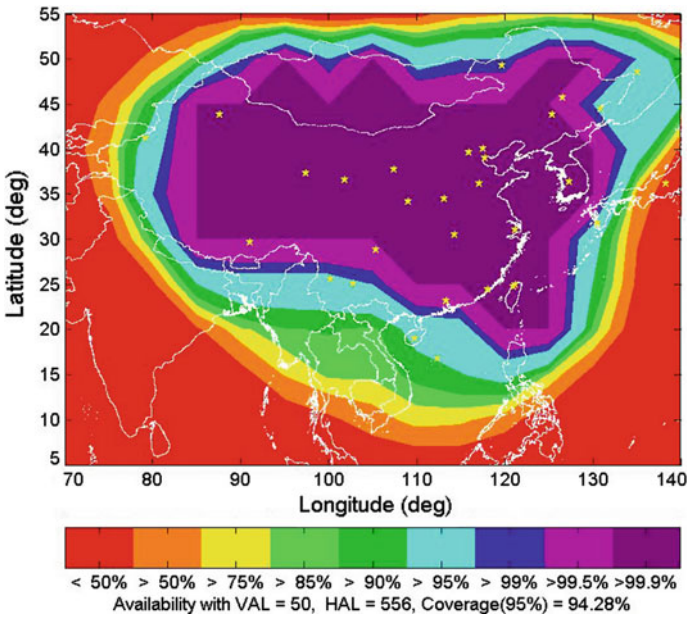


Fig. 7 Availability with dynamic inflation scheme in GIVE monitor for China SBAS

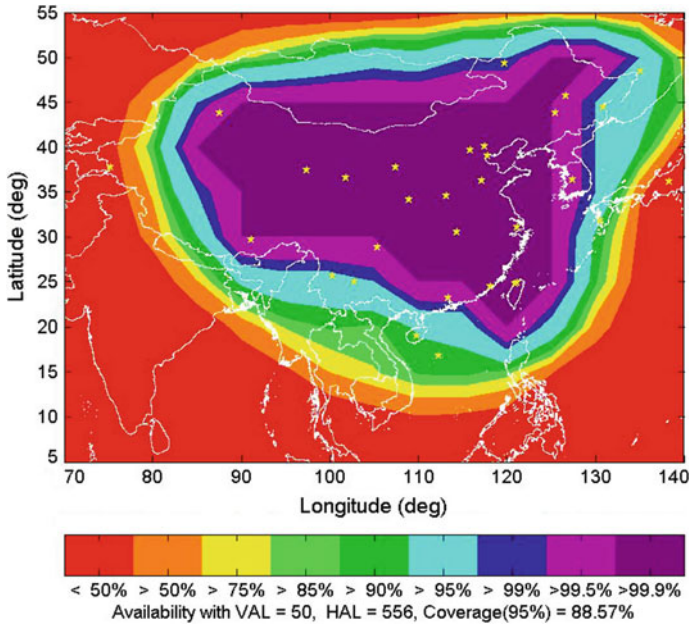


Fig. 8 Availability with static inflation scheme in GIVE monitor for China SBAS

5.3 Undersampled Threat Model

In order to analyze the impact of undersampled threat model, the threat model used in previous analysis was enlarged by 0.8 times and 1.3 times respectively for the error bound estimation. For others, the irregularity detector threshold is taken as 3, and the dynamic inflation factor is adopted. Figures 9 and 10 show the results of Gaussian bounding test with these different undersampled threat model. It can be seen that the result using the 0.8 times threat model cannot pass the threshold test, whereas the result using the 1.3 time threat model satisfy the threshold test with greater integrity margin.

It should be noted that although the ionospheric storm in September 2005 was a strong event (K_p index reached 9), the year of 2005 seen a less active solar activity. So the undersampled threat model previously used in the paper did not constructed with the most strict ways, such as the data depriving strategy [13, 16]. Consequently, this simplified threat model has the maximum value of about 2 m, whereas the maximum value output by the completed threat model (with data depriving strategy) is about 3.8 m [16].

With the simplified threat model, the error bound estimate can effectively overbound the delay error under the conditions of strong ionospheric storm and ionospheric anomaly for a less active solar period, as has shown in previous sections. With the threat model reduced by a factor of 0.8, the Gaussian bounding test

Fig. 9 Gaussian bounding with irregularity detector threshold of 3, dynamic inflation factor and 0.8 times spatial threat model

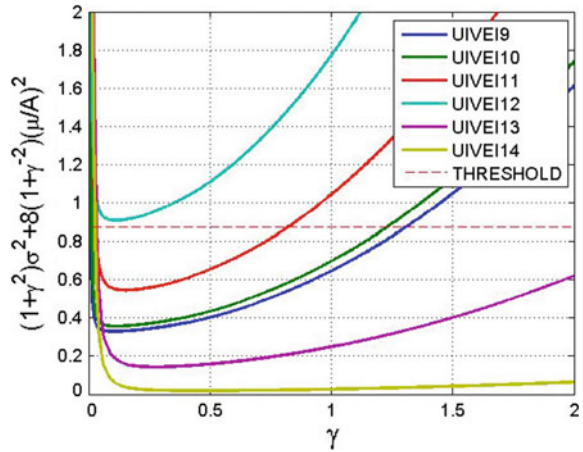
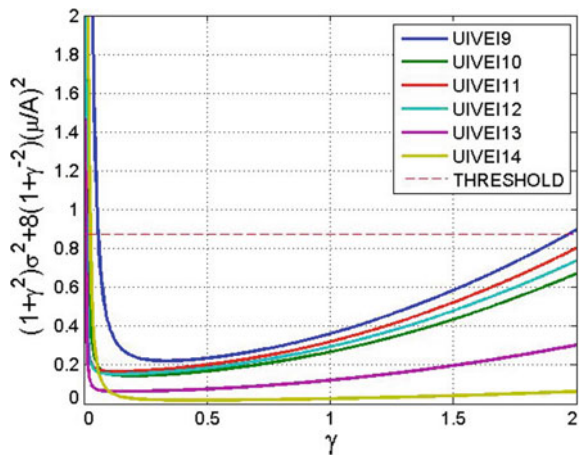


Fig. 10 Gaussian bounding with irregularity detector threshold of 3, dynamic inflation factor and 1.3 times spatial threat model



cannot be passed for this specific ionospheric condition (strong ionospheric storm and ionospheric anomaly under less solar activity). However, it can be seen that the test result is not far from a satisfied one as only the curve of UIVEI = 12 critically lies above the threshold. It could be expected the reduced threat model may still meet the Gaussian bounding test when used for ionospheric conditions of quiet period or low solar activity. On the other hand, although the enlarged threat model has a larger bounding margin for this specific event, considering the enlarged maximum value of the simplified threat model (about 2.6 m) is much smaller than that of the completed one (about 3.8 m). One can expected that even with this enlarged threat model the Gaussian bounding test could not be satisfied during the ionospheric storms in the year of high solar activity.

Therefore, the choice of the undersampled threat model has to be designed carefully in error bound estimation. If a completed threat model was adopted at all

times (high and low solar activity, quiet and stormy ionosphere), the estimated error bound would have excess margin for most of the time. It will improve the effectiveness of error overbounding, but at the cost of system availability. One way for improvement is to establish diverse undersampled threat models based on the ionospheric activities. A specific one would be selected depending on the actual ionospheric activity to achieve enough margin in bounding effectiveness and availability as well.

6 Conclusions

SBAS protects the users from ionospheric error by broadcasting delay correction and its error bound, and assures the error is effectively overbounded with the allocated probability of HMI. Generally, assumption of zero-mean normal error distribution has to be made in the integrity risk analysis.

Grid model under low-latitude ionospheric anomaly and storms would have larger delay error and eventually a biased error distribution, making the error effective overbounding a difficult work. Gaussian bounding theorem provides a criteria for overbounding problem in integrity risk analysis. The theorem is used to analyze the overbounding effectiveness of the estimated bound on error for SBAS in China under the typical condition of storm and ionospheric anomaly in modest solar activity. The results show that the error can be effectively overbounded.

The effectiveness of different methods in error bound estimate are further analyzed. The results show that the irregularity detector is an integral part of SBAS integrity associated with ionospheric error. A reasonable detecting threshold can reduce the error bound estimate and improve the system availability. The dynamic inflation factor will result in a smaller error bound estimate and improve the system availability when compared with the static factor. Undersampled threat model is a tricky factor to cope with, a variety of threat models should be established to effectively overbound the error and maintain the system availability as well for different solar activities.

Acknowledgements The author acknowledges the GNSS data made available from the CMONCO.

References

1. RTCA Special Committee 159 (2001) Minimum operational performance standards for airborne equipment using global positioning system/wide area augmentation system, RTCA/DO-229C, Nov 2001
2. Blanch J (2003) Using Kriging to bound satellite ranging errors due to the ionosphere. PhD thesis, Department of Aeronautics and Astronautics, Stanford University

3. Timothy RS, Authur LR (2002) Application of Gaussian overbounding for the WAAS fault free error analysis. In: Proceedings of the ION GPS 2002, Institute of Navigation, Portland, OR, pp 766–772
4. Liu D, Chen L, Zhen WM (2013) Ionospheric spatial correlation analysis for China area. In: Proceedings CSNC 2013, Wuhan, China, 15–17 May 2013
5. Liu D, Yu X, Chen L, Zhen WM (2017) Analysis on ionospheric delay variogram realization in China area. In: Proceedings CSNC 2017, Shanghai China, 15–17 May 2017, pp 119–130
6. Bruce D (2000) Defining pseudorange integrity—overbounding. In: Proceedings ION GPS 2000, Institute of Navigation, Salt Lake City, UT, pp 1916–1924
7. Timothy RS (2003) WAAS error bounding during ionospheric storms. In: Proceedings ION NTM 2003, Institute of Navigation, Anaheim, CA, pp 175–182
8. Sparks L, Blanch J, Pandya N (2011) Estimating ionospheric delay using kriging: 1. Methodology, Radio Sci 46:RS0D21. <https://doi.org/10.1029/2011rs004667>
9. Sparks L, Blanch J, Pandya N (2011) Estimating ionospheric delay using kriging: 2. Impact on satellite-based augmentation system availability. Radio Sci 46:RS0D22. <https://doi.org/10.1029/2011rs004781>
10. Sparks L, Blanch J, Pandya N (2013) Kriging as a means of improving WAAS availability. In: Proceedings ION GNSS 2013, Institute of Navigation, Portland, OR, pp 2013–2020
11. Pandya N, Sheng F, Castaneda O et al (2012) Using Kriging to optimize WAAS performance over the entire solar cycle. In: Proceedings ION GNSS 2012, Institute of Navigation, Portland, OR, pp 1310–1333
12. Walter T, Hansen A, Blanch J et al (2000) Robust detection of ionospheric irregularities. In: Proceedings ION GPS 2000, Institute of Navigation, Salt Lake City, Utah, pp 209–218
13. Sparks L, Komjathy A, Manucci A (2005) Extreme ionospheric storms and their impact on WAAS. In: Proceedings of the ionospheric effect symposium 2005, Alexandria, VA, May 2005
14. Eric A, Daniel C, Helena G (2002) Improvements to the WAAS ionospheric algorithms. In: Proceedings ION GPS 2002, Institute of Navigation, Portland, OR, pp 2256–2261
15. Altshuler ES et al (2001) The WAAS ionospheric spatial threat model. In: Proceedings ION GPS 2000, Institute of Navigation, Salt Lake City, UT, pp 2463–2467
16. Liu D, Feng J, Chen L, Zhen WM (2017) A study on construction of ionospheric spatial threat model over China area. In: Proceedings CSNC 2017, Shanghai, China, 15–17 May 2017, pp 195–208

Analysis on Characteristics of Delay Errors Under Ionospheric Anomaly in China Area



Dun Liu, Liang Chen, Xiao Yu and Weimin Zhen

Abstract Ionospheric anomaly in south of China will degrade the performance of SBAS grid models severely. Characteristics of ionospheric delay errors are analyzed with observations from the mid and low latitudes area in China for various solar activities and ionospheric conditions. Results show the error distribution is symmetric, unimodal, and overbounded by a biased normal error distribution. The biased distributions would reduce the overbounding capability of the estimated bounds. Ionospheric anomaly contribute to the complicated spatial and temporal variations of the delay errors as storms, and even is the main influencing factors under certain condition.

Keywords Ionosphere · Grid model · Space-based augmentation system (SBAS) Integrity · Error overbounding

1 Introduction

Space-based augmentation system (SBAS) has strict requirements on integrity. Ionospheric grid model is adopted in SBAS to broadcast delay correction GIVD (Grid Ionospheric Vertical Delay) and corresponding error threshold GIVE (Grid Ionospheric Vertical Error) to users. The integrity requirement associated with ionospheric errors is achieved with the delay error overbounded effectively by the broadcast threshold [1].

A variety of error overbounding methods are defined in SBAS. Generally the zero-mean normal distribution has been assumed for the error in these methods, such as the ones defined on “Probability Distribution Function (PDF)” overbounding or “Cumulative Distribution Function (CDF)” overbounding [2, 3]. In the pseudorange integrity overbounding definition developed by Bruce DeCleene, the

D. Liu (✉) · L. Chen · X. Yu · W. Zhen
No. 22nd Research Institute, CETC, Qingdao 266107, Shandong, China
e-mail: Dun.L@163.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_53

requirement is relaxed to the error distribution that “is symmetric, unimodal, and whose cumulative distribution function is bounded by a normal error distribution” [4].

Grid model error follows the normal distribution well for a quiet ionosphere in the mid-latitudes, but it will not be the case under ionosphere storms [3, 5]. The ionospheric anomaly at low latitude impacts the grid model in the similar way of a modest storm. Previous work has shown that the ionospheric delay error does not conform to the normal distribution for SBAS in China where the ionospheric anomaly exists in its southern low latitude region [6, 7]. It is necessary to make a research on the errors of grid model under the conditions of ionospheric anomaly and storms to find effective overbounding ways for SBAS integrity application.

In this paper, the behaviors of the grid model errors are studied with GNSS observations under different solar activities and ionosphere conditions for China area. The second section makes a brief introduction to the ionospheric grid model realized with Kriging method. The third section gives the ways of analysis. The fourth section makes the analysis and a further discussion. Finally, the conclusion is given.

2 GIVE Monitor

In SBAS, the GIVE monitor is responsible to calculate the GIVDs and GIVEs to be broadcast and to ensure that the vertical delay error bound, computed by the user at an ionospheric pierce point (IPP) associated with a GNSS signal ray path, bounds the actual delay error with a sufficiently high probability for each GNSS signal detected by the user’s receiver [8–11].

2.1 Grid Model with Kriging Method

GIVE monitor implemented with Kriging method defines the delay at IPP x_k as:

$$I_{\text{meas}}(x_k) = a_0 + a_1 x_k^{(\text{east})} + a_2 x_k^{(\text{North})} + r(x_k) + m(x_k) \quad (1)$$

The first three terms define the underlying planar trend (deterministic component), the fourth term, r , depicts the residual field (the process noise), and the fifth term, m , is the measurement noise [3, 8]. The delay $I_{\text{est}}(x)$ at the ionospheric grid point (IGP) x is estimated with weighting delays at the surrounding IPPs.

$$I_{\text{est}}(x) = \sum_{k=1}^n \lambda_k I_{\text{meas}}(x_k) \quad (2)$$

Coefficient $\lambda = [\lambda_1 \dots \lambda_n]^T$ is calculated as:

$$\lambda = \left(W - WG(G^T WG)^{-1}G^T W \right) C(x, x_k) + WG(G^T WG)^{-1}X \quad (3)$$

where, W is the weighting matrix, C the covariance matrix, M the measurement noise matrix, G the observation matrix, and $X = [1, 0, 0]^T$.

The formal estimation error at IGP with Kriging method is:

$$\sigma_{IGP}^2 = \sigma_{process}^2(\lambda) + \sigma_{meas}^2(\lambda) \quad (4)$$

$$\sigma_{process}^2(\lambda) = \lambda^T C(x_k, x_l)\lambda - 2\lambda^T C(x_k, x) + C(x, x) \quad (5)$$

$$\sigma_{meas}^2(\lambda) = \lambda^T M(x_k, x_l)\lambda \quad (6)$$

It can be seen that the statistical uncertainty σ_{IGP}^2 includes two terms, $\sigma_{process}^2$ depicting the error from the limitation of ionospheric shell model, and σ_{meas}^2 originating from the noise in measurements.

2.2 Delay Error Bound Estimation

To overbound the errors effectively, GIVE monitor adopts the irregularity detector to find the possible ionospheric disturbances. In addition, inflating the errors and adopting an undersampled threat model are also used in GIVE monitor [9–11].

2.2.1 Ionospheric Irregularity Detector

The ionospheric irregularity detector is designed in GIVE monitor to set the error bound to the maximum (45 m) when it is triggered. A local irregularity detector is designed to handle the threat that local ionospheric behavior does not conform well to the assumed model around IPPs. A system-wide extreme storm detector is also developed to protect users from the most serious disturbances possibly impacts the overall system [12, 13]. Only the local irregularity detector is considered in the work.

In the latest development, SBAS used an improved chi-square test statistic χ_{irreg}^2 to detect the disturbance [9].

$$\chi_{irreg}^2 = \frac{R_{noise}\chi^2}{\chi_{threshold}^2} \quad (7)$$

where, $\chi_{threshold}^2$ is calculated with the inverse of cumulative density function for the allowable false alarm rate P_{fd} . χ^2 is the statistic providing a means of establishing whether the ionospheric measurements are consistent with the grid model. R_{noise} is a factor that prevents the presence of measurement noise from concealing the magnitude of an ionospheric irregularity. For more details, one can refer to Ref. [9]. If χ_{irreg}^2 exceeds the threshold, the grid point is disturbed by storms, and the GIVE is raised to the maximum (45 m).

2.2.2 Inflation of Formal Estimation Error

There is some probability that the statistical uncertainty as given by Eq. (4) significantly underestimates the actual error in the estimated vertical delay. To allow for this possibility, the GIVE is constructed from an inflated version of the formal error [9–14].

Two definitions are developed for inflation factor. The static inflation factor R_{irreg}^2 is determined by Eq. (8), in which χ_p^2 is the p th quantile of the chi-square distribution, P_{fd} and P_{md} are the probability of false alarm and missed detection respectively [14]:

$$\text{Static } R_{irreg}^2(P_{fd}, P_{md}) = \frac{\chi_{1-P_{fd}}^2}{\chi_{P_{md}}^2} \quad (8)$$

The formula for the dynamic inflation factor R_{irreg}^2 is shown in Eq. (9), in which $\chi_{N,lowerbound}^2$ is a system parameter determined by the number of the observation and the probability of hazardous misleading information (HMI) allocated to ionospheric errors [9, 13].

$$\text{Dynamic } R_{irreg}^2 = \frac{R_{noise}\chi^2}{\chi_{N,lowerbound}^2} \quad (9)$$

The inflated version of the formal error at IGP is:

$$\tilde{\sigma}_{IGP}^2 = R_{irreg}^2 [\lambda^T C(x_k, x_l) \lambda - 2\lambda^T C(x_k, x) + C(x, x)] + \lambda^T M \lambda \quad (10)$$

2.2.3 The Ionospheric Undersampled Threat Model

In addition to inflating GIVEs to account for the statistical uncertainty of the formal error variance, SBAS also increases the GIVEs to provide protection from delay estimation error due to the undersampling of irregularities in the ionosphere. Such irregularities can increase the delay experienced by a signal propagating toward a

user while not affecting the signals received by SBAS, causing the user's computed position integrity bound to underestimate significantly the true error in the user's position estimate.

The values of the undersampled ionospheric threat model $\sigma_{undersampled}^2$ is tabulated as a function of the two metrics characterizing the IPPs distribution in the neighborhood of IGPs, namely the fitting radius R_{fit} and the relative centroid metric (RCM) [9, 13, 15]. With the threat model, the overbounding error variance used to define the GIVEs at IGPs can be expressed formally as:

$$\sigma_{GIVE}^2 = \tilde{\sigma}_{IGP}^2 + \sigma_{undersampled}^2 \quad (11)$$

3 Analysis of Ionospheric Delay Error

The key to realize SBAS integrity associated with ionospheric error is to effectively bound the residual after the grid model correction applied. No matter what overbounding mechanism will be adopted, however, the full knowledge of the error distribution is required.

3.1 GNSS Data

GNSS data of the CMONCO (Crustal Movement Observation Network of China) and the IGS sites in China and surrounding areas are used (Fig. 1) in the work.

The data span a variety of solar activities and ionospheric conditions (quiet and stormy). The Kp is adopted as the metric for comparing the magnitude of ionospheric storms. Table 1 details the typical disturbance events in the analysis.

3.2 Analysis Method

The method of analysis is as follows:

- (1) XIAM (24.4°E) in low latitude region and TAIN (36.2°E) in mid latitude region are selected as the test sites (hereinafter referred to as test site), other stations are used for modeling (modeling site).
- (2) Slant ionospheric delays are extracted from GNSS dual-frequency observations, and then converted to the vertical delays at the IPPs for the satellite-to-site signal paths for each station in the network.
- (3) Construct the grid model with the data of modeling sites, and then estimate $\hat{I}_{v,IPP}$ at IPPs for test sites at each epoch with the grid interpolating algorithm.

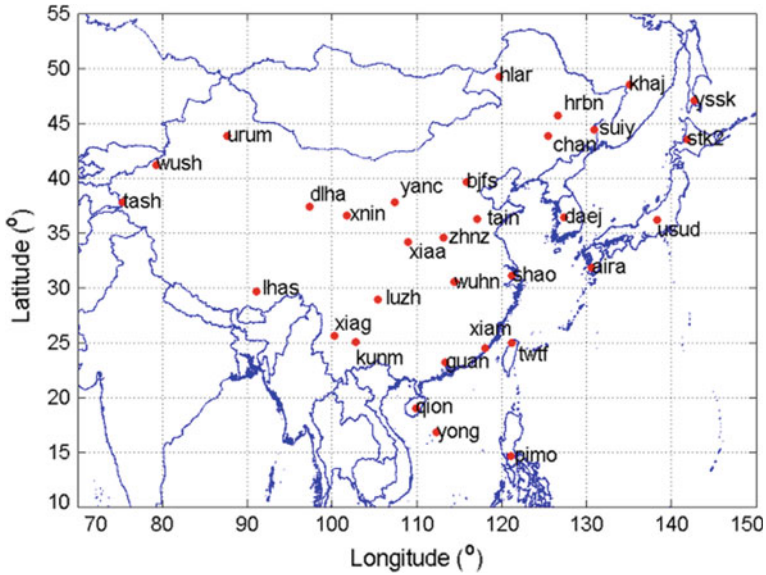


Fig. 1 GNSS network in China and surrounding areas

Table 1 Data used in the analysis

Period of data	Comments
March 18th to 21st, 2001, (DOY 77–80)	(1) Most strong solar activity (2) Kp was 3, 5, 7 for Mar. 18th, 19th, and 20th respectively, indicating a stormy period
October 29th to November 1st, 2003, (DOY 302–305)	(1) Less strong solar activity (2) Kp reached 9 during the October 29th to 31st, indicating the occurrence of strong storm (known as ‘Halloween Event’)
September 10th to 13th, 2005, (DOY 253–256)	(1) Moderate solar activity year (2) Kp reached 9 on September 11th, and was 6 for the other time, indicating a strong storm during the period

- (4) Calculate the delay errors $I_{v,IPP} - \hat{I}_{v,IPP}$ at IPPs with $I_{v,IPP}$ from observations at the test sites and the estimated $\hat{I}_{v,IPP}$.
- (5) Estimate the bounds of delay errors $\sigma_{bound,IPP}$ with Eqs. (4), (10) and (11) respectively. Detection threshold (DT) of 3 and NaN (with irregularity detector inactive) are used for each estimation.
- (6) Normalize ionospheric delay errors as follows with different threshold estimations $\sigma_{bound,IPP}$.

$$\frac{I_{v,IPP} - \hat{I}_{v,IPP}}{\sigma_{bound,IPP}} \tag{12}$$

- (7) Fit the normalized delay errors with normal distribution, and analyze the mean, variance of errors, and the effectiveness of different thresholds on delay errors overbounding.

4 Results and Discussion

4.1 Results of Statistics Analysis

As an example, Figs. 2 and 3 show results of the XIAM site on March 18th 2001. The (a), (b) and (c) panel presents the distributions of normalized errors for different threshold value (σ_{IGP} , $\tilde{\sigma}_{IGP}$ and σ_{GIVE} respectively). Figure 2 is the result for DT = 3, while Fig. 3 gives the result for DT = NaN.

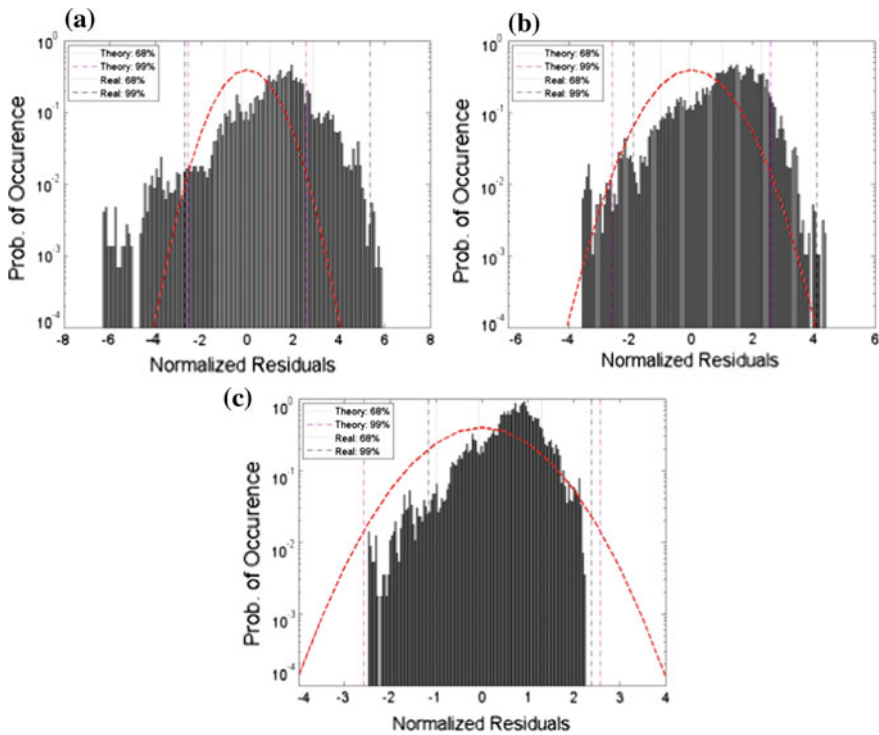


Fig. 2 Distribution of Iono. delay error and its overbounding threshold (XIAM, DT = 3, March 20th, 2001)

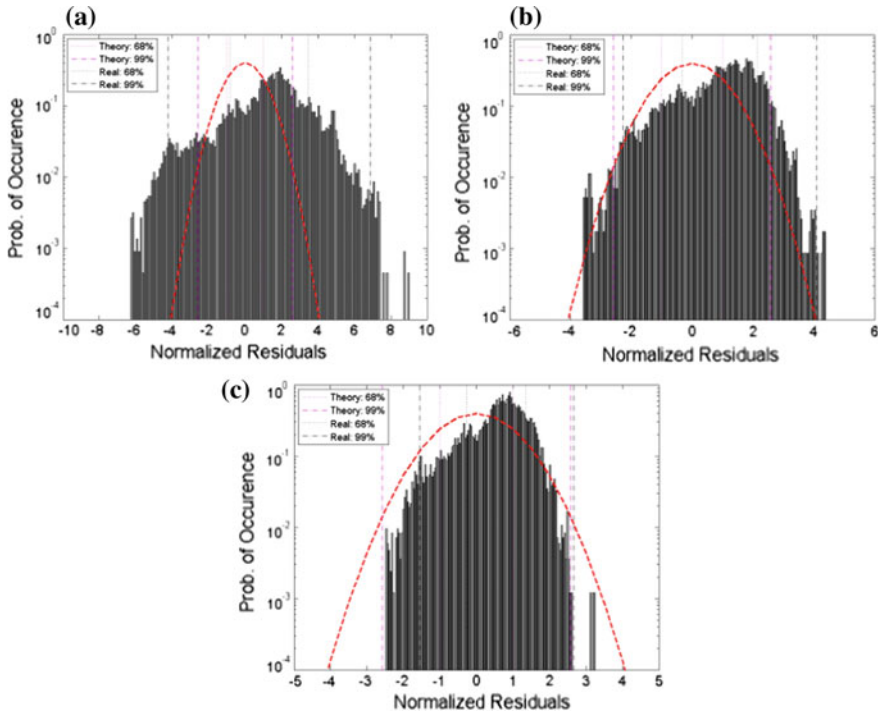


Fig. 3 Distribution of Iono. delay error and its overbounding threshold (TAIN, DT = NaN, March 20th, 2001)

To reduce the influence of ionospheric correlation on the significance of analysis results, delay errors and the thresholds are estimated every 15 min. This will result in a reduction in the available analysis data. So the overbounding probability of 99% is used instead here to illustrate the effect of error distribution on the overbounding effectiveness. As the results, the corresponding thresholds are ± 2.58 . Bounding results of each threshold are given in the figure, with the pink dotted line for the standard normal distribution and the black dotted line for the empirical normal distribution from the real data. The standard normal distribution is also shown (red curve).

Tables 2 and 3 give more results for different DTs and estimated thresholds for sites of XIAM and TAIN on days in Table 1. As can be seen:

(1) Biases exist obviously in empirical delay error distributions

The distributions of delay errors are unimode and symmetric, and biases exist in the fitted normal distribution for the two sites.

For XIAM site, the fitted normal distributions all show a clearly biased feature for data in 2001, 2003 and 2005. The deviation of the mean values during 2001

Table 2 Results of XIAM site, case 1,2,3 for $\sigma_{IGP}, \tilde{\sigma}_{IGP}$ and σ_{GIVE} respectively

		DT = 3			DT = NaN		
		Mean	Var.	Range of 99%	Mean	Var.	Range of 99%
2001.3.18	1	0.52	1.59	[-3.58 4.62]	-0.03	2.86	[-7.41 7.35]
	2	0.47	1.27	[-2.80 3.75]	0.19	1.39	[-3.38 3.77]
	3	0.24	0.74	[-1.65 2.14]	0.05	0.97	[-2.44 2.54]
2001.3.19	1	0.49	1.52	[-3.41 4.40]	-0.09	2.66	[-6.93 6.76]
	2	0.46	1.27	[-2.81 3.73]	0.18	1.26	[-3.08 3.44]
	3	0.23	0.71	[-1.61 2.06]	0.05	0.86	[-2.18 2.27]
2001.3.20	1	1.33	1.58	[-2.74 5.40]	1.32	2.16	[-4.24 6.89]
	2	1.10	1.17	[-1.91 4.11]	0.92	1.24	[-2.27 4.11]
	3	0.61	0.69	[-1.17 2.40]	0.55	0.82	[-1.56 2.66]
2001.3.21	1	1.55	1.67	[-2.76 5.86]	1.14	2.97	[-6.52 8.79]
	2	1.19	1.09	[-1.61 3.99]	0.79	1.33	[-2.64 4.23]
	3	0.71	0.72	[-1.14 2.55]	0.49	0.99	[-2.06 3.04]
2003.10.29	1	-0.24	0.86	[-2.45 1.97]	-0.56	2.65	[-7.38 6.27]
	2	-0.22	0.66	[-1.93 1.48]	-0.25	0.91	[-2.59 2.10]
	3	-0.12	0.38	[-1.09 0.86]	-0.17	0.71	[-1.98 1.65]
2003.10.30	1	-0.35	0.95	[-2.80 2.11]	-0.35	0.96	[-2.82 2.12]
	2	-0.30	0.82	[-2.40 1.81]	-0.30	0.82	[-2.40 1.81]
	3	-0.16	0.44	[-1.30 0.98]	-0.16	0.44	[-1.31 0.98]
2003.10.31	1	-0.39	1.44	[-4.10 3.32]	-0.64	1.74	[-5.12 3.83]
	2	-0.26	1.03	[-2.91 2.39]	-0.36	1.08	[-3.14 2.41]
	3	-0.17	0.63	[-1.78 1.44]	-0.25	0.70	[-2.05 1.54]
2003.11.1	1	-0.15	1.03	[-2.79 2.49]	-0.93	3.13	[-8.99 7.12]
	2	-0.15	0.84	[-2.32 2.01]	-0.32	1.09	[-3.13 2.50]
	3	-0.08	0.46	[-1.27 1.12]	-0.25	0.84	[-2.41 1.92]
2005.9.10	1	0.14	0.72	[-1.72 2.00]	0.14	0.72	[-1.72 2.00]
	2	0.16	0.66	[-1.55 1.87]	0.16	0.66	[-1.55 1.87]
	3	0.07	0.35	[-0.83 0.97]	0.07	0.35	[-0.83 0.97]
2005.9.11	1	0.31	0.67	[-1.43 2.05]	0.03	0.71	[-1.52 2.12]
	2	0.32	0.57	[-1.15 1.80]	0.32	0.58	[-1.18 1.82]
	3	0.15	0.31	[-0.66 0.96]	0.15	0.32	[-0.68 0.98]
2005.9.12	1	0.18	0.99	[-2.36 2.72]	0.15	1.04	[-2.53 2.83]
	2	0.21	0.76	[-1.74 2.16]	0.19	0.77	[-1.79 2.17]
	3	0.09	0.44	[-1.05 1.23]	0.08	0.46	[-1.10 1.25]
2005.9.13	1	0.32	0.66	[-1.38 2.03]	0.32	0.66	[-1.38 2.03]
	2	0.33	0.59	[-1.18 1.84]	0.33	0.59	[-1.18 1.84]
	3	0.16	0.32	[-0.66 0.97]	0.16	0.32	[-0.66 0.97]

Table 3 Results of TAIN site, case 1,2,3 for σ_{IGP} , $\tilde{\sigma}_{IGP}$ and σ_{GIVE} respectively

		DT = 3			DT = NaN		
		Mean	Var.	Range of 99%	Mean	Var.	Range of 99%
2001.3.18	1	0.43	0.65	[-1.23 2.09]	0.44	1.3	[-2.92 3.80]
	2	0.39	0.55	[-1.04 1.81]	0.34	0.61	[-1.23 1.90]
	3	0.20	0.29	[-0.55 0.95]	0.18	0.39	[-0.83 1.19]
2001.3.19	1	0.38	0.67	[-1.33 2.10]	0.29	1.46	[-3.46 4.04]
	2	0.33	0.55	[-1.09 1.75]	0.27	0.67	[-1.45 1.99]
	3	0.17	0.30	[-0.59 0.94]	0.14	0.45	[-1.03 1.31]
2001.3.20	1	0.20	0.76	[-1.762.15]	0.19	0.84	[-1.96 2.34]
	2	0.17	0.65	[-1.50 1.85]	0.16	0.64	[-1.491.80]
	3	0.09	0.34	[-0.79 0.97]	0.08	0.35	[-0.82 0.99]
2001.3.21	1	0.25	0.73	[-1.612.12]	0.23	0.83	[-1.92 2.38]
	2	0.21	0.59	[-1.311.72]	0.19	0.61	[-1.37 1.76]
	3	0.11	0.32	[-0.71 0.94]	0.10	0.35	[-0.79 1.00]
2003.10.29	1	0.12	0.56	[-1.31 1.55]	0.08	0.84	[-2.08 2.24]
	2	0.09	0.48	[-1.14 1.32]	0.08	0.49	[-1.20 1.35]
	3	0.05	0.25	[-0.59 0.69]	0.04	0.29	[-0.70 0.78]
2003.10.30	1	-0.10	0.42	[-1.18 0.97]	0.06	0.60	[-1.49 1.60]
	2	-0.10	0.42	[-1.18 0.97]	-0.05	0.40	[-1.09 0.99]
	3	-0.05	0.20	[-0.56 0.46]	-0.01	0.22	[-0.56 0.55]
2003.10.31	1	0.02	0.67	[-1.71 1.76]	0.02	0.67	[-1.71 1.76]
	2	0.04	0.62	[-1.56 1.63]	0.04	0.62	[-1.56 1.63]
	3	0.01	0.31	[-0.79 0.82]	0.01	0.31	[-0.79 0.82]
2003.11.1	1	0.08	0.47	[-1.12 1.29]	0.10	0.58	[-1.39 1.59]
	2	0.08	0.44	[-1.05 1.21]	0.09	0.45	[-1.07 1.24]
	3	0.04	0.22	[-0.52 0.60]	0.04	0.24	[-0.57 0.66]
2005.9.10	1	-0.07	0.39	[-1.08 0.93]	-0.07	0.39	[-1.08 0.93]
	2	-0.07	0.39	[-1.08 0.93]	-0.07	0.39	[-1.08 0.93]
	3	-0.03	0.19	[-0.51 0.45]	-0.03	0.19	[-0.51 0.45]
2005.9.11	1	-0.04	0.39	[-1.06 0.97]	-0.04	0.39	[-1.06 0.97]
	2	-0.04	0.39	[-1.06 0.97]	-0.04	0.39	[-1.06 0.97]
	3	-0.02	0.19	[-0.51 0.46]	-0.02	0.19	[-0.51 0.46]
2005.9.12	1	0.08	0.42	[-1.01 1.17]	0.08	0.42	[-1.01 1.17]
	2	0.08	0.42	[-1.01 1.17]	0.08	0.42	[-1.01 1.17]
	3	0.04	0.20	[-0.48 0.56]	0.04	0.20	[-0.48 0.56]
2005.9.13	1	0.16	0.42	[-0.93 1.26]	0.16	0.42	[-0.93 1.26]
	2	0.16	0.42	[-0.93 1.26]	0.16	0.42	[-0.93 1.26]
	3	0.08	0.20	[-0.44 0.60]	0.08	0.20	[-0.44 0.60]

(most strong active solar period) are greater than the ones during a strong ionospheric storm in 2003 when the solar activity was declining. For the variance of the delay errors, the results in 2001 are similar to that of the 2003 stormy period. While at other times the variance of the delay errors during 2003 is smaller than that of 2001. The variance of error during 2005 (modest solar activity) is significantly reduced.

For TAIN site, only the delay errors in 2001 obviously show the biased feature, while the deviation of delay errors in 2003 and 2005 is very small. At the same time, it can be seen that the variance of delay errors during the period of strong ionospheric storm in 2003 is similar to the ones during 2001, or even smaller at some times. During 2005, the variance of error significantly decreased.

The XIAM site is located in the area where the low-latitude ionospheric anomaly impacts the grid model most seriously. Consequently, the delay error from the model is affected most seriously to show the characteristics of long-term biased distribution (for year of 2001, 2003 and 2005). Although the TAIN site is located in the mid-latitude region, the estimated correction could also be affected resulting a biased error distribution due to the southward observations with low-elevation angles passing through the low latitude ionosphere anomaly.

(2) Deviation of delay error affects effectiveness of overbounding

For both sites, 99% of the delay errors could fall within the $[-2.58, 2.58]$ range to show an effective overbounding by the normal distribution, but the effective threshold estimates were achieved with different algorithms. For XIAM site, the delay errors can be effectively overbounded by increasing the error threshold with inflating factor or taking the undersampled threat model into account after the irregularity detection was made. For TAIN site however, the delay errors can be effectively overbounded by the formal error with the irregularity detector is applied.

The biased error distribution impacts adversely the overbounding effectiveness. Taking the data of XIAM site on March 20th 2001 as an example, 99% of the delay errors falls in the range $[-1.91, 4.11]$ when the bound is estimated with the irregularity detection and error inflating algorithms. It can be seen the lower bound has been less than the threshold of -2.58 , but the upper bound is much greater than the corresponding threshold of 2.58 , meaning an effective error overbounding at the left but not the right. The overbounding achieved at both sides only when the error bound is more inflated with the threat model added.

As an integral part of GIVE monitor, the irregularity detection mechanism plays an important role in the error overbounding as the detector is designed to find out the possible delay estimations with large error in the grid model, and when triggered increase the bound to the maximum to maintain the allocated probability of PHI.

It should be noted that the effective overbounding of the delay errors is achieved by increasing the bound with a variety of error sources accounted. On the other hand, however, a larger bound means a reduced system availability. Therefore, a trade-off should be made between the effective overbounding and the system

availability [3]. In addition, overbounding with 99% probability is used here only for illustration of the effect a biased error distribution on the effective bound estimation. In real system however, more stringent approaches, such as Gaussian-bounding theorem, should be resorted to validate the bound effectiveness.

4.2 Further Discussion

The essence of the ionospheric grid model is planer interpolation based on the shell model. When there is large spatial gradients in the regional delays, the assumption is broken down, resulting large errors in the estimated ionospheric delay corrections. Situations causing large spatial delay gradients include ionospheric storms, and the ionosphere anomaly in low latitude areas.

In years of high solar activity, large delay errors could also be present in areas where ionospheric anomaly exist. The error caused by large spatial delay gradients in ionospheric anomaly at high solar activity, is even greater than that of a strong ionospheric storm in a modest solar activity year, as has been shown in the event of January 19th, 2001.

Ionospheric storm happens occasionally, while the ionospheric anomaly is an inherent feature of low latitude area. It is expected that the ionospheric anomaly in south of China would be a more influencing factor for SBAS in China.

5 Conclusion

The grid model performance at sites in low latitude region of China degraded by the ionospheric anomaly. Sites in the mid-latitude region could also be impacted when the southward low-elevation observations cross the anomaly area of the ionosphere. The effects make the model error distribution biased. Variation of the biased normal distribution depends on the intensity of solar activity, as well as the ionospheric conditions (quite or stormy).

Bound could be established to cover the delay errors effectively when various error sources were taken into account for SBAS in China. But its effectiveness will be impact by the biased error distribution. Increasing bound leads to more margin in integrity realization, but it achieves at the cost the system availability. Tradeoff has to be made between the effective overbounding and the system availability. In addition, more stringent test method, such as the Gaussian bounding theorem, is also required for the bounding validation analysis.

Both the ionospheric anomaly and storms could exert serious influence on SBAS in China, making delay errors exhibit complex temporal and spatial variations. Under certain conditions, the ionospheric anomaly become a more serious factor to be dealt with.

Acknowledgements The author acknowledges the GNSS data made available from the CMONCO.

References

1. RTCA Special Committee 159 (2001) Minimum operational performance standards for airborne equipment using global positioning system/wide area augmentation system, RTCA/DO-229C, Nov 2001
2. Timothy RS, Authur LR (2002) Application of Gaussian overbounding for the WAAS fault free error analysis. In: Proceedings ION GPS 2002, Institute of Navigation, Portland, OR, pp 766–772
3. Blanch J (2003) Using Kriging to bound satellite ranging errors due to the ionosphere. PhD thesis, Department of Aeronautics and Astronautics, Stanford University
4. Bruce D (2000) Defining pseudorange integrity—overbounding. In: Proceedings ION GPS 2000, Institute of Navigation, Salt Lake City, UT, pp 1916–1924
5. Timothy RS (2003) WAAS error bounding during ionospheric storms. In: Proceedings ION NTM 2003, Institute of Navigation, Anaheim, CA, pp 175–182
6. Liu D, Chen L, Zhen WM (2013) Ionospheric spatial correlation analysis for China area. In: Proceedings CSNC 2013, Wuhan, China, 15–17 May 2013
7. Liu D, Yu X, Chen L, Zhen WM (2017) Analysis on ionospheric delay variogram realization in China area. In: Proceedings CSNC 2017, Shanghai, China, 15–17 May 2017
8. Sparks L, Blanch J, Pandya N (2011a) Estimating ionospheric delay using kriging: 1. Methodology. *Radio Sci* 46:RS0D21. <https://doi.org/10.1029/2011rs004667>
9. Sparks L, Blanch J, Pandya N (2011b) Estimating ionospheric delay using kriging: 2. Impact on satellite-based augmentation system availability. *Radio Sci* 46:RS0D22. <https://doi.org/10.1029/2011rs004781>
10. Pandya N, Sheng F, Castaneda O et al (2012) Using Kriging to optimize WAAS performance over the entire solar cycle. In: Proceedings ION GNSS 2012, Institute of Navigation, Portland, OR, pp 1310–1333
11. Sparks L, Blanch J, Pandya N (2013) Kriging as a means of improving WAAS availability. In: Proceedings ION GNSS 2013, Institute of Navigation, Portland, OR, pp 2013–2020
12. Sparks L, Komjathy A, Manucci A (2005) Extreme ionospheric storms and their impact on WAAS. In: Proceedings of the ionospheric effect symposium 2005, Alexandria, VA, May 2005
13. Liu D, Feng J, Chen L, Zhen WM (2017) A study on construction of ionospheric spatial threat model over China area. In: Proceedings CSNC 2017, Shanghai, China, 15–17 May 2017
14. Walter T, Hansen A, Blanch J et al (2000) Robust Detection of Ionospheric Irregularities. In: Proceedings ION GPS 2000, Institute of Navigation, Salt Lake City, Utah, pp 209–218
15. Altshuler ES et al (2001) The WAAS ionospheric spatial threat model. In: Proceedings ION GPS 2000, Institute of Navigation, Salt Lake City, UT, pp 2463–2467

Design of Mega-Constellations of LEO Satellites for Positioning



Xingchi He and Urs Hugentobler

Abstract With the huge potential of the GNSS market and thousands of LEO satellites in space in the future, a novel way is to use these LEO mega-constellations for positioning. In order to fulfil this purpose, the first step is to build a constellation that could be used for both communication and positioning. This paper discusses about some critical issues when designing LEO mega-constellations for positioning. By statistically analyzing from different perspectives, it is shown that in order to get a more uniform distribution of the number of visible satellites along latitude, the combination of different constellations together as a whole could be much more appealing. Meanwhile, in this paper, we did some preliminary work about how to determine some important parameters to get the most suitable combination. The selection of inclinations, orbit heights, number of combined constellations, and ratio of the number of satellites in each constellation will be discussed.

Keywords Constellation design · LEO satellite · GNSS · Combined constellation Walker constellation

1 Introduction

With the development of the space science, global navigation satellite system (GNSS) has now become an important part of people's daily life. There are more and more applications based on GNSS, from the daily used smartphones, to the automatic pilot in the future. Therefore, on the one hand, traditional spacefaring powers, the U.S. and Russia, begin to develop new phase of space segment for GPS and GLONASS, in order to improve their accuracy and performance. On the other hand, other countries and organizations are actively entering the market. Global navigation systems like Chinese BeiDou and European Galileo, as well as regional

X. He (✉) · U. Hugentobler
Technische Universität München, Arcisstr. 21, 80333 Munich, Germany
e-mail: xingchi.he@tum.de

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_54

navigation systems like Indian IRNSS and Japanese QZSS, are starting to compete with GPS and to share the promising market.

Meanwhile, many companies announced their plans to build low earth orbit (LEO) mega-constellations with massive small satellites, which weigh only several hundreds of kilograms. The purpose of these constellations mainly focuses on communication and high-speed internet coverage.

For instance, the aerospace company SpaceX has announced to launch 4425 satellites into space to provide global internet service [1]. A state-owned Chinese company called China Aerospace & Industry Corporation (CASIC) also announced its plan to build a network of 156 small satellites for global broadband coverage [2]. Other companies such as OneWeb, Boeing, LeoSat, Telesat as well as Samsung also published similar plans to build their own constellations for global internet service and/or communication [3–7].

With such big potential and thousands of LEO satellites in space in the future, one begins to wonder whether it is possible to combine these two fields together. Or more specifically, could we use these LEO mega-constellations for positioning?

Theoretically, a GNSS is nothing but a passive autonomous positioning system which is based on the measurement of signal travel time. If these LEO satellites could carry some essential payload to broadcast the necessary information for navigation, they could form an effective global navigation satellite system.

In this way, the value of these massive LEO constellations can be maximized, since they could not only provide telecom and internet services, but also serve as a high-quality navigation system. Considering the number of these LEO satellites and constellations, they might become powerful competitors against traditional GNSS like GPS and open the opportunities for regular companies entering the market, which has long been monopolized by governments and military.

Therefore, in order to fulfil this purpose, the first step is to build a constellation that could be used for both communication and positioning. This paper discusses some critical issues related to designing LEO mega-constellations for positioning. Traditionally, most of the navigation satellites are in medium earth orbit (MEO). However, due to different priorities of MEO and LEO, some conditions and parameters for a LEO positioning constellation must be reconsidered.

2 State of the Art

Reference [8] gives some general concept about the satellite constellation design. The purpose of its design is mainly for the robust QoS routing technology. Nevertheless, it proposes some basic principles about how to select certain criteria such as orbit type, number of orbit planes and satellites, as well as orbit altitude.

References [9–11] study the coverage of the constellation over the Earth. In references [9, 10], different constellation design algorithms are proposed and compared. For the purpose of multiple levels of continuous coverage, it is proven that the symmetric, inclined constellations are more efficient. Reference [11] seeks

to find optimal constellations with continuous coverage of the 20–60° of latitude band for LEO satellites.

One way to estimate the coverage of Earth of the constellation is by calculating the satellite visibility. Reference [12] provides some methods to analyse the statistics of low-orbit satellite visibility. A bounding equation is given to determine the percentage of time that a low-orbit satellite would occur in certain regions over long periods of time. What's more, it also gives the probability density function of the position of a low-orbiting satellite on the orbit shell. It represents the probability density if there are infinite number of satellites sampling randomly over the Earth. While reference [13] is more interested in the mean visibility time of a non-geostationary satellite with non-repeating ground tracks seen by a fixed ground station. They analyse the visibility time along the latitude mathematically, and find out that it coincides with the expected results perfectly.

3 Preliminary Work

For the purpose of using LEO mega-constellations for positioning, firstly one has to find the basic requirement for the constellation. In other words, it has to be considered how a constellation that minimizes the infrastructure costs to fulfill the global positioning service shall be defined. Parameters such as altitude, inclination, number of satellites, number of orbit planes, elevation angle and so on need to be well defined and analyzed.

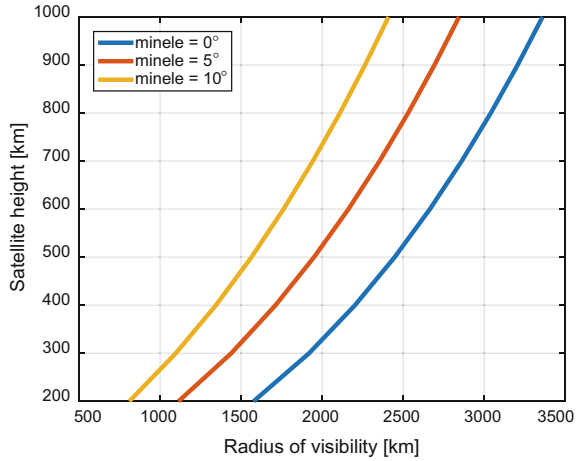
3.1 *Visibility w.r.t. Orbit Height and Elevation Angle*

Figure 1 shows the relationship between the radius of visibility and orbit height, as well as elevation angle. For the same orbit height, one would notice that with smaller elevation angle, the radius of visibility is larger. While it also shows the increase of radius of visibility for increasing orbit height. Increased visibility radius requires fewer satellites but higher launch costs per satellite mass.

3.2 *Visibility w.r.t. Latitude*

According to the principle of GNSS, in general, the ground receiver should receive the signals from at least four satellites to determine the position. In order to provide the global service, this minimum number of visible satellites should be fulfilled at everywhere on Earth. Due to convergence of the orbits towards northern and southern latitudes, the number of observable satellites is a function of latitude. Thus, it is necessary to take this factor into account for the constellation design.

