Lecture Notes in Electrical Engineering 562

Jiadong Sun
Changfeng Yang
Yuanxi Yang *Editors*

China Satellite Navigation Conference (CSNC) 2019 Proceedings

Volume I

Springer

Lecture Notes in Electrical Engineering

Volume 562

Series Editors

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India

Swati Meherishi, Executive Editor (swati.meherishi@springer.com)

Aninda Bose, Senior Editor (aninda.bose@springer.com)

Japan

Takeyuki Yonezawa, Editorial Director (takeyuki.yonezawa@springer.com)

South Korea

Smith (Ahram) Chae, Editor (smith.chae@springer.com)

Southeast Asia

Ramesh Nath Premnath, Editor (ramesh.premnath@springer.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

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

Christoph Baumann, Executive Editor (christoph.baumann@springer.com)

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

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

Jiadong Sun · Changfeng Yang ·
Yuanxi Yang
Editors

China Satellite Navigation Conference (CSNC) 2019 Proceedings

Volume I

Springer

Editors
Jiadong Sun
China Aerospace Science and Technology Corporation
Chinese Academy of Sciences
Beijing, China

Changfeng Yang
China Satellite Navigation Engineering Center
Beijing, China

Yuanxi Yang
State Key Laboratory of Geo-information Engineering
Xi'an Research Institute of Surveying and Mapping
Xi'an, Shaanxi, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-13-7750-1 ISBN 978-981-13-7751-8 (eBook)
https://doi.org/10.1007/978-981-13-7751-8

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

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 navigation satellite systems (GNSS) worldwide. BDS will provide highly reliable and precise positioning, navigation and timing (PNT) services as well as 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 10th China Satellite Navigation Conference (CSNC2019) is held during May 22–25, 2019, Beijing, China. The theme of CSNC2019 is "Navigation, 10 Years and Beyond", including technical seminars, academic exchanges, forums, exhibitions and lectures. The main topics are as follows:

Conference Topics

S01 Satellite Navigation Applications
S02 Navigation and Location-based Service
S03 Satellite Navigation Signal and Signal Processing
S04 Satellite Orbit and System Error Processing
S05 Spatial Frames and Precise Positioning
S06 Time Primary Standard and Precision Time Service
S07 Satellite Navigation Augmentation Technology
S08 Test and Assessment Technology
S09 User Terminal Technology
S10 PNT System and Multi-source Fusion Navigation
S11 Anti-interference and Anti-spoofing Technology
S12 Policies, Regulations, Standards and Intellectual Properties

The proceedings (Lecture Notes in Electrical Engineering) has 114 papers in 12 topics of the conference, which were selected through a strict peer review process from 371 papers presented at CSNC2019. In addition, another 156 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 251 referees and 53 session chairmen who are listed as members of editorial board. The assistance of CNSC2019's organizing committees and the Springer editorial office is highly appreciated.

Editorial Board

Topic: S01: Satellite Navigation Applications

Chairman

Qin Zhang	Chang'an University, Shaanxi, China

Vice-chairmen

Shuanggen Jin	Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China
Geyang Li	Qianxun Spatial Intelligence Inc., Beijing, China
Dangwei Wang	Beijing UniStrong Science & Technology Co., Ltd., Shaanxi, China
Xiaolin Meng	University of Nottingham, Nottingham, England

Topic: S02: Navigation and Location-Based Service

Chairman

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

Vice-chairmen

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: S03: Satellite Navigation Signal and Signal Processing

Chairman

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

Vice-chairmen

Yanhong Kou	Beihang University, Beijing, China
Zheng Yao	Tsinghua University, Beijing, China
Tianxing Chu	Texas A&M University-Corpus Christi, Corpus Christi, Texas, USA

Topic: S04: Satellite Orbit and System Error Processing

Chairman

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

Vice-chairmen

Geshi Tang	Beijing Aerospace Control Center, Beijing, China
Rongzhi Zhang	Xi'an Satellite Control Center, Shaanxi, China
Maorong Ge	Geo Forschungs Zentrum (GFZ), Potsdam, Germany

Topic: S05: Spatial Frames and Precise Positioning

Chairman

Qile Zhao	Wuhan University, Hubei, China

Vice-chairmen

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: S06: Time Primary Standard and Precision Time Service

Chairman

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

Vice-chairmen

Chunhao Han	Beijing Satellite Navigation Center, Beijing, China
Xiaohui Li	National Time Service Center, Chinese Academy of Sciences, Shaanxi, China
Aimin Zhang	National Institute of Metrology, Beijing, China
Zhiheng Jiang	International Bureau of Weights and Measures (BIPM), Paris, France

Topic: S07: Satellite Navigation Augmentation Technology

Chairman

Rui Li	Beihang University, Beijing, China

Vice-chairmen

Jinping Chen	Beijing Satellite Navigation Center, 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
Liwen Dai	John Deere, Torrance CA, USA

Topic: S08: Test and Assessment Technology

Chairman

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

Vice-chairmen

Jianping Cao	Air Force Equipment Institute, Beijing, China
Jun Yang	National University of Defense Technology, Hunan, China
Dexiang Ming	Changsha Technology Research Institute of BeiDou Industry Safety, Hunan, China
Yang Gao	University of Calgary, Alberta, Canada

Topic: S09: User Terminal Technology

Chairman

Mingquan Lu	Tsinghua University, Beijing, China

Vice-chairmen

Wenjun Zhao	Beijing Satellite Navigation Center, Beijing, China
Baowang Lian	Northwestern Polytechnic University, Shaanxi, China
Ying Xu	Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China
Sang Jeong Lee	Chungnam National University, Daejeon, South Korea

Topic: S10: PNT System and Multi-source Fusion Navigation

Chairman

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

Vice-chairmen

Hong Yuan	Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, China
Yongbin Zhou	National University of Defense Technology, Hunan, China
Jinling Wang	University of New South Wales, Australia

Topic: S11: Anti-interference and Anti-spoofing Technology

Chairman

Hong Li	Tsinghua University, Beijing, China

Vice-chairmen

Wei Wang	The 20th Research Institute of China Electronics Technology Group Corporation Shaanxi, China
Feixue Wang	National University of Defense Technology, Hunan, China
Xuebin Zhuang	The First Research and Development Center of China Aerospace Science and Technology Corporation Ltd., Beijing, China

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

Chairman

Junlin Yang	Beihang University, Beijing, China

Vice-chairmen

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

Scientific Committee

Chairman

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

Vice-chairmen

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 Chairmen

Jingnan Liu	Wuhan University, Hubei, China
Yuanxi Yang	China National Administration of GNSS and Applications, Beijing, China

Executive Members

Organizing Committee

Directors

Deputy Directors

Jiaqing Ma	China Satellite Navigation Office, Beijing, China
Guangzhi Jiang	Beijing Municipal Bureau of Economy and Information Technology, Beijing, China
Xianwei Zhi	Beijing Shunyi District People's Government, Beijing, China

Secretary-General

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

Deputy Secretary-General

Weina Hao	Satellite Navigation Headquarters, Chinese Academy of Sciences, Beijing, China
Wenhai Jiao	China Satellite Navigation Engineering Center, Beijing, China
Jing You	Beijing Municipal Bureau of Economy and Information Technology, Beijing, China
Xiongjing Lan	Beijing Shunyi District Economic and Information Commission, Beijing, China

Committee Members

Qun Ding	The 20th Research Institute of China Electronics Technology Group Corporation, Shaanxi, China
Qiang Ma	Beijing Airport Economic Core Area Management Committee, Beijing, China
Ze Wang	Beijing Shunxin Holding Group Co., Ltd., Beijing, China
Chunsheng Wang	Haidian Investment Promotion Bureau, Beijing, China
Li Wang	International Cooperation Research Center, China Satellite Navigation Office, Beijing, China
Bo Wang	Academic Exchange Center, China Satellite Navigation Office, Beijing, China

Contents

Satellite Navigation Application

The Single Point Positioning and Program Application of BDS/GPS/GLONASS Based on HELMERT Variance Component Estimation

Junpeng Shi[1] and Kezhao Li[1,2(✉)]

[1] School of Surveying and Landing Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China
lkznwpu@126.com
[2] Collaborative Innovation Center of BDS Research Application, Zhengzhou 450052, China

Abstract. At present, BDS/GPS/GLONASS operates independently. Because of the difference of orbit accuracy, observation noise and other factors, the number of satellites received and the geometric configuration of each system are different. Based on the RINEX format observation data, we can not simply add the fixed weight ratio. Real time dynamic adjustment is needed based on the data quality of the receiver. In this paper, based on the HELMERT variance component estimation model, the optimal weight ratio is calculated by real-time iteration to improve the positioning stability and accuracy. The iteration times of the previous epoch are used as the initial values of the next epoch to reduce the number of iterations. The calculation of the measured data shows that compared with the traditional fixed weight ratio, the positioning accuracy of this method is significantly improved in all directions. Especially under the condition of poor observation conditions, the positioning accuracy and stability are ensured.

Keywords: Beidou navigation satellite system · Data fusion · HELMERT variance component estimation

1 Introduction

With the improvement of the three generations of China's Beidou (BDS-3), China's Beidou gradually has global services. At present, GNSS systems providing global services have been changed from GPS and GLONASS to GPS, GLONASS, BDS. These three independent global positioning systems have different characteristics, such as satellite constellation design, carrier band, navigation satellite coordinate solution and coordinate frame, which make these three independent positioning systems have their respective advantages [1]. GNSS data fusion has become a current trend [2]. Multi-system working simultaneously can increase the number of visible satellites, which is helpful to improve the geometric distribution structure of satellite systems. It can not only improve the positioning accuracy and stability, but also work steadily under the condition of poor observation environment.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 3–13, 2019.
https://doi.org/10.1007/978-981-13-7751-8_1

In the data fusion of GNSS, how to better optimize the weight matrix between various navigation systems is a problem that needs to be explored [3]. Many experiments have proved that the HELMERT mathematical model can improve the weight of each system in GNSS data fusion, but it is mostly a fixed weight matrix, lacking practicality [4–6]. As the number and geometrical position of visible satellites are constantly changing, the fixed weight matrix cannot dynamically adjust the relationship between each system in the GNSS data fusion, failing to achieve better fusion results.

The single-point localization of GNSS multi-system fusion based on Helmert's variance component estimation is based on the geometric configuration of the visible satellite and the observation error generated by the pseudo-range noise. The optimal weight matrix is iterated through the HELMERT estimation formula. Currently, in the same epoch, the geometric configuration of the visible satellite is fixed. In long-term observations, the geometric configuration of visible satellites and the observation errors produced by pseudorange observations are dynamically changing. The weight matrix between the systems should be dynamically adjusted. Since the geometrical configuration and other correlations exist between adjacent epochs, the last epoch HELMERT weight matrix is used as the initial value of the next epoch, which has obvious effect on reducing the number of iterations. This paper selects GPS, GLONASS, BDS system data, and carries out multi-system data fusion single-point positioning experiment based on RINEX 3.02 data format through MATLAB 2016a programming system.

2 Time and Space Unification of BDS/GPS/GLONASS Data

Now GNSS receivers can receive signals from GPS, GLONASS, BDS and other positioning systems at the same time. The time system in data files based on RINEX (Global Satellite Navigation System Receiver Data Autonomous Exchange Format) is the key. Time system is the key to determine the geometric distance between satellite and observation station, and also to calculate the satellite coordinates at observation time by satellite ephemeris [7] (Table 1).

Table 1 GPS/BDS/GLONASS time base comparison

Positioning system	GPS	BDS	GLONASS
Time system	GPST	BST	GLONASST
Starting time	2006.01.01 (0:0:0)	1980.01.06 (0:0:0)	The period from 1996 is 4 years
With or without leap seconds	NO	NO	YES
Time maintenance	UTC(USNO)	UTC(NTSC)	UTC(SU)

The satellite coordinates calculated in RINEX format are determined in the coordinate frame of each positioning system, and the coordinates of the same point are expressed differently in different coordinate frames. In data fusion, the target coordinate system needs to be determined. In multi-constellation navigation, Coordinate Datum

conversion is needed for each constellation system. The following table describes the coordinate frame of each navigation system and its important ellipsoidal parameters, as shown in Table 2.

Table 2 Comparison of GPS/BDS/GLONASS coordinate systems

Positioning system	Coordinate frame	Coordinate origin	Long half shaft	Earth gravity constant	Angle of rotation
BDS	CGCS2000	Earth center of mass	6378 137 m	398600.5×10^9	7.292115×10^{-5}
GPS	WGS-84	Earth center of mass	6378 137 m	398600.4418×10^9	7.292115×10^{-5}
GLONASS	PZ-90	Earth center of mass	6378 136 m	398600×10^9	7.292115×10^{-5}

At present, when receivers receive multiple positioning system data at the same time, there is usually only one reference clock in the receiver. All the data settlement based on RINEX needs to be revised based on "LEAP SECONDS".

3 The Single Point Location of BDS/GPS/GLONASS Multi-system Based on Variance Data Fusion

BDS/GPS/GLONASS pseudorange single point positioning observation equation can be written as

$$\rho_i^{'s}(t) = \rho_i^s(t) + c * \delta t_i(t) - c * \delta t^s(t) + \delta I_i^s(t) + \delta T_i^s(t) \tag{1}$$

$$\rho_i^s(t) = \sqrt{(X^s - X_i)^2 + (Y^s - Y_i)^2 + (Z^s - Z_i)^2} \tag{2}$$

Where: t is the observation epoch; s is the satellite number; c is the speed of light; $\rho_i^s(t)$ is the pseudorange observation at time t; $\delta t_i(t)$ is the receiver clock error at time t; $\delta t^s(t)$ is the satellite clock at time t; $\delta I_i^s(t)$ and $\delta T_i^s(t)$ is the ionosphere and troposphere delay at time t. $[X_i \quad Y_i \quad Z_i]$ is receiver coordinates; $[X^s \quad Y^s \quad Z^s]$ is satellite coordinates;

After unifying the system s time system and coordinate system of any system satellite, the Taylor equation is developed for the pseudorange observation equation to obtain the error equation. For the three systems of BDS/GPS/GLONASS, the GNSS receivers have different receiver clocks for the three systems, so six parameters to be

sought are required and three receiver position parameters and receiver errors of the three corresponding systems are required. At least 6 satellites are required to solve 6 parameters. When more than 6 satellites are received, the least squares method can be used to solve the parameters. Assuming that the epochs are observed at t, and the number of BDS, GPS, and GLONASS satellites is α, β, and γ, respectively, the error equation is:

$$V_i(t) = A_i(t)\delta G_i(t) - L_i(t) \tag{3}$$

In Eq. (3)

$$A_i(t) = \begin{bmatrix} k_i^1(t) & l_i^1(t) & m_i^1(t) & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_i^{\alpha}(t) & l_i^{\alpha}(t) & m_i^{\alpha}(t) & -1 & 0 & 0 \\ k_i^{\alpha+1}(t) & l_i^{\alpha+1}(t) & m_i^{\alpha+1}(t) & 0 & -1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_i^{\alpha+\beta}(t) & l_i^{\alpha+\beta}(t) & m_i^{\alpha+\beta}(t) & 0 & -1 & 0 \\ k_i^{\alpha+\beta+1}(t) & l_i^{\alpha+\beta+1}(t) & m_i^{\alpha+\beta+1}(t) & 0 & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_i^{\alpha+\beta+\gamma}(t) & l_i^{\alpha+\beta+\gamma}(t) & m_i^{\alpha+\beta+\gamma}(t) & 0 & 0 & -1 \end{bmatrix}$$

$$\delta G_i(t) = \begin{bmatrix} \delta X_i \\ \delta Y_i \\ \delta Z_i \\ \delta T_{BDSi} \\ \delta T_{GPSi} \\ \delta T_{GLOi} \end{bmatrix}$$

$$L_i(t) = \begin{bmatrix} \rho_i^{'1}(t) - \rho_i^1(t) + c * \delta t_i(t) + \delta I_i^1(t) + \delta T_i^1(t) \\ \vdots \\ \rho_i^{'\alpha}(t) - \rho_i^{\alpha}(t) + c * \delta t_i(t) + \delta I_i^{\alpha}(t) + \delta T_i^{\alpha}(t) \\ \rho_i^{'\alpha+1}(t) - \rho_i^{\alpha+1}(t) + c * \delta t_i(t) + \delta I_i^{\alpha+1}(t) + \delta T_i^{\alpha+1}(t) \\ \vdots \\ \rho_i^{'\alpha+\beta}(t) - \rho_i^{\alpha+\beta}(t) + c * \delta t_i(t) + \delta I_i^{\alpha+\beta}(t) + \delta T_i^{\alpha+\beta}(t) \\ \rho_i^{'\alpha+\beta+1}(t) - \rho_i^{\alpha+\beta+1}(t) + c * \delta t_i(t) + \delta I_i^{\alpha+\beta+1}(t) + \delta T_i^{\alpha+\beta+1}(t) \\ \vdots \\ \rho_i^{'\alpha+\beta+\gamma}(t) - \rho_i^{\alpha+\beta+\gamma}(t) + c * \delta t_i(t) + \delta I_i^{\alpha+\beta+\gamma}(t) + \delta T_i^{\alpha+\beta+\gamma}(t) \end{bmatrix}$$

The strict estimation formula of Helmert variance component is [8]:

$$S * \hat{\theta} = W \tag{4}$$

In Eq. (4)

$$\hat{\theta} = \begin{bmatrix} \hat{\sigma}_{01}^2 \\ \hat{\sigma}_{02}^2 \\ \hat{\sigma}_{03}^2 \end{bmatrix}, w = \begin{bmatrix} V_1^T P_1 V_1 \\ V_2^T P_2 V_2 \\ V_3^T P_3 V_3 \end{bmatrix}$$

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix}$$

$$s_{ii} = n_i - 2tr\left(N^{-1}N_i\right) + tr\left(N^{-1}N_i\right)^2 i = 1, 2, 3$$

$$s_{ij} = tr\left(N^{-1}N_i N^{-1}N_j\right) i = 1, 2, 3; j = 1, 2, 3; i \neq j$$

4 Program Implementation and Experimental Process

This program is based on the MATLAB2016a programming platform for ephemeris data and observation data solution.

4.1 Time System Correction

In the observation data file, the header records "TIME OF FIRST OBS" and "TIME OF LAST OBS" (if any) contain the time system identifier; different types of receivers use different RINEX time systems, possibly UTC, It may also be GPS. All must be corrected according to the information provided by "LEAP SECONDS".

Ignore the small deviation between time systems. The GLONASS time system is consistent with UTC in the RINEX observation file. The relationship between UTC and BDS and GPS can be expressed by formula (5) (6):

$$UTC = GPST - \Delta T_{LS-GPS} \tag{5}$$

$$UTC = BDT - \Delta T_{LS-BDT} \tag{6}$$

In Eq. (5)

ΔT_{LS-GPS}—The leap seconds of between GPST and UTC are given by GPS navigation message.

In Eq. (6)

ΔT_{LS-BDT}—The leap seconds of between BDST and UTC are given by GPS navigation message.

4.2 Calculation of Satellite Coordinates Based on Broadcast Ephemeris

The GPS broadcast ephemeris uses the Kepler orbital number and its perturbation parameters to represent the satellite position; the broadcast ephemeris update frequency is 2 h. The broadcast ephemeris mainly includes 1 reference time, 6 corresponding reference time Kepler orbit parameters and 9 orbital perturbation correction parameters.

The GLONASS broadcast ephemeris is a state vector given by the satellite in the PZ-90 coordinate frame, $(X(t_b)Y(t_b)Z(t_b))$ is the satellite position at a given time, $(U(t_b)V(t_b)W(t_b))$ is the speed, $(\ddot{x}(t_b)\ddot{y}(t_b)\ddot{z}(t_b))$ is the perturbation of the sun and the moon, etc. Location [9]. The update frequency is half an hour.

The BDS broadcast ephemeris is similar to the GPS ephemeris, with an update period of one hour. The IGSO is the same as the MEO navigation satellite in its solution process, and is different from the GEO navigation satellite solution. The difference is reference [10].

4.3 Propagation Error Correction

Satellite signals passing through the atmosphere to the receiver will have many factors, including ionospheric errors and tropospheric errors. Corrected by the corresponding mathematical model.

4.4 Coordinate Solution and Accuracy Assessment

Based on the HELMERT variance component model, the mathematical model of single point positioning is optimized. Firstly, the solution of unknown parameters is solved by the principle of least squares.

Then, based on the variance component, the weighted optimization is performed according to the HELMERT variance component model. The optimal weight ratio is achieved through multiple iterations, and finally the data fusion is performed according to the optimal weight ratio. If, in an observation epoch, a system occupies less than one percent of its weight, its observations are automatically discarded.

5 Experimental Examples and Analysis

The experimental data of this experiment is selected on August 22, 2018 for static observation data. The observations have received three mainstream navigation and positioning system dual-frequency receivers of GPS, BDS and GLONASS. The observation time is about 50 min, the sampling interval is 5 s, and MATLAB R2016a is adopted. The programming software writes a variance-based BDS/GPS/GLONASS data fusion program. And accurately calculate the receiver coordinates and positioning accuracy.

As shown in Fig. 1, it can be obtained that the number of satellites received by the receivers of the three systems is much larger than that received by a single system, and the number of visible satellites per single epoch reaches an average of 25. The solution of the equation can be obtained by the number of visible satellites exceeding a certain

number in a single epoch. The more the number of visible satellites, the more stable the least squares solution of the observation equation, and the stronger the system reliability (Figs. 2, 3 and 4).

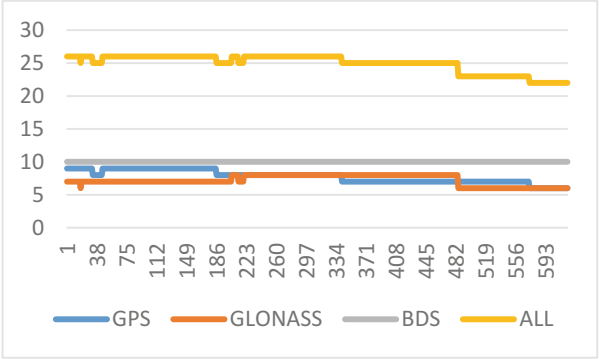


Fig. 1. The number of satellite for GNSS system

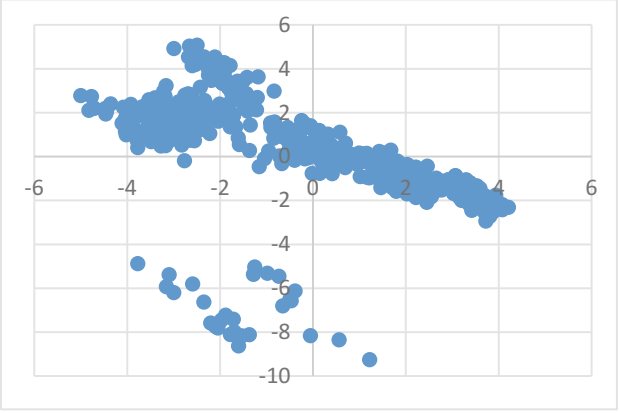


Fig. 2. Residual error components of GPS system in X/Y direction(M)

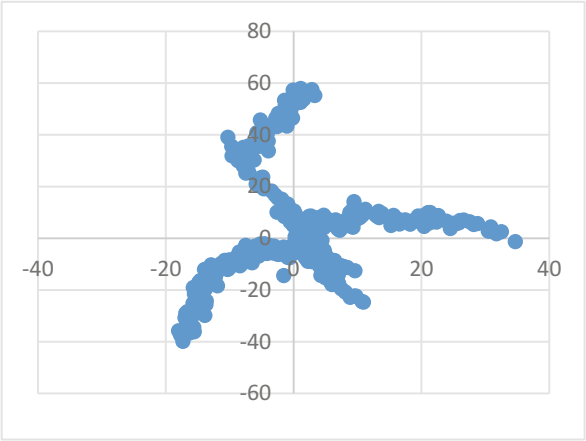


Fig. 3. Residual error components of GLONASS system in X/Y direction(M)

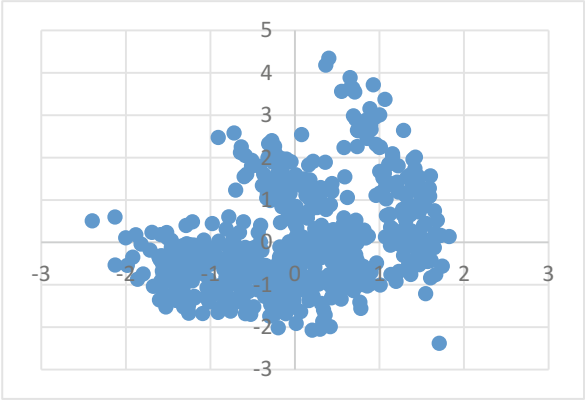


Fig. 4. Residual error components of BDS system in X/Y direction(M)

It can be seen from Fig. 5 that the number of iterations can be significantly reduced by the result of the previous epoch iteration as the initial value and substituted into the HELMERT iterative calculation of the next epoch. The reason is that in the previous epoch, the geometric configuration of the visible satellite has a great similarity with the geometric configuration formed in the next epoch, so the number of iterations can be reduced.

As can be seen from Fig. 6, The variance based on equal-weight data fusion is greatly affected by the observation errors in a single system. Through the original data analysis, the GLONASS positioning system is obtained. The iterative calculation of the Runge-Kutta algorithm results in the influence of the visible satellite coordinates on the step size and the distance ephemeris reference time, and the longer the iterative calculation step, the lower the satellite coordinate precision. However, the step size is too

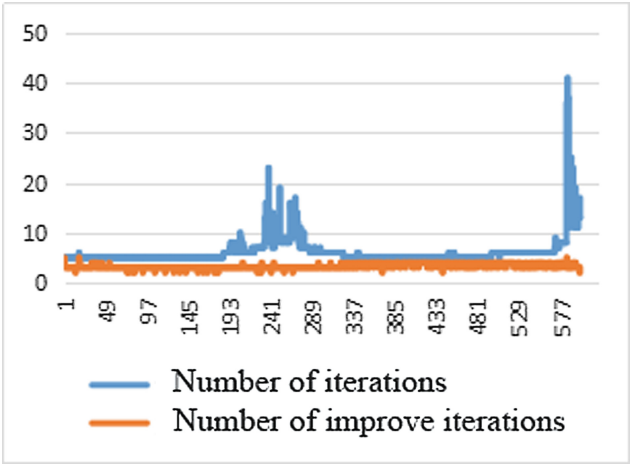


Fig. 5. The number of iterations of Helmert variance component model

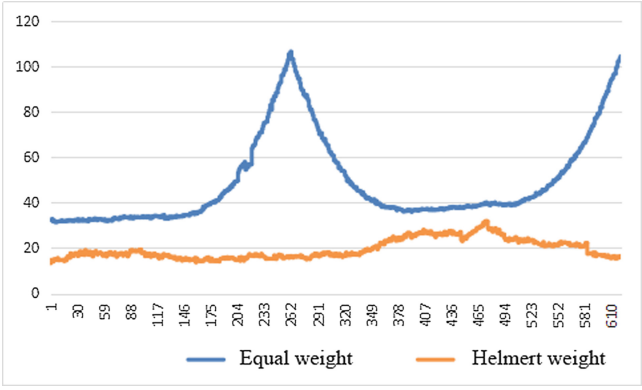


Fig. 6. Multi-system single-point positioning standard difference component in the sum of X/Y/Z directions

small, which affects the calculation speed too much; affected by the Runge-Kutta algorithm, the farther away from the ephemeris reference time, the worse the integration accuracy is. Therefore, the standard deviation of the three directions is cyclically changed. The HELMERT variance component estimation model performs kernel estimation based on the geometry and variance of each system satellite. If a geometric configuration such as a sudden change occurs in a system, the calculation model will reduce the weight ratio of the system, thus ensuring the stability of the entire system.

6 Conclusions

Based on the MATLAB2016a programming platform, this paper implements GPS/BDS/GLONASS GPS/BDS/GLONASS three global positioning system pseudo-range single point positioning data fusion program based on RINEX3.02 data format. In the experiment, the coordinates of the point in the three positioning systems are calculated, and the accuracy of the positioning results is analyzed. It is iteratively weighted by Helmert's variance component mathematical model, and its positioning accuracy is analyzed. The following conclusions were obtained:

1. After GPS/BDS/GLONASS multi-system pseudo-range single-point positioning, the number of visible satellites is greatly increased in single epoch. At the same time, because the number of visual satellites is much larger than the number of unknown ones, the results obtained by least square theory have higher fault-tolerant rate.
2. Due to the increased number of visible satellites, the better geometry of visible satellites in single epoch observations effectively reduces the risk of increased GDOP values.
3. Based on the Helmert variance component model, dynamically adjust the correlation weight matrix between the three systems. The stability of positioning accuracy is guaranteed.

Acknowledgment. This research was funded by the National Natural Science Foundation of China (Nos. 41774039, 41272373).

References

1. Tan S (2017) Innovative development and forecast of BeiDou system. Acta Geodaetica et Cartographica Sinica 46(10):1284–1289. https://doi.org/10.11947/j.agcs.2017.20170329 (China)
2. Yang Y, Lu M, Han C (2016) Some notes on interoperability of GNSS. Acta Geodaetica et Cartographica Sinica 45(3):253–259 (China)
3. He J, Yuan X, Zeng Q, Liu W (2014) Study on single point positioning of GPS/BDS/GLONASS combination. Sci Surv Mapp 39(8):164–170 (China)
4. Du W, Liu Y, Yan J (2017) Multi-GNSS fusion pseudorange single point location algorithm and its program implementation. Geomatics Spat Inf Technol 40(9):47–51
5. Zhang G, Yu X et al (2017) Analysis of the influence of reasonable weight ratio on GPS/BDS pseudorange single point positioning results. Gnss World China 42(3):48–53
6. Fu X (1997) Initiality problem of iterative adjustment using Helmert variance estimation. J Hebei Inst Archit Sci Technol 14(4):53–56
7. Li H, Dang Y, Mi J, Yang F (2013) BDS and GPS, GLONASS multi-mode fusion navigation positioning space-time unification. J Geodesy Geodyn 38(4):189–194

8. Hu S (2017) Theory and Application of Modern Measurement Data Processing. Surveying and Mapping Press, Beijing
9. Ge M, Guo J, Ge S (1999) Calculation method of GLONASS satellite coordinates. Bull Surv Mapp 45(2):2–4
10. China Satellite Savigation Office. Beidou navigation satellite system space signal in space interface control document service signal B3I (version 1.0)

Ground-Based GPS for Soil Moisture Monitoring

Guisheng Zhao[1(✉)], Shuangcheng Zhang[1,2(✉)], Qin Zhang[1], Jingjiang Zhang[3], Lifu Wang[4], and Tao Wang[1]

[1] College of Geology Engineering and Geomantic, Chang'an University, Xi'an 710054, China
zgs@chd.edu.cn, shuangcheng369@vip.163.com
[2] China Meteorological Administration Urumqi Institute of Desert Meteorology, Urumqi 830002, China
[3] Institute of Urban Meteorology, CMA, Beijing, China
[4] Xinjiang Altai Meteorological Bureau, Altai 836500, China

Abstract. This Soil moisture is an important part of the surface water cycle. Effective monitoring of soil moisture is of great significance for weather forecasting, flood forecasting and crop growth. Existing soil moisture monitoring methods (such as drying weighing, remote sensing observation, hygrometer measurement, etc.) have high cost, low spatial and temporal resolution, damage to observation objects, time-consuming and laborious, and long repeated observation periods. With the rapid development of GNSS remote sensing, the GPS signal based on microwave L-band is used for soil moisture monitoring with the advantages of low cost, high time resolution, strong real-time and high automation, which has attracted the attention of many scholars. This paper intends to study the soil moisture inversion algorithm based on GPS signal-to-noise ratio. Firstly, based on the basic principle, the inversion process is given. The fitting phase is zero-processed and the fusion is performed according to the correlation coefficient. Finally, a linear model of humidity is established. The inversion of the PBO plan station data was carried out, and the soil moisture data provided by PBO was compared. The results confirmed that the accuracy and stability of soil moisture inversion using GPS signal-to-noise ratio were improved by zero processing and weighted fusion.

Keywords: GPS · SNR · Soil moisture · Multipath

1 Introduction

Soil moisture is one of the important criteria for measuring water cycle, and it is an important physical quantity in agricultural, meteorological and hydrological research [1]. In order to overcome the shortcomings of traditional measurement methods (such as drying weighing, hygrometer, remote sensing observation, etc.) on the damage of observation objects and long observation period and low temporal and spatial resolution, Larson et al. proposed to use GPS signal-to-noise ratio (SNR). Perform soil moisture changes [2]. The variation of frequency, amplitude and phase in a certain cutoff height angle is used to invert the change of soil moisture using the multipath

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 14–22, 2019.
https://doi.org/10.1007/978-981-13-7751-8_2

reflection component of SNR. Chew et al. theoretically deduced that the correlation coefficient between phase and soil moisture is over 0.99 [3]. The study in [3] reveals that the role of soil within 1 cm of the surface is particularly critical. Under the same soil water content, the influence of soil type on the interference parameters is negligible. Based on the single scattering, the phase and soil content are established. A regression model of water volume. Vey et al. used this model to calculate the soil water content of the area where the SUTM station was located, and the results were in good agreement with the time domain reflectometry [4].

In China, Liu and others began to study the field of GNSS remote sensing [5]. Zhang, Jin et al. used GNSS signal-to-noise ratio to effectively monitor volcanic ash [6], and prospected the latest development and application prospects of GNSS + R [7]. Yan Yuhua and others first carried out GNSS-R measurements on surface soil moisture [8], and Zou and Yang also used the Beidou GEO satellite to perform soil moisture inversion [9]. Min Minsi and others pioneered the use of signal-to-noise ratio to invert soil moisture [10].

Inversion of soil moisture using GPS signal-to-noise ratio can effectively extend the application of geodetic receivers, but how to better integrate the inversion results of multiple satellite arcs and improve the accuracy and reliability of inversion of soil moisture is still worth studying. This paper studies and analyzes the above problems by using the measured data of the P041 station of the US PBO program through zero processing and empowerment.

2 Principles and Methods

2.1 Soil Moisture Inversion Principle

The SNR data is affected by the reflected signal at low elevation angles, and there is significant oscillation. After the direct component is removed by the second-order polynomial, the reflection component can be expressed by the following publicity under the assumption of one reflection [11]:

$$SNR_m = A_m \cos\left(\frac{4\pi H_0}{\lambda} \sin\theta + \varphi_m\right) \tag{1}$$

Where SNR_m denotes the residual sequence after the trend term is removed, λ is the carrier wavelength, H_0 is the station reflection high, A_m and φ_m are the relative amplitude and delay phase of the reflection component, and θ is the reflection elevation angle. Usually, the antenna height H_0 measured in the field is not directly used, but the long-term effective reflection high time series is used to take the median, which is recorded as H_{eff} [3]. Use the following equation to find the effective reflection height [12]:

$$H_{eff} = \frac{\lambda}{2} f_m \tag{2}$$

Where f_m is the dominant frequency obtained by L-S spectral analysis and λ is the carrier wavelength. The frequency is converted into an effective reflection height by the above equation, which can be used for snow monitoring [13] and sea level inversion [14].

The reflection component SNR_m is resampled to obtain its relationship with the $\sin\theta$. According to Eq. (1), A_m and φ_m can be obtained by nonlinear least squares fitting. The change of $-\varphi_m$ is obviously consistent with the trend of soil moisture [15]. Zavorotny, Larson et al. showed a linear relationship between phase change and soil moisture. The relationship between phase and soil moisture can be expressed by:

$$SMC_t = S * \Delta\varphi_m + SMC_{res} \tag{3}$$

Where SMC_t is the soil moisture value of a certain day, $\Delta\varphi_m$ is the phase after zero treatment, and SMC_{res} is the soil moisture base. Different soil and surface environments affect S and SMC_{res}, so this paper uses the soil moisture provided by PBO as a known value to analyze the relationship between phase and soil moisture using the following equation:

$$SMC_t = S' * \Delta\varphi + C \tag{4}$$

S' is the slope of phase and soil moisture, and C is a constant term.

2.2 Soil Moisture Inversion Process

First, the direct and reflected components are separated by a second-order polynomial, and the corresponding relationship between SNR_m and $\sin\theta$ is obtained by resampling. H_{eff} is obtained by L-S spectrum analysis in combination with Eq. (2). Then, according to Eq. (1), A_m and φ_m are obtained by nonlinear least squares fitting.

Next, zero processing is performed on φ_m. The phase value of each satellite per day is subtracted from the fitting phase of the satellite on the 101st day of the year to obtain the amount of phase change per day, which is recorded as $\Delta\varphi_m^{'}$. Φ refers to $-\Delta\varphi_m^{'}$.

Because the influence of various environmental factors (vegetation, topography, etc.) on the fitting result is finally reflected in the correlation coefficient between the fitting phase and the true value of soil moisture, this paper attempts to use the correlation coefficient between the two to perform multi-star inversion. The fusion of results.

Φ is weighted according to the correlation coefficient of each satellite Φ and SMC_t, and all the weighted Φ are fused to obtain $\Delta\varphi$. If a satellite has missing data on a certain day, the satellite does not participate in the fusion of the day.

Finally, a linear model of $\Delta\varphi$ and SMC_t is established according to Eq. (4). The overall inversion process is shown below (Fig. 1).

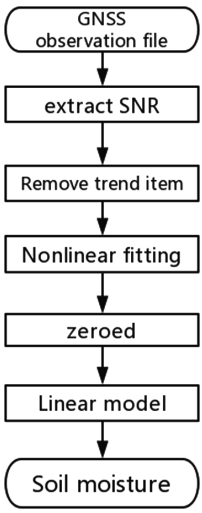


Fig. 1. Soil moisture inversion process

3 Case Analysis

3.1 Data Source

This paper selects the L2 data SNR from the 101st to 220 days of the 2012 PBO station P041 station in the United States for soil moisture inversion. SMC_t also comes from the data provided by the PBO program. Satellite data in the azimuth range from 0° to 90° of the station is selected, and the cutoff height angle is set to 5° to 30°. Under the assumption that the station height is 2 m, the surrounding environment of the P041 station and part of the Fresnel reflection area are shown in Fig. 2.

Fig. 2. The surrounding environment of the P041 station and some Fresnel reflection

The station is located at 39.9495° north latitude and 105.1943° west longitude. The surrounding terrain is relatively flat and open, with no large obstructions and a small amount of seasonal vegetation, which is very suitable for soil moisture inversion related research.

3.2 Correlation Coefficient

Taking the selected data as an example, the relationship between some satellite arc segments Φ and soil moisture is shown in Fig. 3. The black dots in the figure indicate the SMC_t of each day, and the blue dots correspond to the Φ obtained by the inversion of each satellite arc. The figure shows the inversion results of six satellite arcs. Among them, PRN26 has the best inversion results and can respond well to changes in soil moisture almost every day. The inversion result of PRN4 is generally low, and it has almost no response to the rain from the 130th day to the 150th day of the annual accumulation date. The inversion results of the remaining satellites can reflect the change of soil moisture to some extent, but there are also different degrees of hopping and migration, especially during the period after rainfall.

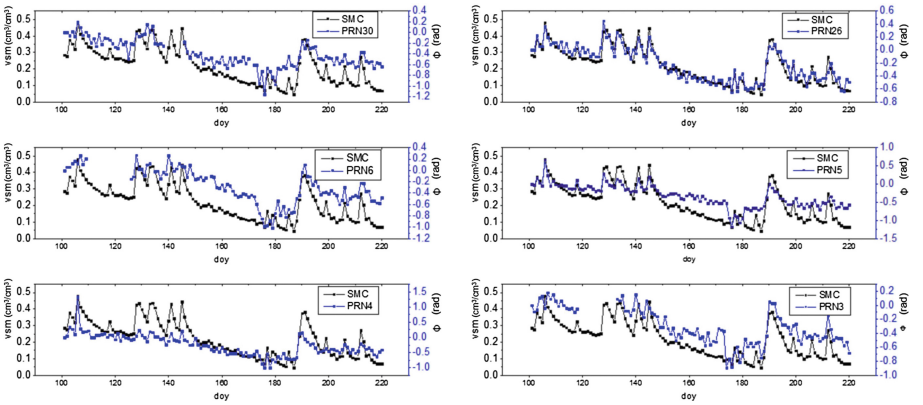


Fig. 3. Part of the satellite arc Φ and *VSM* of P041 station

It can be seen from the figure that the trend of change of Φ is basically consistent with the trend of change of SMC_t. Differences in the inversion results of different satellites may be due to environmental impacts such as vegetation cover and topographic fluctuations.

The correlation coefficients of Φ and $\Delta\varphi$ and SMC_t for each satellite arc are shown in the table below (Table 1).

Table 1. Correlation coefficients of Φ and SMC_t for each satellite

Satellite number	PRN30	PRN26	PRN19	PRN16	PRN15	PRN11	PRN10
Correlation coefficient	0.8784	0.9478	0.9296	0.8421	0.9036	0.9021	0.8797
Satellite number	PRN9	PRN6	PRN5	PRN4	PRN3	PRN1	$\Delta\varphi$
Correlation coefficient	0.9429	0.8580	0.8899	0.8743	0.8829	0.8828	0.9427

It can be seen that the correlation coefficient between Φ and SMC_t after the weighted fusion is 0.9427, which is generally higher than the inversion result of a single satellite, only slightly lower than PRN26 and PRN9. The relationship between $\Delta\varphi$ and SMC_t is shown in Fig. 3.

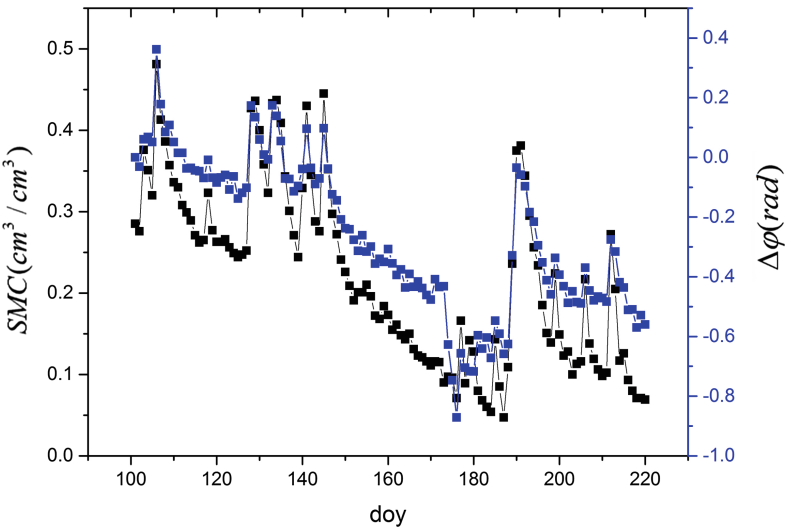


Fig. 4. $\Delta\varphi$ and SMC_t of P041 station

The black dots in the figure represent *SMC*, and the blue dots represent $\Delta\varphi$ obtained by fusion. In the 120 days of the inversion, it can be seen from Fig. 4 that the change trend of $\Delta\varphi$ is basically consistent with the change trend of *SMC*. The poor performance of the doy150–doy180 is due to the poor inversion results of most satellites during this time period. It is speculated that there may be changes in the surface environment during this period, resulting in poor inversion results of satellites.

3.3 Linear Model

According to Eq. (4), into $\Delta\varphi$ and SMC_t, the following linear model obtained at a 95% confidence interval.

$$SMC = 0.407\Delta\varphi + 0.3278$$

The linear model fitting results are shown in Fig. 5.

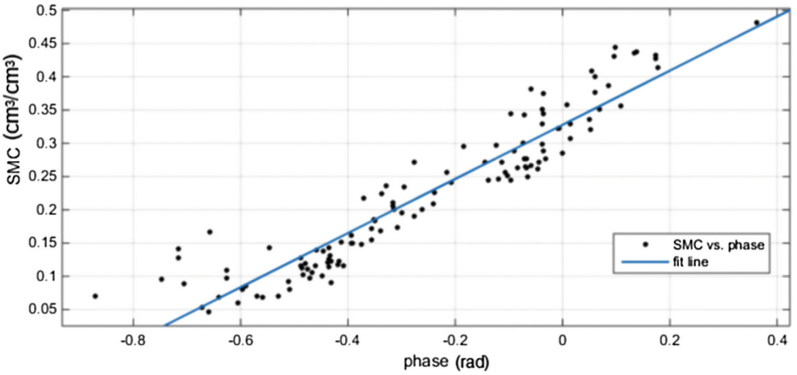


Fig. 5. Linear model fitting results

In the figure, phase refers to $\Delta\varphi$, the blue line is a linear model fitting curve, and the black point is the $\Delta\varphi$ coordinate point corresponding to *SMC*. It can be seen that the overall fitting effect is good, but there is a large deviation when *SMC* is low and high. The fitting result had an R^2 of 0.8887 and an RMSE of 0.037.

In order to verify the feasibility of the linear model, the corresponding $\Delta\varphi$ is obtained from the data from the 221st to the 250th day of the P041 station in 2012, and *SMC′* is obtained according to the established linear model. The relationship between the obtained *SMC′* and the corresponding true value *SMC* is as shown in Fig. 6. The blue dot represents the linear model predicted value *SMC′*, and the black dot represents the soil moisture true value *SMC*.

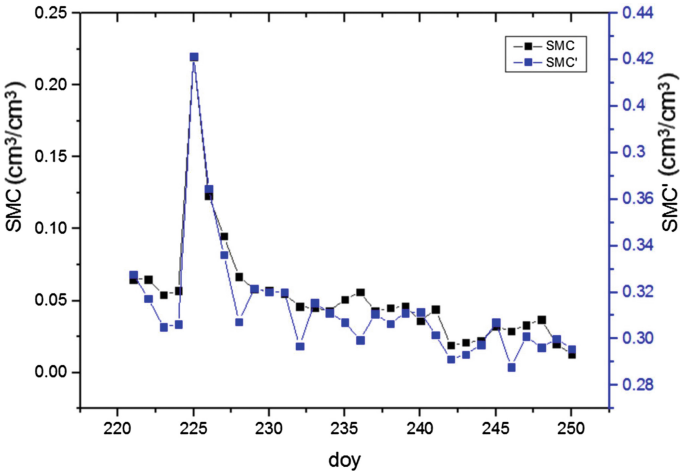


Fig. 6. Linear model verification

The correlation coefficient between SMC' and SMC in the figure is 0.9581, which has a strong correlation and can well describe the trend of soil moisture within the 30 days. However, due to the relatively short data length of the linear model, the obtained SMC' is deviated from SMC. It is expected that the establishment of a linear model through longer-term data can effectively improve the quality of SMC'.