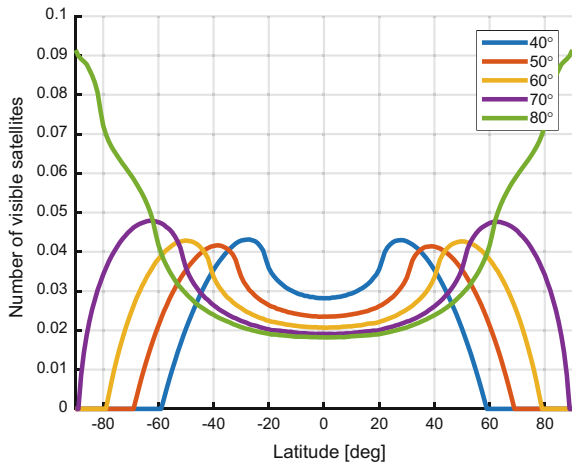
Fig. 1 Radius of visibility w.r.t. satellite height and elevation angle



For example, Fig. 2 is a way to consider this requirement. It shows the average number of visible satellites along the latitude in one revolution for one satellite.

As one could see, for certain orbit height and elevation angle, with larger inclination, the number of visible satellites decreases near the equator, while increases near the latitude corresponding to the inclination. Usually, for a certain inclination, the number of visible satellites first increases with the latitude. It reaches the maximum value at the latitude of $i - \psi$. Here, i is the inclination, ψ means the coverage angle, which is the Earth central angle calculated from the sub-satellite point to the edge of the coverage area. This coverage angle can be calculated from:

Fig. 2 Average number of visible satellites along latitude for orbit height of 800 km and minimum elevation of 10°



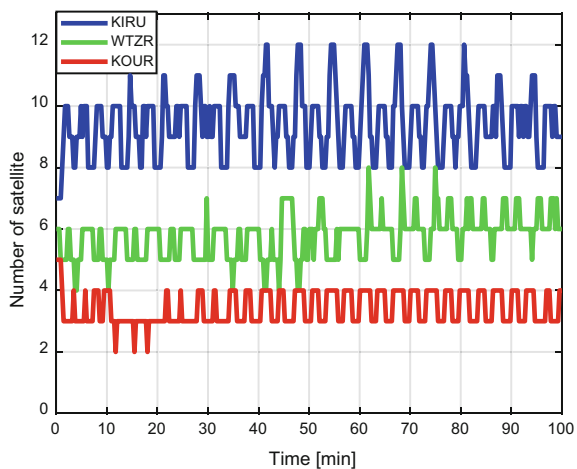
$$\begin{aligned} \psi &= \arcsin \frac{\cos \varepsilon \cdot (\sqrt{r^2 - \cos^2 \varepsilon \cdot R_e^2} - R_e \cdot \sin \varepsilon)}{r} \\ &= \arccos \left(\frac{R_e}{r} \cdot \cos \varepsilon \right) - \varepsilon \end{aligned} \tag{1}$$

Here, the Earth is assumed to be spherical, and its radius is $R_e = 6378$ km. The radius of the satellite orbit is denoted as r . There is $r = R_e + h$, where h is the orbit height. While the minimal elevation angle is ε .

In fact, this is easy to understand since the maximal latitude that the sub-satellite point can reach is the angle of inclination. When the satellite reaches the latitude of inclination, the lower edge of the coverage area happens to be at this latitude corresponding to the $i - \psi$. Meanwhile, the satellite tends to reside longer at higher latitude and the covered longitude is larger with the same coverage area, due to the smaller latitude circle at higher latitude. As a result, the average number of satellites tends to be larger. The only exception is at the inclination near 80° . The average number of visible satellites continues to increase even beyond the latitude of $i - \psi$, which in this case is 61.1° . This is due to the fact that the upper boundary of the coverage area of the satellite $i + \psi$ exceeds 90° . In other words, the satellite could be observed almost all the time when it crosses the polar region. As a consequence, the average number of visible satellites continues to increase up to the poles.

Another perspective could be seen in Fig. 3 which displays the number of satellites of a Walker constellation $74^\circ: 160/10/0$ (inclination of 74° ; total number of satellites 160; 10 equally spaced planes; relative spacing between satellites in adjacent planes 0°) at the orbit height 900 km and above 10° elevation for three sites, one close to the equator (Kourou, red), one in middle Europe (Wetzell, green), and one in northern Europe (Kiruna, blue). The number of visible satellites differs by a factor of nearly three for these examples. The variance among these

Fig. 3 Number of visible satellites as function of time for three stations: orbit height of 900 km, minimum elevation of 10° , Walker constellation of $74^\circ:160/10/0$



three stations is quite obvious. There are more visibilities near the polar region, while less visibilities near the equator. This means that the distribution of the number of the visible satellites is extremely inconsistent.

4 Constellation Design

In order to get a more uniform distribution of the number of visible satellites around the world, it is reasonable to consider the combination of constellations with different inclinations. In this way, the constellation with larger inclination could cover the polar areas, while the constellation with smaller inclination could increase the visibility near the equator. Figure 4 shows an example of the combination of two constellations. It can be seen that with this combination, the new constellation would provide a more homogeneous coverage compared with a single one. It could at the same time maintain a high number of visible satellites at higher latitude and increase the number at lower latitude to a much higher level than any of the two original constellations. Figure 5 gives the result of Fig. 3 for the combined constellation. Comparison of these two figures shows that the combination of two constellations with different inclinations obviously improves the average visibility on Earth. Generally speaking, the number of visible satellites at the three stations are now much closer than the result of a single constellation. This means that the new combined constellation would result in a much uniform coverage than the original one.

Figure 6a shows the number of visible satellites above 10° elevation for the Walker constellation 74°: 80/10/10 at an altitude of 900 km as a function of geographic location at a given time.

Fig. 4 Average number of visible satellites along latitude for combined constellation: orbit height of 900 km, minimum elevation of 10°, Walker constellation of 45°: 1/1/0; with orbit height of 800 km, minimum elevation of 10°, Walker constellation of 75°:1/1/0

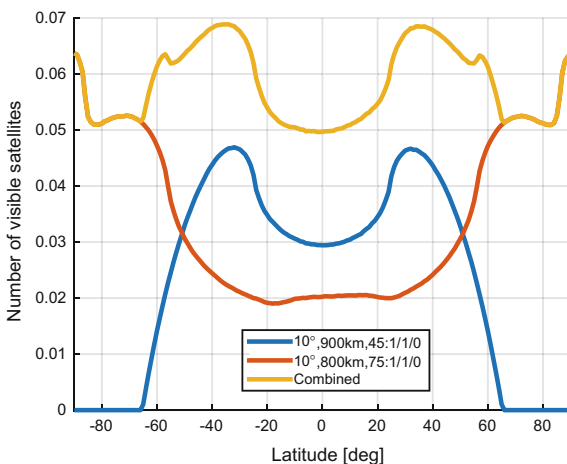
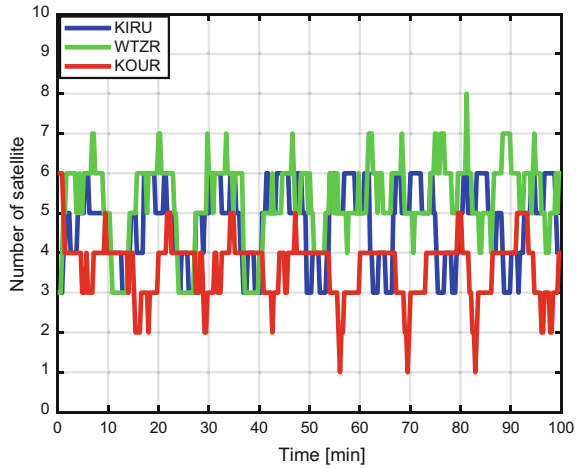


Fig. 5 Number of visible satellites as function of time at three stations for combined constellation: orbit height of 900 km, minimum elevation of 10°, Walker constellation of 74°: 80/10/0 and orbit height of 800 km, minimum elevation of 10°, Walker constellation of 45°: 80/10/0



Again, the uneven distribution as function of latitude can be observed ranging from less than one satellite at low latitudes to up to ten satellites at very high latitudes. Figure 6c on the other hand gives an example of the combined constellation of two Walker constellations (1st: altitude at 900 km, elevation angle 10°, Walker 74°: 80/10/10; 2nd: altitude at 800 km, elevation angle 10°, Walker 45°: 80/10/10) that differ in inclination and orbit height. As one could observe, the combined constellation has a more uniform distribution. One may argue that the number of satellites in the combined one is doubled. However, by comparing Fig. 6b, c, one could notice that even with the same number of satellites, the combined solution still shows a very uniform result. Figure 6c clearly shows more visibility in the lower latitudes. While the visibility of 160 satellites of the single constellation solution mostly accumulates at the higher latitudes. As a summary, the advantage of the combined constellation is not only the increased number of satellites, but also thanks to the improving coverage at the lower latitudes.

From the results shown above it is clear that using a combined constellation leads to much more efficient and economical space segments. Thus, finding the most suitable combinations is the major task of constellation design. However, the work could be quite complex since there is a large number of parameters that need to be determined. For a single Walker constellation, these variables include among others, the altitude of the orbit, inclination, total number of satellites, number of the orbit planes, and spacing angle of the true anomaly between neighbouring planes. When combined two constellations, the variables that need to be determined are doubled, and the number of combination is much larger.

As an example, Fig. 7a shows a preliminary selection for the inclination of the combined constellation. Here, the standard deviation (STD) is used to characterize the smoothness of the number of visible satellites along latitude, in order to find a solution which is as uniform as possible. This figure gives the combination of two satellites that are at the altitude of 900 and 800 km respectively. Both satellites are

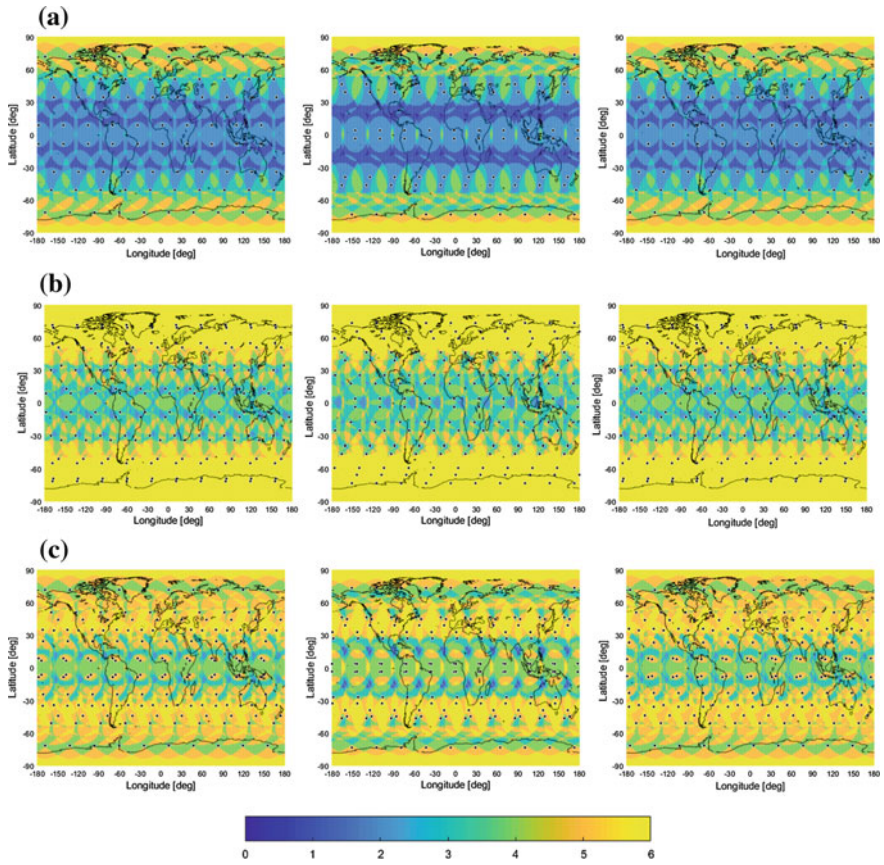


Fig. 6 Number of visible satellites for: **a** orbit height of 900 km, minimum elevation of 10°, Walker constellation of 74°: 80/10/10; **b** orbit height of 900 km, minimum elevation of 10°, Walker constellation of 74°: 160/10/10; **c** combined constellation of orbit height of 900 km, minimum elevation of 10°, Walker constellation of 74°: 80/10/10 and orbit height of 800 km, minimum elevation of 10°, Walker constellation of 45°: 80/10/10. Figures in every row from left to right represent the specific epoch that the argument of latitude of satellites moves 10°, 50° and 100° from the initial position, respectively

observed at a minimum elevation angle of 10°. In this case, their inclinations are varying from 0° to 90°. According to the figure it is obviously that the combination of a higher inclination (about 65°–85°) and a lower inclination (about 30°–60°) leads to a minimum STD. Of course, in order to cover the poles, the satellite with higher inclination should have an upper boundary of the coverage area of at least 90°. In this case, the higher inclination should be larger than 71.1°. The best case is the combination of 73° (900 km) and 38° (800 km).

Similarly, Fig. 7b shows the dependency on orbit height. The inclinations of the two satellites are set to 73° and 38°. Their orbit heights vary from 500 to 2000 km. The figure shows that the orbit heights of the two constellations should be more or

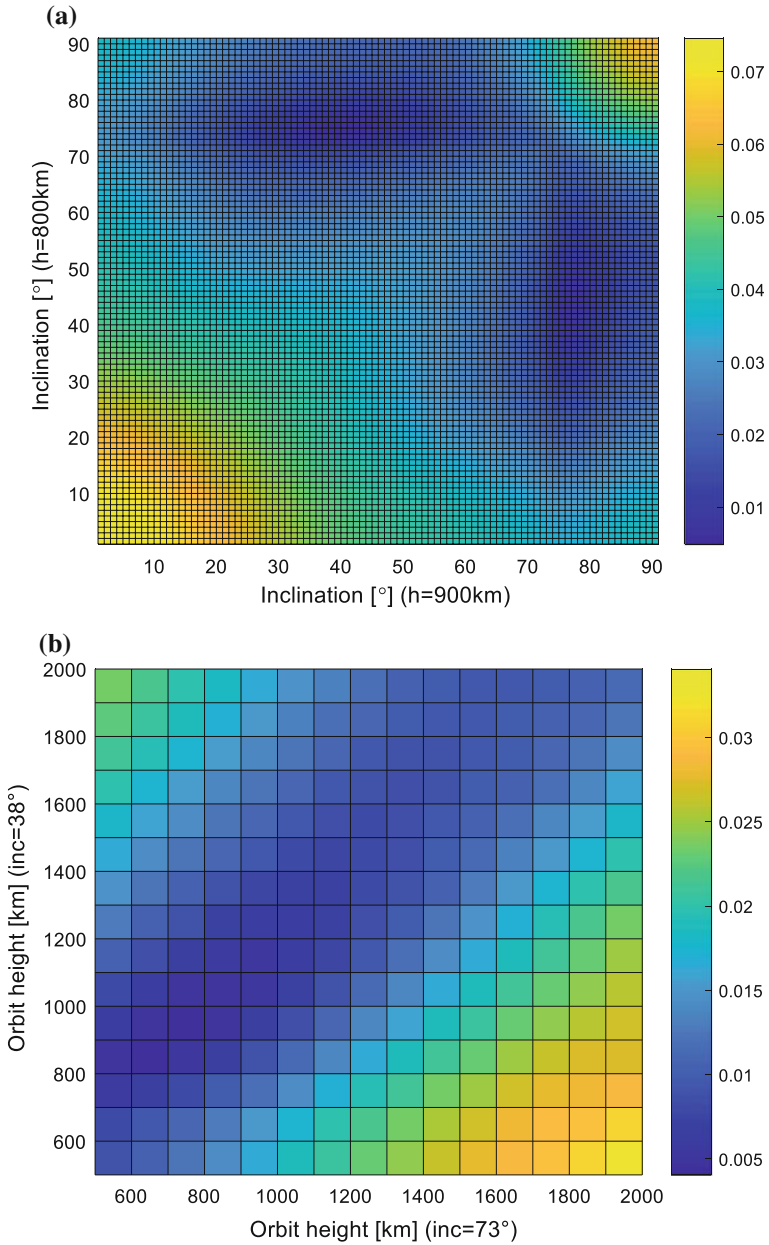


Fig. 7 Standard deviation of the number of visible satellites along latitude for different combination of: **a** inclination; **b** orbit height

less similar, in order to get smaller standard deviation. The minimum STD occurs at the combination of 900 km (73°) and 700 km (38°).

Apart from the combination of two constellations, we also studied the possibility of the combination of three constellations. The result shows that for selected orbit heights (700, 800 and 900 km) and elevation angle (10°), the best combination is 82° (700 km), 33° (800 km) and 63° (900 km). But from another perspective, there is no big difference of the standard deviation value between the combination of three constellations (STD = 0.0040) and the combination of two constellations shown in Fig. 7b (STD = 0.0041). Therefore, other factors like the costs should be taken into account when deciding the combination of two or three constellations.

This is, of course, not the optimum case for all possible combinations, since some parameters are fixed. In particular, the number of satellites in each constellation is assumed to be equal, which may however be varied too. As a result, more variables may be introduced to determine a more general ideal combination. In the next step, 5 parameters are selected as variables: the orbit heights and inclinations of the two constellations, as well as the ratio of the number of satellites in each constellation. The inclinations range from 0° to 90°; the orbit heights are selected between 700 and 900 km; and the ratio of the number of satellites is varying from 10:1 to 10:20. The minimum elevation angle is 10° for both constellations. The result shows that the combination of 700 km + 38° and 900 km + 73° with the same number of satellites in each constellation has the smallest standard deviation.

5 Conclusion and Outlook

This paper shows that the combination of several constellations with different inclinations leads to a distribution of the number of visible satellites along latitude which is more uniform than for a single constellation. We did some preliminary analysis to determine relevant parameters to get the most suitable combination. Parameters such as inclination, orbit height, as well as the ratio of the number of satellites in each constellation under certain conditions are well defined through numerical analysis.

In the future, in order to form the most suitable constellation, other relevant parameters like the number of orbit planes, the spacing angle of the true anomaly between neighbouring planes, and dilution of precision shall also be considered. In addition, it will be studied whether other choices than Walker constellations may further improve the global coverage.

References

1. Mosher D (2016) SpaceX just asked permission to launch 4425 satellites—more than orbit Earth today. Available via Business Insider. <http://www.businessinsider.de/spacex-internet-satellite-constellation-2016-11?r=US&IR=T>. Cited 17 Nov 2016
2. Xinhua (2017) State-owned aerospace giant plans to launch 156 minisatellites. Available via China Daily. http://www.chinadaily.com.cn/china/2017-03/31/content_28755292.htm. Cited 31 Mar 2017
3. OneWeb (2017) Official website of OneWeb. Available via OneWeb. <http://oneweb.world/>. Cited 13 Mar 2017
4. Boyle A (2016) 5G or not 5G? Boeing joins the battle over broadband satellite spectrum. Available via GeekWire. <https://www.geekwire.com/2016/boeing-battle-broadband-internet-satellite/>. Cited 23 Jun 2016
5. Dickie M (2016) LeoSat enterprises contracts first customer. Available via Business Wire. <http://www.businesswire.com/news/home/20160906005657/en/LeoSat-Enterprises-Contracts-Customer>. Cited 6 Sep 2016
6. Selding PB (2016) Telesat prepares shareholder payday, outlines 117-satellite constellation. Available via Space News. <http://spacenews.com/telesat-prepares-shareholder-payday-outlines-117-satellite-constellation/>. Cited 17 Nov 2016
7. Gershgorin D (2015) Samsung wants to blanket The Earth in satellite internet. Available via Popular Science. <https://www.popsoci.com/samsung-wants-launch-thousands-satellites-bring-everyone-earth-internet>. Cited 18 Aug 2015
8. Long L (2014) Satellite network robust QoS-aware routing. Springer, Berlin, Heidelberg, pp 21–40
9. Asvial M, Tafazolli R, Evans BG (2002) Non-GEO satellite constellation design with satellite diversity using genetic algorithm. In: 20th AIAA international communication satellite systems conference and exhibit. <https://doi.org/10.2514/6.2002-2018>
10. Beech TW, Cornara S, Mora MB, Lecohier GD (1999) A study of three satellite constellation design algorithms. In: 14th international symposium on space flight dynamics
11. Lang TJ (1996) Low Earth orbit satellite constellations for continuous coverage of the mid-latitudes. Collection of technical papers of AIAA/AAS astrodynamics conference, pp 595–607
12. ITU (1995) Methods of calculating low-orbit satellite visibility statistics. Rec. ITU-R SA. 1156
13. Katona Z, Donner A (2008) On mean visibility time of non-repeating satellite orbits. In: IEEE international workshop on satellite and space communications. <https://doi.org/10.1109/iwssc.2008.4656802>

A Required Navigation Performance Based Approach to Monitor the Accuracy and Integrity Performance of UAVs for Delivery Applications



Rui Sun, Yucheng Zhang, Bojia Ye and Washington Yotto Ochieng

Abstract As an emerging commercial application, deliveries using Unmanned Aerial Vehicles (UAVs) have the advantages of low cost and flexibility over traditional logistic operations, and are therefore being increasingly adopted by logistics companies. In order to ensure the success of delivery missions, effective traffic management measures, including UAV safety or integrity monitoring, are essential. Currently, however, there is a gap in research on the concept and framework for the evaluation of integrity performance for the delivery application. This paper addresses the issue by proposing a UAV Navigation Monitoring Requirement (UNMR) specification for the monitoring of the accuracy and integrity performance of UAVs for use in delivery applications, based on the concept of Required Navigation Performance (RNP). The method used to determine the position accuracy requirement for en-route and terminal UNMR operations is based on the related factors in the context of UAV-borne delivery. This is represented by the width of the route and radius of the docking or landing spot. Simulation and initial field tests show that the flight technical errors and path definition errors should be considered for the total system error calculation for the UAV delivery. Effective application of the proposed framework for integrity performance evaluation has shown to be of potential benefits for the operational management of UAV deliveries.

Keywords RNP · UAV delivery · Accuracy · Integrity monitoring

R. Sun (✉) · Y. Zhang · B. Ye
College of Civil Aviation, Nanjing University of Aeronautics
and Astronautics, Nanjing, China
e-mail: rui.sun@nuaa.edu.cn

W. Y. Ochieng
Center for Transport Studies, Imperial College London, London, UK

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_55

1 Introduction

Unmanned Aerial Vehicle (UAV) deliveries have the potential to significantly to reduce the cost and time involved in logistics operations, and are therefore being considered by an increasing number of logistics companies. Compared to the traditional delivery methods, such as truck delivery, UAVs are less expensive since the labour costs are replaced by automated operations. In addition, delivering with UAVs are generally due to their flexibility and minimal constraints by infrastructure such as specific roads [1]. Due to these advantages, a number of large organizations have shown interest in drone delivery. In 2013, Amazon announced Prime Air, a service that utilizes multirotor drones to deliver packages from Amazon to customers [2]. DHL also started its Parcelcopter project in 2013 transporting medicine to the island of Juist in the North Sea [3]. In 2014, Google initiated Project Wing, to produce drones that can deliver larger items than Prime Air and Parcelcopter [4]. The United Arab Emirates (UAE) announced a plan to use drones to distribute official government documents such as permits and ID cards [5], while the start-up Matternet collaborated with Swiss Post to test a lightweight package delivery quadcopter [6]. These growing UAV operations have the potential to bring great benefits for their stakeholders. However, deploying UAV delivery operations on a large scale will bring significant challenges for the UAV operators. Ensuring collision-free and seamless operations, along with precise and punctual landing at the terminus are all crucial to the success of delivery missions. Therefore, an effective framework for the real-time evaluation of the safety of each individual UAV delivery operation is essential for the maintenance of safety-of-flight levels [7].

Currently, there are no established rules, infrastructure or framework to enable and safely manage the widespread use of low-altitude airspace and UAV operations, such as UAV-borne deliveries. The US National Aeronautics and Space Administration (NASA), is actively exploring a prototype for an unmanned aircraft traffic management system to enable safe and efficient low-altitude civilian UAV operations. In order to achieve this envisioned system, NASA is doing research and testing in collaboration with the US Federal Aviation Authority (FAA) across four activity series, called “technology capability levels (TCL)”, with each level increasing in complexity. TCL4, in particular, will focus on package delivery with UAVs in high-density urban areas and will be announced after extensive testing in the near future [8]. Jun et al. [9], meanwhile, have developed a framework to correlate UAV operational risks with trajectory conformance requirements based on the International Civil Aviation Organization’s concept of Required Navigation Performance (RNP). RNP is a type of navigation specification that defines the accuracy required for the direction of lateral and flight paths [10], and aircraft conducting RNP operations must remain within that lateral accuracy for 95% of the trajectory [11]. Jun et al.’s work, however, only addressed the application of RNP concept for the operation of small UAVs. The evaluation of the accuracy and integrity performance of the system, which is critical for the safety of UAV delivery, is still to be addressed.

In order to bridge the current research gap in respect to the concept and framework for the evaluation of accuracy and integrity performance of UAV flights in the context of UAV-borne delivery operations, this paper proposes a UAV Navigation Monitoring Requirement (UNMR) specification for accuracy and integrity evaluation that is based on the concept of RNP. In particular, the position accuracy for the terminal and on-route UNMR operations is defined based on the related factors in the context of UAV-borne delivery. The size design of the route and terminal dock are further proposed based on the determined UNMR, up rounded from position accuracy. The value of UNMR operational level for the UAV-borne delivery has been investigated by simulation and initial field tests with the intention of offering significant benefits for the operation and management of UAV-borne deliveries. The contributions of this paper are as follows.

- Clarification of the Service level and navigation level requirements in the context of UAV-borne delivery (Sect. 2).
- Definition of the UNMR for UAV delivery integrity evaluation based on the concept of RNP and proposal of the horizontal position accuracy for use in en-route UNMR operation, and the horizontal and altitude position accuracies for use in terminal UNMR operation (Sect. 3.1).
- Proposal of the position accuracy calculation framework for the en-route operations of UNMR by considering related factors as well as a method for size of the route design and docking area design of the UAV borne delivery (Sect. 3.2).
- Demonstration of feasibility by simulation and field testing (Sects. 4 and 5).

The rest of the paper is organized as follows. The concept of using UAVs for delivery applications is presented in Sect. 2. The UAV Navigation Monitoring Requirements (UNMR) for UAV-borne deliveries is presented in Sect. 3. The simulation and initial field tests for the proposed UNMR based framework for UAV-borne deliveries is discussed in Sects. 4 and 5, respectively. The paper is concluded in Sect. 6.

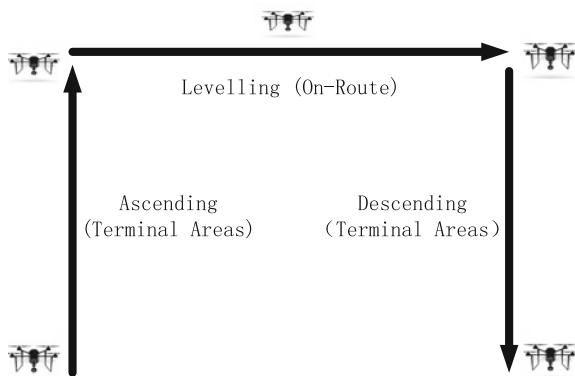
2 The Concept of UAV Delivery System and Its Requirement

Delivery involves using a delivery drone to transport goods packages. As no universal definition exists, a UAV delivery system is defined in this paper as: a system of people, procedures and interacting telematic elements forming a coherent unit that aims to deliver parcels by use of unmanned aerial systems. Depending on the delivery objective, such a system may be designed in various ways, based on the concept of systems engineering [12]. Since little research has been done on defining a conceptual framework for UAV delivery systems, the entities and related requirements are clarified here as follows.

The mode of operation for a UAV is defined in this paper flying from a regional distribution center to a small local distribution point. This type of UAV operation depends on considers battery capability, and size and weight of the parcels. The regional to local operation is the most simple and basic operation mode for the UAV delivery [3–5]. The operator manages in real-time the UAV flight and tracks the status of the parcel status based on control platform. The three phases of flight are defined as ascending, levelling and descending, where ascending and descending are further defined as operations in terminal areas and levelling is defined as an en-route operation. In this paper, the UAV is assumed to follow a process of straight ascent, level flight and then straight descent, as in Fig. 1.

A UAV delivery system may have two primary entities, defined as the system operator and the user. One common desire for both entities is to monitor the UAV to deliver the goods safely to the desired location on time. The typical requirement from the user and operator could be defined as the service level requirements. These requirements mainly include the delivery areas, delivery routes, equipment configuration, timeliness, lack of damage and cost requirement. The levels of these requirements will vary according to different customers’ needs. The navigation systems, which are the underpinning technologies to support this user level requirement, also have their own navigation performance requirements, including accuracy, integrity, continuity and availability for the different stages of the delivery operation [13]. Accuracy refers to the statistical distribution of position errors, velocity errors or speed errors. Integrity is defined as the ability of a system to provide timely and valid warnings if the position errors exceed a certain alarm limit. Alert limit, integrity risk and time to alert are three parameters of integrity. In particular, integrity risk (loss of integrity) is defined as the probability that a user will experience a position error larger than the alert limit (AL) without an alert being raised within the specified time-to-alert (TTA) at any instant in time and at any location in the coverage area [14]. Continuity risk is the probability that a navigation service available at the start of a specific operation is interrupted during the period of operation. The availability is the percentage of time that the services of

Fig. 1 The flight profile of the defined UAV-borne delivery operations



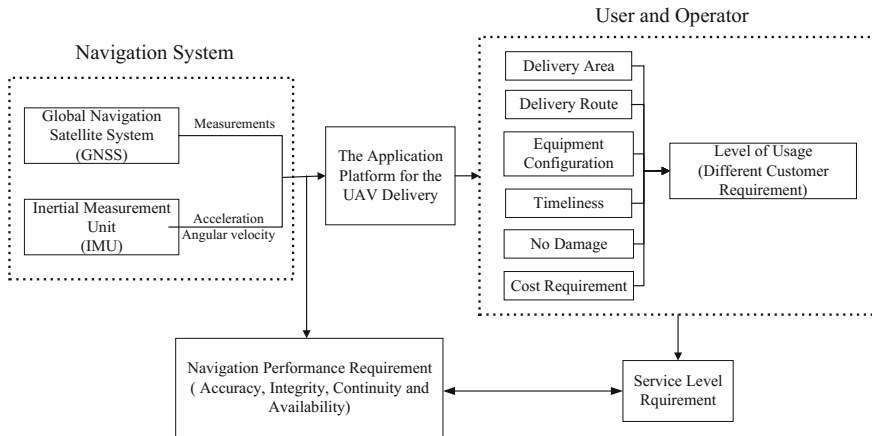


Fig. 2 The navigation and service level requirements for UAV-borne delivery

the system are usable by the navigator. Figure 2 describes the link between the service and navigation level requirements for UAV-borne delivery operations.

As mentioned in Sect. 1, however, the quantitative and testable performance metrics for these requirements are still under investigation, and the required navigation performance metrics are, therefore, still to be derived and standardized from the service requirements. The requirement of accuracy, which is the fundamental requirement for the UAV delivery, is the essential parameter to be determined. Afterwards, the integrity (i.e. the trust that can be placed in the information supplied by the navigation system), which is able to provide warnings to the user when of the position error exceeds the alarm limit, is also a crucial element for monitoring UAV safety. The evaluation of the accuracy and integrity performance is therefore critical. In the following section, an accuracy and integrity performance evaluation framework is presented.

3 UAV Navigation Monitoring Requirement for UAV-Borne Deliveries

3.1 UNMR Framework Definition

The proposed UNMR framework is for the integrity performance monitoring and alerting capability of the in-flight UAV-borne delivery process based on the concept of RNP [11]. All aircraft conducting UNMR operations must have a position error of less than the specified distance for 95% of the trajectory during the delivery borne operations. The position accuracy are defined variously based on the en-route and terminal operations. In the concept of RNP, the position accuracy is the Total

System Error, including Navigation System Error (NSE), Flight Technical Error (FTE), and Path Definition Error (PDE) [10]. Therefore, for the UNMR, different factors related to the TSE are considered for the en-route and terminal operations.

In the context of UAV-borne delivery, a fleet of UAV operations is likely to be needed together, with relatively close intervals to the same destination. The en-route operation therefore, needs to be designed to allow UAVs to operate at the same altitude but with sufficient horizontal space between each of them to avoid lateral and longitudinal collisions with the other aircraft or obstacles during the delivery. Accordingly, the horizontal position accuracy is the central parameter for use in en-route UNMR. For terminal area operations, especially for the descending phase, the horizontal position error should be within the specific designed dock. Furthermore, the vertical position error need to be measured accurately to ensure safe landing. Therefore, the horizontal and vertical errors are both vital for use in terminal areas UNMR operation. In this paper, only the horizontal accuracy is considered for terminal areas UNMR operation in order to simplify the operation requirement issue. The design of one half of the total route width for en-route and the radius for the dock are highly related to the UNMR. The details for calculation of the position accuracy and design the related en-route half width and dock radius are in Sect. 3.2.

The defined UNMR level, refers to the position accuracy in meters, which is expected to be achieved 95% of the operation time. The value of UNMR level will be used for the application of en-route width and dock radius design of the UAV borne delivery. The level is defined as UNMR N , where N is integer value to be set from 1 to N with the interval of 1. For example, UNMR 6 means that for at least 95% the flight time, the UAV, while following the predefined flight route, should have a position accuracy of within 6 m, with this accuracy horizontally during en-route operations or during operations in terminal areas. The concept of UAV Actual Navigation Performance (UANP) is also introduced to evaluate the UAV real time flight position. UANP creates a link between the real flight location and the flight location estimated by the on-board navigation system. The relationship between UNMR and UANP are further discussed in Sect. 3.3.

3.2 Accuracy Calculation for the Size of Route and Dock Design

The width of the UAV route design is highly related to the en-route horizontal position accuracy. The en-route horizontal position accuracy is determined by several factors. The Navigation System Error (NSE) are one critical factor from the measurements of on-board navigation sensors. UAV Flight Technical Error (UFTE), which reflects the recovery ability of the UAV once the irregular course happens, including the delays caused by the communication link, the deviations caused by the wind, air disturbances caused by high speed rotation of the UAV

rotors and jamming etc. In order to ensure the safety operation of the fleet UAV, interval size should also be considered for the following or the leading UAVs. However, the value of the interval size is much less than NSE and UFTE, therefore interval size is considered negligible. The horizontal accuracy for the en-route operation $L_{en-route}$ could be calculated based on the error propagation rule as follows.

$$L_{en-route} = \sqrt{NSE^2 + UFTE^2} \tag{1}$$

The UNMR for the en-route operation, will be further defined based on the round up value of $L_{en-route}$. Assume AW is one half of the total designed route width, A coefficient e is proposed to multiply the $L_{en-route}$ for the delivery safety. A buffer value δ is also proposed for the calculation. Based on the ICAO DOC8168 [15], the half width for the designed path could be expressed as follows.

$$AW = e * UNMR_{en-route} + \delta \tag{2}$$

The dock design is highly related to the horizontal positioning accuracy in the terminal for the UAV delivery. According to ICAO DOC9613, the longitudinal accuracy is mainly determined by Navigation System Errors (NSE) [10]. In addition, the UFTE will also contribute errors for the UAV horizontal position during landing. We consider the parcel as a cuboid with length, width and height of a, b, c . Where, $a \geq b \geq c$. The catercorner of the horizontal projection for the parcel is $\sqrt{(a^2 + b^2)}$. The successful delivery is recognized as the whole of the parcel within of the circle. Therefore, the horizontal accuracy for the terminal operation $L_{Terminal}$ could therefore be defined as.

$$L_{Terminal} = \sqrt{NSE^2 + UFTE^2} + \frac{\sqrt{(a^2 + b^2)}}{2} \tag{3}$$

Similarly, the UNMR for the terminal operation, will also be further defined based on the round up value of $L_{Terminal}$. Assume the R is the radius required for the dock in the terminal. A coefficient w is proposed to multiply the $L_{Terminal}$ for the delivery safety. A buffer value ε is also proposed for the calculation. Based on the ICAO DOC8168 [15], the radius for the designed dock could be expressed as follows.

$$R = w * UNMR_{Terminal} + \varepsilon \tag{4}$$

3.3 The Relationship of UNMR and UANP

The Uncertainty Radius model is proposed to define that, with a 95% probability, the true position of the UAV is located in the centre of an error circle with a radius set at the UANP value. The relationship between the UNMR and UANP is depicted in Fig. 3.

The UNMR and UANP will be compared based on the defined error circle model in order to evaluate the integrity of the UAV. If the UANP is greater than the specific level of the UNMR, the alarm will be issued from the control centre, directing the operator to take effective measures to avoid potential threats of flight safety. Therefore, the UANP could be recognized as protection level and UNMR is the alarm limit. Table 1 shows the relationship between UANP, UNMR and the requirement of integrity.

The derivation of 95% Uncertainty Radius from 1σ uncertainty ellipse could be found in [15]. It is indicated that the radius of the error circle R_p could be calculated in (5).

$$R_p = k\sigma_x = 2.4477\sigma_x \tag{5}$$

The uncertainty radius, R_p , which is related to the value set for the UNMR operational level could be linked to the σ_x , which is the major axis of the standard 1σ uncertainty ellipse and is related to UANP.

Overall, in order to meet the specific integrity monitoring requirement during the flight, 95% circular for the estimated position of the UAV should be kept within the level specified for the UNMR, which means $UANP \leq UNMR$. Once

Fig. 3 The relationship between the UNMR and UANP

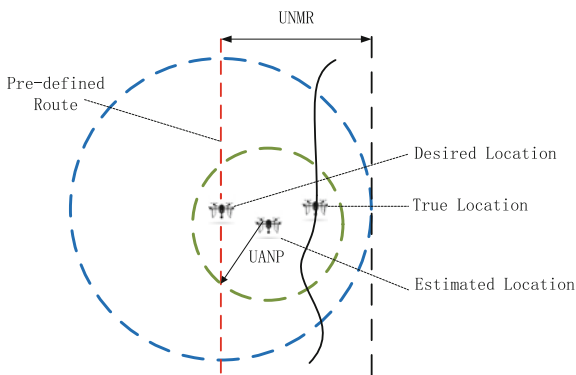


Table 1 UANP, UNMR and the relationship between them to meet the integrity requirement

Scenarios	Meet the requirement for integrity
$UANP \leq UNMR$	Yes
$UANP > UNMR$	No

UANP > UNMR, it is indicated that the navigation performance has not met the requirement. The operator should then issue a warning to the UAV pilot or to the function. In the next section, a set of extensive simulations have been carried out to generate data across different flight stages to allow the investigation of the UNMR operational level for the suitable size of the route and dock design in the context of defined UAV delivery operations.

4 Simulation

The simulation is to investigate the UNMR operational level for the size of the route and dock design in the context of UAV delivery operations. The procedure of the simulation will be as follows. The scenarios were first designed for the UAV delivery in different continents. Afterwards, the horizontal position accuracy for on-route and terminal will be calculated to determine the UNMR level. Accordingly, the size of the route and dock suggested are further calculated by the corresponding level of UNMR. The comparison of UANP and UNMR are then discussed to justify the feasible of the proposed framework applied.

From the study of real statistics from the real UAV delivery test sites all over the world, four typical areas were selected in different continents [2-4]. These are European-Germany-Rhein (6.5524°E, 51.3546°N), America-the U.S.- Menlo park (-122.1817°W, 37.4533°N), Oceania-Australia-Queensland (142.7028°E, -20.9163°S) and Asia-China-Huizhou (114.4109°E, 23.1154°N). The simulations

Table 2 Simulated scenarios

Location	Time (h:min)	Flight direction	Cruise altitude (m)
Rhein	8:00-8:05	W	20
	8:15-8:20	S	35
	8:30-8:35	E	50
	8:45-8:50	N	65
Menlo park	9:00-9:05	W	20
	9:15-9:20	S	35
	9:30-9:35	E	50
	9:45-9:50	N	65
Queensland	10:00-10:05	W	20
	10:15-10:20	S	35
	10:30-10:35	E	50
	10:45-10:50	N	65
Huizhou	11:00-11:05	W	20
	11:15-11:20	S	35
	11:30-11:35	E	50
	11:45-11:50	N	65

started from these four locations and the flight direction were variously from the west (W), south (S), east (E) and north (N) with 10 Hz sample rate. The cruise altitude was set as 20, 35, 50 and 65 m. The flight duration for each time of flight was 5 min and 16 sets of simulations are carried out. Table 2 has shown the designed scenarios.

The simulation data for GNSS outputs, (i.e. Easting and Northing coordinates and velocity) and the Inertial Measurement Unit (IMU) output (i.e. acceleration and yaw rate) are generated. The reference flight trajectory was first generate for the three stages. In the ascent stage, the UAV climbed with changing velocities, decelerating as it approached the desired altitude. The velocity was 0 when it reached the desired height, and it hovered for 5 s prior to the next stage. In the level flight stage, the UAV first accelerated to reach a speed of 8 m/s before randomly accelerating or decelerating between -0.3 m/s^2 and 0.3 m/s^2 . At the end of this stage the UAV decelerated until it was overhead the destination before hovering for 5 s in order to be ready for the descent. The descent stage can be portrayed as the inverse of the ascent stage. The GNSS measurements are further generated by the Spirent GSS8000 simulator. The IMU measurements are generated by adding the zero drift error, the first order Markov process noise and Gaussian White noise on the reference data.

The UANP for the climbing, leveling and ascending stages is calculated from the CKF based GNSS/IMU fusion algorithm presented in our previous research [16]. Based on the experiment results of 16 sets of data for 4 directions, it is indicated that the biggest error and mean value is in the scenario of Queensland. Figure 4 has shown an example of the relationship between the UNMR and UANP for the scenario of Queensland. The values for calculating the horizontal position errors on-route are as follows. NSE is the accuracy of the CKF estimated results for GNSS/IMU fusion, which is 5.48 m and UFTE is set as 0.3 m according to the statistical results of the experienced data. The experienced data is collected from the field test of the TOPXGON T1-A flight controller, which is popular used for the delivery application. The UFTE is estimated based on the difference between the UAV total system error and navigation system error data. We used bootstrap algorithm to estimate the statistic characteristics of the UFTE with the mean error of 0.3 m in the 95% confidence level. Therefore, we have calculated $L_{\text{En-route}} = 5.56 \text{ m}$ based on Eq. (1). The level of UNMR for the en-route is therefore determined as UNMR6. e is set as 1.5 and δ is set as 0, the suggested half width of the route is determined as $AW = 9 \text{ m}$ based on Eq. (2). Similarly, for the terminal areas operation, the value for NSE is 4.6 m and the 1/2 length of the catercorner projection of the parcel is set as 0.2m. The value of UFTE is set as 0.24 m from previous study. Therefore, the $L_{\text{Terminal}} = 4.81\text{m}$ is calculated based on Eq. (3). The level of UNMR for the terminal operation is therefore determined as UNMR5. By setting the $w = 1.5$, $\varepsilon = 0$, the suggested dock radius $R = 7.5 \text{ m}$ could be calculated based on Eq. (4).

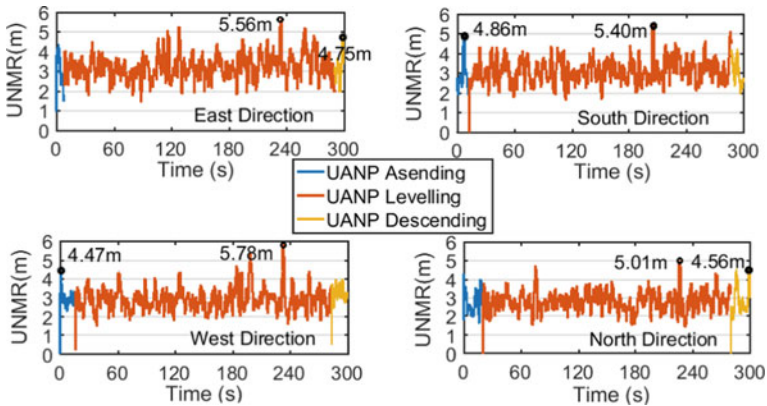


Fig. 4 An example for the UNMR determination based on UANP (the scenario of Queensland)

The alarm limit is the value of UNMR level, which is 6 m for the en-route and 5 m for the terminal operations. The integrity risk is the product of the probability of sensor/system failure and the probability of misdetection.

$$P_{Integrity Risk} = P_{sensor/system Failure} * P_{Misdetection} \tag{6}$$

Our research has shown that for the UAV delivery application, the misdetection rate is 0.08 from the fault detection algorithm used in [17]. The system failure includes the sensor failure and the damage rate of the delivery. If the sensor failure probability is assumed to be 10^{-4} per hour is assumed as is the case for GPS satellites [18], the damage rate is assumed to be 6×10^{-3} per hour due to the parcel damage rate from Shunfeng express, the required IR is of the order of 10^{-5} per hour. The accuracy and time-to-alert requirements include the nominal performance of a fault-free receiver. From the Signal-in-space performance requirements for the TTA en-route and terminal is 15 s [19].

5 Initial Field Test

The initial field test is used to validate the feasible of the proposed framework for the UAV borne delivery. The fleet test is carried out in Linyi City, Shandong province, China. Figure 5 has shown the UAV used in the experiment from Topgxon. The time duration for the test is about 9 min. The UAV positioning obtained from on-board Real Time Kinematic (RTK) GPS are post-processed with backward and forward smooth as the reference. The single point positioning and MEMS integrated results are used as the UAV positioning results.



Fig. 5 The fleet UAV used for test

The flight profile for the leading UAV and the integrity performance has been shown in Fig. 6. Based on the calculated UANP value, we could validate that UNMR5 is the proper level required for the terminal operation during the field test. For the en-route operation, if we use UNMR6 as the requirement, we will find out that the $UANP > UNMR$ around the time epoch 11:53. That will result in a warning sent from the control center to report the integrity could not be met. It is also indicated that unlike the calculation of the total system error for civil aviation operation, which could ignore the flight technical error and path definition error. The UFTE and path definition error could not be ignored for the UAV on-borne operation as the wind condition for the real test environment is complicated and the commercial maps for the path plan are containing errors. In general, it is validated that the proposed framework could be used to evaluate the accuracy and integrity for the UAV delivery.

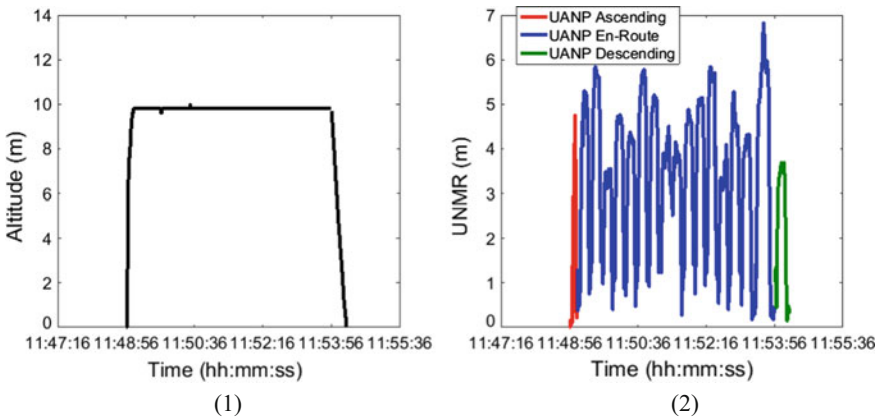


Fig. 6 The flight profile for the UAVs and the integrity performance evaluation

6 Conclusions

This paper investigates an application of the Required Navigation Performance (RNP) concept for UAV delivery by proposing an UNMR framework for the accuracy and integrity performance evaluation. The simulation and initial field test results have shown the proposed framework for integrity performance evaluation could be used for the UAV delivery and therefore will be benefit for the air traffic management. In further, the details for some key factors, such as the Flight Technical Error (UFTE) need to be further analyzed based on different flight models. To determine the value of the UFTE, the control ability for the UAV controller under various wind effects need to be analysed in the simulation and field test. In addition, the service level requirements for the UAV delivery will be investigated based on different usage levels and then derive more details accuracy, integrity, continuity and availability requirement, which could support better UAV delivery operation.

References

1. Dorling K, Heinrichs J, Messier GG et al (2017) Vehicle routing problems for drone delivery. *IEEE Trans Syst Man Cybern Syst* 47(99):1–16. <https://doi.org/10.1109/TSMC.2016.2582745>
2. Amazon.com, Inc. or its affiliates (2016) Amazon Prime Air. Amazon.com, Inc. or its affiliates. Available via DIALOG. www.amazon.com/primeair. Cited 2 Apr 2016
3. Bonn (2014) DHL Parcelcopter launches initial operations for research purposes. DHL International GmbH. Available via DIALOG. http://www.dhl.com/en/press/releases/releases_2014/group/dhl_parcelcopter_launches_initial_operations_for_research_purposes.html. Cited 24 Sep 2014
4. Stewart J (2014) Google tests drone deliveries in project wing trials. BBC, London, U.K. Available via DIALOG. <http://www.bbc.com/news/technology-28964260>. Cited 28 Aug 2014
5. Aljazeera (2014) UAE to use drones for citizen services. Al Jazeera Media Network. Available via DIALOG. <http://www.aljazeera.com/news/middleeast/2014/02/uae-use-drones-government-services-20142121717319272.html>. Cited 13 Feb 2014
6. Dillow C (2015) Meet Matternet, the Drone delivery startup thats actually delivering. *Fortune*. New York, NY, USA. Available via DIALOG. <http://fortune.com/2015/05/01/matternet-drone-delivery>. Cited 1 May 2015
7. ITU (2009) Characteristics of unmanned aircraft systems and spectrum requirements to support their safe operation in nonsegregated airspace. Mobile, Radio Determination, Amateur Related Satellite Services. Geneva, Switzerland Available via DIALOG. <http://www.itu.int/pub/R-REP-M.2171/fr>. Cited 11 Nov 2011
8. Santos VDL, Rios J (2015) NASA UTM: Home. National Aeronautics and Space Administration. Available via DIALOG. <http://utm.arc.nasa.gov/index.shtml>. Cited 2 Jan 2018
9. Jung J, D'Souza SN, Johnson MA et al (2016) Applying required navigation performance concept for traffic management of small unmanned aircraft systems. In: 30th Congress of the international council of the aeronautical sciences (ICAS 2016)

10. ICAO (2008) Performance based navigation (PBN) manual. Doc. 9613. Montréal, Quebec, Canada
11. FAA (2015) Aeronautical information manual
12. Faulconbridge R, Ryan MJ, Faulconbridge R et al (2005) Engineering a system: managing complex technical projects. Argos Press, Canberra. ISBN 1-921138-00-9
13. ICAO (1995) Report of the special communications/operations divisional meeting, Doc. 9650
14. Ochieng WY, Sauer K, Walsh D et al (2003) Gps integrity and potential impact on aviation safety. *J Navig* 56(1):51–65
15. ICAO (2014) Aircraft operations volume II construction of visual and instrument flight procedures. Doc. 8168
16. Sun R, Xie F, Xue D et al (2017) A novel rear-end collision detection algorithm based on GNSS fusion and ANFIS. *J Adv Transp* 2017:1–10. <https://doi.org/10.1155/2017/9620831>
17. Sun R, Cheng Q, Wang G et al (2017) A novel online data-driven algorithm for detecting UAV navigation sensor faults. *Sensors* 17(10):2243. <https://doi.org/10.3390/s17102243>
18. U.S. Department of Defense (2008) Global positioning system standard positioning service performance standard
19. ICAO (2004) Annex 10 vol V AMD 79 Spanish-2004

SBAS Integrity Verification Based on the Extreme Value Theory



Yan Zhang, Zengjun Liu, Caihua Li, Xiaomei Tang and Gang Ou

Abstract In many areas such as safety of life application, the integrity of the navigation system is a very important indicator to indicate the trustworthiness of the system services. According to different scenarios, system integrity requirements are also different, but are usually in the order of 10^{-7} or even less. It is impossible to evaluate the performance by the traditional statistics which using the frequency of occurrence to simulate the probability. The Extreme Value Theory (EVT), which enables extrapolation of the observed error distribution's tail into the misleading information domain, regardless of the underlying error distribution. In this paper, EVT is applied to the SBAS integrity evaluation field, which can effectively solve the problem of insufficient tail samples and realize a reasonable inference of extreme abnormal events. The paper first briefly introduces the basic concept of SBAS integrity and EVT, and then uses the mixture extreme model to solve the problem of threshold choosing. Finally, the paper uses EGNOS message data provided by EMS, and wtza station observation data to evaluate its integrity performance. The result shows that the EGNOS is compliant with LPV requirements in the vicinity of wtza station.