4 Conclusion and Outlook

Although the GPS interference phase is theoretically linear with soil moisture, in actual observations, changes in the observed environment such as receiver noise, topographic changes, vegetation cover, and snow cover may cause deviations between the two. In this paper, by performing zero processing on the inversion phase, the weighted fusion is performed according to the correlation coefficient between each satellite arc and soil moisture, and a linear model between the fusion phase and soil moisture is established. The experimental results of P041 station data show that the method can effectively improve the instability of single satellite inversion results, and the established linear model can also effectively reflect the trend of soil moisture. However, the calculated linear model has a deviation from the true value. This problem may be caused by the short length of the modeling data. At the same time, it still does not solve the problem of terrain changes, vegetation cover and other inversion results. These issues still need further research and resolution.

Acknowledgments. We gratefully acknowledge the provision of data, equipment, and engineering services by the Plate Boundary Observatory operated by UNAVCO for Earth Scope. This work was supported by China Desert Meteorological Science Research Foundation (Sqj2017002), National Science Foundation of China (41731066, 41674001, 41104019) and the Special Fund for Basic Scientific Research of Central Colleges (310826172202).

References

1. Jackson T, Schmugge J, Engman E (1996) Remote sensing applications to hydrology: oilmoisture. Hydrol Sci 41(4):517–530
2. Larson KM, Small EE et al (2008) Using GPS multipath to measure soil moisture fluctuations: initial results. GPS Solutions 12(3):173–177
3. Chew CC, Small EE, Larson KM (2014) Effects of near-surface soil moisture on GPS SNR data: development of a retrieval algorithm for soil moisture. IEEE Trans Geosci Remote Sens 51(1):537–543
4. Vey S, Wickert J, Guntner A et al (2013) Derivation of near-surface soil moisture from the GNSS site Sutherland, South Africa. In: Workshop of GNSS-Reflectometry, Brest, France
5. Liu J, Shao L, Zhang X et al (2007) Advance in GNSS-R studies and key technologies. Geomatics Inf Sci Wuhan Univ 32(11):955–960
6. Zhang Q, Jin S (2018) Application of GNSS SNR data in detection of volcanicplume. J Navig Positioning 6(1):15–20

7. Jin S, Zhang Q, Qian X (2017) New progress and application prospects of global navigation satellite system reflectometry (GNSS + R). Acta Geodaetica et Cartographica Sinica 46 (10):1389–1398
8. Yan S, Gong J, Zhang X et al (2011) Ground based GNSS-R observations for soil moisture. Chin J Geophys 54(11):2735–2744
9. Zou W, Zhang B, Yang D et al (2016) Soil moisture retrieval using reflected signals of BeiDou GEO satellites. Acta Geodaetica et Cartographica Sinica 45(2):199–204
10. Ao M, Hu Y, Liu Y et al (2012) Inversion of soil moisture fluctuation based on signal-to-noise ratio of global positioning system. J Geomatics Sci Technol 29(2):66–69
11. Larson KM, Braun J, Small EE et al (2010) Use of GPS receivers as a soil moisture network for water cycle studies. IEEE J Sel Top Appl Earth Obs Remote Sens 3(1):91–99
12. 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, pp 2638–2648
13. Dai K, Zhang S, Zhang Q et al (2015) Preliminary results of snow depth monitoring based on GPS SNR. Sci Surv Mapp 40(12):112–115
14. Nan Y, Zhang S, Zhang Q et al (2015) Preliminary results of sea level change monitoring with GNSS-MR. Sci Surv Mapp 40(12):125–129
15. Ao M, Zhu J, Hu Y et al (2015) Comparative experiments on soil moisture monitoring with GPS SNR observations. Geomatics Inf Sci Wuhan Univ 40(1):117–120

BDS-MR for Snow Depth Monitoring in Altai

Shuangcheng Zhang[1,2(✉)], Chenglong Zhang[2], Lifu Wang[4], Tao Che[3], Hao Li[2], and Xuerong Chen[2]

[1] China Meteorological Administration Urumqi Institute of Desert Meteorology, Urumqi, Xinjiang, China
shuangcheng369@vip.163.com
[2] College of Geology Engineering and Geomantic, Chang'an University, Xi'an, China
[3] Northwest Institute of Eco-Environment and Resources, CAS, Lanzhou, China
[4] Xinjiang Altai Meteorological Bureau, Altai, China

Abstract. Snowfall is an important freshwater resource; manual monitoring and laser detection can monitor snow to a certain extent, but there is a lot of uncertainty in time and space. Traditional ground-based geodetic Global Navigation Satellite System (GNSS) receivers are more widely distributed, this article uses the advantages of BDS all-weather, real-time, high time resolution and high automation to monitor snow depth, taking the snow monitoring station in Altai City, Xinjiang as the research object, Firstly, explain the basic principles of BeiDou System-Multipath reflectometry (BDS-MR) snow depth detection, secondly, based on the data of the Altai station from January to March 2017, about 90 days, From the aspects of Beidou satellite type, the feasibility and detection accuracy of BDS-MR are studied by using the signal-to-noise ratio (SNR) of BDS, and compared with the measured snow depth data. The results show that: When the SNR of the IGSO and MEO satellites of BDS is used to invert the snow depth deviation to 0.021 m. It is proved that BDS-MR can be used for snow depth detection, which can better utilize the advantages of microwave remote sensing and promote the application of Beidou navigation system.

Keywords: BDS-MR · Snow depth · Altai · Beidou multipath reflection · SNR

1 Introduction

As one of the most active environmental factors in the world, snow has played an important role in the water cycle. As an important freshwater resource, snow has positive significance for human beings, but in extreme weather, snow will bring disaster to people and seriously affect people's lives [1–3]. With the development of technology, from the initial manual monitoring to the use of remote sensing satellites for snow monitoring in recent years, however, manual monitoring consumes a lot of manpower and material resources, and is not suitable for a wide range of monitoring. Due to the low spatial resolution of optical remote sensing and microwave remote sensing, the accuracy of the snow depth results is not enough [4], and it is still one of the research hotspots in the field of remote sensing.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 23–30, 2019.
https://doi.org/10.1007/978-981-13-7751-8_3

In 2007, Larson first proposed a method for measuring soil moisture using an existing GPS continuous operation observatory, namely Global Positioning System-Multipath reflectometry (GPS-MR) technology [5]. In 2009, Larson et al. used the GPS-MR technology to monitor the variation of the snowfall thickness of the two spring blizzards at American Marshall Field site, the accuracy is around 5 cm [6]. In 2014, Nievinski and Larson used GPS-MR to perform a series of surface environmental monitoring [7], which laid the foundation for BDS-MR technology for snow research. Since 2012, a series of studies have been carried out in China [8, 9]. In 2014, Jin et al. used the GPS-L4 method to study the snow surface temperature and snow depth, and proposed a non-parametric statistical estimation model for snow surface temperature and snow depth estimation. The accuracy of the SMM1 station is 0.18° and 0.23 m, MARG station accuracy is 3.8° and 0.13 m [10]. In 2015, Yu et al. used the linear combination of multipath reflection and GPS tri-band (L1, L2, and L5) signal phase measurements to perform snow depth inversion on the data of STK2 station in Japan, and compared with the inversion results of L4 and SNR, the accuracy is about 7 cm [11]. In 2016, Zhang and others used the GPS SNR observations of the PBO network NWOT station to perform snow depth inversion with an accuracy of about 7 cm [12].

At present, in the aspect of snow depth detection, the domestic use of GPS stations in foreign countries is based on GPS-MR technology for snow depth detection [13]. In China, there are many snow-rich areas such as Northeast China, Tibet, Xinjiang, etc. Few scholars have carried out GNSS-MR snow monitoring work in these areas. We have established a GNSS snow monitoring station in the Altai region of Xinjiang, using the Beidou data of the monitoring station to invert the snow depth around the site based on BDS-MR technology, Discuss the impact on the results from satellite types of Beidou, Finally, an appropriate inversion method is obtained based on the error analysis.

2 BDS-MR Detection of Snow Depth of the Basic Theory

Traditional ground Geodetic GNSS receiver receives satellite signals that comprise direct signals and reflective signals in the surface environment, through the reflection of the surface environment, the amplitude and phase of the reflected signals change, resulting in multipath effects. The SNR (representing the ratio of signal strength to noise intensity) is inversely related to multipath, low satellite elevation angle SNR can reflect multipath effects and multipath error accuracy, and in contrast, multipath effects directly affect the SNR. When performing BDS-MR inversion of snow depth, it is necessary to select effective experimental satellites according to the environmental information around the station, the first Fresnel zone [14] and the reflection trajectory of the satellite ground point.

In Fig. 1, the X-axis represents the distance from the receiver antenna to the direction of the satellite being measured, and the Y-axis represents the detected longitudinal distance. Different colors correspond to the first Fresnel zone of different satellite elevation angle, black is 5°, red is 7°, blue is 10°, purple is 15°, and green is 20°. The upper part of Fig. 1 is when there is no snow, and the lower part is when the snow depth is 1 m. It can be seen from Fig. 1 that as the satellite elevation angle increases, the area

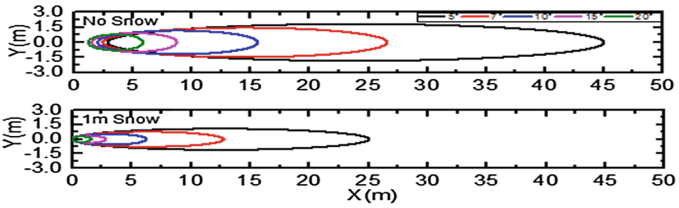


Fig. 1. Fresnel zone corresponding to different satellite elevation angle

of the first Fresnel zone on the ground continuously shrinks; at the same satellite elevation angle, as the thickness of the snow increases, the area of the Fresnel zone becomes significantly smaller. The ground first Fresnel zone can be considered as the spatial resolution of BDS-MR snow depth detection. After the first Fresnel zone on the ground is used to roughly determine the detection range of the snow depth of the station, the trajectory map of the reflection point is used to select the corresponding effective satellite arc, the satellite reflection point trajectory is related to the satellite elevation angle, azimuth angle and antenna height, and represents the position of the satellite relative to the receiver at a certain epoch time, thereby obtaining snow space distribution information. The reflection point trajectory [15] can be expressed as:

$$X = \frac{h}{tan\theta} sin\varphi \tag{1}$$

$$Y = \frac{h}{tan\theta} cos\varphi \tag{2}$$

In Eqs. (1) and (2), X is the component of the projection trajectory on the X axis, Y is the component of the projection trajectory on the Y axis, h is the distance from the antenna phase center to the ground, θ is the satellite elevation angle, and φ is the satellite orientation angle. Trajectory map of satellite reflection point is obtained by the trajectory function of the satellite reflection point (Fig. 2):

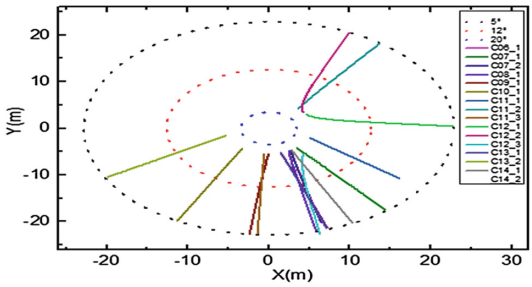


Fig. 2. BDS effective satellite reflection trajectory

The origin in Fig. 2 represents the projection of the antenna phase center to the ground, the horizontal and vertical axes are the distribution of satellite signal reflection points around the station, The solid line of each color is the effective arc of the Beidou satellite system on a certain day, the dotted circles from the outside to the inside represent the satellite elevation angles of 5°, 12°, and 20°, respectively, the effective reflection arc of this day and the spatial distribution information of the snow can be clearly seen from the figure.

BDS-MR technology essentially uses SNR spectrum analysis to obtain surface environmental parameters that cause multipath effects. Snow is also one of the surface environments [16].

$$\mathrm{SNR} \infty P_d + P_r + \sqrt{P_d P_r} cos\varphi \tag{3}$$

In Eq. (3), P_d is the direct signal energy P_r is the reflected signal energy, and φ is the angle between the direct signal and the reflected signal. In order to obtain the change information of BDS multipath caused by surface reflection in SNR, multipath effects need to be separated from the received SNR observations (Fig. 3), since P_d and P_r differ greatly in value, the reflected signal energy can be obtained by formula (4):

$$\mathrm{SNR}(Volts/volts) = 10^{\frac{SNR(dB-Hz)}{20}} \tag{4}$$

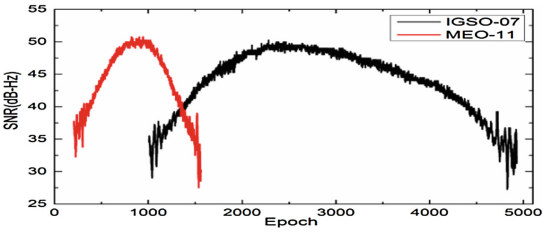


Fig. 3. SNR variation of Beidou satellite MEO IGSO

Figure 3 shows the original SNR data of the MEO (black) and IGSO (red) satellites of the BDS. The horizontal axis is the epoch and the vertical axis is the SNR value, it can be seen from the figure that the SNR is generally parabolic due to the influence of the direct signal, and due to the influence of the multipath effect, it appears as a local periodic oscillation at a low satellite elevation angle, after linearization, the residual sequence amplitude of the multipath reflected signal is obtained by the method of eliminating the trend term (Fig. 4):

In Fig. 4, the horizontal axis represents the resampling time point as a function of the height angle, and the vertical axis represents the linear change value of the SNR value in volts, it can be seen from this figure that the multi-path information of low satellite elevation angle has a serious impact on SNR. The amplitude of the residual sequence can be expressed as Eq. (5):

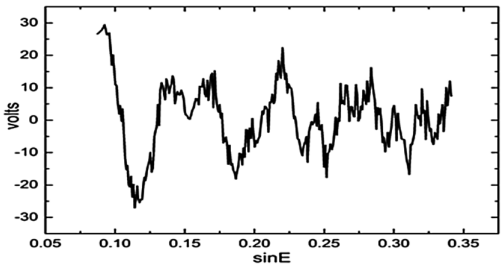


Fig. 4. Remove the SNR residual sequence of the trend direction

$$\mathrm{dSNR} = \mathrm{A}\cos\left(4\pi \mathrm{H}\lambda H^{-1}\mathrm{sinE} + \phi\right) \tag{5}$$

In Eq. (5), A is the amplitude, H is the reflection height, λ is the wavelength of the Beidou satellite band, E is the satellite elevation angle, and ϕ is the phase, t = sinE, $f = 2H/\lambda$ which can be obtained by Lomb-Scargle spectrum analysis (Fig. 5):

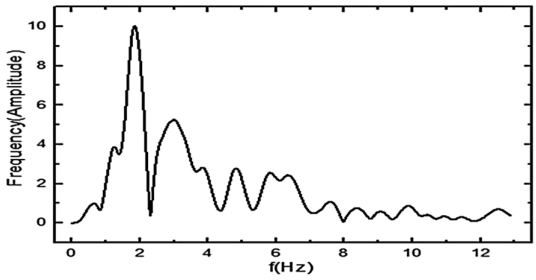


Fig. 5. Lomb-Scargle spectrum analysis

In Fig. 5, the horizontal axis is the frequency spectrum analysis frequency f, and the vertical axis is the frequency amplitude, the vertical reflection distance H is obtained from $f = 2H/\lambda$ (Fig. 5), and finally the snow depth is obtained from the difference between the height and the vertical reflection distance of the antenna.

3 Construction of Altai BDS Snow Monitoring Station

On December 28th, 2016, BDS Snow Monitoring Station (ALTA) was set up in the snow monitoring station of Xinjiang Altai Weather Station at 47°44′26″ north latitude and 88°4′17″ east longitude with an average altitude of 750 m. This is the first GNSS snow monitoring station in China, The BDS snow monitoring station adopts TRIMBLE NETR9, the antenna type is TRIMBLE59900, and the data is formally collected on January 1, 2017, the satellite data sampling interval is 1 s, and the satellite elevation angle is 0° (Fig. 6).

Fig. 6. Xinjiang Altai snow monitoring station

4 Monitoring and Analysis of BDS-MR Snow Depth by Different Beidou Satellite Types

Since the official launch of the Beidou satellite system in 1993, the Beidou satellite system has been continuously improved, mainly consisting of the geostationary orbit satellite GEO and the non-geostationary orbit satellite MEO and IGSO [17], based on the principle of BDS-MR snow depth detection, since the GEO residual sequence does not generate periodic oscillations, the MEO and IGSO satellite selections are analyzed. Select Beidou data from January 1 to March 30, 2017 in Altai, Xinjiang, since most of the current ground-based GNSS receivers have a satellite elevation angle of 5° and a sampling rate of 30 s; the same settings are made in this experiment.

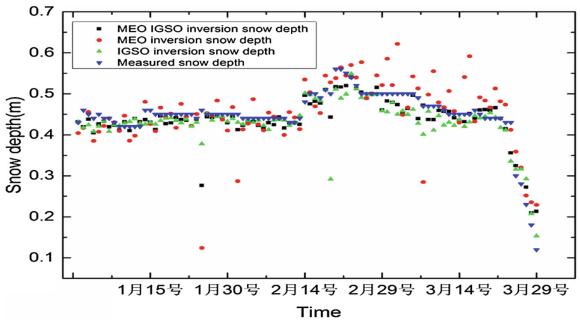


Fig. 7. Comparison of inversion snow depth of different Beidou satellite types and measured snow depth

In Fig. 7, blue represents the measured snow depth data, black is MEO IGSO, red is MEO, and green is the inversion snow depth of IGSO BDS-MR technology. Combining the Fig. 7 and correlation coefficient, the worst mean, and the Root Mean Squared Error (RMSE) in Table 1, it can be clearly seen that the results obtained by combining MEO with IGSO are the closest to the measured values, the results obtained by the MEO are more discrete and not stable enough, the results obtained by the IGSO

Table 1. Inversion of snow depth of different Beidou satellite types and measured snow depth statistics snow monitoring station

	MEOIGSO	IGSO	MEO
R	0.945	0.869	0.686
Mean	0.021	0.024	0.038
RMSE	0.031	0.037	0.059

are better, but the correlation coefficient with the measured snow depth is slightly lower than that obtained by the MEO IGSO, This may be related to satellite performance and the number of satellites.

5 Conclusions

In this paper, the BDS snow monitoring station in Altai is used to analyze the influence of different satellite type on the BDS-MR snow depth results, using 90 days data, the statistical analysis of correlation coefficient, poor mean, and RMSE proves the feasibility of BDS-MR, the following conclusions were obtained:

1. BDS-MR technology can use Beidou data to accurately invert snow depth.
2. Using the MEOIGSO satellites, the BDS-MR technology utilizes more satellites, has a larger base of residual sequence oscillations, has sufficient information to detect snow depth, and improves the accuracy of the inversion.

This paper uses BDS-MR technology to invert the 90 days data of the Altai BDS snow monitoring station in Xinjiang, and analyzes the experimental results from three different satellite type the accuracy of the experimental results obtained by selecting the MEOIGSO satellite and the snow depth accuracy can reach 2.1 cm, it proves that BDS-MR technology has good feasibility, is beneficial to promote the application of Beidou system, and fully utilizes the advantages of traditional ground-based geodetic GNSS receiver and microwave remote sensing.

Future research will focus on making full use of the multi-system and multi-band data of the Xinjiang GNSS snow monitoring station for integration, in the process of processing, it can be found that there are many systematic errors in the method, and the error correction model is proposed to be further explored. The impact of the environment around the site is also one of the research priorities.

Acknowledgments. We gratefully acknowledge the Northwest Institute of Ecology and Environmental Resources, Chinese Academy of Sciences and the Altay Meteorological Bureau for providing experimental data. This work was supported by China Desert Meteorological Science Research Foundation (Sqj2017002), National Science Foundation of China (41731066, 41674001, 41104019) and the Special Fund for Basic Scientific Research of Central Colleges (310826172202).

References

1. Zhang J, Deng Z (1987) Introduction to Precipitation in Xinjiang. Meteorological Press, Beijing, pp 205–243
2. Yang T, Lu G, Li H et al (2011) Progress in the prediction of hydrological extreme events under climate change. Adv Water Sci 22(2):279–286
3. Xu J, Shu H, Liu Y (2014) Analysis of spatial and temporal autocorrelation of snow disasters in Xinjiang from 2000 to 2010. J Catastrophology 29(01):221–227
4. Sha MW (2016) Research on high-precision inversion method based on AMSR2 passive microwave snow parameters. J Glaciol Geocryol 38(01):145–158
5. Bilich A, Larson KM (2007) Mapping the GPS multipath environment using the signal-to-noise ratio (SNR). Radio Sci 42(6):1–16
6. Larson KM et al (2009) Can we measure snow depth with GPS receivers? Geophys Res Lett 36(17), Article ID L17502
7. Nievinski FG, Larson KM (2014) Inverse modeling of GPS multipath for snow depth estimation—part i: formulation and simulations. IEEE Trans Geosci Remote Sens 52 (10):6555–6563
8. Ao M, Hu Y, Liu Y et al (2012) Inversion of soil moisture change trend observed by GPS signal-to-noise ratio. J Surv Mapp 29(2):140–143
9. Wu J, Yang R (2012) Measurement of water surface height using GPS receiver reflected signals. Geodesy Geodyn 32(6):135–138
10. Jin S, Najibi N (2014) Sensing snow height and surface temperature variations in Greenland from GPS reflected signals. Adv Space Res 53(11):1623–1633
11. 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
12. Zhang S, Dai K, Liu K, Hou X, Zhao Y (2016) GPS-MR technology for snow thickness monitoring research. Prog Geophys 31(04):1879–1884
13. Dai K, Zhang S, Li Z, Zhao Y, Nan Y (2016) Study on snow depth monitoring using GPS SNR observations. J Geodesy Geodyn 36(06):525–528
14. Katzberg SJ, Torres O, Grant MS, Masters D (2006) Utilizing calibrated GPS reflected signals to estimate soil reflectivity and dielectric constant: results from SMEX02. Remote Sens Environ 100(1):17–28
15. Liu Z, Huang Z, Jin J (2008) Experimental analysis of improving GPS positioning accuracy using satellite height angle and signal-to-noise ratio. Surv Mapp Eng 17(4):54–58
16. Larson KM, Nievinski FG (2013) GPS snow sensing: results from the earth scope plate boundary observatory. GPS Solutions 17:41–52
17. Tang J (2013) Overview of the development and application of Beidou satellite navigation regional system. Glob Position Syst 38(05):47–52

Monitoring the Study of Offshore Sea Level Changes with GPS-MR Technology

Shuangcheng Zhang[1,2(✉)], Xuerong Chen[2,3(✉)], Yang Nan[4], and Qi Liu[2,3]

[1] China Meteorological Administration Urumqi Institute of Desert Meteorology, Urumqi, China
shuangcheng369@vip.163.com
[2] College of Geology Engineering and Geomantic, Chang'an University, Xi'an, China
821263323@qq.com
[3] Xinjiang Water Resources and Hydropower Survey and Design Institute, Urumqi, China
[4] Wuhan University Satellite Navigation and Positioning Technology Research Center, Wuhan, China

Abstract. The monitoring of sea level change is a hot issue in the field of marine environment and global climate change research. With the rapid development of GNSS theory and application, GPS-MR technology based on multipath effect has been proved by scholars to be used for tidal level monitoring. In order to further expand the application space of GPS-MR technology in the field of marine environment monitoring, this paper conducts research on offshore sea level change based on tidal level change data acquired by GNSS-MR. This paper first gives the calculation process of GPS-MR technology to obtain the daily average sea surface, and then obtains the high-precision daily average sea surface and month by using the measured data of GNSS at different time intervals and different frequencies on the shore of the Harbour Harbor in Washington, USA. The average sea surface is in good agreement with the results of the tide check station. The experimental results show that GPS-MR technology has higher precision for obtaining tidal level change and regional mean sea surface change, and lays a foundation for GPS-MR technology to supplement satellite altimetry for offshore sea level change and to carry out offshore sea tide model refinement.

Keywords: GPS multipath reflection · SNR · Tidal level change · Daily average sea surface

1 Introduction

Sea level is a very sensitive indicator of climate change that reflects the components of the climate system, such as heat, melting of glaciers and ice sheets. Sea level rise poses enormous challenges and threats to human survival and economic development in coastal areas. Monitoring sea level changes not only plays an important role in understanding the climate, but also helps to understand the social and economic

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 31–39, 2019.
https://doi.org/10.1007/978-981-13-7751-8_4

consequences of sea level rise. The relationship between tides and navigation is also very important. It will directly affect the implementation of the navigation plan and the safety of navigation. If it is necessary to pass the shallow water area, it is necessary to calculate the local tide height and tide time based on the tide data and adjust the draft difference correctly; Sailing safely on planned routes requires constant knowledge of local tides and trends. Even in Hong Kong, it is impossible to ignore the impact of tides and trends on ship safety. In the coastal navigation, the captain's navigational order, the company's navigational rules and regulations, and the guiding documents of the international organization's pilots on the navigational watch will all be aware of the current and future tides and tides as the safe work of the bridge important content [1–4].

As early as the beginning of the 21st century, the interferometric measurement of GPS signals proposed by Anderson et al. and can be used to obtain the distance of the GPS antenna from the water surface and get the experimental result of RMS of 12 cm [5]; Larson et al. use GPS-MR technology in Two stations conducted experiments on tidal level monitoring and found that stations with smaller changes in tidal level have higher accuracy [6, 7]; Nakashima and Heki used SNR-based GPS-MR technology to detect sea level change using a single satellite, and get the result of RMS of 27 cm [8]; Löfgren et al. used the wind speed as an indicator of sea surface roughness, SNR analysis performed better than phase delay analysis under rough sea conditions [9]. Taking into account the dramatic changes in the sea surface, Löfgren et al. also used the improved LSP spectrum analysis method to reflect the sea surface tidal level of five GNSS stations in different environments around the world, and found that the improved method When used in stations with large tidal range changes, the results are significantly better than the simple LSP spectrum analysis method [10]; Strandberg et al. use the B-spline method to further enhance the accuracy of tidal level monitoring based on the fusion of multi-mode data. Santamaría-Gómez et al. added a tropospheric correction by laying side antennas, improved Kalman filtering and smoothing algorithm to provide a RMS of 3 cm [12]; Jin et al. based on SNR observations, tri-band combination and code combination BDS-R technology, and used for tidal monitoring [13]; Yu et al. proposed a three-frequency combined observation snow depth detection technology, and use different station environment the experiment is verified [14]; Wang et al. proposed a cGNSS-R epoch measuring algorithm based on single or double difference observation combination, which can effectively weaken the clock error and the ionospheric and tropospheric error to reflect the impact of water surface height [15].

In this paper, the long-term tidal sequence obtained by GNSS-MR technology is analyzed and processed to obtain the average sea surface of the inversion region, which lays a foundation for the subsequent use of the shore-based GNSS station to monitor the regional sea level change and carry out the research on the refined offshore tide model.

2 Basic Theory of Monitoring Tidal Level Changes with GNSS-MR

GNSS multipath effect is regarded as the main error source of high precision positioning. The generation of GNSS multipath is mainly related to the structure and dielectric parameters of the reflective surface. When the GNSS station is located near

the sea surface, the signal received by the GNSS is actually a direct signal and a composite signal reflected by the sea surface.

The GPS receives the signal synthesized by the direct signal and the sea surface reflected signal. The amplitudes of the direct signal and the reflected signal are A_d and A_m respectively. For the measurement type GNSS receiver antenna, the amplitudes of the direct signal and the reflected signal have the following relationship

$$A_d \gg A_m \tag{1}$$

Combined with Fig. 1, among the composite signals A_c captured by the GPS receiver, the direct signal A_d determines the overall trend of the composite signal, that is the overall trend term of the SNR observation, while the reflected signal A_m appears as a local periodic oscillation, which is considered to be mainly due to the low Height angle multipath effect. The SNR observation is the accompanying observation data of the geodetic GPS receiver. The relationship between SNR and signal amplitude is as follows.

$$SNR = A_c^2 = A_d^2 + A_m^2 + 2A_dA_m\cos\theta \tag{2}$$

A_c is the amplitude of the composite signal, and $\cos\theta$ is the cosine of 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, the multipath effect needs to be separated from the received SNR observations. Combining Eq. (1), the influence of multipath effect on signal-to-noise ratio is small (mainly at low elevation angle), that is, A_d and A_m are greatly different in value, and low-order polynomial can usually be used to eliminate trend term A_d.

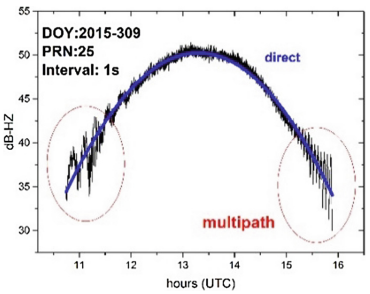


Fig. 1. SNR variation map of PRN25 GPS satellite.

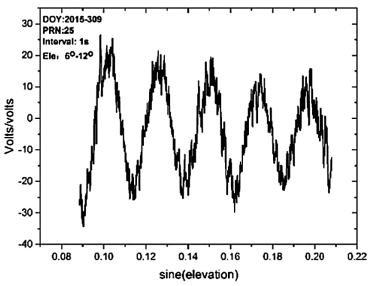


Fig. 2. Remove the SNR residual sequence of the trend term

Figure 1 shows the SNR variation of the 2015 lakeside GPS station PRN25 satellite with a sampling rate of 1 s. It can be seen from Fig. 1 that the overall trend term of SNR is in parabolic form and can be fitted by quadratic polynomial; the fluctuation of both ends of SNR (GPS satellite rising or falling) is mainly caused by multipath effect at low elevation angle. After the overall trend term in Fig. 1 is removed, the SNR residual sequence caused by multipath effects under low elevation angle conditions can be obtained.

Assume that the amplitude of the SNR residual sequence affected by multipath in Fig. 2 can be expressed as

$$A_m = A\,cos\left(\frac{4\pi h}{\lambda} sinE + \phi\right) \tag{3}$$

In Eq. (3), λ is the carrier wavelength, E is the satellite elevation angle, and h is the vertical reflection distance. If $t = sinE$, $f = \frac{2h}{\lambda}$, then Eq. (3) can be simplified to the standard cosine function expression

$$A_m = A\,cos(2\pi ft + \phi) \tag{4}$$

In the Eq. (4), the vertical reflection distance parameter h is included in the frequency f. If the Lomb-Scargle spectrum analysis is performed on the Eq. (4), the frequency f can be obtained, and the antenna phase center to the instantaneous tide level can be obtained vertical distance h.

By performing LSP spectrum analysis on the SNR residual sequence in Fig. 2, the frequency f of the SNR residual sequence can be obtained, and the vertical reflection distance h can be obtained by $f = 2h/\lambda$, and then the change of h is converted into the change of the tidal level. Thereby, real-time monitoring of tidal level changes from SNR observations is achieved.

With the GNSS system providing more than 100 satellite multi-band signals and the number of shore-based CORS stations, the use of GNSS-MR technology to obtain tidal level changes has a high spatial and temporal resolution. The tide gauge station needs to combine GNSS positioning technology and leveling technology to obtain the absolute change of the tide level under the ITRF framework, while the shore-based GNSS station can also measure the vertical deformation of the earth's crust, so GNSS-MR technology has the potential to obtain absolute tide level changes [16–18]. At the same time, GNSS-MR technology can also obtain the tidal wave coefficient, and has the advantages of not contacting with seawater.

3 Basic Principles for Obtaining Sea Level Change Using GNSS-MR Tidal Level Changes

To obtain the daily average sea surface, the long-term trend $a(t - t_0)$ and noise r of sea surface change can be ignored, so Eq. (1) can be expressed as:

$$\zeta(\mathrm{t}) = h_0 + \sum\nolimits_i R_i \cos[\sigma_i t - \mathrm{g}_i] \tag{5}$$

In Eq. (5), h_0 is the average sea surface to be calculated, the second term is the individual tides, R is the average amplitude in the calculated time range, σ is the angular rate of each tide, and t is the first from the observation. The number of hours counted from zero in day, g indicates the initial phase of the tide in the time range to be calculated, and the average sea surface is calculated based on the data of the actual tide

station $\zeta(t)$. The main principle is to use the difference of the tide-dividing cycle to eliminate by linear combination various tides. In this paper, the Dudson method is used to calculate the daily mean sea surface. This method can theoretically eliminate the influence of the sun and the Taiyue tidal system on the daily average water level.

Long-term tidal level variation sequences can be obtained by using GNSS-MR technology. Since the tidal bit sequences obtained by GNSS-MR technology are not equally spaced samples, the sequence needs to be interpolated and resampled into a new sequence of 1 h sampling interval.

4 Experimental Analysis of GNSS-MR Offshore Sea Level Change

4.1 Experimental Data Source

In order to verify the accuracy of the GNSS-MR technology to obtain the average sea surface, this paper uses the observation data of the GNSS continuous operation tracking station SC02 station (in Fig. 3) located on the shore of the Harbour Harbor in Washington, USA. The SC02 observatory is affiliated with the United States. The PBO network is observed at the edge of the plate in the Earth Scope. The SC02 station is built adjacent to the sea (in Fig. 4). It can receive GPS reflections from the sea surface in a large space. The station has more than ten years of continuous observation data. Figure 3 shows the receiver layout and observation environment of the SC02 station. The receiver used in the SC02 station is the TRIMBLE NETR9 geodetic receiver. The GPS antenna is the Trimble company's choke coil antenna with a fairing (SCIT) (TRM59800.80).

In order to verify the accuracy of GNSS-MR technology to obtain the average sea surface and tidal wave coefficient, this paper makes a comparative analysis using the measured data of the Friday Harbor tide station at 359 meters from the SC02 station. The Friday Harbor Tidal Checkpoint is a continuous operation tide gauge station maintained by the National Oceanic and Atmospheric Administration's Marine Products and Service Center. The tide gauge station was built in 1934 and was equipped with Aquatrak's acoustic test in 1996. The tide gauge can provide tidal level data with a sampling interval of 6 min. Figure 4 shows the relative positional relationship between the tide gauge station and the GNSS station.

Fig. 3. SC02 station observation environment

Fig. 4. SC02 station and tide gauge station relative position

In order to use the GPS data of different time spans for experimental analysis, this paper uses the GPS data of 30 s sampling interval of one month and twelve years to compare with the data of tide gauge station.

4.2 Experimental Analysis of Daily Tide Level Change Using GPS-MR

In Fig. 5, the signal-to-noise ratio data of different frequency bands L1 and L2 of the SC02 station in May 2013 (day 121 to day 151) are used to obtain the daily tide level change and the Dudson method, and with the Friday Harbor tide station. The results obtained by the Dudson method are compared.

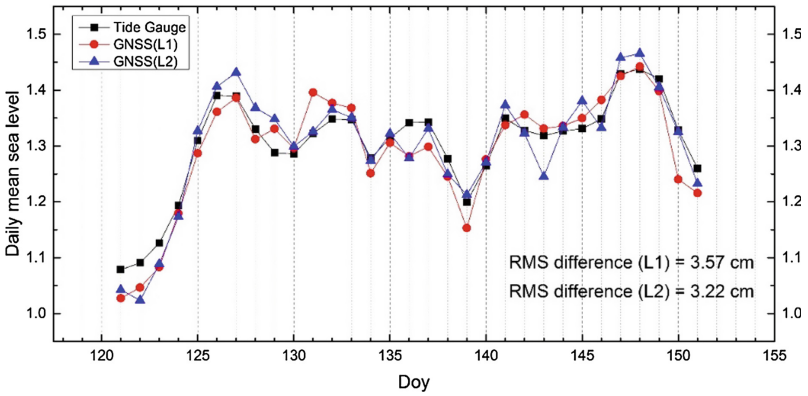


Fig. 5. Comparison of sea level changes between tide station and GNSS-MR in May 2013

In Fig. 5, the horizontal axis represents the annual product day, the vertical axis represents the daily average sea surface relative to the lowest low tide surface, the black point is the daily average sea surface acquired by the Friday Harbor tide station, and the red point is the day obtained by the L1 band of the GPS data of the SC02 station. The average sea surface, the blue point is the daily average sea surface obtained in the L2 band. It can be seen from Fig. 7 that the average daily sea surface obtained in the L1 and L2 bands is in good agreement with the daily average sea surface obtained by the tide station, and the poor average of the results obtained by the L1 band and the tide station is 2.93 cm, the RMS of the difference is 3.57 cm. The poor average of the L2 band acquisition results compared with the tide station results is 2.53 cm, and the difference RMS is 3.22 cm.

4.3 Experimental Analysis of Mean Sea Level Change Based on GNSS-MR Tidal Level Acquisition Area

In order to further verify the effectiveness of GNSS-MR technology to monitor regional sea level change, the monthly average sea surface from 2005 to 2017 was obtained by using the signal-to-noise ratio of L1 frequency band according to the flow of Fig. 4,

and compared with the results of tide gauge station. (in Figs. 6 and 7). In order to obtain the average sea level change rate, the monthly average sea surface sequence is generally fitted by the least squares method, and the slope of the fitted straight line is the average sea level change rate. Figure 6 shows the monthly average sea surface change of the GNSS-MR from SC02 station from 2005 to 2017, and Fig. 7 shows the monthly mean sea level change of the Friday Harbor tide station.

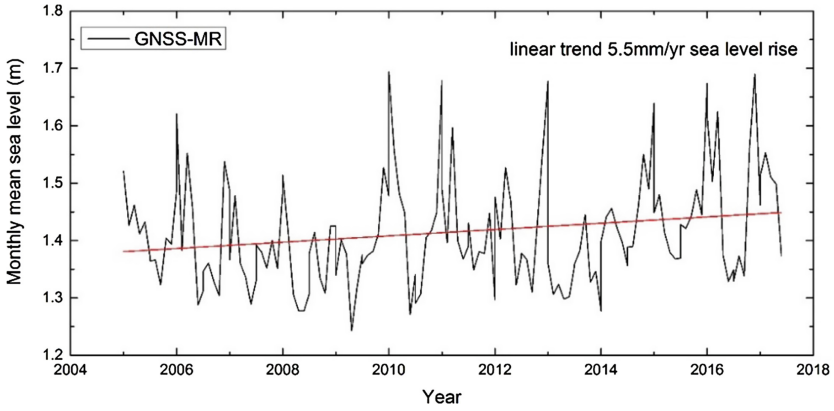


Fig. 6. Monthly mean sea level change of GNSS-MR from 2005 to 2017

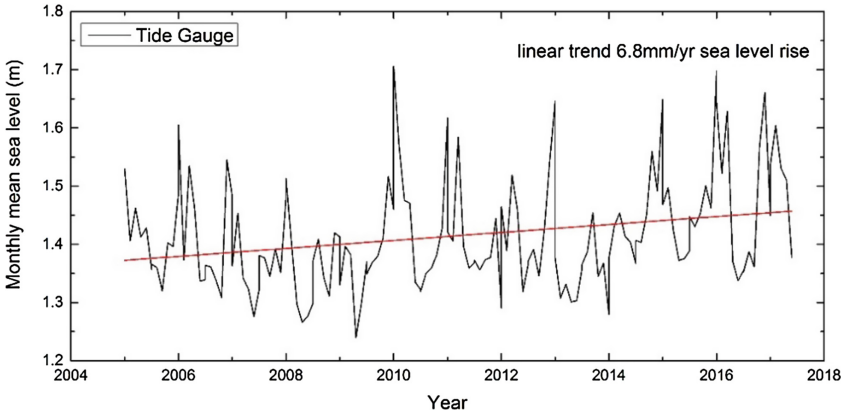


Fig. 7. Comparison of monthly mean sea level changes at tide stations from 2005 to 2017

The black lines in Figs. 6 and 7 are the monthly mean sea surface sequences obtained by the SC02 station and the Friday Harbor station respectively, and the red line is the straight line fitted by the least squares method. The monthly mean sea surface change rate obtained by the SC02 station is 5.5 mm/yr, and the monthly average sea surface change rate obtained by the Friday Harbor station is 6.8 mm/yr. From the

experimental results of Figs. 6 and 7, it can be seen that the GNSS-MR technology agrees well with the monthly average sea surface change rate obtained by the tide gauge station, and verifies the feasibility of the GNSS-MR technique to obtain the regional average sea surface change rate.

5 Conclusion

Based on the existing research results, and the basic theory of daily mean sea surface acquisition, this paper gives the method of GNSS-MR technology to obtain the average sea surface change. The accuracy of GNSS-MR technique for obtaining average sea surface change is analyzed by an example. The following conclusions are drawn:

(1) GNSS-MR technology can effectively obtain long time series of tidal level, which is in good agreement with the results of tide gauge station, so GNSS-MR technology provides an effective supplementary means for obtaining tidal level changes.
(2) The average sea surface obtained by GNSS-MR technology obtained from the average sea surface and the tide gauge station is in good agreement with the overall trend. The long-period monthly average sea surface change rate obtained by GNSS-MR was solved by least squares method, and compared with the monthly mean sea surface change rate obtained from the Friday Harbor tide station at 359 meters, the GNSS-MR technique was used to obtain the regional mean sea surface change. The effectiveness of the rate. To lay the foundation for the GNSS-MR technology to obtain the absolute change of regional average sea surface.

GNSS-MR technology has the advantages of high time resolution, no contact with seawater, and multi-purpose use in obtaining tidal level and regional average sea surface. At the same time, the high-precision acquisition of regional sea level changes complements the satellite altimetry for GNSS-MR technology and lays a foundation for the refinement of offshore tide models.

Acknowledgments. We gratefully acknowledge the provision of data, equipment, and engineering services by the Plate Boundary Observatory operated by UNAVCO for EarthScope. This work was supported by China Desert Meteorological Science Research Foundation (Sqj2017002), the National Science Foundation of China (41731066, 41674001, 41104019) and the Special Fund for Basic Scientific Research of Central Colleges (310826172202).

References

1. Chen Z (1980) Tideology. Science Press
2. Li D, Li J, Jin T et al (2012) Monitoring global sea level change from 1993 to 2011 using multi-generation satellite altimetry data. J Wuhan Univ: Inf Sci Ed 37(12):1421–1424
3. Jin T, Li J, Jiang W et al (2011) A new generation global average sea surface height model based on multi-source satellite altimetry data. J Surv Mapp 40(06):723–729

4. Zhang H, Bao J, Zhou X et al (2012) Discussion on the construction method of seamless vertical reference in china's sea area. Surv Sci 37(1):18–19+22
5. Anderson KD (2000) Determination of water level and tides using interferometric observations of GPS signals. J Atmos Oceanic Technol 17(8):1118–1127
6. Larson KM, Löfgren JS, Haas R (2013) Coastal sea level measurements using a single geodetic GPS receiver. Adv Space Res 51(8):1301–1310
7. 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
8. Nakashima Y, Heki K (2013) GPS tide gauges using multipath signatures. J Geodetic Soc Jpn 59(4):157–162
9. Löfgren JS, Haas R (2014) Sea level measurements using multi-frequency GPS and GLONASS observations. EURASIP J Adv Sig Process 2014(1):1–13
10. Löfgren JS, Haas R, Scherneck HG (2014) Sea level time series and ocean tide analysis from multipath signals at five GPS sites in different parts of the world. J Geodyn 80:66–80
11. Strandberg J, Hobiger T, Haas R (2016) Improving GNSS-R sea level determination through inverse modeling of SNR data. Radio Sci 51(8):1286–1296
12. Santamaría-Gómez A, Watson C (2017) Remote leveling of tide gauges using GNSS reflectometry: case study at Spring Bay, Australia. GPS Solutions 2017:1–9
13. Jin S, Qian X, Wu X (2017) Sea level change from BeiDou Navigation Satellite System-Reflectometry (BDS-R): first results and evaluation. Global Planet Change 149:20–25
14. Yu K, Ban W, Zhang X 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
15. Wang N, Bao L, Gao F (2016) Evangelical GNSS-R altimetry single difference and double difference algorithm. J Surv Mapp 45(7):795–802
16. Wu F, Wei Z, Li Y et al (2015) Tidal analysis of Dagang tide station and national elevation datum change. J Surv Mapp 44(7):709–716
17. Jiao W, Wei Z, Guo H et al (2004) Joint GPS base station and tide gauge data to determine absolute changes in sea level. J Wuhan Univ (Inf Sci Ed) 29(10):901–904
18. Zhou D, Zhou X, Zhang X et al (2016) Vertical deformation analysis of the crust of the coastal tide station in China using GPS continuous observation. J Wuhan Univ (Inf Sci Ed) 41(4):516–522
19. Zhang S, Nan Y, Li Z et al (2016) Monitoring and analyzing tidal level changes with GNSS-MR technology. J Surv Mapp 45(9):1042–1049
20. Munekane H (2013) Sub-daily noise in horizontal GPS kinematic time series due to thermal tilt of GPS monuments. J Geodesy 87(4):393–401

Retrieval of Precipitable Water Vapor Using GNSS Data Under Conditions Without Collocated Meteorological Observations

Zhaozhe Li(✉), Wujiao Dai, Biyan Chen, and Yaxin Wen

Department of Survey Engineering and Geomatics, Central South University, Changsha, China
812005520@qq.com

Abstract. Water vapor is the most active component of the atmosphere, playing a significant role in formation of severe weather such as thunderstorms and hail. Therefore, accurate water vapor information is of great significance for studying its temporal and spatial variation and regional weather forecasting. Obtaining atmospheric PWV from GNSS observations requires two key meteorological parameters weighted mean temperature and surface pressure. Meteorological parameters are typically measured by sensors deployed at the GNSS station. However, a large number of GNSS stations are mainly used for positioning, and are not equipped with meteorological sensors, so that the water vapor cannot be accurately estimated. In this paper, the GPT2w model and the method of meteorological data interpolation are used to simulate the meteorological data observation. It is found that the GPT2w model and the meteorological data interpolation can accurately retrieve the water vapor content, but the method of interpolating meteorological data is more accurate than the GPT2w model. In addition, the paper also analyzes the interpolation accuracy from four aspects: the number of interpolated stations, the distance between the interpolated stations, the height difference between the interpolated stations, and the weather conditions. It is found that when the height difference between the interpolated and target stations is small, the meteorological data obtained by interpolating from a short distance station is more accurate. In terms of weather conditions, the interpolated value of sunny day is more accurate; the interpolation accuracy is not directly related to the number of interpolated stations used.

Keywords: GNSS PWV · Meteorological observation · Meteorological data interpolation · Weighted mean temperature

1 Introduction

Ground-based GNSS water vapor retrieval technology [8] was proposed by Bevis in the 1990s, and it has been proved to be a means to effectively detect atmospheric water vapor. Ground-based GNSS observations have the advantages of high temporal resolution, all-weather, high precision, and low cost. And thanks to the ground-based GNSS technology, measurement precision can be obtained comparable to that of

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 40–51, 2019.
https://doi.org/10.1007/978-981-13-7751-8_5

radiosonde stations and ground-based microwave radiometers. Ground-based GNSS technology can be used for weather forecasting, climate analysis and other applications.

Regarding the water vapor retrieval without meteorological data, scholars have established various tropospheric empirical models based on years of measured meteorological data, that is, without prior astronomical data, the local meteorological data is obtained directly by inputting regional location and time information. Schüler et al. [6] established the Trop Grid2 model based on meteorological data from the National Center for Environmental Prediction (NCEP); Böhm et al. [1] established the most accurate tropospheric empirical model based on re-analysis meteorological data from the European Centre for Medium-Range Weather Forecasting (ECMWF). Chen et al. [2] constructed a map of precipitable water vapor (PWV) in Hunan Province based on data from 58 GNSS states, radiosonde data and ECMWF reanalysis data, using meteorological data interpolation methods. While the influence of the distance between the interpolation station and the target station, the height difference, the number of interpolation stations and the weather conditions on the precision of the interpolation results are not evaluated.

And this paper conducts research based on the above four aspects.

2 Zenith Tropospheric Delay (ZTD), Zenith Hydrostatic Delay (ZHD), Zenith Wet Delay (ZWD) and PWV

GNSS signals are affected by the ionosphere and troposphere as they traverse the atmosphere. The tropospheric delay of the GNSS signal during the propagation process [9] is closely related to the atmospheric water vapor content. Therefore, the atmospheric water vapor content can be inverted by studying its delay [5].

As stated above, atmospheric delays can be divided into ZTD, ZHD and ZWD. ZTD can be calculated by Formula (1):

$$\Delta L^0 = 10^{-6} \int N_d ds + 10^{-6} \int N_w ds = \Delta L_h^0 + \Delta L_w^0 = ZTD \tag{1}$$

ZHD is generally established by a model. This paper uses the Saastamoinen dry delay model to calculate the dry delay:

$$ZHD = 0.2277 \cdot \frac{P_s}{F(\phi, H)} \tag{2}$$

$$F(\phi, H) = 1 - 0.0026 \cdot \cos(2\phi) - 0.00028H \tag{3}$$

ZWD is multiplied by the conversion coefficient Π to get PWV, which is expressed by the following formula:

$$ZWD = ZTD - ZHD \tag{4}$$

$$PWV = \Pi \cdot ZWD \tag{5}$$

3 Experimental Program and Result Analysis

Firstly, the Bernese software is used to solve ZTD, and then the weighted mean temperature is seasonally fitted according to the radiosonde data, and then ZHD is calculated according to the Saastamoinen dry delay model combined with the measured surface pressure. Then ZWD is calculated, and the conversion coefficient is calculated in combination with the weighted mean temperature fitted in the season, and finally the atmospheric precipitation PWV is obtained.

3.1 Experiment Data

This paper uses the observations and meteorological data of 9 GNSS stations in Hong Kong (time interval of observation is 5 s, and that of meteorological data is 1 min) over the year of 2017 as the basic data. The distribution of 9 GNSS stations and 45004 radio sounding stations in Hong Kong are shown in Fig. 1:

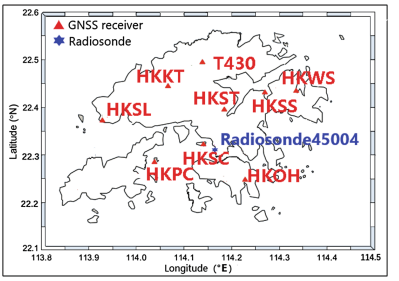


Fig. 1. Distribution of 9 GNSS stations and radiosonde45004 stations in Hong Kong

3.2 Weighted Mean Temperature Fitting

Since the weighted mean temperature T_m is important in solving PWV. The following focuses on the acquisition of the weighted mean temperature. The weighted mean temperature [7] is mainly obtained by the absolute temperature and the absolute pressure e, which is expressed by:

$$T_m = \int \frac{e(h)}{T(h)} dh \Big/ \int \frac{e(h)}{T^2(h)} dh \tag{6}$$

Where $e(h), T(h)$ represent temperature and surface pressure at different heights.

Although the Formula (6) is the formula for accurately calculating the weighted mean temperature, it requires the water vapor pressure and absolute temperature of many different heights in a certain area to be obtained. The calculation is very

inconvenient, so in 2007, Chen [3] used 8-year radiosonde data of Hong Kong to fit the relationship between the weighted mean temperature T_m (K) and the absolute temperature T (K) in Hong Kong:

$$T_m = 106.7 + 0.605T_S \tag{7}$$

Additionally, this paper uses the data of the Hong Kong Sounding Station (called radiosonde45004) downloaded from the University of Wyoming's radiosonde website to calculate the weighted mean temperature based on the absolute temperature between the years of 2007 and 2016 by Formula (6). Moreover, this paper fits the relationship between T_m and T in four seasons these 10 years (Fig. 2):

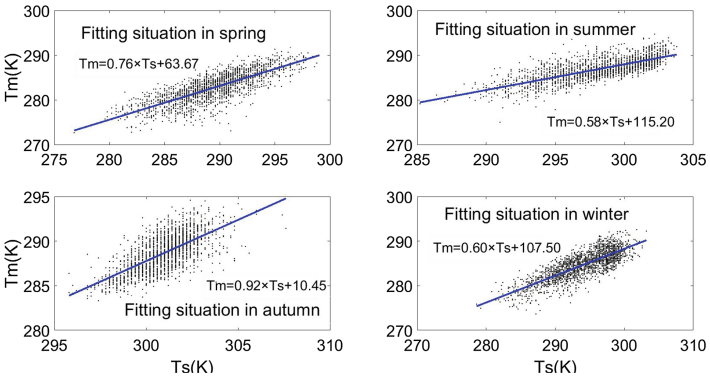


Fig. 2. The seasonal fitting of the weighted mean temperature

Since there was only meteorological data at 0:00 and 12:00 per day at Radiosonde45004, this paper uses the data at 0'clock to compute meteorological parameters. Additionally, this paper uses calculated value of the seasonal fitting formula and Eq. (7) to compare the precision. The comparison of the calculation results of each formula is shown in Fig. 3 and Table 1:

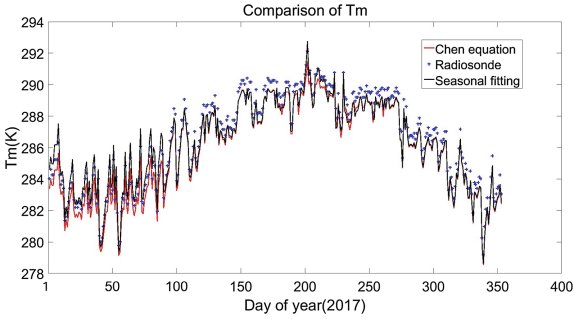


Fig. 3. Time series of Tm measured by Radiosonde45004, Chen's classic formula fitting value and seasonal fitting value over the whole year of 2017

Table 1. Precision comparison of the weighted mean temperature of radiosonde station in Hong Kong with that of seasonal fitting and Chen's classical formula

Comparison	Bias(K)	RMS(K)	Max(K)	Min(K)
Radiosonde vs Chen's equation	−0.68	1.85	3.89	−5.98
Radiosonde vs Seasonal fitting	−0.30	1.63	4.35	−5.59

It can be seen from Fig. 3 and Table 1 that the weighted mean temperature precision of the seasonal fitting is higher, the RMS error of the seasonal fitting is 1.63 K, and the RMS error of the Eq. (7) is 1.85 K. The Bias of the seasonal fitting is also better than that calculated by Eq. (7). In the following, the seasonal fitting formula is used to obtain the weighted mean temperature for the nine GNSS stations in Hong Kong. In addition, the GPT2w model [4] is the most accurate tropospheric empirical model at present, and in this paper, its calculated meteorological data are used to invert PWV as a comparative reference.

For Radiosonde45004 is closer to the hksc station (about 2 km away) and farther away from the other 8 stations. Therefore, it is unreasonable to use the data of radiosonde station as the true value to compare with that of the other eight stations. Therefore, PWV detected is firstly compared by the hksc station with that of Radiosonde45004:

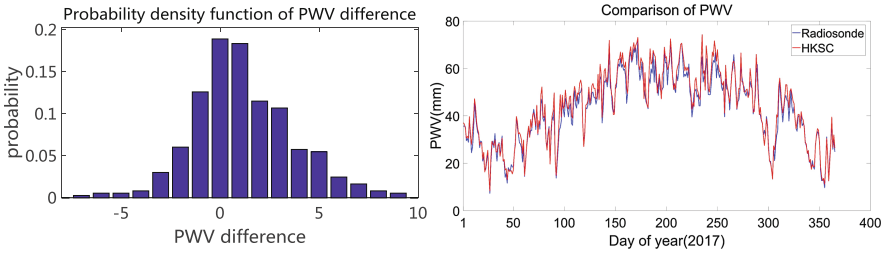


Fig. 4. (a) Time series of PWV measured by Radiosonde45004 and that of hksc over the whole year of 2017 (b) probability density function (PDF) of PWV difference

Table 2. Comparison of precision between PWV of Radio45004 and that of hksc station

GNSS station	Bias(K)	RMS(K)	Max(K)	Min(K)
hksc	1.82	3.33	9.62	−7.64

It can be seen from Table 2, Fig. 4(a) and (b) PWV obtained by the GNSS station inversion has good consistency with the value of the radiosonde station.

According to the numerical values and trends, it can be seen the Bias is 1.82 mm, the RMS value is 3.33 mm, and the PWV difference is also mainly distributed between −4 mm and 6 mm, and the frequency is about 95.62%.

3.3 Interpolation of Meteorological Data

In order to obtain PWV without meteorological observation data, the meteorological data firstly need to be obtained by the GNSS station near the area to interpolate T_m and P_s on the plane according to the interpolation formula, and then according to the fitted meteorological data. And calculate T_m and P_s of the height of the area based on the elevation fitting relationship. This paper selects the weighted mean temperature and pressure (true value) of hkoh, hkpc, hksl, hkws, t430 stations, and interpolates the meteorological data of the other four stations hkkt, hksc, hkss, hkst station, the interpolation formula is as follows [2]:

$$y_G = \frac{\sum_{i=1}^{n} \exp(-d_i^2) \cdot y_i}{\sum_{i=1}^{n} \exp(-d_i^2)} \tag{8}$$

Where n is the number of stations of known meteorological data, d is the distance of the known station from the unknown station, y is the interpolated data, y_G is the result of the interpolation, this is the weighted mean temperature T_m and the surface pressure P_s.

Since the formula is performed on the same plane, and the geodesic heights of 9 GNSS stations are different, it is necessary to firstly calculate the meteorological data to the same plane and then perform the interpolation calculation. In this paper, using the radiosonde data during 2007 from 2016, T_m and P_s of different heights are highly fitted within 1500 m, and their relationships are obtained.

The relationship between T_m and altitude is as follows:

$$T_m = T_{m_0} + 0.003068 \times (H_0 - H_S) \tag{9}$$

The relationship between P_s and altitude is as follows:

$$P_S = P_0 \times e^{-0.0001179 \times (H_S - H_0)} \tag{10}$$

The fitting situation is shown in Fig. 5(a) and (b):

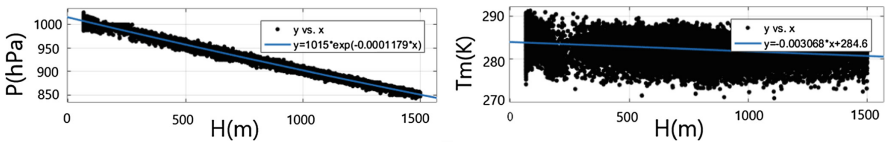


Fig. 5. (a) Fitting of surface pressure under 1500 m (b) Fitting of weighted mean temperature under 1500 m

In addition, this paper also uses the GPT2w model results, the interpolation results and the precision of the GNSS station for the precision comparison shown in Fig. 6, Table 3:

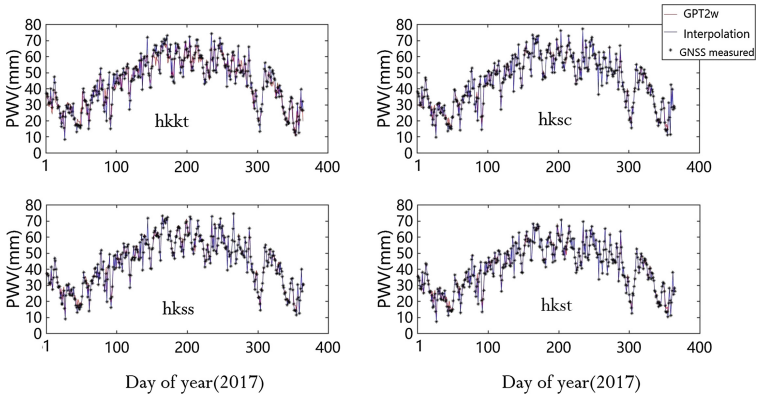


Fig. 6. Time series of true value, PWV interpolated by GNSS and measured by GPT2w model at 4 GNSS stations over the whole year of 2017

It can be seen from Fig. 6 and Table 3 that the precision indexes of the PWV values calculated by the interpolated meteorological data are better than the GPT2w model values, especially the Rms value and the Bias value. This is mainly because the GPT2w model is only an empirical model, and the simulated meteorological data is different from the local measured meteorological data.

Table 3. Precision comparison of PWV between interpolation and GPT2w measured

Comparison		hkkt	hksc	hkss	hkst
		PWV (mm)	PWV (mm)	PWV (mm)	PWV (mm)
Interpolation vs True value	Bias	−0.10	0.22	0.18	−0.36
	Rms	0.22	0.07	0.11	0.47
	Max	0.37	1.31	1.05	0.47
	Min	−1.03	−0.63	−0.21	−1.18
GPT2w vs True value	Bias	−0.54	−0.45	−0.61	−0.64
	Rms	1.42	1.38	1.43	1.37
	Max	4.80	5.85	4.77	5.16
	Min	−4.04	−4.04	−4.23	−4.05

3.4 Influence of the Number of Stations on the Precision of the Interpolation Result

Considering only the influence of the number of stations on the interpolation results, this paper uses data of other 2–5 GNSS stations to interpolate data of hkkt, hksc and hkss stations. The comparison of the interpolation results of meteorological data is shown in Table 4. It can be seen the number of stations used in meteorological data interpolation has no direct impact on the precision of interpolation.

Table 4. Precision comparison between meteorological data obtained by different interpolation stations and GNSS observed

Number of stations		hkkt		hksc		hkss	
		Ps (hPa)	Tm (K)	Ps (hPa)	Tm (K)	Ps (hPa)	Tm (K)
2	Bias	0.12	−0.32	0.01	0.11	0.39	−0.08
	RMS	0.17	0.65	0.13	0.37	0.48	0.68
	Max	2.58	1.66	0.82	1.41	1.19	2.29
	Min	−0.64	−2.95	−0.87	−1.62	−1.31	−2.58
3	Bias	0.12	−0.32	0.01	0.11	−0.17	0.31
	RMS	0.17	0.65	0.13	0.37	0.29	0.55
	Max	2.58	1.66	0.82	1.41	0.67	2.50
	Min	−0.64	−2.95	−0.87	−1.62	−2.02	−1.50
4	Bias	0.12	−0.32	−0.03	0.48	−0.17	0.31
	RMS	0.17	0.65	0.34	0.86	0.29	0.55
	Max	2.58	1.66	0.91	3.08	0.67	2.50
	Min	−0.64	−2.95	−2.77	−2.54	−2.02	−1.50
5	Bias	0.12	−0.32	−0.03	0.74	−0.17	0.31
	RMS	0.17	0.65	0.34	0.48	0.29	0.55
	Max	2.58	1.66	0.91	3.08	0.67	2.50
	Min	−0.64	−2.95	−2.77	−2.54	−2.02	−1.50

3.5 Influence of the Distance on the Precision of the Interpolation Result

Since only the influence of the inter-station distance on the interpolation result is considered. Table 5 shows the comparison of the PWV calculated by interpolated results of meteorological of different inter-station distances with the true values.

Table 5. Precision comparison between PWV measured by GNSS stations and that interpolated by different distance stations over the whole year of 2017 at 3 GNSS stations

Comparison		hkkt	hksc	hkss
		PWV (mm)	PWV (mm)	PWV (mm)
GNSS measured vs PWV interpolated by closer distance stations	Bias	−0.05	0.44	0.11
	RMS	0.11	0.49	0.18
	Max	0.31	1.14	1.05
	Min	−0.56	−0.81	−0.92
GNSS measured vs PWV interpolated by further distance stations	Bias	0.04	0.49	0.09
	RMS	0.16	0.56	0.35
	Max	0.60	1.31	1.58
	Min	−0.44	−0.83	−1.39

It can be seen from Table 5 the precision of the interpolation result of closer stations' data is better. The reason should be that as the distance between target station and interpolated station increases, the error of the value calculated by the Formula (8) will increase. Or it may be caused by inaccurate meteorological data of the interpolated station.

3.6 Influence of Different Elevation on the Precision of the Interpolation Result

Since only the inter-station height difference is considered for the result, only the hkst station and the hkoh station of the 9 stations' geographical heights are higher there, respectively 258.70 m and 166.40 m, so they are used as stations having the bigger inter-station height difference. And the data of other stations having smaller inter-station height difference is also used for interpolation. The results are shown in Fig. 7 and Table 6:

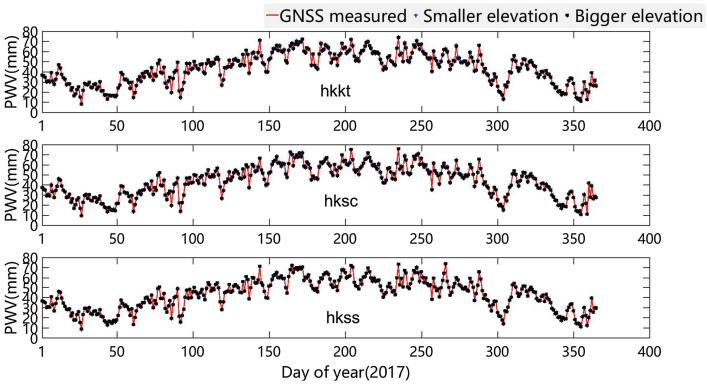


Fig. 7. Time series of PWV measured by GNSS station and interpolated by different elevation stations over the whole year of 2017 at 3 GNSS stations

It can be seen from Fig. 7 and Table 6 that PWV obtained by the meteorological data or the inversion calculation is interpolated using stations with small differences between the target stations, and the precision of the obtained results is better than that of the high difference. The precision of station interpolation results should be due to the high degree of fitting of meteorological data, which causes the error of the fitting value of meteorological data to be amplified due to the increase of the height difference.

Table 6. Precision comparison between PWV measured by GNSS station and that interpolated by different elevation stations over the whole year of 2017 at 3 GNSS stations

Comparison		hkkt	hksc	hkss
		PWV (mm)	PWV (mm)	PWV (mm)
GNSS measured vs PWV interpolated by smaller elevation difference stations	Bias	−0.10	0.12	0.11
	RMS	0.21	0.24	0.18
	Max	1.03	1.01	1.05
	Min	−0.87	−1.12	−0.97
GNSS measured vs PWV interpolated by bigger elevation difference stations	Bias	0.32	0.44	0.13
	RMS	0.45	0.49	0.34
	Max	1.28	1.14	1.65
	Min	−0.99	−0.91	−1.28

3.7 Influence of Weather Condition of the Interpolation Result

In the above analysis, the impact of weather conditions on the interpolation results is not considered. This paper uses the meteorological data of hksc, hkoh, hkst, hkpc stations on May 23rd to May 25th, 2017 (heavy rainstorm) and November 1st to November 3rd (sunny days) to conduct meteorological data for other 3 stations.

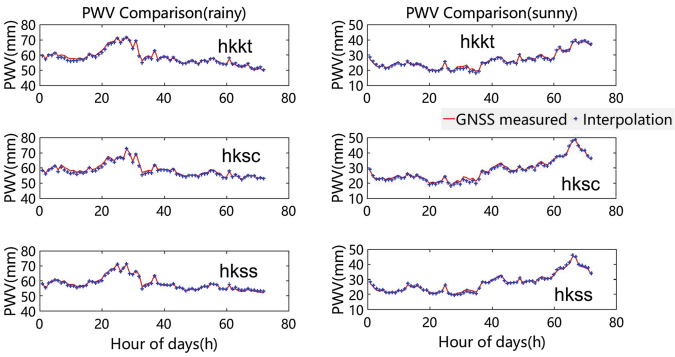


Fig. 8. Time series of PWV measured by GNSS station and interpolated on different weather conditions at 3 GNSS stations

It can be seen from Fig. 8 that the value of the PWV calculated by interpolation and its time trend are very close to the true value of the GNSS station, regardless of the weather. Maybe the distance between stations in Hong Kong GNSS station is short and the height difference is small, so that the interpolation algorithm can be used to obtain high-precision meteorological data (Table 7).

Moreover, the RMS errors are about 0.5 mm, and the Bias value are almost below 0.5 mm regardless of the weather. However, considering the four precision indicators

Table 7. Precision comparison between PWV interpolated on different weather conditions and GNSS measured at 3 GNSS stations

Station/weather	hkkt (sunny)	hkkt (rainy)	hksc (sunny)	hksc (rainy)	hkss (sunny)	hkss (rainy)
Bias (mm)	0.04	0.20	−0.38	−0.56	0.22	0.13
RMS (mm)	0.38	0.61	0.28	0.33	0.34	0.48
Max (mm)	1.58	1.63	1.62	1.97	1.59	1.85
Min (mm)	−0.92	−0.97	−1.89	−1.09	−1.68	−1.12

of the 3 GNSS stations, the precision of the interpolated values on sunny days is slightly better than that of rainstorm days.

4 Conclusions

This paper uses the observation files and meteorological data of Hong Kong GNSS station as the basic data in 2017, combines the weighted mean temperature seasonal fitting situation of Hong Kong radiosonde station from 2007 to 2016, and uses the GPT2w model and the method of weather data interpolation. The simulation of non-collocated meteorological observations also fits the relationship between the weighted mean temperature and the same height of the surface pressure at the altitude of 1500 m in Hong Kong Sound Station from 2007 to 2016 during the experiment of interpolating meteorological data. After the experiment, the following conclusions were obtained:

(1) When there is no meteorological data observation at all stations, the GPT2w model can be used to accurately invert the atmospheric precipitation.
(2) When there are meteorological data observations in some stations, and no meteorological data observations at other stations. The method of interpolating meteorological data can be used to reverse PWV more accurately.
(3) When interfering with meteorological data, the meteorological data obtained by interpolation with the target station is smaller and the distance is shorter, so that the obtained water vapor is closer to the true value.
(4) Precision of PWV obtained by interpolating meteorological data is relatively high, whether in rainy or sunny days. However, the PWV precision obtained by interpolating the weather data is higher at sunny days.

References

1. 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. https://doi.org/10.1007/s10291-014-0403-7
2. Chen B, Dai W, Liu Z et al (2018) Constructing a precipitable water vapor map from regional GNSS network observations without collocated meteorological data for weather forecasting. Atmos Measur Tech 11(9):5153–5166. https://doi.org/10.5194/amt-11-5153-2018

3. Chen YQ, Liu YX, Wang XY et al (2007) GPS real-time estimation of precipitable water vapor-Hong Kong experiences. Acta Geodaetica Cartogr Sin 36(1):9–12
4. Zhonghao H, Lintao L, Xinghui L (2017) An assessment of GPT2w model and fusion of a troposphere model with in situ data. Geomatics Inf Sci Wuhan Univ 42(10):1468–1473. https://doi.org/10.13203/j.whugis20150758
5. Jin S, Li Z, Cho J (2008) Integrated water vapor field and multiscale variations over china from GPS measurements. J Appl Meteorol Climatol 47(11):3008–3015. https://doi.org/10.1175/2008JAMC1920.1
6. Schüler T (2014) The TropGrid2 standard tropospheric correction model. GPS Solutions 18 (1):123–131. https://doi.org/10.1007/s10291-013-0316-x
7. Yao Y, Liu J, Zhang B et al (2015) Nonlinear relationships between the surface temperature and the weighted mean temperature. Geomatics Inf Sci Wuhan Univ 40(1):112–116
8. Zhao Q (2018) Studies on the key technologies in water vapor inversion using ground-based GNSS and its applications. Acta Geodaetica Et Cartographica Sinica 47(3):424. https://doi.org/10.11947/j.agcs.2018.20170427
9. Zhang EH, Cao YC, Wang XY et al (2015) Characteristics of water vapor in a heavy rainstorm based on ground-based GPS measurements in Beijing. Meteorol Sci Technol 43 (06):1157–1163

A Original Observation Data Compression Method for Space-Based GNSS Receiver Based on Sparse Representation

Xiangyu Li(✉), Liheng Zhao, Zhang Cheng, Yong Ding, and Xianyang Liu

Space Star Technology Co., Ltd., Beijing, China
lixiangyul989@126.com

Abstract. The Space-based GNSS navigation receiver has the advantages of high precision, low cost and global coverage. It mainly provides information such as time reference, position and speed for spacecraft. With the release of GPS IGS products, it is based on the after-the-fact measurement of satellites. The development of precision post-processing technology is becoming more and more perfect. The Beidou system independently constructed in China has initially completed the verification of post-processing technology in the Asia-Pacific region. With the completion of the Beidou global navigation system in 2020, global coverage will be realized, and the application of precision post-processing technology will be implemented afterwards. The satellite orbit determination accuracy and load pointing accuracy are greatly improved, and a new implementation path for spaceborne gravity measurement is also provided. However, more and more satellites have proposed to be compatible with the Beidou system on the basis of the original GPS system, so as to realize the after-the-fact precision processing of the Beidou. In order to ensure the accuracy of the original observation measurement, the amount of data that needs to be generated and transmitted is often large. Star data storage, transmission and downlink channel capacity have caused a relatively large burden. Therefore, it is of great significance to provide a theory of raw measurement data compression for on-board GNSS receivers. The method of sparse representation and compressed sensing can effectively A set of primitive observations is represented by three parameters and a set of observation matrix methods, effectively reducing the satellite storage space and reducing the pressure of the downstream channel.

Keywords: Sparse representation · GNSS raw observation · Compression method

1 Introduction

In the research, scientists found that most of the information is expressed in a certain degree of redundancy. By using certain models and coding methods, this redundancy can be reduced. Claude Shannon of Bell Labs and RMFano of MIT presented the

This work was supported in part by National Key Research Grant 2016YFB0501302 and Grant 2016YFB0501304.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 52–62, 2019.
https://doi.org/10.1007/978-981-13-7751-8_6

earliest Shannon-Fano coding method for efficient coding of symbols to achieve data compression. DAHuffman first published his paper "Minimum Redundancy" in 1952. A Method for the Construction of Minimum Redundancy Codes. In the 1980s, mathematicians were not satisfied with some of the Achilles heel in Huffman coding. They started from a new perspective and designed another coding method that is more accurate and closer to the "entropy" limit in information theory-arithmetic coding. It can be proved that the compression effect obtained by arithmetic coding can minimize the redundancy of information, and accurately express the original information content with the least amount of symbols. Although the arithmetic coding can obtain the best compression effect, it may consume dozens of times. In the field of lossy compression algorithms, one of the most common methods is the discrete cosine transform algorithm (DCT) applied to JPEG image compression. Usually, the main energy of natural signals (including sound and image) is concentrated in the low frequency after discrete cosine transform. Partly, and when a person receives a voice signal or an image signal, there is a physiological process of downsampling, thus losing high-frequency part information without affecting the human intuitive feeling. However, for the on-board GNSS raw observation data, if too much high frequency information is lost, the ability of the signal to indicate high dynamic characteristics will be lost. Therefore, if the DCT algorithm is used to directly compress the original measurement data of the on-board GNSS navigation, the compression efficiency will be greatly reduced when discarding less DCT high-frequency atoms. When more DCT high-frequency atoms are discarded, the signal representation will be seriously lost. High dynamic performance capability, based on the shortcomings of the above compression algorithm and the characteristics of the original measurement data of the on-board GNSS navigation receiver, a raw measurement data compression algorithm based on sparse representation for spaceborne navigation receiver is proposed. The compression method still includes model and coding. In the two parts, and referring to the principle of DCT compression algorithm, the sparse representation idea is introduced. In the sparse domain, the dictionary model which characterizes the original measurement data of the spaceborne navigation receiver is obtained, and then the pre-compressed data is sparsely coded to complete the effective compression of the data.

The IGS data center is currently known to use the Hatanaka RINEX format to store all GPS observations. It is a "compressed" RINEX format that can compress 25–30%, which reduces network transmission review and storage space, but is limited to Ground compression, the compression algorithm is more complex and cannot be implemented in an onboard receiver.

2 Original Measurement Data Compression Algorithm Based on Sparse Representation

2.1 Discrete Cosine Transform (DCT) Coding

The discrete cosine transform (DCT for Discrete Cosine Transform) is a transform related to the Fourier transform. It belongs to a kind of orthogonal transform. It is similar to the discrete Fourier transform, but only retains the real part. The discrete

cosine transform can be used. The pre-transformed signal is transformed from the spatial domain to the frequency domain. After the transform, the energy of the signal will be mainly concentrated in the frequency portion corresponding to the larger code value, and the frequency information corresponding to the smaller code value is the non-principal component information of the signal, and the value is eliminated. The code value close to 0 achieves the purpose of data compression, and its mathematical expression is:

$$F(u) = c(u)\sum_{i=0}^{N-1} f(i)\cos\left[\frac{(i+0.5)\pi}{N}\mathrm{u}\right]$$
$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, u = 0 \\ \sqrt{\frac{2}{N}}, u \neq 0 \end{cases} \tag{1.1}$$

The original signal is the DCT transformed coefficient, which is the dimension of the original signal, and can be regarded as a compensation coefficient. The DCT transform matrix can be made into an orthogonal matrix, indicating the row number of the column signal. The encoding process is equivalent to solving the optimal function, where is a set of DCT basis functions (hereinafter collectively referred to as a dictionary), which is a set of cosine transform values at different frequencies. The DCT dictionary is:

$$\mathrm{D} = \left(\cos[\tfrac{(1+0.5)\pi}{N}u] \quad \cos[\tfrac{(2+0.5)\pi}{N}u] \quad \ldots\ldots \quad \cos[\tfrac{(N+0.5)\pi}{N}u]\right) \tag{1.2}$$

X is the linear representation coefficient of the original measurement signal in the basis function DCT dictionary.

2.2 Non-learning Sparse Representation Dictionary Construction Method

In the discrete cosine transform DCT dictionary, the dictionary atoms are perpendicular to each other. The set of basis functions in the dictionary is not specially constructed according to the characteristics of the original measurement data of the on-board GNSS navigation receiver, but a general transformation method. The DCT dictionary is for such signals. As a non-redundant dictionary, fewer atoms cannot accurately represent the raw measurement data of the on-board GNSS navigation receiver. Therefore, according to the unique attributes of the original measurement data of the on-board GNSS navigation receiver, a dedicated redundant dictionary set is searched for, and the original measurement data is effectively retained. Under the premise of high dynamic characteristics, it is especially important to implement sparse coding.

By analyzing the on-board GNSS navigation receiver measurement data (measurement information characterizing the motion situation), its principal components include distance, velocity, acceleration, jerk, etc., through which the custom dictionary $D_{N\times L} = [d_0 \quad d_1 \quad \ldots \quad d_L]$ is built.

$$d_i = \left[1 \quad t^i \quad \ldots\ldots \quad (L*t)^i\right]^T \tag{1.3}$$

2.3 Adaptive Sparse Representation Dictionary Learning Method

The purpose of signal processing is to better extract or reveal the information contained in the signal. In order to achieve a more flexible, concise and adaptive representation of the signal, Coifman and Hauser [2, 3] proposed the concept of sparse decomposition. The sparse decomposition theory holds that the more sparse the decomposition result is, the closer it is to the intrinsic or intrinsic structure of the signal. If the basis function chosen in the decomposition can make the decomposition result more sparse, it is considered that the basis function is better. If the basis function chosen to represent the signal is similar to the intrinsic structure of the signal, then the signal can be represented by only a few basis functions. The information in the signal is concentrated on these few basis functions, which is easier to extract and interpret. Sparse signal representation can substantially reduce the cost of signal processing and increase compression efficiency. Sparse representation of the mathematical model is:

$$\arg\min\|Y - DX\| + \lambda\|X\|_{0,1} \tag{1.4}$$

Where Y is the raw measurement data, D is the model, and X is the code. It can be known from the formula that the optimal sparse coding is obtained by using the L0 norm or the L1 norm penalty term, but the premise guarantees that the set of basis functions D is a redundant dictionary characterizing such pre-represented signals.

2.4 Sparse Coding

The coding formula follows the optimal formula for solving the following, and performs sparse coding.

$$\arg\min\|y_i - Dx_i\| + \lambda\|x_i\|_2^2 \tag{1.5}$$

Among them, $D = [d_0 \quad d_1 \quad \ldots \quad d_L]$ is one of the above three kinds of dictionaries, L is the sparsity degree sparsity, x_i is the sparse coding coefficient of the original measurement data y_i to be sparsely coded.

In the actual application of the orbit, by transmitting the sparse coded value, the ground station receives the sparse code value and then inversely encodes $\hat{y}_i$ it according to the known dictionary $D = [d_0 \quad d_1 \quad \ldots \quad d_L]$ and the sparsity L parameter s number.

3 Adaptive Sparse Representation Dictionary Learning Method

3.1 Dictionary Learning Method

The training dictionary method is a method that can adaptively generate a dictionary according to the characteristics of the target signal. The method adaptively trains the dictionary by inputting some signal samples, so that similar signal samples can realize sparse decomposition under the dictionary. In general, the number of input signal samples should be as large as possible, and contain enough information so that the trained dictionary can represent other unknown signals. Here, an over-complete dictionary training method based on sparse representation-K-SVD method [2] is used.

The K-SVD training dictionary method was first used to achieve image denoising and improve image resolution. The method combines singular value decomposition (SVD) and K-means clustering algorithm to adaptively train an overcomplete dictionary of sparse representations for training signals. When the signal required to be thinned has similar characteristics to the training signal in the dictionary. The dictionary can then sparsely encode the target signal. The method mainly uses the sparse constrained tracking and the singular value decomposition algorithm to alternately apply, so as to achieve adaptive adjustment of the signal.

The specific steps of K-SVD are as follows:

Initialization: random dictionary, $D(0) \in R_{n\times \mathrm{k}}$, iteration count statistics $J = 1$, maximum number of iterations *Iter*.

Step 1: Calculate the sparse coding using the matching pursuit algorithm $\mathrm{x}_i, i = 1, 2, \ldots, \mathrm{N}$

$$\underset{\mathbf{x}}{\operatorname{argmin}}\left\{ \| \mathbf{y}_i - \mathrm{Dx} \|_2^2 \right\} \text{Subject to } \|\mathbf{x}\|_0 \le T_0 \tag{1.6}$$

Step 2: Dictionary update:

For k = 1, 2…, K, determine the position of the kth atom used in the sample, i.e. the position of the kth row of non-zero elements in x

$$\mathbf{w}_k = \{i | 1 \le i \le N, x_i(k) \ne 0\} \tag{1.7}$$

Calculating residuals $E_k = Y - \sum_{j\ne k} d_j x^j$

Select E_k and d_k in the corresponding column to get E_k^R.

Use singular value decomposition $E_k^R = U\Sigma V^T$, make $d_k = u_1$, $x_k^R = \Sigma(1,1)v_1$, complete a loop.

Step 3: Let $J = J + 1$, judge whether it is satisfied $J > Iter$, if it is satisfied, stop iteration, otherwise, continue to perform the second step.

3.2 Sparse Coding

The matching pursuit algorithm [4] (OMP) is a greedy algorithm for finding this optimal optimal sparse decomposition locally. It is proposed by Tropp and Gilbert and successfully applied to signal recovery. The specific steps of OMP are as follows:

Input: redundant dictionary $D(0) \in R_{n\times \mathrm{k}}$, precoded signal $\mathbf{y}_i$, signal sparsity L.
Output: signal sparse representation coefficient estimation x_i
Initialization: residual $r_0 = y_i$, index set $\Lambda_0 = \phi, t = 1$

Cycle through steps 1 to 5:

Step 1: Find the atomic position λ most relevant to the residual r from D;
Step 2: Update the index set $\Lambda_t = \Lambda_{t-1} \cup \{\lambda_{\mathrm{t}}\}$ and record the set of reconstructed atoms in the found dictionary $D_t = [D_{t-1}, d_{\lambda_{\mathrm{t}}}]$;
Step 3: Get by least squares $\hat{\mathrm{x}}_{\mathrm{t}} = \arg\min_x \|y_i - D_t\hat{\mathrm{x}}\|_2$;
Step 4: Update the residual $r_t = y_i - D_t\hat{\mathrm{x}}_{\mathrm{t}}, t = t + 1$;
Step 5: Determine whether it is satisfied $t > L$. If it is satisfied, stop iteration. Otherwise, perform step 1.

3.3 Feasibility Analysis of Data Compression Algorithm

Using the original measurement data (L1CA code pseudorange, L1P pseudorange, L2P code pseudorange, L1 carrier, L2 carrier) transmitted by the onboard GNSS navigation receiver, using the DCT dictionary, the custom dictionary, and the proposed sparse representation dictionary pair This raw data (time series signal) is subjected to sparse coding compression, and the compression accuracy analysis [1] is compared. The result is shown in the Figs. 1, 2, 3, 4, 5 and 6 (Table 1):

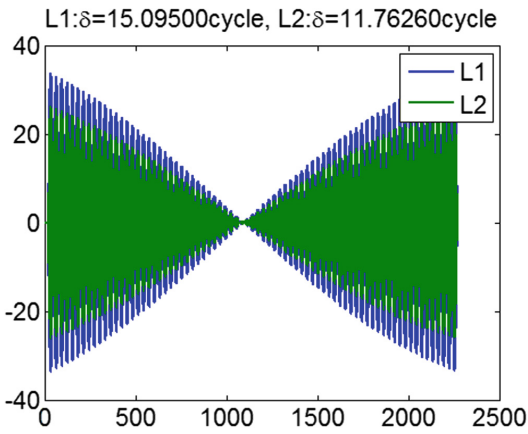


Fig. 1. DCT carrier compression accuracy analysis results

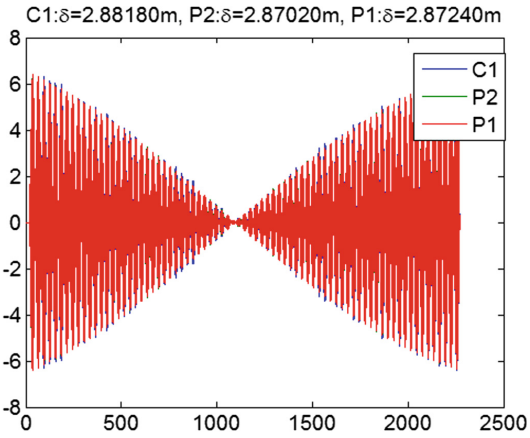


Fig. 2. DCT pseudorange compression accuracy analysis results

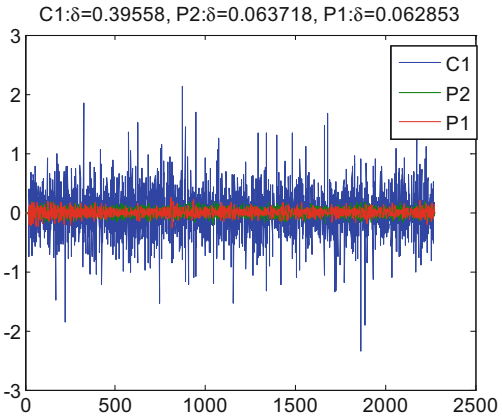


Fig. 3. Custom DCT pseudorange compression accuracy analysis results

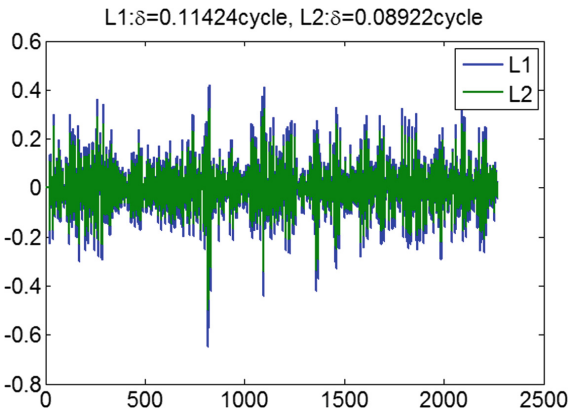


Fig. 4. CustomDCT carrier compression accuracy analysis results

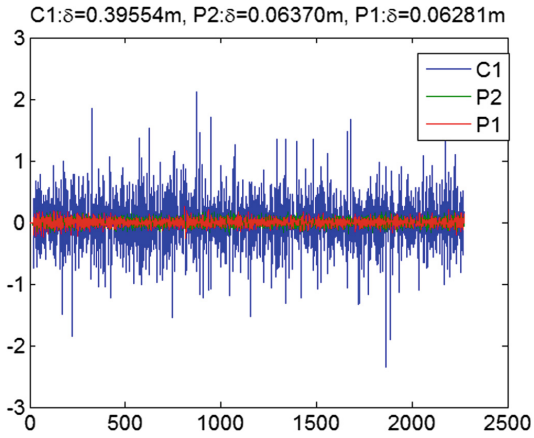


Fig. 5. Sparse representation pseudorange compression accuracy analysis results

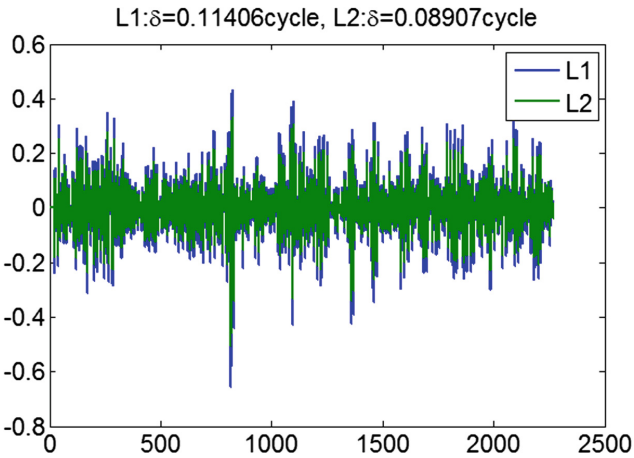


Fig. 6. Sparse representation carrier compression accuracy analysis results

By comparing and analyzing the above three methods, it can be seen that the sparse representation method can retain the original measurement quantity characteristics to the greatest extent and ensure the minimum process loss.

Table 1. Accuracy comparison results

Serial number	Compression algorithm	Items	Accuracy comparison results (m/cycle)
1.	DCT	L1Cpseudorange	2.88180
		L1 Ppseudorange	2.87240
		L2Ppseudorange	2.87020
		L1carrier	15.0950
		L2carrier	11.7626
2.	Custom DCT	L1Cpseudorange	0.39558
		L1 Ppseudorange	0.06285
		L2Ppseudorange	0.06372
		L1carrier	0.11424
		L2carrier	0.08922
3.	Sparse representation	L1Cpseudorange	0.39554
		L1 Ppseudorange	0.06281
		L2Ppseudorange	0.06370
		L1carrier	0.11406
		L2carrier	0.08907

4 Precision Orbit Verification

By analyzing the accuracy loss of the three compression coding methods, it can be seen that the DCT dictionary compression method can not meet the post-precision orbit determination requirements. The accuracy of the self-defined DCT method is slightly lower than that of the sparse representation compression method. The proposed sparse

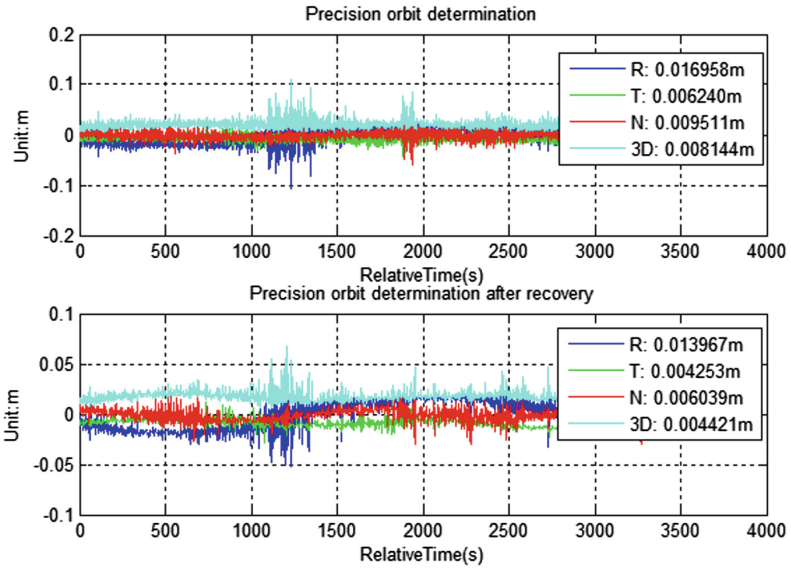


Fig. 7. Orbit determination results of Sparse representation 1

representation compression method uses two satellite transmits data under the orbit to carry out post-accurate orbit determination experiments. The orbital accuracy of the track overlap segment is shown in Figs. 7 and 8 (Table 2).