Keywords SBAS EVT · Integrity verification

1 Introduction

In many areas, such as safety of life application, the integrity of the navigation system is a very important indicator. It refers to the ability to provide timely and effective warning to the users when the system cannot be used, and reveals the extent of the system to provide reliable service. At the very beginning, the design of GPS system did not consider the requirements of integrity, although the system can

Y. Zhang · Z. Liu (✉) · C. Li · X. Tang · G. Ou
College of Electronic Science, National University of Defense Technology,
Changsha, Hunan, People's Republic of China
e-mail: zjunliu@126.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_56

monitor the fault of the satellite and send it to the user through the navigation message, the time to alarm is usually an hour scale. With the continuous expansion of its application, especially in the aerospace and other occasions with high security requirements, system integrity is playing an increasingly important role.

With different scenarios, system integrity requirements are also different, but are usually in the order of 10^{-7} or even less. As 150 s is used as the unit in the definition of integrity, it takes about 24 years to observe an abnormal event. Further, with the consideration of the statistical requirement, single or several observations may not sufficient to describe the probability characteristics. Therefore, it is unrealistic to use traditional methods to calculate the integrity risk of SBAS. It is necessary to extrapolate to the extreme field with limited observations in finite time.

The Extreme Value Theory (EVT) [1], which enables extrapolation of the observed error into tail characteristics, and regardless of the underlying error distribution. It has been widely used in areas which sensitive to extreme exceptions, such as disaster prediction, financial and insurance actuarial. In this paper, EVT is applied to the SBAS integrity evaluation field, which can overcome the drawback of traditional statistical method that depends highly on the number of samples and make a reasonable extrapolation. At the same time, the mixture extreme model is used to estimate the threshold to avoid the adverse effects of the subjective threshold selection method. By using the designed SBAS software receiver, this paper makes a thorough data analysis with the actually measured data from EMS and IGS wtza station, and proves that the EGNOS integrity meets the system design requirements.

2 SBAS Integrity

2.1 The Concept of SBAS Integrity

SBAS integrity is usually described by three indicators, namely, the Alarm Limit (AL), the Probability of Integrity Risk (P_{req}) and the Time To Alarm (TTA). SBAS requires that the probability of Hazardous Misleading Information (HMI) is less than the probability of integrity risk specified in a certain scenario, which can be expressed as follows,

$$P_{HMI} = P_r\{xPE > xAL | xPL < xAL\} \leq P_{req} \quad (1)$$

Among them, x stands for H/V, representing horizontal/vertical, xPE is stands for positioning error, xPL is stands for the protection level, and xAL is stands for the alarm limit.

The PL is calculated by received SBAS integrity parameters and local error characteristics of the receiver. Its calculation formula [2] is as follows. Here and later in this paper, we take VPL as an example.

$$VPL = K \sqrt{\sum_{i=1}^n S_{v,i}^2 \sigma_i^2} \tag{2}$$

$S_{v,i}$ is a spatial projection matrix of pseudo range domain to position domain. σ_i is the pseudo range error of the i satellite. K is the coefficient corresponding to the integrity requirements of a specified scenario. n is the number of visible satellite. The satellite pseudo range error is further composed of the ephemeris error, the clock error, the ionosphere delay error, the troposphere delay error and the error from the receiver itself. That is:

$$\sigma_i^2 = \sigma_{i,udre}^2 + \sigma_{i,uire}^2 + \sigma_{i,tropo}^2 + \sigma_{i,rev}^2 \tag{3}$$

$\sigma_{i,udre}$ and $\sigma_{i,uire}$ are calculated by the broadcast SBAS integrity parameters. $\sigma_{i,tropo}$ is calculated according to the tropospheric error model given in the MOPS.

2.2 SBAS Integrity Estimation

The key to SBAS integrity analysis is to estimate the probability of HMI events:

$$P_{HMI} = \int_0^{VAL} \int_{VAL}^{+\infty} f(VPE, VPL) \cdot dVPE \cdot dVPL \tag{4}$$

$f(VPE, VPL)$ represents the two-dimensional density function composed of random variables VPE and VPL. Usually, when analysing system integrity, MI (Misleading Information) is a more conservative indicator which is used to indicate the probability that VPE is greater than VPL:

$$\begin{aligned} P_{MI} &= P_r\{VPE > VPL\} = \int_0^{+\infty} \int_{VAL}^{+\infty} p(VPE, VPL) \cdot dVPE \cdot dVPL \\ &= \int_0^{+\infty} \int_y^{+\infty} p(x, y) \cdot dx \cdot dy \end{aligned} \tag{5}$$

On the one hand, using MI as a more conservative estimate of integrity risk does not have an impact on the correctness of the evaluation results. On the other hand, the MI evaluation results are independent of the alarm threshold, which means that

the results are not dependent on specific application scenarios, and become more applicable. For simplicity, we uses X to represent VPE, and Y represents VPL in formula (4).

Since the two-dimensional probability density is not known accurately and it is difficult to estimate, the concept of SI (Safety Index) is introduced to convert it into the probability calculation of a one-dimensional random variables.

$$SI = VPE/VPL \tag{6}$$

$$P_{MI} = \int_0^{+\infty} \int_y^{+\infty} p(x, y) \cdot dx \cdot dy = \int_1^{+\infty} \int_0^{+\infty} f(z, y)|y| \cdot dy \cdot dz = \int_0^{+\infty} f_{\frac{z}{y}}(z) \cdot dz \tag{7}$$

That is, the probability of MI occurrence can be obtained by calculating P ($SI > 1$). In the text below, the system integrity risk evaluation is based on the SI .

3 EVT and EVT Application in Integrity Verification

EVT describes the extreme characters of random variables, and it takes samples which seriously deviate from the mean as the research object. There are many different methods to characterize the extreme behaviour of stochastic processes, among which the most important two models are Block Maximum Method (BMM) and Peak Over Threshold (POT).

BMM Model: Suppose that $\{X_1, X_2, \dots, X_n\}$ is a sequence of independent and identically distributed random variables, and if there is a sequence of normalized constants, $\{a_n > 0\}$ and $\{b_n\}$, the following EVT formula can be established:

$$P_r\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow H(x) \tag{9}$$

Then, the nondegenerate distribution $H(x)$ must have the following form (Generalized Extreme Value, GEV):

$$H(x, \mu, \sigma, \zeta) = \exp\left\{-\left(1 + \zeta \frac{x - \mu}{\sigma}\right)^{-1/\zeta}\right\}, \quad \left(1 + \zeta \frac{x - \mu}{\sigma} > 0\right) \tag{10}$$

μ , σ and ζ are position parameter, scale parameter and shape parameter respectively. The shape parameter ζ plays a major role in the distribution characteristics, which determines the tail characteristics of the random variables. When BMM is used, it is necessary to divide the samples into several non-overlapping intervals, and then select the maximum (minimum) values of each

interval to fit the GEV distribution, and estimate the parameters μ , σ and ξ . The problem of BMM is that it is difficult to make full use of the extreme value data whose number is originally very scarce. In order to make up for these deficiencies, POT model is developed. This method uses all extreme values which exceed a critical threshold to model the extreme character of the variable.

POT model: Suppose that $\{X_1, X_2, \dots, X_n\}$ is a sequence of independent identically distributed random variables, whose distribution function is F . u is chosen as the threshold, and the excess of the threshold is defined as $Y_i = X_i - u$. Therefore, the distribution function of Y_i can be expressed as:

$$F_u(y) = P_r\{Y < y | X > u\} = P_r\{X - u < y | X > u\} = \frac{F(X) - F(u)}{1 - F(u)} \quad (11)$$

When u is large enough, $F_u(y)$ (the function of the threshold excess) follows the Generalized Pareto Distribution (GPD):

$$G(y; \xi, \sigma_u) = 1 - \left(1 + \xi \frac{y}{\sigma_u}\right)^{-1/\xi} \quad (12)$$

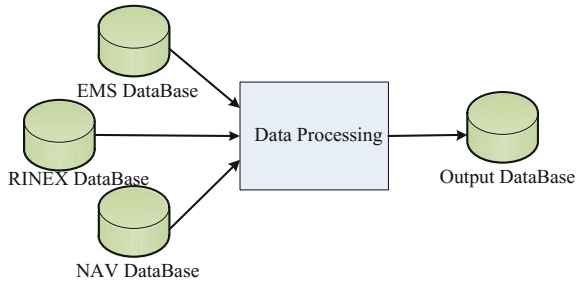
At present, many researchers have applied extreme value theory to SBAS integrity evaluation, of which the most important papers come from Holland NLR (National Aerospace Laboratory) and Delft University. The two gave the basic steps of using EVT to assess the SBAS integrity, and developed the software GIMAT (GNSS Integrity Monitoring and Analysis Tool) [3–5] for assessing the performance of EGNOS. Although above papers have proved that the effectiveness of this method to a certain extent, there are still many problems to be solved, including: selection of threshold in function fitting and so on. Especially when the model fitting effect is poor, we don't know how to further judge the reason of the anomaly, whether it is due to the integrity fault or the sample does not conform to the EVT hypothesis.

This paper chooses SI as SBAS integrity indicator and uses POT model to fit the extreme value sample. A mixture extreme model is used to determine the threshold. At the same time, the real data from German wtza station is used to evaluate the integrity of EGNOS, and the abnormal data in the test results is also analysed in depth.

4 Measurement and Verification of EGNOS Integrity

4.1 Data Processing

GIMAT has not been released online. In this paper, the differential and integrity information processing function is added in the open source software receiver,

Fig. 1 Processing flow chart

goGPS [6]. The input of the software is the data from EMS, RINEX and NAV database. The output of the software includes the position results at the observation time and the VPL calculation results (Fig. 1).

EMS [7] (EGNOS Message Server) is an Internet based offline SBAS data broadcasting service which provided by EGNOS. Both RINEX and NAV data files come from the IGS website.

The core part of the Data Processing is as follows (read and process files in database by cycle of days):

Step1: Read the ephemeris data file from NAV database and store the corresponding ephemeris information of each record.

Step2: Read the data files from RINEX database and store the corresponding observation information of each record (including the receiving time, pseudorange of L1/L2 frequency point, carrier phase, Doppler and carrier to noise ratio).

Step3: Read EGNOS data file from EMS database, obtain the EGNOS message type, the original message and the broadcast time corresponding to each message, and analyze the EGNOS message to get PRN_mask, FC, UDRE, LTC, GIVE and IGP parameters, etc. The specific process is shown in the following picture (Fig. 2).

Step4: Data matching includes two aspects: find the corresponding FC parameters for UDREs in MT6 and match the differential and integrity data obtained from Step3 with the receiver observations read from Step2.

Step5: The receiver position is calculated by the least square algorithm at each observation time. Then, by comparing the calculated position and the known real position of the receiver, the Vertical Positioning Error (VPE) can be obtained. Vertical Protection Level (VPL) is calculated by using the method specified in MOPS.

Step6: Using Rstudio[®] to analyze the statistical characteristics of them, and use EVT to infer integrity risk. There are some conventions in the implementation, which mainly involves the principle of using augmentation information.

1. When the number of satellites which meet threshold requirements of the satellite elevation or carrier noise ratio is less than 4, the record is not used for integrity analysis.

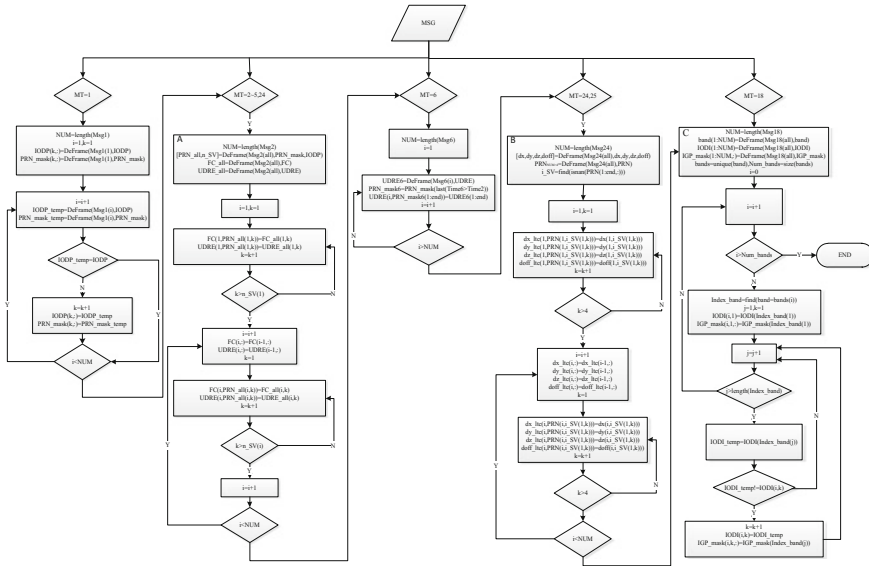


Fig. 2 Processing of SBAS differential and integrity parameter

2. When the number of available UDRE (UDRE_i < 14) is less than 4, the record is not used for integrity analysis.
3. When the number of available GIVE (GIVE < 15) is less than 3 around an IPP (Ionospheric Pierce Point), it does not used in the position calculation. When the number of satellites that meet the GIVE requirement is less than 4, the record is not used for integrity analysis.

4.2 Measured Data Analysis

This paper analyzes the integrity of the EGNOS system by the data obtained from EMS and IGS. Data during 30/4/2017 (day120) to 4/6/2017 (day155) in Germany wtza station was used. Firstly, the abnormal samples are removed in pre-processing. Then, the statistical characteristics of the data are analyzed. Finally, the HMI probability is obtained by EVT extrapolation and the conclusion is explained in detail. As shown in Fig. 3, the curve reflects the change of VPE/VPL with time.

In the graph, the blue point represents the VPE and the red point represents the VPL. The black vertical dotted line divides the daily results. There is a total of 30 days data, which from day120 to day155, and those badly missing data in EMS files (day126–day127, day132–day134 and day137) are removed. As can be seen from the graph, there are some obvious abnormal data in the original VPE/VPL result. Table 1 classifies these anomalies and gives the cause of them.

Fig. 3 Time-varying graph of VPE and VPL during day120 to day155

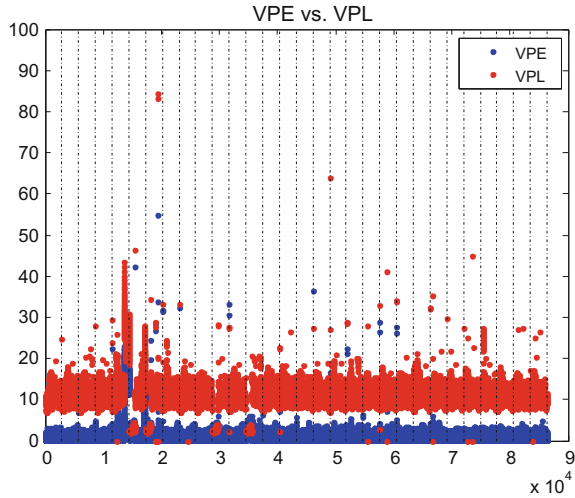


Table 1 Record of abnormal data

Outlier position	Abnormal cause
day124	Only have 0–15 h data, data in 02, 06, 08, 15 h are incomplete and errors exist in 09 h
day125	Only have 08–19 h
day128	1690–1696, 1837–1838, 1845–1852, 1868–1961, 2132–2133: all satellite’s UDRE is ‘Not Monitor’ 1697–1698, 1961–1967: less than 4 satellites meet the UIRE requirement
day130	1448–1449: all satellite’s UDRE is ‘Not Monitor’
day135	There is no 0–06 h data
day138	There is no 0–07 h data
day145	768–769: all satellite’s UDRE is ‘Not Monitor’
day146	1108–1182: all satellite’s UDRE is ‘Not Monitor’
day149	507–509: all satellite’s UDRE is ‘Not Monitor’ 510–515: less than 4 satellites meet the UIRE requirement
day151	726–727, 1554–1555: less than 4 satellites meet the UIRE requirement 728, 1556–1557: all satellite’s UDRE is ‘Not Monitor’
day155	445–447: all satellite’s UDRE is ‘Not Monitor’

In addition, the last hour data of the EMS file in each day is incomplete, affecting the VPE and VPL calculations at the start of the second day. The integrity Stanford graph [8] after the elimination of the outlier data is shown in Fig. 4.

It can be seen that there are no intact abnormal events in the test period. The distribution of SI is further fitted by EVT, and the results of the system integrity are deduced. The following graph gives the probability density function of SI (Fig. 5).

The statistical properties of SI are given in Table 2.

Fig. 4 The Stanford plot of SBAS integrity after abnormal data elimination

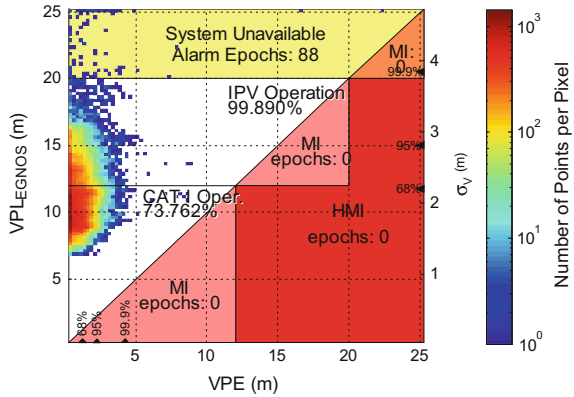


Fig. 5 The probability distribution curve of SI

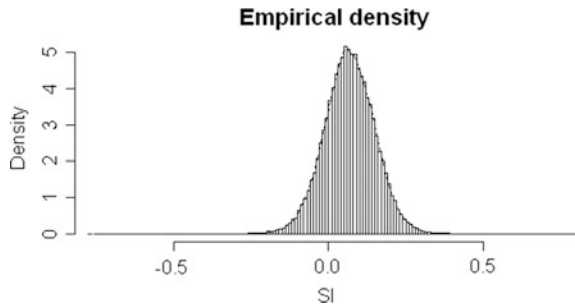


Table 2 The statistical result of SI

Min	Max	Med	Mean	Std	Skew	Kurt
-0.758	0.795	0.067	0.067	0.082	0.030	3.99

Min, Max, Med, Mean, Std, Skew and Kurt respectively describe the minimum, maximum, median, mean, standard deviation, skewness and kurtosis coefficient of the data set. The tail characteristic of random variable distribution is usually determined by kurtosis coefficient. The normal distribution (kurtosis value is 3) is used as the basis of discrimination, and the kurtosis coefficient greater than 3 indicates that the random variable has the characteristic of thick tail. It can be seen that the approximation error of Gauss distribution is not suitable for the integrity analysis which concerns tail characteristics.

The mixture extreme model given above is used to process SI data. The core part uses normal distribution and the tail part uses GPD distribution. The maximum likelihood estimation method is used to estimate the parameters (σ , ξ , u). The results are as follows (the 95% confidence interval is given in parentheses) (Table 3).

Table 3 Estimation of SI fitting results

GPD _l	σ_l	0.0429(0.0419, 0.0439)
	ξ_l	-0.024(-0.0342, -0.0144)
	μ_l	-0.0436(-0.0502, -0.0370)
Norm	Mean	0.0676(0.0673, 0.0679)
	Std	0.0799(0.0797, 0.0801)
GPD _u	σ_r	0.0411(0.0405, 0.0417)
	ξ_r	0.0154(0.0063, 0.0245)
	μ_r	0.1717(0.1707, 0.1727)

The integrity risk can be extrapolated to be:

$$P = \int_{-\infty}^{-1} GPD(x; \sigma_l, \xi_l, \mu_l) \cdot dx + \int_1^{+\infty} GPD(x; \sigma_r, \xi_r, \mu_r) \cdot dx = 2.383 \times 10^{-8} \tag{14}$$

In conclusion, by comparing the ICAO LPV integrity requirements, EGNOS can provide LPV service in the vicinity of wtza station. In the future analysis, it is necessary to study how to extend the integrity evaluation results to the whole service area.

5 Summary

In this paper, the extreme value theory is applied to SBAS integrity evaluation. The samples of limited observation time are reasonably extrapolated, which overcomes the dependence of traditional statistical estimation methods on the number of samples. At the same time, the mixture extreme model is used to estimate the threshold to avoid the adverse effects of the subjective threshold selection method. By using the designed SBAS software receiver, this paper makes a thorough data analysis with the actually measured data from EMS and IGS wtza station, and proves that the EGNOS integrity meets the system design requirements.

Acknowledgements This study is supported by National Science Foundation of China (61601485).

References

1. Kotz S, Nadarajah S (2002) Extreme value distributions. Theory and applications. Price theory and applications. Prentice Hall, London, pp 207–243

2. Minimum operational performance standards for Global Positioning System/Wide Area Augmentation System airborne equipment. RTCA/DO-229C (2001)
3. Veerman H, Kleef A, Wokke F et al (2012) A tool for GNSS integrity verification based on statistical extreme value theory. *Inst Navig* 1655–1665
4. Veerman H, Kleef AV, Imparato D et al (2012) Verifying EGNOS integrity in practice: GIMAT—a tool based on statistical extreme value theory. *The workshop on satellite navigation technologies & European workshop on GNSS signals & signal processing*. IEEE, pp 1–8
5. Ober PB, Imparato D, Verhagen S et al (2014) Empirical integrity verification of GNSS and SBAS based on the extreme value theory. *Navigation* 61(1):23–38
6. Herrera AM, Realini E, Realini E et al (2016) goGPS: open-source MATLAB software. *GPS Solutions* 20(3):595–603
7. Vazquez J, Lacarra E, Sanchez MA (2014) EDAS (EGNOS data access service) for added value applications
8. Tossaint M, Samson J, Toran F et al (2005) The stanford-ESA integrity diagram: focusing on SBAS integrity. *Proc Int Tech Meet Satell Div Inst Navig* 894–905

Reduced ARAIM Subsets Method Determined by Threshold for Integrity Risk



Qian Meng, Jianye Liu, Qinghua Zeng, Shaojun Feng and Rui Xu

Abstract Advanced Receiver Autonomous Integrity Monitoring (ARAIM), based on multi-constellation and multi-frequency of global navigation satellites system, which can provide vertical navigation in terminal approach, has attracted more attentions. The ARAIM user algorithm is based on the multiple hypothesis solution separation within the Integrity Support Message. The total probability of the monitored fault modes subsets must meet the integrity budget. In ARAIM baseline algorithm, the fault mode determination is an independently sequential structure. The number of subsets determined by the maximum simultaneous faults number is redundant which increases computational burdens in fault-tolerant position. A refined subsets determination method is proposed to optimize the calculation. Instead of determining the maximum simultaneous faults number, the proposed method accumulates the fault mode in descending order according to the probability. The threshold for the integrity risk coming from the unmonitored faults is the only parameter to terminate this accumulation and determine the number of subsets. Every subset will be detected the observability and continuity further in this process to make sure that ARAIM can work normally. The structure is more compact and reasonable. The tests results shown that the proposed method is more rigorous as the accumulation processing can make sure the monitored subsets meet the threshold exactly and the accuracy of not monitored risk is increased by three orders of magnitude. The global 99.5% coverage under LPV-200 requirements can be up to 95.94%. The number of subsets is reduced more than 80% significantly without any impacts to the worldwide performance of ARAIM.

Keywords ARAIM · Monitored subsets · Continuity · ISM

Q. Meng (✉) · J. Liu · Q. Zeng · R. Xu
Navigation Research Center, Nanjing University of Aeronautics
and Astronautics, Nanjing, China
e-mail: mengqian@nuaa.edu.cn

Q. Meng · J. Liu · Q. Zeng · R. Xu
Satellite Communication and Navigation Collaborative Innovation Center, Nanjing, China

S. Feng
Centre for Transport Studies, Imperial College London, London, UK

1 Introduction

Advanced Receiver Autonomous Integrity Monitoring (ARAIM) is outlined by the GPS Evolutionary Architecture Study (GEAS) and has been further developed within the EU-U.S. Cooperation on Satellite Navigation Working Group C ARAIM Technical subgroup (ARAIM TSG) [1, 2]. ARAIM aims at providing localizer precision with vertical guidance down to 200 ft altitude (LPV-200) for worldwide aircraft landing navigation [3, 4]. The baseline ARAIM user algorithm is based on multiple hypothesis solution separation (MHSS). The number of subsets and the corresponding fault modes probability are calculated within the satellite prior fault probability and constellation prior fault probability. These information is provided by the ISM. But ISM does not specify which fault modes need to be monitored or their corresponding prior probabilities. This determination must be made by the receiver based on the contents of ISM and positioning resolution. A classical subsets determination method based on the maximum simultaneous faults number is proposed in the baseline ARAIM algorithm [1]. The process includes calculating maximum simultaneous faults number, forming all subsets and filtering the subsets. The method is only sufficient and redundant subsets are involved which will increase the computation burden. Two monitored subset reducing methods: constellation-removed and orbit-removed are introduced into the MHSS processing which can evaluate many fault modes at once and reduce computation load dramatically [5, 6]. However, they are subsets consolidation substantially and do not change the monitored probability and the number of initial subsets as they do not refer to the determination of maximum simultaneous faults number. All the references above do not mention the continuity detection in the fault modes determination, neither.

A refined method is proposed in this paper and the fault subsets are determined based on the threshold of integrity risk. The number of subsets is reduced greatly as the probability of not monitored risk is more accurate. Continuity detection is added in the fault mode calculation which make sure that ARAIM can work normally. Section 2 briefly describes the baseline ARAIM algorithm and the necessary of continuity detection. Section 3 gives the reduced subsets method in details. Section 4 compares and evaluates the worldwide performance of the proposed and baseline methods under dual-constellation. Section 5 is the conclusion.

2 ARAIM Subsets Determination

The baseline ARAIM subsets determination scheme is summarized here briefly and an analysis of the necessary of continuity detection is given then.

2.1 Baseline ARAIM Algorithm

The baseline ARAIM user algorithm is a sequence-structure which based on MHSS. It can be summarized in three steps.

- (1) Calculating maximum simultaneous faults number. The maximum simultaneous faults number $N_{fault,max}$ is defined by the probability of events that not monitored is less than the threshold for the integrity risk.

$$N_{fault,max} = \max\{r \in 1, \dots, N_{sat} | P_{multiple}(r + 1, P_{event,1}, \dots, P_{event,N_{sat} + N_{const}}) \leq P_{FAULT_THRES}\} \tag{1}$$

where $P_{multiple,not_monitored} = P_{multiple}(r + 1, P_{event,1}, \dots, P_{event,N_{sat} + N_{const}})$.

An upper bound algorithm is given in the baseline algorithm.

$$P_{multiple}(r, P_{event,1}, \dots, P_{event,N_{sat} + N_{const}}) = \sum_{i_1 < i_2 < \dots < i_r} P_{event,i_1}, \dots, P_{event,i_r} \leq \frac{(\sum_{k=1}^{N_{sat} + N_{const}} P_{event,k})^r}{r!} \tag{2}$$

Then $N_{fault,max}$ is determined by the following:

$$N_{fault,max} = \varphi(u) = \min\left\{r \mid \frac{u^{r+1}}{(r+1)!} \leq P_{FAULT_THRES}\right\} \tag{3}$$

where $u = \sum_{k=1}^{N_{sat} + N_{const}} P_{event,k}$.

- (2) Forming all subsets. Once the $N_{fault,max}$ is determined, the corresponding fault modes can be searched one by one and from the lower-order fault to the higher-order fault. The probability of the corresponding subset is

$$P_{fault,k} = P_{no_fault} \prod_{i \in idx} \frac{P_{event,i}}{1 - P_{event,i}} \tag{4}$$

where idx is the fault events in the above fault mode. The probability of mode that no fault exists is also the probability of all-in-view subset:

$$P_{no_fault} = \prod_{k=1}^{N_{sat} + N_{const}} (1 - P_{event,k}) \tag{5}$$

- (3) Filtering the subsets. This step is to move the subsets that cannot be monitored (because the remaining satellites do not allow the receiver to compute a position). In this case, these events must be removed from the list of faults (and their integrity risk subtracted from the available budget). This is true of all subsets with three satellites or less belonging to one constellation, or four satellites or less belonging to two or more constellations. Their total integrity risk is noted as $P_{unobservable}$.

Finally the total integrity risk of the modes that are not monitored is:

$$P_{fault,not_monitored} = P_{multiple,not_monitored} + P_{unobservable} \quad (6)$$

The total probability of monitored modes is:

$$P_{monitored} = P_{no_fault} + \sum_{k=1}^{N_{fault}} P_{fault,k} \quad (7)$$

The fault modes determination and the corresponding subsets in baseline ARAIM algorithm is summarized above. The redundancy is large using the mathematical upper bound algorithm to estimate the maximum simultaneous faults number. With the increase of the maximum simultaneous faults number, the number of subsets is exploding. Every subset need to be calculated in the fault-tolerant positioning and protect levels, which increases great computational burdens. If the baseline determination method is a ruler, there's no doubt that the scale is too rough.

At the same time, the probability of unobservable risk is lack of further analysis. Especially whether the threshold for integrity risk can still be met after filtering the subsets need to be judged cautiously. Further, the three steps are in sequential structure which lacks of feedback. Even the system exists severe integrity risk that caused by unobservable subsets, it only be found after the first two steps are finished. The processing is not efficient.

2.2 Continuity Risk

Integrity is defined as the probability that an undetected navigation system error result in hazardous misleading information. Furthermore, continuity risk, or probability of loss of continuity (LOC), is the probability of a detected but unscheduled navigation function interruption after an operation has been initiated [7, 8]. Additional navigation means must be found when LOC occurs. As a consequence, pilots would face increased workloads and more stress would be brought to air traffic controllers.

Normally the probability of LOC is defined as the sum of probability of false alarm, the probability of fault detection when a fault occurs and all other causes of

continuity loss. The critical satellite probability where whose removal is expected to lead to LOC during an approach is analysed in [8] and the conclusion is that the impact of critical satellite outages on the continuity of ARAIM is small. However, the constellation fault is not considered in above analysis. But in ARAIM MHSS, the constellation faults and satellite faults are independent events and the probability of constellation faults that cause critical poor geometry is much more destructive than the critical satellite mentioned above.

And at the same time, according to the definition about the unobservable subsets in baseline algorithm, we can conclude that the unobservable subsets are caused by the constellation fault in most cases. For example, under the current GNSS status, where both Galileo and BeiDou are still under construction, whether GPS constellation fault could be monitored is a serious problem. If GPS constellation fault cannot be monitored, the probability of corresponding subset will cause a serious risk to the ARAIM continuity. Not only the constellation faults, but also any unobservable subsets caused by other cases, for example challenging circumstances, may increase the continuity risk. The continuity detection is essential in fault modes determination.

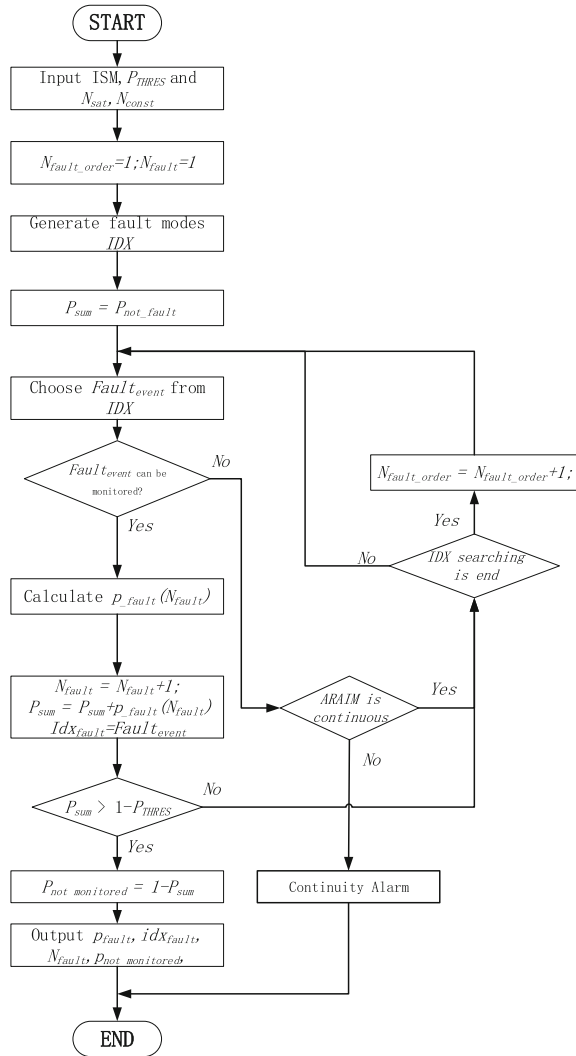
3 Reduced Subsets Determined by Threshold for Integrity Risk

Different from the baseline ARAIM scheme, whose structure is sequential and the core idea is to determine the maximum simultaneous fault number firstly, the proposed refined method here is based on feedback and the number of subsets is determined by the threshold for integrity risk coming from the unmonitored faults. The monitored subsets will be searched from the top probability to the bottom probability. And this search will end immediately once the probability accumulation meets the threshold. Figure 1 gives the detailed flowchart illustrating the proposed subsets determination. The all-in-view subsets, corresponding to the fault-free mode, will be accumulated at first. It is a number close to 1. Then a fault modes search index will be generated which will determine the search order. It is obvious that the probability of low-order fault is bigger than the high-order fault. So the fault modes index can include the low-order fault first, then update to the high-order one. As the probability of events that have no fault approximately equals to 1, to reduce the computation burden, we can only calculate the fault events as follows:

$$P_{\text{fault_order},k} = \prod_{i \in \text{idx}} P_{\text{event},i} \quad (8)$$

After the index is generated, the feedback-structure is working. We will detect whether the subset corresponding to the fault mode can be observed one by one.

Fig. 1 Flowchart illustrating the proposed fault modes determination



Only the observable subsets can be added to the fault mode monitoring index Idx_{fault} and the probability of the fault mode according to Eq. (4) is accumulated to the total monitored probability P_{sum} . The threshold for the integrity risk is tested every time one fault mode has been added to the Idx_{fault} . Once the threshold P_{THRES} is met, the search is terminated and the total integrity risk of the faults that are not monitored is:

$$P_{not_monitored} = 1 - P_{sum} \quad (9)$$

In the protection level calculation, the PHMI is adjusted by the $P_{fault,not_monitored}$:

$$PHMI_{VERT/HOR,ADJ} = PHMI_{VERT/HOR} \times \left(1 - \frac{P_{fault,not_monitored}}{PHMI_{VERT} + PHMI_{HOR}} \right) \quad (10)$$

As shown above, $P_{fault,not_monitored}$ is used as a whole value, it is more concise in the proposed method compared to the baseline method where $P_{fault,not_monitored}$ is calculate as the sum of $P_{multiple,not_monitored}$ and $P_{unobservable}$.

If the subset cannot be observed, the search will not turn to the next fault mode immediately. The continuity detection is triggered. Here gives a concise judgment condition:

$$\sum P_{fault_order,k} \leq P_{THRES} \quad (11)$$

If Eq. (11) holds, then turn to the next fault mode search. Otherwise, the continuity alarm is triggered and ARAIM is unavailable at this epoch.

4 Worldwide Simulation and Evaluation

In this section, we will test the proposed method under GPS and Galileo dual-constellation. The experiments are divided into two parts: the first one will simulate and test the number of subsets and the corresponding probability of not monitored risk under different number of satellites in view, and the second one will evaluate the worldwide performance under GPS and Galileo dual-constellation.

As the ARAIM is under research, determining the parameters in ISM is still one major challenge to enabling ARAIM service. Simulation conditions in ARAIM research are assigned to some experiences. The TSG Milestone 3 report inherits the ISM parameters presented in Milestone 2 report [9]. In this paper, we respect and accept the above settings. The numerical combinations about satellites and constellations prior faults probability are set as $P_{const,j} = 10^{-4}$ and $P_{sat,i} = 10^{-5}$. About the constellation configurations, the ‘optimistic’ GPS and Galileo constellations are chosen to simulate the worldwide circumstance which include 27 GPS satellites in six orbits and 27 Galileo satellites in three orbits. User mask angle is 5° . The simulation results come from a 5 by 5° user grid and 10 sidereal days with 600 s time steps. Figure 2 shows the test results under different number of satellites in view. The number of satellites are from 10 to 30. The blue point line and red solid line represent the baseline method and the proposed method respectively.

As Fig. 2a shown, under 22 satellites, the performance of baseline and proposed methods are the same. The number of subsets that need to be monitored increase with the number of satellite in view. The total number is no more than 50. But there

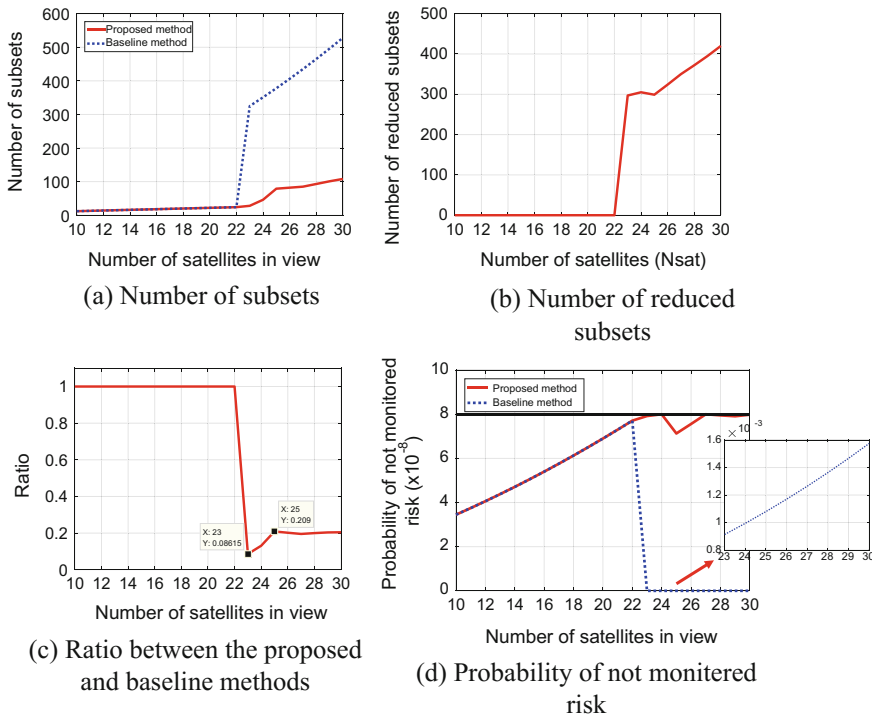


Fig. 2 The test results of number of satellites in view

is a big difference from 23 satellites. The baseline method have a big step as it is determined by the maximum simultaneous faults number. The total number jumps to more than 300 and begins to increase quickly with the number of satellites. But on the other hand, the increase of number of subsets calculated by the proposed method is not obvious. The total number doesn't exceed 100 until 30 satellites. Figure 2b, c show the number of reduced subsets and the ratio between the subsets calculated by the proposed method and the baseline method. They can divided into three stages: from 10 to 22 satellites, the results are exactly the same; 23–25 satellites, due to the big step of the baseline method, the reduced subsets keep around 300 and the ratio increase to 0.2 from only 0.08; from 26 to 30 satellites, the number of reduced subsets increase quickly from 300 to over 400 and the ratio maintains at about 0.2. The proposed method shows obvious superiority. Figure 2d is about the probability of not monitored risk. It shows that for both the above methods, the probabilities of not monitored risk can meet the threshold requirement. But the redundancy of baseline method has exceeded nearly three orders of magnitude from 23 satellites. On the other hand, the proposed method can meet the threshold exactly.

Figure 3 shows the worldwide performance of two methods. The global ARAIM coverages of two methods are essentially identical and that of the proposed method

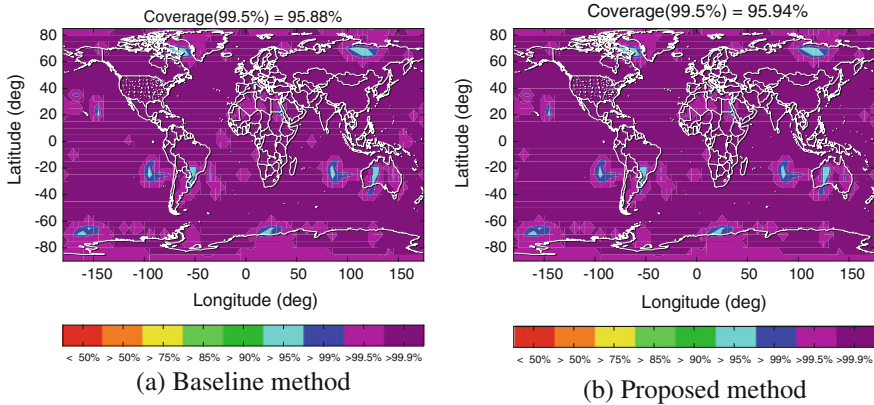


Fig. 3 Global ARAIM coverage of two methods

is even slightly higher (95.94 vs. 95.88%). The reductions of monitored risk and subsets don't affect the availability. Instead, the consistency of subsets solutions is a little improved.

Focus on the number of subsets, Fig. 4 is the distribution map around the world. The color interval is 30. The range of the baseline map is 190–400 and that of the proposed map is 10–220. The southern and western hemispheres are shown in negative coordinates. The subsets of baseline method are much more than those of proposed method. The total number exceed 280 in most parts of the earth, corresponding to over 23 satellites in view. The low latitudes near the equator have the most subsets than any other parts, where most areas have subsets more than 340, even above 370. The middle latitudes have the less subsets. In sharp contrast, the subsets of proposed method are below 40 in most parts of the earth. The low latitudes have the most subsets, which is similar to the trend of baseline method.

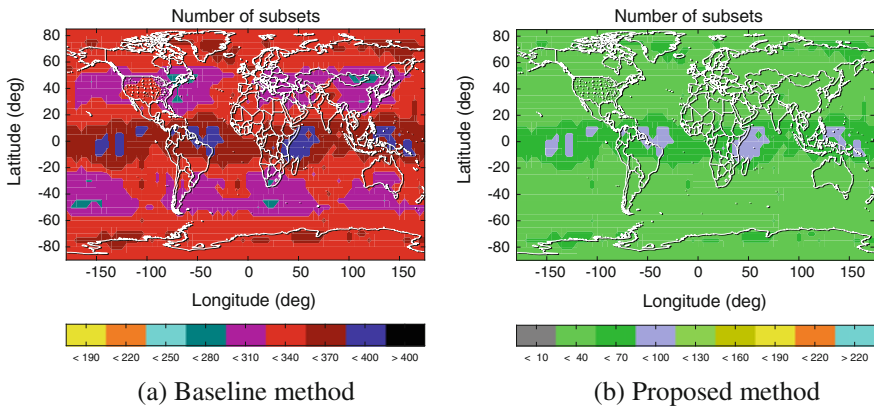


Fig. 4 Number of subsets of two methods

But the total number is less than 100, which is far less than the top number obtained by baseline method. The numbers of subsets at other areas fall into the same range which is 10–40.

Globally, more than 80% and 300 subsets are reduced and the proposed method shows amazing performance. Although the baseline method monitors more fault modes, it is not wise and necessary to calculate so many subsets solutions. The ARAIM availability doesn't benefit from a higher monitored probability but the calculation pressure is exacerbated.

5 Conclusions

The number of subsets and the probability of monitored risk determined by the maximum simultaneous faults number is much redundant. The results of global coverage tests also show that more monitored risk won't improve the ARAIM availability. A reduced subsets method with a full respect to the integrity risk threshold is proposed in this paper. Compared to the baseline method, the number of subsets can be refined to the least (not consider the subsets mержence mentioned in Introduction). The unobservable subset is detected in every fault mode search. The unobservable fault mode won't add to the monitoring index and the probability of not monitored risk can be got clearly after the threshold for the integrity risk is met. Further, continuity detection is considered after the subset is detected as unobservable and a concise judgment condition is given.

The simulations respect to the number of satellites show that the proposed method has dramatic superiority once more than 22 satellites in view. The average reduced subsets are above 300 and more than 80% of the subsets can be reduced. The ratio between subsets obtained by the proposed method and baseline method can be as low as 0.08 at its best. The worldwide test results under GPS and Galileo dual-constellation also verify the effectiveness of the proposed method. More than 80% subsets can be reduced without any impacts to the ARAIM availability under LPV-200. The calculation load and memory resources can be released greatly.

It can be expected that with the increase of GNSS constellations and satellites in space, the number of satellites in view will grow further. The proposed method will provide more advantages in the development of ARAIM.

Acknowledgements This research has been funded by the National Natural Science Foundation of China under Grant No. 61328301 & 61374115 & 61533008 & 61603181; the Funding of Jiangsu Innovation Program for Graduate Education under Grant No. KYLX16_0379), Fundamental Research Funds for the Central Universities (NS2015037).

References

1. Blanch J, Walker T, Enge P et al (2015) Baseline advanced RAIM user algorithm and possible improvements. *IEEE Trans Aerosp Electron Syst* 51(1):713–732
2. EU-U.S. Cooperation on Satellite Navigation, Working Group C (2016) ARAIM technical subgroup milestone 3 report
3. El-Mowafy A, Yang C (2016) Limited sensitivity analysis of ARAIM availability for LPV-200 over Australia using real data. *Adv Space Res* 57(2):659–670
4. El-Mowafy A (2017) Advanced receiver autonomous integrity monitoring using triple frequency data with a focus on treatment of biases. *Adv Space Res* 59(8):2148–2157
5. Walter T, Blanch J, Enge P (2014) Reduced subset analysis for multi-constellation ARAIM. In: *Proceedings of the international technical meeting of the institute of navigation*. San Diego, CA, USA
6. Ge Y, Wang Z, Zhu Y (2017) Reduced ARAIM monitoring subset method based on satellites in different orbital planes. *GPS Solutions* 5:1–14
7. Joerger M, Pervan B (2016) Fault detection and exclusion using solution separation and chi-squared araim. *IEEE Trans Aerosp Electron Syst* 52(2):726–742
8. Zhai Y, Joerger M, Pervan B (2015) Continuity and availability in dual-frequency multi-constellation ARAIM. In: *Proceedings of ION GNSS + 2015*. Tampa, Florida, USA
9. EU-U.S. Cooperation on Satellite Navigation, Working Group C (2015) ARAIM technical subgroup milestone 2 report

Comparison and Analysis of the Optimized AGC Setting and the Improved CNMC Method for the Ground Pseudolite Enhanced Beidou System



Jingyang Fu, Li Wang, Guangyun Li, Yanglin Zhou and Shuaixin Li

Abstract Beidou The pseudolites can be used as a backup signal source of the Beidou System on orbit while in the special geographical environment. It also provides foundation enhancement service for the Beidou system. When the ground-based satellite works on the ground, its observation is seriously affected by the multipath effect. At present, there are mainly signal methods and data processing methods to deal with the multipath in pseudolites pseudo range observation values. In this paper, the optimized AGC settings in signal processing methods and the improved CNMC method in data processing are selected for comparison and analysis. The comparative analysis of the results shows that the improved CNMC method is slightly better than the AGC location algorithm after the pseudolite pseudo range observation.

Keywords Beidou system · Ground-based augmentation · Pseudolites Multipath effect · Improved CNMC

1 Introduction

Pseudo range multipath error is an important error source that affects the positioning accuracy of Beidou satellite navigation system [1]. Is also the same for pseudo-random satellites. Even due to the influence of propagation path and surrounding environment, the pseudo range is affected more seriously by multipath. Pseudo satellite pseudo range multipath error is a non-random error, and it is difficult to eliminate the error by direct differential correction. Through the data analysis of pseudorandom satellite monitoring station, it is found that pseudo range data contain large multipath errors, which is strongly correlated with the layout of

J. Fu (✉) · L. Wang · G. Li · Y. Zhou · S. Li
Information Engineering University, Zhengzhou 450001, China
e-mail: fjynoob@sina.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_58

pseudorandom satellites. When receiving the Beidou satellite visible machine number is less than 4, the pseudo satellite as a required signal source positioning solution. Therefore, how to effectively reduce the error of the pseudo range multipath is the key problem in the cooperative positioning of the Beidou/pseudolites.

For the suppression and elimination of multipath mainly focused on hardware method and data post processing method, the hardware is mainly divided into two categories: one is through the improvement of antenna design and location to reduce received multipath signals of [2], the other is through the signal tracking loop design and multipath estimation method to deal with [3] on digital signal receiver. The hardware method has a good effect on the multi path suppression of the navigation signal, but it needs to use expensive anti multipath antenna and special digital signal circuit design. Data post-processing methods are more easily implemented. There are several common methods: carrier phase smoothing pseudo range [4], wavelet method [5], digital filter, stellar time filtering [6], signal to noise ratio method [7] and antenna array data processing [8]. Among them, the CNMC method has good effect on eliminating the pseudo range multipath and noise, and it is easy to implement. The CNMC method proposed by [9] can effectively weaken the influence of the Beidou satellite's pseudo range multipath. References [10, 11] has proved that the CNMC method is superior to the classical dual frequency Hatch filtering method in smoothing the Beidou satellite pseudo range. Reference [12] research shows that in the GNSS three system combination location, the CNMC smooth Beidou data can be used to improve the positioning performance. The above papers have proved the feasibility and advantages of CNMC for the Beidou satellite pseudo range processing. However, how to apply the CNMC method to pseudo satellite pseudo range multipath processing is seldom studied, and most of the research is based on the simulation data.

In this paper, the optimized AGC setting method and the improved CNMC method for data processing are selected respectively, and the specific formulas are given respectively. At last, using the actual data of pseudo satellite test field, this paper compares and analyzes the positioning accuracy of the two different methods under the same experimental conditions.

2 Optimal AGC Setting of Pseudolites

2.1 *The Concept and Significance of AGC*

AGC is a closed-loop feedback circuit whose function is to adjust the amplitude of the received signal. Because the amplitude of the received signal has a certain volatility, the feedback circuit will gain the amplitude value to maintain the stability of the output signal, and make ADC get the correct quantization value. The output signal of the AGC can be expressed as:

$$r_g(t) = A_g \cdot r_h(t) \tag{1}$$

$$A_g = \sqrt{\frac{1}{\frac{1}{RT} \int_{t-RT}^t r_h^2(t) dt}} \tag{2}$$

It is AGC recovery time. From type 2, the gain size is inversely proportional to the received signal power in the time window, so the setting directly affects the performance of the AGC.

2.2 Optimal Setting of AGC

This section will study and optimize the setting of AGC gain size and recovery time, reduce the SNR loss caused by the quantization operation, and improve the maximum duty cycle of the receiver.