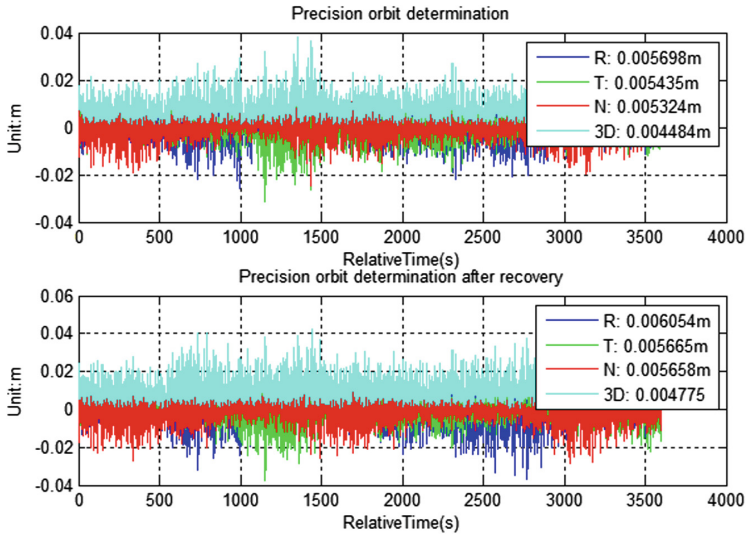


Fig. 8. Orbit determination results of Sparse representation 2

Table 2. Precise orbit determination accuracy comparison results

Data types	Data volume	State	RMS (m)
Satellite1	4000 s	Before compression	0.0081
		After compression	0.0044
Satellite2	4000 s	Before compression	0.0045
		After compression	0.0048

5 Summary

Sparse representation data compression algorithm is a new type of data processing algorithm, which has the advantages of easy implementation and high degree of reduction. As the number of fully-built GNSS receivers in the BD Global System will continue to increase, the ground will receive more and more GNSS raw observation data, which will provide important sample data for subsequent orbital measurements, auxiliary load device imaging, etc. In view of the fact that the real on-orbit situation may not be able to transmit a large amount of original observation data for post-processing due to limitations of the satellite link resources, this paper demonstrates a new type of data compression algorithm, namely on-board GNSS reception based on

sparse representation. In this paper, the original observation data of the 500 km LEO orbit satellite is transmitted to the simulation verification, and the compressed data is used for post-precision orbit determination. The results show that the data of the original observation data compression algorithm of the on-board GNSS receiver with sparse representation is highly accurate. It can meet the post-processing requirements of on-orbit raw observation data, and provide theoretical support and system verification for real-time data compression of subsequent satellite receivers.

References

1. Guo L, Li J, Liu R et al (2017) GNSS measurement receiver observation data quality assessment. In: China Satellite Navigation Academic Annual Conference
2. Aharon M, Elad M (2006) K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Sig Process 54(11):4311
3. Bruckstein AM, Donoho DL, Elad M (2009) From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Rev 51(1):34–81
4. Tropp JA, Gilbert AC (2007) Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory 53(12):4655–4666

Ground-Based GPS Used for Snowfall Weather Monitoring Research

Yajie Wang[1(✉)], Shuangcheng Zhang[1,2(✉)], Lifu Wang[3], and Chenglong Zhang[1]

[1] College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, Shanxi, China
l50132.986@qq.com, shuangcheng369@vip.163.com
[2] China Meteorological Administration Urumqi Institute of Desert Meteorology, Urumqi, Xinjiang, China
[3] Xinjiang Altai Meteorological Bureau, Altai, China

Abstract. With the continuous improvement of ground-based GPS meteorology and GPS remote sensing snow cover theory, combined with the increasing global GPS tracking stations, ground-based GPS is used as a new sensor in space to better predict and monitor snowfall weather. Based on the basic principles of ground-based GPS remote sensing water vapour and GPS remote sensing snow thickness theory, GPS observation data from the P360 site of the California PBO network from July 2012 to October 2017 in the United States were used to obtain the PWV combined the meteorological file through the GAMIT solution. And based on GPS-MR technology retrieve snowfall thickness. The analysis of the results shows that the ground-based GPS site can effectively obtain the continuous, real-time, high-time-resolution precipitable water PWV and snow depth around the site moreover obtain the relationship between snow depth and PWV and actual snowfall, which further expands the application of ground-based GPS in snowfall weather monitoring.

Keywords: GPS remote sensing technology · Precipitable water · GPS-MR · Snow depth · Weather monitoring

1 Introduction

Water vapor is one of the indispensable gas components in the atmosphere. It affects the radiation balance, energy transfer, cloud formation and precipitation, and is the most active factor affecting climate and weather change. It has a strong indication for the disaster weather on a small scale. Therefore, timely and accurate monitoring of water vapor content and its variation in the atmosphere have a very important realistic meaning. However, at present, the space-time distribution means of detecting water vapor, such as radiosonde technology, meteorological satellite, water vapor radiometer, radar detection technology and surface hygrometer observation, all have some problems, such as low temporal or spatial resolution and high cost [1]. With the development of ground-based GPS and GPS remote sensing meteorology, because of its advantages of real-time, all-weather, high precision, high resolution and no need for

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 63–71, 2019.
https://doi.org/10.1007/978-981-13-7751-8_7

calibration, it provides a new way to obtain atmospheric water vapor information with extremely high space-time variation in real time. The principle and method of retrieving water vapor by measuring GPS receiver was first put forward by Bevis of the United States [2]. On the basis of this, related studies were also carried out in China Li et al. [3] proved by GPS remote sensing water vapor experiment in Chengdu area that the accuracy of GPS/PWV can reach 1–2 mm. Fan and others studied the retrieve of ocean water vapor information and three-dimensional water vapor tomography and applied them to practice in the Bohai Sea region [4]. Biyan League [5], Yu [6], Cao [7] and other scholars all carried out the retrieval of water vapor spatial distribution algorithm by remote sensing tomography. The research has promoted the rapid development of GPS remote sensing meteorology in the direction of business application. Chen and Liu used the shipborne global satellite navigation system to detect ocean water vapor content in 2017 [8]. Analyzing the influence of different algorithms on the estimation of reducible water content were analyzed. Zhang and Zhao also studied the application of GPS remote sensing water vapor in haze weather monitoring. It is found that GPS water vapor and haze have correlation, and can be used to monitor and forecast haze weather to some extent [9].

Snow is one of the key components of hydrological system. To grasp the trend of snowfall and improve the monitoring ability of snow disaster can reduce the serious threat and heavy loss to agriculture, transportation and even electricity supply. Early monitoring of snowfall was mainly manual, echo detection of snow depth and radiation measurement were mainly ground measurements, because of the high cost, low temporal or spatial resolution greatly limit progress in this field. The NASA Langley Research Center and the University of Colorado jointly proposed the retrieval study that is to use the multipath effects of the SNR information in 1997, and then it is applied to snow depth monitoring [10, 11]. However, It has the strict requirements in the hardware [12], In 2009, the American scholar Larson proposed the GPS Multipath Reflection based on SNR for GPS at the first time [13]. In Europe and America, GPS-MR has a certain research foundation. However, the snow depth detection of GPS-MR is relatively late in China, and the related research results are relatively few. In 2015, Dai et al. described the theory and flow of GPS-MR in detail, and gave the relationship between multipath and the SNR. And proved the reliability of this technique [14]. Through comparing and analyzing the observational data of PBO (US Plate Boundary observation Program) in 2016, Jin et al. proved that the SNR data of L2P can be used to retrieval snow depth [15]. Zhang in 2017, Dai and others used GPS data from NWOT, Colorado, to retrieve the snowfall thickness. The measured data are in good agreement with each other, which verifies the feasibility of using GPS-MR technique to detect snow depth, and further improves the theory of GPS-MR retrieval of snow depth [16].

Based on the theory of GPS meteorology and remote sensing snow, this paper takes the P360 sites in PBO as an example to study the relationship between water vapor and snowfall, and to provide some thoughts for the extensive application of ground-based GPS.

2 Theoretical System of Ground GPS Remote Sensing Rain and Snow Monitoring

2.1 Monitoring Principle of GPS Water Vapor in Foundation

Ground GPS/MET is widely used in many fields, such as space geodesy, atmospheric spatial structure, weather forecast, global climate change monitoring and so on. It is spreading across the globe. After more than 20 years of research on ground-based GPS atmospheric water vapor, many national observatories have been put into production, and our country also develops actively in the field of GPS/MET. The related research work has been carried out, and it has been further promoted to develop towards the direction of specialization.

The current GPS remote sensing water vapor detection technology is also relatively mature. Based on the GAMIT calculation of the PWV value of the ground-based GPS remote sensing water vapor is calculated. It is one of the more commonly used methods at present. The basic flow of this method [17, 18] is

(1) Get ZTD
Based on the baseline solution of GPS observation data using GAMIT software to obtain the total zenith delay parameter ZTD

(2) Separate ZWD
Using the empirical model to calculate the ZHD, and separate the ZWD. The tropospheric empirical model can be summarized as follows:

$$ZTD \bullet M(E) = ZHD \bullet M_{ZHD}(E) + ZWD \bullet M_{ZWD}(E) \tag{1}$$

M_{ZHD}(E) and M_{ZWD}(E) are mapping functions of zenith dry and wet delay components, respectively. The common mapping functions are CFA2.2 function model, Chao function model, MTT function model, VMF1 model, NMF function model and GMF function model [19].

(3) Get PWV
PWV and ZWD have the following relationship:

$$PWV = \frac{10^6}{\rho_w R_v (K_2' + K_3/T_m)} \bullet ZWD = \Pi \bullet ZWD \tag{2}$$

K2′, K3, Rv are constants, Liquid water density, usually taken; It is called dimensionless proportional factor, which is used to express the proportional relation between precipitable water and zenith wet delay. Tm (K) is the atmospheric weighted mean temperature.

Taken together, the process for getting a PWV is as shown in the following figure (Fig. 1).

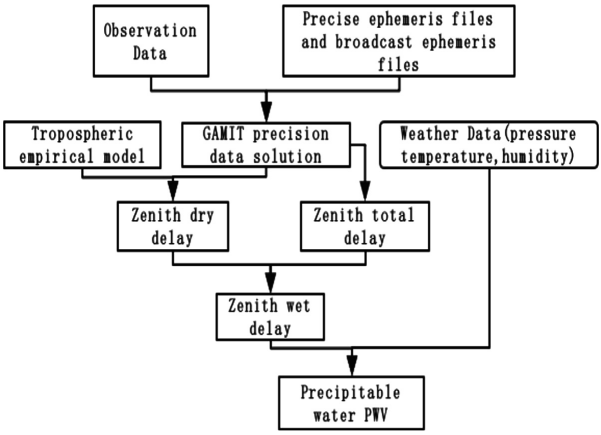


Fig. 1. Ground-based GPS acquisition PWV flowchart

2.2 The Basic Principle of GPS Remote Sensing for Snow Thickness

In the process of signal propagation, the effect of the reflectors near the station, the direction of propagation, the amplitude, the polarization and the phase will change. When the reflected signal arrives at the receiver and superposed with the signal coming directly from the satellite arriving at the receiver, the interference will occur between the two signals, resulting in a so-called "multipath error" when the observed value deviates from the true value. Reduce the accuracy of navigation positioning and baseline solution [20]. Even if the choke coil antenna of the GPS receiver antenna is suppressed, or the multipath effect is mitigated by improving the circuit design of the GNSS receiver, when the satellite altitude angle is low (5°–25°), the multipath error still exists.

In GPS measurement, SNR will increase with the increase of satellite altitude angle. From its value, you can intuitively see the impact of the multipath. According to the snow depth inversion theory put forward by Larson [21] as follow:

$$SNR \infty P_{\mathrm{d}} + P_r + \sqrt{P_{\mathrm{d}} P_r} \cos \varphi \tag{3}$$

Where P_d is direct signal energy, P_r is reflected signal energy, ϕ stands for the angle between the direct signal and the reflected signal.

$$SNR(Volts/volts) = 10^{\frac{SNR(db-Hz)}{20}} \tag{4}$$

The multipath effect information is separated from the SNR signal by linearing the Eq. (4) expression and using the low order polynomial to eliminate the trend term. After remove the trend term of the direct-emitting signal, the residual sequence amplitude of the multipath reflection signal can be expressed as:

$$dSNR = Acos(4\pi H\lambda H^{-1} sinE + \phi) \tag{5}$$

Where A is the amplitude, H is the reflection height, λ is the wavelength of the GPS carrier. (in this paper, L1), E is used as the satellite altitude angle. ϕ is angle between reflect signal and direct signal. If recorded $t = \sin E, f = 2H/\lambda$, The frequency f can be obtained by Lomb-Scargle spectrum analysis, and then can obtain the reflection height H.

The relationship between the reflection height and the depth of the retrieve snow cover can be expressed as the following:

$$H_{\text{snow}} = H_{\text{station}} - H \tag{6}$$

Where H_{snow} is the depth of snow, $H_{station}$ is the antenna height of the station, H is the reflection height obtained from the spectrum analysis.

The flow chart for snow depth research based on the GPS SNR see Fig. 2.

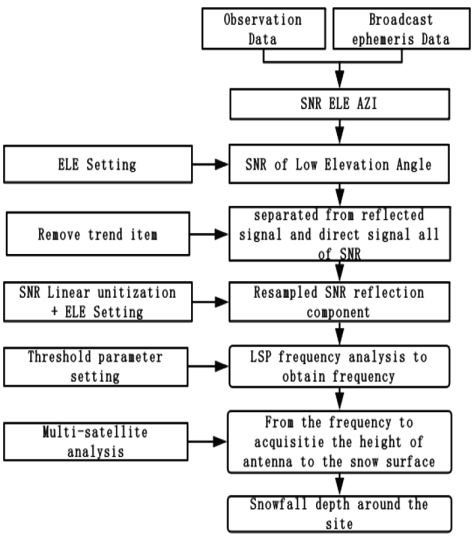


Fig. 2. Flow chart for snow depth detection based on GPS signal-to-noise ratio

3 Example Analysis

3.1 Data Sources

In this paper, we take the observation data of P360 from July 2012 to October 2017 in PBO Network, California, USA as an example to obtain the long time series values of PWV and snow depth. The P360 site (see Fig. 3), built in 2005, is located in the Island Park area of Montana, United States of America, at an average altitude of 1857.861 meters. With a snowfall of 1.4 m in winter and 40 mm rainfall in summer. The current receiver type is TRIMBLE NETRS. Antenna type TRIMBLE29659.00.

Fig. 3. Observational environment in winter and summer at P360 site

Based on the theoretical basis of ground GPS for monitoring rain and snow, the PWV two hours value of long time series and the snow thickness daily value of long time series can be obtained. In addition, Actual snowfall data around the P360 site was provided by the USDA Natural Resources Conservation Service.

3.2 Results

After data processing, the PWV values and snow depth values at P360 site were obtained. The actual snow depth value is compared with the PWV value and snow depth value obtained by GPS remote sensing theory (see Figs. 4 and 5).

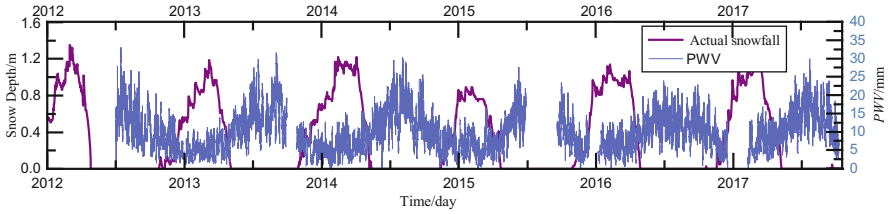


Fig. 4. Comparison of PWV obtained from ground-based GPS and actual snowfall

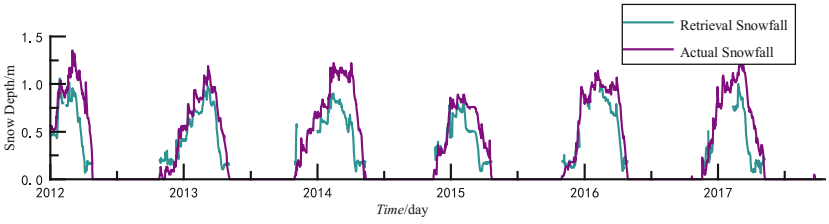


Fig. 5. Ground-based GPS compares snow depth with actual snow depth

In Fig. 4, the PWV and actual snowfall depth peaks and troughs show a reverse correlation. Snow weather mainly occurs in the seasons when the water vapor content is small. The formation process of snow has a certain negative correlation with the change of PWV, that is, a sharp decrease in PWV creates favorable conditions for the snowfall weather.

In Fig. 5, the trend is the same in the same time range. The difference between the result of GPS-MR and the actual snow depth is smaller and the accuracy is higher.

From the analysis of Figs. 4 and 5, we can see that in the long time series:

1. The measured snow depth is in good agreement with the inversion result and the trend is consistent.
2. There is a nonlinear negative correlation between the PWV value and the snowfall depth value. The PWV value is correlated with the trough and peak of snow thickness. This may be due to the fact that water vapor in the air is not falling to the ground as expected.

In order to find out a more detailed correlation between the PWV and the actual snow depth values, we analyze the short time series.

As shown in Fig. 6, the range of analysis is reduced to a short time range before and after snowfall.

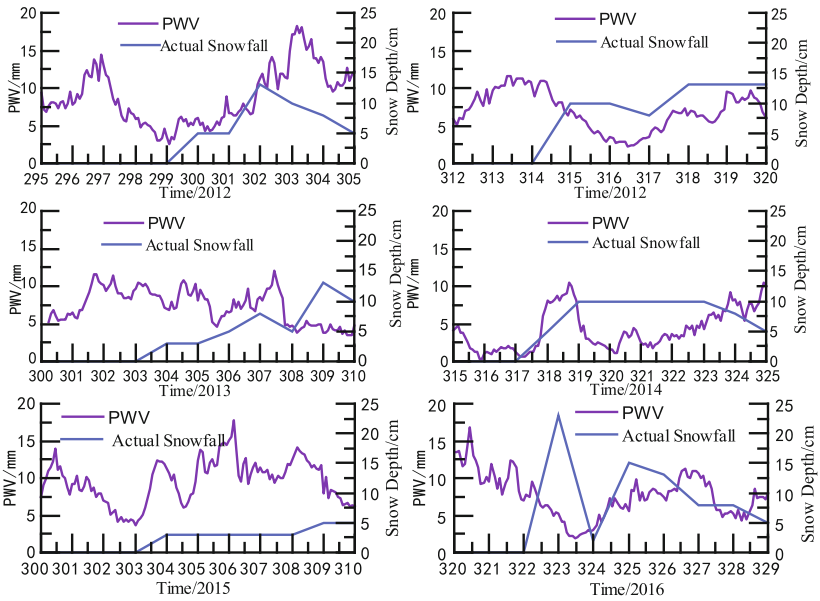


Fig. 6. Comparison of PWV and actual snowfall of 2015–2016

In this short time series, PWV will have a small peak, a reserve of water vapor required for snowfall. When the temperature is reached, PWV will become solid snow.

The snowfall will only occur in the following three days, and when the snowfall begins in a short time, the snowfall gradually accumulates, and the PWV value also presents a rising state, but the latter two have no obvious relationship, which may be related to the temperature or atmospheric circulation at that time. From Fig. 6 the analysis shows that:

From the point of view of the short time series, before snowfall is studied separately, and the relationship between the change of PWV value and the time of snowfall indicates that the snowfall has a certain lag.

Through the above analysis, it can be concluded that the continuous operation of ground GPS and the theory of snow remote sensing in the region can obtain high precision real-time PWV and snow depth value, and the appearance of snow fall is closely related to the PWV.

4 Conclusion

The ground-based GPS remote sensing can effectively utilize the resources and data of the existing continuous operation of GNSS observation station, without other hardware costs, and can be constructed only through the software deployment of the data center, which is cheap and easy to realize. With the support of this theoretical foundation, the ground-based GPS tracking site can obtain continuous, real-time, high-resolution precipitable PWV and snow thickness data around the sites, and objectively reflect the variation of the two data. It detects of regional disasters and becomes a powerful supplement to traditional meteorological analysis and research. Through the results, the feasibility of obtaining high accuracy snow cover thickness by GPS and the reliability of GPS remote sensing theory are verified. And It is shown that PWV and snow thickness have certain correlation. Applying the relationship between the two to the research and analysis of snowfall weather, can further expanding the ground GPS falling. The application of snow weather monitoring improves the practical utilization value of foundation GPS.

In this paper, we preliminary know that, GPS water vapor may contribute to the monitoring of snowfall weather to some extent. and then the more accurate and reliable data processing results are used for in-depth analysis to obtain the evolution characteristics of PWV and snowfall depth values in different regions which will be further improve the accuracy of snowfall forecast. In order to achieve the purpose of real-time application of PWV, we need to localize the model and parameters in data processing. At the same time, we need to further explore how to use the foundation GPS-MR technology to obtain more accurate snow depth detection data, and add this technique to the ground-based GPS remote sensing water vapor processing system. So that better combination of the two to improve the monitoring of snow weather.

Acknowledgments. We gratefully acknowledge the provision of data, equipment, and engineering services by the Plate Boundary Observatory operated by UNAVCO for Earth Scope. This work was supported by China Desert Meteorological Science Research Foundation (Sqj2017002), National Science Foundation of China (41731066, 41674001, 41104019) and the Special Fund for Basic Scientific Research of Central Colleges (310826172202).

References

1. Jiang G (2011) Study of the Theory and Application to Real time Remote Sensing Using Ground-based GPS. Chang'an University, Xi'an
2. Bevis M, Businger S, Herring TA et al (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. Chin. J. Geophsics 97 (D14):15787–15801
3. Li G (2007) On the remote Sensing PWV Using Ground-based GPS Technique and Application in Meteorology. Southwest Jiao tong University, Si'chuan
4. Fan S (2013) Research on GPS Marine Water Vapor Inversion and Three Dimensional Water Vapor Tomography. Wuhan University, Wuhan
5. Bi Y, Yang G, Nie J (2010) GPS water vapor tomography method based on Kalman filter and its application. Plateau Meteorol. 30(1):109–114
6. Yu S, Liu L, Liang X (2010) Analysis of the influence of constraints on GPS vapor tomography. J. Surv. Mapp. 39(5):491–496
7. Cao Y (2012) Ground-based GPS tomography three-dimensional vapor and its application in meteorology. Chinese Academy of Meteorological Sciences, Beijing
8. Chen G, Liu Y, Liu X (2017) Analysis of influence factors of marine water vapor information detection by ship-borne GNSS. J. Wuhan Univ. (Inf. Sci. Ed.) 42(2):270–276
9. Zhang S, Zhao L, Lü X (2018) Application of GPS water vapor in haze weather monitoring. J Wuhan Univ (Inf Sci Ed) 43(03):451–456
10. Wan W, Chen X et al GPS-R remote sensing research progress at home and abroad. Remote Sens Inf 27(3):112–117
11. Liu J, Shao L, Zhang X (2007) GPS-R research progress and key technologies. J Wuhan Univ (Inf Sci Ed) 32(11):955–960
12. Zhang S, Dai K, Liu K (2016) GPS-MR technology for snowfall thickness monitoring research. Prog Geophys 31(4):1879–1884
13. Larson KM, Gutmann ED, Zavorotny VU, Braun JJ, Williams MW, Nievinski FG (2009) Can we measure snow depth with GPS receivers. Geophys Res Lett 17(1):1–5
14. Dai KY, Zhang SC, Zhang Q et al (2016) Preliminary research on snow depth monitoring with GPS SNR. In: China Satellite Navigation Conference (CSNC) 2016 Proceedings: Volume II. Springer, Heidelberg, p 71–84
15. Jin S, Qian X, Kutoglu H (2016) Snow depth variations estimated from GPS-R: a case study in Alaska from L2P SNR data. Remote Sens 8(1):63
16. Zhang S, Dai K (2018) A preliminary study of snow depth detection using GNSS-MR technology. J Wuhan Univ (Inf Sci Ed) 43(2):234–240
17. Zhang S (2009) Research and Application of Remote Sensing Water Vapor Using Ground_Based GPS/Met. Wuhan University, Wuhan
18. Zhang Q, Li J et al (2005) GPS Measurement Principle & Application. China Science Publishing & Media Ltd, Beijing
19. Zhang S, Ye S, Liu J et al (2009) The latest development of dynamic mapping functions such as and its application in GNSS remote sensing water vapor. J Wuhan Univ (Inf Sci Ed) 34(3):280–328
20. Wang X (2000) The influence of multipath errors in GPS measurement. Gradient Geo-Dyn 20(1):56–59
21. Larson KM, Nievinski FG (2013) GPS snow sensing: results from the earth scope plate boundary observatory. GPS Solute 17:41–52

Pseudo-Range Single Point and Differential Positioning Accuracy Test Based on Android Smartphone

Yongsheng Liu(✉), Chengfa Gao(✉), Bo Chen, and Ruicheng Zhang

School of Transportation, Southeast University, Nanjing 211189, China
liuyongsheng_nj@163.com, gaochfa@163.com

Abstract. In the era of the Internet of Everything, providing high-precision location services for the public is one of the key directions for the future development of surveying and mapping science. Under this background, Google opened the APIs of Android intelligent terminal to obtain GNSS raw data in 2016. Based on this, this paper mainly explains how to obtain the GNSS raw data of Android smartphone, and uses the Huawei P10 and Xiaomi 8 smartphones installed with self-developed real-time positioning APP to positioning test and analysis accuracy of single-frequency pseudo-range single-point, dual-frequency pseudo-range single point and pseudo-range differential positioning. The problem of inconsistent pseudo-range and carrier phase data in the GNSS raw data of the smartphone is considered to be caused by different clocks of the smartphone; Although there is an inconsistency problem, the positioning accuracy of the smartphone can be improved by the carrier phase smoothing method; In the case of no smoothing, the differential positioning effect is limited due to excessive noise of the pseudo-range; When the carrier phase is smoothed 10–30 times, the pseudo-distance difference method can be used to improve the plane accuracy of Huawei P10 by 10–30%, the plane accuracy of Xiaomi 8 can be increased by 20–40%, and the elevation direction can be increased by 60–70%. Finally, when smoothing 30 epochs for differential positioning, the RMSE in P10 plane positioning is ±3.9 m, the elevation is ±4.3 m, the RMSE in the plane positioning of Xiaomi 8 is ±1.1 m, and the elevation is ±1.4 m. Although the dual-frequency de-ionospheric combination of Xiaomi 8 can effectively eliminate the influence of the ionosphere, it can't effectively improve the positioning accuracy because it can amplify the noise by 2.59 times, which needs further analysis.

Keywords: GNSS · Android smartphone · Pseudo-range single point positioning · Pseudo-range differential positioning

1 Introduction

The low-cost GNSS antennas and chips are widely used in smart phones to provide low-precision position and navigation services for the masses, but their positioning output is only a three-dimensional position result, and it is impossible to obtain the original pseudo-range and carrier phase observation data of the positioning chip. The

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 72–81, 2019.
https://doi.org/10.1007/978-981-13-7751-8_8

positioning accuracy is about 10 m. With the development of science and technology, users need GNSS to provide users with higher-precision navigation and positioning services in real time to meet the needs of the public for location services. At the Google I/O conference in May 2016, Google announced that developers can access raw GNSS measurements using smartphones and tablets running the Android N (version: 7.0) operating system. This allows us to get rid of the black-box concept of GNSS receivers, and pseudo-range, Doppler, carrier phase data can be used to calculate more accurate positions, which is of great significance for the research of precision positioning based on intelligent terminals [1].

This paper mainly introduces how Android smart phones acquire GNSS raw data, and then uses Huawei P10 and Xiaomi 8 smart phones installed with self-developed real-time positioning APP to test and analyse the accuracy of single-frequency pseudo-range single point, dual-frequency pseudo-range single point and pseudo-range differential positioning.

2 Obtaining Smartphone GNSS Raw Data

According to the statistics of Android developers' official website, as of 2018.10.26, the market share of Android N and its versions is about 49.7%, and different types of smartphones have different access to GNSS raw data. For details, see the official website [2]. The interface for Android terminal development to obtain GNSS raw data is mainly included in the two major categories of GnssMeasurement and GnssClock, as shown in Tables 1 and 2. It can be known from the two categories of GnssMeasurement and GnssClock that the pseudo-range cannot be directly obtained, and it needs to be obtained by calculating the corresponding time, namely:

$$\rho = (t_{Rx} - t_{Tx}) \cdot c \tag{1.1}$$

Where: ρ denotes a pseudo-range, t_{Rx} denotes the time at which the receiver receives the signal, and t_{Tx} (*ReceivedSvTimeNanos*) indicates the moment when the satellite transmits the signal. t_{Rx} cannot be obtained directly, and needs to be obtained through calculation: Note that the time system of different GNSS systems is different, and the method of obtaining t_{Rx} is also different.

$$t_{Rx} = t_{Rx_{GNSS}} - GNSS_{week} \tag{1.2}$$

$$t_{Rx_{GNSS}} = TimeNanos + TimeOffsetNanos - (FullBiasNanos + BiasNanos) \tag{1.3}$$

In the formula, *TimeNanos*, *TimeOffsetNanos*, *FullBiasNanos* and *BiasNanos* are specified in Tables 1 and 2, and $GNSS_{week}$ is the time of each system from 0:00 on January 6, 1980 to the current GPS week. By calculation, t_{Rx} can be obtained, and then pseudo-range data [3] can be obtained.

Table 1. GnssMeasurement class of major GNSS raw data

Raw data	Remark
AccumulatedDeltaRangeMeters	AccumulatedDeltaRangeMeters = −k * CarrierPhase, k = c/f, where c is speed of light, f is the carrier frequency
AccumulatedDeltaRangeState	0: The state is invalid or unknown; 1: The state is valid; 2: A reset has been detected; 4: A cycle slip has been detected
CarrierFrequencyHz	
Cn0DbHz	
ConstellationType	1: GPS; 2: SBAS; 3: GLONASS; 4: QZSS; 5: BDS; 6: Galileo; 9: unknown
MultipathIndicator	0: Unknown state; 1: Multipath; 2:No multipath
PseudorangeRateMetersPerSecond	PseudorangeRateMetersPerSecond = −k * DopplerShift, k = c/f
ReceivedSvTimeNanos	That is, starting from the current GPS week, the value range [0,604800] * 10^9
State	Cumulative sum of per-satellite sync state
Svid	
TimeOffsetNanos	

Table 2. GnssClock class of major GNSS raw data

Raw data	Remark
BiasNanos	
FullBiasNanos	That is, the difference between GPS at 0:00 on January 6, 1980 and the local clock of the mobile phone
HardwareClockDiscontinuityCount	
LeapSecond	
TimeNanos	

3 Obtaining Smartphone GNSS Raw Data

The mobile APP is developed in the Android Studio development environment using the Java language. Its main functions include: 1. Get the original data of smart phone by calling GnssMeasurement and GnssClock; 2. Download the latest ultra-fast precision ephemeris (1 h update once) and broadcast ephemeris [4] from IGS data center of Wuhan University; 3. Real-time decoding of RTCM3 format data (mainly decoding 1074, 1084, 1094, 1124 type data in MSM4) [5]. The APP is used for single point and differential positioning and standardizes the positioning results into text format. The main interface of the APP is shown in Fig. 1.

Fig. 1. The main interface of the self-developed APP

Four suitable control points GE01, GE03, GE04, GJ02 are selected in the Jiulonghu Campus of Southeast University of Nanjing. The coordinates of each point are known. The plane error is about 2 cm, and the elevation error is about 5 cm. The m-level positioning result of the smartphone can be regarded as the true value. The Huawei P10 and Xiaomi 8 smartphones installed with the self-developed APP were placed on the control point for static observation. The sampling rate is 1 s. The observation time of each point is about 10–15 min, and the cut-off carrier-to-noise ratio is 10 dB.

4 Raw Data Analysis

Huawei P10 uses a single-frequency GNSS chip. Under good conditions, it can receive about 10–20 satellites signals. Xiaomi 8 uses Broadcom BCM47755 dual-band GNSS chip to support GPS (L1 and L5) and Galileo (E1 and E5a). Dual-frequency data can receive about 15-25 satellites signals, of which dual-frequency data satellites account for about 2/5 of the total. Table 3 is obtained by analyzing the carrier-to-noise ratios of several satellites with long synchronization observation time.

Table 3. Mean C/N of part satellites (unit: dB)

Satellite	G12	G20	R04	C09	G32	G25	E07
Huawei P10	28.0	22.4	28.9	23.0	31.1	26.3	29.1
Xiaomi 8	37.1	33.8	31.7	27.5	34.8 26.5 (L5)	31.7 29.2 (L5)	28.5 26.4 (E5a)

The C/N of Xiaomi 8 is about 19% higher than that of Huawei P10. Only some satellites are slightly worse. At the same time, the L1 (E1) band carrier-to-noise ratio of Xiaomi 8 is about 16% higher than that of L5 (E5a).

In the process of processing the original data, it is found that the pseudo-range change rate and the carrier phase change rate in the original GNSS observation data are inconsistent, and the system deviation occurs, and the carrier phase change rate is consistent with the Doppler data. In theory, the three should be consistent. All the satellites of Huawei P10 and Xiaomi 8 show this phenomenon. The difference between adjacent epochs can obtain the rate of change of pseudo-range and carrier phase (reduction to distance), taking G12 and other satellites as an example. The details are shown in Fig. 2 and Table 4 (A).

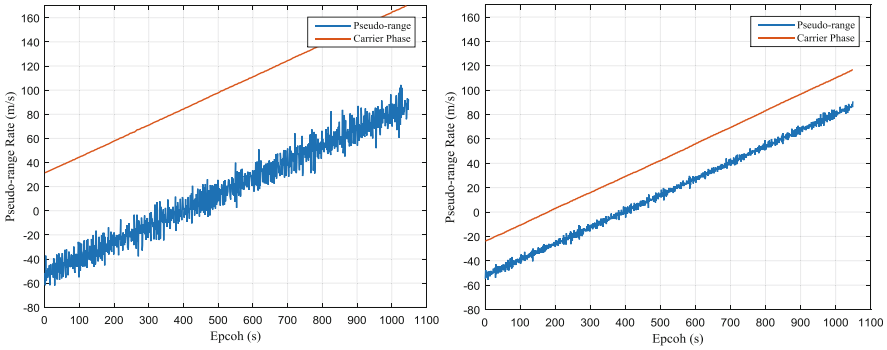


Fig. 2. Pseudo-range rate of Huawei P10 (Left) and Xiaomi 8 (Right) of G12 satellite

In response to this phenomenon, the "White Paper on using GNSS Raw Measurements on Android devices" [3] released by the European GNSS organization pointed out that some smart phones use different power consumption clocks to use together to reduce power consumption. Therefore, the above phenomenon can be interpreted as a system deviation caused by the use of different clocks for the pseudo-range and the carrier phase.

Regression analysis is carried out on the pseudo-range change rate and carrier phase change rate of the above satellites, and the standard deviation is obtained to evaluate the quality (see Table 4: B_1 and B_2). The following conclusions can be drawn: 1. smartphone pseudo-range quality: Galileo > GPS = BDS > GLONASS; 2. smartphone carrier phase quality: GPS = BDS > Galileo = GLONASS; 3. The pseudo-range quality of Xiaomi 8 is much higher than that of Huawei P10, which is 2–4 times; 4. GPS data quality differences of Huawei P10 is relatively large.

Table 4. Analysis table of smartphone GNSS raw data

Smartphone	Time (s)	Satellite	A	B_1	B_2
Huawei P10	1049	G12	83.75	7.389	0.146
	1049	G20	83.80	8.263	0.532
	1049	G32	83.79	10.711	0.169
	1049	R04	83.78	12.250	0.277
	948	R15	83.57	15.735	0.593
	1047	C09	83.74	5.246	0.299
	1049	C25	83.78	3.847	0.386
	983	E07	83.75	3.650	0.407
	849	E08	83.91	3.940	0.580
Xiaomi 8	1049	G12	28.81	1.998	0.295
	1049	G20	28.81	2.828	0.374
	1049	G32	28.81	2.706	0.367
	1041	R04	28.81	4.983	0.477
	1049	R15	28.82	7.789	0.448
	1047	C03	28.82	2.902	0.252
	1049	C09	28.81	2.061	0.281
	1049	E07	28.81	1.549	0.452
	1049	E08	28.81	1.227	0.358

Note: Xiaomi 8's dual-frequency satellite data uses L1 (E1) band data; A: Average of the difference between the carrier phase change rate and the pseudo-range change rate (m/s); B_1: Pseudo-range change rate regression analysis standard deviation (m/s); B_2: Carrier phase change rate regression analysis standard deviation (m/s).

5 Positioning Test Result

5.1 Positioning Model

According to the experimental scheme, single-frequency pseudo-range single point positioning, dual-frequency pseudo-range single point positioning and pseudo-range differential positioning [6] experiments and analysis accuracy, single-frequency pseudo-range single point positioning results as a benchmark. The positioning accuracy was characterized by RMSE of the deviation between the positioning result and the truth value (see Sect. 3).

The pseudo-range single-point positioning must correct the satellite's various errors: the satellite ephemeris and the satellite clock difference can use the ephemeris and clock difference data in the ultra-fast precision ephemeris; the weighting method uses the carrier-to-noise ratio weighting model [7] The tropospheric delay correction uses the UNB3 M tropospheric correction model; the relativistic influence uses the classical relativistic correction model; the Earth rotation effect uses the Sagnac effect correction

model, the single-frequency ionospheric delay correction uses the 8-parameter ionospheric correction model, and the dual-frequency ionospheric delay correction uses the dual-frequency de-ionospheric combination model.

The differential positioning base station is about 3 kilometres away from the observation point. It is a Trimble NetR9 GNSS receiver, which broadcasts RTCM3 format observation information (GPS, GLONASS and BDS in real time. The table data below are calculated by the three systems, except for others). The pseudo-range data of the base station is smoothed by 3 epochs of carrier phase to reduce noise.

5.2 Single Frequency Differential Positioning

Because of the inconsistency between carrier phase and pseudo-range of smart phones, to use carrier phase to smooth pseudo-range, it is necessary to ensure that the smoothing times of each satellite are the same, so as to ensure that the deviation caused by inconsistency is absorbed by the receiver clock difference in the positioning process, so as to improve the positioning accuracy. Huawei P10 can increase its elevation by 48% and 33% by smoothing five epoch planes, as shown in Table 5.

Table 5. Positioning results of Huawei P10 at GE01 (unit: m)

Smoothing epoch	1 (C)			5 (C)			10 (D)			20 (D)		
RMSE	N	E	U	N	E	U	N	E	U	N	E	U
Single point	6.338	3.595	9.285	4.345	2.201	4.825	3.129	1.698	4.216	2.680	1.348	3.329
Difference	5.453	3.428	9.324	2.898	1.940	4.889	2.384	1.486	4.262	1.811	1.053	3.390
Promotion	13.96%	4.64%	−0.42%	33.31%	11.85%	−1.32%	23.82%	12.46%	−1.09%	32.41%	21.85%	−1.83%

Note: The above C and D represents different time periods.

Table 6. Positioning results of smooth 30 epochs of Huawei P10 at GE03, GE04, and GJ02 (unit: m)

Control point	GE03			GE04			GJ02		
RMSE	N	E	U	N	E	U	N	E	U
Single point	3.883	1.864	4.322	2.156	2.504	6.033	3.575	5.797	4.989
Difference	2.307	1.900	3.422	1.869	1.810	3.096	2.216	5.652	6.503
Promotion	40.59%	−1.95%	20.83%	13.34%	27.71%	48.68%	38.02%	2.50%	−30.34%

When Huawei P10 does not carry out carrier phase smoothing, the difference can improve the positioning accuracy, but because of its large pseudo-range noise, the improvement effect is not obvious. However, by smoothing pseudo-range through carrier phase, the noise is greatly reduced, and the original hidden system error will be exposed through the difference. These errors can be corrected effectively, so that the positioning accuracy can be improved. After smoothing 20–30 epochs, the plane positioning accuracy can be improved by 10–30% using difference, but the elevation direction fluctuates greatly, as shown in Tables 5 and 6.

Since the pseudo-range quality of Xiaomi 8 itself is high, the carrier phase smoothing is not performed, the plane positioning accuracy has reached 3.5 m, and the

elevation is 5.7 m. After the differential correction, the plane positioning accuracy is 3.0 m and the elevation is 3.8 m, see Table 7. After smoothing 10–30 epochs, the difference can be used to improve the plane positioning accuracy by 20–40% and the elevation direction by 60–70%. The effect is significant, and the difference is smooth after 30 epochs. The plane positioning accuracy can reach 1.1 m. The elevation is 1.4 m, as shown in Table 8, which provides a solution for realizing sub-meter positioning of intelligent mobile terminals.

5.3 Dual Frequency Single Positioning

Xiaomi 8 has dual-frequency data of GPS and Galileo. When using four-system positioning, the dual-frequency deionization combined model is used for GNSS satellite data with dual frequency [8], see the formula below.

$$\rho_{1,5} = \frac{f_1^2}{f_1^2 - f_5^2}\rho_1 - \frac{f_5^2}{f_1^2 - f_5^2}\rho_5 \tag{1.4}$$

Among them, $\rho_{1,2}$ is the corrected pseudo-range observation value, ρ_1 and ρ_5 are the dual frequency pseudo-range observation values, and f_1 and f_5 are the frequencies of the carrier L1 (E1) and L5 (E5a), which are 1575.42 MHz and 1176.45 MHz respectively.

In order to highlight the dual-frequency effect, its weight is expanded to twice the original, but the final positioning result is much worse than that of single-frequency positioning. The combination of L1 (E1) and L5 (E5a) bands enlarges the pseudo-range noise to $\sqrt{1.2606^2 + 2.2606^2} = 2.59$ times of the original. That means the influence of pseudo-range noise is much greater than that of system errors such as the ionosphere, which makes the effect of using the dual-frequency de-ionospheric model less effective, see Table 7.

Table 7. Positioning results of no smoothing of Xiaomi 8 at GE01 (unit: m)

GNSS systems	Three systems(GPS, GLONASS and BDS)				Four systems (GPS, GLONASS, BDS and Galileo)		
RMSE	N	E	U		N	E	U
Single-frequency single point	2.736	2.236	5.655	Single-frequency single point	2.311	2.322	5.556
Difference	2.575	1.582	3.767	Dual-frequency single point	4.550	3.789	12.841
Promotion	5.90%	29.24%	33.38%	Promotion	——	——	——

Table 8. Positioning results of Xiaomi 8 at GE01 (unit: m)

Smoothing epoch	10			30		
RMSE	N	E	U	N	E	U
Single point	1.194	1.408	5.877	1.068	1.420	5.925
Difference	1.104	0.786	1.908	0.881	0.651	1.417
Promotion	7.60%	44.23%	67.53%	17.52%	54.13%	76.09%

6 Conclusion and Prospect

1. Although there is a problem of inconsistency between pseudo-range and carrier phase in smart phones, the positioning accuracy can be improved by carrier phase smoothing when the smoothing times of each GNSS reference satellite are the same. After smoothing 5 epochs, the plane accuracy and elevation direction of Huawei P10 can be improved by 33% and 48%. In the future, the causes of inconsistencies can be further analyzed and considered to establish models for correction.
2. In the case of no smoothing, the differential lifting positioning accuracy is limited due to the excessively large pseudo-range noise. However, when the carrier phase is smoothed 10-30 times, the pseudo-range difference method is adopted, and the Huawei P10 plane accuracy can be improved by 10–30%. The plane precision of Xiaomi 8 can be increased by 20–40%, and the elevation direction can be increased by 60–70%. The effect is remarkable. Finally, after smoothing 30 epochs, the P10 plane positioning accuracy is 3.9 m, the elevation is 4.3 m, the Xiaomi 8 plane positioning accuracy is 1.1 m, and the elevation is 1.4 m. This paper does not analyse and use Doppler data. In the future, different filtering methods can be tried for dynamic and static positioning analysis.
3. Although the dual-frequency de-ionospheric combination can effectively eliminate the influence of the ionosphere, the noise amplification of Xiaomi 8 is 2.59 times that of the original, which leads to the inability to effectively improve the positioning accuracy. This problem needs further analysis. Because of space, this paper does not make a further analysis of the distribution characteristics of pseudo-range noise, which can be further studied in the future.

References

1. Simon B, Frank VD (2016) Precise positioning using raw GPS measurements from Android smartphones. GPS World 27(11):43–48
2. Google, Inc. Raw GNSS Measurements [EB/OL]. https://developer.android.google.cn/guide/topics/sensors/gnss. Accessed 20 Aug 2018
3. European GNSS Agency. White Paper on using GNSS Raw Measurements on Android devices (2017)
4. IGS data center of Wuhan university. ftp://igs.gnsswhu.cn. Accessed 10 Sept 2018
5. RTCM STANDARD 10403.3, Differential GNSS (Global Navigation Satellite Systems) Services–Version 3 (2016)

6. Gao C, Hu W (2011) The Principle and Application of Satellite Navigation and Positioning. China Communications Press, Beijing
7. Liu C, Li F (2018) Comparison and analysis of different GNSS weighting methods. Sci Surv Mapp 43(08):39–44
8. Xie G (2017) Principles of GPS and Receiver Design. Publishing House of Electronics Industry, Beijing

Detecting Slow-Slip Events in GNSS Position Time Series Using Relative Strength Index

Jiugang Xie(✉), Cuilin Kuang, Hongke Hou, and Xiaotao Bai

School of Geosciences and Info-Physics,
Central South University, Changsha, China
xjgcsu@163.com

Abstract. Slow slip (slow earthquake, silent earthquake) is an important way to release the pressure caused by the movement of the tectonic plate, and the other way is the earthquake in the conventional sense. Compared with conventional earthquakes, slow slip occurs deeper, lasts longer, and recurs periodically, but the stress intensity released is much weaker, and even does not generate seismic waves which is difficult to be detected by seismographs. The subduction zone where slow slip happens may breed super large earthquakes of magnitude 8 or higher. Studying slow slip can not only deepen the understanding of the movement of plate boundaries, but also increase the ability to predict earthquakes. GNSS technology has been widely used in monitoring plate motion because its high precision and high time resolution, we can detect slow-slip events from GNSS position time series. In this paper, the single-site transient signal detection method based on relative strength index (RSI) proposed by Brendan is used to detect the slow-slip events of the GNSS sites on the HIKURANGI subduction zone in New Zealand. The results show that the method can detect most slow-slip events, but exist missed and wrong detections within acceptable limits, and some GNSS sites have detection anomalies. This paper explored the reasons why the method detected anomalies at some sites, and got some beneficial conclusions, and gave some corresponding strategies which can strengthen the applicability and capability of the method.

Keywords: GNSS position time series · Transient signal detection · Relative strength index · Slow slip

1 Introduction

GNSS technology has become one of the main means of monitoring crustal deformation. Lots of GNSS sites around the world provide reliable data for studying plate motion, fault slip, strong earthquake deformation field, etc. How to distinguish and extract GNSS transient signals effectively has become an international frontier research [1]. Signals in the GNSS position time series include tectonic signals (deformations caused by earthquakes, volcanic eruptions, etc.) and untectonic signals (changes caused by sea tide, polar tide, solid tide, snow, etc.). Transient signals generally refer to

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 82–97, 2019.
https://doi.org/10.1007/978-981-13-7751-8_9

tectonic signals with distinct beginnings and ends, which are caused by geophysical driving sources. The transient signal is very important information, which is conducive to the cognition and occurrence mechanism of natural disasters, the movement and deformation modes of plate boundaries, and contributes to the development of disciplines such as seismology, geodynamics and geophysics.