According to Fig. 1, the quantization threshold for any bit ADC is as follows:

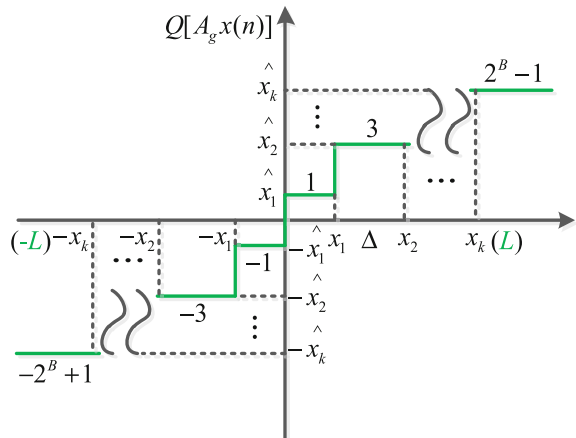
$$B = \{-(2^B - 1), \dots, -3, -1, 1, 3, \dots, 2^B - 1\} \tag{3}$$

Therefore, the quantized output results can be expressed as:

$$r_q[n] = \{2i + 1\}_{i=-2^{B-1}, -2^{B-1} + 1, \dots, 2^{B-1} - 1} \tag{4}$$

It represents the threshold of the decision, the minimum quantized value is, the maximum value is, the intermediate value is evenly distributed between the two. The probability of its output results can be expressed as:

Fig. 1 Hit scalar quantization scheme



$$\begin{aligned}
 p_{B,2i+1} &= P(r_q[n] = 2i + 1) \\
 &= \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2_{IF}}} \int_{\frac{A_g}{2^{B-1-1}} - y[n]}^{+\infty} \exp\left\{-\frac{z^2}{2\sigma^2_{IF}}\right\} dz, & i = 2^{B-1} - 1 \\ \frac{1}{\sqrt{2\pi\sigma^2_{IF}}} \int_{-\infty}^{\frac{A_g}{2^{B-1-1}} - y[n]} \exp\left\{-\frac{z^2}{2\sigma^2_{IF}}\right\} dz, & i = -2^{B-1} \\ \frac{1}{\sqrt{2\pi\sigma^2_{IF}}} \int_{\frac{A_g}{2^{B-1-1}} - y[n]}^{\frac{i+1}{A_g} - y[n]} \exp\left\{-\frac{z^2}{2\sigma^2_{IF}}\right\} dz, & i = \text{else} \end{cases} \quad (5)
 \end{aligned}$$

For the gain, the target signal is the noise variance. According to the type 5, the mean value can be expressed as:

$$\begin{aligned}
 E\{r_q[n]\} &= \sum_{i=-2^{B-1}}^{2^{B-1}-1} (2i+1)p_{B,2i+1} = \sum_{i=0}^{2^{B-1}-1} (2i+1)[p_{B,2i+1} - p_{B,-(2i+1)}] \\
 &= \sum_{i=0}^{2^{B-1}-1} [p_{B,2i+1} - p_{B,-(2i+1)}] + 2 \sum_{i=1}^{2^{B-1}-2} i[p_{B,2i+1} - p_{B,-(2i+1)}] \\
 &\quad + (2^B - 2)[p_{B,2^{B-1}-} - p_{B,-(2^{B-1})}] \\
 &= \frac{1}{\sqrt{2\pi\sigma^2_{IF}}} \int_{-y[n]}^{y[n]} \exp\left\{-\frac{z^2}{2\pi\sigma^2_{IF}}\right\} dz + \frac{2(2^{B-1}-1)}{\sqrt{2\pi\sigma^2_{IF}}} \int_{\frac{(2^{B-1}-1)/A_g - y[n]}{(2^{B-1}-1)/A_g - y[n]}}^{(2^{B-1}-1)/A_g + y[n]} \exp\left\{-\frac{z^2}{2\sigma^2_{IF}}\right\} dz \\
 &\quad + \frac{2}{\sqrt{2\pi\sigma^2_{IF}}} \sum_{i=1}^{2^{B-1}-2} i \left[\int_{i/A_g - y[n]}^{i/A_g + y[n]} \exp\left\{-\frac{z^2}{2\pi\sigma^2_{IF}}\right\} dz - \int_{i/A_g - y[n]}^{(i+1)/A_g + y[n]} \exp\left\{-\frac{z^2}{2\pi\sigma^2_{IF}}\right\} dz \right] \\
 &= \frac{1}{\sqrt{2\pi\sigma^2_{IF}}} \int_{-y[n]}^{y[n]} \exp\left\{-\frac{z^2}{2\pi\sigma^2_{IF}}\right\} dz + \frac{2}{\sqrt{2\pi\sigma^2_{IF}}} \sum_{i=1}^{2^{B-1}-1} \int_{i/A_g - y[n]}^{i/A_g + y[n]} \exp\left\{-\frac{z^2}{2\sigma^2_{IF}}\right\} dz \\
 &= \frac{\text{sign}\{y[n]\}}{\sqrt{2\pi\sigma^2_{IF}}} \left[\sum_{i=-2^{B-1}+1}^{2^{B-1}-1} \int_{i/A_g - |y[n]|}^{i/A_g + |y[n]|} \exp\left\{-\frac{z^2}{2\sigma^2_{IF}}\right\} dz \right] \quad (6)
 \end{aligned}$$

It is a symbol position. The integral result of type 6 can be expressed by the sum of the rectangular area in Fig. 2.

Because the received Beidou satellite signal is smaller than the noise 26 dB, Available $\frac{y[n]}{\sigma_{IF}} \ll 1$, Type 6 is variable:

$$\begin{aligned}
 E\{r_q[n]\} &\approx \frac{\text{sign}\{y[n]\}}{\sqrt{2\pi\sigma^2_{IF}}} \left[\sum_{i=-2^{B-1}+1}^{2^{B-1}-1} 2|y[n]| \exp\left\{-\frac{(i/A_g)^2}{2\sigma^2_{IF}}\right\} dz \right] \\
 &= \frac{2y[n]}{\sqrt{2\pi\sigma^2_{IF}}} \left[1 + 2 \sum_{i=1}^{2^{B-1}-1} \exp\left\{-\frac{(i/A_g)^2}{2\sigma^2_{IF}}\right\} dz \right] \quad (7)
 \end{aligned}$$

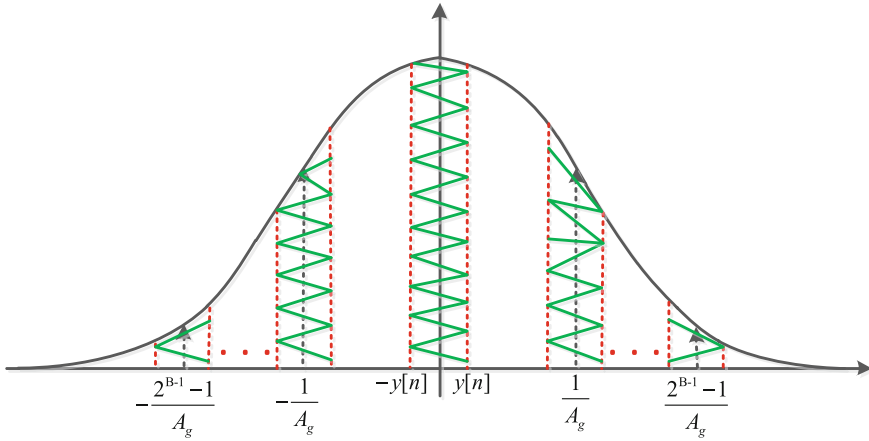


Fig. 2 Representation of the mean probability distribution area of quantized output

In the same way, the square of the mean value can be obtained.

$$\begin{aligned}
 E\{r_q^2[n]\} &\approx \frac{2}{\sqrt{2\pi\sigma_{IF}^2}} \sum_{i=0}^{2^{B-1}-2} (2i+1)^2 \int_{i/A_g}^{(i+1)/A_g} \exp\left\{-\frac{z^2}{2\sigma_{IF}^2}\right\} dz \\
 &\quad + \frac{(2^B-1)^2}{\sqrt{2\pi\sigma_{IF}^2}} \int_{(2^{B-1}-1)/A_g}^{+\infty} \exp\left\{-\frac{z^2}{2\sigma_{IF}^2}\right\} dz \\
 &= \sum_{i=0}^{2^{B-1}-1} (2i+1)^2 \operatorname{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right) - \sum_{i=0}^{2^{B-1}-2} (2i+1)^2 \operatorname{erfc}\left(\frac{(i+1)/A_g}{\sqrt{2}\sigma_{IF}}\right) \\
 &= 1 + \left[\sum_{i=1}^{2^{B-1}-1} (2i+1)^2 \operatorname{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right) - \sum_{i=1}^{2^{B-1}-1} (2i-1)^2 \operatorname{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right) \right] \\
 &= 1 + \sum_{i=1}^{2^{B-1}-1} \left[(2i+1)^2 - (2i-1)^2 \right] \operatorname{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right) \\
 &= 1 + 8 \sum_{i=1}^{2^{B-1}-1} i \cdot \operatorname{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right)
 \end{aligned} \tag{8}$$

Among them, the residual error function in that $\frac{|y[n]|}{\sigma_{IF}} \ll 1$, we know that $E^2[r_b[n]] \ll E[r_b^2[n]]$, $\operatorname{Var}\{r_q[n]\} = E[r_q^2[n]] - E^2[r_q[n]]$, Therefore, the variance can be obtained:

$$\text{Var}\{r_q[n]\} \approx 1 + 8 \sum_{i=1}^{2^{B-1}-1} i \cdot \text{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right) \tag{9}$$

According to the formula 7 and 9, the available bit ADC output SNR is as follows:

$$\text{SNR}_{nc,B} = \frac{2}{\pi} \frac{\left[1 + 2 \sum_{i=1}^{2^{B-1}-1} \exp\left\{-\frac{(i/A_g)^2}{2\sigma_{IF}^2}\right\}\right]^2}{1 + 8 \sum_{i=1}^{2^{B-1}-1} i \cdot \text{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right)} \frac{A^2 N}{2\sigma_{IF}^2} \tag{10}$$

The quantization and quantization before dividing the quantization formula for SNR loss:

$$L_B(A_g) = \frac{2}{\pi} \frac{\left[1 + 2 \sum_{i=1}^{2^{B-1}-1} \exp\left\{-\frac{(i/A_g)^2}{2\sigma_{IF}^2}\right\}\right]^2}{1 + 8 \sum_{i=1}^{2^{B-1}-1} i \cdot \text{erfc}\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right)} \tag{11}$$

Therefore, According to the formula, the SNR loss curve is calculated respectively, as shown in Fig. 3. As can be seen from this diagram, as the number of quantiles increases, the loss of SNR is smaller and the corresponding superiority is greater. Table 1 summarizes the optimal AGC gain.

According to the optimal AGC gain results in the table, set the pseudo satellite signal transmitting terminal, signal source will use pseudo satellite optimal AGC is set after the experiment required, positioning experiment behind, will be carried out in detail below.

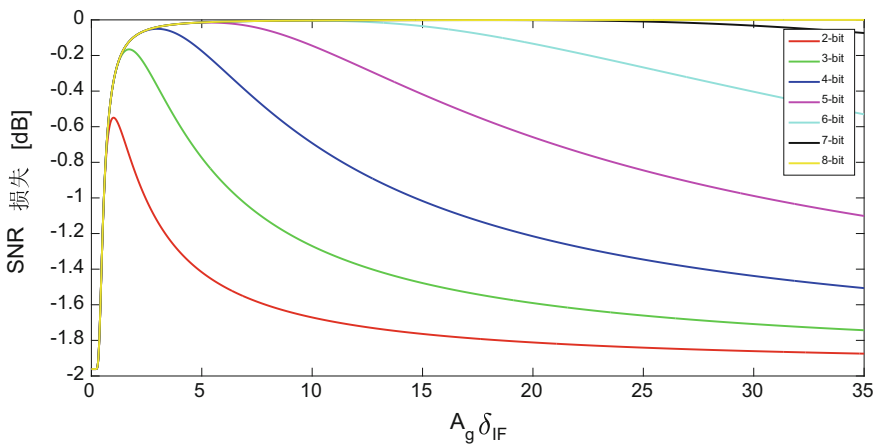


Fig. 3 AGC gain and SNR loss curve

3 Improved CNMC Method for Pseudolites

The multipath error is due to signal propagation path in the process of signal interference caused by the different time delay, positioning terminal not only receive navigation signals directly transmitted, received at the same time around various reflecting material reflection signal, this signal and direct reflection to receive navigation signals generated interference delay, the measured value and the true value deviation between production. The ground fixed pseudorandom signal multipath is different from that of the Beidou satellite on orbit. Compared with the pseudo satellite's error, it has the following characteristics.

- (1) in the Beidou satellite positioning, can choose the star, the elevation is low generally do not participate in satellite positioning solution, this method can weaken the multipath effect; and in the use of pseudo satellite positioning, the elevation angle is low, in addition to considering the special situation of Beidou satellite constellation available the lack of the multipath error to give pseudo satellite measurements for processing;
- (2) a reflection of the multipath error of Beidou satellite signals from the main terminal, and in addition to pseudo satellite multipath signals caused by the reflection of the end, due to the close distance and signal transmitting stations on the ground, so also will be multi path image signals generated by the transmitter;
- (3) the multipath of the Beidou satellite can be analyzed by statistical analysis after periodic observation, so as to get average and weaken.

Pseudorandom satellites are usually installed at known fixed points in precise coordinates. Therefore, when static measurements are made, this multipath error is difficult to eliminate and weaken through periodic observation.

Taking the B1 frequency as an example, the pseudo range and phase observation equations of the pseudo satellite B1 frequency can be expressed as:

$$P_1(t) = R(t) + \varepsilon_{trop}(t) + \varepsilon_{sta}(t) + \varepsilon_{clk}(t) + \eta_P(t) \quad (12)$$

$$L_1(t) = R(t) + \varepsilon_{trop}(t) + \varepsilon_{sta}(t) + \varepsilon_{clk}(t) + \lambda_1 N_1 + \eta_L(t) \quad (13)$$

The error of the station coordinates is expressed in the formula, and the rest of the meaning is similar to the CNMC method of the Beidou satellite, which can be seen in the document [13]. Obviously, the pseudo range and phase observations of pseudorandom satellites are no longer affected by ionospheric delay. Therefore, the Beidou satellite's CNMC method can not be directly applied to pseudo range multipath processing, which needs improvement. The difference between pseudo satellite's pseudo range and phase observation can be expressed as:

Table 1 Optimal AGC gain (unit: dB)

Quantization bits B	Optimal $A_g \sigma_{IF}$	Minimum SNR loss [dB]
2	1.0	-0.5495
3	1.7	-0.1657
4	2.9	-0.0504
5	5.3	-0.0152
6	9.6	-0.0045
7	17.6	-0.0013
8	32.5	-0.0003

$$\begin{aligned} CMC_1(t) &= P_1(t) - L_1(t) \\ &= -N_1 \lambda_1 + \eta_P(t) - \eta_L(t) \end{aligned} \quad (14)$$

Similar to formula (3), which can be ignored. If required, it is required to be taken first, and the item is recorded as the following two formulas.

$$Nc(t_0) = P_1(t_0) - L_1(t_0) \quad (15)$$

$$Nc(t_i) = Nc(t_{i-1}) + \frac{1}{N} (P_1(t_i) - L_1(t_i) - Nc(t_{i-1})) \quad (16)$$

The first use of type (15) to initialize, and then use the type (16) of the each epoch, with the prolongation of time, calculation accuracy will be higher.

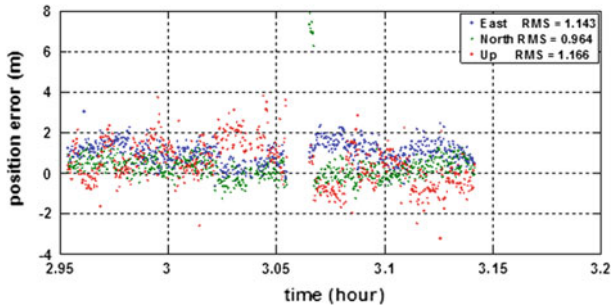
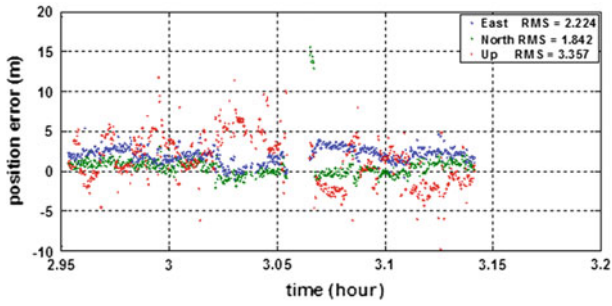
Then, the pseudo range multipath and noise error can be obtained according to the formula (14).

$$\eta_P(t) = P_1(t) - L_1(t) - Nc(t) \quad (17)$$

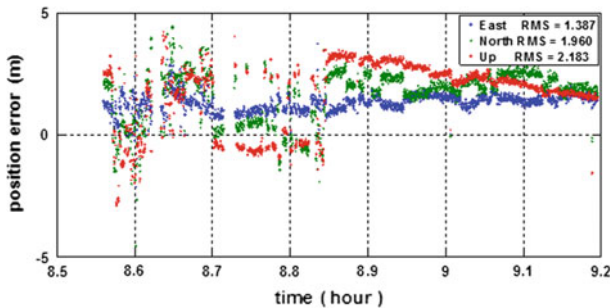
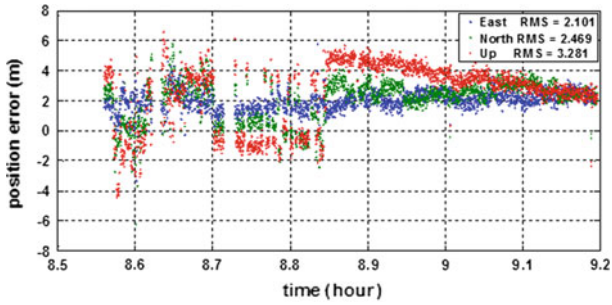
In the original pseudo range observation value of the pseudo satellite, the high precision pseudo range observation value can be obtained by eliminating the formula (17). It is still necessary to note that when the week jump is found, the solution is required to be initialized. Comparing the CNMC method and related formula of the pseudo satellite and the Beidou satellite, it is found that the pseudo satellite CNMC method is more concise, and only the single frequency data can be realized.

4 Comparison and Analysis of the Two Methods

In July 2016, the test vehicle terminal dynamic positioning in Beidou ground-based pseudolite demonstration area, the Beidou satellite and pseudo satellite positioning mode in common, the base station set up dual frequency base station equipment in

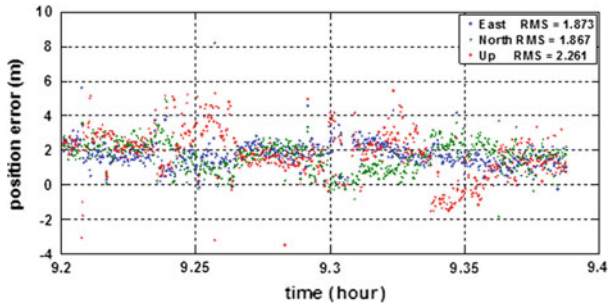
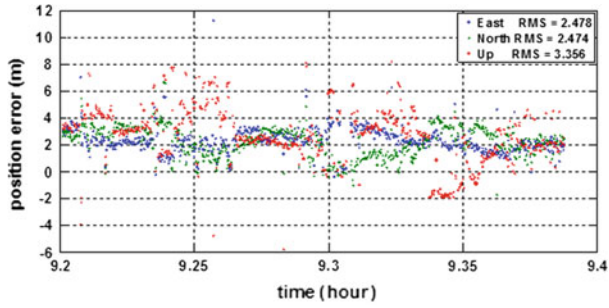


(a) comparison of first groups of positioning results

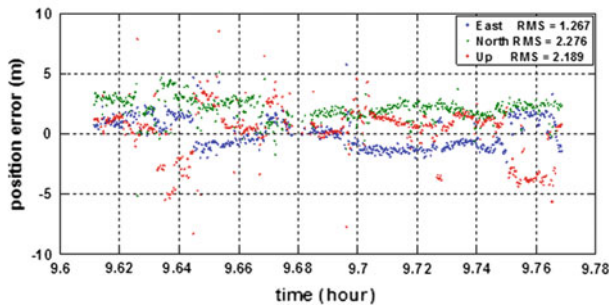
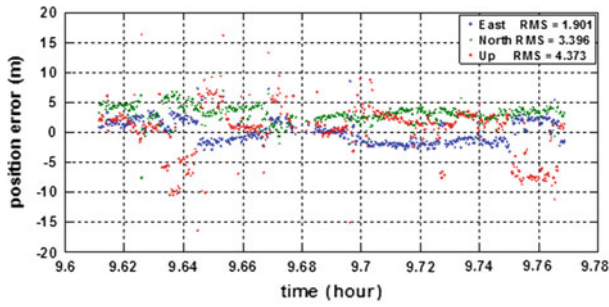


(b) comparison of second groups of positioning results

Fig. 4 Comparison of positioning results by different method

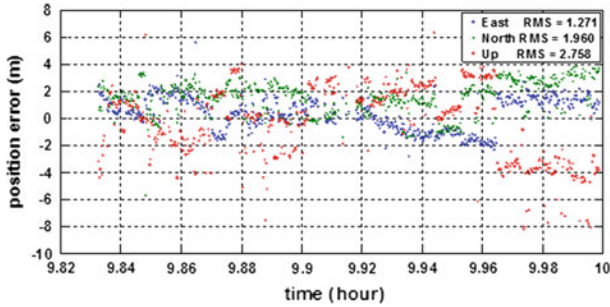
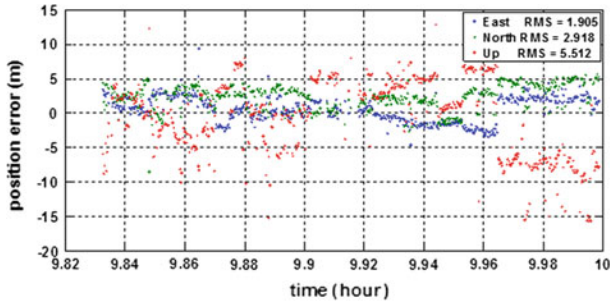


(c) comparison of third groups of positioning results

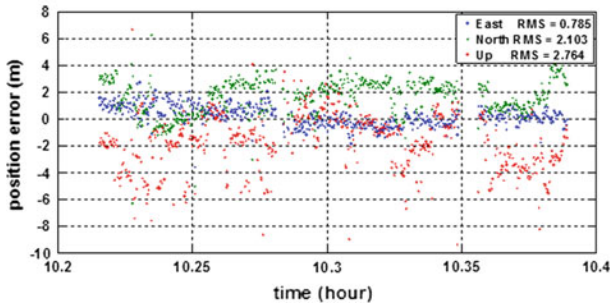
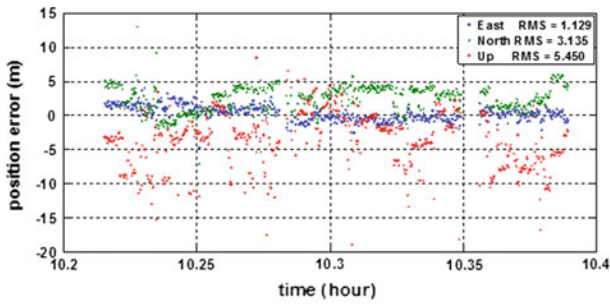


(d) comparison of fourth groups of positioning results

Fig. 4 (continued)



(e) comparison of fifth groups of positioning results



(f) comparison of sixth groups of positioning results

Fig. 4 (continued)

Table 2 Comparison of positioning accuracy between AGC and CNMC (RMS, unit m)

Test group number	Horizontal precision			Elevation precision		
	AGC	CNMC	Difference	AGC	CNMC	Difference
1	2.888	1.495	1.393	3.357	1.166	2.191
2	3.242	2.401	0.841	3.281	2.183	1.098
3	3.502	2.645	0.857	3.356	2.261	1.095
4	3.892	2.605	1.287	4.373	2.189	2.184
5	3.485	2.336	1.149	5.512	2.758	2.754
6	3.332	2.245	1.087	5.451	2.764	2.686
Avg	3.390	2.288	1.102	4.222	2.220	2.001

Beidou; demonstration car were set up two sets of enhanced positioning terminal, pseudo signal optimization were recorded after AGC, and pseudo distance no signal optimization AGC; according to the chosen route, continuous testing; sampling interval dynamic test for the second one, the original observation data recorded two sets of pseudo base stations, satellite receiver demo car. Take into account the time reduction and coordinate conversion, the two group enhanced positioning terminal positioning results of real-time preliminary comparison test; measurement data using base station and demonstration vehicle for high precision calculation (relative positioning solution), and the two positioning results and enhanced positioning terminal positioning results by the time reduction after alignment test.

There are a total of 6 sets of experiments, observation data by post-processing method, measurement is given and the Beidou pseudo satellite positioning solution joint positioning results under the mode of positioning accuracy evaluation of the two methods, each of the above diagram is to optimize the AGC method, the following is a modified CNMC method. The concrete results are as follows (Fig. 4).

Table 2 shows, for ground users, the experimental conditions in this paper, by the 6 groups of experimental data statistics shows that the level of accuracy using the improved CNMC method because the optimization setting method of AGC to improve 1.102 m, elevation accuracy using improved CNMC method because the optimization methods to enhance the 2.001 m AGC settings.

5 Conclusion

This paper briefly introduces the research and development process of pseudorange multipath processing, describes the AGC optimization setting method of pseudo satellite augmented Beidou system and the improved CNMC method, designs and conducts comparative experiments. According to the data processing and analysis of the experimental results, we can draw the following conclusions:

- (1) when the pseudorange is used as a supplementary signal source for the Beidou system and used as a ground-based augmented signal transmitting device, the

- optimization of AGC settings and the improvement of CNMC method can effectively improve the positioning accuracy.
- (2) under the experimental conditions described in this paper, it can be concluded that the improved CNMC method with the use of pseudorandom satellites is better than the AGC optimization settings alone.
 - (3) the next step will be to study the effect of optimizing the AGC setting and improving CNMC on the positioning accuracy and other performance indicators.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 41274014) and (No. 41501491).

References

1. Zhou W (2013) Research and realization on theories and methods of precise positioning base on BeiDou navigation satellite system. PLA Information Engineering University, Zhengzhou, China
2. Cobb HS (1997) GPS pseudolites: theory, design, and applications. Stanford University
3. Seo S, Park J, Suk J et al (2014) A design of dual frequency bands time synchronization system for synchronized-pseudolite navigation system. *J Pos Navig Timing* 3:71–81
4. Hatch R (1983) The synergism of GPS code and carrier measurements. In: International geodetic symposium on satellite doppler positioning, vol 1, pp 1213–1231
5. Wang M, Chai H, Liu M et al (2016) BDS relative static positioning over long baseline improved by GEO multipath mitigation. *Adv Space Res* 57(3):782–793
6. Zheng DW, Zhong P, Ding XL et al (2005) Filtering GPS time-series using a Vondrak filter and cross-validation. *J Geodesy* 79(6–7):363–369
7. Wu Y, Chen X, Wu C (2008) Mitigation of multi-path effect using SNR values. *Geomatics Inf Sci Wuhan Univ* 8(018):842–845
8. Nie J (2012) Study on GNSS antenna array anti-jamming algorithm and performance evaluation key techniques. National University of Defense Technology, Changsha
9. Wu XL, Zhou JH, Wang G et al (2012) Multipath error detection and correction for GEO/IGSO satellites. *Sci China Phys Mech Astron* 55(7):1297–1306
10. Chang Z, Hu X, Guo R et al (2015) Comparison between CNMC and hatch filter & its precision analysis for. *Sci Sin Phys Mech Astron* 45(7):079508
11. Cao Y, Hu X, WU B et al (2012) The wide-area difference system for the regional satellite navigation system of COMPASS. *Sci China Phys Mech Astron* 55(7):1307–1315
12. Zhang Y, Chen JC, Wu B et al (2014) GPS/GLONASS/COMPASS combined positioning based on CNMC. In: China satellite navigation conference (CSNC) 2014 proceedings: vol III. Springer, Berlin, pp 523–532
13. Fu J, Zhou J, Li G (2016) The positioning performance analysis of BeiDou/Pseudolites collaboration by CNMC method. *Acta Geod Cartogr Sin* 45(S2):188–193

Research on BDS/GPS Integrated Navigation Satellite Selection Algorithm Based on Particle Swarm Optimization



Ershen Wang, Chaoying Jia, Tao Pang, Pingping Qu
and Zhixian Zhang

Abstract In the multi-constellation satellite navigation system, all the visible satellites are used for positioning, which will increase the computation amount of the receiver and affect the real-time positioning. How to quickly and effectively select visible satellites for positioning is a research topic. For this problem, a satellite selection algorithm based on Particle Swarm Optimization (PSO) is proposed. In this method, the visible satellite is numbered, random grouping, and each group as a particle; the velocity-displacement model in the PSO keeps the particles gradually close to the minimum value of the GDOP. Under a series of simulation experiments, the key parameters such as inertia weight factor, acceleration coefficient and maximum velocity of PSO are determined. Besides, local search based on chaos mechanism are introduced, which can avoid the results of PSO algorithm into local optimum. Finally, the performance of satellite selection with PSO is confirmed to be remarkable by the simulation experiments under different numbers of selected satellites. The results show that this method can quickly select satellites under BDS and GPS system, and the result meets receiver positioning accuracy.

Keywords BeiDou navigation satellite system (BDS) · Global positioning system (GPS) · Satellite selection · Particle swarm optimization (PSO)
Geometric dilution of precision

E. Wang (✉) · C. Jia · T. Pang · P. Qu · Z. Zhang
School of Electronic and Information Engineering, Shenyang Aerospace University,
Shenyang, China
e-mail: wanges_2016@126.com

E. Wang
Liaoning General Aviation Key Laboratory, Shenyang Aerospace University,
Shenyang, China

1 Introduction

With the development of the Global Positioning System (GPS) of the United States, GLONASS of Russia, Galileo of the EU and BeiDou of China, the number of visible satellites has increased significantly, and the combination of constellations will improve the user's satellite geometry. Abundant navigation signal source provides the conditions that can improve the performance of the navigation and positioning, but it also brings signal processing burden of the receiver. If all visible satellites are used for positioning, it will be difficult to guarantee the real-time positioning. Therefore, selecting the best set of visible satellites is necessary from the multiple visible satellites, which can improve the accuracy and real-time performance of navigation and positioning.

The measure standard of satellite selection algorithm is the spatial geometry distribution between users and visible satellites. According to the characterization of the spatial distribution between user and visible satellites, dilution of precision is the most widely used in multiple parameters [1]. Among them, the Geometric Dilution of Precision (GDOP) is related to geometry between user and satellites. The smaller its value, the better the spatial geometric distribution and the higher the positioning accuracy. At present, researchers mainly focus on three aspects: the calculation of GDOP [2, 3], the fast satellites selection algorithm based on geometric distribution [4] and the satellites selection algorithm based on genetic algorithm [5, 6]. All of algorithms above mentioned can reduce the calculation of satellites selection to some extent, and get an effective result of satellites selection in shorter time. Similar to genetic algorithm, particle swarm optimization (PSO) is a bionic algorithm. The system forms the initial population by randomly generating a set of solutions. And then, according to the adaptation information of individuals, the algorithm starts to search among the population. Finally, PSO system will find the optimal solution in the search space. The PSO algorithm is well applied in many fields. This paper attempts to apply the PSO algorithm to satellites selection, and the PSO satellites selection model is established. By introducing a local search based on chaos mechanism, the global search ability of the algorithm is improved. The performance of the PSO was also verified by comparing with the results of satellites selection of traversal method. In the case of satellites selection of BDS/GPS, the number of satellites selection is more than 5, so the PSO satellites selection takes less time and has higher efficiency.

2 Calculation Method on DOP

The position of the navigation receiver position consists of the satellite position (x_u, y_u, z_u) and the pseudo-range ρ_i ($j = 1, 2, \dots, n$, n is the total of visible satellites received by the receiver at the current time), the pseudo-range observation equation is as follows:

$$\rho_j = \sqrt{(x_j - x_u)^2 + (y_j - y_u)^2 + (z_j - z_u)^2} + c(t_u - t_j) + \varepsilon_\rho \tag{1}$$

ρ_j is the value of the pseudo-range observations, c is the speed of the light, t_u is the prior time of the receiver relative to the GPS time, t_j is the prior time that the j th satellite's clock relative to the GPS system, ε_ρ is the observed noise of the pseudo-range (including the ionospheric, tropospheric delay and delay caused by multipath effects).

In the process of positioning, the approximate position of the receiver is assumed to be $\hat{\rho}_j = (\hat{x}_u, \hat{y}_u, \hat{z}_u)$, the estimated clock bias is \hat{t}_u . The first-order Taylor series expansion of (1) at the receiver approximate position and the clock bias yields the linearized equation:

$$\Delta\rho_j = -\frac{x_j - \hat{x}_u}{\rho_0} \Delta x_u - \frac{y_j - \hat{y}_u}{\rho_0} \Delta y_u - \frac{z_j - \hat{z}_u}{\rho_0} \Delta z_u - c\Delta t_u + \varepsilon_\rho \tag{2}$$

where, $\rho_0 = \sqrt{(x_j - \hat{x}_u)^2 + (y_j - \hat{y}_u)^2 + (z_j - \hat{z}_u)^2}$, $\Delta\rho_j = \rho_j - \hat{\rho}_j$, $\Delta x_u = x_u - \hat{x}_u$, $\Delta y_u = y_u - \hat{y}_u$, $\Delta z_u = z_u - \hat{z}_u$, $\Delta t_u = t_u - \hat{t}_u$.

$l_i = -\frac{x_j - \hat{x}_u}{\rho_0}$, $m_i = -\frac{y_j - \hat{y}_u}{\rho_0}$, $n_i = -\frac{z_j - \hat{z}_u}{\rho_0}$, respectively represent the direction cosine of the j th satellite unit vector from the approximate position. Equation (2) can be written in matrix form:

$$\Delta\rho = H\Delta X \tag{3}$$

where,

$$H = \begin{bmatrix} l_1 & m_1 & n_1 & 1 \\ l_2 & m_2 & n_2 & 1 \\ \dots & \dots & \dots & \dots \\ l_n & m_n & n_n & 1 \end{bmatrix}, \Delta X = \begin{bmatrix} \Delta x_u \\ \Delta y_u \\ \Delta z_u \\ -c\Delta t_u \end{bmatrix}, \Delta\rho = \rho_j - \hat{\rho}_j \tag{4}$$

The typical implementation of the solution algorithm is an iterative linear least-squares method, assuming that the pseudo-range error is $d\rho$ [1], and the components of the same distribution and independent of each other, the mean squared error equal to the satellite user range error, that is:

$$\text{cov}(d\rho) = I_{n \times n} \sigma_{URE}^2 \tag{5}$$

So,

$$\text{cov}(dX) = (H^T H)^{-1} H^T \text{cov}(d\rho) H (H^T H)^{-1} = (H^T H)^{-1} \sigma_{URE}^2 \quad (6)$$

The dilution of the precision can be defined by the symmetric matrix $(H^T H)^{-1}$, the geometric dilution of the precision GDOP can be defined as the square root of the covariance matrix, and the other commonly used performance indicators are HDOP (Horizontal Dilution of Precision), PDOP (Position Dilution of Precision), TDOP (Time Dilution of Precision).

$$GDOP = \sqrt{\text{trace}\{(H^T H)^{-1}\}} \quad (7)$$

For BDS/GPS dual constellation, due to the different range error performance, it is usually done by appending extra columns in Eq. (4).

$$H = \begin{bmatrix} l_{G,1} & m_{G,1} & n_{G,1} & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{G,m} & m_{G,m} & n_{G,m} & 1 & 0 \\ l_{B,1} & m_{B,1} & n_{B,1} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{B,n} & m_{B,n} & n_{B,n} & 0 & 1 \end{bmatrix} \quad (8)$$

In Eq. (8), the subscript G represents the GPS constellation, the number of visible satellites is 1..m, subscript B represents the BDS constellation, the number of visible satellites is 1..n. GDOP is still calculated using Eq. (7), which is a function only related to the satellite and user geometry layout. The value of GDOP changes with the change of the number of satellites involved in navigation, and decreased monotonously with the increase of satellites. The details are given in [7].

DOP is an important basis for satellites selection algorithm. The smaller the DOP is, the higher the accuracy of positioning resolution is. The satellites selection according to DOP can ensure the positioning accuracy [1].

3 Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO), as a hot research topic in many fields, has been successfully applied in multiple fields (including neural network training, power system, communication, image processing, etc.) [8]. The PSO algorithm is a stochastic optimization algorithm based on swarm intelligence. It mimics the foraging behavior of birds. The problem search space is analogized to the flight space of birds, and each bird is abstracted into a particle to represent a candidate solution of the problem, the optimization needs to find the optimal solution is equivalent to

looking for food. As a group optimization algorithm proposed by the simulation of biological predation in nature, the PSO algorithm can solve the optimization problem through the cooperation and competition among individuals in the population.

3.1 Basic Principle of PSO Algorithm

PSO is initialized to a random set of random particles (random solutions). In each iteration, the particle updates itself by tracking two “extrema” [9–11]. The first one is the optimal solution found by the particle itself, and called the individual extrema (pbest). The other is the optimal solution found by the entire population, and called the global extrema (gbest). After a limited number of iterations, each particle gradually approaches the optimal solution, and eventually finds the optimal solution. Taking the i th particle in d -dimensional search space as an example, its position and velocity are expressed as $X_i = [x_{i1}, x_{i2}, \dots, x_{id}]$, $V_i = [v_{i1}, v_{i2}, \dots, v_{id}]$.

The fitness value of each particle is evaluated by the fitness function, and the best position (pbest) $P_i = [p_{i1}, p_{i2}, \dots, p_{id}]$ after t iterations, and the best position (gbest) found in the population. Then, PSO algorithm updates velocity and position of each particle by follow equations:

$$v_{i,j}(t+1) = \omega v_{i,j}(t) + c_1 r_1 [p_{i,j} - x_{i,j}(t)] + c_2 r_2 [p_{g,j} - x_{i,j}(t)] \quad (9)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1), \quad j = 1, \dots, d \quad (10)$$

where, ω is the inertial weight factor, c_1 and c_2 is acceleration coefficient, r_1 and r_2 is a random number that is evenly distributed between 0 and 1. In addition, by setting the velocity limitations $[v_{\min}, v_{\max}]$ and the position range $[x_{\min}, x_{\max}]$ of the particles, the movement of the particles can be appropriately restricted.

Different from genetic algorithm, PSO algorithm uses real number coding, standard PSO algorithm does not have more parameters to be adjusted. The search performance depends on the balance between global search and local improvement, which depends largely on the control parameters of the algorithm, including the size of population, maximum velocity, maximum iteration, inertia weight factor and acceleration coefficient. Standard PSO does not have many parameters to adjust, usually select according to experience [12].

The velocity limitations of particle V_{\max} determine the particle’s maximum move distance in an iteration cycle. Shi gives the simulation results of the PSO system optimization experiment, when V_{\max} is the alone value of $[2, V_{\max}]$ and the velocity inertia within $[0.1, 1.05]$.

The acceleration factor c_1 and c_2 is usually chosen as 2, Clerc recommended the value of $(c_1 + c_1)/2$ is 1.494.

Inertia factor ω , the value is generally between (0, 1), Clerc suggested that its value is 0.729, and Riget adopted strategy which decreasing ω with the different number of iteration.

Termination conditions generally choose to reach the maximum cycle setting, or to meet the predetermined error requirements.

3.2 Chaotic Particle Swarm Optimization Algorithm

In order to make the particle swarm optimization algorithm faster at the optimal search time and avoid falling into the local optimum, the chaos concept, the time-varying inertia weight and the time-varying acceleration coefficient are introduced in PSO [13]. For a given optimization function, CPSO correspond to the search process as chaotic orbit traversal process, the search process has the ability to avoid local extremum, and finally obtain the global optimal solution or satisfactory solution. Optimization based on chaotic sequence includes two key steps:

- (1) Mapping chaos space to the solution space of optimization problem;
- (2) The use of chaotic dynamic characteristics to achieve the solution space search.

Logistic maps are one-dimensional discrete-time nonlinear systems and exhibit quadratic nonlinearity [12]. Logistic mapping is expressed as

$$x_{i+1} = f(x_i) = \mu x_i(1 - x_i) \quad i = 0, 1, 2, \dots \quad (11)$$

where, $x_i \in (0, 1)$, when $x_0 \notin \{0, 0.25, 0.5, 0.75, 1\}$ and $\mu = 4$, the sequence generated by Logistic mapping presents the chaotic dynamic characteristics. Small changes of the initial variables will lead to huge differences in the subsequent tracks.

4 The Steps of PSO Satellites Selection

Suppose at a certain moment, the number of visible satellites is n , and select a certain number of m from them. The GDOP value as small as possible. In the PSO satellites selection, each particle represents a combination of visible satellites, the location of the particle is determined by m elements, each element represents a visible satellite.

- (1) Coding

Randomly array all the visible satellites observed by the receiver at the current time, and then sequentially code the visible satellites from $1, \dots, n$ to the next, and the serial numbers correspond to the visible satellites one by one and remain unchanged in subsequent operations.

(2) Generated initial population randomly

Set the population with M , the initial population of $G_0 = \{g_{0i}\} (i = 1, \dots, M)$, each particle $g_{0i} = [g_{0,1}, g_{0,2}, \dots, g_{0,m}]$ is taken from all visible satellites in initial population. In order to make the number of selected satellites is the same, we must ensure that each element is different in g_{0i} . Initialization velocity $v_0 = [v_{0,1}, v_{0,2}, \dots, v_{0,m}]$, the subscript "0" means that the particles go through 0th iterations, that means the initial position and velocity. After repeated experiments, it is concluded that, in order to avoid the search results of PSO falling into local optimum, a chaotic sequence is introduced into the initial population. First, a random number z_i of (0, 1) is generated, and z_i is updated according to formula (11). the chaotic space is mapped to the solution space to be optimized according to the formula $x_i = x_{imin} + z_i(x_{imax} - x_{imin})$.

(3) Fitness calculation

The fitness used in this paper is the GDOP value which is the combination of visible satellites with encoding, and denoted as the target value $f_{t,i}$ of the particle, and the subscript "t" means that the particle go through t th iterations. The particles in the initial population are substituted into the fitness function in turn, and the fitness value of each particle is obtained. choose the smallest fitness value in population as the initial global optimal position g_{best} and each particle itself position as the initial local optimal position p_{best} .

(4) Updated

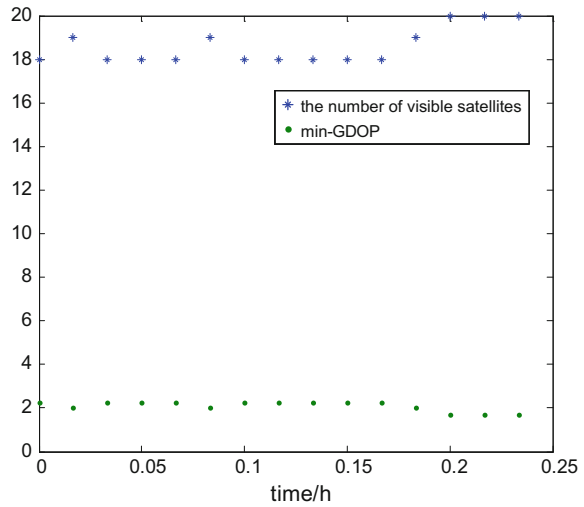
For each particle, according to the position and velocity update formulas of PSO, the position $g_{t,i}$ and the velocity $v_{t,i}$ of the particle in the population are continuously corrected, and the target value corresponding to $f_{t,i}$ is calculated respectively.

5 Experiment and Simulation Analysis

PSO satellites selection is to solve the GNSS constellations, and the number of satellites selected more than 5. This paper select BDS/GPS dual constellations, the receiver coordinates are $[-2279827.3156, 5004704.3094, 3219776.2093]$, the number of satellites selected is 6. The satellite position is calculated by ephemeris, and the cut-off height angle of the satellite is set as 5° . The simulation time start at 2016.7.31.00:00:00, the simulation duration is 900 s, and the simulation step is 60 s.

In this paper, we use the traversal method to get the GDOP value, and choose its value as a reference. Assuming that the number of all visible satellites is n , then select m visible satellites from them. Thus, the number of combinations are C_n^m . The so-called traversal method is to calculate the GDOP values of combinations one by one, and obtains the minimum value of GDOP. The number of BDS/GPS visible satellites and their corresponding minimum GDOP values are shown in Fig. 1.

Fig. 1 The number of BDS/GPS visible satellite and their min-GDOP



As can be seen in Fig. 1, at the same time, the receiver receives about 18 BDS/GPS visible satellites, take selecting 6 visible satellites from 18 visible satellites as an example, it needs to calculate GDOP values $C_{18}^6 = 18,564$ times. Each time-consuming about 2.6 s or so.

Set the algorithm parameters according to the PSO satellites selection procedure in Sect. 4: Number of iterations $MaxIt = 50$, Inertia coefficient $\omega = 0.7298$, Acceleration coefficient $c_1 = c_2 = 1.4962$, $V_{max} = 2$. GDOP value is calculated at a certain moment. As the number of iterations increases, the GDOP value changes as shown in Fig. 2.

By traversing method, the minimum value of GDOP at current moment is 2.25. From the results in the figure, the convergence velocity of PSO algorithm is fast,

Fig. 2 GDOP values change with different

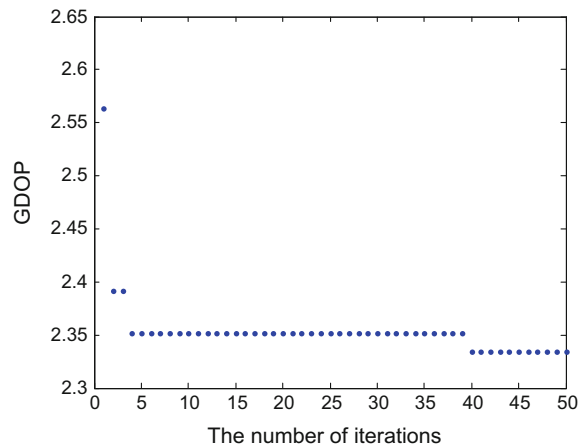
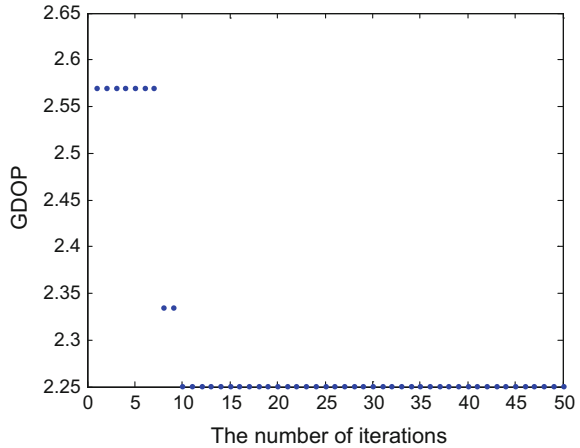


Fig. 3 GDOP values change with different number of PSO iterations number of CPSO iterations



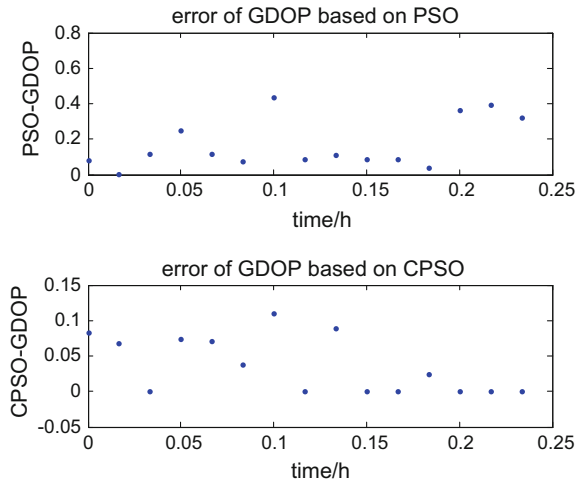
and the number of iterations is stable around 2.34 when it is less than 10th iterations. Obviously, PSO algorithm emerges “premature” phenomenon (that is trapped in the local optimum). The following uses CPSO algorithm, and the introduction of time-varying inertia weight and time-varying acceleration coefficient. Time-varying inertia weight ω usually decrease linearly from 0.9 to 0.4, and are expressed as $\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min}) \times \frac{it}{\text{MaxIt}}$ in the iterative process. $\omega_{\max}, \omega_{\min}$ is the maximum and minimum inertia coefficients, respectively; it, MaxIt is the current iteration number and the maximum iteration number. The maximum number of iterations MaxIt is 50. Simulation test at the same moment, the results is obtained in Fig. 3.

It can be seen from this figure that the CPSO algorithm also converges when the number of iterations is less than 10, and effectively improves the disadvantage that PSO satellites selection algorithm fall into the local optimum. When the number of iterations is 50, the time-consuming is 1.512039 s. If the number of iterations is reduced to 15 times, the satellites selection time will be shorter.

The PSO and CPSO is applied to select the satellite, and each result is compared with the results of the traversal method. GDOP calculation error of the new algorithm is shown in Fig. 4.

It can be seen from the figure that the CPSO algorithm has a more stable GDOP error than the standard PSO algorithm. The calculation error of PSO-GDOP is between 0 and 0.5, while the calculation error of CPSO-GDOP is basically stable at 0–0.15.

Fig. 4 The results error of PSO and CPSO select satellites



6 Conclusions

In this paper, we present a method of selecting satellites by the Particle Swarm Optimization (PSO) algorithm. A satellites selection model of PSO is constructed and introducing the concept of chaos, effectively avoids the limitation of the PSO algorithm which is easy to get into the local optimum and improves the system stability. The simulation results show that under the BDS/GPS dual constellations, and the number of satellites selection is more than 5, the satellites selection time of proposed algorithm is shortened to 42.3% than traverse method, and the minimum GDOP calculation error is stable within a certain range, which can quickly select the satellites.

Acknowledgements This study is funded by the National Natural Science Foundation of China (61571309 & 61101161), the Liaoning BaiQianWan Talents Program, Scientific Study Project for Liaoning Province Ministry of Education (L201716), Liaoning Excellent Talents in University (LR2016069) and the Aerospace Science Foundation of Avic (2015ZC54010).

References

1. Zhang J (2004) Air-ground collaborative airspace surveillance. Aviation Industry Press, Beijing
2. Phatak MS (2001) Recursive method for optimum GPS satellite selection. *IEEE Trans Aerosp Electron Syst* 37(2):751–754
3. Meng F, Wang S, Zhu B (2015) GNSS reliability and positioning accuracy enhancement based on fast satellite selection algorithm and RAIM in multi-constellation. *Aerosp Electr Syst Magazine IEEE* 30(10):14–27

4. Liu S, Zhao GR, Gao C, Zeng B (2017) A fast satellite selection algorithm for GPS/BD integrated navigation system. *Electr Opt Control* 24(3):32–35
5. Song D, Xu CD, Hu CS, Zhang PF (2015) Satellite selection with genetic algorithm under multi-constellation. *J Astronaut* 36(3):300–308
6. Huo HY, Zhang XL (2015) Fast satellite selection method for integrated navigation systems. *J Beijing Univ Aeronaut Astronaut* 41(2):273–282
7. Han TX (2014) Research on GNSS multi-system satellite selection strategy. Shanghai Jiaotong Univ 03:28–45
8. Wang L, Liu B (2008) Particle swarm optimization and scheduling algorithm. Tsinghua University Press, Beijing
9. Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization: an overview. *Swarm Intell* 1(1):33–57
10. Eberhart RC, Shi Y (2002) Particle swarm optimization: developments, applications and resources. In: Proceedings of the 2001 congress on evolutionary computation, 2001, vol 1. IEEE, pp 81–86
11. Jiang Y, Hu T, Huang CC et al (2017) An improved particle swarm optimization algorithm. *Appl Mech Mater* 193(1):231–239
12. Pan F, Li WX, Gao Q (2013) Particle swarm optimizer and multi-object optimization. Beijing Institute of Press, Beijing
13. Kuru L, Ozturk A, Kuru E et al (2015) Determination of voltage stability boundary values in electrical power systems by using the chaotic particle swarm optimization algorithm. *Int J Electr Power Energy Syst* 64(15):873–879

A Minor Fault Detection Algorithm Based on Sliding-Window Accumulated Parity Vector



Rong Wang, Zhi Xiong, Jianye Liu, Chuanyi Li and Hangshuai Ma

Abstract The receiver autonomous integrity monitoring (RAIM) technology, which with the advantage of quick response and completed independent, plays an important role in the integrity monitoring system of GNSS. With the development of BDS and other navigation satellite systems, the dual-mode and the multi-mode satellite navigation receivers are paid more and more attention, which provide not only superior positioning accuracy, but also sufficient redundancy for RAIM. Among the fault detection algorithms, the “snapshot” detection methods show satisfied rapidity for detecting large fault in pseudorange measurement, while their abilities to detect a minor fault is poor. Although minor faults usually not serious enough to cause unavailable positioning of GNSS receiver immediately, but their influence on positioning error accuracy will accumulate with time to a significant level. In order to overcome the shortcoming of the traditional RAIM algorithm in detecting minor faults, an improved minor fault detection algorithm based on sliding-window accumulated parity vector is proposed in this paper. In the proposed algorithm, a new statistics for fault detection is obtained by accumulating the parity vector in a sliding window, so as to make an optimized fault diagnosis results online. A simulation system is established for validating the proposed algorithm. The simulation results show that the proposed algorithm could shorten the time delay in detecting minor faults.

Keywords Fault detection · Parity vector · Sliding-window accumulation
Minor fault

R. Wang · Z. Xiong (✉) · J. Liu · C. Li
Navigation Research Center, College of Automation Engineering,
Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
e-mail: xiongzhi@nuaa.edu.cn

R. Wang
e-mail: rongwang@nuaa.edu.cn

H. Ma
Science and Technology on Avionics Integration Laboratory, China National
Aeronautical Radio Electronics Research Institute, Shanghai 200233, China

1 Introduction

The receiver autonomous integrity monitoring (RAIM) technology, which with the advantage of quick response and completed independent, plays an important role in the integrity monitoring system of GNSS [1, 2]. With the development of BDS and other navigation satellite systems, the dual-mode and the multi-mode satellite navigation receivers are paid more and more attention [3], which provide not only superior positioning accuracy, but also sufficient redundancy for RAIM [4].

A kind of “snapshot” detection methods for single fault detection, such as parity vector method, least square method and distance comparison method, are commonly used in multi-mode receiver. Among the fault detection algorithms, the “snapshot” detection methods show satisfied rapidity for detecting large fault in pseudorange measurement, while their abilities to detect a minor fault is poor [5, 6]. Although minor faults usually not serious enough to cause unavailable positioning of GNSS receiver immediately, but their influence on positioning error accuracy will accumulate with time to a significant level [7].

In order to overcome the limitations of traditional RAIM algorithm for small incremental minor faults in pseudo-ranges, a minor fault detection algorithm based on sliding-window accumulated parity vector is proposed for BDS/GPS dual-mode receiver. By accumulating the parity vectors of BDS/GPS measurements in a sliding window, a novel detection statistic based on the parity vectors on multiple epochs is built, which improves the sensibility to slowly varying faults. The simulation is carried out to verify the proposed algorithm.

2 Improved RAIM Scheme with On-line Sliding Window

In order to improve the sensitivity of BDS/GPS dual-mode receiver in detecting minor faults in pseudorange, the traditional parity-vector fault detection method is improved by introducing a sliding-window accumulation strategy, so as to enlarge the uniformed statistics in fault detection. The improved RAIM scheme with on-line sliding window is shown in Fig. 1.

In Fig. 1, the observational equation for BDS and GPS joint positioning is adopted in positioning calculation. On this basis, the parity vectors are calculated on each epoch, and a sliding window is constructed to accumulate them on-line and generate a novel statistics. Then the conclusion of fault detection is made by comparing the statistics with the detection threshold determined by the false alarm rate. Finally, the fault pseudo-range measurements is removed so as to obtain pollution-free BDS/GPS observations.

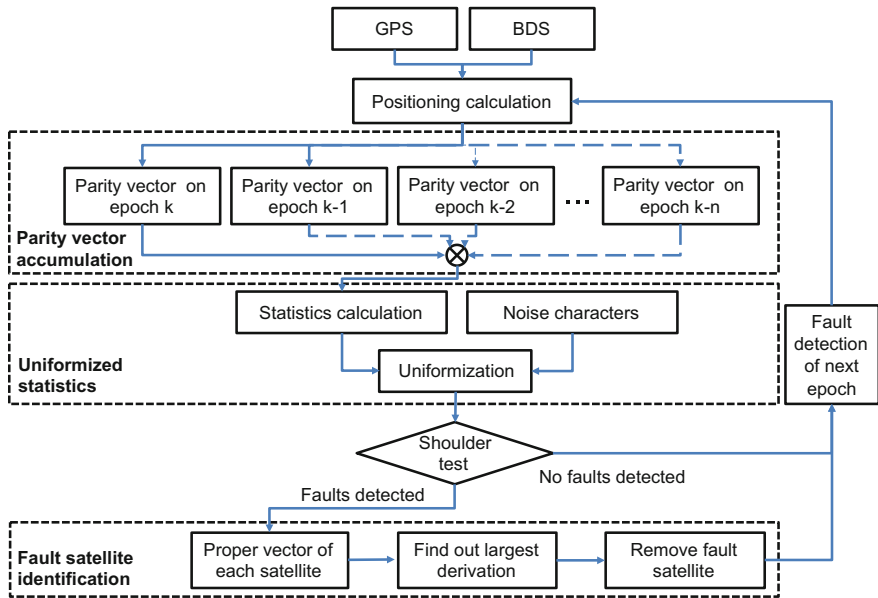


Fig. 1 Improved RAIM scheme with on-line sliding window

3 Improved RAIM Algorithm Based on Accumulated Parity Vector

Considering the correlation of minor fault’s influence between epochs, the detection performance for minor fault in pseudorange could be improved by accumulating parity vectors of multiple epochs.

3.1 Measurement Model for BDS/GPS Joint Positioning

Since the dual-mode receiver contains both BDS and GPS signal receiving units, there are two different kinds of pseudorange measurements. Here, the receiver clock error act on BDS signal receiving unit δ_b and act on GPS signal receiving unit δ_g are taken in consideration as states, as Eq. (1) shows.

$$X = [\Delta x \ \Delta y \ \Delta z \ \delta_b \ \delta_g]^T \tag{1}$$

where $\Delta x, \Delta y, \Delta z$ are the position error of receiver along the three axes of the earth-centred-earth-fixed (ECEF) coordinates.

For the situation that n satellites are observed totally, following measurement equation could be obtained:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \boldsymbol{\varepsilon} \quad (2)$$

where \mathbf{Y} is the observation qualities, which constitute by the difference between the estimated and measured pseudoranges; $\boldsymbol{\varepsilon}$ is the error of pseudorange measurements; \mathbf{H} is the measurement matrix, which reflecting the geometrical relationship between the receiver and the observed satellites. Suppose among n observed satellites, there are j BDS satellites and $n - j$ GPS satellites, then

$$\mathbf{H} = \begin{bmatrix} h_{1x} & h_{1y} & h_{1z} & 1 & 0 \\ h_{2x} & h_{2y} & h_{2z} & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{jx} & h_{jy} & h_{jz} & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{nx} & h_{ny} & h_{nz} & 0 & 1 \end{bmatrix} \quad (3)$$

The corresponding matrix elements in Eq. (3) are

$$h_{ix} = -\frac{x_i - x_u}{r_i} \quad (4)$$

$$h_{iy} = -\frac{y_i - y_u}{r_i} \quad (5)$$

$$h_{iz} = -\frac{z_i - z_u}{r_i} \quad (6)$$

where x_i, y_i, z_i are the position of the satellite i in ECEF coordinates, respectively; x_u, y_u, z_u are the position of receiver in ECEF coordinates; $r_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2}$ is the distance from the receiver position to satellite i .

In Eq. (3), the first three columns correspond to the direction cosine between the user receiver and the corresponding satellites, and the fourth and fifth columns respectively correspond to the clock skew correlation coefficients act on the BDS signal receiving unit and act on the GPS signal receiving unit.

3.2 Fault Detection Based on Parity Vector

In the traditional RAIM algorithm, a QR decomposition on the measurement matrix \mathbf{H} is made after the least square positioning is completed [8, 9]:

$$Q^T = \begin{bmatrix} Q_m \\ Q_p \end{bmatrix} \tag{7}$$

where Q_m is the first 5 rows of Q , Q_p is the last $n - 5$ rows. Then the observation space is divided to a state space with 5 dimensions and a parity spaces with $n - 5$ dimensions, then the parity vector is

$$p = Q_p \varepsilon \tag{8}$$

In the traditional RAIM algorithm, the detection statistics is constructed from the parity vectors of each epoch [10, 11]

$$fd = \frac{|p|^2}{\sigma^2} \tag{9}$$

where σ is the standard deviation of pseudorange noise.

Equation (9) is the normalized detection statistics related to the pseudorange measurement errors on single epoch, which follows a centre or off-centre Chi-square distribution with $n - 5$ degrees of freedom. The obtained detection statistics is compared with the detection threshold to determine whether there is a fault.

The probability of detecting fault successfully has a positive correlation with the off-centre degree of the detection statistics. If a large deviation occurs in pseudoranges, the off-centre degree of detection statistics also raise to a high level, so it is easy to detect large faults. If a small deviation and gradually increasing minor fault occur in pseudoranges, there may not significant changes showed in the off-centre degree, which indicates that a minor faults is difficult to be detected.

3.3 Detection Statistics from Accumulation of Parity Vectors

Under a constant fault alarm rate, the probability of detecting fault is only related to the off-centre parameter [12, 13]. Therefore, this paper constructs a sliding window to accumulate the parity vectors of N epochs and rebuilds the fault detection statistics. The accumulated parity vectors is

$$sum_p = \begin{cases} \sum_{k=1}^N p^k & T > N \\ \sum_{k=1}^T p^k & T < N \end{cases} \tag{10}$$

where T is the discretized time of current epoch.

By uniformizing the accumulated parity vectors, the new fault detection statistics is

$$fd = \frac{|sum_p|^2}{N\sigma^2} \quad (11)$$

Since the measurement noise of pseudorange follows a zero-mean Gaussian random distribution, the variance of the total noise in the cumulative time is N times of the variance of the single-epoch noise. After uniformizing the accumulated parity vectors, the constructed statistics still follows a centre or off-centre Chi-square distribution with $n - 5$ degrees of freedom. So the detection threshold under constant false alarm rate is

$$p_r(fd < sh) = \int_0^{sh} f_{\chi^2(n-5)}(x)dx = 1 - P_{FA} \quad (12)$$

where P_{FA} is the false alarm rate.

If $fd > sh$, a fault is detected in pseudoranges and then the identification of fault pseudorange is needed; else if $fd < sh$, no fault is detected and joint positioning calculation will be carried out sequentially.

4 Simulation and Analysis

4.1 Simulation Description

In order to verify the performance of dual-modal receiver RAIM algorithm based on sliding-window accumulated parity vector, this paper simulate the satellite constellation according to the BDS and GPS satellite orbit parameters. The output frequency of BDS/GPS receiver is defined as 1 Hz. The standard deviation of pseudorange noise is 20 m, and the lowest elevation angles of BDS and GPS are defined as 15° . In addition, an aircraft dynamic flight trajectory is simulated, including maneuvering actions such as acceleration climb, cruise flight and maneuver, as shown in Fig. 2. The simulated flight time is 3000 s. The number of observed satellites by BDS/GPS receivers during flight is shown in Fig. 3.

In Fig. 3, the total number of observed BDS and GPS satellites during flight is between 16 and 10. The false alarm rate is 10^{-6} and the sliding window length N is set as 8. Both hard fault and soft minor faults are injected into pseudoranges of different satellites respectively. The parameters of the injected faults are shown in Table 1.

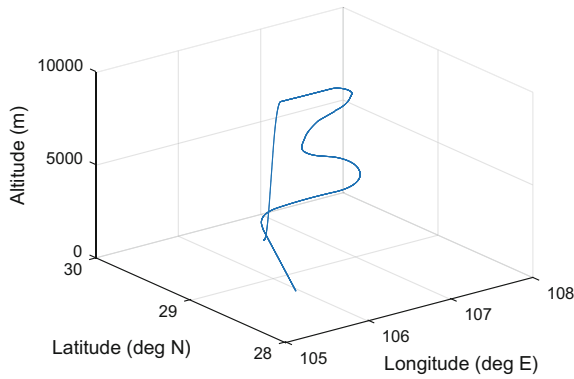


Fig. 2 Simulated flight trajectory

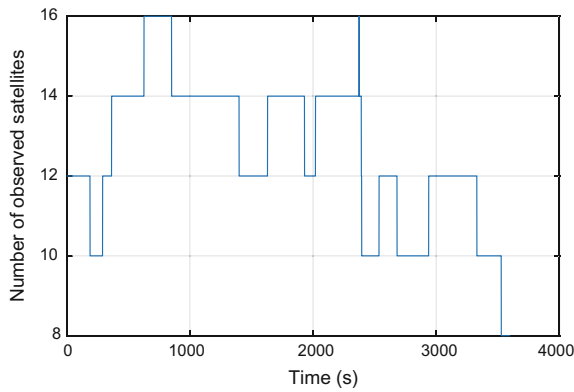


Fig. 3 Number of observed satellites during flight

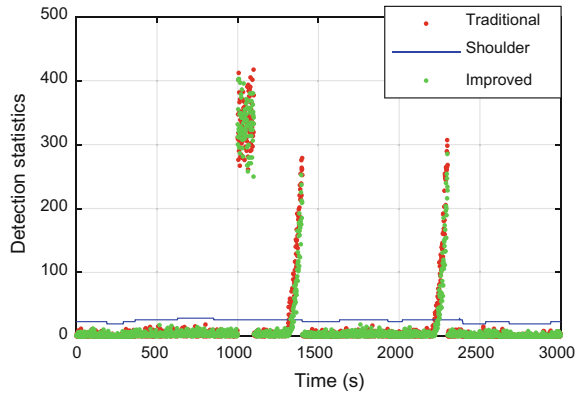
Table 1 Parameters of injected faults

Faults style	Fault satellite No.	Fault injected epochs (s)	Faults value (m)
Hard fault	13, 18	1000–1100	100
Soft minor fault	16	1300–1400	$5 + 1 * t$
Soft minor fault	3	2200–2300	$2.5 + 1.2 * t$

4.2 Simulation Results and Analysis

According to the simulation conditions set in Sect. 4.1, the traditional and the improved RAIM algorithm are compared by simulation. The curves of the fault detection statistic during the simulated flight are shown in Fig. 4.

Fig. 4 Comparison of fault detection statistics during flight



In Fig. 4, the red dot indicates fault detection statistics of the improvement of RAIM, and the green dot indicates the detection of traditional parity vector method. It can be seen that both fault detection statistics of the two methods are significantly higher than the non-fault time during the hard fault injection time. Furthermore, the fault detection statistics are compared in detail during the two soft faults occurred periods, that is from 1300 to 1400 s and from 2200 to 2300 s, as shown in Figs. 5 and 6.

In Fig. 5, the improved RAIM algorithm detects slight changes in soft faults at 1317 s, whereas the traditional algorithm delay to 1326 s to detect the same faults. As shown in Fig. 6, the improved RAIM algorithm detects soft minor faults at 2210 s, whereas the traditional algorithm detected failures late until 2224 s. Therefore, it could be concluded that the improved RAIM algorithm detects minor faults earlier before the traditional one. This indicates that the improved RAIM algorithm based on sliding-window accumulated parity vector is more sensitive to slowly changing minor soft faults than traditional algorithm, so as to remove the corresponding fault satellite earlier and improve the integrity of the dual-mode receiver.

Fig. 5 Comparison of fault detection statistics between 1300 and 1400 s

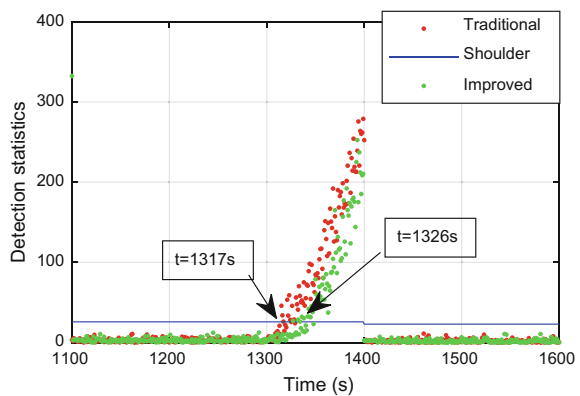
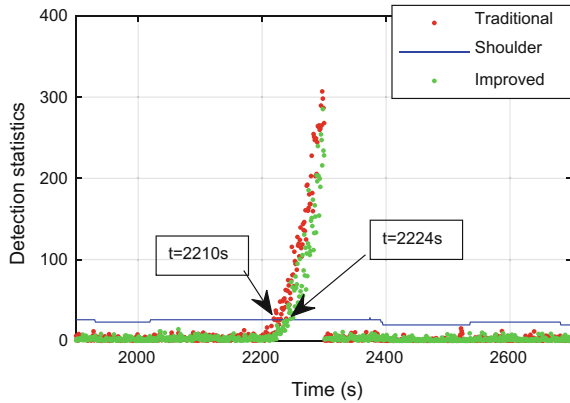


Fig. 6 Comparison of fault detection statistics between 2200 and 2300 s



5 Conclusions

In order to overcome the shortage of the traditional RAIM algorithm in detecting slowly change minor faults, a minor fault detection algorithm based on sliding-window accumulated parity vector is proposed. The simulation of the BDS/GPS dual-mode positioning during dynamic flight is carried out to validate the proposed algorithm. The simulation results indicates that the performance of the improved RAIM algorithm based on sliding-window accumulated parity vector shows superior performance than traditional RAIM algorithm based on single-epoch test, especially shortens the detection delay time of soft minor faults.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (Grant No. 61703208, 61673208, 61533008, 61533009, 61374115), Foundation Research Project of Jiangsu Province (The Natural Science Fund of Jiangsu Province, Grant No. BK20170815, BK20170767), the Aeronautic Science Foundation of China (Grant No. 2016552043,20165852052), the Fundamental Research Funds for the Central Universities (Grant No. NS2017016, NZ2016104, NP2017209, NZ2017001, NJ20170005), advanced research project of the equipment development (30102080101), the “333 project” in Jiangsu Province (Grant No. BRA2016405), the Scientific Research Foundation for the Selected Returned Overseas Chinese Scholars (Grant No. 2016), the peak of six personnel in Jiangsu Province (Grant No. 2013-JY-013), Foundation of Jiangsu Key Laboratory “Internet of Things and Control Technologies” and the Priority Academic Program Development of Jiangsu Higher Education Institutions, Science and Technology on Avionics Integration Laboratory.

References

1. Xu X, Yang C, Liu R (2013) Review and prospect of GNSS receiver autonomous integrity monitoring. *Acta Aeronaut ET Astronaut Sin* 34(3):451–463
2. Lu H, Lian J (2010) Research on RAIM algorithm of Berdou receiver. *J Air Force Eng Univ (Nat Sci Ed)* 11(3):53–57

3. Li X, Ge M, Dai X, Ren X, Fritsche M, Wickert J, Schuh H (2015) Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BDS, and Galileo. *J Geodesy* 89 (6):607–635
4. Xu L, Zhang H, Li X (2011) Study on RAIM algorithm of multi-mode navigation system. *J Time Freq* 34(02):131–138
5. Bhatti UI, Ochieng WY (2007) Integrity of an integrated GPS/INS system in the presence of slowly growing errors. Part I: a critical review. *GPS Solutions* 11(3):173–181
6. Kang CW, Park CG (2016) A soft-failure detection and identification algorithm for the integrated navigation system of lunar lander. *Proc Inst Mech Eng Part G J Aerosp Eng* 230 (11):2023–2035
7. Zhou Y, Wang Y, Dou S, Meng Z (2016) Innovative soft fault diagnosis method for dual-redundancy sensors. *Sens Rev* 36(1):14–22
8. Li Y, Li M (2012) Study on RAIM algorithm of GPS receiver based on parity vector. *Geomatics Spat Inf Technol* 35(1):158–160
9. Song K, Fan S, Liu J (2016) An optimized RAIM algorithm based on kalman filter and parity vector. *Navig Control* 15(6):101–106
10. Han W, Wang Z, Shen Y (2017) Fault estimation for a quadrotor unmanned aerial vehicle by integrating the parity space approach with recursive least squares. *Proc Inst Mech Eng Part G J Aerosp Eng* 0954410017691794
11. Peng X, Huang G, Guo C (2011) The research on failure detection of parity vector RAIM algorithm. *J Projectiles, Rockets, Missiles Guidance* 31(6):199–202
12. Liu W, Li Z, Wang F (2010) A new RAIM method for detecting and correcting weak pseudo-range bias under gradual change. *J Astronaut* 31(04):1024–1029
13. Huang G, Chang C, Hao S (2013) Weight correction cumulative epoch RAIM algorithm based on adaptive strategy. *J Comput Appl* 33(08):2366–2369

The Construction Method of Ionospheric Delay Model with Combined BDS Phase and Range Observation



Zongpeng Pan, Hongzhou Chai, Yulong Kong, Rui Wang and Chunhe Liu

Abstract The ionospheric delay model is a reliable method to reduce the influence of ionospheric delay. The phase smoothed pseudorange observations are always used to construct ionospheric delay model. However, the accuracy of the constructed ionosphere model is affected by precision of phase smoothed pseudorange observations and the number of base stations. This contribution presents a construction method of ionospheric delay model with combined BDS phase and range observation. The observations of ionospheric delay is constructed by precise point positioning (PPP) with original observations. At the same time, the polynomial model is used to ionosphere modeling. Finally, the validity of this method is verified using BDS and GPS data. The bias of ionosphere model is less than 1 m and the standard deviation is better than 0.4 m.

Keywords Ionospheric delay · BDS · Precise point positioning
Original observations

1 Introduction

The ionosphere is an important part of the Earth's atmosphere. The time delay when the satellite navigation signal passes the ionosphere, that is, the ionospheric delay. Ionospheric delay has always been one of the main error sources that affect the accuracy of GNSS navigation and positioning. Ionospheric delay correction must be performed in navigation and positioning [1]. Especially for single-frequency users, the accuracy of ionospheric delay correction directly affects the positioning accuracy of pseudo-range single-point positioning and single-frequency PPP [2–6]. The ionospheric delay correction model is an effective method to weaken the

Z. Pan (✉) · H. Chai · Y. Kong · R. Wang · C. Liu
Institute of Surveying and Mapping, Information Engineering University,
Zhengzhou 450001, China
e-mail: panzongpeng@yeah.net

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_61

influence of ionospheric delay [7–13], such as the Klobuchar model, the NeQuick model, the global ionosphere model (GIM) and so on.

Ionospheric delay correction models are usually constructed using a non-geometric combination of pseudo-range observations (P4) smoothed by carrier-phase. However, the accuracy of phase smoothed pseudo-range observations is limited by smoothing arc. On the other hand, when the carrier measurements occurs cycle slip or data interruption, it will lead to many smooth segments, which will affect the accuracy of smoothed observation. Therefore, the accuracy of the ionospheric model obtained by this method is influenced by the accuracy of the phase smoothed pseudo-range observations.

This contribution presents a construction method of ionospheric delay model with combined phase and range observation. Different from the phase smoothed pseudo-range observation, the ionospheric delay observations are constructed by precise point positioning (PPP) with original phase and pseudo-range observations. The precise ephemeris correction and known station coordinates are constrained to reduce the parameters in PPP model [14, 15]. Then, the ionospheric delay observations extracted at each station are modeled using a polynomial function. Compared with the phase smoothed pseudo-range observation, this method can directly use high-precision carrier phase observation. When the PPP is initialized for a short period of time, the carrier phase observation in the PPP model plays a major role, and then the high-precision ionospheric delay can be obtained. On the other hand, this method can ensure the rapid convergence of the ionosphere delay parameters when a cycle slip occurs, and does not have a significant impact on the accuracy of the estimated ionospheric delay. The second part of this paper is the ionospheric delay estimation model and the ionospheric model construction method, the third part is the experimental analysis. Finally, the conclusion of this paper is given.

2 Mathematical Model

2.1 Ionospheric Delay Estimation Model

The ionospheric delay observations are obtained by PPP with original phase and pseudo-range observations. GNSS original pseudo-range and carrier phase observation equations can be expressed as,

$$P_{i,r}^s = \rho_r^s + \delta t_r - \delta t^s + T_r^s + \gamma_i I_{1,r}^s + b_{i,r}^s - b_i^s + \varepsilon_{i,r}^s \quad (1)$$

$$L_{i,r}^s = \rho_r^s + \delta t_r - \delta t^s + T_r^s - \gamma_i I_{1,r}^s + \lambda_i^s N_{i,r}^s + \delta_{i,r}^s \quad (2)$$

where, $L_{i,r}^s$ and $P_{i,r}^s$ are the original carrier phase and pseudo-range observations, in meters; the subscript r for station number, i for the frequency number, superscript

s for any satellite, $\rho_r^{s,j}$ is geometric component which include geometric distance between satellite and receiver, δt_r and δt^s for clock error of receiver and satellite, respectively, T_r^s is troposphere delay, $\gamma_i = f_1^2/f_i^2$, $I_{1,r}^s$ is ionosphere path delay on the frequency L1. $b_{i,r}^s$, b_i^s are hardware delays in code observations of receiver and satellite. $N_{i,r}^s$ is phase ambiguity which include initial phase bias. $\varepsilon_{i,r}^s$, $\delta_{i,r}^s$ are carrier phase and code measurement noise and other un-modeled errors, λ_i^s is the wavelength of frequency i .

When PPP is used to estimate the ionospheric delay, the precise ephemeris correction and known station coordinates are constrained to reduce the parameters in PPP model. At the same time, the error corrections such as satellite and receiver antenna phase center offsets (PCO) and variations (PCVs), relativistic corrections, phase windup, a priori model tropospheric correction and earth tides etc. must be considered. The satellite clock correction provided by IGS analysis centers contains the hardware delay of the satellites. When satellite clock correction is replaced in Eqs. (1) and (2), the PPP model can be expressed as

$$P_{i,r}^s = \rho_r^s + \delta \tilde{t}_r - \delta \tilde{t}^s + T_r^s + \gamma_i \tilde{I}_{1,r}^s + \varepsilon_{i,r}^s \tag{3}$$

$$L_{i,r}^s = \rho_r^s + \delta \tilde{t}_r - \delta \tilde{t}^s + T_r^s - \gamma_i \tilde{I}_{1,r}^s + \lambda_i^s \tilde{N}_{i,r}^s + \delta_{i,r}^s \tag{4}$$

Among them, the parameter of receiver clock, the ionospheric delay and the ambiguity are shown as

$$\begin{aligned} \delta \tilde{t}_r &= \delta t_r + b_{IF,r}; \delta \tilde{t}^{s,j} = \delta t^s + b_{IF}^s \\ \tilde{I}_{1,r}^s &= I_{1,r}^s + (b_{IF}^s - b_1^s) - (b_{IF,r} - b_{1,r}) \\ \lambda_i^s \tilde{N}_{i,r}^s &= (1 + \gamma_i)(b_{IF}^s - b_{IF,r}) - \gamma_i(b_1^s - b_{1,r}) + \lambda_i^s N_{i,r}^s \end{aligned} \tag{5}$$

where, the subscript IF is ionosphere-free combination. The hardware delay in the above ionospheric delay parameter can be simplified. In this case, the estimated ionospheric delay is,

$$\tilde{I}_{1,r}^s = I_{1,r}^s + \beta(DCB_r - DCB^s) \quad \beta = f_2^2/(f_1^2 - f_2^2) \tag{6}$$

where, $DCB_r = b_{2,r} - b_{1,r}$ and $DCB^s = b_2^s - b_1^s$, for the receiver and satellite differential code deviation, respectively. As can be seen from the Eq. (6), the ionospheric delay estimated in PPP moded is consistent with that of the traditional carrier phase smoothed pseudo-range. According to Eqs. (3)–(6), we can obtain the ionospheric delay contained hardware delay with Kalman Filter. The processing method of the parameters is consistent with the estimation of PPP, which can be referred to [5, 6] and will not be repeated here.

2.2 Ionospheric Model Construction Method

Ionospheric model can be constructed by taking the above estimated ionospheric delay as a virtual observation. Based on the assumption that the ionosphere can be regarded as a single layer, we can choose different mathematical functions to describe the ionospheric variation according to different temporal and spatial scales (global, regional or local). Commonly used ionospheric model includes the spherical harmonic function model, the generalized triangular series and the polynomial function model [7–13]. At present, the service area of BeiDou satellite navigation system (BDS) is mainly in the Asia-Pacific region. Therefore, the experiments select the data from the regional stations, and the polynomial function model is applied to ionospheric modeling. Then, Eq. (6) can be further expressed as

$$\tilde{I}_{1,r}^s = MF(z) * \sum_{j=0}^n \sum_{k=0}^m A_{jk} (\varphi - \varphi_0)^j (S - S_0)^k + \beta(DCB_r - DCB^s) \quad (7)$$

In the formula, $MF(z)$ is the mapping function at ionosphere pierce point (IPP), z is the azimuth of IPP, $S - S_0 = (\lambda - \lambda_0) + (t - t_0)$, A_{jk} is the polynomial model coefficient, λ is the longitude of the IPP, t is the local time, S_0 is sun hour angle for the center of the measuring zone (φ_0, λ_0) at noon t_0 , φ_0 is the latitude for the geographical center of the area.

In the above equation, there is a linear correlation between the satellite and the receiver DCB. The DCB parameters of the satellite and receiver cannot be directly separated. Therefore, the constraints should also be added to separate the DCB of the satellite and receiver. In this paper, the sum of DCB of all satellites is set to zero as a constraint [16, 17]. Then, the polynomial model coefficients and all satellite and receiver DCBs are obtained by least squares adjustment.

3 Experiments and Results Analysis

3.1 Data and Processing Strategies

In order to verify the proposed method, the experiment was conducted using the observed data from 14 Australian CORS stations of 2016, DOY 250-254, and each station can observe BDS/GPS data, data sampling interval is 30 s, satellite cutoff elevation set to 10° . The distribution of stations is shown in Fig. 1.

In the experiment, the PPP with original BDS and GPS observations is used to estimate the ionospheric delay, and then the ionosphere model constructed by the proposed method are compared and analyzed. In the data processing, precision satellite orbit and clock error correction is provided by GFZ multi-system precision ephemeris. The station coordinates are constrained to CORS network adjustment results. Also, all types of PPP error correction must be taken into account. The

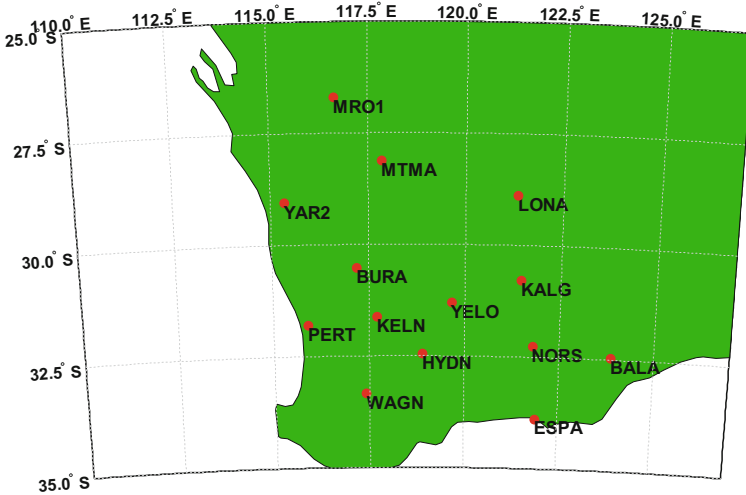


Fig. 1 Station distribution

ionospheric delay observations extracted by PPP at each station can be used to construct ionospheric model. A low-order polynomial function is selected to separate satellites and receivers DCB in ionospheric delay, due to the selected stations in smaller area. In this paper, we use a first order polynomial model and solve a set of polynomial model coefficients every two hours, while a group of the DCB parameters are estimated in one day. Firstly, the stability of multi-day DCB values is analyzed. Secondly, the accuracy of the constructed ionospheric polynomial model is verified. The constructed ionospheric model is compared with the GIM released by CODE [17].

3.2 DCB Stability Analysis

Generally, satellites and receivers DCB have good stability. Therefore, analyzing the stability of receivers and satellites DCB obtained in the above experiment can verify the feasibility of the proposed method. It also reflects the accuracy of ionospheric polynomial model.

Figures 2 and 3 shows the standard deviation (std) of the multi-day BDS satellites DCB values and GPS satellites DCB values, respectively.

As can be seen from Fig. 2, the std of all BDS satellites DCB are less than 0.6 ns. The stability of DCB of different orbit type is different. The DCB of IGSO satellites have the smallest std, which are better than 0.2 ns. The std of the GEO satellites DCB and the MEO satellites DCB are worse than IGSO satellites DCB. This is mainly due to the number of BDS MEO satellites and the limited observable

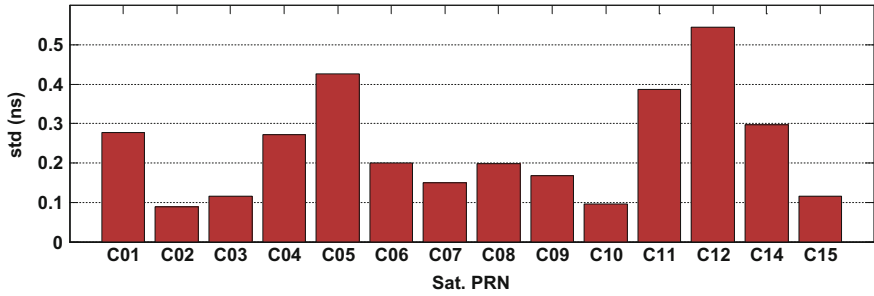


Fig. 2 The std of the BDS satellites DCB for 5 days

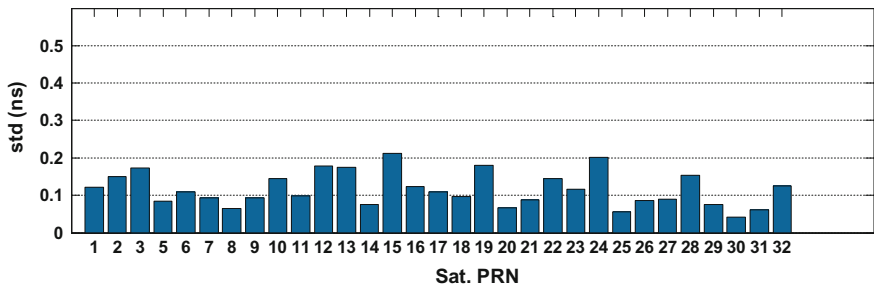


Fig. 3 The std of the GPS satellites DCB for 5 days

period of MEO satellites in a day, so its DCB stability is a bit poor. Although the GEO has a longer observable period, it basically does not move. On the other hand, the precision of the orbit and clock error correction of the BDS GEO satellites is currently poor, which is also the reason for the poor stability of the GEO DCB.

As can be seen from Fig. 3, the std of all GPS satellites DCB are less than 0.2 ns. The stability of most satellites DCB is better than 0.1 ns. The stability of GPS satellites DCB is obviously better than that of BDS. The reason may be the accuracy of GPS ephemeris is better than BDS ones and the errors of GPS PPP can be effectively corrected. However, the antenna phase center correction of BDS satellite and receiver has not accurate correction model [5, 18].

Figure 4 gives a stability analysis of the receiver DCB, in which the red bar indicates the std of the receiver DCB estimated using BDS data and blue bar indicates the std of the receiver DCB estimated using GPS data.

Figure 4 shows that, the std of receiver DCB using GPS data is significantly better than that of receiver DCB using BDS data estimation. The std of receiver DCB estimated using BDS data is better than 0.8 ns. The std of receiver DCB estimated by GPS data is better than 0.4 ns. Comparing Figs. 2 and 3 with Fig. 4, we can see that the stability of satellite DCB is better than the stability of receiver DCB.

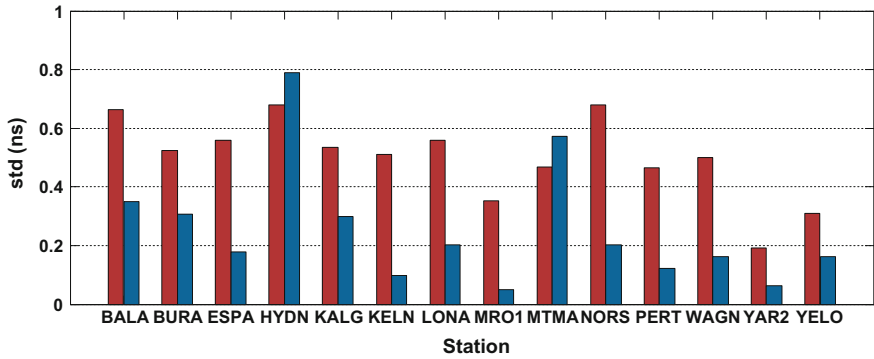


Fig. 4 The std of the receiver DCB for BDS (red) and GPS (blue)

3.3 Accuracy Analysis of Ionospheric Model

In order to analyze the accuracy of the constructed ionospheric polynomial model, it is compared with the GIM ionospheric model released by CODE. The data of CUT0 station, which is not involved in ionospheric modeling, is selected to verify the accuracy of ionospheric polynomial model. The ionospheric delay at the IPP is obtained by using polynomial model and GIM, respectively. The time series bias of BDS and GPS ionospheric polynomial model with respect to GIM are shown in Figs. 5 and 6, respectively.

By analyzing the time series of bias in Figs. 5 and 6, it can be seen that the constructed BDS and GPS ionospheric models using the proposed method agree well with the CODE GIM. The biases between them are small and less than 1 m within a day. The accuracy of the constructed BDS ionospheric delay model is the same as that of the GPS ionospheric delay model. However, the ionospheric delay between the constructed ionospheric model and the GIM shows a system bias of

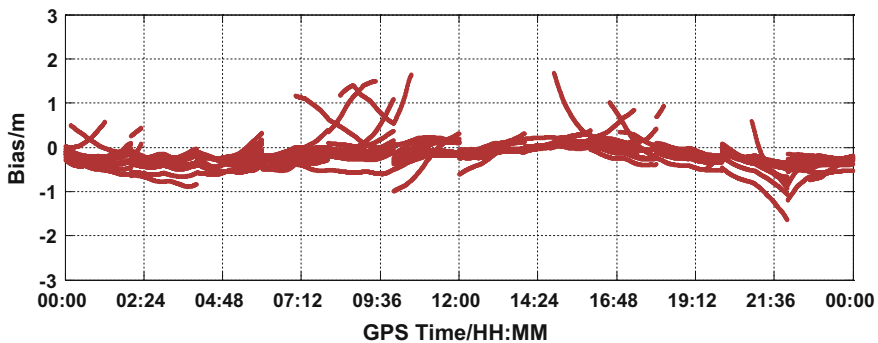


Fig. 5 The bias of BDS ionospheric polynomial model with respect to GIM

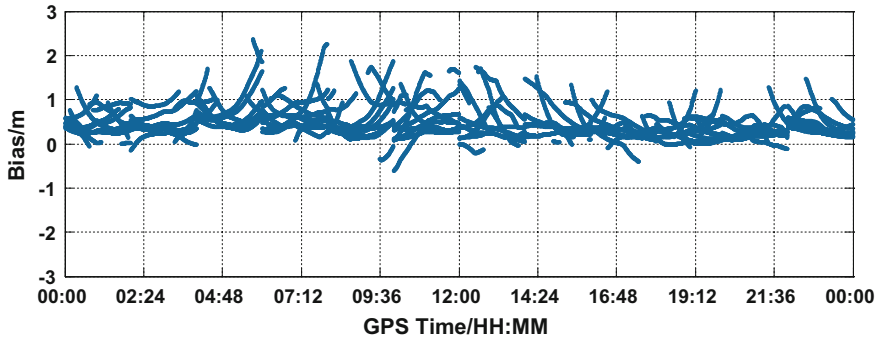


Fig. 6 The bias of GPS ionospheric polynomial model with respect to GIM

decimeter level, as can be clearly seen from Fig. 7. It can be seen from Fig. 7 that ionospheric model bias are less than 1 m, but the center of the bias distribution is not at zero, indicating that there is systematic bias in the ionospheric model.

Table 1 shows the statistics of ionospheric model bias for all days, and gives four indicators as mean, std, maximum and minimum respectively. It can be seen from the Table 1 that the mean value of ionospheric model bias between BDS and GPS is not zero, and there is a systematic bias of 0.3–0.6 m. The reason may be related to the accuracy of receiver DCB and the DCB bias will be reflected in the ionospheric model.

However, from the statistics std of ionospheric model bias, we can see that the constructed BDS and GPS ionospheric model are agree with GIM. The std of model bias is less than 0.4 ns, and the accuracy of BDS ionosphere model is better than that of the GPS ionosphere model. At the same time, the maximum and the minimum ionospheric delay bias are between ± 3 m. This phenomenon appears at the local noon, which the ionosphere changes more actively. Through the above analysis shows that the method proposed in this paper can effectively constructed high-precision ionospheric model.

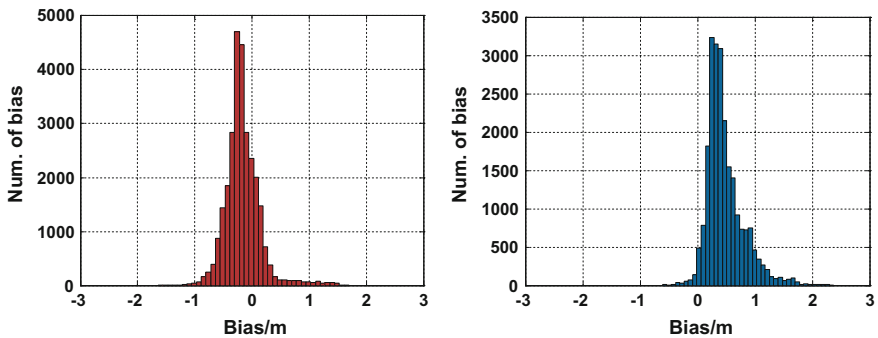


Fig. 7 Frequency distribution of bias for BDS and GPS ionospheric polynomial model

Table. 1 Statistics of ionospheric model bias

DOY	BDS				GPS			
	Mean	std	Max	Min	Mean	std	Max	Min
250	-0.164	0.316	1.695	-1.638	0.480	0.329	2.371	-0.608
251	0.373	0.267	2.169	-0.217	0.408	0.308	3.111	-0.424
252	0.610	0.331	2.609	-0.145	0.479	0.365	2.344	-0.592
253	0.463	0.3427	3.189	-0.268	0.442	0.349	2.617	-0.777
254	0.564	0.359	3.697	-0.198	0.428	0.387	3.147	-0.697

Unit m

4 Conclusion

In this paper, a construction method of ionospheric delay model with combined phase and range observation is proposed. Different from the phase smoothed pseudo-range observation, the ionospheric delay observations are constructed by precise point positioning (PPP) with original phase and pseudo-range observations. At the same time, the polynomial model is used to ionosphere modeling. The BDS and GPS observations in Australia for five days are used to verify the proposed method. Firstly, the std of all BDS satellites DCB are less than 0.6 ns. The stability of DCB of different orbit type is different. The std of all GPS satellites DCB are less than 0.2 ns. Secondly, the std of receiver DCB using GPS data is significantly better than that of receiver DCB using BDS data. The std of receiver DCB estimated using BDS data is better than 0.8 ns. The std of receiver DCB estimated by GPS data is better than 0.4 ns. Finally, compared with the GIM ionospheric model released by CODE, the bias of ionosphere model is less than 1 m and the STD is better than 0.4 m, which verifies the effectiveness of the proposed method.

Acknowledgements Thanks to the IGS MGEX and Australia CORS for providing GNSS observations. This work was supported by National Natural Science Foundation of China (Grant No. 41274045, No. 41574010) and State Key Laboratory of Geo-information engineering (No. SKLGIE2015-Z-1-1).

References

1. Leick A, Rapoport L, Tatarnikov D (2014) GPS satellite surveying, 4th edn
2. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. *GPS Solutions* 5(2):12–28
3. Bisnath S, Gao Y (2009) Current state of precise point positioning and future prospects and limitations observing our changing earth. Springer, Berlin, pp 615–623
4. Kouba J (2009) A guide to using International GNSS Service (IGS) products. *Maryland Biol Stream Surv Data Versar Inc* 4(3):106
5. Pan Z, Chai H, Kong Y (2017) Integrating multi-GNSS to improve the performance of precise point positioning. *Adv Space Res*

6. Pan Z, Chai H, Kong Y (2017) Integrating multi-GNSS to improve the performance of precise point positioning. *Adv Space Res*
7. Li Z, Ou J, Huo X (2012) Two-step method for the determination of the differential code biases of COMPASS satellites. *J Geodesy* 86(11):1059–1076
8. Hernández-Pajares M, Juan JM, Sanz J et al (2009) The IGS VTEC maps: a reliable source of ionospheric information since 1998. *J Geodesy* 83(3–4):263–275
9. Zhang R, Yao YB, Hu YM et al (2017) A two-step ionospheric modeling algorithm considering the impact of GLONASS pseudo-range inter-channel biases. *J Geodesy* 2:1–12
10. Hoque MM, Jakowski N, Berdermann J (2017) Ionospheric correction using NTCM driven by GPS Klobuchar coefficients for GNSS applications. *GPS Solutions* 3–4:1–10
11. Zhang R, Song WW, Yao YB et al (2015) Modeling regional ionospheric delay with ground-based BeiDou and GPS observations in China. *GPS Solutions* 19(4):649–658
12. Li Z, Yuan Y, Wang N et al (2015) SHPTS: towards a new method for generating precise global ionospheric TEC map based on spherical harmonic and generalized trigonometric series functions. *J Geodesy* 89(4):331–345
13. Hoque MM, Jakowski N (2015) An alternative ionospheric correction model for global navigation satellite systems. *J Geodesy* 89(4):391–406
14. Tu R, Zhang H, Ge M et al (2013) A real-time ionospheric model based on GNSS precise point positioning. *Adv Space Res* 52(6):1125–1134
15. Yao Y, Zhang R, Song W et al (2013) An improved approach to model regional ionosphere and accelerate convergence for precise point positioning. *Adv Space Res* 52(8):1406–1415
16. Wu X, Zhou J, Tang B et al (2014) Evaluation of COMPASS ionospheric grid. *GPS Solutions* 18(4):639–649
17. Hernández-Pajares M, Roma-Dollase D, Krankowski A et al (2017) Methodology and consistency of slant and vertical assessments for ionospheric electron content models. *J Geodesy* 2:1–10
18. Lou Y, Liu Y, Shi C et al (2016) Precise orbit determination of BeiDou constellation: method comparison. *GPS Solutions* 20(2):259–268

Improved RAIM Algorithm Based on Kalman Innovation Monitoring Method



Zhengen Yang, Huaijian Li and Xiaojing Du

Abstract Integrity monitoring is an important means to guarantee the integrity of satellite navigation system. Receiver autonomous integrity monitoring (RAIM), as a client integrity monitoring method, has a lot of advantages such as not dependent on external equipment, low cost and easy to implement. Therefore, it is widely used in integrity monitoring. Traditional RAIM methods comprise RAIM algorithm based on Kalman filter and snapshot algorithm based on pseudorange observation. Compared with the snapshot algorithm, the Kalman filter innovation monitoring method is not limited to use the current measurement. Therefore it has advantages of independent detection and less calculation. It also can be used under the condition of few satellites. Unfortunately, the Kalman filter-based method is not sensitive to slowly varying pseudorange fault. Thus we propose an integrated algorithm combining the parity vector method based on the non-coherent accumulation with Kalman filter-based detection method. The result shows that, compared with the traditional parity vector method and the Kalman filter-based detection method, the proposed algorithm has a better result in fault detection and monitoring delay.

Keywords Receiver autonomous integrity monitoring · Kalman filter
Parity vector method

1 Introduction

Integrity is a function to notify the users in time or terminate the signal when the error of navigation system exceeds the alert limits and makes the system unavailable [1]. With the development of the global navigation satellite system (GNSS), the requirements for the stability are becoming more and more stringent. Integrity monitoring is becoming more and more important. The traditional integrity moni-

Z. Yang · H. Li · X. Du (✉)

Beijing Institute of Technology, University of Beijing, Beijing, China
e-mail: duxiaojing2633@163.com

© Springer Nature Singapore Pte Ltd. 2018
J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018 Proceedings*, Lecture Notes in Electrical Engineering 497,
https://doi.org/10.1007/978-981-13-0005-9_62

759

toring algorithm is gradually difficult to meet the needs, so it is particularly important to improve the traditional algorithm and its performance. The traditional integrity monitoring algorithms include the integrity monitoring based on the system itself and the receiver autonomous integrity monitoring (RAIM). System based integrity monitoring cannot guarantee the time to alert to fit the needs of high reliability in some cases, so it is not widely used [2]. While RAIM algorithm uses the internal redundancy information of the receiver to detect and identify faults with the characteristics of not relying on external equipment, low cost and easy to implement, making it a widely used integrity monitoring algorithm.

RAIM algorithm is divided into the snapshot algorithm based on the pseudorange residual and the algorithm based on the Kalman filter. As a typical algorithm of snapshot algorithm, parity vector method is simple to calculate and has shown good detection performance for large pseudo range deviation. And its improved algorithm is sensitive for small and slowly-changing pseudo range deviation [3]. Compared with the snapshot algorithm, RAIM algorithm based on Kalman filter overcomes the restrictions of satellite quantity requirement and can effectively complete the independent detection. But it is not sensitive to the slow variable fault detection because of the memory effect of the filter.

This paper analyzes the advantages and disadvantages of the two kinds of RAIM algorithms. We realize an improved algorithm of RAIM by combining the two basic algorithms which can detect faults when the satellites visible are not enough while reduce the possibility of missing detection of small and slowly-changing faults at the same time. The above detection performance is proved in the simulation experiment.

2 Snapshot Algorithm

2.1 Usability Analysis

To meet the needs of fault detection, visible satellite quantity and satellite geometry at each time must be examined before the detection, that is, the usability judgement of RAIM.

At present, the most common RAIM availability judgement methods are ARP method (approximate radial error protection method), δH_{max} method and HPL analysis. The above three methods have been proved to be equivalent [4]. In this paper, the HPL analysis method is selected.

When the detection start, the detection statistic SSE/σ_0 will be compared with the detection threshold T^2 . If the detection statistic is below the threshold, there is a missed detection. The missed detection rate P_{MD} is expressed as:

$$P_{MD} = P(T|(n - 4), \lambda) \tag{2.1}$$

where P_{MD} should satisfy the following probability formula:

$$P_r(SSE/\sigma_0^2 < T^2) = \int_0^{T^2} f_{\chi^2(n-4)}(x)dx = P_{MD} \tag{2.2}$$

Given the false alarm rate P_{FA} and the missed detection rate P_{MD} , the non centralization parameters λ can be calculated.

The error protection level PL is calculated to guarantee the false alarm rate and the missed detection rate of RAIM, as well as the quality of the satellite geometry.

Where the formulas for calculating the horizontal error protection level and the vertical error protection level are:

$$HPL = \delta HDOP_{max} \times \sigma_0 \times \sqrt{\lambda} \tag{2.3}$$

$$VPL = \delta VDOP_{max} \times \sigma_0 \times \sqrt{\lambda} \tag{2.4}$$

The maximum of all such $\delta HDOPs$ is the maximum horizontal dilution of precision factor, which is calculated before the RAIM fault detection. The calculation formula is:

$$\delta HDOP_{max} = \max \left(\sqrt{\frac{A_{1i}^{*2} + A_{2i}^{*2}}{Q_{vii}}} \right) = \sqrt{HDOP_i^2 - HDOP^2} \tag{2.5}$$

If $\delta HDOP_{max}$ is below a certain threshold of the horizontal precision factor, it means in the worst case with a difficult fault to detect, it also can ensure that the missed detection rate is below P_{MD} .

The maximum vertical dilution of precision factor $\delta VDOP_{max}$ is:

$$\delta VDOP_{max} = \max \left(\sqrt{\frac{A_{3i}^{*2}}{Q_{vii}}} \right) = \sqrt{VDOP_i^2 - VDOP^2} \tag{2.6}$$

where $A^* = (A^T A)^{-1} A^T$; A_{ji}^* ($j = 1, 2, 3$) equals the j th line of A^* , $HDOP_i$ and $VDOP_i$ are $HDOP$ and $VDOP$ factors after removing the i th satellite. Q_v represents the inverse matrix of the pseudorange residual vector.

The purpose of calculating the protection limit value PL is to evaluate the reliability of the positioning results, which depends on the accuracy estimation of each pseudorandom measurement and the geometric distribution of the satellites.

Compare the PL with the alarm limit AL. When PL is below AL, the integrity of the fault can be guaranteed, and the RAIM algorithm is available.