However, the GNSS positioning results are interfered by various error sources. There are complex spatiotemporal correlations in the GNSS position time series, the signals and noise are mixed together and difficult to distinguish and separate. Nikolaidis [2] pointed out that GNSS single-site, single-component position time series usually satisfy the formula (1.1).

$$y_{(t_i)} = a + bt_i + c\ \sin(2\pi t_i + \varphi_{ann}) + d\ \sin(4\pi t_i + \varphi_{semi}) + \sum_{j=1}^{N_{offset}} e_j H(t_i - T_j^{offset}) + \sum_{j=1}^{N_{quake}} f_j H(t_i - T_j^{quake}) t_i + \sum_{j=1}^{N_{quake}} g_j \ln[1 + (t_i - T_j^{quake})/\tau_j] H(t_i - T_j^{quake}) + v_i \tag{1.1}$$

Where y is GNSS site single-component coordinates; t indicates time, i indicates numbers of days; a is initial position; b is velocity; c, d, φ_{ann}, φ_{semi} are the amplitudes and phases of the annual and semi year terms respectively; N_{offset} is the total number of offset; T_j^{offset} is the moment when jth offset occurs; e_j is the intensity of jth offset; N_{quake} is the number of strong earthquakes; T_j^{quake} is the moment of jth strong earthquake; f_j is the velocity change after earthquake; g_j is the amplitude of post-earthquake deformation with logarithmic decay; H is Heaviside function; v_i is error.

Slow slip is a phenomenon of plate fracture and slip, which is a typical tectonic movement, generally occurring near the subduction zone that may be pregnant with super-large earthquakes of magnitude 8 or bigger. As shown in Fig. 1 [3], the simplest slow-slip model consists of two plates in the subduction zone, one tectonic plate subducts under the other. In the vicinity of the surface, the plates are locked together to accumulate potential energy, which only moves in the form of earthquakes. In the deep, geologists believe that the faults are stably slipping all year round. In the middle, slow slip occurs [3]. At present, people's understanding of slow slip is very limited. Studying slow slip is helpful to reveal the movement state of plate boundary and the complexity of earthquake displacement, and to the development of focal physics. Although the inevitability of the relationship between the occurrence of slow-slip events and earthquakes needs further study, it still preserve possibility for earthquake prediction [4].

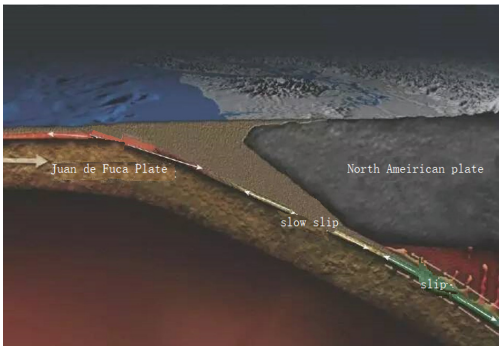


Fig. 1. Diagram of slow slip

In order to detect and extract transient signals in GNSS position time series, Dong et al. [5] and Tian et al. [6] used principal Component analysis (PCA) and correlation weighted stacking filtering to remove spatial correlation (common-mode error) in GNSS observation networks, thus improving the ability to detect weak tectonic signals. However, PCA is a method of processing information signals using second-order statistics of signals, and is not suitable for cases while signals have strong non-Gaussian properties. Because PCA requires that the principal components are not correlated, the extracted results are multi-source mixing, which makes it difficult to interpret geophysical events. Correlation weighted stacking filtering requires that GNSS observe sites have a certain density and spatial correlation, and it is possible to remove common transient signals as common-mode error. Kedar [7] and Ji [8] have successively proposed PCA-based transient signal detection methods that can improve the signal-to-noise ratio and detect possible inconspicuous transient signals mixed in noise, but they still do not solve the drawbacks of PCA. Walwer [9] proposed the Multichannel Singular Spectrum Analysis (M-SSA) method considering spatiotemporal correlations. However, the method cannot reveal the relationship between the transient deformation signal and the intrinsic geophysical driving mechanism when multiple geophysical driving sources exist.

With the increasing number of GNSS observatories and the accumulation of observation data, how to automatically, efficiently and objectively distinguish and extract the transient signals becomes a very challenging task. Common methods, such as PCA, still have some limitations, and other network-based transient signal detection methods can not be applied to the situation that lack of sites or exist high spatial correlation of local sites. Crowell [10] proposed a single-site automated detection method of transient signal based on relative strength index (RSI) and used it to detect slow-slip events in Cascadia subduction zone. The results show that the detected slow-slip events were in good agreement with previous research, and the feasibility of using this method to detect transient signals is proved theoretically and practically. This paper focuses on the method, analyzes the performance of the method from the theory and examples, and draws many useful conclusions, which can enhance the applicability and the ability to detect weak signals of the method.

2 Research Method

2.1 Relative Strength Index

Relative strength index is often used to describe the momentum of stock rising or falling, providing a reference for buying or selling. As shown in Fig. 2, there are 4 points. If the coordinates of a point are smaller than the next point, the red arrow indicates the coordinate increment *UP*. If the coordinates of one point are larger than the next point, the coordinate decrement *DOWN* is indicated by a green arrow.

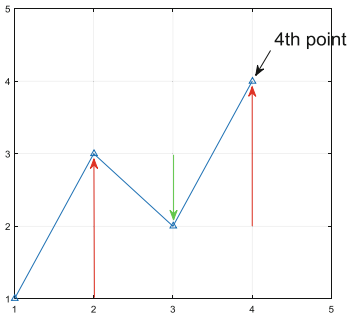


Fig. 2. Diagram of calculating RSI

The calculation formula of the RSI (80 in this example) whose window length is 3 at the 4th point is the formula (2.1), thus we can obtain the RSI of any point while the window length is arbitrary. The characteristics of RSI make it have the potential to detect transient signals: the average value is always 50 when there is no long-term trend; it does not depend on the absolute value of noise, but is sensitive to the relative changes between the rising and falling periods; for a given noise model, the standard deviation of RSI is always the same, which allows us to objectively distinguish noise and signals [10].

$$RSI_4(length = 3) = 100 - \frac{100}{1 + \frac{\sum UP}{\sum DOWN}} \tag{2.1}$$

2.2 Kurtosis

Kurtosis is the fourth central moment of distribution, which is used to measure the aggregation degree of data in the center. It can reflect the sharp or flat degree of the top of the frequency distribution curve. When the data are added at ends of the distribution, the kurtosis increases.

$$Kurtosis = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \overline{x}}{\sigma}\right)^4 \tag{2.2}$$

Where N is the length of the distribution, x_i represents the i th data, σ is the standard deviation.

If there is only random noise in the GNSS position time series, the corresponding RSI sets should be approximately normal distribution. There is a point to be noticed, the RSI series directly computed is affected by noise. In order to stabilize the solution and reduce the number of outliers, Brendan proposes to perform a center moving average process, but this process with smoothing property will cause a change in the originally normalized RSI set. Brendan simulated a large number of data of different noise types, after obtaining the RSI series and the smooth RSI series after the central moving average processing, he pointed out that the smoothing process has little effect on the normal distribution characteristics of the RSI set, and the difference is acceptable. In order to distinguish the rising and falling of coordinates, the RSI set is divided into two RSI subsets with a threshold of 50, each RSI subset obeys an approximate half-normal distribution. Since the smoothed RSI subset is also approximately half-normally distributed in the case of only random noise, their kurtosis should be a relatively fixed value, but because of the randomness of the noise, this fixed value will have a fluctuation range. GNSS position time series mainly mixes three kinds of noise: white noise, flicker noise and random walk noise, different types of noise will make this fixed value have different results, and the fluctuation range is also different. Brendan used simulated data to count the fixed values u and the standard deviation σ indicating the fluctuation range corresponding to the three kinds of noise, the specific values are shown in Table 1 [10].

Table 1. Statistics of the RSI-Kurtosis for different noise sources

Noise source	Mean u	Median	Standard deviation σ
White	4.6946	4.6523	0.3184
Flicker	4.4352	4.3998	0.3108
Random-walk	3.2199	3.2199	0.1195

In fact, due to signals, two RSI subsets will have more data at the end of the distribution (far from 50) which make them deviate from half-normal distribution, and their kurtosis may increase correspondingly. Whether the degree of increasing is within the normal fluctuation range caused by the randomness of noise, and how much probability is it regarded as a signal, we can calculate probability using to the parameters u and σ corresponding to the three kinds of noise according to calculating formula (2.3). Because the corresponding parameters of white noise are the largest, if the corresponding parameters of white noise are used to calculate the probability, it is suitable to detect the strongest signal. The corresponding parameters of random walk noise are the smallest, if the corresponding parameters of random walk noise are used to calculate the probability, weak signals can be detected.

$$probability = \frac{1}{2}[1 + \frac{2}{\sqrt{\pi}} \int_0^{\frac{k-u}{\sqrt{2}\sigma}} e^{-n^2} dn] \tag{2.3}$$

2.3 Specific Steps

The specific steps of the single-site automated detection method of transient signal based on RSI proposed by Brendan are: (1) calculating the RSI series of the original GNSS position time series with a window length of 21 days. The reason for choosing 21 days as the length of the window: more interested in slow-slip events lasting for several weeks, the usual duration of slow-slip events is also a few days to a few weeks although there is evidence that there are much longer slow-slip events in some subduction zones; if the window is too short, the RSI calculation will be affected by noise and outliers, (2) performing a 21-day center moving average for the RSI series, (3) separating the smoothed RSI values into two subsets, one greater than 50 and one less than 50 (for GPS, this is separating east/west, north/south, up/down) and sorting them by distance from 50, (4) RSI reduction-Kurtosis scheme: computing the kurtosis for each RSI subset, and then removing the most anomalous RSI value (furthest from 50) and recomputing the kurtosis of new subset. This is repeated until only 200 data points remain, the 200-point requirement is done to ensure the kurtosis does not become unstable. The result is kurtosis as a function of RSI exceedance above or below 50. The kurtosis k of the beginning (closest to 50) of the new subset to the end point (farthest away from 50) of the new subset is substituted into formula (2.3), and the probability of the end point as a signal is obtained, (5) a point with a probability of 0.95 or more is regarded as a signal point.

3 Experimental Data

The experiment uses data from the east direction of the coastal GNSS observation sites in the HIKURANGI subduction zone in New Zealand, the spatial distribution of the sites is shown in Fig. 3. This subduction zone is located at the junction of the Pacific plate and the Australian plate. The crustal movement is active, and the slow-slip event usually occurs every 1 to 2 years, and lasts for 2 to 6 weeks. Slow slip appears as a sudden, rapid and continuous rise or fall in the GNSS position time series, and its continuity distinguishes it from the offset. The dashed box in Fig. 4 shows typical slow-slip events of site GISB.

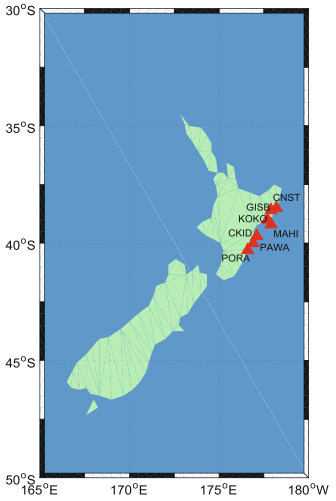


Fig. 3. Spatial distribution of experimental sites

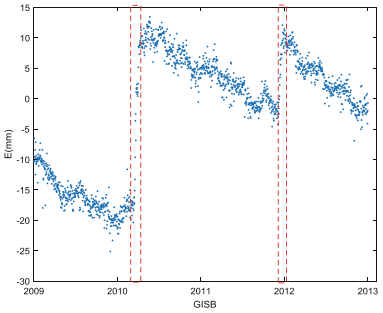


Fig. 4. Performance characteristics of slow slip in time series

4 Experimental Results Analysis

Figure 5 shows the results of choosing the corresponding parameters of white noise. The red lines correspond to the right y axis 'Probability', and the positive and negative symbols of probability indicate the rise and fall of coordinates. For the convenience of observation, the point with probability less than 0.95 is classified as 0. Except for CNST and KOKO, most of the anomalies in the time series of sites are detected, including the series with obvious slow-slip characteristics and some series which are

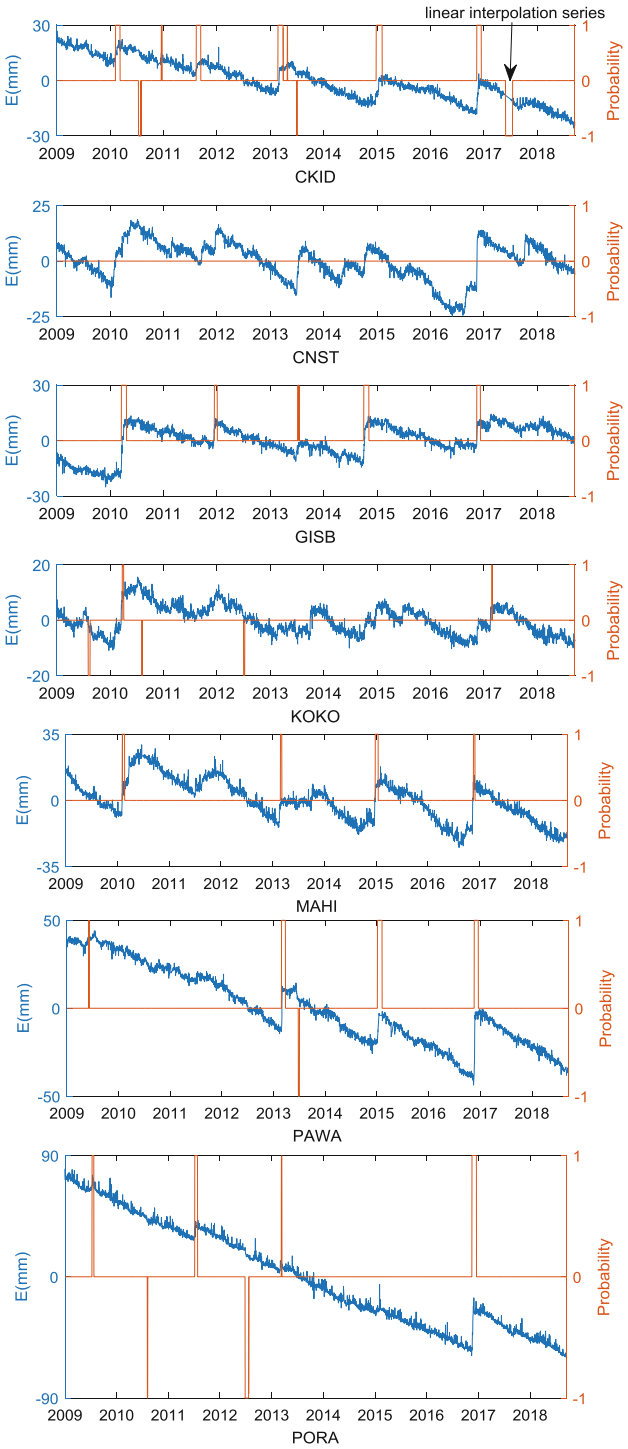


Fig. 5. Results of choosing parameters of white noise

difficult to distinguish whether or not they are signals. What these sites have in common is that the boundaries between the abnormal series (which may contain signals) and the normal series are clear, signals are basically mixed in relatively long and stable series. Another advantage of this method is the ability to find the duration of the signal. However, due to the relatively long window of RSI calculation, the sudden change in time series may not lead to the rapid change of RSI series. After the sudden change, RSI may still be prominent in a period of time, which will lead to a certain delay in the time window of detected signal. The delay can also be seen in the figure above, but the influence is not obvious. Notice that a linear series is deliberately interpolated between 2017 and 2018 (point 3060 to point 3108) of the first site as an analog signal, and the time window detected is point 3068 to point 3220. There is a certain delay at the detected time, but the duration is about the same. On the contrary, the detection results of CNST and KOKO are poor, especially the site CNST which detected no signals but exists obvious signals with slow-slip characteristics. The characteristics of two sites is that the signal and the noise are mixed together and difficult to distinguish, resulting in a relatively concentrated distribution of the RSI subsets. Therefor, the kurtosis is not large enough, and the signal is annihilated by noise. In order to improve the detection results, the corresponding parameters of random walk noise can be used to enhance the detection ability of weak signals.

As shown in Fig. 6, using the corresponding parameters of random walk noise to enhance the detection ability of weak signals, all sites detected more signal points although error detection has increased too, but the overall effect is not bad. However, the result of the site CNST is still not good enough, because the station has the lowest signal-to-noise ratio and it is very difficult to distinguish the signal and noise.

Even if a signal is very obvious, the RSI obtained is very prominent (far from 50), but if the noise level of the whole series is high, the RSI prominence value is more, the RSI distribution is not dispersed enough, the kurtosis is naturally not high, and so the probability of calculation is low. Even if a signal is weak, the RSI obtained is not prominent enough, but if the noise level of the whole time series is low, the RSI distribution is dispersed and the kurtosis is high, the probability of calculation is also high.

Therefore, the method detects relatively obvious signals. After all, signals and non-signals are relative in themselves. This can also be reflected from the comparison between the detection results of the WUHN site and the CNST site as shown in Figs. 7 and 8. The crustal activity at WUHN site is not so active, the range of change is small, and the position time series is relatively stable, so the slight change will be obvious, and the detection results are also consistent with the visible changes in the dotted frame of Fig. 7.

The comparison of RSI subset distribution between the two sites in Fig. 9 reveals the reasons for the inconsistency of detection effects: the RSI subsets distribution of WUHN site is dispersed, the kurtosis is high, and more signal points (probability 0.95 or more) can be detected in step (4).

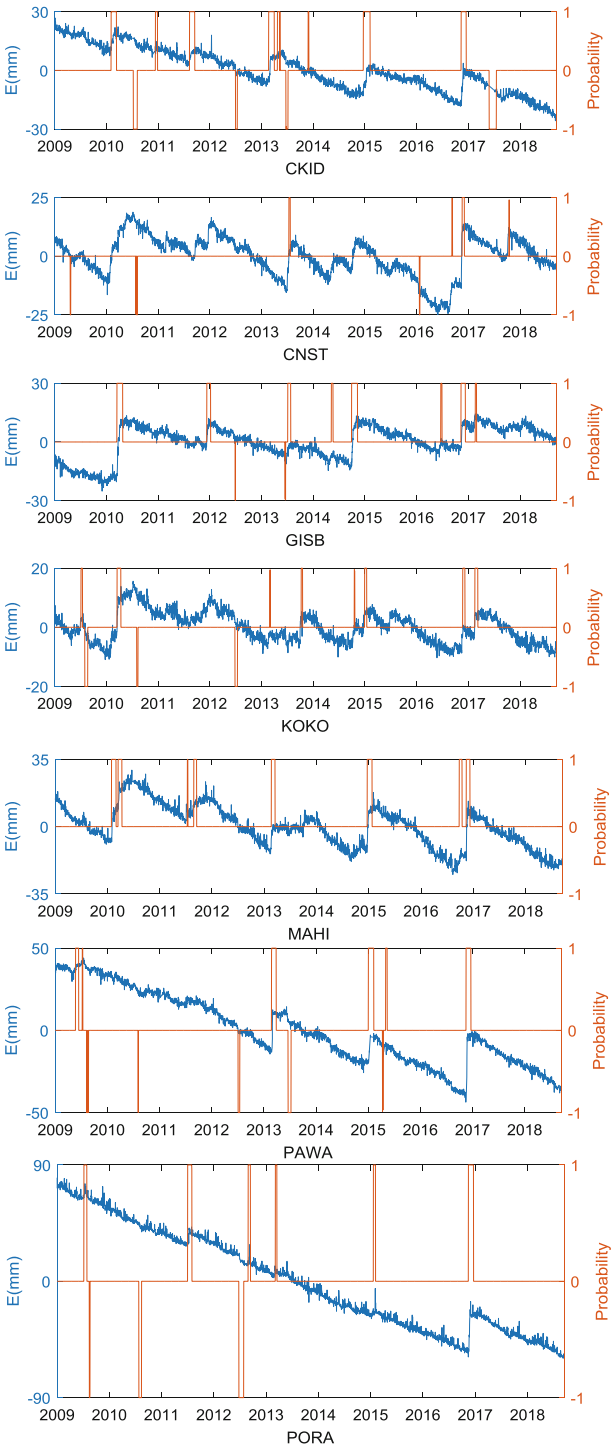


Fig. 6. Results of choosing parameters of random walk noise

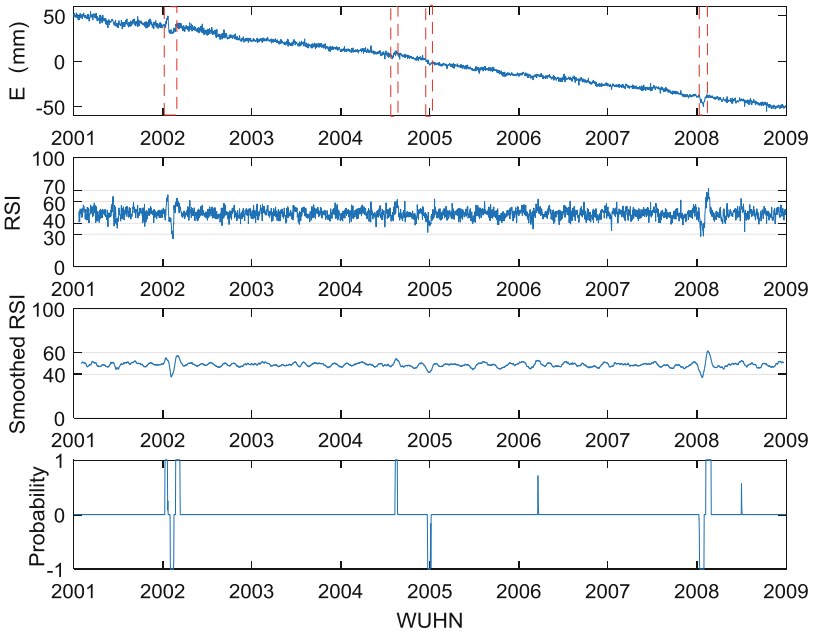


Fig. 7. Experimental results of WUHN site

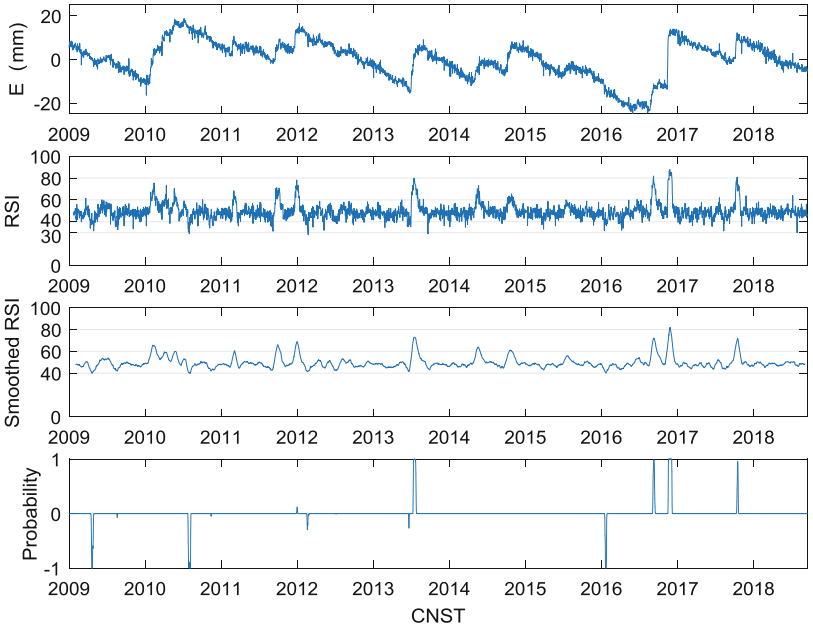


Fig. 8. Experimental results of CNST site

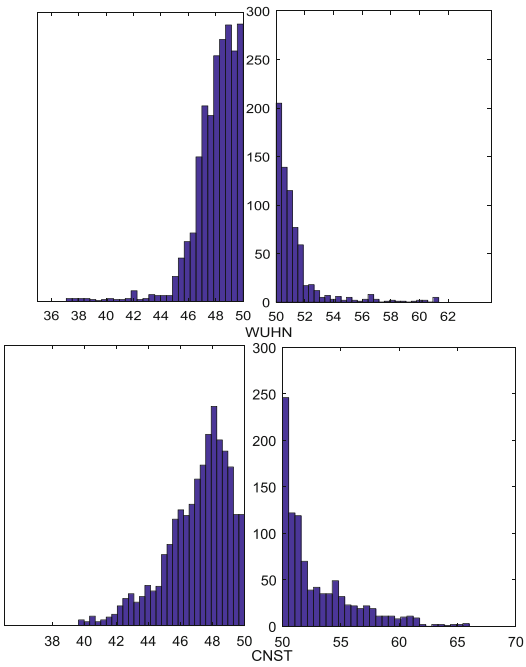


Fig. 9. Frequency distribution histograms of RSI subsets of two sites after smoothing

Considering that Brendan proposed to perform a 21-day central moving average processing in order to reduce the influence of noise on the RSI calculation, and pointed out that this smoothing processing has little effect on the characteristics of the normal distribution of the RSI set, but this conclusion is based on the assumption that only random noise exists. In fact, the noise composition in GNSS time series is mixed, and there are complex signals, which lead to the RSI set deviating from normal distribution. If we perform the central moving average, the impact on the original RSI distribution is difficult to estimate. If this is ignored, the impact of smoothing will be underestimated. In fact, changing the window length of the central moving average processing will change the distribution of the RSI subsets in varying degrees, which may result in very different detection results. The Fig. 10 shows the effect of the central moving average processing window length (from top to bottom, 17, 13, 9, and 5 respectively) on the distribution of the RSI subsets. As the window length decreases, the data at both ends of the distribution increases, and the distribution of RSI tends to stretch from the middle to both ends. When the RSI reduction-Kurtosis step performed, higher kurtosis will be obtained and more signal points will be detected. Although the parameters in Table 1 are based on the central moving average of the window length of 21 days, subjectively reducing the window length of the smoothing process to change the distribution of the RSI will result in less matching of parameters, but it do enhance the ability to detect weak signals.

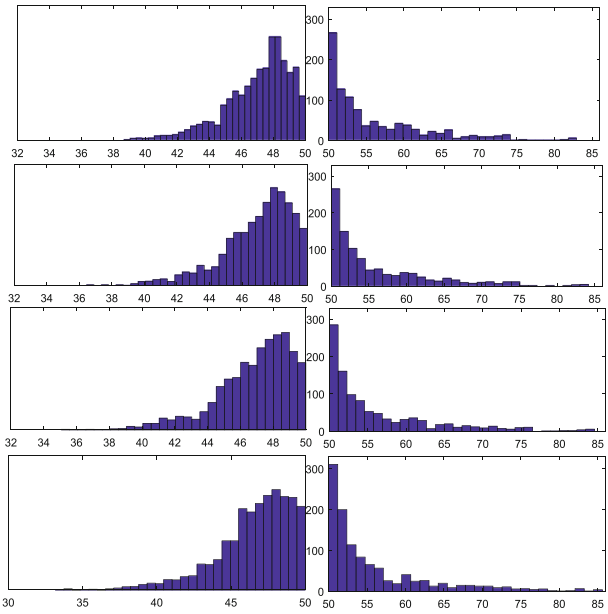


Fig. 10. The changes of RSI subsets distribution with window length

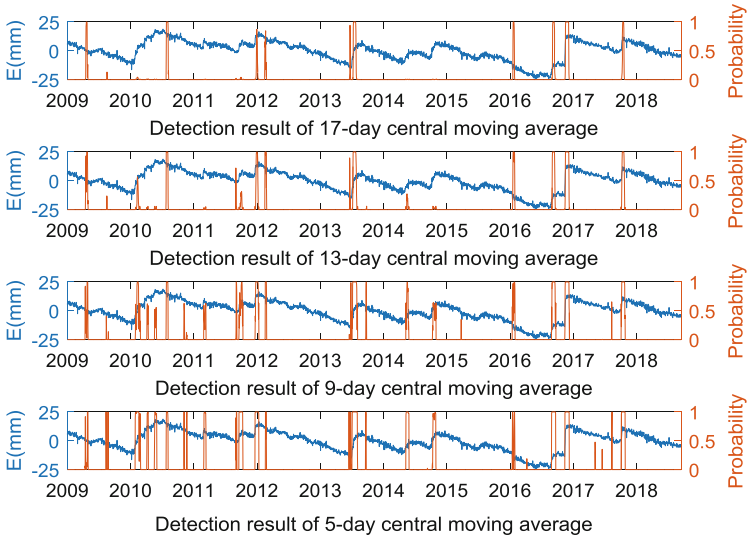


Fig. 11. Detection results with different window lengths

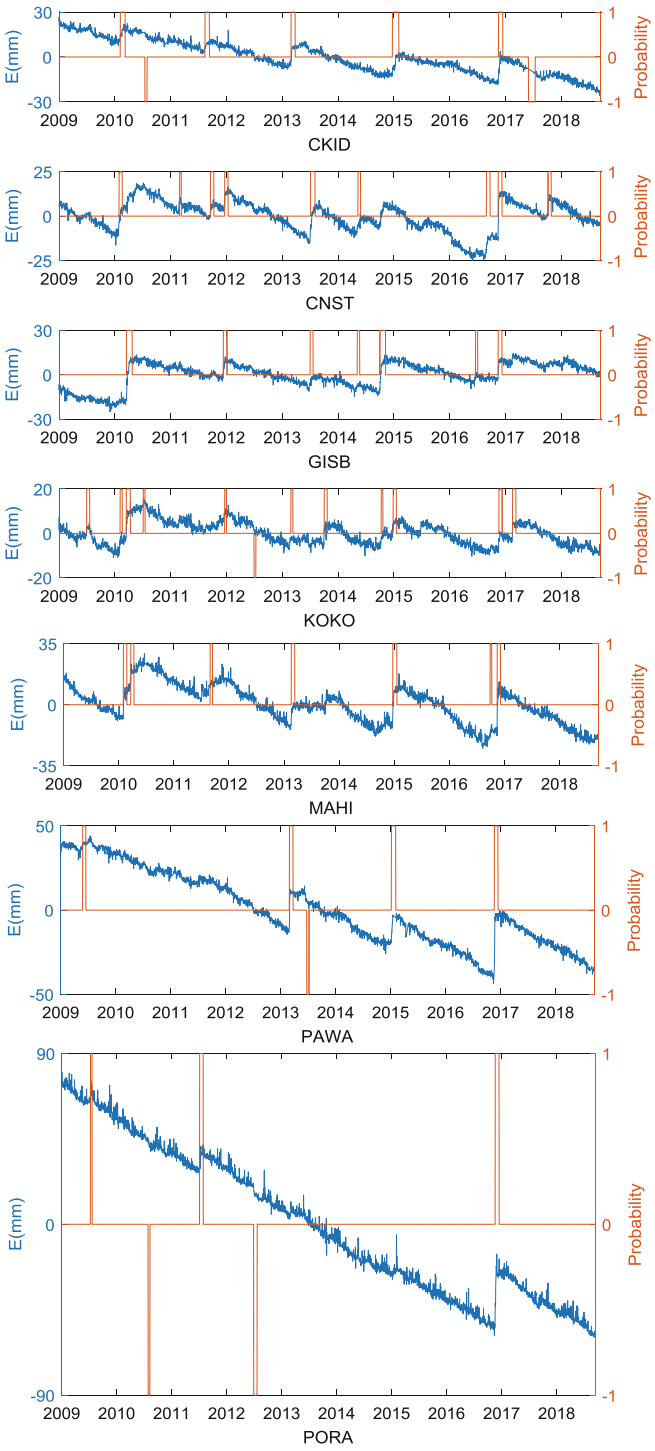


Fig. 12. Results after further enhancement of detection ability

As shown in the Fig. 11 (in order to show the change of probability of each point, the point with probability less than 0.95 is not is classified as 0), with the window length of the smoothing process reduced, the number of signal points is gradually increased, and weaker signals can be detected, but the number of error detections is also increasing. In the case we known that the slow-slip event of this subduction zone usually lasts for 2 to 6 weeks, if we want to further improve the ability of detecting weak signals by reducing the window length of smoothing processing and suppress the increase of error detections, we can eliminate the detected signals whose duration is less than 10 days. In order to further improve the ability to detect weak signals and suppress the error detections and uninterested signals, the corresponding parameters of random walk noise are selected, and 5 days are used as the window of center moving average processing, and the detection points whose duration is less than 10 days are deleted. The final result is shown in Fig. 12. Unfortunately, most of the research is focused on a single slow-slip event, although there is insufficient evidence that the detected signal is a slow-slip event, at least most of the detected signals have obvious characteristics of slow-slip events. The detection result of site GISB from September to October 2014 coincide with the time of slow-slip event recorded in paper [11].

5 Conclusion

This paper deeply analyzes the performance of the single-site transient signal detection method based on RSI proposed by Brendan from the theory and examples, and draws the following conclusions: the method can objectively detect the transient signals in GNSS time series and give the time window of their occurrence and end although there may be some delay. There are also cases of false detection and missed detection, but it is acceptable, so we can use it to detect slow-slip events; using the corresponding parameters of white noise to detect signals is suitable for time series with low overall noise level and distinct signals; using the corresponding parameters of random walk noise to detect signals can enhance the detection ability of weak signals, but also increase error detection; reducing the time window of the central moving average can further enhance the ability to detect weak signals, and is suitable for time series with high overall noise level, but also increases error detection. The duration threshold of all detected signals must be set in combination with the signal of interest (the target of this experiment is slow-slip event lasting for 2 to 6 weeks) to reject some of the false and uninterested detected signals.

Acknowledgments. This study is supported by National Science Foundation of China (41774040).

References

1. Tian Y (2011) Study on medium and long term errors in GPS position time series. Institute of Geology, China Earthquake Administration
2. Nikolaidis R (2002) Observation of geodetic and seismic deformation with the global positioning system. Cancer Res 71(8 Suppl.):714
3. Yin F, Lai G (2017) Slow slip and earthquake. Physics 46(7):462–463
4. Zhang Y, Hou M (2012) Slow slip: a new type of earthquake. Physics 41(1):39–40
5. Dong D, Fang P, Bock Y et al (2006) Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis. J Geophys Res: Solid Earth 111(B3)
6. Tian Y, Shen Z (2011) Correlation weighted stacking filtering of common-mode component in GPS observation networks. Acta Seismol Sin 33(2):198–208
7. Kedar S, Granat R, Dong D (2010) Detection of anomalous strain transients using principal component analysis and covariance descriptor analysis methods (invited). American Geophysical Union
8. Ji KH, Herring TA (2013) A method for detecting transient signals in GPS position time-series: smoothing and principal component analysis. Geophys J Int 193(1):171–186
9. Walwer D, Calais E, Ghil M (2016) Data-adaptive detection of transient deformation in geodetic networks. J Geophys Res Solid Earth 121(3):2129–2152
10. Crowell BW, Bock Y, Liu Z (2016) Single-site automated detection of transient deformation in GPS time series with the relative strength index: a case study of Cascadian slow-slip: automated transient detection. J Geophys Res Solid Earth 121(12):9077–9094
11. Wallace LM, Webb SC, Ito Y et al (2016) Slow slip near the trench at the Hikurangi subduction zone, New Zealand. Science 352(6286):701–704

Snow Depth Determination Based on GNSS-IR

Yifan Zhu[1,2,3,4(✉)] and Fei Shen[1,2,3,4]

[1] School of Geographical Sciences, Nanjing Normal University, Nanjing, China
zyfl99608@163.com
[2] Key Laboratory of Virtual Geographic Environment
(Nanjing Normal University), Ministry of Education, Nanjing, China
[3] State Key Laboratory Cultivation Base of Geographical Environment
Evolution (Jiangsu Province), Nanjing, China
[4] Jiangsu Center for Collaborative Innovation in Geographical Information
Resource Development and Application, Nanjing, China

Abstract. Snow is a critical component of the climate system and a key storage component in the hydrological cycle. Under the condition of global warming, changes in snow cover not only influence the water runoff, but may also affect the local ecosystem. Therefore, how to accurately measure the depth of snow and predict melting rate plays an important role in studying ecological changes and flood control system. However, the traditional ground-based observation method has low temporal and spatial resolution and cannot observe in real time. In recent years, with the continuous development of GNSS (Global Navigation Satellite System), GNSS-IR (GNSS interferometric reflectometry) technology, which does not require the use of special receivers and inverts surface environmental information based on ordinary geodetic receivers, has become a hot research issue. In this paper, the principle of GNSS based on SNR (signal-to-noise ratio) observation using GNSS-IR technology to detect snow depth is given in detail firstly, then the feasibility and the accuracy of GNSS-IR technology for detecting snow depth is analyzed by using Lomb-Scargle spectrum analysis method with related date by several GPS satellites. The results show that the value of the inversion is in good agreement with the measured value. The accuracy of the technique for snow-depth detection can reach about 6 cm, which can better reflect the snow depth and the variation of snow depth with time. The application of this technology is of great significance for extending the field of GNSS application and for the observation of high spatial and temporal resolution of snow depth. It has important reference value for inverting other types of surface environmental monitoring information.

Keywords: GNSS-IR · Snow depth · SNR · Lomb-Scargle algorithm

1 Introduction

Snow is an important, multi-purpose water source that is closely related to human life. More than a quarter of the world's population live in river banks and basins formed by the melting of ice, snow and glaciers. The sea level rise caused by floods and melting glaciers bring huge economic losses to humans every year. Therefore, melting glaciers

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 98–105, 2019.
https://doi.org/10.1007/978-981-13-7751-8_10

or melting snow may have a huge impact on human life and ecosystems. For managers of flood control systems, it is very important to measure the amount of water stored in snow and predict the rate of melting [1]. However, the accurate measurement of snow depth is quite difficult because the spatial and temporal variability of snow is very complicated [2, 3]. At present, the traditional methods of measuring snow depth mainly include manual measurement and snow sensor detection. However, due to the lack of sufficient spatial and temporal resolution of these methods, the measurement of snow depth is not accurate and timely to some extent, which cannot be satisfied the requirement of accuracy measurement [4].

In recent years, with the continuous development of GNSS, its application fields have become more and more extensive. Compared with its traditional applications such as positioning, navigation and timing, GNSS-IR based on SNR observation is becoming a hot research field. GNSS-IR is a remote sensing technology based on ordinary geodetic receivers for environmental element measurement. This technology has gradually gained favor from domestic and foreign scholars for its low cost and high spatial and temporal resolution. Nowadays this technology has a certain research foundation in foreign countries, but it is still in its infancy in China and requires a lot of research [5]. Larson first proposed that the detection of snow depth can be performed by SNR data and studied the problem based on snow densities and satellite elevation, and a series of valuable research results have been obtained [6, 7]. Jin combined the snow depth detection with the surface temperature, and studied the detection of snow depth based on the L4 signal by using a nonparametric statistical model [8]. In addition, Zhang focused on comparing the accuracy of detecting snow depth with single and multiple satellites [5, 9]. Wang focused on the analysis of the results of snow depth detection by arc length, satellite elevation and azimuth [10].

Because there are few studies on snow depth detection based on GNSS-IR in China, and it is still in the stage of repeated demonstration and preliminary application, lacking research on different frequency signals. Therefore, this paper discusses the theoretical methods of the technology and effects of signals of different frequencies.

2 Basic Theory of Snow Depth Detection Based on GNSS-IR

This paper takes the observation data of PRN29 of NWOT station in PBO network on the 60th doy in 2011 as an example to illustrate the process of detecting the snow depth. Figure 1 shows the erection diagram of GNSS instrument. H (antenna height) is the vertical dimension from the antenna phase center to the soil surface, h (detecting distance) is the vertical dimension from the antenna phase center to the snow surface, h_{snow} (snow depth) is the vertical dimension from the snow surface to the soil surface, O is the reflection point of the signal reflected to the antenna phase center on the snow surface, E (satellite elevation) is the angle between the incident signal and the snow surface.

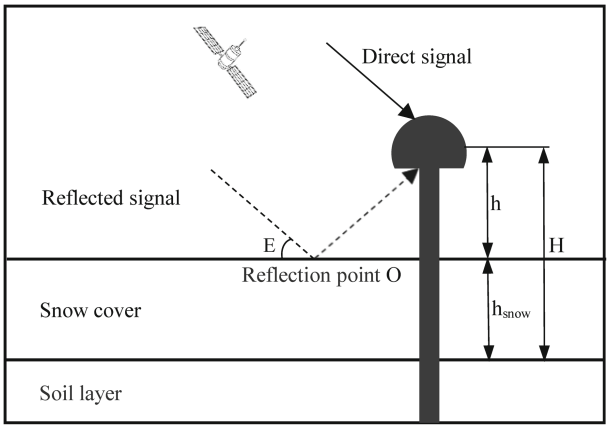


Fig. 1. Erection diagram of GNSS instrument

For precise positioning of GNSS, the direct signal received from satellites is considered as effective signal and the reflected signal is deemed as multipath signal while the snow depth detection based on GNSS-IR uses this reflected signal [11, 12]. The SNR observation indicates the signal quality received by the antenna. The value of the SNR is mainly related to the antenna gain, multipath effect, and satellite signal power [13]. In the case of low elevation angle, the SNR value will fluctuate greatly due to the influence of multipath effect. The data shown in Fig. 2 is the change of SNR value obtained based on the L1 and L2 channel, where value in the low elevation angle is in the dotted lines. The required snow depth value could be provided by processing the SNR data.

The synthesis signal in Fig. 2 is composed of direct signal and reflected signal. For ordinary geodetic receivers, direct signal that determines the trend is the main signal and its value is much larger than the reflected signal. Since the GNSS-IR uses the reflected signal to detect the snow depth, it is necessary to process the synthesized signal to remove the direct signal and extract the desired reflected signal. This paper uses second-order polynomial to deal with the signal.

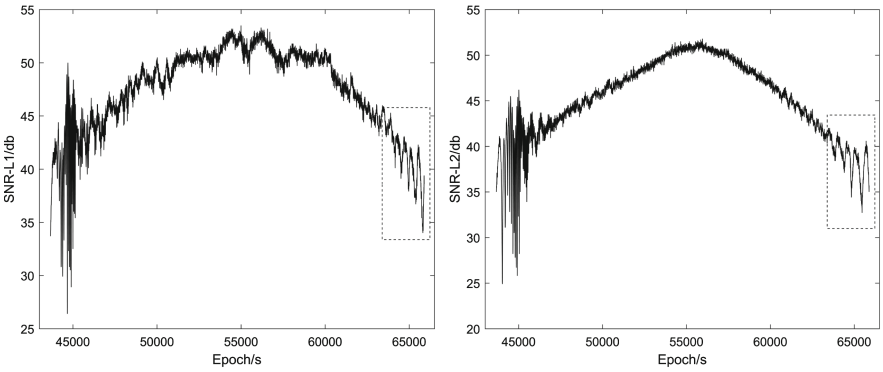


Fig. 2. SNR observation obtained based on L1 signal (left) and L2 signal (right)

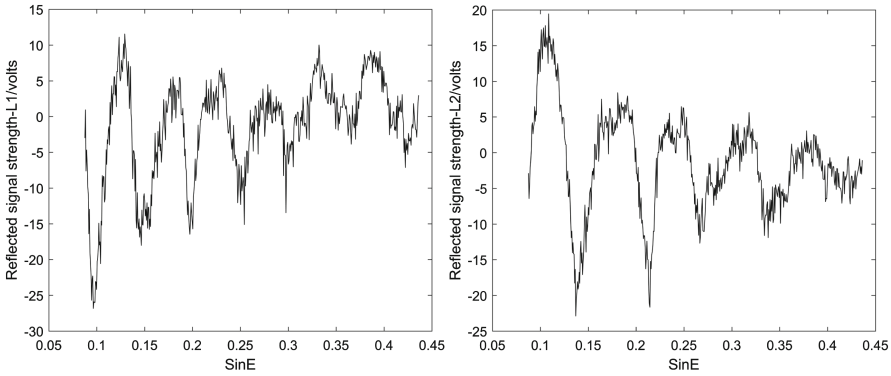


Fig. 3. Reflection signal sequences obtained based on L1 signal (left) and L2 signal (right)

The data shown in Fig. 3 is the reflected signal extracted by processing the synthesized signal. The cut-off elevation angle of satellites are taken as 5°–25°. The value of abscissa in the two figures is resampled a non-equal time intervals based on elevation angle and the linear value of ordinate is obtained by SNR data. For this signal, it can be approximated as sequence consisting entirely of reflected signal by the snow surface. From the figure, we can see that the fluctuation of the reflected signal based on L1 is obviously more complicated than that based on L2, which may increase the difficulty of calculating the effective snow depth value by using the L1 data and effect the accuracy of the results.

The reflected signal can be modeled as [14]:

$$A_r = A\ \cos\left(\frac{4\pi h}{\lambda}\sin\ E + \varphi\right) \tag{1}$$

Where,

A_r is the linear value of reflected signal,
A is the amplitude of reflected signal,
h is the detecting distance of the receiver,
λ is carrier wave length,
E is elevation angle of satellite,
Φ is initial phase

Assuming $2\pi h/\lambda = f$, the formula (1) can be written as:

$$A_r = A\ \cos(2\pi f\ \sin\ E + \varphi) \tag{2}$$

Where the frequency f is a function that the independent variable is h and reflects the trend of h. Then the value of h can be calculated by the value of f. Since the function is resampled at non-equal time intervals, the frequency f can be obtained by using Lomb-Scargle periodogram or Fourier transformation with cubic spline interpolation. The paper uses Lomb-Scargle periodogram to estimate f. The data shown in Fig. 4 is

the value of h obtained by f and the maximum value is the actual effective distance. Then the h_{snow} is given by:

$$h_{snow} = H - h \tag{3}$$

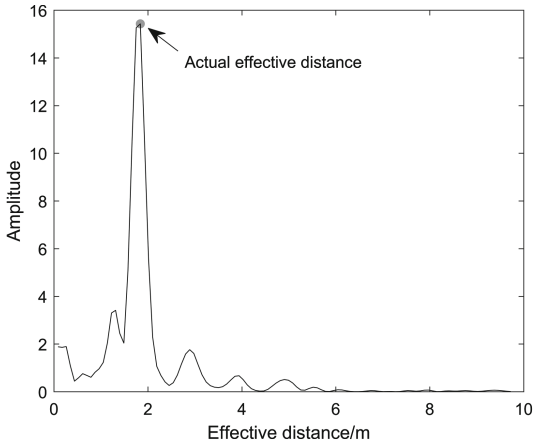


Fig. 4. L-S spectrum obtained based on L2 signal

3 Experiments and Results

In order to explain how GNSS-IR is applied to snow depth detection, this paper takes the observation data (PRN13, 21, 29, 32) of NWOT station in PBO network on the 35th, 60th, 90th, 130th, 175th, 225th and 300th doy in 2011 as examples to illustrate the process of detecting the snow depth (https://www.unavco.org/data/gps-gnss/data-access-methods/dai2/app/dai2.html#). Moreover, the contrastive analysis is made with the meteorological data from PBO H_2O studying team (http://xenon.colorado.edu/portal/index.php?product=snow). The antenna model of this station is TRM41249.00 and the data sampling rate is 1 s. The receiver will not be covered by snow due to the influence of perennial wind. Figure 5 shows the environment around the station.

Here the results of 175th day and above 7 days based on L1 and L2 are shown in Fig. 6. From the results in Fig. 6 (left), it can be seen that the results obtained based on L2 are in good agreement with the measured results and using the average of values obtained by multiple satellites can effectively avoid the influence of outliers, but the results acquired based on L1 are quite different from the measured results. From Figs. 2 and 3, the larger fluctuations of the SNR of L1 may contribute to this situation. From the results in Fig. 6 (right), the detection results based on L2 agree well with the measured results and could reflect the change of snow depth with the number of days. However, there is a large deviation in the data of 130th doy. Combined with meteorological data, it is confirmed to be affected by snowfall, which enhanced the multipath effect. Also, the results of these days obtained based on L1 also show big difference

Fig. 5. Surrounding environment of the NWOT station (http://xenon.colorado.edu/portal/index.php?product=snow&station=nwot)

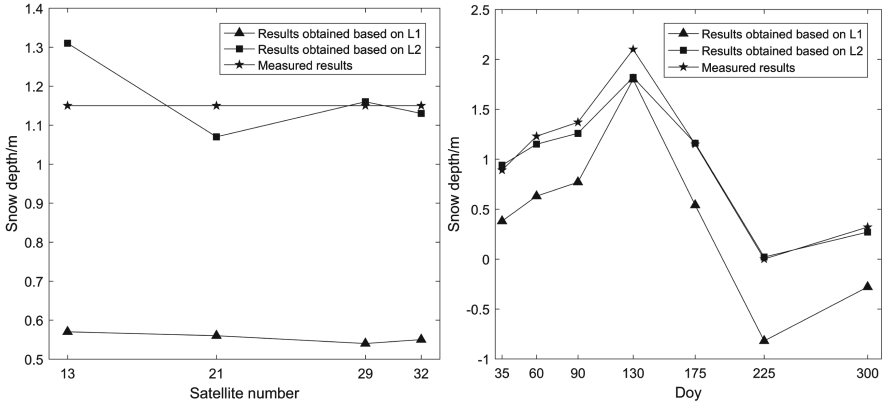


Fig. 6. Comparison of snow depth on the 175th doy obtained based on L1 and L2 signals (left) and Comparison of seven-day snow depth obtained based on L1 and L2 signals (right)

from measured results, which further illustrates the need for more in-depth research on how to use the L1 signal to detect snow depth. In summary, the following conclusions can be drawn:

(1) It confirms the feasibility of using GNSS-IR to detect snow depth and the results can better reflect the change of snow depth with time.
(2) GNSS signals with different frequencies will have a great impact on the detection results. The maximum value of the difference between the measured results and results obtained based on L1 is 0.76 m, the average value of the difference is 0.55 m, and the standard deviation of the difference is 0.14 m. The maximum value of the difference between the measured results and results obtained based on

L2 signal is 0.28 m, the average value of the difference is 0.06 m, and the standard deviation of the difference is 0.11 m. The detection results obtained based on L2 match the measured results well and it can achieve centimeter-level accuracy. The detection results of the L1 signal are quite different from the measured results and cannot be used for snow depth detection.

(3) The influence of outliers can be effectively avoided by using the average of values obtained by multiple satellites.

(4) There is a big difference between the detection results and the measured results in the snowy weather, which indicates that the multipath effect would enhance in the snowy weather and would have a great impact on detecting the snow depth.

4 Discussion

Different from the traditional GNSS applications, snow depth detection based on GNSS-IR utilize the multipath signals that are not necessary in traditional applications. It makes full use of the advantages of easy access and low cost of the ordinary geodetic receiver. Compared to the conventional instruments, its result has a higher spatial and temporal resolution. In summary, snow depth detection based on GNSS-IR has attractive application prospects.

This paper verified the feasibility of GNSS-IR for snow depth detection based on numerical analysis and found that the following problems need further research: (1) The impacts that different frequencies of GNSS signal have on GNSS-IR to detect snow depth. (2)The optimization for the model of snow depth detection based on GNSS-IR. (3) Snow depth detection on different slopes of ground. (4) Construction of GNSS-IR model for snow depth detection with different snow density.

With the unique advantages, GNSS-IR is gradually attracting attention of scholars all over the world. Especially with the improvement of BDS, it can further expand the potential of applications of remote sensing and add new impetus to the development of BDS.

References

1. Shi J, Dozier J (2000) Estimation of snow water equivalence using SIR-C/X-SAR. I. Inferring snow density and subsurface properties. IEEE Trans Geosci Remote Sens 38 (6):2465–2474
2. Erickson TA, Williams MW, Winstral A (2005) Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. Water Resour Res 41(4)
3. Armstrong RL, Brun E (2008) Snow and climate: physical processes, surface energy exchange and modeling. Cambridge University Press, Cambridge
4. Eisen O, Frezzotti M, Genthon C et al (2008) Ground-based measurements of spatial and temporal variability of snow accumulation in East Antarctica. Rev Geophys 46(2)
5. Zhang S, Dai K, Nan Y et al (2018) Preliminary research on GNSS-MR for Snow depth. Geomat Inf Sci Wuhan Univ 43(2):234–240

6. Larson KM, Gutmann ED, Zavorotny VU et al (2009) Can we measure snow depth with GPS receivers? Geophys Res Lett 36(17):L17502
7. Larson KM, Nievinski FG (2013) GPS snow sensing: results from the EarthScope Plate Boundary Observatory. GPS Solutions 17(1):41–52
8. Jin S, Najibi N (2014) Sensing snow height and surface temperature variations in Greenland from GPS reflected signals. Adv Space Res 53(11):1623–1633
9. Zhang S, Dai K, Liu K et al (2016) Research of GPS-MR on snow depth monitoring. Prog Geophys (in Chin) 31(4):1879–1884
10. Wang Z, Liu Z, An J et al (2018) Snow depth detection and error analysis derived from SNR of GPS and BDS. Acta Geod Cartogr Sin 47(1):8–16
11. Auber J (1994) Characterization of multipath on land and sea at GPS frequencies
12. Liu J, Shao L, Zhang X (2007) Advances in GNSS-R studies and key technologies. Geomat Inf Sci Wuhan Univ 32(11):955–960
13. Bilich AL, Larson KM, Axelrad P (2004) Observations of signal-to-noise ratios (SNR) at geodetic GPS site CASA: implications for phase multipath. Proc Cent Eur Geodyn Seism 23:77–83
14. 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)

Multi-GNSS Kinematic Relative Positioning of Multiple Base Stations and Its Application in Deformation Monitoring

Honglin Tang[1] and Tianhe Xu[2,3](✉)

[1] School of Geology Engineering and Surveying, Chang'an University, No. 126, Yanta Road, Xi'an, China
954106049@qq.com
[2] Institute of Space Science, Shandong University, Weihai, Shandong, China
thxugfz@163.com
[3] State Key Laboratory of Geo-Information Engineering, Xi'an, Shanxi, China

Abstract. The traditional single base station GNSS deformation monitoring system is greatly influenced by the base station stability and data quality. In order to improve the accuracy and reliability of monitoring system, this paper develops a new method of real-time kinematic relative positioning based on multiple base stations and multi-GNSS. Aiming at the problem of weight allocation of multiple base station and different GNSS, the idea of Helmert variance component estimation and robust estimation are introduced. Through a real example analysis, the robust variance component weighting method is more efficient and reliable than the traditional range-based (baseline length) weighting method for different base station and empirical weighting method for different GNSS. The accuracy of the deformation monitoring of the proposed method is higher and more reliable than those of single base station with GNSS or multi-GNSS.

Keywords: Deformation monitoring · Real-time relative positioning · Robust estimation · Variance component estimation · Multiple base station

1 Introduction

With the continuous improvement of BeiDou Satellite Navigation System (BDS), the situation of multi-GNSS coexistence and development has taken shape. Multi-GNSS has become a trend of development and brought new opportunities and challenges to GNSS theoretical algorithms, data processing and applications [1]. Since both BDS and Global Position System (GPS) adopt the code division multiple access communication mode, the combined positioning effect of the two systems is more significant [2], which can increase the number of observation satellites and improve the spatial distribution geometry of the satellite. Multi-GNSS combination positioning needs to reasonably determine the weight ratio between different GNSS observations. If the weight ratio between different GNSS observations is incorrect, it will affect the positioning accuracy and hard to meet the requirements of high-precision data processing.

The base station serves as the initial benchmark for the GNSS deformation monitoring network and requires long-term reuse. For traditional single-base station GNSS

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 106–113, 2019.
https://doi.org/10.1007/978-981-13-7751-8_11

deformation monitoring system, monitoring accuracy is greatly affected by the stability of the base station and data quality. Deformation monitoring network based on single GNSS system is becoming more and more inadequate, which brings hidden risk for deformation monitoring of dam, bridge and other deformation bodies. More and more people pay attention to it and propose an effective solution of adopting multi-base station GNSS deformation monitoring system. Real-time kinematic relative positioning of multiple base stations also needs to reasonably determine the weight ratio between different base station observations. Otherwise, the accuracy of multi-base stations data fusion results will be affected.

In view of the above two problems, this paper introduces the idea of robust estimation and variance component estimation into multi-base stations and multi-GNSS relative positioning. For the weighting problem between different observations, the Helmert variance component estimation based on the least squares principle is adopted. However, the least squares does not have robustness. When the observations contain gross errors, it will affect the optimal unbiased estimation of the parameters, which seriously affects the accuracy of the solution. To this end, this paper adopts IGG-III scheme to construct the equivalence weight function, and reduces the weight of the observation of gross error by the weight reduction factor. The method is applied to the multi-GNSS real-time kinematic relative positioning of multi-base stations. The experimental results verify the effectiveness of the proposed method.

2 Data Processing Theory

2.1 Unification of Space-Time Datum

The BDS adopts the time reference of BDT and the GPS adopts GPST. Before the combined positioning, the time reference should be unified. The relationship between the weekly and the second conversion of the time system is as follows:

$$\begin{aligned} \mathrm{BDS} &= \mathrm{GPS} - 1356(\mathrm{weekly}) \\ \mathrm{BDS} &= \mathrm{GPS} - 14(\mathrm{second}) \end{aligned} \tag{1}$$

The GPS adopts the coordinate system is the WGS-84 coordinate system, and the BDS adopts the CGCS2000 coordinate system. There are slight differences between the two coordinate system, which can be ignored in the relative positioning at short baseline [3].

2.2 Helmert Variance Component Estimation

The basic idea of Helmert variance component estimation is that: firstly, the initial weight of various types of observation is determined for example by elevation angle model; and adjustment is conducted according to the least squares principle to obtain the residual square sum $V^T PV$ and re-weight the observations under certain criteria; the above step is repeated through iterative calculation and stopped until the difference of the unit weight variance for different types of observations is less than the specified threshold (the threshold is set as 0.001 in this paper).

Let the observations L1 and L2 are the GPS and BDS observations respectively. Taking the indirect adjustment as an example, the error equation can be expressed as:

$$\begin{aligned} V_1 &= A_1\hat{X} - L_1 \\ V_2 &= A_2\hat{X} - L_2 \end{aligned} \tag{2}$$

The formula for estimating the variance of two types of observations is

$$S_{2\times 2}\hat{\sigma}^2_{2\times 1} = W_{2\times 1} \tag{3}$$

The k + 1 weight matrix calculation formula can be expressed as:

$$P_i^{k+1} = c/\sigma_{0i}^2 \cdot P_i^k \tag{4}$$

where $\hat{\sigma}^2_{2\times 1} = \begin{bmatrix} \hat{\sigma}^2_{01} \\ \hat{\sigma}^2_{02} \end{bmatrix}$, $W_{2\times 1} = \begin{bmatrix} V_1^T P_1 V_1 \\ V_2^T P_2 V_2 \end{bmatrix}$, $\mathrm{S} = \begin{bmatrix} n_1 - 2tr(N^{-1}N_1 + tr(N^{-1}N_1)^2 & tr(N^{-1}N_1N^{-1}N_1) \\ tr(N^{-1}N_1N^{-1}N_1 & n_2 - 2tr(N^{-1}N_2 + tr(N^{-1}N_2)^2 \end{bmatrix}$.

$\mathrm{N} = A_1^T P_1 A_1 + A_2^T P_2 A_2 = N_1 + N_2$, n1 and n2 are the observation numbers of GPS and BDS, respectively. Through Eqs. (3) and (4), the unit weight variance is iteratively calculated until the unit weight variances of different types of observations are equal [4, 5]. The above method can be similarly applied to the determination of the weight ratio between different base station data, which is not repeated here for the length limitation.

2.3 IGG-III Equivalent Weight Function

Robust estimation puts the gross errors into a stochastic model. It has both stability and robustness. There are many robust estimation methods. In this paper, we choose the weighted iteration method in the generalized maximum likelihood estimation method. The formula is as follows:

$$\bar{P}_{ig} = P_{ig}\omega_i \tag{5}$$

Because it is simple and easy to understand, it is widely used and calculated by the appropriate weight function. The IGG-III related equivalence factor function is used as follows [6]:

$$\omega_i = \begin{cases} 1 & |u_i| \le k_0 \\ \frac{k_0}{|u_i|}\frac{(k_1 - |u_i|)^2}{(k_1 - k_0)^2} & k_0 \le |u_i| \le k_1 \\ 0 & |u_i| \ge k_1 \end{cases} \tag{6}$$

where $u_i = v_i/\sigma$, and it is the medium error of the correction number v, K_0 is generally taken from 1.0 to 1.5, and K_1 is usually valued from 2.5 to 3.0.

2.4 Flow Chart of Robust Helmert Variance Component Estimation Algorithm

See Fig. 1.

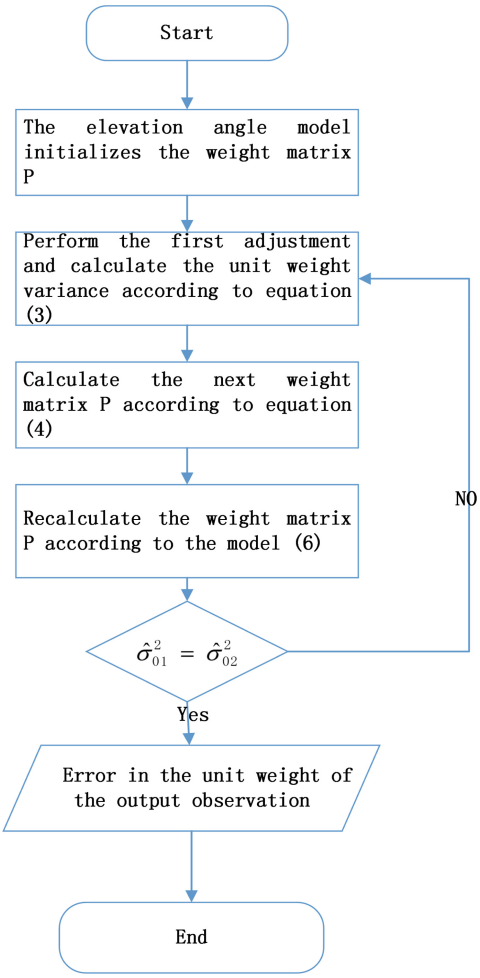


Fig. 1. Flow chart of the robust Helmert variance component estimation algorithm

3 Case Analysis

3.1 Multiple Base Stations Data Fusion

This paper uses GNSS monitoring network data of Xi'an Jingyang County for computation and analysis. The data sampling rate is 5 s with a total of 24 h of observation, and the number of satellites in each system during the observation period is as follows:

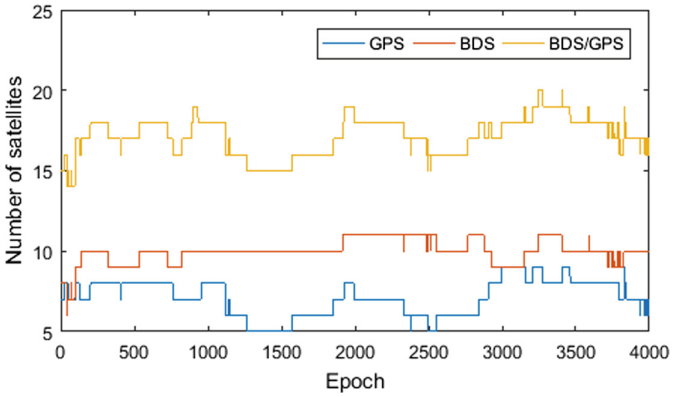


Fig. 2. BDS/GPS visible satellite number

It can be seen from Fig. 2 that the visible number of satellites during the observation period is normal, and the experimental requirements are met. The weight of the multi-base station using the traditional baseline length cannot fully reflect the quality of the baseline solution at the observation time. The quality of the baseline solution is determined by the baseline length and the base station. For the impact of multiple factors such as reliability [6], it is obviously unreasonable to use a single standard for data fusion. Baselines of different lengths should be regarded as different types of observations in a broad sense. Therefore, the robust Helmert variance component estimation is used to determine the weights, which not only takes into account the influence of gross error, but also makes full use of all the observations. The real-time kinematic relative positioning accuracy is very high, and the resolution of the single base station is in the same magnitude, so the data fusion can adopt the weighted average. Before data fusion, the base station needs to be benchmarked to eliminate the contradiction between stations. Then, the single base station is separately decentralized, and then the data is merged. The initial datum selection is to combine the dual system to get the coordinate average value of one day. JY01 is used as the monitoring point, three base stations are selected for weighted averaging, whose weight is determined by the baseline length. The multi-base station data fusion is performed according to the robust Helmert variance component. The accuracy of the fusion result is listed as follows:

Table 1. Data fusion accuracy statistics

Baseline name	Baseline length (m)	E/mm STD	N/mm STD	U/mm STD
JY01-JY02	106.53	1.11	1.41	3.81
JY01-JY03	227.41	1.10	1.43	3.46
JY01-JY05	699.35	1.65	1.88	4.21
Baseline length fusion	—	1.16	1.46	3.75
Robust component fusion	—	1.14	1.44	3.46

It can be seen from Table 1 that the robustness of the Helmert variance component is improved compared with the traditional weighting method. The accuracy of E, N, and U directions of the proposed method are better than those by the weighting method of the baseline length, Because the new model considers the influence of gross errors and random model errors on the experimental results, the real-time kinematic relative positioning of multi-base stations overcomes the shortcomings of single base station deformation monitoring system, and makes the positioning results more reliable and the positioning accuracy more higher. More realistic reflection of the actual displacement of the monitoring point can be found by the proposed method.

3.2 Multi-GNSS Combined Baseline Solution

The BDS/GPS dual-frequency observation data of Xi'an Jingyang County was selected. The JY01-JY05 baseline is selected for experimental analysis. The data solution is adopted as kinematic relative positioning mode, and the integer ambiguity solution is solved by the real-time mode based on floate solution. The EKF (Extended Kalman Filter) is used for parameter estimation [7, 8], and the variance component estimation and robust estimation are added in the software. For the empirical weighting method, the weight ratio of BDS and GPS is selected as 1:2 [9]. The coordinate difference sequence with 3000 epochs are calculated and compared relative to the initial datum. The residual sequences of the E, N and U direction are shown in Fig. 3. Table 2 summarizes the accuracy statistics of the two weighting methods.

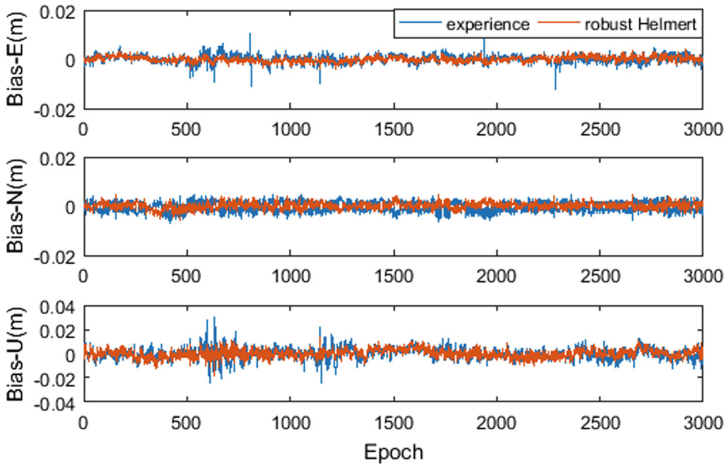


Fig. 3. Comparison of different weighted Methods positioning results

It can be seen from Fig. 3 and Table 2 that the positioning results of the two weighting method are different, and the precision of the robust Helmert variance component method is much higher, because the empirical weighting based on the elevation angle model has no ability to resist the gross error, and the actual quality of

the observations cannot be truly reflected. In some observation epochs, the data oscillating phenomenon is more significant due to the factors such as multipath, observation noise and sidereal day error. The robust weighting data is more smooth. There is no obvious oscillating phenomenon, and it has strong resistance to gross error. Therefore, the robust Helmert variance component estimation can not only properly balance the two types of observation weights in BDS/GPS combined positioning, but also resist the influence of gross error. The positioning accuracy is higher and the positioning result is more reliable.

Table 2. Accuracy statistics of different weighting methods

Stochastic model	E/mm	N/mm	U/mm
Elevation angle	1.81	1.83	4.92
Robust Helmert variance component	1.10	1.81	3.62

4 Conclusion

Due to the systematic difference between BDS and GPS and the multiple base stations fusion, it is not optimal to use the empirical or baseline length weighting method. The weight of robust Helmert variance component has good resistance to gross error and good balance between different GNSS and different baseline. It overcomes the defects of least squares without any robustness for gross error, and can more reasonably determine the weight of observation data between different base station and different GNSS. The results of the actual deformation monitoring example show that the multi-GNSS kinematic relative positioning with multi-base station has higher positioning accuracy and more reliable positioning results than the traditional single GNSS with single base station or multi-GNSS with single base station. All in all, the proposed method can effectively improve the accuracy and capability of GNSS deformation monitoring.

Acknowledgements. The study is funded by National Natural Science Foundation of China (41574013, 41731069 and 41874032) and National Key Research & Development Program of China (2016YFB0501701).

References

1. Zhou F (2018) Theory and methodology of multi-GNSS undifferenced and uncombined precise point positioning. East China Normal University
2. Gao X, Dai W (2014) Application of robust Helmert variance estimation in GPS/BDS combined positioning. J Geod Geodyn 34(01):173–176
3. Han J (2017) Research on BDS/GPS relative positioning algorithm and its application in landslide monitoring. Chang'an University
4. Zhang Q (2011) Modern measurement data processing and application. Surveying and Mapping Press, Beijing

5. Wen C, Gao L, Fang J et al (2014) The high-precision weighing system based on the improved amplitude-limiting and average filtering algorithm. Chin J Sens Actuators 27 (5):649–653
6. Yang Y (1994) Anomaly influence diagnosis and robust estimation. Surv Mapp Bull 5:34–36
7. Chen Y (2014) Research on position of BD-2/GPS integrated navigation based on school bus safety service system. Xi'an University of Science and Technology
8. Yang Y, Xu T (2003) An adaptive Kalman filter combining variance component estimation with covariance matrix estimation based on moving window. J Wuhan Univ (Inf Sci Ed) 28 (06):714–718
9. Liu J, Tu R, Zhang R, Zhang P, Lu X (2018) Application of Helmert's variance estimation in GPS/GLONASS/BDS combined positioning weight ratio. J Geod Geodyn 38(06):568

Fusion of Radiosonde and GPS Data for Water Vapor Tomography

Na Liu[1,2,3(✉)], Pengfei Zhang[1,2,3], Yuping Gao[1,2,3], Pingli Wang[1,2,3], and Chengshi Zhao[1,3]

[1] National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China
liuna@ntsc.ac.cn
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Key Laboratory of Time and Frequency Primary Standards, Xi'an 710600, China

Abstract. The technique of ground-based GPS tomography is a newly spatial tool to sense the water vapor profile in atmosphere area. It has become a complementary technique for other conventional atmospheric sounding techniques. In this study, the radiosonde data was utilized to improve the accuracy of GPS tomography by forming the vertical constraint information. The experiment of Xi'an area tomography campaign was employed to assess the performance. The result shows that approach of fusion of radiosonde and GPS data for water vapor tomography can improve the accuracy of wet reflectivity, increase utilization of radiosonde data.

Keywords: Ground-based GPS · Sequential least square · Tomography water vapor · Wet refractivity

1 Introduction

Recently, the GPS technique is widely applied in the area of atmospheric sounding due to its advantage of low-cost, high accuracy, as well as better spatial and temporal resolution than other conventional sounding technique [1]. In fact, Flores proved the validation of ground-based GPS water vapor tomography and performed a preliminary research on water vapor tomography algorithms in earlier year based on experimental work and data collected from the European Centre for Medium-Range Weather Forecasts (ECMWF) [2]. Skone proposed the approach that moving voxels to perform water vapor tomography, which has a sufficient level of accuracy to estimate the three-dimensional (3D) spatial distribution of atmospheric vapour [3]. In 2006, the systematic investigation on tomography algorithms using additional constraints and validated these algorithms in experiment was performed; and then, the 3D spatial distribution of water vapor over the Shanghai region was also obtained [4]. Meanwhile, the tomographic approach with additional constraints and Kalman filtering was studied [5]. In 2010, the mechanism on different constraint effects on GPS water vapor tomography was analysed [6]. In order to assign weight of different observation and suppress the outliers, the approach of robust estimation of variance components vapour

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 114–121, 2019.
https://doi.org/10.1007/978-981-13-7751-8_12

tomography was studied [7]. In 2017, the virtual slant signals approach was proposed to improve the accuracy of water vapor tomography models, and then the high-quality tomographic results were obtained based on Hong Kong region observation [8]. To improve the utilization of signal rays cross through the addition of an assisted tomographic area, the approach of water vapor inversion was proposed, which showed better effectively when utilized the GPS information that cross through the sides of an area and enhanced the accuracy of tomographic result [9].

The aforementioned studies have greatly advanced the development of ground-based water vapor tomography techniques and laid a solid foundation for their extensive application. In actual practice, water vapor detection is carried out predominately via the observational data from the Continuously Operating Reference Station (CORS) network, along with the meteorological data collected by automatic weather stations. However, in most circumstances, the height difference between each station within the CORS network is insufficient to reflect the actual vertical distribution of water vapor. Therefore, to obtain a higher accuracy, calculations should also incorporate water vapor data acquired from other sources, such as prior radio sounding information [10]. On the other hand, although the traditional sounding data presents a high accuracy, it acutely lacks temporal resolution, as measurements are often performed only at UTC 0:00 and 12:00.

To address the issues of radio sounding data (high vertical distribution resolution with few measurements) and GPS tomography (high sampling rate with insufficient vertical distribution resolution), a novel water vapor tomography approach of fusion radiosonde and GPS data was proposed in this study, which applied radio sounding data to the vertical constraints and adopted sequential least squares for parameter estimation. This approach not only effectively improved the accuracy of ground-based GPS water vapor tomography, but also overcome the issue of insufficient temporal resolution in traditional radio sounding data. Thereby, it helps in promoting the routine operation of water vapor spatial distribution tomography.

2 The Approach of Ground-Based GPS Tomography Water Vapor with Additional Constraints

The technique of ground-based GPS water vapor tomography generally adopts a three-dimensional voxel model. The observation equation is established according to the cross length of each GPS signal in passing grids. Since the water vapor density is regarded unaltered in ideal hypothesis case, the number of observations usually exceeds the number of unknown parameters. However, there is no guarantee that all voxels will have GPS signals, which will lead the observation equation rank deficiency and undetermined [7]. In order to solve the problem, a certain constraints were usually applied. Generally, those constraints included horizontal and vertical constraints, which were determined by mathematical relationship and physical laws. In this study, Gaussian weighted constraint was adopted to establish the horizontal constraint, and the latest moment radiosonde data was used to form the vertical constraint, which are processed as virtual observations when the least squares is used for calculation.

Therefore, the three observation equations in the tomography with additional constraints will be discussed as follows.

2.1 Observation Equation

The observation equation for GPS slant water vapor can be expressed as follows:

$$Swd = \sum \left(L_{i,j,k} \cdot x_{i,j,k} \right) \tag{1.1}$$

where Swd denotes the GPS slant water vapor; i, j, k represent the index of grids in three coordinate components; $L_{i,j,k}$ is the distance of GPS satellite passing through the established voxels; $x_{i,j,k}$ is the water vapor density in every voxels.

2.2 Horizontal Constraints

With respect to horizontal constraint, the constraint was established according to the characteristics of water vapor distribution in the atmosphere, and then the relationship between adjacent grids is used. Flores et al. [2] proposed that the relationship between the grids can be expressed as a weighted average form, which can be express as:

$$0 = \omega_1 \cdot x_1 + \cdots + \omega_{j-1} \cdot x_{j-1} - x_j + \omega_{j+1} \cdot x_{j+1} + \cdots \tag{1.2}$$

where x_j, ω_j represent the refractivity and coefficient of grid j, respectively. In fact, the coefficient ω_j can be determined based on Gaussian weighted function [4], the formula is:

$$\omega_i = \begin{cases} 0 & , \quad i,j \text{ in same layer} \\ \dfrac{e^{-\frac{d_{i,j}^2}{2\sigma^2}}}{\sum e^{-\frac{d_{ii,j}^2}{2\sigma^2}}} & , \quad i,j \text{ in different layer} \end{cases} \tag{1.3}$$

where, $d_{i,j}$ denotes the distance between grid i and j; σ is the smoothing factor, it is usually determined by the horizontal length of grid.

2.3 Vertical Constraints

In the vertical constraints, the vertical constraint equation was established using radiosonde data.

$$x = I_r x_r \tag{1.4}$$

where x_r is the wet refractivity in different vertical layers deprived from radiosonde data. I_r denotes the coefficient matrix.

3 Fusion of Radiosonde Data and GPS for Water Vapor Tomography

As discussed above, the observation equation of fusion of radiosonde and GPS data for water vapor tomography can be further formed as follows.

$$\begin{bmatrix} Swd \\ {\scriptstyle m\times 1} \\ Horiz \\ {\scriptstyle s\times 1} \\ \mathrm{Re}f \\ {\scriptstyle s\times 1} \end{bmatrix} = \begin{bmatrix} H_s \\ {\scriptstyle m\times s} \\ H_h \\ {\scriptstyle s\times s} \\ I_r \\ {\scriptstyle s\times s} \end{bmatrix} \underset{s\times 1}{x} + \begin{bmatrix} e_s \\ {\scriptstyle m\times 1} \\ e_h \\ {\scriptstyle s\times 1} \\ e_r \\ {\scriptstyle s\times 1} \end{bmatrix} \tag{1.5}$$

where *Swd* denotes the slant water vapor, *Horiz* is horizontal constraint, and R*ef* is vertical constraint. x is the wet refractivity matrix in all established voxels, H_s, H_h and I_r is the corresponding coefficient matrix; e_s, e_h and e_r is the corresponding error matrix. The observation equation can be further written as error equation, the compact form can be expressed as

$$V = B\hat{x} - l \tag{1.6}$$

The Eq. (1.6) includes three parts, of which the weight can be determined by the approach of variance component estimation. In the data processing, the least square estimation (LSQ) in sequential mode was employed. As for the initial adjustment process, vertical constraints is directly formed based on radiosonde data, the formula can be expressed as

$$B_0^T P_0 B_0 \cdot \hat{x}' = B_0^T P_0 l \tag{1.7}$$

considering that different temporal resolution for radiosonde and GPS technique, the following adjustment process utilised the parameter information from the lasted adjustment result based on LSQ in sequential mode, the wet refractivity parameter in the lasted result was treated as virtual observation in the following adjustment, the mathematic principle can be written as:

$$\begin{bmatrix} I \\ B_i \end{bmatrix}^T \begin{bmatrix} P_{\hat{x}'} & 0 \\ 0 & P_i \end{bmatrix} \begin{bmatrix} I \\ B_i \end{bmatrix} \hat{x}_i'' - \begin{bmatrix} I \\ B_2 \end{bmatrix}^T \begin{bmatrix} P_{\hat{x}'} & 0 \\ 0 & P_i \end{bmatrix} \begin{bmatrix} 0 \\ l_i \end{bmatrix} = 0 \tag{1.8}$$

When the radiosonde data in new temporal arc was collected, the LSQ in sequential iterative process is reinitialized. Otherwise, the wet refractivity between the two temporal arcs can be determined in recurrence mode based on Eq. (1.8).

4 Tomography Results

In order to verify the effectiveness of the approach, the data from Xi'an GPS meteorological service network as the tomography experimental area was collected. There are 7 stations involving in the network, the highest difference of elevation between these

stations is only 140 m, what's worse, they are located at the edge of the study area. In tomography process, the objective area is always divided into some three-dimensional grids to obtain each discrete part of atmosphere information. In our grids partitioning, the horizontal interval is 0.13° (Fig. 1). In vertical direction, the research altitude is set as 10 km according to the general distribution of water vapor in atmosphere, and 1 km is taken as the vertical interval. Ultimately, Xi'an area is divided into 5 × 7 × 10 grids.

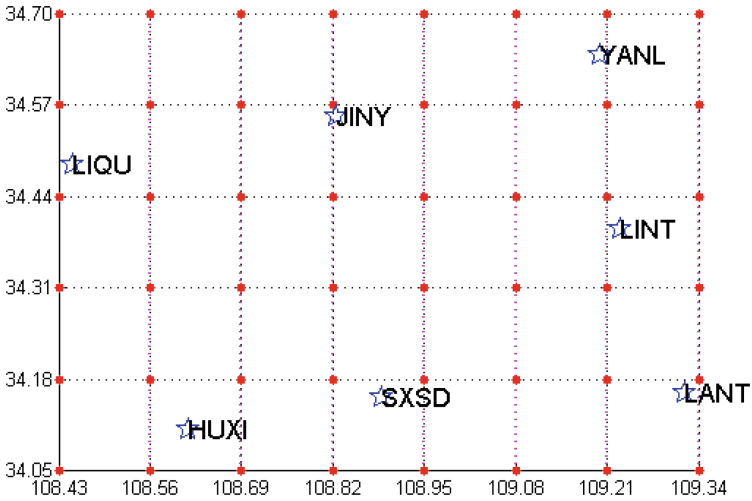


Fig. 1. Plan view of Xi'an tomography voxel model

In the experiment, the observation is collected from 26 August, 2011 for water vapor tomography. Firstly, radiosonde data from Xi'an meteorological station at two time points (UTC 0:00 and 12:00) are solved, and stratified calculation is conducted with the interval of 1 km to obtain the average water vapor at each altitude, as shown in Fig. 2. Secondly, high-precision GPS data processing software GAMIT is utilized to obtain the zenith wet delay (with interval of 30 min) of each station in the grids, the correlation between those stations was decreased by introduced three IGS stations with distances of more than 500 km (incl. BJFS, KUNM, and WUHN). The average water vapor of each altitude layer at UTC 0:00 are treated as vertical constraints, and sequential least square is employed to estimate with interval of 30 min. Except that the first computation session is directly constrained by sounding data, each adjacent session is solved sequentially with parameter adjusted values of the previous session by means of the sequential least square. The vertical constraint is updated till calculating to UTC 12:00. Furthermore, sounding data at UTC 12:00 are regarded as reference values of water vapour in this experiment for verifying the tomography method. The average of wet refractivity in each layer retrieved by tomography technique and sounding data at UTC 12:00 are presented in Fig. 3. The abscissa axes in Figs. 2 and 3 represent atmospheric altitude in kilometer, while the ordinate axes represent wet refractivity with the unit of mm/km. Figure 4 shows the spatial distribution of wet refractivity, the unit of color zone is mm/km.

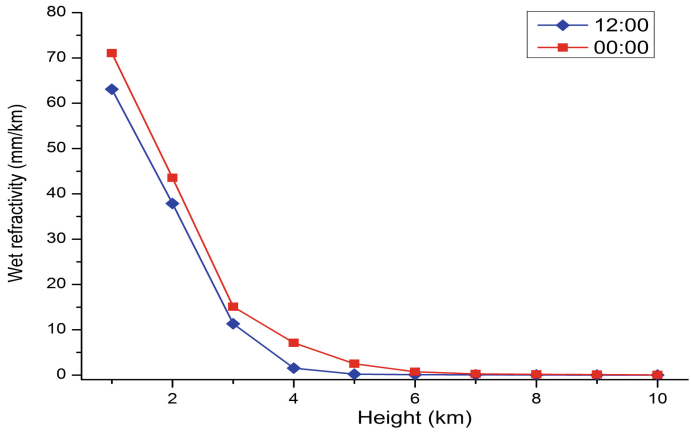


Fig. 2. Comparison of wet refractivity based on sounding technique at UTC0 and UTC12.

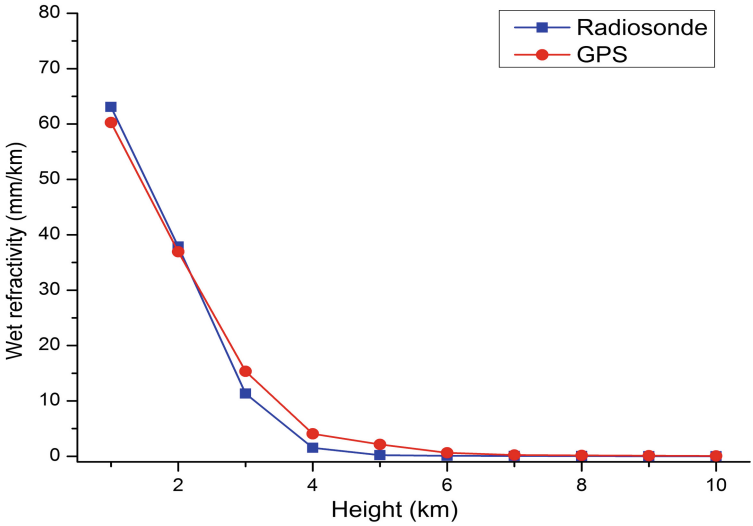


Fig. 3. Comparison of wet refractivity between GPS and radiosonde.

It can be seen from Figs. 2 and 3, the wet refractivity from both radio sounding and GPS tomography consistently provided a good reflection of the three-dimensional spatial distribution of the regional water vapor wet refractivity and its trend, i.e., the bottom layer of the troposphere was the most influenced by the water vapor. The water vapor subsequently decreased with increasing height, while refractivity gradually reduced to approximately zero. From the sounding data of the two sessions listed in Fig. 2, it was noted that on the day of the experiment, the water vapor level changed relatively significantly. Within 12 h, the water vapor varied by nearly 10 mm/km at the bottom layer in both the sessions. This provided a solid basis for the accurate evaluation of the GPS detection of water vapor.

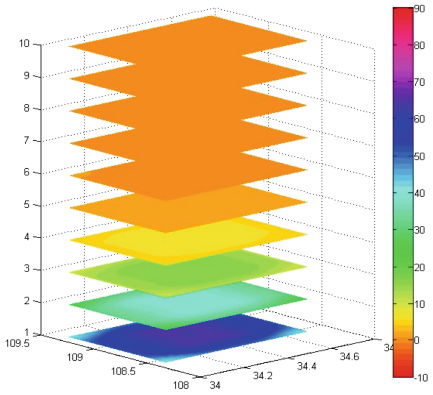


Fig. 4. 3D view of wet refractivity based on fusion sounding data and GPS at UTC 12.

As shown in Fig. 3, using the radio sounding data at 0:00 as the vertical constraints, the wet refractivity of each voxel was calculated by sequential least squares estimation at an interval of 30 min until 12:00. When averaged by height, the wet refractivity of each voxel was highly consistent with the sounding data at 12:00, and the RMS reached 1.8 mm/km. This is because although only the sounding data at 0:00 was directly constrained, during other computation sessions, the adopted adjusted value from the previous session already included the original constraint information. As the computation proceeded, the prior sounding constraint information gradually weakened. Therefore, the adjusted value reflects the water vapor information measured by GPS. On the other hand, the adoption of the measurement data as the initial constraint resulted in the effective utilization of its highly accurate water vapor information in practical applications. Consequently, the problem of insufficient temporal resolution in traditional sounding data was successfully solved when combined with the GPS water vapor tomography results.

Although the sounding data enhanced the accuracy of tomography when used as the vertical constraint, it resulted in inferior solutions at the edge of the voxels. This was mainly caused by the relatively fewer GPS rays penetrating through the edge of the GPS network. Therefore, in future studies, the division rules of the ground-based GPS tomography network should be investigated in depth. In addition, due to the limited experimental conditions, the experiment in this paper only selected data acquired in a single day to validate the algorithm. In future applications, further analysis should be conducted with a larger amount of observation data.

5 Conclusions

In this paper, the latest moment radiosonde data is adopted as vertical constraints, and parameter estimations are performed using the sequential least squares algorithm. This approach effectively compensates for the lack of temporal resolution in radiosonde data. The GPS station can be treated as a radiosonde station, which has the character of

operated uninterruptedly and strongly safeguard the real-time monitoring severe weather. The findings of this work will contribute to the operational application of GPS water vapor tomography.

References

1. Bi Y (2006) The study of global positioning system (GPS) remote sensing of atmospheric water vapor. Beijing University
2. Flores A, Ruffini G, Rius A (2000) 4D tropospheric tomography using GPS slant wet delays. Ann Geophys 18(2):223–234
3. Skone S, Hoyle V (2005) Troposphere modeling in a regional GPS network. Positioning 4 (1&2):230–239
4. Song S, Zhu W, Ding J et al (2006) 3D water-vapor tomography with Shanghai GPS network to improve forecasted moisture field. Chin Sci Bull 51(5):607–614
5. Cao Y, Chen Y, Li B (2006) Determining atmospheric absolute humidity by GPS. Chin J Sci Instrument 27(s2):1132–1135
6. Yu S, Liu L, Liang X (2010) Influence analysis of conditions on GPS water vapor tomography. Acta Geod Cartogr Sin 39(5):491–496
7. Zhang S, Zhang P, Zhang Q et al (2013) Ground-based GPS tomography spatial water vapor distribution with robust variance components estimation. Geomat Inf Sci Wuhan Univ 38 (2):144–147
8. Yao Y, Zhao Q, Luo Y (2017) An approach of imposing virtual signals to sophisticate water vapor tomographic model. Geomat Inf Sci Wuhan Univ 42(11):1658–1664
9. Zhao Q, Yao Y, Luo Y (2017) A method to improve the utilization of observation for water vapor tomography by adding assisted tomographic area. Geomat Inf Sci Wuhan Univ 42 (9):1203–1208
10. Zhang S, Ye S, Wan R et al (2008) Preliminary tomography spatial wet refractivity distribution based on Kalman filter. Geomat Inf Sci Wuhan Univ 33(8):797–799

Denoising Analysis of Different Data Domains Based on EEMD for Landslide Monitoring

Zhe Li(✉), Guanwen Huang, Junqiang Han, Zhezhe Lei, and Peng Hu

College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, Shanxi, China
1058696947@qq.com

Abstract. There are complex non-modeling errors and random noises that are difficult to effectively separate in the landslide monitoring data, and it is difficult to eliminate the influence by difference method. These noises exist in each satellite separately, and their integrated effects are expressed in the coordinate residuals. However, denoising only in a single data domain has the problem of residual noise. In this paper, the EMD and EEMD methods are used to denoise the measured landslide monitoring data in the double-difference observation domain, the coordinate domain and the integrated data domain of the two. The results show that compared with the EMD method, the EEMD method can effectively reduce the occurrence of modal aliasing, improve the automation level of data processing, and is more suitable for complex monitoring environments; using the EEMD method to simultaneously denoise in the double-difference observation domain and the coordinate domain, the root mean square error is slightly improved, and the standard deviation is, compared with the results of not denoising, increased by 12.3%, 46.9%, and 10.1% in the three directions of E, N, and U, respectively, and the denoising is increased by 8.8%, 9.5%, and 8.7%, respectively, compared with the single data domain. Therefore, using the EEMD method to synthesize the denoising methods in different data domain can effectively reduce the effects of random noise and instantaneous strong noise, and more effectively reflect the true landslide monitoring displacement changes.

Keywords: EEMD · Observation domain · Coordinate domain · Noise reduction · Landslide monitoring

1 Introduction

China is one of the countries with serious geological disasters, causing a large number of deaths and property losses due to various landslide disasters every year [1]. With the development of Global Navigation Satellite System (GNSS), the use of GNSS technology for landslide monitoring has become a hot topic in geological engineering applications [2, 3]. The large-scale physical model tests combined with landslides conducted by Wang [4, 5] and others show that GPS-RTK technology can be used for dynamic real-time deformation monitoring of landslide disasters under certain conditions. Liu [6] and others realized the practical application of high-precision landslide disaster deformation monitoring based on GNSS-RTK technology. Although the

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 122–133, 2019.
https://doi.org/10.1007/978-981-13-7751-8_13

method of short baseline relative positioning can well weaken the influence of satellite ephemeris error, atmospheric error, etc., in the landslide monitoring, there is a noise and multipath error closely related to the surrounding environment, which is difficult to be weakened by differential means. For the noise problem existing in the monitoring data, the commonly used noise reduction methods are wavelet denoising [7], empirical mode decomposition (EMD) [8], etc. The EMD method decomposes the signal into multiple components from high frequency to low frequency according to the characteristics of the data itself. It has the advantage of being unaffected by the selection of wavelet basis function and eliminating the strong instantaneous noise. Therefore, it can more effectively reduce the influence of noise in practical applications [9, 10]. In view of the modal aliasing phenomenon that often occurs in EMD method, Wu [11] and others proposed the Ensemble Empirical Mode Decomposition (EEMD) method, which weakens the influences of modal aliasing problem by adding the method of Gaussian white noise. In the selection of denoising data field, there are some non-modeling errors and random noise in the relative positioning observation domain, which affects the reliability of the baseline solution [12–16]. Although the denoising theory in the observation range is more rigorous, it has higher requirements for the filtering denoising method, otherwise it is difficult to achieve better denoising effect [17]. In GNSS data processing, the combined effects of multiple errors will eventually be reflected in the coordinate residuals. Therefore, filter denoising is generally performed in the coordinate domain. However, considering only the combined effect of noise and ignoring the influence of single observation noise is not tight enough, there is noise residue [18–20]. In summary, denoising in a single data domain cannot effectively eliminate the effects of residual noise, and there is no comprehensive research on denoising of different data domains. Therefore, the EMD and EEMD methods are used to perform denoising analysis in the observation range, coordinate domain, and integrated data fields of the measured landslide monitoring data.