2.2 Parity Vector Method

The snapshot algorithm includes the parity vector method, the pseudorange comparison method and the least square residual method. The three methods have been proved to be equivalent [5]. Parity vector method is the most widely used method and recommended by the Radio Technical Commission for Aeronautics (RTCA) SC-159 group.

Linear measurement equation of the system can be expressed as:

$$Y = GX + \varepsilon \quad (2.7)$$

where, Y denotes the measurement vector. X is the state vector and G denotes the transition matrix of them. ε is the Gaussian white noise in measurement.

If there is no measurement error, the formula (2.7) can be written as:

$$y = Gx \quad (2.8)$$

G is a $n \times 4$ dimensional matrix. The QR decomposition of G can be written as:

$$G = QR \quad (2.9)$$

Left multiplicative matrix Q^T on both sides:

$$Q^T y = Rx, Q^T = \begin{bmatrix} Q_x \\ Q_p \end{bmatrix}, R = \begin{bmatrix} R_x \\ 0 \end{bmatrix} \quad (2.10)$$

where, Q_x is the first 4 lines of Q^T , Q_p is the rest part of it, and the R_x is the first 4 lines of R .

We can obtain from formula (2.10) that:

$$\hat{x} = R_x^{-1} Q_x y \quad (2.11)$$

Considering the effect of observation error, where $Y = GX + \varepsilon$, the parity vector can be written as:

$$p = Q_p \varepsilon \quad (2.12)$$

where Q_p is the parity space matrix, vector p is projected by parity space matrix. It can directly reflect the bias information of fault satellites [6].

Suppose $F_{SSE} = p^T p$ and the pseudorange error is σ . When there is no fault in the system, the statistics F_{SSE}/σ^2 will obey the distribution of $\chi^2(n-4, 0)$. When there is a fault, it will obey the non centralization distribution of $\chi^2(n-4, \lambda)$, where λ represents the non centralization parameter.

Assuming a false alarm rate P_{FA} , according to the above analysis, the probability of the statistics F_{SSE}/σ^2 below the detection threshold is:

$$P_r(SSE/\sigma^2 < a) = \int_0^a f_{\chi^2(n-4)}(t)dt = 1 - P_{FA} \tag{2.13}$$

The statistics T is:

$$T = \sqrt{F_{SSE}/(n - 4)} \tag{2.14}$$

The detection threshold is:

$$T_D = \sigma\sqrt{a/(n - 4)} \tag{2.15}$$

Compare the statistics T with the threshold T_D , if T is below T_D , there is no fault detected. On the other hand, there is a fault.

If the i th satellite has a fault and the error is b_i , then the non centralization parameter can be expressed as [7]:

$$\lambda = E(P_{SSE})/\sigma^2 = b_i^2 S_{ii}/\sigma^2 \tag{2.16}$$

where S_{ii} is the i th diagonal term of $S = I + H(H^T H)^{-1} H^T$.

2.3 Parity Vector Method Based on Non-coherent Accumulation

The accumulating parity vector method can be used to solve the problem of poor detection performance in small and slow-changing fault detection. Under the same request of false alarm rate, the fault detection rate of RAIM algorithm is affected by the non central parameter, while the central parameter is influenced by the normalized range deviation. By accumulating detection statistics, we can increase the amplitude of normalized range deviation, improving the fault detection rate. However, direct coherent accumulation needs known deviation sign, so the parity vector method based on non-coherent accumulation is adopted [3].

As P^k is the parity vector at epoch k moment, we can structure the non-coherent detection statistics as:

$$F = \sum_{k=1}^n \|P^k\|_2^2 \tag{2.17}$$

where, n is the cumulate number of epochs.

Under the condition of no fault, the non-coherent accumulation statistics is normalized and the normalized detection statistics will satisfy the distribution of χ^2 .

When there is a fault, it will obey the distribution of $\chi^2(n - 4, \lambda)$. Because the satellite constellation changes a little during short time, Q_p can be considered a constant during the n epoch, and the parameter of non centralization can be expressed as:

$$\lambda = \sum_{k=1}^n \sum_{i=1}^{n-4} \left(\frac{Q_p b}{\sigma^2} \right) \tag{2.18}$$

3 Kalman Innovation Monitoring Method

3.1 Kalman Filter-Based RAIM Algorithm

The linearized equation of discrete Kalman filter is modeled as [8]:

$$\begin{aligned} X_k &= \Phi_{k/k-1} X_{k-1} + W_k \\ Z_k &= H_k X_k + V_k \end{aligned} \tag{3.1}$$

X_k is the filter state vector at k moment with parameters of position, acceleration, clock bias and clock frequency bias. $\Phi_{k/k-1}$ is the state transition matrix. W_k is the system noise matrix. Z_k is the measurement vector at k moment. H_k is the measurement matrix. V_k is the measurement noise covariance matrix.

It is conditional to keep the characteristic of zero bias during the estimation in the Kalman filter Algorithm. One of the necessary conditions is the measurement error must show the characteristic of zero-mean Gaussian white noise. When the fault exists, the characteristic of zero-mean in measurement error will be damaged as well as the state estimation. We present two opposing hypotheses to detect faults:

H_0 no fault exists in the system, $E(V_k) = 0$.

H_1 faults exist in the system, $E(V_k) = b$ (b is the fault bias).

Construct T as the statistical test quantity of fault, and compare it with certain threshold. We can determine whether there is a fault in the system.

$T \geq T_d$ then hypothesis H_0 is correct.

$T < T_d$ then hypothesis H_1 is correct.

In practical application, the measurement noise is a random variable, and the real position of user is unknown. So we cannot get the bias of the state estimation. Therefore, the key is that how to extract the test statistic in the Kalman filter which reflects the noise bias and the estimation bias.

When there is no fault in system, it is known from the optimal valuation theory that the result of the filter is unbiased estimation. Hence:

$$\begin{aligned}
 E\{\tilde{X}\} &= E\{X - \tilde{X}(Z)\} = E(HX_k + V_k - H\hat{X}_{k/k-1}) \\
 &= H\Phi_{k/k-1}E(X_{k-1}) + HE(W_k) + E(V_k) = 0
 \end{aligned}
 \tag{3.2}$$

where the residual of Kalman filter is also found unbiased:

$$E(\Delta Z_k) = E(Z_k - \hat{Z}_k)
 \tag{3.3}$$

If there is a fault at k moment, the expectation of residual can be written as:

$$E(\Delta Z_k) = H\Phi_{k/k-1}E(X_{k-1}) + HE(W_k) + E(V_k + \mathbf{b}) = \mathbf{b}
 \tag{3.4}$$

Therefore, the system fault variation is shown in the statistical property of Kalman filter. ΔZ of the i th satellite is selected as the test statistic. It can be expressed as:

$$T = \Delta Z_{k,i}
 \tag{3.5}$$

3.2 Detection Threshold

It is known from (3.5) that the residual in Kalman filter follows the standard distribution of $N(0, 1)$ [9]. Where:

$$\sigma = \sqrt{H_k P_{k/k-1} H_k^T + R_k}
 \tag{3.6}$$

Assumed that the false alarm rate is P_{FA} . According to the above analysis, the probability of statistic T below the detection threshold can be expressed as:

$$P_r(\Delta Z / \sigma^2 < |a|) = \int_{-a}^a \Phi(t) dt = 1 - P_{FA}
 \tag{3.7}$$

So we can get the detection threshold as:

$$T_d = a\sigma^2
 \tag{3.8}$$

4 Improved RAIM Algorithm

Kalman innovation monitoring is used to estimate the current measurement by historical information, comparing with the current actual measurement to determine whether there is a range fault. The fault of each satellite can be independently detected with less computation, fewer stars and multiple faults, overcoming the shortcoming of the snapshot algorithm. But note that Kalman filter method is not sensitive to slowly-changing faults, which is mainly due to the fact that the test statistic is not only related to the fault deviation of the current time, but also to the estimated deviation before the current time. The improved parity vector method, however, reduces the noise interference through the accumulation of test statistic and has strong detection ability for slowly varying faults. Thus, it can be seen that the combination of Kalman filter and improved parity vector method can effectively compensate for the detection of slowly-changing faults in Kalman filter detection. And it will achieve a better result in integrity monitoring. The process can be illustrated as:

- (1) The Kalman filter detection method is used at first for all satellites visible.
- (2) When there is a fault detected in the previous step, the result of the Kalman filter detection is considered correct.
- (3) When there is no fault detected in the previous step, the non-coherent accumulative parity vector method is used to detect and identify the fault.

When the satellites visible are not enough, the Kalman filter detection will be selected to monitor the fault independently, which effectively avoids the requirement of the non-coherent accumulation parity vector method in the visible satellite quantity, and ensures the availability of the RAIM algorithm. When the satellites visible are enough, the algorithm is mainly based on the non-coherent accumulated parity vector method, supplemented by Kalman filter monitoring to achieve a better detection results.

5 Experiment and Analysis

To evaluate the performance of the improved algorithm, Beidou navigation system satellites signals are simulated. There are 10 satellites, including 5 IGSO/MEO satellites and 5 GEO satellites. The simulation time is 10 min, and the sampling interval is 1 s, with a total of 630 groups of data.

The Gauss white noise with a standard deviation of 0.2 m is added to the pseudorange of all the visible satellites. While one of them has added a slowly-changing error from 0 to 3 m. The fault detection rate is selected as 10^{-6} .

The simulation experiments of the parity vector method, the non-coherent accumulation parity vector method and the improved method proposed in this paper are carried out. The experimental results are shown in Fig. 1.

Fig. 1 Fault detection rate of parity vector method, parity vector method based on non-coherent accumulation and the improved method proposed in this paper

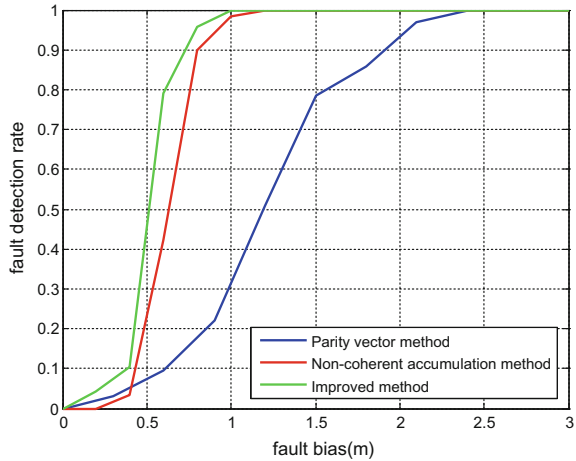
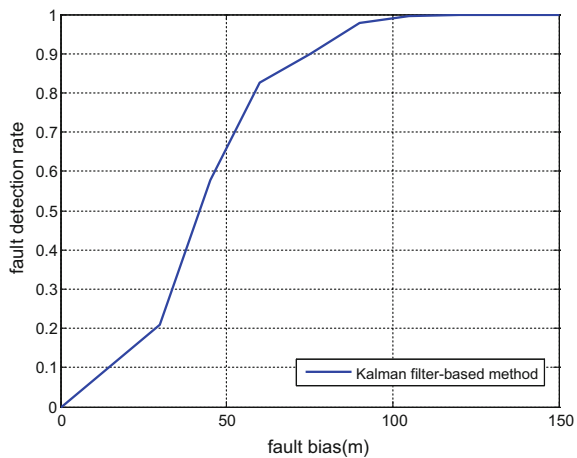


Figure 2 shows that the parity vector method can reach a detection rate of about 99% when the fault is about 2.4 m. When fault is above 2.5 m, the fault detection rate is 100%. The non-coherent accumulative parity vector method reaches a fault detection rate of about 99% in the detection when the pseudorange error is about 1 m. At this time, the fault detection rate of the traditional parity vector method is only 30% of it, so we can see that the improved non-coherent accumulated parity vector method has a high sensitivity to the slowly-changing faults.

The method proposed in this paper has a certain effect of slight and slow-changing fault detection. And the non-coherent cumulative parity vector method is improved with the help of Kalman filter. And when the satellite visible is not enough, the improved method has the same result with Kalman filter-based RAIM method which is shown in Fig. 2, while the other methods are unusable.

Fig. 2 Fault detection rate of Kalman filter-based RAIM method when there are four satellites available



6 Conclusion

In this paper, the two algorithms of RAIM are synthesized and optimized by analyzing the characteristics of them, forming a new algorithm which combines the Kalman filtering method and the improved parity vector method. The algorithm effectively reduces the detection problems of Kalman filter in small and slow-changing faults as well as the strict requirement of the parity vector method. At the same time, the amount of calculation is increased. But the time-consuming is below the sampling step. So it has no actual effect on simulation results. In conclusion, the proposed algorithm is suitable for the scene that the number of visible satellite is hard to be guaranteed and the requirement of fault detection is strict. At last the performance of the algorithm is verified by simulation experiments.

References

1. Xiaohao X, Chuansen Y, Ruihua L (2013) Review and prospect of GNSS receiver autonomous integrity monitoring. *Acta Aeronaut et Astronaut Sinica* 03:451–463. <https://doi.org/10.7527/S1000-6893.2013.0081>
2. Mathieu J, Fangcheng C, Steven L, Boris P (2012) RAIM detector and estimator design to minimize the integrity risk. In: *Proceedings of the 25th international technical meeting of the satellite division of the Institute of Navigation, Tennessee, USA, ION*, pp 3856–3867
3. Hai S, Xinming H, Wenxiang L, Bo X, Gang O (2014) Research on the RAIM method based on non-coherent accumulation for tiny pseudo-range bias. *J Astronaut* 06:708–712. <https://doi.org/10.3873/j.issn.1000-1328.2014.06.013>
4. Minghan S, Guoliang S, Yugui Z (2006) Research on RAIM availability algorithm. *J Telem Track Command* (02) 24–27 + 36
5. Yongchao W, Zhigang H, Guoliang S (2006) Research on the equivalence among three satellite fault identification algorithms. *J Telem Track Command* 01:51–55
6. Shangyue L, Ji X, Shipin H, Jianhua L, Feng L (2014) Research on receiver autonomous integrity monitoring. *Navig Positioning Timing* 01:40–44
7. Cheng C, Guorong H, Xingzhao P, Gang X, Dong X (2013) Research on fault detection through parity vector RAIM algorithm in complex environment. *Missiles Space Veh* 05:37–40. <https://doi.org/10.7654/j.issn.1004-7182.20130509>
8. Jiancai S, Chunping H, Guixiang X (2017) GNSS receiver autonomous integrity monitoring based on EKF. *J Tianjin Univ (Sci Technol)* (04) 405–410
9. Guoliang S, Minghan S, Jinping C (2006) A study on set combined method for receiver integrity autonomous monitoring. *Acta Aeronaut et Astronaut Sinica* 06:1171–1175

LEO Constellation Optimization Model with Non-uniformly Distributed RAAN for Global Navigation Enhancement



Deng Pan, Dan Sun, Jing Ren and Mingtao Li

Abstract This paper established a constellation optimization model for global navigation enhancement, of which RAAN is non-uniformly distributed, aimed at GDOP (geometric dilution of precision) optimization of important cities around the world, using genetic algorithm to solve the constellation optimization problem and using STK to analyse the optimization results. The simulation results show, compared with the classic Walker constellation of which global uniformly distributed RAAN, the Leo constellation of global non-uniformly distributed RAAN not only finishes high precision worldwide global navigation, but reduce the number of satellite constellation, which is beneficial to reduce the cost and deploy rapidly.

Keywords Navigation enhancement · GDOP · LEO Walker constellation Non-uniformly distributed RAAN · Genetic algorithm

1 Introduction

Using Leo satellite constellation to enhance navigation can effectively shorten positing time and information transmission time. The characteristics of low path loss and low orbit can improve the positioning accuracy and anti-interference ability. Therefore, it is becoming a trend to use LEO satellite to build communication system to navigation enhancement.

In the navigation and positioning, the geometric dilution of precision (GDOP) is an important index to calculate position and estimate constellation, which is of practical significance in constellation design and multi-system combinatorial optimization. It is generally believed that the Walker constellation with uniformly

D. Pan · D. Sun · J. Ren · M. Li (✉)

National Space Science Center. CAS, NSSC, Beijing, China
e-mail: limingtao@nssc.ac.cn

D. Pan · D. Sun · J. Ren · M. Li

University of Chinese Academy of Sciences. UCAS, Beijing, China

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_63

distributed RAAN is capable of providing optimal GDOP. For example, the classic LEO mobile satellite system: the Iridium satellite system and the Globalstar mobile satellite system, both are the Walker constellation of global uniform distributed RAAN.

The Iridium satellite system consists of an interconnected network of 66 satellites (plus 6 spares) in LEO. The satellite orbits the Earth at an altitude of 781 km, and the spare satellites are flown at a slightly lower altitude of 649 km. Their orbits are inclined at an angle of 86.4° with respect to the equator, providing global Earth coverage including the north and south poles. The constellation is configured as 6 planes with 11 operational satellites and 1 spare satellite in each plane. The RAAN of adjacent planes are 30° apart. The satellites are equally spaced and arranged [1].

The Globalstar mobile satellite system is composed of 48 operational satellites and 8 spare satellites. The satellite orbits the Earth at an altitude of 1414 km. Their orbits are inclined at an angle of 52° with respect to the equator. The constellation is configured as 8 planes spacing between 45° . Each plane consists of 6 satellites which are equally spaced and arranged. The beam of the Globalstar system can cover the range within latitude of 70° , it means that the Globalstar system isn't applicable in polar regions [2].

The RAAN of constellation used for global navigation is commonly 180° or 360° uniformly distributed. It is generally believed that this kind of distribution mode is more suitable for global navigation. The paper is aiming at problem B of the 9th China Trajectory Optimization Competition (CTOC), put forward a kind of LEO Walker constellation of which RAAN is not 180° or 360° uniformly distributed. By using genetic algorithm (GA), constellation parameters are optimized. We won the championship of CTOC by the lowest cost when ensuring the performance of GDOP. The optimization result has decent GDOP performance worldwide so that be able to perform global navigation enhancement tasks.

2 Constellation Design and Optimization

2.1 Design Indexes and Constraints

The constellation is aimed at enhancing global navigation by using LEO satellite. At any time during construction and service phases, the orbit altitude of all satellites shall be no less than 500 km. The construction time of the constellation is three months. At the end of construction time, the precision of position for injection shall be less than $1e-3$ km/s, the velocity precision shall be less than $1e-6$ km/s. There are two evaluation indexes to the performance of the constellation. The first one is to maximize the total navigation revenue of all key cities; the second index is to minimize the total cost of constellation construction.

- Total navigation revenue

The first index is to maximize the total navigation revenue of all key cities,

$$\begin{aligned}
 \text{Obj}_1 &= \max \left(\sum_{k=1}^K \delta_k w_k \right) \\
 \delta_k &= \begin{cases} 1, & \text{if } \max(\text{GDOP}_{k,l}) \leq 10 \\ 0, & \text{if } \max(\text{GDOP}_{k,l}) \geq 10 \end{cases}, \quad l = 1, 2, \dots, L
 \end{aligned} \tag{1}$$

where, δ_k is 0–1 variable, and it is used to measure whether the k-th city can meet the navigation accuracy requirements, w_k is the weight of the k-th city. $\text{GDOP}_{k,l}$ ($k = 1, 2, \dots, K; l = 1, 2, \dots, L$) is the GDOP of the k-th city at the l-th sampling point.

GDOP is an important factor to measure navigation precision and it is determined by the geometrical relationship between city and satellites. Lower GDOP value indicates higher precision. Suppose there are N visible satellites for the k-th city at the l-th sampling point, denote a unit vector from k-th city to n-th satellite as $\bar{r}_{k,n} = [(x_n - x)/R_n, (y_n - y)/R_n, (z_n - z)/R_n]$, where $[x_n, y_n, z_n]$ is position components of the n-th satellite, $[x, y, z]$ is the position components of the k-th city, $R_n = \sqrt{(x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2}$, create the following matrix H

$$H = \begin{bmatrix} \frac{(x_1-x)}{R_1} & \frac{(y_1-x)}{R_1} & \frac{(z_1-x)}{R_1} & 1 \\ \frac{(x_2-x)}{R_2} & \frac{(y_2-x)}{R_2} & \frac{(z_2-x)}{R_2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{(x_N-x)}{R_N} & \frac{(y_N-x)}{R_N} & \frac{(z_N-x)}{R_N} & 1 \end{bmatrix}_{N \times 4} \tag{2}$$

define $Q = (H^T H)^{-1}$, and its principal diagonal elements are $Q_{11}, Q_{22}, Q_{33}, Q_{44}$ respectively, then the GDOP can be evaluated by $\text{GDOP}_{k,l} = \sqrt{Q_{11} + Q_{22} + Q_{33} + Q_{44}}$, note that at least 4 visible satellites are required to evaluate GDOP [3].

- Total cost of constellation construction

The second index is the total cost of constellation construction. It is related to satellite numbers and the way satellite deployed. The navigation constellation shall be constructed by two ways. The first one is launching new rockets. The distributor carries at most 16 navigation satellites and located at a circular parking orbit with altitude equal to 900 km. Another way is utilizing piggyback missions. The distributor carries at most 8 navigation satellites and stays in an orbit of piggyback missions (carry orbits). For saving the cost, we should choose the second way to deploy satellite as far as possible.

$$C_{\text{all}} = N_{\text{launch}}C_{\text{launch}} + N_{\text{carry}}C_{\text{carry}} + N_{\text{naviSat}}C_{\text{naviSat}}$$

$$\text{Obj}_2 = \min(N_{\text{launch}}C_{\text{launch}} + N_{\text{carry}}C_{\text{carry}} + N_{\text{naviSat}}C_{\text{naviSat}})$$

where, N_{launch} , N_{carry} , N_{naviSat} are the number of new launches, the number of piggyback missions and the number of navigation satellites of the constellation. C_{launch} , C_{carry} , C_{naviSat} are single launch cost, single piggyback cost and cost of single navigation satellite, set as 1.2, 0.2 and 0.05 respectively (virtual currency unit). The total cost should meet the constraints $C_{\text{all}} < 10$.

2.2 Constellation Configuration Optimization

The optimization of navigation constellation is actually to search a parameters' combination to get the best constellation performance and economic indicators [4]. The parameters mainly consist of type of orbit, number of satellite, number of planes, distribution of RAAN, distribution of satellites in one orbit, orbital altitude and inclination. All parameters influenced the results need to be considered comprehensively to realize the optimal constellation.

2.2.1 Walker Constellation Modeling

Because of good coverage performance, Walker constellation is usually used in large constellations, such as GPS, as well as the Iridium and the Globalstar mobile communication system mentioned earlier. The Walker constellation have the following characteristics: the orbital altitude and inclination of every orbit are the same. Besides, the constellation can be represented by three parameters T/P/F. The parameter T means number of total satellites and P means number of planes. Assuming that N is the number of satellites in one orbit, we have $T/P = N$. In Walker constellation, the first orbit is called nominal orbit and the first satellite in one orbit is called nominal satellite. The parameter F represent the latitude angle of the nominal satellite whose left orbit's nominal satellite passing ascending node exactly.

The initial phase of nominal satellite in the first orbit and phase difference F determine the initial phase of nominal satellite in each orbit. If Ω_0 which means the RAAN of nominal orbit at the initial moment, M_0 which means the phase of nominal satellite in nominal orbit and $\Delta\Omega$ which means the range of RAAN are already known, the RAAN Ω_{ij} and phase M_{ij} of the j-th star in the i-th orbit are calculated by [5]:

$$\begin{cases} \Omega_{ij} = \Omega_0 + (i-1) \frac{\Delta\Omega}{P-1} \\ M_{ij}(0) = M_0 + (i-1) \cdot F \cdot \frac{2\pi}{T} + (j-1) \cdot P \cdot \frac{2\pi}{T} \end{cases} \quad (3)$$

where, $i = 1, 2, \dots, P$; $j = 1, 2, \dots, N$.

2.2.2 Satellite Dynamics Modeling

During the construction and service phases, only earth’s central gravitational field and the perturbation of J_2 are considered. The Cartesian coordinate description is implemented during the orbit transfer, and the dynamic model is shown below

$$\begin{cases} \ddot{x} = -\frac{\mu x}{r^3} \left(1 + \frac{3}{2} J_2 \left(\frac{Re}{r} \right)^2 \left(1 - 5 \frac{z^2}{r^2} \right) \right) \\ \ddot{y} = -\frac{\mu y}{r^3} \left(1 + \frac{3}{2} J_2 \left(\frac{Re}{r} \right)^2 \left(1 - 5 \frac{z^2}{r^2} \right) \right) \\ \ddot{z} = -\frac{\mu z}{r^3} \left(1 + \frac{3}{2} J_2 \left(\frac{Re}{r} \right)^2 \left(1 - 5 \frac{z^2}{r^2} \right) \right) \end{cases} \quad (4)$$

where, μ is gravitational constant of the earth, x, y, z are position components of the satellite in the ECI coordinate system, $r = \sqrt{x^2 + y^2 + z^2}$ is geocentric distance, Re is equatorial radius, J_2 is earth oblateness gravity coefficient [6].

The orbit propagation with mean orbital elements is implemented before the first impulse and after the last impulse (including the service phase) due to its efficiency. The dynamic model of mean orbital elements under J_2 perturbation is shown below

$$\begin{cases} \dot{a} = 0 \\ \dot{e} = 0 \\ \dot{i} = 0 \\ \dot{\Omega} = -\frac{3A_n J_2}{2(1-e^2)^2} \left(\frac{Re}{a} \right)^2 \cos i \\ \dot{\omega} = \frac{3A_n J_2}{4(1-e^2)^2} \left(\frac{Re}{a} \right)^2 (5 \cos^2 i - 1) \\ \dot{M} = \frac{3A_n J_2}{4(1-e^2)^2} \left(\frac{Re}{a} \right)^2 (3 \cos^2 i - 1) \end{cases} \quad (5)$$

where, a is semi-major axis, e is eccentricity, i is inclination, Ω is right ascension of ascending node, ω is argument of perigee, M is mean anomaly. $A_n = \sqrt{\mu/a^3}$ is the mean motion of the orbit [6].

2.2.3 Genetic Algorithm (GA)

Genetic Algorithm (GA) is a random search algorithm which is formed by simulating natural selection and natural genetic mechanisms. Its basic idea is from Darwin’s theory of evolution and Mendel’s theory of genetics. It uses coding space to replace parameter space of the problems, and use the fitness function as evaluation of an individual. Meanwhile, GA uses choosing, cross and mutation to build an iteration to simulate the natural selection and genetic mechanisms [7]. This process will result in that offspring have bigger fitness than their parents. The

optimal individual of last generation which is decoded can be regarded as the approximate optimal solution of the problem.

2.2.4 Optimize Variables and Range

The model of Walker constellation contains two integer parameters and six orbit parameters: the number of planes P , the number of satellites in each orbit N , orbital radius a , orbital eccentricity e , orbital inclination i , RAAN of nominal orbit Ω_0 , phase angle of nominal orbit u_0 , phase difference of adjacent orbit F .

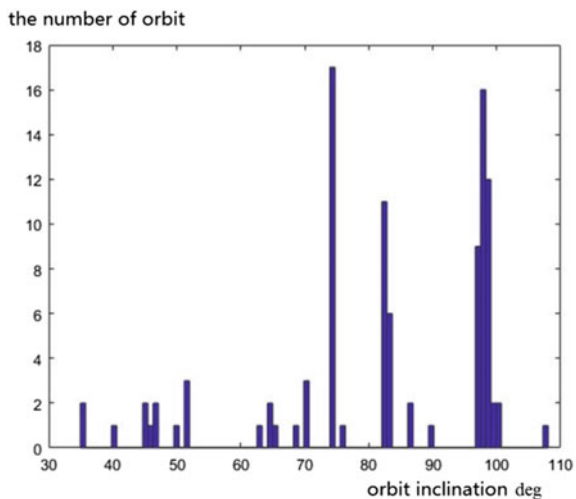
This paper uses Walker constellation with circle orbits, so the number of input variables becomes 7 and the input variables are the number of constellation orbits P , the number of satellites in each orbit N , orbital radius a , orbital inclination i , RAAN of nominal orbit Ω_0 , the coverage of RAANs $\Delta\Omega$, phase difference of adjacent orbit F .

It consumes a lot of fuel to change the inclination and RAAN of a satellite. Therefore, orbital inclination should be in a range where the carry orbits concentrate. Figure 1 shows the inclination of carry orbits mostly focuses on 74° , 83° and 98° . According to the experimental results, when inclination greater than 55° , the navigation performance of Walker becomes worse as the inclination increases. So 74° is chosen as orbital inclination.

Due to circular Walker orbits have more uniform coverage, set the orbital eccentricity e to 0. Considering that one piggyback mission can carry 8 satellites, set $N \leq 8, N \in N^+$.

The coverage performance is better as the orbit higher, so the fuel is main used to adjust the semi-major axis. According to the fuel, the optimization range of the orbit radius is set up at 9100–9400 km.

Fig. 1 Orbit inclination distribution of given carry orbit. There are a total of 100 orbitals



In order to save fuel, we use J2 perturbation to deploy RAAN instead of using fuel. Therefore, we compute the range of RAAN which J2 perturbation leads to. The Fig. 2 shows the RAAN range of carry orbits at the end of construction under J2 perturbation. We chooses 100° – 130° as the range of Ω_0 and 120° – 320° as the range of $\Delta\Omega$ to optimize.

2.2.5 Optimization Procedure and Result

Taking obj1 from (1) as objective function, solve the constellation model using GA. The parameters in GA were set as follows: Population Size = 100, interval for population = [0, 1], crossover probability = 0.8, mutation probability = 0.2, Generations = 50, StallGenLimit = 30.

The optimization solution is, when $i = 74^{\circ}$, $T = 77$, $P = 11$, $F = 5.914$, $a = 9337.001$ km, $\Delta\Omega = 202.924^{\circ}$, $\Omega_0 = 105.311^{\circ}$. That means a 77/11/5.914 Walker constellation. Specifically, 77 satellites were uniformly distributed in 11 circular orbit planes of which radius is 9337.001 km, inclination is 74° . The RAAN of the constellation spans from 105.31° to 308.23° .

This non-globally distributed Walker constellation reduces the number of satellites and the costs of construction. Also, it achieves global high precision navigation, which leads to best navigation performance.

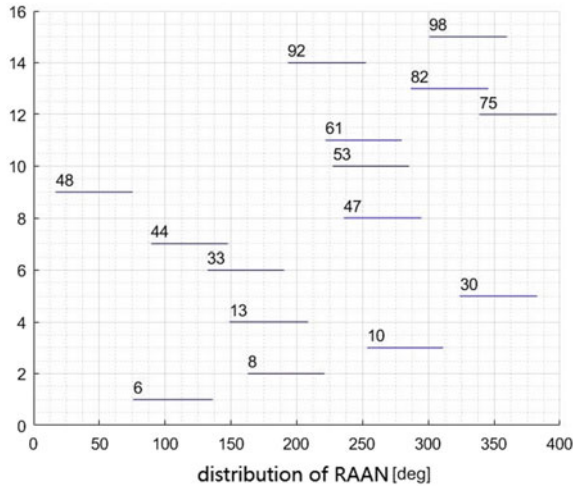


Fig. 2 Accessible RAAN range of carry orbits at the end of construction under J2 perturbation

3 Simulation and Analysis

Figure 3 shows the constellation configuration, simulated by STK. We set 500 observation points in key cities around the world and special positions like the equator and two poles. The corresponding weight values are given according to the importance of observation points ($w_k = 1, 2, \dots, 5$). Calculated in STK, within two years, each observation point at each instant can be covered by 4 satellites at least. It means the optimization result can well fulfill the navigation enhancement task.

As Fig. 3 shows, though the RAAN of constellation doesn't span 360° or 180° , the orbits form a uniform envelope over Earth. This is the reason that it can achieve global high precision navigation.

Figure 4 shows the GDOP value of several cities and the North Pole within 1 month. Obviously, GDOP value is lower than 5 whether in high or low latitude area.

4 Constellation Deployment

To save cost, we construct the constellation by carrying. Figure 5 shows which carry orbit can carry satellites to a specific target orbit plane. The vertical lines represent target orbit planes in sequence, horizontal lines represent the carry orbits. In case the vertical and horizontal lines cross, that carry orbit can carry satellites to that orbit plane.

The accuracy requirement of construction is pre-determined as follows, the precision of position for injection shall be less than $1e-3$ km/s, the velocity precision shall be less than $1e-6$ km/s and the mass precision shall be under $1e-4$ kg. To meet the requirement, we apply a small-pulse at initial moment, utilizing perturbations to phase, and then adjust radius and eccentricity by Hohmann rendezvous,

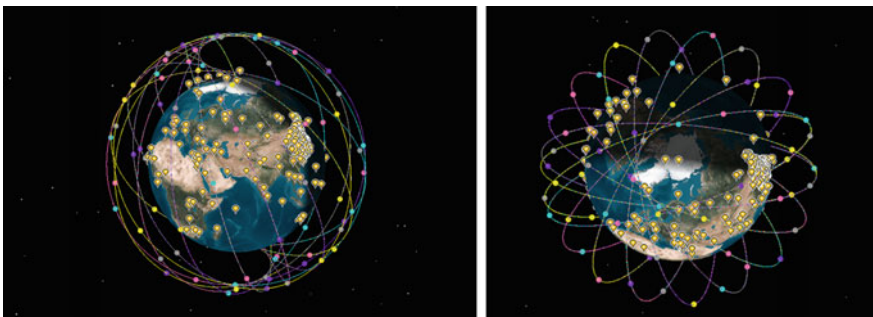


Fig. 3 The Walker constellation configuration, simulated by STK. The yellow dots on earth surface represent observation points

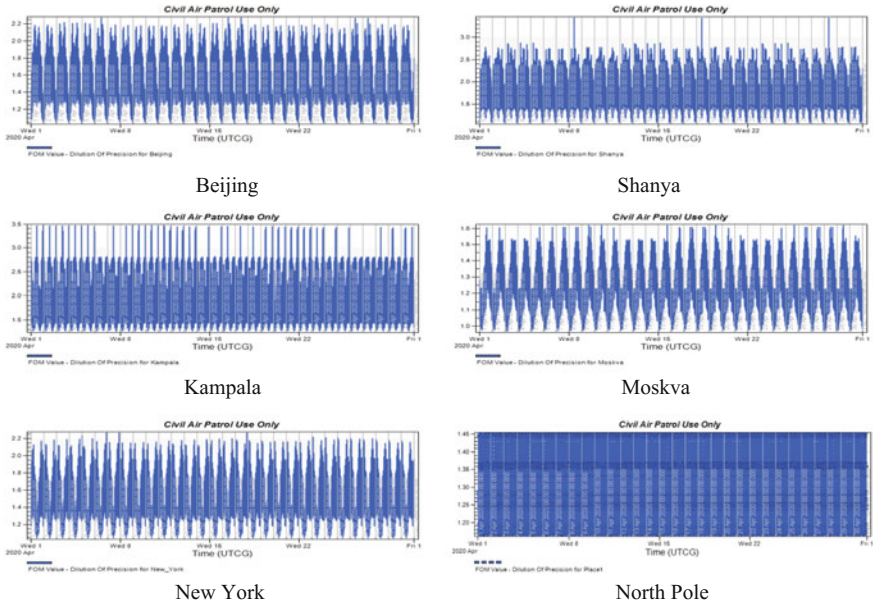
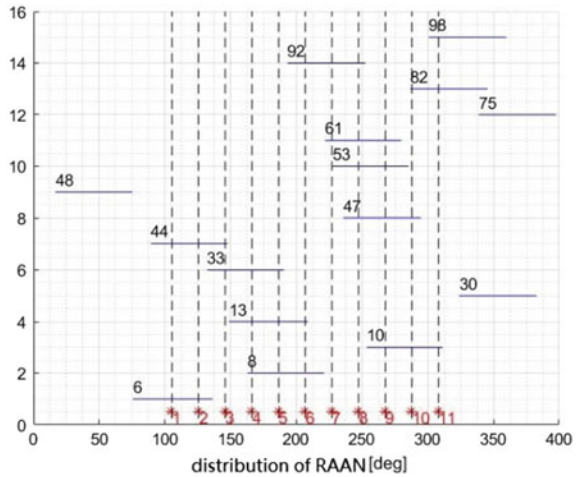


Fig. 4 GDOP of several cities and North Pole within 1 month

Fig. 5 Optimized result of RNA distribution. At the end of construction period



adjusting RAAN under the influence of J2 perturbations; finally, the satellites enter orbits through Lambert transfer.

Upon inspection, all the 77 satellites can finish injection at completion time of construction phase, and fulfill the accuracy requirement.

5 Conclusion

This paper improve the traditional Walker constellation, take the navigation performance of global key cities and cost of construction as objective function, establish an RAAN non-global uniformly LEO Walker constellation model, and solve it using GA. Reduces the number of satellites as much as possible, which is beneficial to reduce the cost and deploy rapidly.

The modified Walker constellation can achieve a high precision navigation performance, can fulfill the global navigation enhancement task. Meanwhile, we explore the constellation deployment pattern of using carry orbits, which is beneficial to save cost and construct rapidly, provide references for deployment of LEO satellite.

References

1. Garrison TP, Ince M, Pizzicaroli J et al (2015) Systems engineering trades for the IRIDIUM constellation. *J Spacecr Rockets* 34(5):675–680
2. Laixing L (1997) Globalstar constellation and its attitude and orbital control system. *Erospace China* (1):17–18 + 21 + 2
3. Langley RB (1999) Dilution of precision. *Gps World* 10(5):52–59
4. Ping S, Guangji Q, Zhonggui C (2006) Studies on the design and analysis of regional navigation constellations. *Chin J Space Sci* 26(4):268–276
5. Song Z-m, Dai G-m, Wang M-c, Lei PFNG (2014) Theoretical analysis of Walker constellation coverage to area target. *Comput Eng Des* 10:3639–3644
6. Zeng Y (2007) Constellation design based on genetic algorithm. Huazhong University of Science and Technology
7. Zhang Y, Fan L, Zhang Y (2008) Theory and design of satellite constellation. Science Press

Design of Navigation Message Authentication Assisted by Ground Based Augmentation Systems



Muzi Yuan, Zhe Liu, Xiaomei Tang, Shengqiang Lou and Gang Ou

Abstract Spoofing attack organized by generating fake navigation signal can precisely manipulate PNT output of target receiver. In this paper an authenticated augmentation message for GBAS is designed to obtain the authentication for both navigation message from satellites and augmentation message from GBAS broadcasters. The proposed navigation message authentication method has an advantage in efficiency and authentication delay compared with legacy schemes integrated into satellite signal while has an equally high security level as other schemes.

Keywords Anti-spoofing · Ground based augmentation system
Navigation message authentication

1 Introduction

While an increasing number of critical infrastructures rely on the integrity of GNSS service, life security may be threatened by spoofing attacks [1]. GNSS spoofing attack refers to the attack performed by broadcasting a counterfeit GNSS signal to the target receiver. Receivers cannot distinguish an authorized signal from its fake replica without extra information. Navigation message authentication (NMA) is an effective method to verify the integrity of navigation message by signature or other message authentication codes (MACs).

A number of NMA methods integrated in satellite signal has been proposed in recent years [2–7]. The efficiency of these methods are measured by time to first authenticated fix (TTFAF) and time between authentications (TBA) [8]. In latest research, NMA based on satellite signal can achieve a TTFAF and TBA of 10 s in Galileo open service [6]. Main drawbacks of these NMA methods is the relatively

M. Yuan · Z. Liu · X. Tang · S. Lou · G. Ou (✉)
School of Electronic Science, National University of Defense Technology,
410073 Changsha, China
e-mail: ougangcs@139.com

low bitrate and the weak adjustability of navigation message in satellite signal. Low bitrate requires long transition time and weak adjustability leads to huge amount of bits involved in a single NMA fix.

Signal of ground based augmentation system (GBAS) has the advantage of higher bitrate and more agile message arrangement compared with satellite signal. There are many methods raised for GBAS to improve the performance of GNSS [9–11]. However, there are few literature focusing on the possibility of implementing an NMA assisted by GBAS broadcasters. We proposed a basic design of NMA assisted by GBAS broadcasters to obtain a high efficiency and a high robustness against spoofing attacks.

The paper is organized as follows: the principle and model of NMA protection against spoofing attack are discussed in Sect. 2; the proposed implementation of NMA assisted by GBAS is demonstrated in Sect. 3; an analysis of performance is performed in Sect. 4.

2 Model of NMA Protection Against Spoofing Attack

Spoofing attacks neutralized by NMA are categorized as intermediate spoofing attack and estimation and replay attack [12]. The first category is performed by generating counterfeit GNSS signal based on the real position and the spoofing position of target and estimating the navigation message for spoofing attack. The second category is performed by recording and tampering the authorized GNSS signal and replaying it in spoofing scenarios. NMA is able to protect the integrity of GNSS service against spoofing attack by recognizing those falsify in navigation message. In this chapter model of spoofing attack and NMA are discussed.

2.1 Model of Spoofing Attacks

Receivers perform a position, velocity and time (PVT) solution by implying parameters (x_i, y_i, z_i, ρ_i) form GNSS signal into Eq. (2.1).

$$\sqrt{(x_i - x_r)^2 + (y_i - y_r)^2 + (z_i - z_r)^2} + c\Delta t_{r-s} = \rho_i, i = 1, \dots, N \quad (2.1)$$

Here in Eq. (2.1) (x_i, y_i, z_i) and (x_r, y_r, z_r) are coordinate of satellite and receiver, c is the speed of light, Δt_{r-s} is the clock differential between receiver and satellite, ρ_i is the pseudorange measurement from satellite to receiver.

Since parameter (x_i, y_i, z_i) is determined by navigation message, spoofing attackers can generate fake navigation message to lead the receiver believe in spoofed parameter (x'_i, y'_i, z'_i) . If the spoofed parameter (for example in satellite

No. 1) fulfills Eq. (2.2), the receiver will make a PVT solution $(x'_r, y'_r, z'_r, \Delta'_{r-s})$ under the manipulation of spoofing attackers.

$$\begin{aligned} \sqrt{(x'_1 - x'_r)^2 + (y'_1 + y'_r)^2 + (z'_1 + z'_r)^2} + c\Delta'_{r-s} &= \rho'_1 \\ \sqrt{(x_i - x'_r)^2 + (y_i + y'_r)^2 + (z_i + z'_r)^2} + c\Delta'_{r-s} &= \rho_i, i = 2, \dots, N \end{aligned} \quad (2.2)$$

Hence, the combination of authorized signal and spoofed signal can be modeled as Eq. (2.3).

$$Y_k = \alpha \hat{w}_{k-d} c_{k-d} + w_k c_k + N_k \quad (2.3)$$

Here in the model, \hat{w}_{k-d} is the manipulated navigation message fulfills Eq. (2.2), c_{k-d} is the delayed pseudorange code, w_k and c_k are navigation message and pseudorange code broadcasted by authorized satellite. α is the energy gain of spoofing signal and N_k is noise.

2.2 Principle of Navigation Message Authentication

The basic function of NMA is to distinguish \hat{w}_{k-d} from w_k in Eq. (2.3). Most NMA schemes are based on signature or MAC algorithms such as elliptic curve digital signature algorithm (ECDSA) [13] and timed efficient stream loss-tolerant authentication (TESLA) [14].

ECDSA uses key pairs to ensure message authentication. Private keys which are only held by satellites are employed to sign message while public keys which are published to public are employed to verify the signature attached to the message. TESLA uses key stream to derive MACs form message. Keys to verify MACs are announced after MACs are broadcasted to protect privacy.

In NMA schemes previously proposed the combination of NMA protected signal and spoofed signal can be modeled as Eq. (2.4).

$$Y_k = \alpha(\hat{w}_{k-d} || \hat{s}_{k-d})c_{k-d} + (w_k || s_k)c_k + N_k \quad (2.4)$$

Here in Eq. (2.4) operator $||$ is to merge to strings, s_k is signature or MAC derived by NMA schemes. Since spoofing attackers have no access to private keys or unannounced MAC keys, fake signature or MAC \hat{s}_{k-d} will not pass any verification in receivers. Hence receivers can identify authorized signal under protection of NMA.

In the proposed scenario of this paper, receivers are under the coverage of GBAS broadcaster. Hence NMA messages are transmitted through GBAS signals. The structure of proposed NMA is shown in Fig. 1.

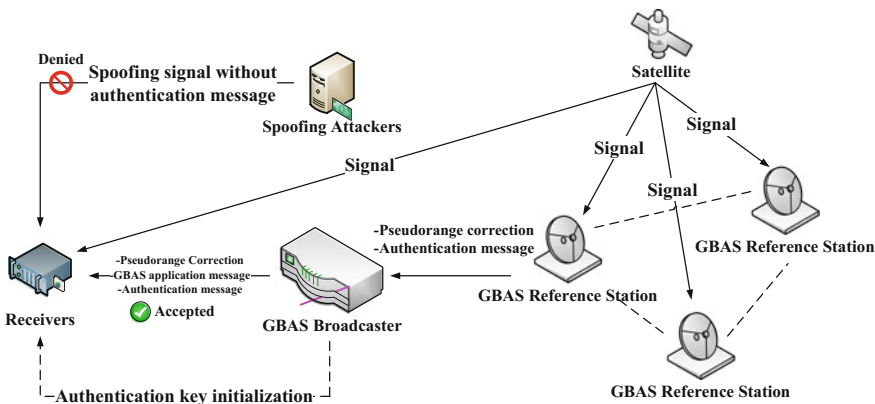


Fig. 1 The structure of proposed NMA

3 NMA Implementation with GBAS Broadcasters

While the NMA is integrated into satellite signal, those signatures or MACs are extra information of legacy navigation message which require modification in the navigation message and occupy extra communication bandwidth of satellites. The GBAS broadcaster can provide a data transmission rate up to 40.5 Kbps, which is much higher than the 50 bps (120 bps for Galileo) navigation message. If NMA is provided by third-party information source such as GBAS broadcasters the capacity and agility of authentication message may be extended.

In this chapter we propose an implementation of NMA scheme provided by GBAS broadcasters. The structure of broadcasting information are specified and the behaviour of both GBAS providers and receivers are described.

3.1 Structure of NMA Protected GBAS Broadcasting Message

In previously proposed GBAS implementations, broadcasting messages are composed of various message types including pseudorange corrections, ground based ranging source information and information for other applications [11]. The proposed NMA scheme is integrated into GBAS broadcasting by a new message type composed of three sections: frame information section, navigation message authentication section (NMAS) and GBAS message authentication section (GMAS).

The NMAS is the section to authenticate navigation message broadcasted by satellites. The NMAS is composed of frame ID, page ID, time information and authentication message. The authentication message is calculated via navigation

Table 1 Data structure of NMA and GMAS for D1 navigation message in BeiDou system

Data content	Bits used	Range of values
Message type ID	8	0–255
PRN ID	6	0–63
Frame ID	3	1–5
Page ID	7	1–24
Time identifier	–	–
Week number	13	0–8191
Second of week	20	0–604,800
NMA authentication message	Variable	–
GMAS authentication message	Variable	–

message of the particular frame and page indicated in the head of the NMA. The GMAS is the section to authenticate the integrity of GBAS message itself. The GMAS is an authentication message for all the messages broadcasted in the interval of two frames.

As an example, the data structure of NMA and GMAS for D1 navigation message in BeiDou system is shown in Table 1.

The NMA and GMAS are broadcasted after every frame is delivered. This time interval is 6 s in D1 navigation message of BeiDou system. Message length of both NMA and GMAS are variable for various security requirements. For critical scenarios, message length of NMA may achieve the maximum length of 569 bits.

3.2 Generation of NMA and GMAS Messages

NMA messages are generated through satellite signal after every frame is delivered. GBAS first check the integrity of navigation message by calculating and comparing the positioning solution with its real location. Then frame ID, page ID and time information are extracted from the navigation message. Authentication message of this frame is calculated and truncated via secure hash algorithms (for example SHA-256). The process of NMA messages is shown in Fig. 2.

GMAS message is a signature of all the GBAS messages broadcasted in last interval of navigation message frame, which is generated through private keys and verified through public keys. GBAS calculate and truncate the signature via asymmetric cryptographic algorithms (for example ECDSA). The process of GMAS messages is shown in Fig. 3.

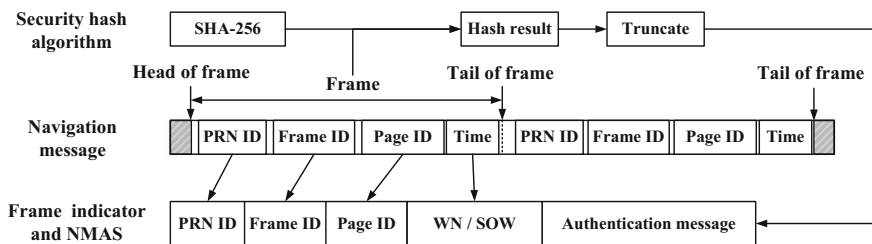


Fig. 2 Generation of frame indicator and NMAS from received navigation message

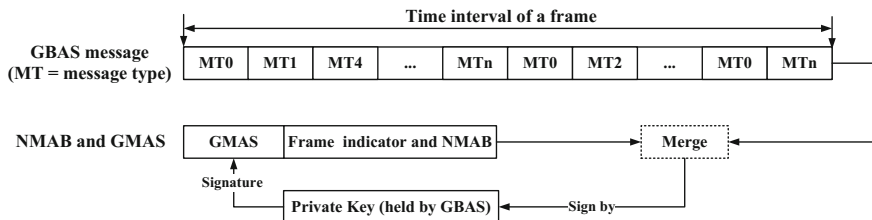


Fig. 3 Generation of GMAS from GBAS message and the form of NMA message type

3.3 Verification of NMAS and GMAS Messages

After receiving a whole frame of navigation message, receivers first verify the integrity of GBAS message by verifying if the signature in GMAS is consistent with GBAS message received in last frame interval. While GBAS message is reliable, receivers calculate hash result of the navigation message frame indicated in the frame information section via the same security hash algorithm. If the result matches with the NMAS, the navigation message frame will pass the authentication. Otherwise a bit error or spoofing attack can be detected.

Verification logic diagram is shown in Fig. 4.

In the proposed scheme, the GBAS augment message is easier to be authenticated than the navigation message. Thus the verification of GBAS augment message is set up before NMA to ensure security.

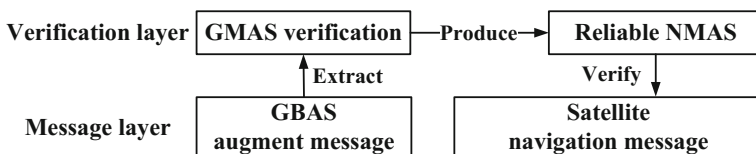


Fig. 4 Logic diagram of authentication verification

4 Performance Analysis

Performance of NMA can be indicated from the aspect of security and efficiency. For security indicators, this paper takes the average attack time (AAT) and the possibility of false alarm (PFA) into consideration. For efficiency indicators, the TTFAF and TBA are taken into consideration.

4.1 Security

The security of the proposed NMA scheme is ensured by secure hash algorithm and asymmetric cryptographic algorithm. Introduced in a 1998 report, it takes 3 months of a network of 50,000 Pentium Pro 200 MHz machines to crack a ECCp-109 challenge in ECDSA [13]. Based on Moore’s law and analysis of hardware associate [15], the estimating time to crack a ECDSA signature system by the network of 50,000 mainstream machines can be shown as Table 2.

For a normal expiration period (e.g. 1 year), ECDSA is secure for NMA in predictable future.

4.2 Efficiency

The GBAS broadcasting data rate (40.5 Kbps) is much higher than civil navigation message (120 bps in Galileo and 50 bps in other systems). Hence a broadcast of 569 bit NMA message every frame only cost a minor partition of broadcasting ability and can be fulfilled easily.