2 Algorithm Principle

2.1 Principles of EMD and EEMD Algorithms

The EMD algorithm decomposes the signal into a series of components of different scales according to the characteristics of the signal itself. These components are called Intrinsic Mode Function (IMF). Each IMF must satisfy two conditions: (1) In the entire sequence, the number of extreme points and the number of zero crossings must be equal or at most 1 difference; (2) at any point of the signal, the average of the upper envelope composed of local maximum and the lower envelope formed by local minimum is zero [8, 21]. Therefore, the EMD decomposition of a set of signals can be expressed as:

$$x(t) = \sum_{i=1}^{n} IMF_i + T_i(t) \tag{1}$$

Where IMF_i is the i^{th} modal component; $T_i(t)$ is the residual after decomposition. After the signal is decomposed, a finite number of IMF components and a residual term from high frequency to low frequency are obtained. Usually, the high frequency part is regarded as noise and is rejected, and the remaining part and the residual item are reconstructed to obtain a main signal.

The landslide monitoring observation data usually contains a large amount of signal interference and environmental noise, causing the main signal to be discontinuous, which leads to the modal aliasing phenomenon, that is, the IMF component obtained by the decomposition contains signals of different scales, or signals of the same scale are incorrectly decomposed into other IMF components. At this point, the decomposed IMF component has no longer a corresponding physical meaning, and signal reconstruction will be misjudged, thus affecting the denoising effect.

In response to this phenomenon, Wu et al. proposed an improved EMD method, namely the EEMD method. White noise is added to the signal to be decomposed. According to the uniform distribution feature of the white noise spectrum, when the signal is applied to the white noise background spread over the entire time-frequency space, the signals of different time scales are automatically distributed to the appropriate reference scale to avoid the modal aliasing. Through averaging the results of the multiple decompositions, the final IMF component [11, 22, 23] can be obtained. The algorithm flow is shown in Fig. 1 below:

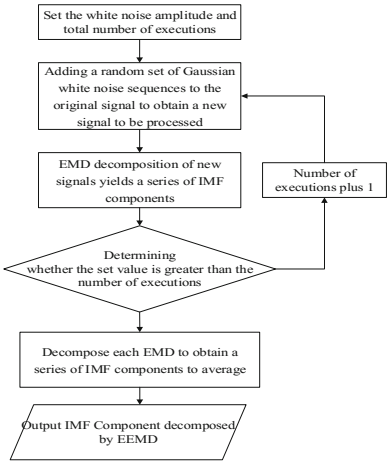


Fig. 1. Flow chart of EEMD algorithm

As a rule of thumb [8], the added Gaussian white noise amplitude is set to 0.2 and the total number of executions is 100.

2.2 Main Signal Extraction

After EEMD decomposition, a series of IMF components from high frequency to low frequency are obtained. The noise rejection and the main signal extraction are achieved by determining the critical point of the noise signal and the main signal. At present, the commonly used methods are correlation coefficient method [22] and modulus cumulative mean method [8, 24]. This paper mainly uses the modulus cumulative mean method to determine the signal-to-noise boundary point. The formula is as follows:

$$\hat{h}_m = mean\left[\sum_{i=1}^{m}(IMF_i(t) - \frac{mean(IMF_i(t))}{std(IMF_i(t))})\right](i \le m) \tag{2}$$

In the formula, $IMF_i(t)$ is the i^{th} scale component. When $\hat{h}_m$ deviates from zero, the component m is considered to be the signal-to-noise demarcation point, and the signal from the m component and the subsequent signal is reconstructed as the main signal.

2.3 Denoising Different Data Fields

Denoising in the observation range means that the observation data is initially processed firstly, and the double-difference observation information and the original coordinate residual sequence of the L1 and L2 frequencies of each satellite are obtained. Then, the double-difference observations of each satellite of these two frequencies are denoised by EMD and EEMD methods respectively. The main residual signal after denoising is extracted and corrected by the epoch, and a new observation equation is reconstructed and then solved to obtain a new coordinate residual. The denoising of different observation domains refers to first denoising in the observation domain, and then using the same denoising method to denoise again the obtained coordinate residual sequence. The overall data processing flow is shown in Fig. 2 below:

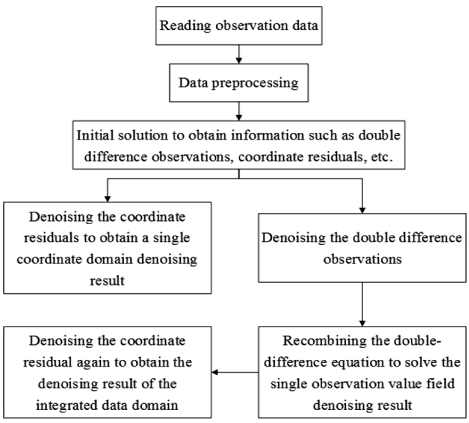


Fig. 2. Flow chart of overall data processing

3 Experiment Analysis

The landslides in the Qinba Mountain Area in Daijiaba Town, Ningqiang County, Hanzhong City were selected as the research objects. The landslide monitoring points were laid in the middle of the slope of the mountain, and there were more earth and stone and vegetation around. There are large environmental noise and transient strong noise in the observation data. The data acquisition equipment is the high-precision UB380 receiver and the HG-GOYH7151 high-precision measuring antenna. The acquisition period is data of one hour from 3:00 to 4:00 on Dec. 8, 2017 (year 342) UTC. The sampling interval is 1 s. In order to verify the advantages of the EEMD method compared to EMD and to analyze the performance of denoising in different data fields, the above two methods are used to denoise in the observation range, the coordinate domain, and the comprehensive observation range and coordinate domain. The residual results are shown in Fig. 3 (take G29 as an example).

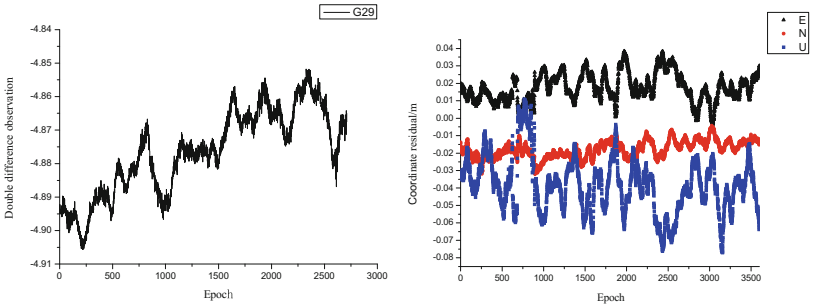


Fig. 3. Residual of G29 satellite

It can be seen from Fig. 3 that after two differentials, the influence of most common errors is weakened, but there are still some random noise and instantaneous strong noise in the double difference observation and the coordinate residual.

3.1 Denoising Analysis

In order to eliminate the influence of noise as much as possible to obtain the true and accurate point changes, it is necessary to perform denoising in data processing. Due to the space limitations, only the data processing of denoising in different data fields using EMD and EEMD methods respectively is shown. Firstly, the EMD and EEMD methods are used to decompose the double-difference observation sequence to obtain a series of IMF components with different scales. The decomposition results of the G29 satellite are shown in Figs. 4 and 5:

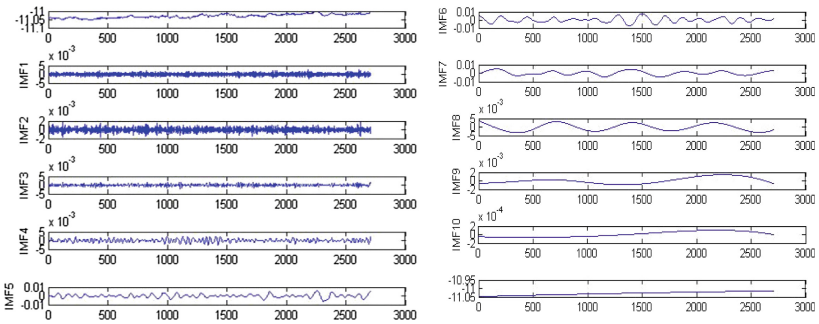


Fig. 4. Decomposition results of EMD algorithm

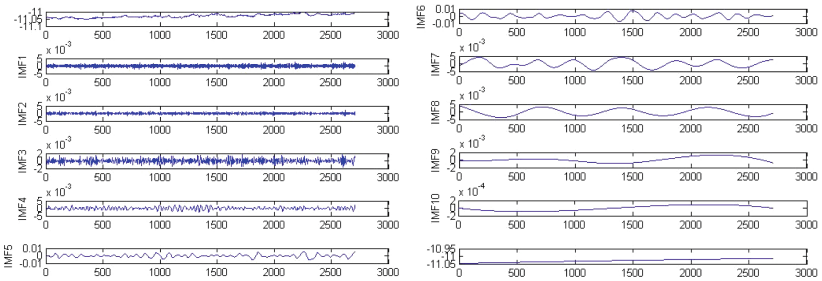


Fig. 5. Decomposition results of EEMD algorithm

It can be seen from the figure that the first half of the decomposition of the two methods is basically the same, and the variation trend of some IMF components in the latter half is slightly different, and different IMF components have a certain influence on the calculation of the subsequent signal-to-noise boundary point and the main signal extraction. The relationship between the cumulative mean and the scale component of the normalized modulus of the two methods is shown in Fig. 6:

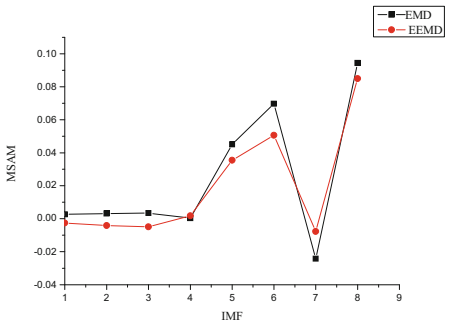


Fig. 6. The relationship between the modulus cumulative mean and the decomposition scale of the observation domain

It can be seen from Fig. 6 that the relationship diagrams obtained by the two methods of EMD and EEMD have the same trend, and both relationship polylines appear to deviate from zero at IMF5, that is, the signal-to-noise critical point is IMF5. The part of IMF1 ~ IMF4 is regarded as high frequency noise and is rejected, and the remaining components and residuals are reconstructed to obtain the main signal. The reconstructed main signal is shown in Fig. 7:

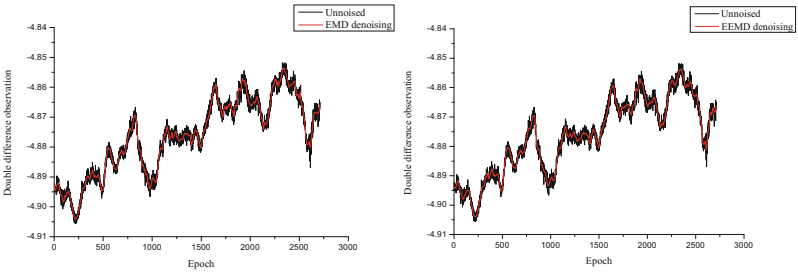


Fig. 7. Denoising result graph of EMD and EEMD algorithm in observation domain

Denoise each satellite in the L1 and L2 frequencies separately, and store the result in a file. When it is calculated again, replace it with the double difference observation of the corresponding epoch to form a new observation equation, and then resolve it to obtain the coordinate residual. Denoising the coordinate residual again, and the relationship between the modulus cumulative mean and the IMF component, and the denoising are compared as shown in Figs. 8 and 9 (only the N direction is an example):

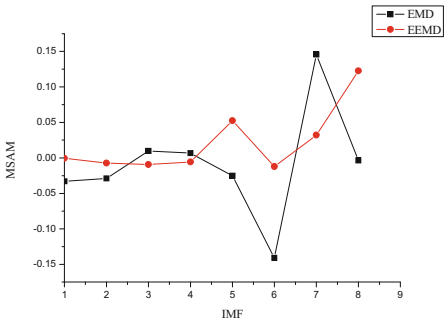


Fig. 8. The relationship between the modulus cumulative mean and the decomposition scale of the coordinate domain

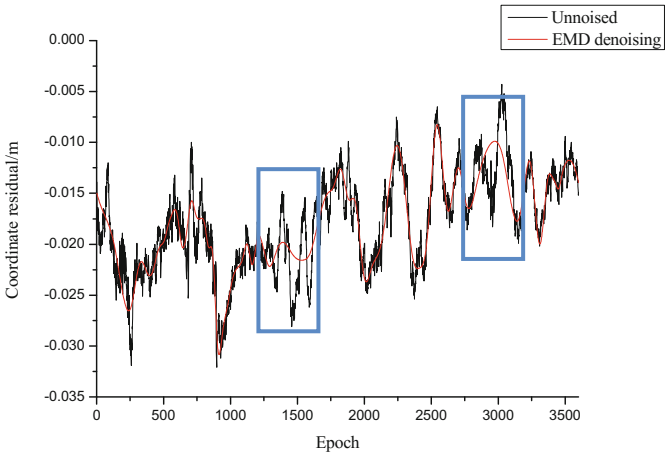


Fig. 9. Denoising result graph of EMD algorithm in coordinate domain

It can be seen from the figure that the signal-to-noise demarcation point can be judged by the graph of modulus cumulative mean and the scale relationship. However, after the signal reconstruction according to the demarcation point, there appears a part of the denoising result of the EMD method that does not match the original trend (at the blue square in Fig. 9), which shows that modal confusion occurs at this point. And the coordinate change after reconstruction of the EEMD method is more consistent with the original sequence change trend (Fig. 10).

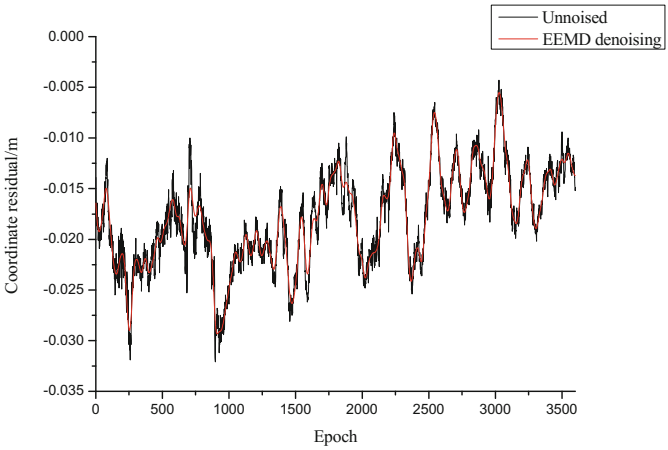


Fig. 10. Denoising result graph of EEMD algorithm in coordinate domain

Modal aliasing not only represents the anomaly of the result after denoising, but sometimes occurs when the signal-to-noise demarcation point is judged, as shown by the G18 satellite in Fig. 11.

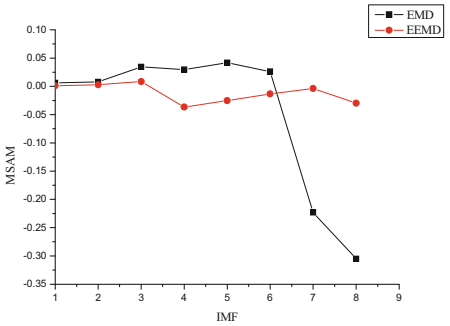


Fig. 11. The relationship between the modulus cumulative mean and the decomposition scale of the observation domain (G18)

It can be seen from the figure that the signal-to-noise demarcation point decomposed by the EEMD method is obviously displayed at the IMF4, while the result obtained by the EMD method is expressed at the IMF7. When the signal is reconstructed by the above-mentioned signal-noise demarcation point, the results achieved by EEMD method are more consistent with the original sequence, and the results by EMD method are completely wrong. When the above phenomenon occurs, the signal-to-noise demarcation point must be manually determined, which greatly reduces the data processing efficiency and the reliability of the result.

The above results can be summarized as follows: (1) Both denoising methods can effectively reduce the influence of noise, especially instantaneous strong noise; (2) For the poor quality data, the EMD method sometimes has the problem of unclear relationship between the cumulative mean value of the standard modulus and the decomposition scale, and it is impossible to directly and effectively judge the signal-to-noise demarcation point.; (3) After obtaining the significant signal-to-noise demarcation point, the main signal generated by the EEMD method has a one-to-one correspondence with the original signal in terms of change trend, peak and trough, while adopting the EMD method sometimes appears different from the original signal. This is mainly because the landslide monitoring station is affected by the surrounding environment, and there are many complex noises that are difficult to separate in the observation data. The EMD cannot identify them when scale decomposition is performed. The occurrence of modal aliasing causes the signal-to-noise boundary point to be unclear and the decomposition scale is confused, which greatly reduces the automation degree and denoising reliability of the denoising data processing. Therefore, the EEMD method denoising performance is better than EMD in the measured data processing.

3.2 Precision Analysis

The root mean square error (RMS) after denoising in the double-difference observation range, the coordinate domain, and the integrated two data fields is shown in Table 1. The standard deviation (STD) is shown in Table 2.

Table 1. RMS results for different scenarios

			E(m)	N(m)	U(m)
RMS	Original precision	Unnoised	0.0194	0.0185	0.0417
	EMD	Observation domain	0.0194	0.0185	0.0416
		Coordinate domain	0.0193	0.0184	0.0416
		Integrated data domain	0.0192	0.0184	0.0417
	EEMD	Unnoised	0.0193	0.0184	0.0416
		Observation domain	0.0193	0.0185	0.0415
		Coordinate domain	0.0189	0.0183	0.0409

Table 2. STD results for different scenarios

			E(m)	N(m)	U(m)
STD	Original precision	Unnoised	0.0081	0.0081	0.0159
	EMD	Observation domain	0.0081	0.0048	0.0157
		Coordinate domain	0.0078	0.0047	0.0157
		Integrated data domain	0.0076	0.0044	0.0146
	EEMD	Unnoised	0.0080	0.0048	0.0157
		Observation domain	0.0079	0.0047	0.0156
		Integrated data domain	0.0071	0.0043	0.0143

It can be seen from Table 1 that the root mean square errors of the six schemes after denoising are very close, and the results of the comprehensive data domain using the EEMD method are slightly dominant, and the E, N, and U directions are increased by 0.5 mm and 0.2 mm, and 0.8 mm respectively. Table 2 reflects the degree of dispersion of the post-coordinate residuals of denoising. It can be clearly seen that the six schemes can greatly improve the stability of the results and are more consistent with the characteristics of smooth change of landslide monitoring points. The accuracy of combining different data fields is obviously higher than that of single data domain. The denoising results of the integrated data domain using EEMD method are improved by 1.0 mm, 3.8 mm and 1.6 mm respectively in the three directions of E, N and U, and the increasing range was 12.3%, 46.9%, and 10.1%, which was 8.8%, 9.5%, and 8.7% higher than the single data field.

4 Summary

In this paper, a denoising method based on the EEMD method to synthesize two data fields is proposed for the complex noise that is difficult to eliminate by modeling or differential methods in landslide monitoring. Through the measured landslide observation data, the advantages of EEMD method in reducing modal aliasing phenomenon compared with EMD are verified, and the denoising efficiency and reliability are improved. Compared with the results after denoising, it can be found that the RMS accuracy is slightly improved, and the STD accuracy is significantly improved. Therefore, the EEMD method is used to synthesize the denoising methods of different data fields, which can effectively reduce the influence of random noise and instantaneous strong noise, and can effectively reflect the real landslide monitoring displacement changes, and provide reliable coordinate change information for the subsequent establishment of multipath error models.

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). We also thank the IGS and iGMAS authorities for providing the data for this study.

References

1. Yin Y (2004) Initial study on the hazard-relief strategy of geological hazard in China. Chin J Geol Hazard Control 15(2):1–7
2. Zhou P, Guo J, Li A et al (2008) Application of No. 1 Beidou navigation satellite communication technology in the landslide real-time monitoring. GNSS World China 33 (5):20–23
3. Tang Y (2016) Application of GNSS technology in emergency deformation monitoring of landslide. Gold 37(5):70–75
4. Wang L (2015) A study on key technology high precision GPS monitoring for geological hazard. Acta Geod Cartogr Sin 44(07):826
5. Wang L, Zhang Q, Li X et al (2011) Dynamic and real time deformation monitoring of landslide with GPS-RTK technology. J Eng Geol 19(2):193–198
6. Liu Y (2014) The research on high-precision deformation monitoring technology of BD-S/GPS combined system. Liaoning Technical University
7. Xiong YL, Ding XL, Dai WJ, Chen W, Huang DF (2004) Mitigation of multipath effects based on GPS phase frequency feature analysis for deformation monitoring applications. In: ION GPS 2004, pp 268–276
8. Huang NE, Shen Z, Long SR et al (1971) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc A 1998(454):903–995
9. Shi C, Luo Q (2003) Hilbert-huang transform and wavelet analysis of time history signal. Acta Seismol Sin 25(4):398–405
10. Dai W, Ding X, Zhu J et al (2006) EMD filter method and its application in GPS multipath. Acta Geod Cartogr Sin 35(4):321–327

11. Wu Z, Huang NE (2004) A study of the characteristics of white noise using the empirical mode decomposition method. Proc R Soc Lond 460A:1597–1611
12. Wang J, Gao J, Wang J (2008) GPS baseline solution based on empirical mode decomposition. Acta Geod Cartogr Sin 37(1):10–14
13. Du Y, Huang G, Zhang Q et al (2018) A novel predictive algorithm for double difference observations of obstructed BeiDou geostationary earth orbit (GEO) satellites. Adv Space Res 63:1554–1565
14. Han J, Huang G, Zhang Q et al (2018) A new azimuth-dependent elevation weight (ADEW) model for real-time deformation monitoring in complex environment by multi-GNSS. Sensors 18(8):2473–2488
15. Qiu X, Liu G (2016) Influence analysis of BDS/GPS observation noise on ambiguity resolution. J Navig Position 4(4):69–76
16. Yi Q, Liu X, Liu W (2017) Multipath mitigation method by combined EMD and wavelet in GPS/BDS deformation monitoring. J Geod Geodyn 37(5):462–466
17. Chen D, Yw S, Liu Y, Liu Z (2014) Applied analysis of GPS multipath errors based on observation domain. Geomat Inf Sci Wuhan Univ 39(2):147–151
18. Zhang S, He Y, Li Z et al (2017) EMD for noise reduction of GPS time series. J Geod Geodyn 37(12):1248–1252
19. Luo C, Feng W, Ding Y et al (2018) GPS multipath correction model research based on CEEMD. J Geod Geodyn 38(4):381–385
20. Liu C, Wang J, Hu H, Gao J (2010) Research on real-time correcting model of multipath in GPS dynamic deformation monitoring. Geomat Inf Sci Wuhan Univ 35(4):481–485
21. Jia R, Zhao T, Sun H et al (2015) Micro-seismic signal denoising method based on empirical mode decomposition and independent analysis. Chin J Geophys 58(3):1013–1023
22. Xue Z, Zhou R, Li G et al (2013) Modeling multipath effect with a noise assisted empirical mode decomposition. Sci Surv Mapp 38(4):55–58
23. Wu C, Wang P, Fang X et al (2017) The time-frequency characteristics of strong earthquake records in Shandong Province using the EEMD method analysis. J Geod Geodyn 37(12):67–70
24. Luo H (2017) Research and analysis of CEEMD in eliminating GPS multipath

Applicability Analysis of Klobuchar Model Based on Short-Term Prediction in Different Latitudes

Yunzhen Yang[1], Lilong Liu[1(✉)], Liangke Huang[1(✉)], Jun Chen[2], Wei Zhou[1], and Qingtong Wan[1]

[1] College of Geomatic Engineering and Geoinformatics, Guilin University of Technology, Guilin, China
hn_liulilong@163.com, lkhuang666@163.com
[2] School of Geodesy and Geomatics, Wuhan University, Wuhan, China

Abstract. The classic Klobuchar model is widely used in navigation and positioning. However, the ionospheric delay correction accuracy is difficult to meet the high-precision positioning. In this paper, the observation data provided by the International Global Navigation Satellite Systems Service Center are used to calculate the Vertical Total Electron Content (VTEC) with the Klobuchar model and the Dual-frequency correction model. Then the Auto-regressive Integrated Moving Average (ARIMA) model is used to forecast the error of the 9th day and 10th day between the Klobuchar model and the dual-frequency correction model base on the error of the former eight days. The forecast results are used to improve the model. Finally, the accuracy of the improved model is to be evaluated in different environment and different latitudes. The results show that the average relative accuracy of the improved Klobuchar model is 71.66% and 69.69% in the ionospheric active period and ionospheric quiet period, respectively. The improved Klobuchar model is more consistent with the dual-frequency correction model, and can better to reflect the temporal evolution characteristics of the ionosphere.

Keywords: Ionosphere delay correction · Klobuchar model · ARIMA model · Precision evaluation

1 Introduction

The ionospheric activity is affected by solar radiation, geomagnetic field and others. When the satellite signals passes through the ionosphere, the direction and speed of the signals will change, which result in ionospheric delay errors in GNSS measurements. This error can be up to more than 10 m in the zenith direction and even 50 m when the elevation angle is small [1]. Therefore, it is important to weaken or eliminate the ionospheric delay error reasonably and accurately.

At present, the ionospheric correction models commonly used in GNSS positioning and navigation applications, include the IRI model [2], the NeQuick model [3] and the Klobuchar model [4]. Due to the simplicity and convenience of model, Klobuchar model is widely used, but it only corrects for 50%–60% of the ionospheric delay in

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 134–142, 2019.
https://doi.org/10.1007/978-981-13-7751-8_14

high-precision GNSS positioning [5]. Therefore, lots of studies have been carried out. The eight parameters have been re-estimated and released by the center for Orbit Determination in Europe using the global GPS data in 1997 [6]; Yuan et al. [7] refined the parameters with the data from hundreds of IGS and Crust Movement Observation Network of Chinese, then the result show that the accuracy of the model have 15% improvement in China. However, the parameter optimization only focuses on improving amplitude and period. In order to take into account the changes with the initial phase and the nighttime delay, domestic and foreign researchers have refined the model by adding parameters to extend the structure [8–11]. Wang et al. [8] proposed a 10-parameters model by considering the nighttime terms as a linear function of geomagnetic latitude, and the accuracy of the model is improved. Based on the least squares surface fitting model, Chenghui et al. [9] used the measured data to improve the initial phase, amplitude and nighttime delay of the Klobuchar model, and established the ionospheric delay regional model. Zhang et al. [11] considered the influence of the initial phase and the nighttime terms, combined with the least squares method to extend the eight parameters to fourteen parameters with GPS dual-frequency correction data in China region, and the model has better correction effect at the middle latitude region. Chen et al. [12, 13] refined the Klobuchar model using the holt-winters exponential smoothing model with China regional peripheral observation station data provided by the IGS Center in 2011, and the accuracy of the refined model has been greatly improved. However, the two methods of parameter optimization and parameter addition can better correct the ionospheric delay, but they are limited in their use in the GNSS positioning and navigation applications because of their low correction accuracy and complex methods.

With the development of GNSS technology, single-frequency users are eager to improve their positioning accuracy. The time series of ARIMA model [14] has the advantages of less sample data, general theory, simple calculation and high-precision of short-term forecasting, and has carried out some applied research on the short-term prediction of the ionosphere. Based on the above questions, this study treats the bias between the Klobuchar model and the Dual-frequency correction model as a times series. The ARIMA model is used to predict the bias. The bias is used to improve the Klobuchar model. Then, we will evaluate the applicability in different environments and different latitudes.

2 The Establishment of Model

2.1 Klobuchar Model

The Klobuchar model is simplified by the Bent ionospheric empirical model, which reflects the temporal evolution characteristics of the ionospheric delay in a day [1]. The mathematical formula of the Klobuchar model is given by Eq. (1):

$$T = \begin{cases} A_1 + A_2 \cos[2\pi(t - A_3)/A_4], |t - A_3| \leq A_4/4 \\ A_1, |t - A_3| \geq A_4/4 \end{cases} \quad (1)$$

Where, T is the vertical ionospheric time delay, t is the local time, when the receiver is located at the intersection of both the satellite connection and the ionosphere (in s), A_1 is the nighttime constant term (in s), A_3 is the initial phase, which indicates the pole of the cosine function (in s), A_2 and A_4 are the eight broadcast coefficients in GPS navigation message.

2.2 ARIMA Model

The ARIMA(p, d, q) model is called the auto-regressive integrated moving average model. It was proposed by Box and Jenkins in the 1970s. The basic idea is to convert the non-stationary time series to stationary time series, and then establish the dependent variable by regression with their hysteresis value and the present and hysteresis value of the random error term [14]. The ARIMA model can be expressed as shown in Eq. (2):

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, \forall_S < t \end{cases} \tag{2}$$

Where, ∇^d is the difference operational, $\Phi(B)$ is the auto-regressive coefficient polynomial of the stationary reversible ARMA (p, q) model, $\Theta(B)$ is the moving smoothing coefficient polynomial of the stationary reversible ARMA (p, q) model, x_t is the time series, ε_t is the Gaussian white noise time series, s and t denote the different moments of the time series, d is the difference order, p and q are the auto-regressive coefficient and the moving average coefficient.

2.3 The Improvement of Klobuchar Model Based on ARIMA Model

In view of the shortcomings of the Klobuchar model mentioned in the previous section, we take advantages of the ARIMA model, and use the ionospheric measured data provided by the IGS Center to calculate the ionospheric VTEC value with the Klobuchar model and the dual-frequency correction model, respectively. Then, the bias of the ionospheric VTEC values can be calculated between the Klobuchar model and Dual-frequency correction model in the former 8 days of each epoch. Further, The ARIMA model is used to predicate the bias for the next two days. Finally, the predicted bias values can be used to improve the Klobuchar model in the 9th day and the 10th day. The method of improvement is as shown in Eq. (3):

$$T = \begin{cases} A_1 + A_2 \cos[2\pi(t - A_3)/A_4] + \Delta_d, |t - A_3| \leq A_4/4 \\ A_1 + \Delta_d, |t - A_3| \geq A_4/4 \end{cases} \tag{3}$$

Where, d is 1 or 2, Δ_1 and Δ_2 are the ionospheric VTEC bias on the 9th day and the 10th day from the ARIMA model prediction, the other parameters are defined as described in Sect. 2.1.

3 Result of Analysis

3.1 The Index of Accuracy Evaluation

Compared with other models, the accuracy of the ionospheric dual-frequency correction model is above 95% [15], it is very close to the true value. Therefore, it can be treated as a reference value, and two statistical indices, including the Precision (PRE) and Root Mean Square Errors (RMSE) are introduced for validation. The statistical indices are defined in Eqs. (4)–(5).

$$PRE = \frac{1}{e-b}\sum_{i=b}^{e}\left[1 - \frac{I_m^i - I_R^i}{I_R^i}\right] \times 100\% \tag{4}$$

$$RMSE = \sqrt{\frac{1}{e-b}\sum_{i=b}^{e}\left(I_m^i - I_R^i\right)^2} \tag{5}$$

Where, m = 1 is the Klobuchar model, m = 2 is the improved Klobuchar model; b and e are the initial epoch and the final epoch of the time period, respectively; I_m^i is the VTEC value of the Klobuchar model or the improved Klobuchar model at the i-th epoch; I_R^i is the VTEC value of the dual-frequency observation model at the i-th epoch.

3.2 The Source of Data

Considering that the 11-a cycles change of sunspot and the geomagnetic activity index, this study chooses six IGS stations including niril, yakt, chan, xian, pimo and iisc stations located at the high latitudes, middle latitudes and low latitudes of the northern hemisphere, respectively. The data of ionospheric quiet period are from the day of the year 016 to 025 in 2011, while the ionospheric active period are from the day of the year 286 to 295 in 2011. The sampling rate of the data is 30 s, and the satellite elevation cut-off angle is 15°.

3.3 The Analysis of Experiment

Figure 1 is a comparison of the basic Klobuchar model (BKM), the dual-frequency correction model (DCM as a reference value), and the improved Klobuchar model (IKM) at different station in the ionospheric active period from October 21 to 22 in 2011. While Fig. 2 is a comparison of BKM, DCM and IKM at different station in the ionospheric quiet period from January 24 to 25 in 2011. In Figs. 1 and 2, the horizontal axis represents Universal Time (UTC) and the unit is h, the vertical axis indicates the VTEC and the unit is TECU.

For Figs. 1 and 2, we can obtained that the variation trend of ionospheric VTEC calculated with BKM, IKM and DCM is difference at different station, and the difference of magnitude is large. Whether the first day or second day, IKM fits the data of DCM well at each station and can better reflect the temporal variation of the ionosphere in both the ionospheric active period and ionospheric quiet period compared with

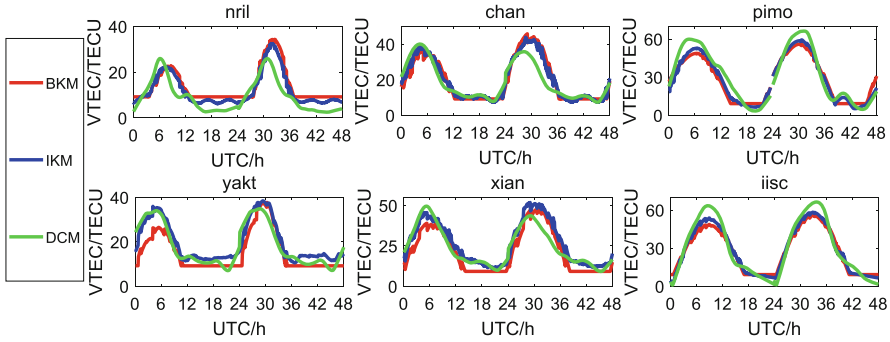


Fig. 1. Comparison of VTEC with BKM, IKM and DCM at different station in ionospheric active period

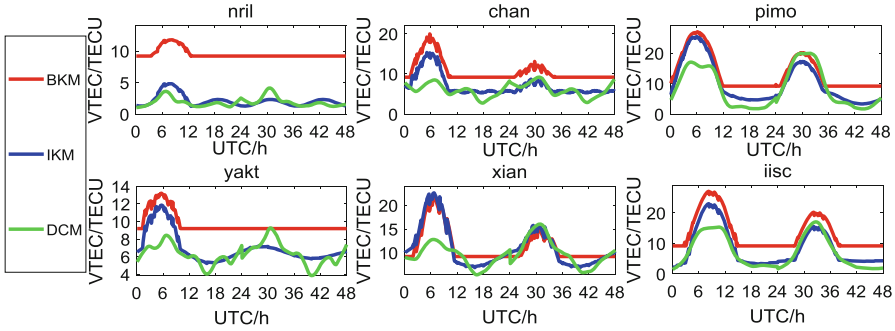


Fig. 2. Comparison of VTEC with BKM, IKM and DCM at different station in ionospheric quiet period

BKM, which show that IKM and DCM have the best conformity at the middle latitudes region, followed by the high latitudes and the lower latitudes. At the same time, the improved effect of the IKM in the ionospheric active period is better than that in the ionospheric quiet period. In addition, Fig. 2 presents that the ionospheric VTEC values at nril and yakt station at the high latitudes are fixed with 9.2TECU on the second day, this is due to the fact that when the amplitude A of cosine function is calculated using the coefficient transmitted in GPS navigation message, sometimes the value is less than zero, that is a negative amplitude, which leads to the ionospheric VTEC fixed with a constant value.

The accuracy of each time period with GKM and BKM relative to DCM in ionospheric action period for the first day and second day is shown in Table 1, the each period of time is 4 h, and there is 6 periods of time in a day. It can be found that the accuracy of the improved Klobuchar model has been improved in high latitudes, middle latitudes and low latitudes. In terms of daytime, the accuracy of IKM improves by 12.66% at the high latitudes, 6.74% at the middle latitudes and 18.25% at the low latitudes on the first day, respectively; while it improves by 10.46% at the high

latitudes, 1.76% at the middle latitudes and 8.10% at the low latitudes on the second day, respectively. For the nighttime, the precision of IKM is better than BKM, which can increase the ionospheric precision by 22.04% at the high latitudes, 13.18% at the middle latitudes and 20.60% at the low latitudes on the first day, respectively, while it increases by 20.70% at the high latitudes, 7.21% at the middle latitudes and 21.87% at the low latitudes on the second day, respectively. Which show that the level with ionospheric correction accuracy of low latitudes is better than that of high and middle latitudes. IKM shows the best precision at the middle latitudes and the worst precision at the high latitudes for the same period of time, which due to that the Klobuchar model based on data from the middle latitudes region, and it is consistent with the strongest applicability in middle latitudes areas. In the same environment and the same station, there is an obvious difference in the correction effect on the second day compared with the first day. in addition, compared with the daytime, the correction effect of the nighttime is better.

Table 1. The statistics of relative accuracy for each time period with IKM and BKM in Ionospheric action period on the first day and second day (%)

Station	Model	[0, 4]	[5, 8]	[9, 12]	[13, 16]	[17, 20]	[21, 24]
		1d/2d	1d/2d	1d/2d	1d/2d	1d/2d	1d/2d
nril	IKM	71.14/75.97	82.32/86.81	46.68/32.75	68.77/54.32	−40.7/−35.09	−0.59/−47.98
	BKM	54.71/63.74	79.22/84.07	28.71/5.51	39.91/18.16	−105.7/−101	−49.9/−103.9
yakt	IKM	87.94/88.78	93.09/83.78	81.26/81.47	88.16/75.87	83.58/71.67	57.77/63.97
	BKM	55.39/69.90	82.43/86.91	85.98/76.66	77.91/82.51	85.90/85.61	76.66/77.07
chan	IKM	85.70/95.53	93.95/68.52	78.83/47.45	89.82/85.47	93.77/93.75	89.29/91.26
	BKM	77.29/89.98	90.32/63.17	72.03/36.10	81.31/87.28	84.50/88.27	83.77/85.60
xian	IKM	89.39/93.17	92.50/76.92	80.26/53.53	85.26/86.64	92.84/89.54	91.43/79.17
	BKM	67.56/80.64	82.12/86.80	90.87/67.89	71.58/71.92	6.51/64.04	78.62/85.47
pimo	IKM	80.41/90.86	87.57/89.02	91.53/90.33	65.65/73.47	58.47/77.76	53.86/66.15
	BKM	83.75/89.28	81.12/84.47	82.78/85.89	52.34/61.55	24.02/63.75	27.62/23.34
iisc	IKM	69.61/28.41	89.56/93.00	85.12/86.85	77.78/93.43	82.69/76.97	13.64/−10.62
	BKM	−12.5/−1.58	81.46/88.67	77.62/83.15	78.46/87.29	78.83/71.99	−32.8/−61.95

The accuracy of each time period with GKM and BKM relative to DCM in ionospheric quiet period for the first day and second day is summarized in Table 2. It can be concluded that IKM can improve by 214.14% at the high latitudes, 20.8% at the middle latitudes and 50.26% at the low latitudes for the daytime of the first day, while it improves by 278.29% at the high latitudes, 40.30% at the middle latitudes and 162.82% at the low latitudes for the nighttime of the first day, respectively. In terms of the daytime with the second day, the rate of precision of IKM increased by150.06%, 19.69% and 29.80% relative to BKM at the high latitudes, middle latitudes and low latitudes, respectively. Concerning the nighttime with the second day, it increased by 284.52% at the high latitudes, 37.80% at the middle latitudes and 181.69% at the low latitudes. These indicate that IKM presents the best improved accuracy effect on high latitudes compared with low latitudes and middle latitudes. In addition, compared with

the reference value, a negative correction rate value occurs of BKM at high latitudes, because the ionospheric VTEC value is small and the Klobuchar model is fixed to a constant value on the day of the station. It can also indicate that the constant ionospheric delay value at night cannot accurately reflect the temporal variation at night.

Table 2. The Statistics of relative accuracy with IKM and BKM for each time period in Ionospheric quiet period on the first day and second day (%)

Station	Model	[0, 4]	[5, 8]	[9, 12]	[13, 16]	[17, 20]	[21, 24]
		1d/2d	1d/2d	1d/2d	1d/2d	1d/2d	1d/2d
nril	IKM	90.18/77.63	68.74/67.11	24.65/86.30	83.14/85.34	66.82/66.57	58.77/66.12
	BKM	−496./−286.1	−174.0/−84.7	−341.9/−248.	−512./−475.4	−360/−358.5	−405.6/−483
yakt	IKM	68.54/96.38	55.04/83.78	85.08/94.63	89.57/83.35	88.17/79.03	93.32/89.45
	BKM	38.15/62.79	33.71/90.75	48.12/65.77	48.12/23.69	19.68/11.11	43.91/42.79
chan	IKM	51.03/86.27	21.92/91.46	75.19/89.36	91.20/91.67	35.68/41.82	82.97/82.86
	BKM	8.11/60.46	−24.00/63.68	11.00/25.51	44.43/41.33	−71.4/−61.12	35.48/42.10
xian	IKM	64.11/87.91	38.26/95.46	61.45/92.86	84.60/85.37	78.04/78.67	98.06/97.56
	BKM	80.23/90.79	39.52/88.98	63. 05/93.59	77.24/83.85	49.86/52.23	93.16/92.42
pimo	IKM	59.87/89.62	49.35/84.43	85.65/72.95	−15.84/63.32	−38.40/38.66	11.47/57.29
	BKM	35.22/67.09	38.71/97.84	64.34/84.49	−135./−17.92	−258./−207.7	−143./−138.5
iisc	IKM	86.89/51.37	70.35/86.30	61.07/91.07	88.00/90.89	75.04/20.74	65.26/−34.76
	BKM	−96.0/−96.87	36.71/74.31	32.68/68.86	18.19/7.38	−125./−190.1	−149./−307.2

Combining Tables 1 with 2, we can find out that the relative accuracy of the IKM in active ionosphere is slightly better than that in quiet ionosphere, the corrective rate of the IKM at night is more obvious than that in daytime, and the improvement effect on the first day is better than that on the second day.

The RMSE statistics both BKM and IKM in different periods relative to DCM are shown in Fig. 3. The horizontal axis represents the time period and the unit of measurement is 4 h, which has 6 periods in a day. And the vertical reflection RMSE (Root Mean Square Error) and the unit is TECU. It can be seen that during the active period of the ionosphere, compared to the BKM, IKM can reduce the RMSE by 1.84 TECU at the high latitudes, 2.21TECU at the middle latitudes and 2.28 TECU at the low latitudes on the first day, while it can reduce the RMSE by 1.16 TECU at the high latitudes, 0.72 TECU at the middle latitudes and 2.12 TECU at the low latitudes on the second day, Among them, the RMSE is less than 3 TECU by 38.89% and 33.33% both on the first day and the second day, respectively. For the ionospheric quiet period, the RMSE of IKM reduce by 4.88TECU, 1.53 TECU and 3.86TECU relative to BKM at the high latitudes, middle latitudes and low latitudes on the first day, respectively, while the RMSE on the second day reduce by 4.33TECU, 1.51 TECU and 2.77 TECU in the high latitudes, middle latitudes and low latitudes on the second day, respectively. Meanwhile, the RMSE value in each period is less than 3TECU accounting for 69.44% and 94.44%, which is 1.79 times and 2.83 times of the ionospheric active period, respectively. Compered with the nighttime, the RMSE value is relatively large in the daytime, which may be influenced by solar activity, the TEC value of ionospheric

model varies greatly in daytime, which makes the residual error is larger. In addition, we analyzed the correlation between IKM and DCM, The results show that the correlation coefficient of the two models is 0.819, which is 0.061 more than the correlation between the basic Klobuchar model and the dual-frequency correction model. Therefore, it can be seen that the improved Klobuchar model is feasible.

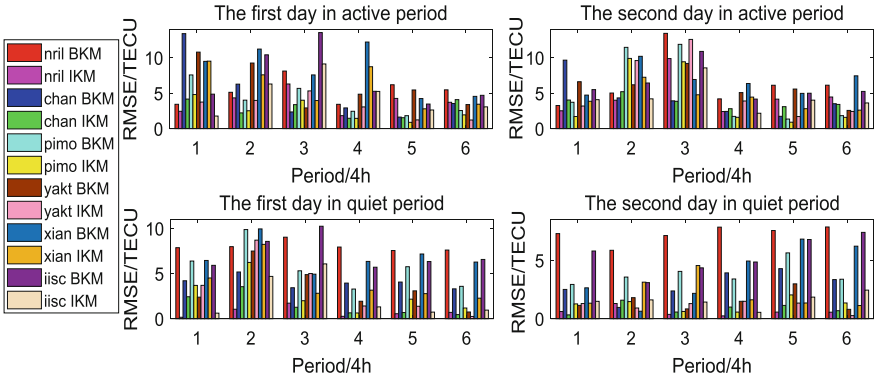


Fig. 3. RMSE of BKM and IKM at different station in different environments

4 Conclusion

In this paper, the observation data provided by the IGS Service Center are used to calculate the VTEC with the Klobuchar model and the Dual-frequency correction model. Then the ARIMA model is used to forecast the error of the 9th day and 10th day between the Klobuchar model and the dual-frequency correction model base on the error of the former eight days. The forecast results are used to improve the model, and the accuracy of the improved Klobuchar model is analyzed in different environments and different latitudes. The results show as follows:

1. No matter what in the ionospheric active period or the ionospheric quiet period, the improved Klobuchar model has better applicability, and the correlation coefficient both the improved Klobuchar model and the dual-frequency correction model is 0.8185, which can better reflect the temporal evolution characteristics of the ionosphere.
2. Both the relative accuracy and the absolute accuracy of the improved Klobuchar model are improved. The improvement effect at the high latitudes is better than the middle latitudes and low latitudes, and the improvement effect of the nighttime is better than the daytime.

Due to the limited data used in this paper, we will carry out the application of the improved Klobuchar model in other areas in the next step and discuss the accuracy of this research.

Acknowledgements. This works 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 (2017GXNSFBA 198139; 2017GXNSFDA198016); the Guangxi Key Laboratory of Spatial Information and Geomatics (14-045-2 4-10; 16-380-25-01). The authors would like to thank the International GNSS Service Center (IGS) for providing the GPS observation data.