In analysis below, we take D1 navigation message in BeiDou system as an example. While public keys have been initialled into receivers, mean TTFAF and TBA can be calculated via Eqs. (4.1) and (4.2).

$$\overline{TTFAF} = \frac{1}{L_f} \cdot \sum_{n=1}^{L_f} \frac{(n + L_f)T_b}{(1 - BER)^{L_f}} \tag{4.1}$$

Table 2 Crack time estimation for different length of key by different platforms

Key length	Pentium M	XC3S1000 [15]	ASIC
96	10 h	3.6 h	–
128	97 years	51 years	5 months
160	7.6×10^6 years	6.2×10^5 years	5.0×10^4 years
233 (estimate)	5.0×10^{16} years	4.1×10^{15} years	3.3×10^{14} years

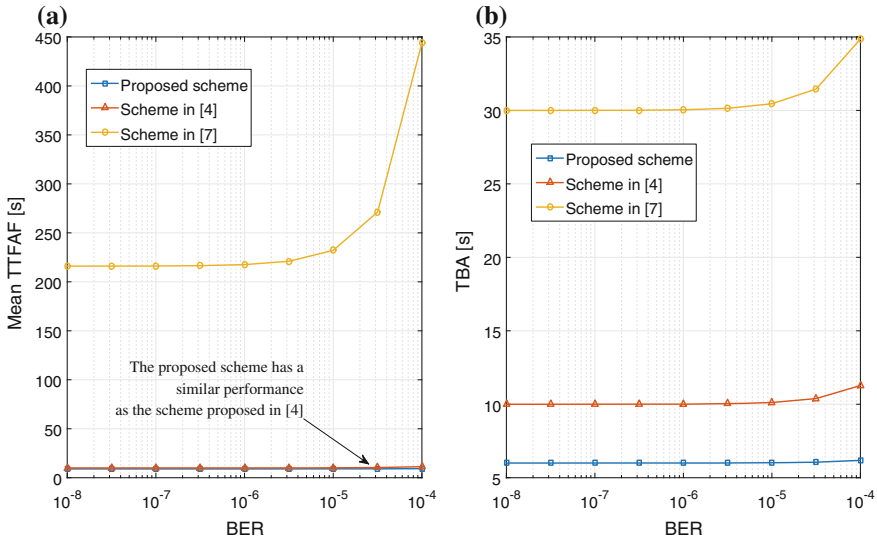


Fig. 5 A compare in TTFAF and TBA among three NMA schemes. Scheme in [4] is a scheme proposed for Galileo with 120 bps data rate. Scheme [7] is a scheme proposed for BeiDou system with 50 bps data rate

$$\overline{TBA} = \frac{T_b L_f}{(1 - BER)^{L_f}} \tag{4.2}$$

Here L_f is length of the navigation message involved in one authentication fix, T_b is interval of navigation message bit, BER is bit error rate of navigation message.

A performance compare in efficiency between legacy NMA scheme integrated in satellite signal [4, 7] and the proposed NMA scheme assisted by GBAS is shown in Fig. 5. The proposed NMA scheme effectively reduce both TTFAF and TBA for the authentication.

In TTFAF aspect, the proposed scheme maintains a similar performance level as the scheme proposed in [4], which is a scheme with high efficiency in the 120 bps Galileo system. In TBA aspect, the proposed scheme is better than other two schemes with the fastest authentication recovery.

5 Conclusion

This paper proposed an NMA scheme assisted by GBAS. Receivers can ensure the integrity of both navigation message and GBAS message via the proposed message type of GBAS broadcasting. The proposed scheme has an advantage in efficiency and authentication delay compared with legacy schemes integrated into satellite signal while has an equally high security level as other schemes.

Acknowledgements This work is supported by National Science Foundation of China (61601485).

References

1. Anonymous (2001) Vulnerability assessment of the transportation infrastructure relying on the global positioning system. John A. Volpe National Transportation Systems Center, U.S
2. Wullems C, Pozzobon O, Kubik K (2005) Signal authentication and integrity schemes for next generation global navigation satellite systems. In: Proceedings of the European navigation conference
3. Wesson KD, Rothlisberger MP, Humphreys TE (2011) A proposed navigation message authentication implementation for civil GPS anti-spoofing. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011), pp 3129–3140
4. Curran JT, Paonni M, Bishop J (2014) Securing the open-service: a candidate navigation message authentication scheme for Galileo E1 OS. In: European navigation conference ENC 2014, Rotterdam
5. Kerns AJ, Wessons K, Humphreys T (2014) A blueprint for civil GPS navigation message authentication. In: Proceedings of IEEE/ION PLANS 2014, Monterey, CA, pp 262–269
6. Hernandez IF, Rijmen V, Granados GS et al (2016) A navigation message authentication proposal for the Galileo open service. *Navig J Inst Navig* 63:85–102
7. Yuan M, Lv Z, Chen H et al (2017) An implementation of navigation message authentication with reserved bits for civil BDS anti-spoofing. In: China satellite navigation conference (CSNC) 2017 proceedings, Vol 2, pp 69–80
8. Hernaandez IF, Rijmen V, Granados GS, et al (2014) Design drivers, solutions and robustness assessment of navigation message authentication for the Galileo open service. In: Proceedings of the 27th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2014), pp 2810–2827
9. Braff R, Shively C (2005) A method of over bounding ground based augmentation system (GBAS) heavy tail error distributions. *J Navig* 58:83–103
10. Dautermann T, Felix M, Grosch A (2012) Approach service type D evaluation of the DLR GBAS testbed. *GPS Solut* 16:375–387
11. Ludwig T, Korn B, Geister R, et al (2011) Towards higher levels of automation in taxi guidance: using GBAS terminal area path (TAP) messages for transmitting taxi routes. In: 30th digital avionics systems conference, Vol 4, No. 5, pp 1–11
12. Margaria D, Motella B, Anghileri M, et al (2017) Signal structure-based authentication for civil GNSSs: recent solutions and perspectives. In: *IEEE signal processing magazine*, pp 27–37
13. Johnson D, Menezes AJ, Vanstone SA (2001) The elliptic curve digital signature algorithm (ECDSA), certicom corporation
14. Perrig A, Canetti R, Tygar JD, Song D (2002) The TESLA broadcast authentication protocol. *CryptoBytes* 5(2):2–13
15. Gueneysu T, Paar C, Pelzl J (2007) Attacking elliptic curve cryptosystems with special-purpose hardware, field programmable gate arrays, pp 207–215

Configuration Method of the Dynamic Navigation Network Based on Near Space



Lizhi Deng, Yingkui Gong, Hongxia Wan and Bingcheng Liu

Abstract This paper focuses on the optimal design of the near space navigation network configuration concerning a specific target area. Optimal design not only ensure navigation service for specific area with high accuracy and low cost; but also enable all users in the target area to obtain a controlled precision. In this design, a double-layer Y configuration is adopted as the optimal basic unit. Adaptive algorithm and genetic algorithm are selected to detect the boundary of the configuration to realize the synchronization optimization of cost and accuracy. Combining with the working principle of aerostat in the stratosphere airspace, we design a platform dynamic trajectory and provide an algorithm optimization of layout for the partial users' worse precision in dynamic scene. The simulation results show that the optimal design method proposed in this paper can meet the configuration design requirements of the dynamic navigation network based on near space with different size area and different precision requirement.

Keywords Near space · Intelligent algorithm · Dynamic navigation network
Aerostat

1 Introduction

Near space is a public space (20–100 km) that the countries all over the world want to seize and utilize, possessing high available value. The navigation network based on near space can effectively improve the regional positioning accuracy on the ground, improve the positioning accuracy of satellite navigation, and improve the

L. Deng

University of Chinese Academy of Sciences, No. 19(A) Yuquan Road,
Shijingshan District, Beijing 100049, China

L. Deng · Y. Gong (✉) · H. Wan · B. Liu

Academy of Opto-Electronics, Chinese Academy of Sciences,
No. 9 Dengzhuang South Road, Haidian District, Beijing 100049, China
e-mail: ykgong@aoe.ac.cn

© Springer Nature Singapore Pte Ltd. 2018

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2018*

Proceedings, Lecture Notes in Electrical Engineering 497,

https://doi.org/10.1007/978-981-13-0005-9_65

stability and reliability of navigation services [1]. The aerostats such as super pressure balloon are developing rapidly, and they can rise or even float for long time in 20–25 km.

Recently, domestic research on near space network is still in the exploratory stage, and there is a lack of systematic optimization analysis for the stratospheric network model, e.g. from unit to whole, edge detection and so on. Meanwhile, they are divorced from reality during the operation of the aerostat and the overall operation of the platform without taking the characteristics of airspace into account [2]. In this paper, we focus the study on the configuration method of dynamic navigation network for wide area and seamless coverage. And we put forward a static network configuration method based on optimal design of the basic unit and the optimization method of the dynamic network configuration which is suitable for the platform runtime environment.

2 Optimal Unit Configuration

The optimal unit configuration is the premise and basis of the whole network configuration optimization, which directly affects the performance of the entire network.

2.1 Useful Location Calculation

Actually, if the elevation angle of the near space platform is too small, a large amount of ground noise, atmospheric refraction and ground scattering will seriously deteriorate the locating signal [3]. Therefore, considering the elevation angle, the computational formula for the actual coverage radius of the near space platform will change accordingly. The formula is shown as follows:

$$r = R_e \times \left[\arccos \left(\frac{R_e \times \cos \alpha}{R_e + h} - \alpha \right) \right] \quad (1)$$

In the formula, R_e is the earth radius, α is the elevation angle, and h is the distance between platform and ground. When calculating the coverage radius of the near space platform, α is 6° , which is in line with the actual situation, and its coverage radius is shown in Table 1.

Table 1 Coverage radius of near space platform

h/km	35	30	28	25	24	20	18
r/km	274.9	240.9	226.9	205.5	198.2	168.5	153.3

It can be seen from above table that the higher the height of the platform is, the wider the vision is, and hence the greater the coverage radius is. The near space platform for user navigation and positioning should be located in the smallest airspace near the edge of wind cutting. Relevant data show that the 18–35 km range is the best empty space for long-term residence of the platform.

It is well-known that the number of user-visible aerostat in satellite navigation must be not less than 4, so that the effective positioning can be achieved. Therefore, the basic unit configuration starts from 4 nodes, such as square, rectangle, Y-type and rhombus, as Fig. 1. Configurations are the projection of double-layer configurations on the ground, which is 3 bottoms, 1 top in Fig. 1.

2.2 Single Configuration Location Analysis

Taking the Y-type as an example to do the configuration analysis. The target center is (39°N, 116°E), the east-west distance and north-south distance of the target area are both 20 km. The structure is two layers (three vertices are located at 20 km height, and the Y center is at 25 km). The simulation result is shown in Fig. 2.

From Fig. 2, the longer the projection length on the ground between the vertices in the Y-type configuration is, the smaller the PDOP (position dilution of precision) maximum value for all the users in the same target area is. So under the constraint of Y-type, 6° elevation and 4-star positioning, the projection length takes 168.5 km

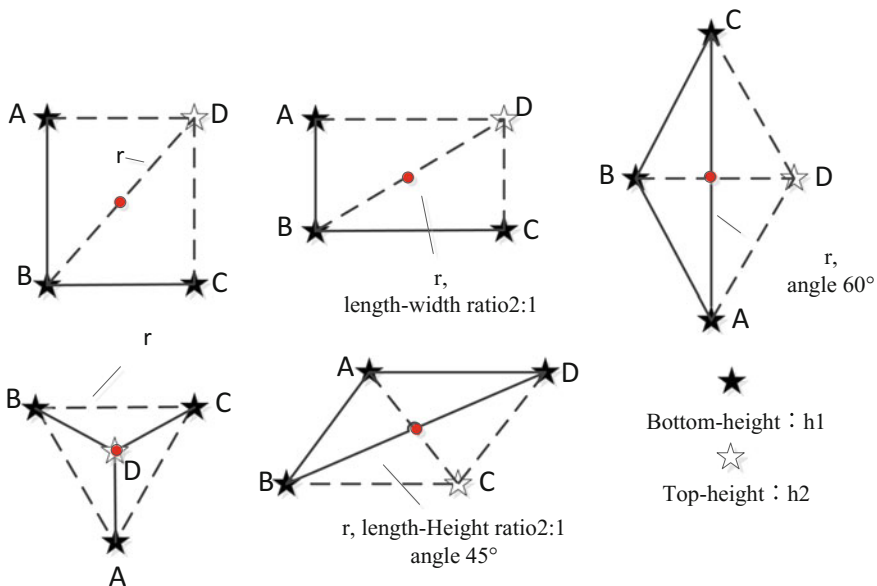
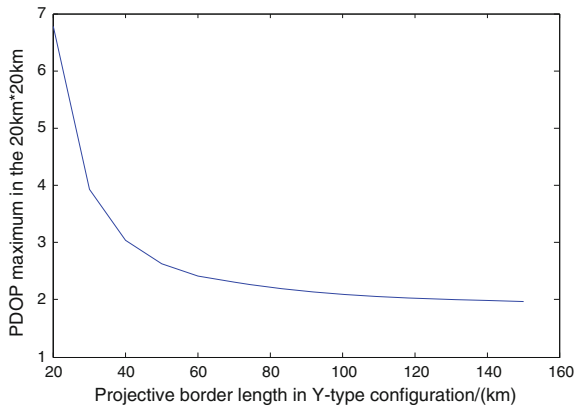


Fig. 1 Different unit configurations

Fig. 2 The influence of different side length on the region



(denoted as r_{max}), which is the coverage radius of the near space platform under the ground level, from Table 1.

Similarly, the Y-type is selected as the basic unit configuration, and the comparison among single-layer 20, 25 km, double-layer 1 with center at 20 km and apex at 25 km, double layer 2 with center at 25 km and apex at 20 km, the users' PDOP is shown in Table 2. The $PDOP_{aver}$ in the table is PDOP average for all users in the target area, and the rest are PDOP intervals, with the target area of 38°N–40°N, 115°E–117°E. The total number of users are 160,801.

As a result, the double-layer 2 has better performance indexes than other configuration in the unit layout. That is, the height of the vertex is h_1 , the height of the center is h_2 , and the mode of $h_1 < h_2$ can satisfy the unit optimality as the basic unit configuration skeleton.

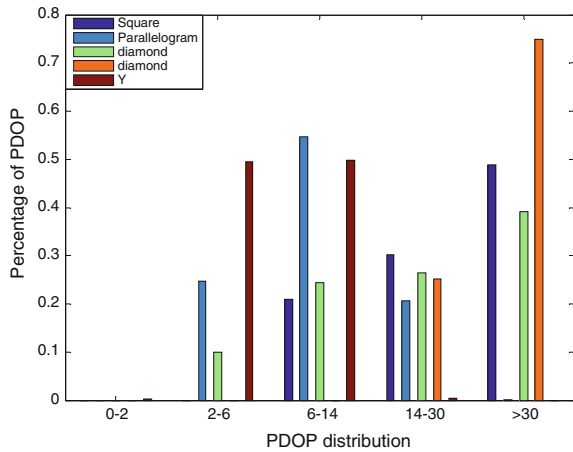
2.3 Comparison of Different Unit Configurations

An analysis based on the influence of different unit configurations on the positioning accuracy can lead to an optimal unit configuration. We select the square, rectangle, rhombus, parallelogram, Y-type and other common unit layout to do comparative analysis, which are shown in Fig. 1. According to the location analysis of the previous unit configuration, the optimal positions of the various

Table 2 Comparison of different layers under the Y unit

Configuration	$PDOP_{aver}$	(0,2]	(2,6]	(6,14]	(14,30]	(30,100)
Single-layer 20 km	8.8152	278	55,478	76,580	28,465	0
Single-layer 25 km	6.7314	332	75,940	83,687	842	0
Double-layer 1	9.64	217	53,284	68,322	38,978	0
Double-layer 2	6.5423	425	79,540	80,105	731	0

Fig. 3 Percentage of PDOP under different unit configurations



configurations at the same height are selected, that is, the distance between the aerostats is the longest. As the basis for evaluation, there is same target area of 38°N–40°N, 115°E–117°E. The total number of users is 160,801. The result of PDOP percentage is shown in Fig. 3.

Summarizing the simulation results, Y-type unit configuration has obvious advantages in performance index compared with other four configurations. Therefore, the Y-type configuration (height of vertex is h_1 , height of center is h_2 , height of $h_1 < h_2$) is selected as basic unit configuration to meet the optimal unit layout. The distribution of PDOP contour in the target area is shown in Fig. 4.

By comparing the single-layer and double-layer of the unit configuration, we get the result that double layer is better than single layer, and then the best unit configuration is obtained by comparing the best coverage of different unit configurations (double layer) type, it's Y-type.

3 Network Extension and Boundary Detection Method

3.1 Navigation Network Model Description

According to the simulation results in Sect. 2, the unit configuration can achieve the best navigation effect. We can use Y-type as the basic unit configuration, and then expand the network from the directions of latitude and longitude, taking the network node multiplexing into account. So the model only need to add two near space platform when expanding one basic unit. On one hand, it ensures the uniform distribution of nodes, on the other hand, reduces the layout cost. Figure 5 shows the unit configuration and extension method.

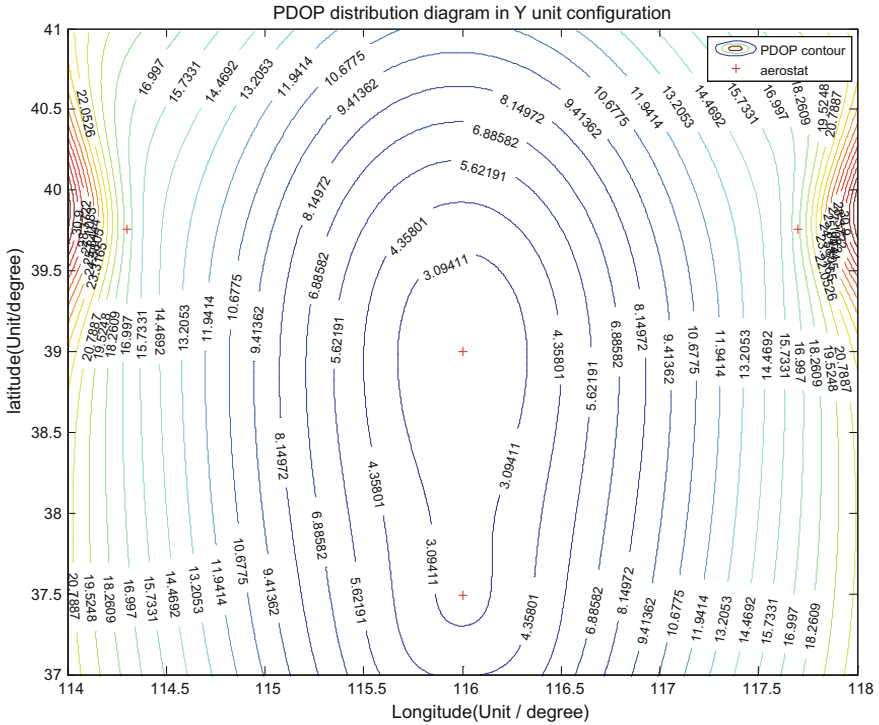


Fig. 4 PDOP contour map with Y configuration

When expanding unit layout, there will be edge redundancy that can affect the overall cost. So further optimization is needed, there is a way of reducing excess aerostats in the platform edge.

This application scene is that no matter how large the target area is, a stable and optimal near-space network configuration can be quickly provided, which can meet the high-precision positioning requirements of all users in the target area. Among the traditional configuration designs, the selected optimization indexes are: GDOP (Geometric Dilution of Precision) average, dispersion degree s and 4-star common view area S in the user area [5]. In this paper, the common view area S was deleted in the network evaluation indexes. Moreover, PDOP instead of GDOP is used. And new parameters are added, they are Num and “one vote” decision index Y/N . Num means the number of near space platform, the latter is whether to meet the needs of users. So there are 5 indexes: the target area PDOP average, PDOP mean square deviation, PDOP maximum, the total number of platforms used in the network, whether it meets the high accuracy and seamless coverage requirement. The indicators are characterized as follows:

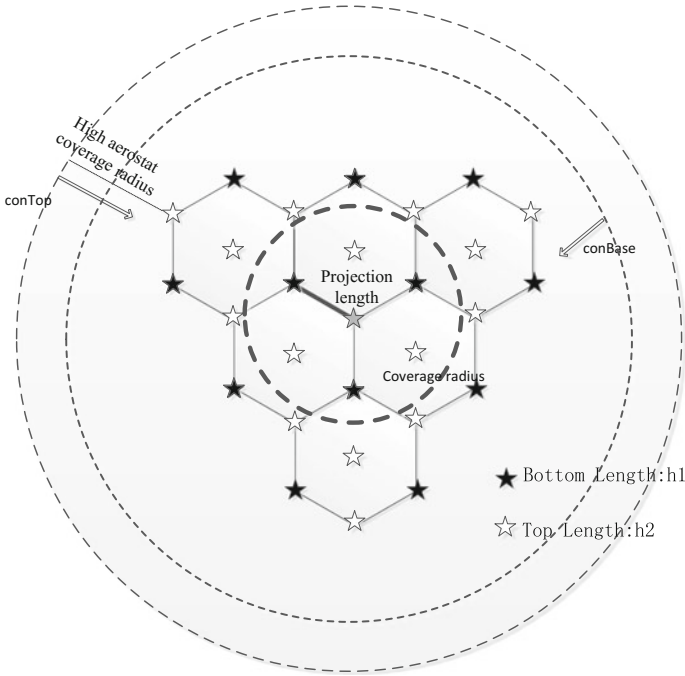


Fig. 5 Schematic diagram of configuration parameters of near space platform

$PDOP_{aver}$ reflects the pros and cons of network positioning performance from the average point of view. The definition is shown as follows:

$$PDOP_{aver} = \sum_{i=1}^n PDOP_i/n \tag{2}$$

In the formula: $i = 1, 2 \dots N$; n users are evenly distributed.

The dispersion s is the mean square deviation of all users' PDOP, which can measure the fluctuation degree of the location accuracy factor on the near space navigation network. The smaller the fluctuation is, the more stable the service quality of the target area is.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (PDOP_i - PDOP_{aver})^2} \tag{3}$$

$$PDOP_{max} = \max(PDOP_i) \tag{4}$$

Formula (4) is the maximum of all users' PDOP in the target area. Num is the total number of platforms, and $N_i \geq 4$. The "one vote" decision index Y/N takes the high precision requirement as the criterion, for example, the requirement in this

paper is $\text{PDOP} \leq 4$ for all users, if it's satisfied, the value is 1; on the contrary, it is recorded as 0.

Target area parameters: the center latitude and longitude of the target area, the distance in the latitude and longitude direction.

Network configuration parameters: the height of the platform in the near space, the length r of the projection on ground between the platforms, the overall contraction length at the bottom is conBase , the value range is $[0, r_{\text{hBase}}]$, the top contraction length is conTop , The value range is $[0, r_{\text{Top}}]$. The parameters are shown in Fig. 5.

3.2 Boundary Detection Methods

3.2.1 Adaptive Detection Method

Because of the negative correlation between the three network configuration parameters and the evaluation parameter, the greater the parameters is, the lower the evaluation parameter is, and the better the effect is.

The adaptive detection algorithm is described as follows:

- (1) When conBase and conTop are 0, the maximum projection edge length r_{max} is established.
- (2) $\text{conTop}_{\text{max}}$ is determined by r_{max} and $\text{conBase}_{\text{min}}$, and $\text{conBase}_{\text{min}}$ is updated on this basis; $\text{conBase}_{\text{max}}$ and $\text{conTop}_{\text{min}}$ are determined conversely; the further boundary ranges of the three networking parameters are obtained.
- (3) Randomly generating a set of values (a, b, c) (a represents the projection side length r , b represents the conBase , and c represents the conTop) within the bounds of the boundary. If the judgment index is satisfied, the current comprehensive evaluation function value is recorded. Since a, b and c are inversely proportional to the comprehensive evaluation function, a increases the random Δa , and if it still satisfies the decision index and the function value decreases, it is updated; on the contrary a reduces Δa , if satisfied, then continue to increase upward.

The program exits the loop when the value of the function cannot be reduced any further for three consecutive changes. Similarly, b, c obtains the final comprehensive evaluation function value according to the same logic, which are the optimized values of network configuration parameters.

3.2.2 Intelligent Detection Algorithm

There are many types of intelligent algorithms. Greedy algorithm and genetic algorithm are applied to the actual detection scenes in the paper. The specific

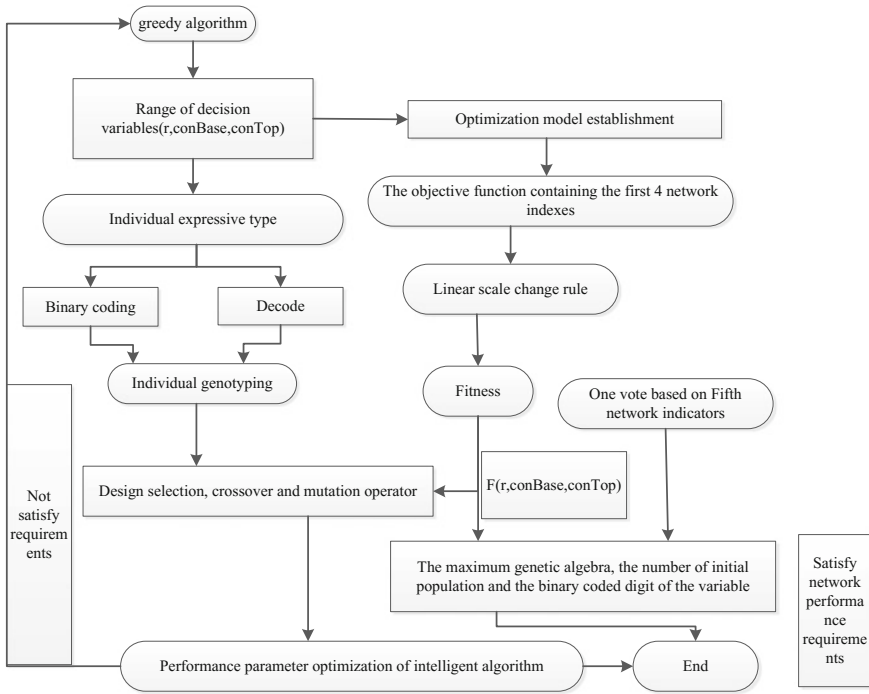


Fig. 6 Intelligent algorithm optimization process

features of the algorithm are not described in detail here, and Fig. 6 shows the optimization flow chart.

The fitness function is shown as follows:

$$F = f(r, conBase, conTop) \tag{5}$$

In the formula:

$$F = \sum_{j=1}^4 f_1 \times \alpha_j, \sum_{j=1}^4 \alpha_j = 1 \tag{6}$$

Under the same target area, the configuration parameters are calculated according to two algorithms. Taking the optimal solution can not only make up for the lack of randomness in the adaptive algorithm, but also improve the local optimization that may fall into the intelligent algorithm. This is the optimal design of the algorithm.

3.3 Algorithm Simulation Analysis

The two algorithms introduced in the previous section are based on the optimization of multi-objective and multi-parameter algorithms, so as to establish the network configuration. The simulation examples of the algorithm are as follows:

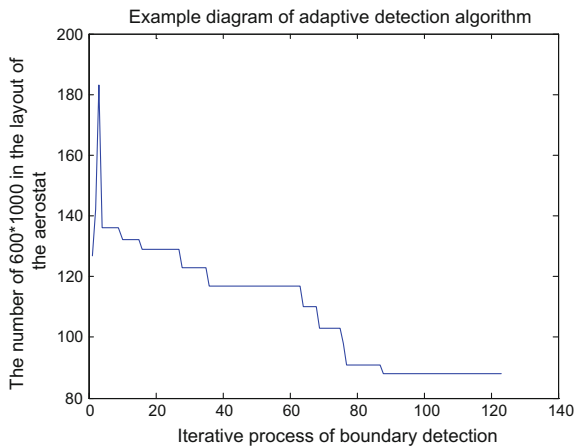
The center of the service area is (39°N, 116°E). The near space platform expands to the four sides with the Y-type as the unit configuration, the double layers' height is 25 and 20 km. Table 3 shows different simulation results which were obtained through the intelligent detection algorithm and the adaptive algorithm.

The data in the table shows that when the service area is within a certain size (e.g. 20 km * 384 km and 320 km * 320 km), two algorithms have the same networking results. However, when the service area is larger, the adaptive detection algorithm can play a more significant role. In terms of computing time, the intelligent detection algorithm has more efficient. So we can select the detection algorithm according to different target area. When the target area is more than 320 km * 320 km, we suggest using adaptive detection algorithm. On the contrary, the intelligent detection algorithm is more efficient. The specific optimization process of adaptive detection algorithm is shown in Fig. 7.

Table 3 Different simulation results between two algorithms

Service area		Adaptive algorithm				Intelligent algorithm			
		Parameters			Index	Parameters			Index
E-W	N-S	r	conBase	conTop	Num	r	conBase	conTop	Num
20	384	149.9	168.5	61.7	10	146.3	95.9	36.4	10
600	1000	106.7	168.5	102.7	88	102.4	145.4	91.7	91
320	320	94.2	168.5	102.8	28	106.5	165.8	75.3	28

Fig. 7 Example diagram of adaptive detection algorithm



This chapter introduces an extended model based on the optimal unit configuration, which can achieve the specific scope of the near space network configuration, and meet the seamless coverage requirements of the target goal. At the same time, we choose the intelligent algorithm represented by greed and genetic algorithm, and adaptive algorithm to solve the problem of layout boundary detection. The results show that the algorithm can achieve synchronous optimization of networking cost and accuracy.

4 Dynamic Network Design and Optimization

4.1 Working Principle of Super-Pressure Balloon

The wind layer above and below the zero wind layer is the opposite of the zonal component, so the flight direction can be changed by active control in the direction of the earth's weft line. Along the direction of the earth's meridian, there is a north-south component of the meridional direction of different wind layers, and the wind speed and wind direction are different. Therefore, the trajectory of the balloon can be controlled to a certain extent by a small range of height, that is, the balloon's latitude remains relatively constant [6].

4.2 Dynamic Network Design

According to the working state of the super pressure balloon in the quasi zero wind layer, the abstract of the mathematical model is carried out and the dynamic path of the network is designed. The movement process is described as: As for the bottom of aerostat, running at $v = 5$ m/s from west to the east in the weft direction, when the distance boundary only 2.5 km, with $a = 0.005$ m/s² acceleration deceleration, and acceleration of a in the vertical direction, when running on the height to the height of 2.5 km, the two direction acceleration steering, through calculation the increase in the total time of 2000s. As for the top of the aerostat, the movement process is just the opposite, the entire network formed "conveyor belt" mode, the trajectory is shown in Fig. 8.

The setting of the boundary is particularly important. First of all, the aerostat in the network must return to the starting position on a "conveyor belt" cycle, need to meet the same operating mileage of aerostat in different dimensions of a cycle, so the boundary is set in km instead of degree. So the extreme value of the longitude in different latitude in the initial layout is counted. Due to different longitude differences in different latitudes, the actual distances represented by longitude differences in latitude need to be calculated. Choosing the maximum actual distance among them as the length of "conveyor belt" The distance difference is converted into the

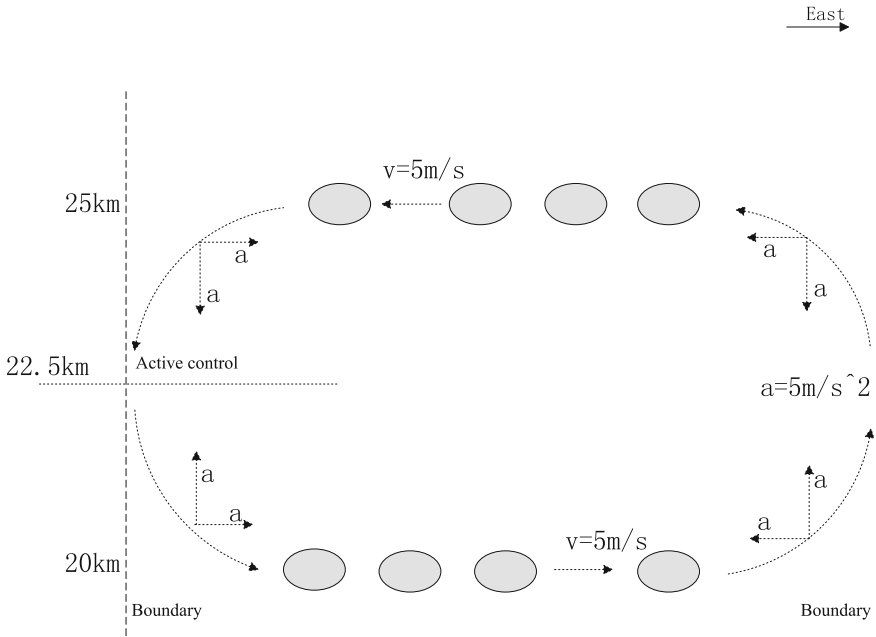


Fig. 8 Platform dynamic trajectory simulation diagram

longitude range corresponding to each latitude circle according to the longitude of the center, which is the boundary value.

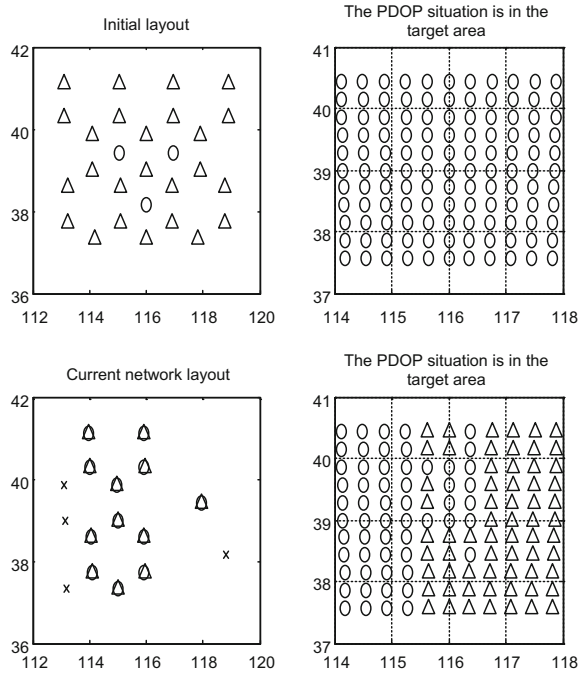
This design has the following characteristics:

1. The model of the working principle of super pressure balloon provides a feasible solution to solve the dynamic running process of the platform.
2. Comparing with the mode that each aerostat are relatively static upper and lower circular motion, it saves the energy to control the high energy consumption. The entire dynamic process is not only in line with the actual, but also maximizing the extension of the hovering time, reducing control costs.

4.3 Dynamic Optimization

The simulation experiment is carried out according to the model of the dynamic trajectory. The center of the service area is (39°N, 116°E). The near space platform expands to the four sides with the Y-type as the unit configuration, the double layers height is 25 and 20 km. The target area is 320 km * 320 km, and the target precision is user PDOP ≤ 4 . The static network configuration in Fig. 9 is shown on the left (Δ represents top, \circ represents bottom), On the right is the PDOP of the

Fig. 9 Missing layout in initial and dynamic



122 users in the target area under the initial layout. The lower left indicates a layout when a partial user PDOP does not meet the positioning requirements in the 50,000s moment, and the lower right is the PDOP situation of the user at that time.

Through the above simulation, it is found that the dynamic path design with the best static layout cannot meet the accuracy requirement, some or even all of the users fail to meet the PDOP requirement, as shown in Fig. 9. Based on this initial layout, the proportion of user PDOP in several dynamic process cycles is shown in Fig. 10. The reason is that the worst configuration happens is top or bottom layer occur coincidence in the process.

As for worse PDOP in the dynamic process, the adaptive algorithm is adapt to this situation. Doing optimization for the worst case of a dynamic process, the case is like the bottom time in the Fig. 10. The simulation is shown in Fig. 11, and dynamic optimal configuration is shown in Fig. 12.

The improve precision between static and dynamic configuration is shown in Fig. 13. After dynamic optimization, it's changed that the layout parameters from (94.2, 168.5, 102.6) to (57.1, 84.3, 164.4), and the aerostat number is increased from 28 to 72. The result meets the requirements of the user's positioning throughout the dynamic process, which is shown in Fig. 14.

The mathematical model of the super-pressure balloon is established according to its movement in the quasi-zero wind. Therefore, a dynamic trajectory of the near space platform was designed, and the dynamic optimization by adaptive algorithm was proposed at the lack of users' PDOP when the static optimal layout is in the

Fig. 10 User PDOP ratio in the dynamic process of 320 * 320 km service area

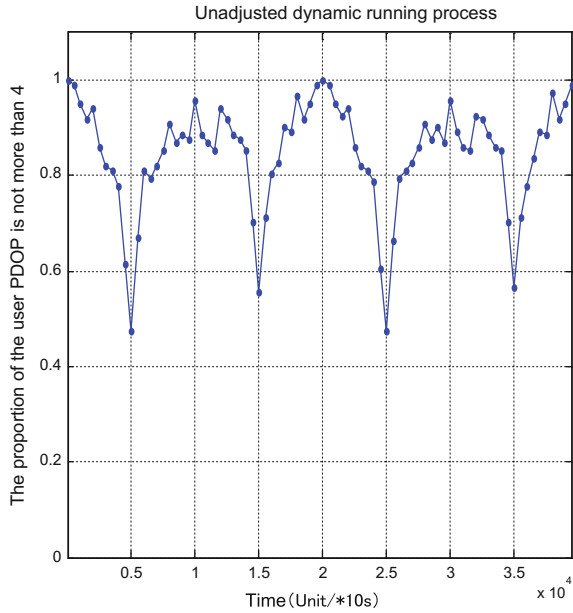
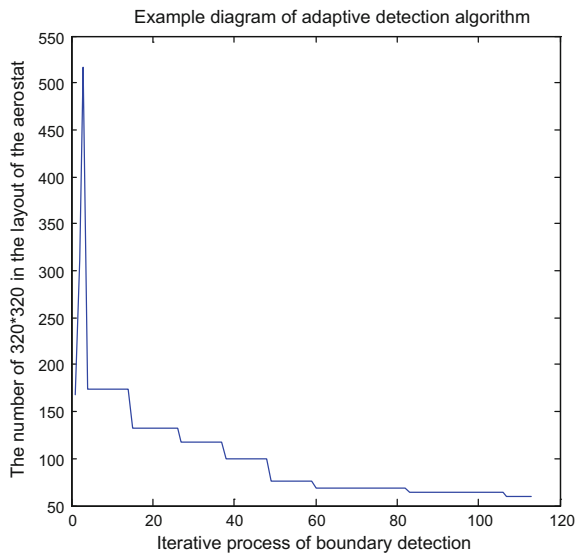


Fig. 11 The process of dynamic optimization



dynamic process. At the same time, the significant increase of the aerostat number is worth discussing in depth. Altogether, the model is adapt to the environment of quasi-zero wind, and also makes the near space navigation network configuration more realistic.

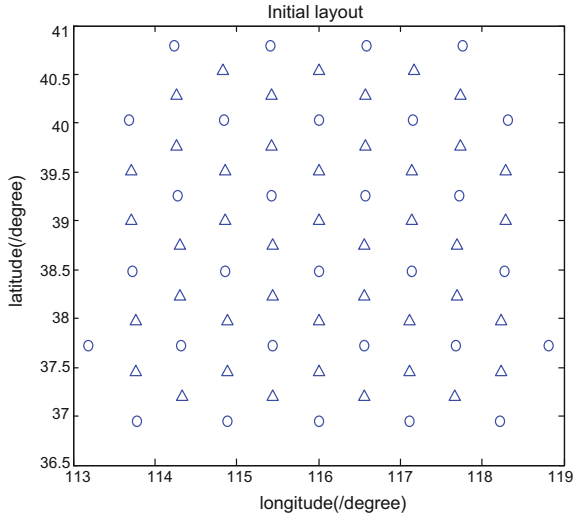


Fig. 12 Configuration after dynamic optimization

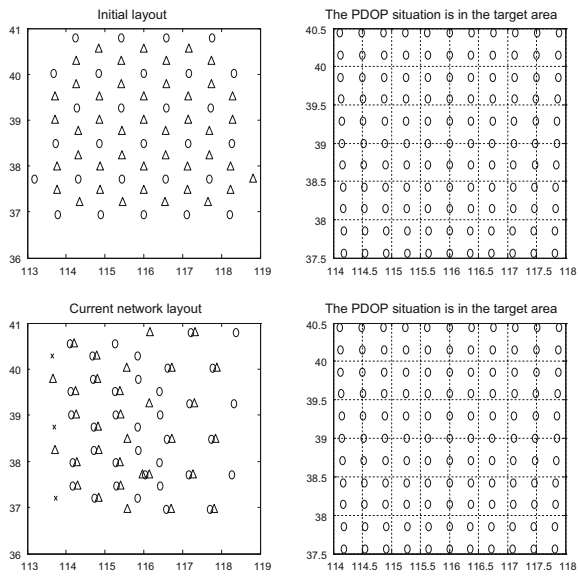
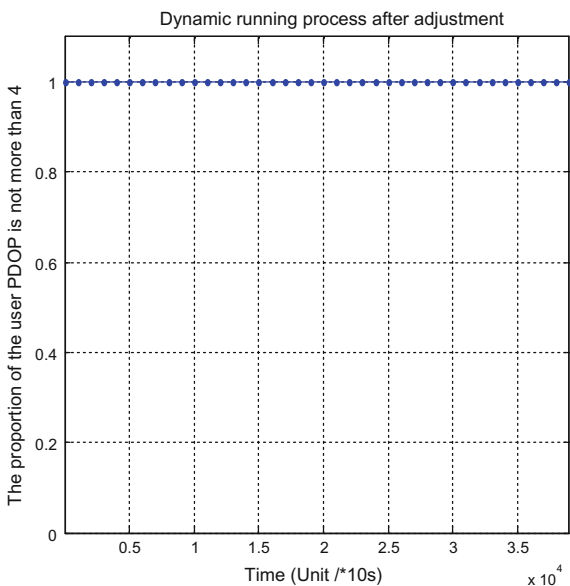


Fig. 13 Comparison after dynamic optimization

Fig. 14 The result of dynamic optimization



5 Conclusions

In this paper, we have proposed adaptive algorithm and intelligent algorithm to detect and optimize the edge of the network platform, and designed the dynamic trajectory and optimization of the platform which is adapt to the actual environment. On this basis, we have completed a new optimization design of platform network based on near space. The design is static and dynamic coordination of the actual scene layout design without rigidly adhere to the traditional geometric configuration model.

The experimental results show the following points. The configuration method can effectively solve the problem that satellite navigation signal cannot be received from complex terrains. It provides high-precision navigation accuracy with different PDOP requirements for target areas. This paper contributes to the research on near space independent network navigation, and it has formed the quantitative configuration result, which can effectively improve the utilization of the near space.

Acknowledgements This work is supported by the National Natural Science Foundation of China (No. 91438207).

References

1. Yang C, Qian Q, Wang F, Sun M (2018) Application of improved adaptive genetic algorithm in function optimization. *Appl Res Comput* (04)
2. Jin C, Qiu Q, Peng W (2016) Trajectory analysis and network performance simulation of near space vehicle. *Chin J Aeronaut* 37(S1):134–138
3. Malleswaran M, Deborah SA, Manjula S (2010) Integration of INS and GPS using radial basis function neural networks for vehicular navigation. *Contra Autom Robot Vis Singap* 7–10:2427–2430
4. Shi H (2014) Research on enhancement network optimization and precise timing synchronization for air-borne pseudo-satellites of Beidou. *Nanjing University of Aeronautics and Astronautics*
5. Tan L (2013) Research on pseudolite system for BeiDou based on independent and dynamic aircrafts configuration. *Nanjing University of Aeronautics and Astronautics*
6. Zong R, Gao X, Wang X, Lv Z (2012) Deployment of high altitude platforms network: a game theoretic approach. In: *Proceedings of IEEE international conference on computing, networking and communications*, Maui, Hawaii, USA

Research on Ionospheric Grid Correction Strategy Based on Chinese Region



Kewu Chang, Rui Li and Haipeng Li

Abstract The estimate of Grid Ionospheric Vertical Delay (GIVD) and Grid Ionospheric Vertical Delay Error (GIVE) is the core content of ionospheric delay processing section in Satellite Based Augmentation Systems (SBAS). At present, WAAS and EGNOS respectively use Kriging method and Inverse Distance Weighting (IDW) method to estimate GIVD and GIVE. The existence of a fixed ionospheric abnormal zone at low latitudes south of the Yangtze River in China led to the advantages and disadvantages of the two algorithms in different regions of China. Aiming at this problem, this paper takes two mature methods of ionospheric grid correction as a reference, combined with the distribution characteristics of ionosphere in China, selected the measured data in China for simulation analysis. According to the comparative analysis of the simulation performance of the algorithm, this paper proposes an algorithm combination of strategies. This paper selects the measured data of 31 monitoring stations and 30 user stations from Crustal Movement Observation Network of China (CMONOC) for verification. Firstly, the spatial correlation characteristics of the ionosphere over China are analyzed and the applicability of the two algorithms to China is adjusted. Secondly, we compare the performance of the two mature algorithms. Finally, based on the comparative analysis, combined with the theoretical analysis of the use of two algorithms, this paper puts forward the combination strategy. In the area north of the Yangtze River, the ionospheric delay error model is stable and Kriging method is adopted to obtain a larger service area and better calibration performance. At the same time, in the area south of the Yangtze River, in order to counter the frequent occurrence of ionospheric abnormal, IDW method was chosen to ensure the integrity. The results

K. Chang
Beijing Institute of Tracking and Communication Technology, Beijing, China

R. Li (✉)
Collaborative Innovation Center of Geospatial Technology, Wuhan, China
e-mail: lee_ruin@263.net

R. Li · H. Li
Beihang University, Beijing, China

show that compared with the IDW method alone, the combination accuracy is improved and the service area in the marginal area is expanded. Compared with the single plane Kriging method, the combined method ensures the integrity of the ionosphere active area in the south of the Yangtze River.

Keywords Ionospheric delay · Algorithm combination strategy
Grid correction · China region

1 Introduction

Ionospheric delay is one of the main error sources of satellite navigation and positioning. It is an important factor that restricts the positioning accuracy of single-frequency users. It is also one of the major problems in the process of building a Beidou Satellite Based Augmentation System (BDSBAS). The estimation of grid point correction is the core part of the ionospheric delay processing at the satellite-based augmentation system. At present, the WAAS system in the United States and the European EGNOS system respectively use the planar Kriging method and the distance reciprocal weighting method to estimate the vertical ionospheric delay correction values and the confidence limits [1]. The fundamental of the grid ionospheric correction method is to make the ionosphere equivalent to a fixed height of the thin layer, the master station using the ground reference station dual-frequency receivers real-time collection of ionospheric delay measurements, estimated coverage area within the latitude and longitude grid point Ionospheric vertical delay correction and correction accuracy, i.e. GIVD and GIVE, moreover correctional data are transmitted on L1 through Geostationary satellites (GEO) to user. The user interpolates IPP ionospheric delay values and confidence limits using 3 or 4 surrounding mesh points based on where the Ionospheric Pierce Point (IPP) is located.

The existence of a fixed ionospheric abnormal zone at low latitudes south of the Yangtze River in China led to the advantages and disadvantages of the two algorithms in different regions of China. Aiming at this problem, this paper takes two mature methods of ionospheric grid correction as a reference, combined with the distribution characteristics of ionosphere in China, selected the measured data in China for simulation analysis. According to the comparative analysis of the simulation performance of the algorithm, this paper proposes an algorithm combination of strategies. The simulation results show that compared with the method of using an algorithm alone, the combination strategy can improve the correction performance under the premise of ensuring the integrity, which provides a reference for the ionospheric delay error grid correction processing in the BDSBAS system.

2 Ionospheric Delay Estimation

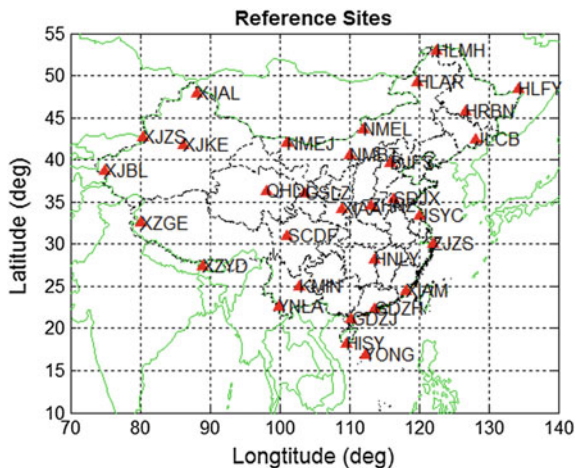
The ionosphere is a region of gaseous ionization (free electrons and ions) at a height of 50–1000 km from the ground. For the propagation of electromagnetic signals, the ionosphere belongs to the dispersion medium. The amount of delay caused by different frequency carrier signals crossing the ionosphere is different. Based on this principle, dual frequency observations can be used to extract the ionospheric delay. In order to facilitate the modeling analysis, SBAS system simplifies the complex three-dimensional ionospheric model to a two-dimensional ionospheric shell model at a fixed height above the ground (350 km in this paper). Based on the ionospheric shell model, the line-of-sight ionospheric delay on the signal propagation path can be converted to a zenith vertical ionospheric delay. According to the provisions of RTCA DO-229E [2], combined with BDSBAS service area considerations, China’s ionosphere grid coverage is E70°–E140°, N10°–55°, the IGP’s are divided by latitude and longitude $5^\circ \times 5^\circ$, the update cycle is 5 min. In order to estimate GIVD, the ionospheric thin layer model should firstly be modeled. Therefore, this paper first studies the ionospheric delay characteristics in China using the dual-frequency observations obtained from monitoring stations.

2.1 Spatial Correlation of Ionospheric Delay

The distribution of 31 monitoring stations in China is shown in the red triangle in Fig. 1, and the observed values of L1 and L2 frequencies of the GPS are obtained by dual-frequency receivers.

Seriously affected by the global ionospheric disturbance are mainly distributed in low latitudes and Polar Regions near the magnetic equator. Compared with the

Fig. 1 31 monitoring station receivers (the red triangles represent receivers)



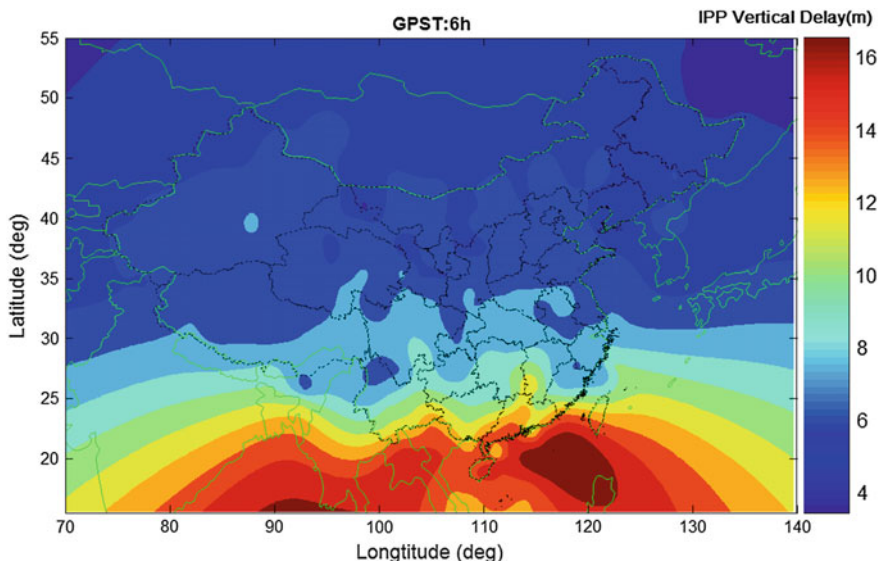


Fig. 2 IPP vertical delay (m) in Chinese Region (March 1, 2015, GPST: 6 h)

United States and Europe and other regions, the ionospheric environment in China has obvious regional characteristics, as shown in Fig. 2. The changes in ionosphere in northern China are relatively stable. The southern China is at low latitudes. The influence of ionospheric abnormal is broader and more frequent, which seriously affects the application performance of GNSS in south of the Yangtze River in China.

2.2 IDW Algorithm

2.2.1 Fundamental

According to the IPPs provided by the monitoring station, the inverse of the distance between the IPP and the IGP is taken as the weight, and the ionospheric delay correction value of the IGP is obtained through weighted interpolation [3]. Secondly, we collect the observation sequence of IPPs in a renewal cycle and calculate the GIVE.

2.2.2 Grid Ionospheric Vertical Delay Estimation

Obtain N sets of IPPs information of all stations and stars at time t , including IPP latitude and longitude $(\varphi_{IPP,i}, \lambda_{IPP,i})$, IPP vertical delay $I_{IPPi,v}(t)$, $1 \leq i \leq N$. For a

Fig. 3 IGP and IPP

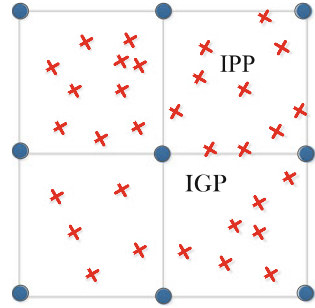
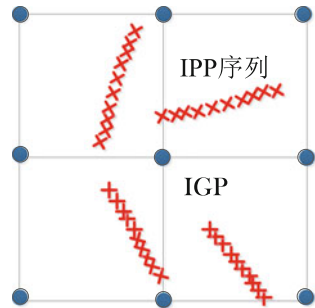


Fig. 4 IPP sequences around IGP



certain IGP j , take the IPPs in the four grids around the grid point as shown in Fig. 3..

Using the vertical delay at the N IPPs, the grid point ionospheric vertical delay correction ($\hat{I}_{IGPj,v}$) is obtained by interpolating from the inverse of the distance weighting (Fig. 4). The formula is as follows:

$$\hat{I}_{IGPj,v} = \sum_{i=1}^N w_i \cdot I_{IPPi,v} \tag{1}$$

The weight of the i th IPP is related to the direct distance from IPP to IGP: $w_i = \frac{\frac{1}{d_i}}{\sum_{k=1}^N \frac{1}{d_k}}$, d_i represents the distance from the i th IPP to the IGP to be sought. In order to ensure the continuity of interpolation, Klobuchar model is introduced into the calculation for adjustment. At this time, the vertical delay estimation value of the grid point is calculated as follows:

$$\hat{I}_{IGPj,v} = \sum_{i=1}^n \left(\frac{M_{IGP}}{M_i} \right) \cdot \frac{\frac{1}{d_i}}{\sum_{k=1}^n \frac{1}{d_k}} \cdot I_{IPPi,v} \tag{2}$$

M_{IGP} and M_i are the Klobuchar model values (vertical direction) at the grid point and the penetration point, respectively, and the eight model parameters required can be extracted from the current navigation message.

2.2.3 Grid Ionospheric Vertical Delay Error Estimation

Assuming that the sampling period of the observed data is 30 s, update period of the central station broadcasts the GIVD and GIVE to the user is $T_0 = 300$ s, In an update cycle, several groups of 10 IPP groups are formed around each IGP. As shown in the figure below, only four sequences are drawn here. In the actual case, the number of sequences is closely related to the number of IPPs around the grid point of the current cycle.

For IPP corresponding to time t in sequence i , two vertical ionospheric delay values are obtained: Smooth delay calculated by the dual-frequency method and the vertical delay estimation value of the IPP at time t $\hat{I}_{IPPi,v}(t)$ is interpolated from the four grid point IGP vertical delay correction values $\hat{I}_{IGPj,v}(t - T_0)$ in the previous update period. Then the error sequences can be expressed as formula (3):

$$e_{IPPi}(t) = I_{IPPi,v}(t) - \hat{I}_{IPPi,v}(t), t = (t_k - 270), (t_k - 240), \dots, (t_k - 30). \quad (3)$$

For each IPP error sequence satisfying the above conditions within 300 s, the error limit is constructed, as i denote the sequence number:

$$E_{IPPi}(t) = \left| \overline{e_{IPPi}(t)} \right| + \kappa(\text{Pr}) \cdot S_i \quad (4)$$

where, $\overline{e_{IPPi}(t)} = \frac{1}{m} \sum_{k=1}^m e_{IPPk}(t_k)$, $m = 10$, $S_i = \sqrt{\frac{1}{m-1} \sum_{k=1}^m (e_{IPPi}(t_k) - \overline{e_{IPPi}(t)})^2}$.

$\kappa(\text{Pr})$ Is the quantile of the 99.9% confidence probability.

The errors in all the error sequences are reciprocally weighted, resulting in an IGP absolute error of:

$$\hat{e}_{IGP(j)} = \frac{\sum_{i=1}^n \left(\frac{1}{d_{ij}} \right) \cdot |e_{IPPi}(t)|}{\sum_{k=1}^n \frac{1}{d_{kj}}} \quad (5)$$

In the formula, d_{ij} is the distance between the i th IPP and the j th IGP, and n is the IPP numbers of the adjacent grid.

In summary, get GIVE expression:

$$GIVE_j = \max(E_{IPPi}) + \hat{e}_{IGP(j)} \quad (6)$$

where, $\max(E_{IPP,j})$ is the maximum value of the error limits of all the IPP errors, and $\hat{e}_{IGP(j)}$ is the absolute error of the ionospheric delay of the grid points.

2.3 Kriging Algorithm

2.3.1 Fundamental

The essence of the Kriging method is a linear, unbiased, Best Linear unbiased Estimator (BLUE) method. Specifically, the Kriging method considers the number of sampling points, the distribution of locations, and the spatial correlation between the sampling points and the points to be estimated. In a limited area of estimation, the linear, unbiased and least mean square Error as a guideline, for each of the sampling points for estimation are given a certain weight coefficient, and finally weighted average to estimate the value of the point to be estimated [4].

2.3.2 Grid Ionospheric Vertical Delay Estimation

Suppose that there are n sampling points x_1, x_2, \dots, x_n around the ionospheric penetrating point x , the corresponding vertical delay observations for the ionosphere are $I(x_1), I(x_2), \dots, I(x_n)$, then the vertical delay estimate at the x -point is as follows:

$$\begin{aligned}
 GIVD = \hat{I}(x) &= \sum_{k=1}^n \lambda_k I(x_k) = \Lambda^T I \\
 I &= [I_1, I_2, \dots, I_n]^T, \Lambda^T = [\lambda_1, \lambda_2, \dots, \lambda_n]
 \end{aligned}
 \tag{7}$$

where, Λ is the coefficient vector. The problem of estimating is to find the optimal value under certain constraint conditions, that is, the Kriging coefficient Λ satisfies the following formula:

$$\begin{cases}
 \text{Min} & (\hat{\sigma}^2 = \Lambda^T C(\mathbf{h}) \Lambda - 2\Lambda^T C(\mathbf{h}_0) + C(0)) \\
 \text{s.t.} & G^T \Lambda = X
 \end{cases}
 \tag{8}$$

where, $\hat{\sigma}^2$ is the estimated error variance, $C(\mathbf{h})$ is a matrix of $n \times n$, representing the spatial covariance matrix between observations. $C(\mathbf{h}_0)$ is a n -dimensional vector, represents the spatial covariance vector between the estimator and the observed quantity. $C(0)$ is the spatial covariance between two observations with a distance of zero, i.e. the prior variance, independent of position x . The G^T is the observation matrix and X is the position vector:

$$G^T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ Lat_1 & Lat_2 & \dots & Lat_n \\ Long_1 & Long_2 & \dots & Long_n \end{bmatrix}, X = \begin{bmatrix} 1 \\ Lat \\ Long \end{bmatrix}$$

Using the Lagrangian method to solve the above equation, we can get the final Kriging coefficient Λ :

$$\begin{aligned} \Lambda &= (W - HG^T W)C(\mathbf{h}_0) + HX \\ H &= WG(G^T WG)^{-1}, W = C^{-1}(\mathbf{h}) \end{aligned} \quad (9)$$

So the GIVD value at the grid point can be obtained.

2.3.3 Grid Ionospheric Vertical Delay Error Estimation

In the ionospheric model, the GIVE value at the grid points represents the accuracy of the GIVD estimate. However, anomalies such as ionospheric disturbances and under sampling reduce the accuracy of GIVD estimates. This paper does not consider the IPP under-sampling conditions of the ionospheric anomaly, only the introduction of expansion model, a reasonable expansion of GIVE value to ensure the integrity of the client. Therefore, the GIVE value can be obtained from the following equation:

$$GIVE = \kappa_{99.9\%} \sigma_{GIVE}, \sigma_{GIVE}^2 = R_{irreg}^2 \hat{\sigma}_{IGP}^2 \quad (10)$$

where, $\kappa_{99.9\%}$ is the quantile of the standard normal distribution 99.9% confidence intervals. $\hat{\sigma}_{IGP}^2$ is the variance of the delay value at IGP obtained by using the covariance function during the calm period. R_{irreg}^2 is the expansion factor to overcome the inaccurate model, the specific expression is as follows:

$$R_{irreg}^2 = \alpha_n I_{meas}^T (W - WG(G^T WG)^{-1} G^T W) I_{meas} \quad (11)$$

$$P = \frac{2}{\pi} \int_{\phi=0}^{\frac{\pi}{2}} \left(K^2 \frac{\alpha_n}{\sin^2(\phi)} + 1 \right)^{-\frac{n-3}{2}} d\phi \quad (12)$$

where I_{meas} is the IPP ionospheric vertical delay observation sequence, α_n can be obtained using the integrity probability assigned to the ionosphere portion, which can be computed from Eq. (12). In Eq. (12), $P = 2.25 \times 10^{-8}$ is the probability of goodness assigned to the ionospheric error part, and $K = 5.592$ is the bilateral quantile of the standard Gaussian distribution corresponding to 1-P [5].

3 Algorithm Implementation and Results

In this paper, the dual-frequency observations of 31 monitoring stations (shown as Fig. 1) and 30 user stations are selected to compare and analyze the performance of the two mature algorithms mentioned above.

3.1 Algorithm Estimation Error Analysis

For IPPs formed by 30 user stations, the ionospheric delay value extracted from the dual-frequency observations is taken as a true value, the ionospheric retardation values calculated by the planar Kriging method and IDW method are used as the estimated value, and then the estimated error of all IPP delay values can be obtained.

In this section, we choose the data of different ionospheric disturbances to analyse the estimation error of the planar Kriging method and the IDW method. Figure 5 is a three-dimensional graph of the estimated error of March 9, 2015 (Kp-index of 2), solved by the Planar Kriging and IDW methods. Figure 6 is a side view of Fig. 5 viewed from the latitudinal direction. Figure 7 is a three-dimensional graph of the estimated error of March 18, 2015 (Kp-index of 8), solved by the Planar Kriging and IDW methods. Figure 8 is a side view of Fig. 7 viewed from the latitudinal direction.

As can be seen from the data analysis in Figs. 5 and 6, the northern errors of the two algorithms are all less than the southern ones. Basically, the demarcation line between north and south can be defined as 30°–35° along the latitude line. This article takes the Yangtze River as the dividing line. Data analysis of Figs. 7 and 8 further verify the above conclusion. Comparing Figs. 6 and 8, it can be seen that the estimation error of Plane Kriging method during ionospheric stability period is less than the IDW method, and the plane Kriging method has a larger service area in the

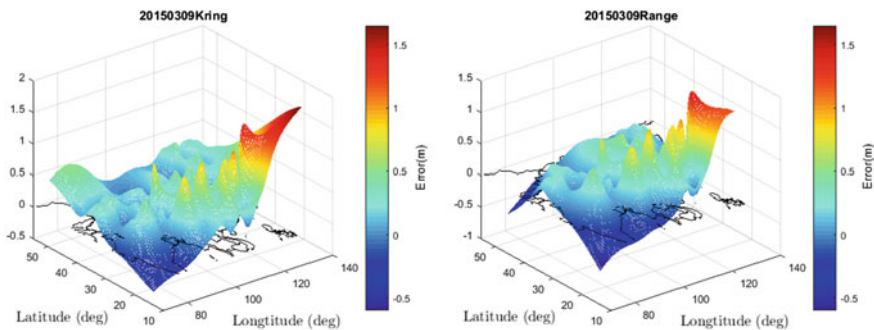


Fig. 5 March 9, 2015 estimated error Kriging (left) IDW (right)

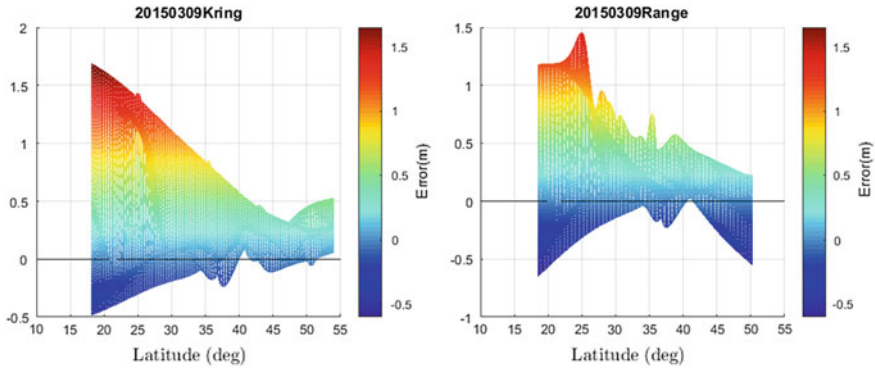


Fig. 6 The side view of Fig. 5 from the latitudinal direction

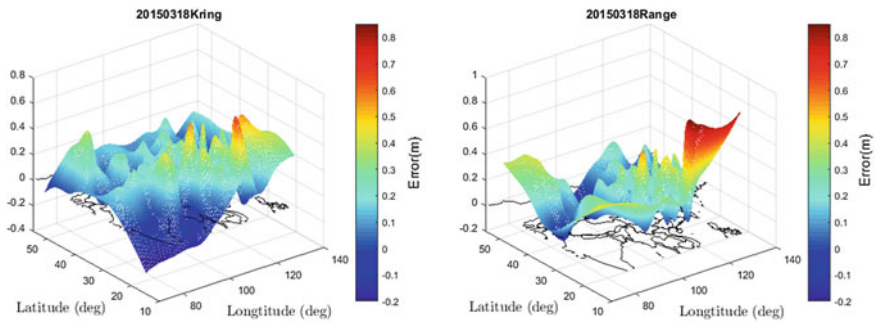


Fig. 7 March 18, 2015 estimated error Kriging (left) IDW (right)

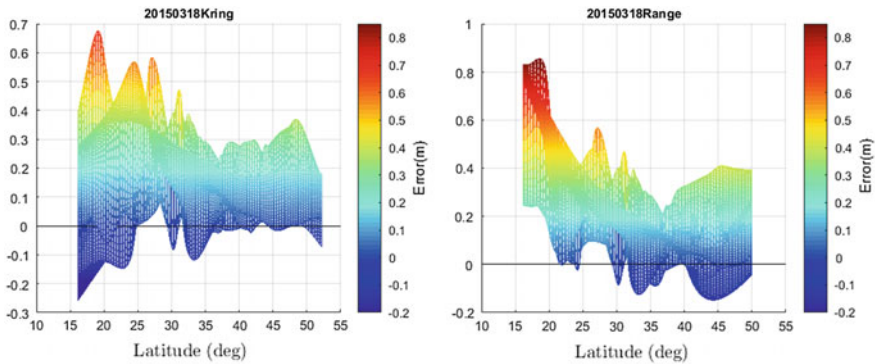


Fig. 8 The side view of Fig. 7 from the latitudinal direction

marginal. However, in the storm of ionosphere, especially in the south of the Yangtze River, the estimated error of the planar Kriging method is slightly larger than the IDW method.

3.2 Algorithm Combination Strategy

Combining with the distribution characteristics of ionosphere of China in Sect. 2 and the theoretical analysis of the two algorithms, based on the comparative analysis of the estimation errors of the two algorithms in Sect. 3.1, the algorithm combination strategy is proposed. In the area north of the Yangtze River, the ionospheric delay error model is stable and Kriging method is adopted to obtain a larger service area and better calibration performance. At the same time, in the area south of the Yangtze River, in order to counter the frequent occurrence of ionospheric abnormal, IDW method was chosen to ensure the integrity.

3.3 UIVE Envelope Analysis and Verification

We select the grids to the north of the Yangtze River and the north of the Yangtze River, respectively, and compare the UIVE envelope of the two algorithms. This section selects the grids of (N20°, E110°) and (N40°, E110°) in the lower left corner on March 18, 2015, as shown in Fig. 9. The results show that the enveloping of IDW in the south of the Yangtze River is more conservative than the Kriging method. In the north of the Yangtze River, the IDW enveloping method is slightly loosely than the Kriging method.

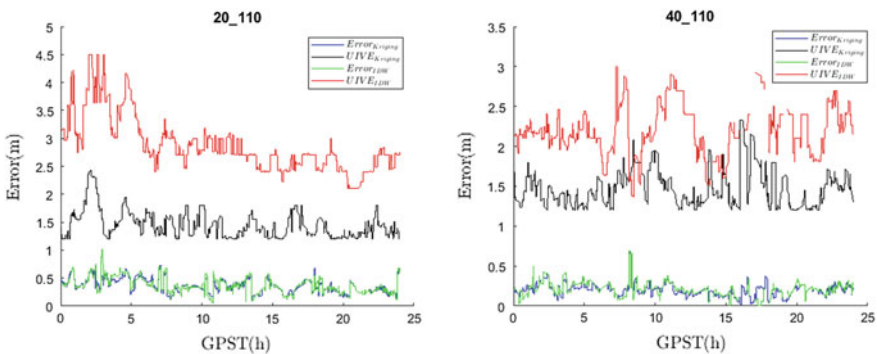


Fig. 9 Results of user ionospheric delay correction error and error bound from two algorithms (Left: south of the Yangtze River, right: north of the Yangtze River)

4 Conclusion

In view of the complicated ionospheric distribution in China, this paper takes the mature ionosphere grid correction method that has been used by the WAAS system in the United States and the European EGNOS system as a reference, combining the simulation results of the two algorithms in different regions of China, algorithm combination strategy was proposed. In the area north of the Yangtze River, the ionospheric delay error model is stable and Kriging method is adopted to obtain a larger service area and better calibration performance. At the same time, in the area south of the Yangtze River, in order to counter the frequent occurrence of ionospheric abnormal, IDW method was chosen to ensure the integrity. The results show that compared with the IDW method alone, the combination accuracy is improved and the service area in the marginal area is expanded. Compared with the single plane Kriging method, the combined method ensures the integrity of the ionosphere active area in the south of the Yangtze River. Therefore, the research work in this paper further reveals the regional distribution of ionosphere in China and the advantages and disadvantages of the two mature algorithms applied in China, and provides a more realistic view of future ionospheric delay correction grid correction algorithm in China's satellite-based augmentation system.

Acknowledgements This data are provided by the Crustal Movement Observation Network of China(CMONOC). Thanks to the CMONOC.

References

1. Sparks L, Blanch J, Pandya N (2011) Estimating ionospheric delay using kriging: 1. Methodology. *Radio Sci* 46(6)
2. RTCA Special Committee 159 (2016) Minimal operational performance standards for global positioning system/wide area augmentation system airborne equipments (RTCA DO-229E). RTCA, Washington D.C
3. Qin B (2013) Research on improvement of the grid ionospheric correction technology. Beihang University, Beijing
4. Li H, Li R, Wang Z, Gao W (2017) An improved algorithm of ionospheric grid correction based on GPS and compass multi-constellation. In: Sun J, Liu J, Yang Y, Fan S, Yu W (eds) *China Satellite Navigation Conference (CSNC) 2017 proceedings: volume II*. CSNC 2017. Lecture notes in electrical engineering, vol 438. Springer, Singapore
5. Blanch J (2003) Using Kriging to bound satellite ranging errors due to the ionosphere. Stanford University, Stanford

Satellite Integrity Autonomous Monitoring (SAIM) of BDS and Onboard Performance Evaluation



Lang Bian, Wenshan Liu, Tao Yan, Yueling Cao, Rui Li, Wei Wang,
Xiao Liu, Wenyin Lei, Yansong Meng and Lixin Zhang

Abstract With the development of satellite navigation technology, the users' demands for the integrity of GNSS are more and more intense. The ground based monitoring system can hardly report an alarm message to the GNSS users during the valid alarming period due to the satellite-earth propagation delay. It is beneficial to monitor the abnormal events and report the corresponding alarms in orbit. Through this way, which is an important branch of the future GNSS integrity monitoring, the time needed to give an alarm is shorter and the capability of system integrity service is strengthened. A new generation of BDS satellites has the capable of satellite autonomous integrity monitoring (SAIM). This paper presents the technology scheme of SAIM, and proposes the integrity monitoring method of navigation signals and clock on-board. The proposed method is verified based on the onboard test of the IGSO satellite. In addition, we analyze the integrity telemetry data from the new generation of BDS satellite, including the signal delay, power, carrier phase measurement, correlation peak, consistency of pseudo-code and carrier phase, clock phase and frequency step. The analysis results indicate that the quality of the data on orbit meets the requirements, and SAIM could monitor effectively the abnormal change of satellite clock and navigation signal, generate rapidly the alarm message, and transmit them to user. The alarm time is less than 6 s through the message, and 2 s through the non-standard code (NSC). In the end, we present the future work for improving the SAIM technology of BDS.

Keywords SAIM · Navigation signal · Satellite clock

L. Bian (✉) · W. Liu · T. Yan · X. Liu · W. Lei · Y. Meng · L. Zhang
China Academy of Space Technology (Xi'an), 504 Chang'an Avenue,
Xi'an 710100, China
e-mail: bianl35@163.com

Y. Cao
Shanghai Astronomical Observatory, 80 Nandan Road, Shanghai 200030, China

R. Li
Beihang University, 37 Xueyuan Road, Beijing 100083, China

W. Wang
Institute of Space Engineering, 26 Beiqing Road, Beijing 100094, China

1 Introduction

With the in-depth applications of global navigation satellite system in military and civilian areas, navigation integrity is demanded increasingly, especially in the field of Safety-of-Life (SoL) [1], and the user is very strict with the requirements of integrity. Therefore, integrity of the global navigation satellite system is a key requirement of upgrading the system. GPS III and GALILEO have proposed the development goal of global CAT I service [2–6].

Because of the limits of territory and other factors, the construction of global monitoring station is very difficult in China, so it is very difficult to implement the global integrity monitoring by the ground facility alone. Only rely on the monitoring station, alarm time will be very long, up to tens of minutes to several hours, and it is difficult to meet the alarm time requirement of users in the field of SoL. So, SAIM will be an effective way to solve the global integrity monitoring of BDS under the condition of limited global station distribution.

Any abnormalities in the space segment, ground segment or user segment may lead to the integrity of the abnormal events, so maintenance of the security of satellite navigation system is an extremely complex systematic work. SAIM needs to ensure that the navigation signal of the satellite is intact, and the alarm information can be sent out immediately after the discovery of the abnormal events. In previous work, CAST Xi'an had carried out the systematic SAIM technology research, designed the scheme of SAIM [7, 8], and a joint space-ground integrity monitoring system research.

The new generation BDS navigation satellite is equipped with the SAIM payload. For the first time, the autonomous integrity monitoring of navigation signal and satellite clock is carried out on orbit, and the performance evaluation and integrity alarm test are verified. This paper puts forward the technical scheme of SAIM and the integrity monitoring method of navigation signal and satellite clock, and analyzes the integrity telemetry data of the new generation BDS navigation satellite. The results show SAIM can monitor the abnormal jump of satellite clock and navigation signal on the orbit, and can quickly generate alarm information to broadcast to users. The alarm time of message is less than 6 s, and the alarm time of non-standard code (NSC) is less than 2 s. This paper finally puts forward the direction of integrity technology development of subsequent BDS system.

2 Monitoring Scheme for Autonomous Integrity of BDS Satellite

SAIM receives navigation signal from downlink transmit antenna front-end through wired link, generates pseudo range, carrier phase, signal power and correlation measurement, and monitors the time and frequency reference signal which generate navigation signals to find out satellite clock phase and frequency jump. Navigation

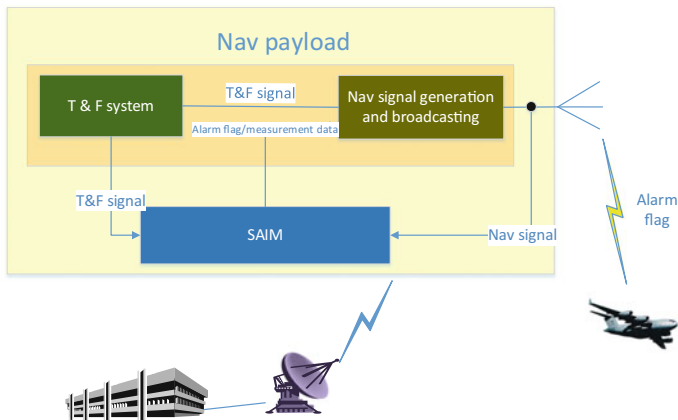


Fig. 1 The overall scheme of SAIM of BDS

signal integrity information is generated by integrated the satellite clock and navigation signal monitoring information. The signal measurement data can be analyzed by telemetry data to support further data analysis and performance evaluation (Fig. 1).

2.1 Selection of Monitoring Measurement

SAIM monitoring measurement should be selected from the user’s application requirements. User positioning accuracy can be defined by the formula (1.1):

$$\sigma_u = GDOP \times \sigma_{URE} \tag{1.1}$$

where σ_u is user position error, GDOP is Geometric Dilution of Precision, and σ_{URE} is user equivalent measurement error.

SAIM is independent of GDOP, and it is only related to the ranging error introduced by the satellite, and the ranging error introduced by the satellite is only part of total error.

The user directly uses the pseudorange and carrier phase to locate. The abnormal decreased power usually means that the payload failure on the satellite, and will cause the carrier to noise ratio (C/N0) decreased. Non ideal characteristic of payloads, such as the temperature variation, AM/AM, AM/PM or other factors, will cause signal delay variation of payload, and the pseudorange measurement system the signal correlation peak deviation; the signal correlation peak distortion can make range bias to the user according to different RF front-end filter and correlator interval measurement; baseband clock and radio frequency local oscillator (LO) frequency coherence changing will make the pseudorange and carrier phase

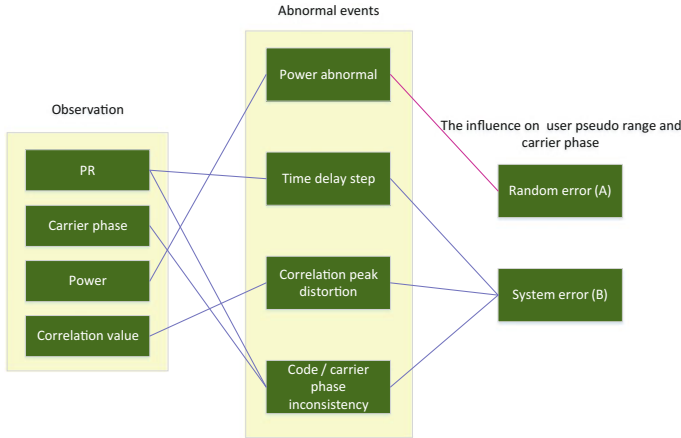


Fig. 2 The relationship between measurement, abnormal event and Influence on receiver measurement error

measurements are not consistent or divergent, influencing the precise positioning user.

The observational measurements of the navigation signals, and the relationship between the abnormal events and the impact on the user are shown in Fig. 2.

Because the reference frequency which generates all the frequency signals of the navigation satellite is provided by the satellite clock, the abnormal jump of the satellite clock will cause all the navigation signals to be abnormal.

Therefore, the measurement of integrity monitoring mainly includes:

- (1) Navigation signal: pseudo range, carrier phase, signal power and correlation value;
- (2) Satellite clock: phase jump and frequency jump.

2.2 Monitoring Process and Alarm Type

SAIM only generates alarm flag to fast changing abnormal events. The slowly changing abnormal events are mainly analyzed through the ground, so as to decide whether or not to issue alarm flag.

Integrity monitoring alarm is mainly decided by the abnormal of navigation signal and clock. If it is found that the satellite clock phase or frequency abnormal jump, and exceeds the preset alarm threshold, all signals are warning; if the delay or power of a branch signal abnormal jump, the branch signal alarm. At the same time, the satellite will transmit the signal measurement information such as pseudo range, carrier phase, power and correlation value to the ground. The ground can be further analyzed and processed, such as long-term monitoring of code/carrier phase

consistency, distortion evaluation of zero crossing of S curve caused by correlation peak distortion.

There are two ways to alert:

- (1) Information alarm: embed the generated alarm into the navigation message and broadcast to the user along with other navigation data.
- (2) Signal alarm: the navigation signal spread spectrum code is set to non-standard code.

The alarm time of the information alarm is mainly limited to the broadcast cycle of the navigation message, and the non-standard code alarm time is shorter, within 2 s.

3 Navigation Signal Monitoring Technology

The navigation signal monitoring range includes signal power, signal delay, signal correlation peak distortion, code/carrier phase inconsistency.

3.1 Signal Power Monitoring

The decrease of signal power caused by various reasons, such as digital DA or amplifier makes the C/N_0 of user receiver decrease. In ideal case, signal power fluctuates around the mean within a certain range, usually less than 0.5 dB. If the signal power is abnormal, the power measurement value will produce abnormal jump. And the power measurement value can be obtained from (1.2):

$$PW = 10 \times \log_{10}(I^2 + Q^2) + \Delta PW \quad (1.2)$$

PW represents the signal power. I represents the correlation value of the same phase branch received by the signal. Q represents the correlation value of the quadrature branch received by the signal, ΔPW is a constant for the calibration of the power zero value.

If the power generates an abnormal jump, it will be monitored immediately. And if the default threshold is exceeded, the alarm information will be generated.

3.2 Signal Delay Monitoring

The signal time delay anomaly observed by users is a pseudo-range abnormal jump, which may be caused by pseudo code fault, instruction anomaly or other payload faults. There is no dynamic on-board, so the pseudo range of the monitoring

receiver in normal condition should be the Gauss distribution around a certain range. Because the frequency source of the receiver is different from the satellite clock, the pseudo range measurement shows the phenomenon of linear drift. Therefore, clock error compensation should be made in the process of signal delay measurement. The actual signal time delay measurement values are obtained from the following formula:

$$TDelay = PR + \Delta T \quad (1.3)$$

$TDelay$ is signal delay, PR is the pseudo range measurement, and ΔT is the clock difference between satellite clock and local oscillator.

3.3 *Abnormal Monitoring of Signal Correlation Peak Distortion*

Two methods can be used to monitor the correlation peak distortion. One is in pseudorange domain monitor, which uses a plurality of receiving channels, with different receiver correlator spacing, then monitoring the dispersion of pseudorange; the second is using a pair of correlator for loop to track the signal, and other correlation value is carried out by multicorrelator, The first one can get the pseudo range value directly, which can directly analyse the influence of the correlation peak distortion on the user ranging. But, the resource consumption is too large when the multiple signals are received. So, the second is the preferred method.

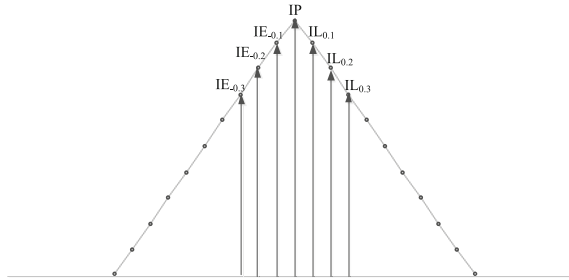
For the incoherent delayed lock loop (DLL), the code loop discrimination error can be expressed as:

$$\begin{aligned} D(\varepsilon) &= \frac{1}{2} (E^2 - L^2) \\ &= \frac{1}{2} \left(|R(\varepsilon + \delta/2)|^2 - |R(\varepsilon - \delta/2)|^2 \right) \end{aligned} \quad (1.4)$$

$D(\varepsilon)$ is discrimination error, E is early correlation value, L is late correlation value, ε is code phase error, δ is the correlator interval. Therefore, the symmetry of the correlation peak directly affects the identification error of the different correlator.

At present, most of the correlators of aviation receivers are equipped with narrow correlators. Therefore, we choose three pairs of narrow correlators (0.1, 0.2 and 0.3 respectively) in order to get three pairs of correlation values, and monitor the symmetry of signal correlation peaks (Fig. 3).

Fig. 3 Multicorrelator of PN code



The monitoring matrix is as follows [9]:

$$\begin{aligned} \Delta_{0,2,0,1}(k) &= \frac{(I\tilde{E}_{-0,2}(k) - I\tilde{L}_{0,2}(k)) - (I\tilde{E}_{-0,1}(k) - I\tilde{L}_{0,1}(k))}{2I\tilde{P}(k)} \\ \Delta_{0,3,0,1}(k) &= \frac{(I\tilde{E}_{-0,3}(k) - I\tilde{L}_{0,3}(k)) - (I\tilde{E}_{-0,1}(k) - I\tilde{L}_{0,1}(k))}{2I\tilde{P}(k)} \end{aligned} \tag{1.5}$$

$$\begin{aligned} r_{\pm 0,1}(k) &= \frac{I\tilde{E}_{-0,1}(k) + I\tilde{L}_{0,1}(k)}{2I\tilde{P}(k)}, \quad r_{\pm 0,2}(k) = \frac{I\tilde{E}_{-0,2}(k) + I\tilde{L}_{0,2}(k)}{2I\tilde{P}(k)}, \\ r_{\pm 0,3}(k) &= \frac{I\tilde{E}_{-0,3}(k) + I\tilde{L}_{0,3}(k)}{2I\tilde{P}(k)} \\ r_{-0,1}(k) &= \frac{I\tilde{E}_{-0,1}(k)}{I\tilde{P}(k)}, \quad r_{-0,2}(k) = \frac{I\tilde{E}_{-0,2}(k)}{I\tilde{P}(k)}, \quad r_{-0,3}(k) = \frac{I\tilde{E}_{-0,3}(k)}{2I\tilde{P}(k)} \\ r_{0,1}(k) &= \frac{I\tilde{L}_{0,1}(k)}{I\tilde{P}(k)}, \quad r_{0,2}(k) = \frac{I\tilde{L}_{0,2}(k)}{I\tilde{P}(k)}, \quad r_{0,3}(k) = \frac{I\tilde{L}_{0,3}(k)}{2I\tilde{P}(k)} \end{aligned} \tag{1.6}$$

Through the statistical analysis of the above 11 monitoring quantities, the abnormal detection is carried out.

3.4 Code/Carrier Phase Inconsistency Monitoring

High precision users will incorporate the carrier phase measurements into the navigation solution. The code/carrier phase inconsistencies observed on the ground are mostly caused by the ionosphere. Code/carrier phase inconsistency is also possible to occur on-board. Although the PN code and carrier are traceable to the on-board atomic clocks, the frequency signals used to generate PN code and carrier phase are through different paths. Any abnormality will lead to pseudo code and carrier phase inconsistency. Usually, the code/carrier phase difference caused by the time and frequency system, is manifested by the divergence between the pseudo range and the carrier phase measurement value. In addition, the relative phase

relationship between the PN code and the carrier will also change because of the abnormal. This kind of fault will affect the ambiguity solution in precision positioning application. Therefore, in the aspect of carrier/pseudo-code phase consistency monitoring, CAST Xi'an has designed two monitoring schemes for two types of fault modes.

(1) CCD (Code/Carrier Divergence)

A two order filter is used to estimate the PN code and carrier divergence rate (equivalent to the series of two first order filters), so as to reduce the influence of code noise. The input of the filter is the original PR measurement value minus the carrier measurement value. The filter is represented as (Laplace domain) [10]:

$$\hat{D}(s) = \frac{s}{(\tau_{d1}s + 1)(\tau_{d1}s + 1)} Z(s) = \frac{1}{(\tau_{d1}s + 1)} \cdot \frac{s}{(\tau_{d2}s + 1)} Z(s) \quad (1.7)$$

The system is converted to the digital domain, using $s = \frac{1}{T} \frac{1-z^{-1}}{z^{-1}}$, at this time:

$$\begin{aligned} d_1(k) &= \frac{\tau_{d1} - T}{\tau_{d1}} d_1(k - 1) + \frac{1}{\tau_{d1}} (z(k) - z(k - 1)) \\ d_2(k) &= \frac{\tau_{d2} - T}{\tau_{d2}} d_2(k - 1) + \frac{T}{\tau_{d2}} d_1(k - 1) \end{aligned} \quad (1.8)$$

In the formula (1.8), $z(k) = PR_r(k) - \varphi(k)$, T is the sampling interval of the code/carrier measurements, τ_{d1} and τ_{d2} are time constant.

First, the difference between the pseudo range of the original code and the carrier pseudo range is calculated, and Z(k) is obtained. Then the two order digital filtering is performed to get $d_2(k)$, and $d_2(k)$ is the code carrier divergence rate.

(2) CCB (Code/Carrier Bias)

The purpose of CCB is to determine whether the increments of the pseudorange and carrier phase are consistent in normal operation.

$$\begin{aligned} \Delta PR(k) &= PR(k) - PR(k - 1) \\ \Delta \varphi(k) &= \varphi(k) - \varphi(k - 1) \end{aligned} \quad (1.9)$$

The difference between the increment of code and the carrier is CCB, as follows:

$$\Delta I(k) = \Delta PR(k) - \Delta \varphi(k) \quad (1.10)$$

4 Satellite Clock Monitoring Technology

High stability crystal oscillator is the local reference of SAIM payload. Oscillator has good short-term stability, but long stability is worse compared with on-board atomic clocks. Thus SAIM only monitors the satellite clock phase or frequency jump abnormally, without long-term deterioration of performance monitoring. The deterioration of the long term performance of the atomic clock can be given by the long-term observation of the ground monitoring network and the inter-satellite link measurements.

The clock difference between the satellite clock and the local reference oscillator for three consecutive periods is t_k , t_{k-1} , and t_{k-2} respectively, and the k represents the current time. It is assumed that the phase clock x is generated at the present time. In order to remove the phase shift introduced by the frequency offset of the satellite clock and the local reference oscillator, we make compare the adjacent measured values, as follows:

$$\Delta T_1 = t_k + x - t_{k-1} = \Delta t_{f(k,k-1)} + x \quad (1.11)$$

$$\Delta T_2 = t_{k-1} - t_{k-2} = \Delta t_{f(k-1,k-2)} \quad (1.12)$$

Δt_f is the phase shift introduced by the frequency difference between the satellite clock and the local oscillator. As the phase shift introduced by the frequency offset of the local oscillator is very small in 3 s of continuous time, it can be ignored. So $\Delta t_{f(k,k-1)}$ and $\Delta t_{f(k-1,k-2)}$ can be approximately considered equal.

The difference between ΔT_1 and ΔT_2 is as follows:

$$\Delta T_1 - \Delta T_2 = (\Delta t_{f(k,k-1)} + x) - \Delta t_{f(k-1,k-2)} \quad (1.13)$$

The phase jump of the satellite clock is obtained from the upper form, and the phase offset is offset by the frequency offset of the local reference oscillator, and the phase jump variable of the satellite clock is obtained. The frequency jump of satellite clock can be derived from phase jump, and it is not redundant.

5 Performance Evaluation Results in Orbit Operation

A new generation of BDS navigation satellite has been launched in the middle of 2015, then test and evaluation were carried out. The pseudorange and carrier phase, signal power, satellite clock phase jump, satellite clock frequency jump, the original concept of measurements quality was statistical analysed. The following figure shows the results of the original observation of the new generation of IGSO satellites in May 12, 2017. The pseudo range accuracy is less than 0.08 ns, carrier

phase accuracy is less than 0.008 weeks, power monitoring accuracy is less than 0.2 dB, satellite clock phase jump and frequency jump are less than 0.01 ns and 0.1 MHz (Figs 4, 5, 6 and 7).

According to the results of the correlation peak distortion monitoring matrix given in the 3.3 section, as shown in Fig. 8, B2a signals can be seen that the correlation peaks are stable and symmetric in the short term, and the long-term stability of the symmetries will be evaluated later.

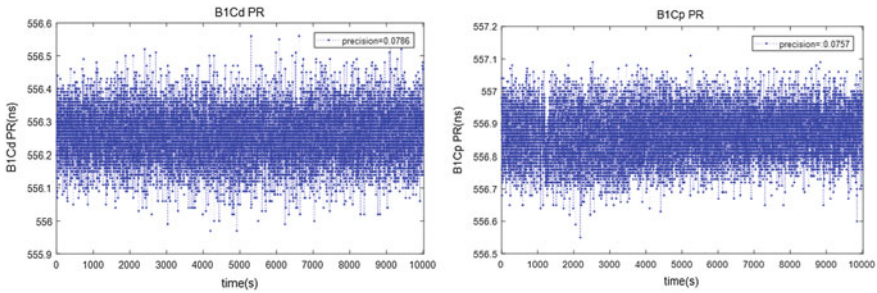


Fig. 4 Evaluation result of time PR measurement

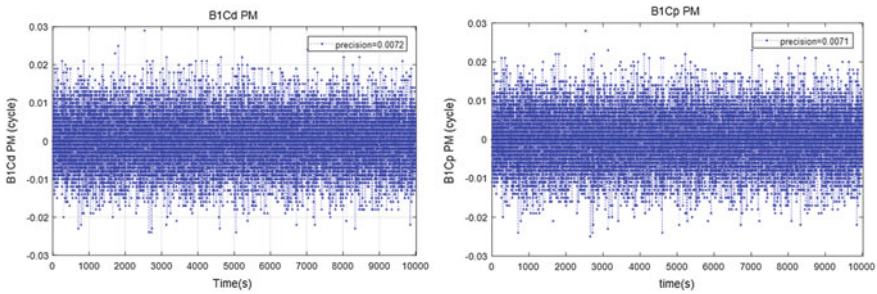


Fig. 5 Result of carrier phase measurement

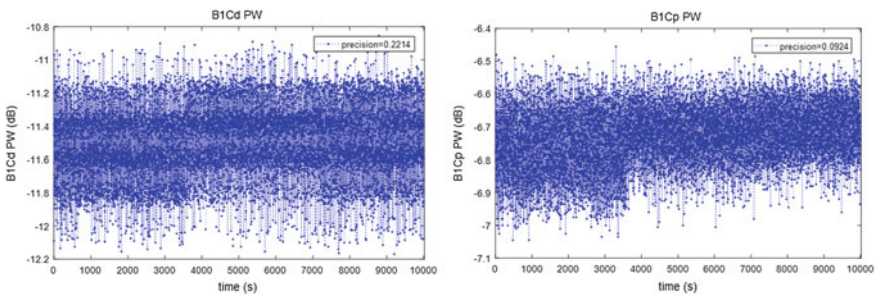


Fig. 6 Result of signal power measurement

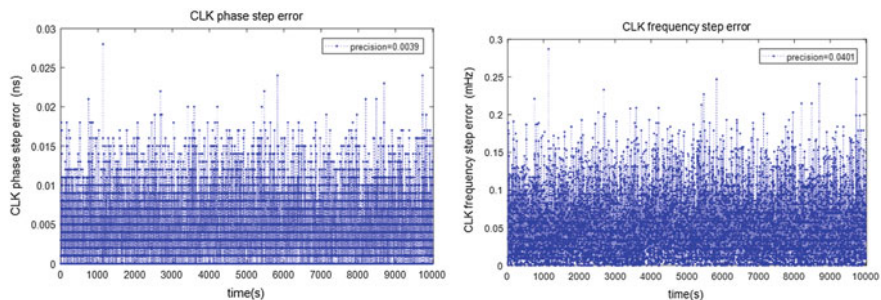


Fig. 7 Result of phase and frequency step measurement

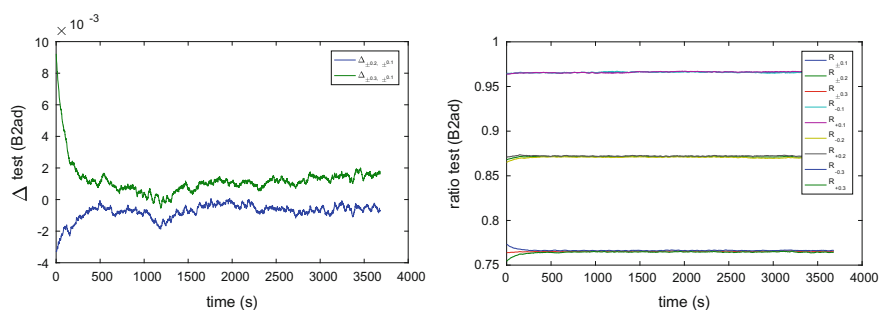


Fig. 8 Result of correlation peak monitoring of B2a

The code/carrier phase consistency monitoring results are as follows: (Fig. 9).

In addition to the performance evaluation under normal operation, in September 2017, the onboard fault injection and the satellite-ground close loop alarm time test were carried out. The signal and satellite clock faults were generated by the ground instructions, for example time delay, power, satellite clock phase and frequency adjustment. The results show that after the fault generated, SAIM can detect the fault in 1 s, and quickly alert to user. The alarm time is less than 6 s while using navigation message. And if NSC warning mode is adopted, the alarm time is less than 2 s. The following figure shows the partial results of the signal abnormally monitoring (Fig. 10).

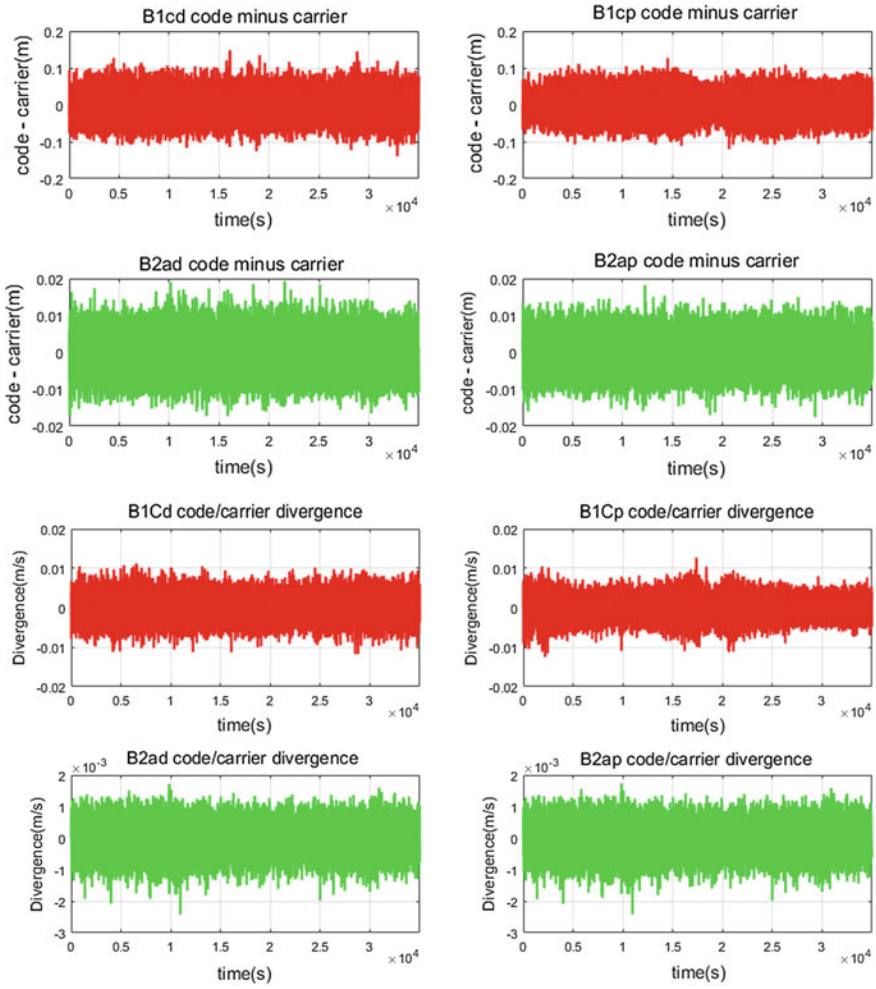


Fig. 9 Result of code and carrier consistence monitoring

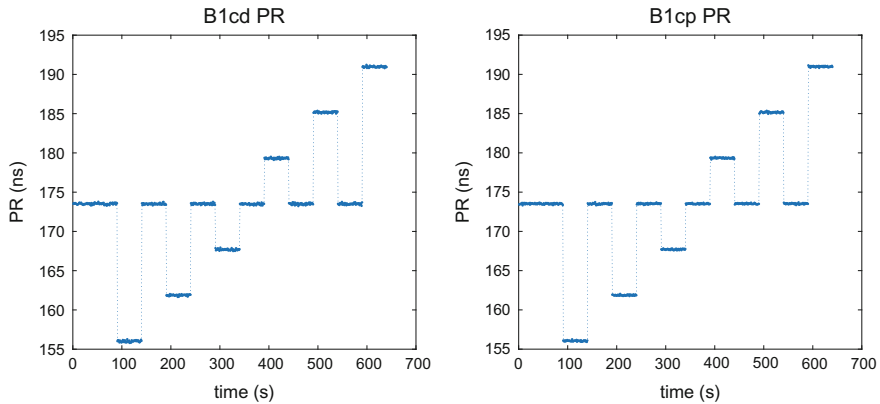


Fig. 10 Result of B1c signal PR step measurement

6 Conclusion

At present, SAIM is carrying out on orbit long-term monitoring experiment, and monitoring on orbit abnormal events to accumulate integrity faults of navigation satellites. The current monitoring focuses on the anomalies of satellite clock and navigation signal, resulting in abnormal jump of receiver measurements, which can alleviate the lack of integrity of users due to the navigation satellite's abnormality. The monitoring accuracy meets the requirements. The alarm time is less than 6 s by navigation message, and less than 2 s by NSC. Until the end of November 2017, false alarm did not occur due to be lack of measurement precision.

From the historical operation of GPS, satellite clock and orbit anomalies are also the main type in-orbit. The highest frequency of satellite is clock failures [11], and the slow failure of satellite clock accounts for a considerable proportion. BDS satellite may also have similar integrity risk. Next, we will further research the monitoring scheme and algorithm of orbit and clock slow change onboard. We will make a comprehensive analysis of main and backup clock difference and inter satellite link observation data, then evaluate the feasibility of autonomous monitoring on orbit.

Acknowledgements This work is supported by National Natural Science Foundation of China (Grant 61627817; Grant 91438107).

References

1. Viearsson L, Pullen S, Green G, Enge P (2001) Satellite autonomous integrity monitoring and its role in enhancing GPS user performance. In: Proceedings of the 14th international technical meeting of the satellite division of The Institute of Navigation (ION GPS 2001), Salt Lake City, UT 2001, pp 11–14
2. McGraw G, Murphy T (2001) Safety of life considerations for GPS modernization architectures. In: Proceedings of the 14th international technical meeting of the satellite division of The Institute of Navigation (ION GPS 2001), Salt Lake City, UT 2001, pp 632–640
3. Rodriguez-Perez I, Garcia-Serrano C, Catalan CC, Garcia AM, Tavella P, Galleani L, Amarillo F (2011) Inter-satellite links for satellite autonomous integrity monitoring. *Adv Space Res* 47(2):197–212
4. GEAS (2008) GNSS evolutionary architecture study, phase I—panel report
5. GEAS (2010) GNSS evolutionary architecture study, phase II—panel report
6. Rodríguez I, García C, Catalán C, Mozo Á (2009) Satellite autonomous integrity monitoring (SAIM) for GNSS systems. In: 22nd international meeting of the satellite division of The Institute of Navigation, Savannah, GA, 22–25 Sept 2009
7. Bian L et al (2012) A principle design for global integrity of COMPASS. *Lect Notes Electr Eng* 161:45–55
8. Bian L et al (2013) Recent research on satellite autonomous integrity monitoring (SAIM) technology. In: 64rd international astronomical congress, Beijing, China (9)
9. Phelts RE (2001) Multicorrelator techniques for robust mitigation of threats to GPS signal quality. Ph.d. dissertation form Stanford GPS LAB
10. Simili DV, Pervan B (2006) Code-carrier divergence monitoring for the GPS local area augmentation system. *IEEE*
11. Heng L (2010) GPS signal-in-space anomalies in the last decade: data mining of 400,000,000 GPS navigation messages. In: Proceedings of the 23rd international technical meeting of the satellite division of the Institute of Navigation (ION GNSS 2010), Portland, OR, Sept 2010, pp 3115–3122