References

1. Luo W, Liu Z, Li M (2014) A preliminary evaluation of the performance of multiple ionospheric models in low-and mid-latitude regions of China in 2010–2011. GPS Solutions 18(2):297–308
2. Bilitza D (2001) International reference ionosphere 2000. Radio Sci. 36(2):261–275
3. European Commission (2016) European GNSS (Galileo) Open Service–Ionospheric Correction Algorithm for Galileo Single Frequency Users, Issue 1.2. European Commission, Brussels, Belgium
4. Zhang HP, Ping JS, Zhu WY et al (2006) Brief review of the ionospheric delay models. Prog Astron 24(1):16–26
5. Klobuchar JA (2007) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst AES-23(3):325–331
6. Dach R, Brockmann E, Schaer S et al (2009) GNSS processing at CODE: status report. J Geodesy 83(3–4):353–365
7. Yuan Y, Huo X, Ou J et al (2008) Refining the Klobuchar ionospheric coefficients based on GPS observations. IEEE Trans Aerosp Electron Syst 44(4):1498–1510
8. Wang N, Yuan Y, Li Z et al (2016) Improvement of Klobuchar model for GNSS single-frequency ionospheric delay corrections. Adv Space Res 57(7):1555–1569
9. Chenghui C, Lilong L, Junyu LI et al (2015) Establishment of region ionospheric delay model in Nanning based on improved Klobuchar model. J Geod Geodyn 35:797–800
10. Libing XU, Chenglin C, Guangxi C et al (2015) A new Klobuchar model and its precision analysis for Beidou navigation satellite system. J Geod Geodyn 118(1):73–82
11. Zhang H (2006) China region ionosphere based on ground-based GPS monitoring and delay corrective research. Astronomical Observatory, CAS, Shanghai
12. Chen J, Huang L, Liu L et al (2017) Applicability analysis of vtec derived from the sophisticated Klobuchar model in China. ISPRS Int J Geo-Inf 6(3):75–89
13. Liu L, Chen J, Huang L et al (2018) A sophisticated Klobuchar model based on the holt exponential smoothing model. Geomat Inf Sci Wuhan Univ 43(04):599–604
14. Zhang X, Ren X, Wu F et al (2014) Short-term TEC prediction of ionosphere based on ARIMA model. Acta Geod Cartogr Sin 43(2):118–124
15. Brunner FK, Gu M (1991) An improved model for dual frequency ionospheric correction of GPS observations. Manuscripta Geod 16(3):205–214

Research on Satellite Selection Algorithm Based on Delaunay Triangulation Projection

Ning Wang[1], Kezhao Li[1,2(✉)], and Jinben Wei[3]

[1] School of Surveying and Landing Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China
[2] Collaborative Innovation Center of BDS Research Application, Zhengzhou 450052, China
lkznwpu@126.com
[3] Zhengzhou Xinda Institute of Advanced Technology, Zhengzhou, China

Abstract. How to improve the speed and accuracy of location information is one of the key issues in current satellite navigation and positioning. This paper proposes a satellite selection algorithm based on the projection of the Delaunay triangulation. The main idea of the algorithm is: Projecting the Delaunay triangulation through the in-plane polyhedral projection method to quickly calculate the volume of the ball inscribed polyhedron. The relationship between polyhedral volume and GDOP is verified. The experimental calculations show that when the number of visible satellites participating in the solution is less than 9, with the increase of polyhedron volume, the GDOP value decreases significantly. When the number of visible satellites reaches 9, the GDOP value does not change significantly with the increase of the polyhedral volume, but the solution aging is significantly reduced. Conclusion: Among the many visible satellites, 9 satellites with better spatial geometric distribution are selected as the solution satellites, and the navigation and positioning accuracy is high, and the solution aging is taken into consideration.

Keywords: Maximum polyhedron volume · Geometric Dilution of Precision (GDOP) · Delaunay triangulation · Satellite selection

1 Introduction

Along with the construction and the upcoming global network operation of BDS-3 and Galileo, as well as the modernization of GPS and GLONASS, by then, there will be hundreds of satellites providing navigation services [1, 2]. The number of visible satellites in near-Earth space target navigation will increase to 40–50. Tacking all the observation data from the visible satellites as the processing data, the position calculation is very large and difficulty. Thus time-consuming of navigation response is significantly reduced. Especially when the least squares solution is solved, the huge data causes the dimension to be difficult and the number of iterations is greatly

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 143–151, 2019.
https://doi.org/10.1007/978-981-13-7751-8_15

increased. How to ensure the accuracy of its solution and ensure its solution rate and stability is one of the key issues for the multi-system satellite navigation positioning.

The current satellite selection algorithm mainly has the largest polyhedral volume method and the determinant minimum method [3]. The maximum polyhedral volume method and the maximum tetrahedral volume of the four satellites have a rigorous theory [4]. For five- satellite, six- satellite or even more satellites, there is no theoretical proof of how to choose a satellite. However, in the current solution environment, the visible satellites reach forty or fifty. If all participate in the solution will reduce the solution speed, and it is easy to cause the matrix dimension to suffer and cause the program to crash. But only choose four satellites and discard the remaining data, which is also a kind of data loss.

Based on the Satellite Tool Kit (STK) simulation platform, this paper simulates the current huge visible navigation satellite environment, and proposes an equivalent evaluation method based on Delaunay triangulation for segmented polyhedron. Related simulation experiments were carried out. The correctness of the equivalent evaluation method is verified, and the distribution and law of visible satellites in the optimal geometry are verified to verify the correctness of the relevant selection ideas.

2 Spatial Geometric Accuracy Index:GDOP

The GNSS pseudorange observation equation is as shown in Eq. (1) [5]:

$$\tilde{\rho}_i^S(t) = \rho_i^S(t) + c \cdot \delta t_i(t) - c \cdot \delta t^S(t) + \delta I_i^S(t) + \delta T_i^S(t) \tag{1}$$

Where

$$\rho_i^S(\mathrm{t}) = \mathrm{c} \cdot \left(t_i(\mathrm{t}) - t^S(\mathrm{t})\right) = \left[(X^s(\mathrm{t}) - X_i)^2 + (Y^s(\mathrm{t}) - Y_i)^2 + (Z^s(\mathrm{t}) - Z_i)^2\right]^{1/2} \tag{2}$$

In Eq. (1): t is the observation epoch; $\delta t^S(t)$ is the S satellite clock difference of the satellite observing the epoch at t; $\delta t_i(t)$ is the receiver clock error of the satellite receiver of the i station at t observation epoch; $\delta I_i^S(t)$ is the ionospheric delay effect of the i-station relative to the S-th satellite in observing the epoch; $\delta T_i^S(t)$ is the influence of i station relative to the S satellite on the tropospheric delay error of t observation epoch.

When the number of satellites tracked by the GNSS receiver is greater than 4 and the number of observed epochs is n_t, there is an error matrix equation:

$$V_i(t) = A_i(t)\,\delta\widehat{G}_\iota(t) - L_i(t) \tag{3}$$

Where

$$A_i(t) = \begin{bmatrix} k_i^1(t) & l_i^1(t) & m_i^1(t) & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_i^{\mathrm{m}}(t) & l_i^m(t) & m_i^m(t) & -1 & 0 & 0 \\ k_i^{m+1}(t) & l_i^{m+1}(t) & m_i^{m+1}(t) & 0 & -1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_i^{m+n}(t) & l_i^{m+n}(t) & m_i^{m+n}(t) & 0 & -1 & 0 \\ k_i^{m+n+1}(t) & l_i^{m+n+1}(t) & m_i^{m+n+1}(t) & 0 & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_i^{m+n+\omega}(t) & l_i^{m+n+\omega}(t) & m_i^{m+n+\omega}(t) & 0 & 0 & -1 \end{bmatrix}$$

$$\delta G_i(t) = \begin{bmatrix} \delta X_i \\ \delta Y_i \\ \delta Z_i \\ \delta T_{BDSi} \\ \delta T_{GPSi} \\ \delta T_{GLOi} \end{bmatrix}$$

$$L_i(t) = \begin{bmatrix} \rho_i^{'1}(t) - \rho_i^1(t) + c * \delta t_i(t) + \delta I_i^1(t) + \delta T_i^1(t) \\ \vdots \\ \rho_i^{'\mathrm{m}}(t) - \rho_i^m(t) + c * \delta t_i(t) + \delta I_i^m(t) + \delta T_i^m(t) \\ \rho_i^{'\mathrm{m}+1}(t) - \rho_i^{\mathrm{m}+1}(t) + c * \delta t_i(t) + \delta I_i^{\mathrm{m}+1}(t) + \delta T_i^{\mathrm{m}+1}(t) \\ \vdots \\ \rho_i^{'\mathrm{m}+\mathrm{n}}(t) - \rho_i^{\mathrm{m}+\mathrm{n}}(t) + c * \delta t_i(t) + \delta I_i^{\mathrm{m}+\mathrm{n}}(t) + \delta T_i^{\mathrm{m}+\mathrm{n}}(t) \\ \rho_i^{'\mathrm{m}+\mathrm{n}+1}(t) - \rho_i^{\mathrm{m}+\mathrm{n}+1}(t) + c * \delta t_i(t) + \delta I_i^{\mathrm{m}+\mathrm{n}+1}(t) + \delta T_i^{\mathrm{m}+\mathrm{n}+1}(t) \\ \vdots \\ \rho_i^{'\mathrm{m}+\mathrm{n}+\omega}(t) - \rho_i^{\mathrm{m}+\mathrm{n}+\omega}(t) + c * \delta t_i(t) + \delta I_i^{\mathrm{m}+\mathrm{n}+\omega}(t) + \delta T_i^{\mathrm{m}+\mathrm{n}+\omega}(t) \end{bmatrix}$$

On the basis of Eq. (3), according to the principle of least squares, we can obtain $\delta\widehat{G_\iota}(t)$ of n_t epochs observed at i station:

$$\delta\widehat{G_\iota}(t) = (A_i^T(\mathrm{t})A_i(t))^{-1}A_i^T(\mathrm{t})\,L_i(t) \tag{4}$$

According to the error propagation law of the adjustment data processing, the cofactor matrix of the unknown parameter optimal estimate $\delta\widehat{G_\iota}(t)$ can be expressed as:

$$\widehat{\mathrm{Q}}_{\widehat{\delta G}\widehat{\delta G}} = (A^T(\mathrm{t})A_i(t))^{-1} = \begin{bmatrix} Q_{\widehat{\delta X_i}\widehat{\delta X_i}} & Q_{\widehat{\delta X_i}\widehat{\delta Y_i}} & Q_{\widehat{\delta X_i}\widehat{\delta Z_i}} & Q_{\widehat{\delta X_i}\widehat{\delta D_i}} \\ Q_{\widehat{\delta Y_i}\widehat{\delta X_i}} & Q_{\widehat{\delta Y_i}\widehat{\delta Y_i}} & Q_{\widehat{\delta Y_i}\widehat{\delta Z_i}} & Q_{\widehat{\delta Y_i}\widehat{\delta D_i}} \\ Q_{\widehat{\delta Z_i}\widehat{\delta X_i}} & Q_{\widehat{\delta Z_i}\widehat{\delta Y_i}} & Q_{\widehat{\delta Z_i}\widehat{\delta Z_i}} & Q_{\widehat{\delta Z_i}\widehat{\delta D_i}} \\ Q_{\widehat{\delta G_i}\widehat{\delta X_i}} & Q_{\widehat{\delta G_i}\widehat{\delta Y_i}} & Q_{\widehat{\delta G_i}\widehat{\delta Z_i}} & Q_{\widehat{\delta G_i}\widehat{\delta D_i}} \end{bmatrix} \tag{5}$$

Description i station t observation epoch visual navigation satellite space geometry on the three-dimensional position and clock difference can be defined as GDOP (Geometric Dilution of Precision), the expression is:

$$\mathrm{GDOP} = \sqrt{Q_{\widehat{\delta X_i}\widehat{\delta X_i}} + Q_{\widehat{\delta Y_i}\widehat{\delta Y_i}} + Q_{\widehat{\delta Z_i}\widehat{\delta Z_i}} + Q_{\widehat{\delta D_i}\widehat{\delta D_i}}} \tag{6}$$

In GNSS systems, the positioning accuracy can be expressed as the product of the geometric precision factor and the observation accuracy:

$$\sigma_A = \delta_0 \,.\, GDOP \tag{7}$$

It can be seen from the definition of the GDOP that when the observation accuracy δ_0 is determined, the smaller the GDOP, the higher the positioning accuracy.

The number of satellites in the GNSS system has increased significantly, making the network of visible satellites more dense. Using the Delaunay triangulation feature "maximum angle maximization" and "arbitrary triangles are not coplanar", the resulting triangulation is unique. A large number of existing literatures have demonstrated the correlation between the maximum multi-area volume method and the GDOP value [6–9]. Therefore, the following simulation experiment ideas are proposed.

3 Simulation Experiment

After analyzing the shortcomings of the maximum polyhedral volume experiment in the existing literature, this paper proposes a simulation experiment based on the maximum multi-area volume method, using the STK platform for simulation experiments.

3.1 Polyhedron Volume Selection Method Based on Dalaunay Triangulation

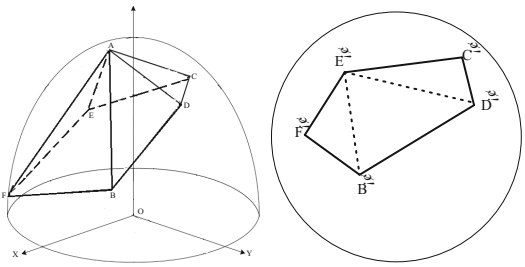
Firstly, establish a unit hemisphere with the test site as the center of the sphere, and project all visible satellites onto this hemisphere; then, select one of the satellites with the highest elevation angle as the "seed satellite " and project the remaining satellites onto the equatorial plane of the hemisphere; and then the tetrahedron corresponding to the "seed satellite" formed by the satellite represented by each triangle vertex in the D-triangle network constitutes the corresponding Hamilton path; finally, all the tetrahedral bodies are accumulated and summed to obtain the sphere inscribed polyhedron volume.

Take Fig. 1 as an example. It is assumed that there are six visible satellites A, B, C, D, E, and F at a certain moment. Since the altitude angle of satellite A is the largest in the station coordinate system, satellite A is used as the "seed satellite ", and then project the remaining satellites onto the XOY surface, with B', C', D', E', and F' points, and generate these discrete points to generate a satellite triangulation to obtain a triangle

B'E'F' and a triangle B'E'D', triangle C'D'E', and then obtain the tetrahedral ABEF, tetrahedral ABED and tetrahedral ACDE, finally can transform the volume of the in-plane polyhedron ABDCEF which is not easy to solve into the easily solved tetrahedral ABEF, solving the volume sum of tetrahedral ABED and tetrahedral ACDE. The tetrahedral volume formula can be expressed by Eq. (8):

$$V_{ABEF} = \frac{1}{6}\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_5 - x_2 & y_5 - y_2 & z_5 - z_2 \\ x_6 - x_5 & y_6 - y_5 & z_F - z_E \end{vmatrix} \tag{8}$$

Calculate V_{ABED} and V_{ACDE} in this way, and then calculate the volume of the inscribed polyhedron ABDCEF.



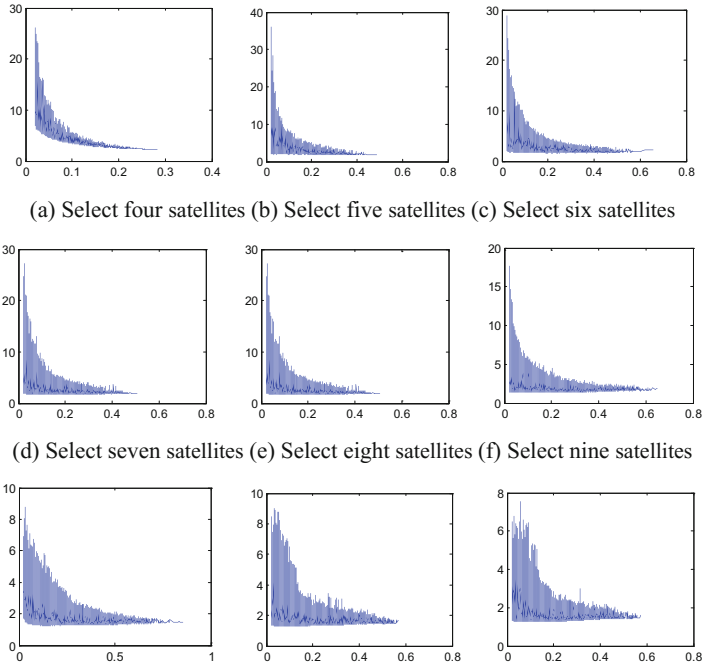
(a) **The space graphics of satellite distribution** (b) **The over head view of Satellite distribution**

Fig. 1. Distribution of satellite

The hemispheres are divided by an azimuth angle of 45 degrees and an elevation angle of 30 degrees. A satellite is randomly generated inside each grid, traversing all combinations, and calculating its GDOP. Since the corresponding GDOP value reaches the order of 10^2 when the volume value is too small, in order to clearly show the relationship between the volume and the GDOP value, the portion whose volume value is less than 0.02 is eliminated. In Fig. 2, the relationship between the polyhedral volume composed of four satellites and the corresponding GDOP value to the polyhedral volume composed of twelve satellites and the corresponding GDOP value are sequentially given. Take the volume of the polyhedron in the sphere formed by the satellite as the horizontal axis, and the corresponding GDOP value is the vertical axis.

Conclusion

(1) In multi-satellite positioning, as the volume value increases, the GDOP value still shows a decreasing trend.
(2) Starting from the 9 satellites positioning, in the figure, the local GDOP value shows a slowly increasing trend with the increase of volume.



(a) Select four satellites (b) Select five satellites (c) Select six satellites

(d) Select seven satellites (e) Select eight satellites (f) Select nine satellites

(g) Select ten satellites (h) Select eleven satellites (i) Select twelve satellites

Fig. 2. Relationship between GDOP and polyhedral volume

3.2 Cluster Method for Satellite Selection

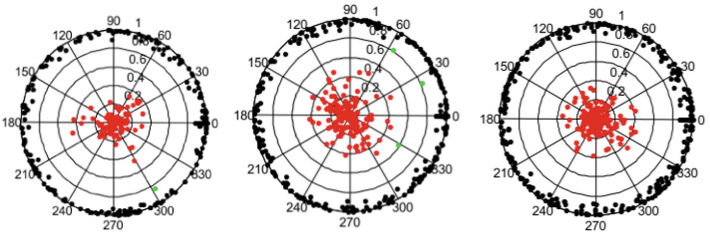
Clustering method selection means that in the process of satellite selection, the visible satellites are divided into different categories by using prior information to reduce the number of GDOP calculations and improve the real-time performance of the satellite selection algorithm. The prior information is mostly used in the following two aspects: (1) Firstly, the satellite with the highest altitude angle and the satellite with the lowest altitude angle are selected; (2) The visible satellites are divided into low elevation angle, medium elevation angle and high elevation angle according to the elevation angle. However, whether these prior information is correct still requires further study.

In view of the more use of the existing satellite selection methods, the satellites are grouped according to the elevation angle. This paper discusses the correctness of the satellite group by height angle.

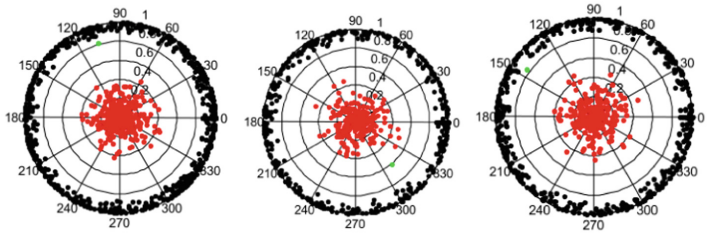
The unit hemispherical surface is divided by an azimuth angle of 45 degrees and an elevation angle of 30 degrees. A satellite is randomly generated in each grid, and the satellite group is selected by the optimal GDOP value selection algorithm, and the corresponding spatial position is recorded. In order to better analyze the satellite arrangement, the following coordinate rotation is adopted for the selected satellite combination (the azimuth of the satellite defining the minimum elevation angle in each satellite group is α)

$$\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

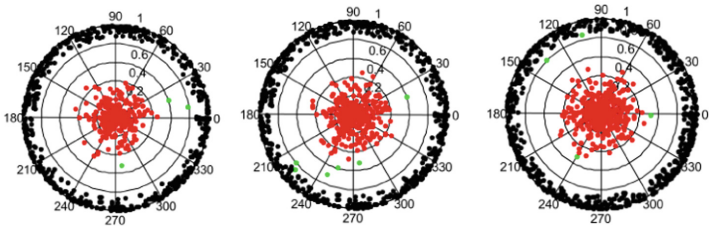
Divided into three regions according to the elevation angle mentioned in the existing literature: low elevation angle region (0°–30°), medium elevation angle region (30°–60°), high elevation angle region (60°–90°). The satellites in the low elevation region of the final selected optimal satellite combination are marked black, the satellites in the middle elevation region are marked in green, and the satellites in the high elevation region are marked in red. Repeat the above experiment 100 times. The projection pattern of the satellite on the xoy plane in the case of optimal GDOP when four satellite combinations are selected, and the projection pattern of the satellite on the xoy plane in the case of optimal GDOP when twelve satellite combinations are selected are given in turn (Fig. 3):



(a) Select four satellites (b) Select five satellites (c) Select six satellites



(d) Select seven satellites (e) Select eight satellites (f) Select nine satellites



(g) Select ten satellites (h) Select eleven satellites (i) Select twelve satellites

Fig. 3. Satellite distribution of minimum GDOP satellite selection algorithm

Conclusion

(1) When the number of selected satellites is small, the number of satellites at the bottom corner and the top-most angle satellite appearing in the minimum GDOP value selection is small. Only when the number of selected satellites is large, the number of satellites at the bottom corner and the top-most angle satellites has been significantly improved.
(2) At present, the visible satellites are divided into three regions of low, medium and high at an elevation angle of 30° and 60°, and the ratio of satellites in the three regions is 1:0:3, which is basically in accordance with the optimal GDOP value satellite combination distribution. However, there are still some satellites with the best GDOP value in the middle elevation domain.
(3) The distribution of the base satellites is not a uniform arrangement as considered by current research. Taking four-satellite positioning as an example, the current research suggests that the optimal combination is to select the satellite with the highest elevation angle as the top seat, and the base satellite is evenly arranged at 120° on the bottom circle, but the spatial structure of the actual optimal visible satellite combination is difficult to meet this condition.

4 Conclusion

In this paper, a set of simulation experiments are given to analyze the problems in the most widely used polyhedral volume method and clustering method. The volume of the hemispherical inscribed polyhedron is calculated by the Delaunay triangulation projection method. The variation of the polyhedral volume change and the GDOP value is analyzed. It is found that the polyhedral volume still has some guiding significance for the selection combination regardless of the number of selected satellites. The optimal GDOP is adopted. The value selection algorithm analyzes the correctness of the satellite region by using prior information in the clustering method, and finds that the spatial structure of the actual optimal visible satellite combination is difficult to meet the conditions. It provides some ideas for the subsequent GNSS satellite selection algorithm research.

Acknowledgment. This research was funded by the National Natural Science Foundation of China (Nos. 41774039,41272373).

References

1. Afifi A, El-Rabbany A (2016) Precise point positioning using triple GNSS constellations in various modes. Sensors 16(6):779
2. Ning J, Yao Y, Zhang X (2013) Summary of the development of global navigation satellite systems. J Navig Position 1(01):3–8
3. Kihara M, Okada T (2005) A satellite selection method and accuracy for the global positioning system. Navigation 31(1):8–20

4. Zheng Z, Huang C, Feng C, Zhang F (2003) The amendment of optimun geometry based on pour satellites. Acta Astron Sin (3):310–317
5. Li K, Yang L, Chai L (2014) Global navigation satellite system. China Coal Industry Publishing House, Beijing
6. Wang W, Huang B (2014) Research on satellite selection algorithm for multi-satellite navigation system. In The fifth annual conference for satellite navigation of China, Nanjing
7. Zhang M, Zhang J (2009) A fast satellite selection algorithm: beyond four satellites. IEEE J Sel Top Signal Process 3(5):740–747
8. Zhao X, Zhang J, Zhu L (2012) The fast satellite selection algorithm for COMPASS. Space Electron Technol 9(2):4–9
9. Chen Y (2014) Research on positioning solution technology for GPS and BD dual-mode GNSS receiver. Xiamen University

Establishment of Regional Tropospheric Delay Model in Australia

Yongchao Ma, Peng Chen(✉), Hang Liu, and Qingshan Ruan

College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China
chenpeng0123@gmail.com

Abstract. Tropospheric delay error is independent of signal frequency and has strong temporal and spatial variation. It is one of the most important sources of error in satellite navigation and positioning. The common tropospheric empirical model is affected by the temporal and spatial characteristics of meteorological parameters, which cannot meet the needs of precision positioning. In this paper, we use the tropospheric delay data obtained from the GNSS continuous operation reference station in 2013–2017 years to analyze its spatiotemporal variation characteristics, and analyze the applicability of several commonly used tropospheric empirical models in Australia. At the same time, a ZTD spatiotemporal modeling method based on improved polyhedral function is proposed, and a regional non-meteorological delay model (A_ZTD) is established for the zenith tropospheric delay calculated by the Australian base station. After testing, the internal coincidence accuracy of the region model (A_ZTD) is 4.1 cm, and the fitting effect is better. The ZTD provided by the reference station, which is not involved in the modeling, the tropospheric delay provided by the GGOS Atmosphere and the ZTD derived from the sounding data solution are used as the external coincidence check. The results show that the accuracy of the A_ZTD model in Australia is better than that of the GZTD model, UNB3, UNB3m and UNB4 models. Compared with the empirical model, the accuracy of the A_ZTD model is improved in the region. This model can describe the spatial variation of the troposphere without meteorological parameters. It is more suitable for the precise correction of the regional tropospheric delay.

Keywords: Tropospheric delay · Regional tropospheric delay model · UNB series model

1 Introduction

The radio waves experience propagation delays when passing through the neutral atmosphere (primarily the troposphere), which are known as the tropospheric delays. The tropospheric delay is one of the main error sources in space geodetic techniques such as GNSS and VLBI. Its value varies from 2 m to 20 m depending on the satellite elevation angle [6]. At present, the methods of eliminating tropospheric delay errors mainly include model correction method, parameter estimation method and external correction method. The model correction method can be divided into two categories. One is the tropospheric delay model that requires meteorological parameters, such as

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 152–162, 2019.
https://doi.org/10.1007/978-981-13-7751-8_16

Hopfield, Saastamoinen, Black. And the other is an empirical model that does not require meteorological parameters, such as the UNB series model (UNB1-UNB4, UNB3m), EGNOS, GPT series models [1, 3, 8]. In recent years, many domestic scholars have established a series of new global tropospheric models to overcome the shortcomings of traditional tropospheric models, such as SHAO, IGGtrop, GZTD. They all achieved centimeter-level accuracy on a global scale [4, 9, 13, 14].

The tropospheric delay has a strong regionally due to the water vapor in the atmosphere of different regions is randomly changed with spatial-temporal distribution. The empirical tropospheric delay models are established from meteorological reanalysis data. The calculated tropospheric delay exist 1–2 cm error with a low time resolution [2]. It is difficult to meet the requirements of high-precision spatial measurement such as regional precision positioning [5], InSAR atmospheric correction [7, 10]. The increase of the number of consecutively operated reference stations in region provides a good opportunity to study the characteristics of tropospheric delay spatial variation and establish regional models. This paper compares the regional tropospheric delay model (A_ZTD model) which is based on the tropospheric delay calculated by the Australian base station observation data from 2013–2017. A_ZTD model is compared with the commonly used UNB series model and GZTD model, and the simulation effect and error characteristics of the above model in Australia are discussed.

2 Data Source and Spatial and Temporal Characteristics of Tropospheric Delay

2.1 Modeling Data and Processing Methods

This paper selects the observation data of 130 CORS stations in Australia from 2013 to 2017 with 30 s as the sampling interval. And it combines the observation data with the data of 4 IGS tracking stations around Australia to calculation the zenith tropospheric delay. The distribution of CORS stations and sounding stations is shown in Fig. 1. The evenly distributed 116 base stations are used as modeling stations, and the remaining 14 base stations serve as checking stations.

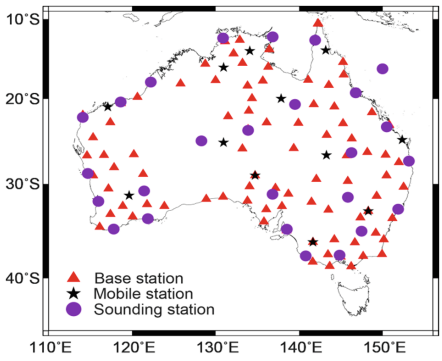


Fig. 1. Distribution of CORS stations and sounding stations in Australia

In this paper, the GAMIT/GLOBK software is used to estimate the Zenith Tropospheric Delay (ZTD) of Australia CORS station data. One estimated parameter is selected per hour when estimating. IGS final ZTD products at four IGS stations located in Australia were taken as reference for comparison. The result shows that the accuracy of ZTD estimated by GAMIT is about 4 mm, so this paper uses the ZTD solved by GAMIT as the true value.

2.2 Spatial-Temporal Variation Characteristics of Tropospheric Delay

Establishing a high-precision delay model requires an accurate understanding of the temporal and spatial variation of tropospheric delay. Figure 2 shows the time series diagrams of tropospheric delays at MAIN, NMTN and NTJN stations in Australia from 2013 to 2017.

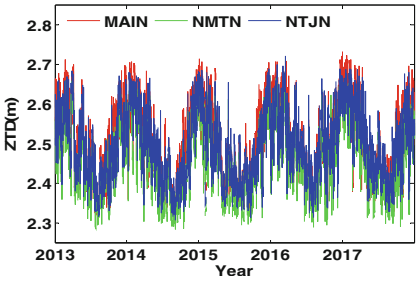


Fig. 2. ZTD Time Series of MAIN, NMTN and NTJN Stations from 2013 to 2017

It can be seen from Fig. 2 that the tropospheric delay amount in this region has obvious annual and semiannual variation. Australia is located in the middle and low latitudes. In order to study whether there are short-period variations, the ZTD data of four CORS stations from 2013 to 2017 in the region were randomly selected to obtain the annual average of each moment at 24 times a day, and their departure were calculated to highlight the short-period variation (Fig. 3). It can be seen from Fig. 3 that the tropospheric delay in this region has a certain daily cycle change and a half-day cycle change, which is consistent with the conclusions of Yao et al. [11].

In order to analyze the spatial distribution of tropospheric delay, this paper draws the distribution of the average annual value of GGOS_ZTD in Australia at 00 UTC over the whole 2015 year (Fig. 4a) and the distribution of the height of GGOS grid point in Australia (Fig. 4b). It can be seen from Fig. 4a that the ZTD increases with the decrease of latitude in the ocean area, and gradually increases with the increase of longitude in the 10°S–40°S region. But this phenomenon is not obvious in the 40°S–50°S region, and the above situation is not met in the land region. It is found from Fig. 4b that the terrain in Australia is fluctuating, and the spatial distribution of ZTD is negatively correlated with the elevation.

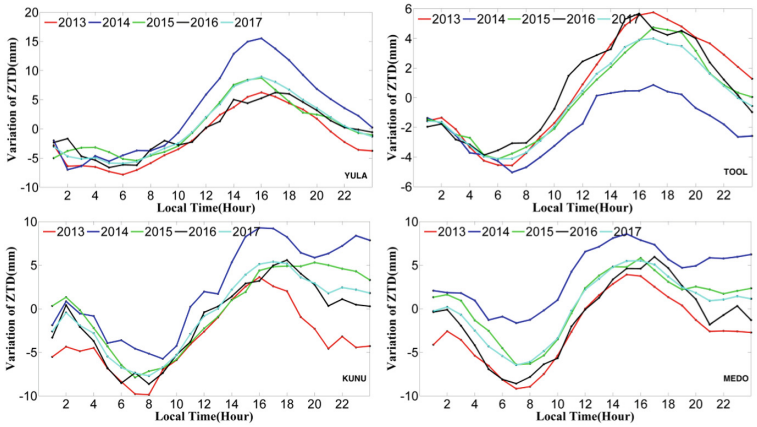


Fig. 3. Annual diurnal variation of ZTD2015 at 4 CORS stations in Australia

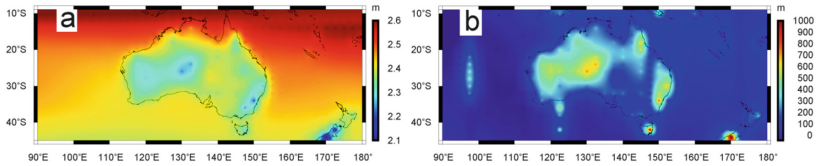


Fig. 4. Distribution of GGOS_ZTD/H (UTC 00) in Australia

3 Establishment of A_ZTD Model

In this paper, the tropospheric delay period variation characteristics are considered simultaneously in the modeling process. The daily cycle and half-day period terms are added based on the GZTD model. And the CORS_ZTD is periodically fitted according to formula (3.1). In order to intuitively study the influence of latitude and longitude and eliminate the influence of elevation on ZTD, this paper normalizes ZTD to mean sea level using formula (3.2).

$$ZTD_t = a_1 + a_2 \cos\left(2\pi \frac{doy - C_1}{365.25}\right) + a_3 \cos\left(4\pi \frac{doy - C_2}{365.25}\right) + a_4 \cos\left(2\pi \frac{hod - C_3}{24}\right) + a_5 \cos\left(4\pi \frac{hod - C_4}{24}\right) \tag{3.1}$$

$$ZTD = ZTD_t \bullet \exp(\beta h) \tag{3.2}$$

Where α_1 and β are the average ZTD of mean sea level and elevation correction coefficient ($\beta = -0.00013137$) [12], and α_2, α_3, α_4, α_5 are the coefficients related to ZTD annual cycle, semi-annual cycle, diurnal variation and semi-diurnal variation. C_1, C_2, C_3, and C_4 are the initial phase of the annual cycle, half-year cycle, diurnal variation and semi-diurnal variation. *doy* is the day of the year. hod is the UTC time; h is the height of the station.

When fitting regional ZTD, the spherical harmonic function has rank deficiency, and the solution of spherical crown harmonic function is complex. The simple polynomial function cannot accurately express the characteristics of the spatial-temporal distribution, and the multi-faceted fitting function which is approximated by a series of regular mathematical surface fitting can be well reflects the spatial distribution characteristics of ZTD [10]. This paper uses the multi-faceted function fitting method to expand the above α_i.

$$\alpha_i = \sum_{j=1}^{u} \delta_j Q(x, y; x_j', y_j') \tag{3.3}$$

Where $Q(x, y; x_j^{'}, y_j^{'})$ is the kernel function, μ is the number of nodes, δ_j is the parameter to be estimated. In this paper, considering that the curvature and projection errors of the earth affect the plane results when studying large areas, the polyhedral function is improved as follows:

$$\alpha_i = \sum_{j=1}^{u} \delta_j Q(\lambda, \varphi; \lambda_j^{'}, \varphi_j^{'}) \tag{3.4}$$

$$Q = [d^2(\lambda, \varphi; \lambda_j^{'}, \varphi_j^{'}) + \sigma^2]^{\gamma} \tag{3.5}$$

Where σ^2 represents the smoothing factor. It has a great influence on the fitting accuracy. In this paper, $\sigma = 0.01$ is determined by the dichotomy. The kernel function used is a positive hyperboloid function, so $\gamma = 1/2$.

After the linearization of the formula (3.1), the least square method is used to calculate the model coefficients in this paper. The A_ZTD model does not require meteorological measurement data, and the ZTD can be calculated according to the latitude, longitude, altitude and time of the point to be sought. The model is easy to use and has good correction accuracy.

4 Test and Analysis of Model Accuracy

In this paper, the tropospheric delay of 30 evenly distributed CORS stations in Australia is used as the internal reference data. The tropospheric delay of 14 non-modeled CORS stations, the zenith tropospheric delay provided by GGOS Atmospheres and the tropospheric delay amount calculated by sounding data are used as the external reference data. In this paper, the location coordinates and time from different data sources are used to estimate ZTD, and the bias and root mean square error (RMS) between them are compared with the true values. The calculation formulas are as follows:

$$Bias = \frac{1}{N} \sum_{j=1}^{u} \left(ZTD_i^{\text{model}} - ZTD_i^{\text{dat}_j} \right) \tag{4.1}$$

$$RMS = \sqrt{\frac{1}{N}\sum_{j=1}^{u}\left(ZTD_i^{\text{model}} - ZTD_i^{\text{dat}_j}\right)^2} \tag{4.2}$$

4.1 Internal Accuracy Verification

Time series diagram (Fig. 5) is drawn using the ZTD data of MEDO, KUNU, TOOL, and YULA stations from 2013 to 2017. It can be seen from Fig. 5 that the UNB3/3m/4 model has the same fitting trend with a low accuracy and poor fitting effect, which reflects the annual cyclical variation of ZTD. This may be due to the fact that the UNB series model has less consideration of regional meteorological changes in Australia when it is established, indicating that the UNB3/3m/4 model has certain limitations in the Australian. The tropospheric delay obtained by the GZTD model and the A_ZTD model are similar to the measurement values, which reflect their annual cyclical changes. The A_ZTD model has a better agreement with the measured ZTD time series.

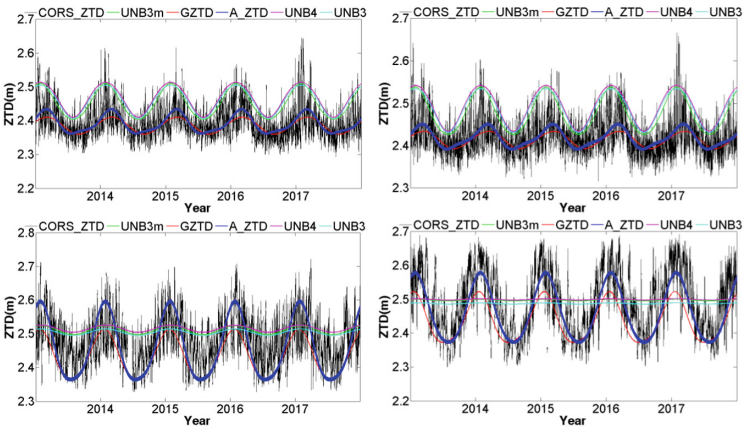


Fig. 5. Comparisons between the measured ZTD of CORS station and the results of different models in 2014

In order to verify the rationality of the model, the CORS_ZTD of 30 CORS stations from 2013 to 2017 and the ZTD solved by the A_ZTD model are used to verify the internal accuracy. Table 1 shows the statistical results of the Bias and RMS of the A_ZTD model. The average deviation of the model is 0.3 cm. The average root mean square error of the model is 4.1 cm, indicating that the new model has a good fitting effect.

Table 1. Statistics of Bias and RMS for A_ZTD model (unit:cm)

	Bias	RMS
Max	5.36	6.6
Min	−4.84	3.1
Mean	0.3	4.1

4.2 External Conformity Verification

In this section, the accuracy of model correction is further validated based on the external coincidence benchmark data from 2016 to 2017.

4.2.1 Verify the Accuracy of the Model with CORS_ZTD

Table 2 shows the statistical results of the deviation and root mean square error of the above model in 14 CORS stations in Australia. It can be seen from Table 2 that the RMS mean of the ZTD calculated by the A_ZTD model in the CORS station is 5.24 cm, which is better than the GZTD model and the UNB3/3m/4 model under the same conditions. It is found from the GZTD model and the UNB series model deviation statistics list that the absolute value of Bias is larger than 2 cm for most stations, indicating that their estimated values have a systematic deviation from the true value. The minimum RMS of the GZTD model is 3.79 cm, the largest value is 8.71 cm and the average value is 6.13 cm. The accuracy of the model does not achieve high precision on a global scale in Australia (Bias: −0.02 cm, RMS: 4.24 cm) [5]. It may be caused by two aspects. On the one hand, the GZTD model is established by grid data that provided by GGOS Atmosphere. And there is systematic deviation between GGOS_ZTD and GNSS_ZTD [12]. On the other hand, the GZTD model itself does not

Table 2. Statistics of Bias and RMS for different models at 14 CORS station (unit:cm)

Site	A_ZTD		GZTD		UNB3		UNB3m		UNB4	
	Bias	RMS	Bias	RMS	Bias	RMS	Bias	RMS	Bias	RMS
NHIL	1.50	3.70	0.60	3.79	5.50	6.98	4.39	6.29	5.77	7.44
PARK	1.94	4.75	2.26	5.17	3.86	5.89	3.28	5.52	4.34	6.39
COOB	−1.24	5.34	1.16	5.88	6.69	8.83	6.68	8.78	7.03	9.40
ERMG	−1.52	5.66	1.13	5.71	6.64	8.69	6.63	8.66	6.95	9.29
BNDY	0.90	5.40	4.35	7.18	2.40	6.22	2.42	6.19	3.16	6.54
KARR	0.98	5.58	0.56	6.28	7.30	10.59	7.60	10.63	8.24	11.23
RKLD	−0.04	5.56	1.32	6.87	5.00	9.25	5.41	9.30	5.86	9.80
TOW2	−2.14	6.28	4.29	7.31	0.99	6.96	1.45	6.88	1.93	7.13
COEN	−3.04	5.89	3.85	8.71	−2.29	7.59	−1.19	7.34	−0.95	7.32
PERT	−1.40	4.46	−0.61	4.53	8.30	9.78	7.77	9.36	8.79	10.29
YELO	1.06	4.57	0.35	4.66	6.37	7.94	6.08	7.70	7.08	8.55
KMAN	−0.10	5.67	2.73	6.90	3.28	10.09	4.20	10.35	4.26	10.50
MAIN	0.29	5.09	3.36	6.58	0.73	8.40	1.84	8.56	1.85	8.58
YULA	0.65	5.40	2.38	6.21	4.72	7.78	4.79	7.70	5.31	8.33
Mean	−0.15	5.24	1.98	6.13	4.25	8.21	4.38	8.09	4.97	8.63

achieve global high precision in Australia. The accuracy of the UNB series model in the Australian region is roughly the same and has a positive average deviation. And the value of the UNB series model in the southern hemisphere is higher than the measurement value. The accuracy of UNB3 is better than that of UNB4, but the improvement is not significant. Compared with other models, the stability of A_ZTD model is better. And the fitting results are more in line with the ZTD environment characteristics of the Australian region.

4.2.2 Verify the Accuracy of the Model with GGOS_ZTD

The global tropospheric delay products provided by GGOS have a high accuracy and can be used as a basis to verify the accuracy of the tropospheric model. Figure 6 shows the Bias and RMS statistics of the GGOS_ZTD data compared with different models. Overall, the accuracy of A_ZTD (Bias: 0.95 cm, RMS: 5.82 cm) is comparable to the accuracy of the GZTD model (Bias: 0.29 cm, RMS: 5.71 cm) and is superior to the UNB3 model (Bias: 6.14 cm, RMS: 9.12 cm), UNB3m model (Bias: 6.01 cm, RMS: 8.95 cm) and UNB4 model (Bias: 6.92 cm, RMS: 9.62 cm). The limitations of UNB series models in the southern hemisphere are further illustrated. The UNB series model deviations are all positive, and the UNB3m overall accuracy is better than the UNB3/4 model, which is consistent with the results of using the CORS_ZTD data. The overall accuracy of the GZTD model using GGOS_ZTD data verification is better than the accuracy of using CORS_ZTD data, while the A_ZTD model is just the opposite. This may be caused by the systematic deviation of ZTD obtained by different data and observation methods.

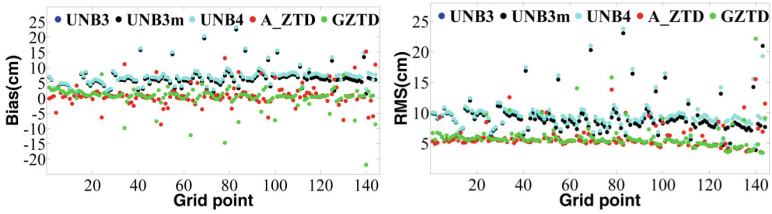


Fig. 6. Error statistics results of different models in GGOS_ZTD

4.2.3 Verify the Accuracy of the Model Using the ZTD Data of the Sounding Station

The accuracy of A_ZTD and GZTD models under two different data sources is different. In order to accurately express the model accuracy of GZTD and A_ZTD, the sounding data is used as the true value for checking. Table 3 is a statistical table of model accuracy validated on the basis of 26 sounding data in the Australian region from January to July 2018. As can be seen from Fig. 1, most of the sounding stations are located in the coastal areas of Australia. The poor extrapolation of the A_ZTD model results in low accuracy (Bias: 2.47 cm, RMS: 5.67 cm), but slightly better than the GZTD model (Bias: 3.04 cm, RMS: 5.95 cm).

Table 3. Statistics of Bias and RMS for A_ZTD GZTD model under sounding data (unit: cm)

Model	Bias	RMS
A_ZTD	2.47(−1.06, 4.63)	5.67(4.80, 8.03)
GZTD	3.04(1.21, 4.40)	5.95(4.06, 9.41)

4.3 The Influence of Latitude, Longitude and Elevation on the Model

In order to verify the influence of elevation and latitude and longitude on the accuracy of each model, the Bias and RMS of each model are solved based on the GGOS_ZTD data. Because the accuracy of the UNB series model in Australia is poor, the UNB3m with better precision is taken as an example to analysis. Figure 7 shows the distribution of deviations and root mean square errors for each model. Figure 7(a, b) (c, d) (e, f) are the Bias and RMS of the A_ZTD, GZTD, and UNB3m models, respectively.

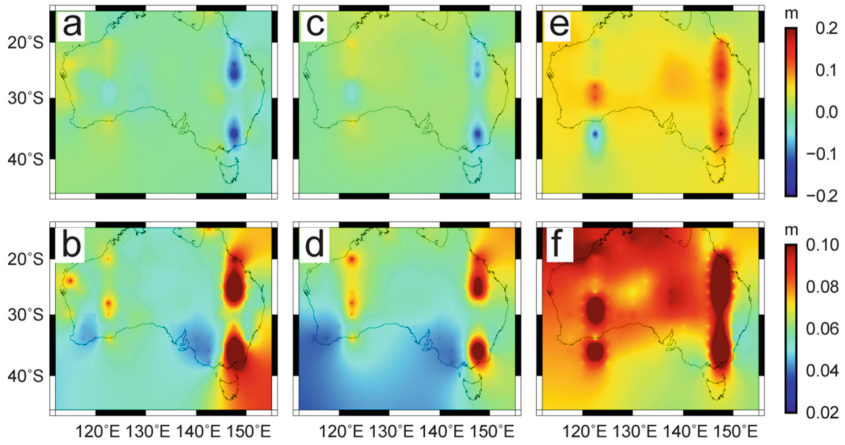


Fig. 7. Distribution of Bias and RMS in different models

It can be seen from Fig. 7 that the overall deviation of the A_ZTD model is small, and most of them are within 3 cm, which is closer to the true value than other models. It can be seen from Fig. 7a/c/e that there is no obvious relationship between the average deviation of the three models and the latitude and longitude. At the same time, it is found that the deviation distributions of the three models are almost the same and the areas where the maximum deviation occurs are consistent. In combination with Fig. 4, it can be found that large deviations occur in the places where the terrain fluctuations are large. The deviation of the A_ZTD and GZTD models may be caused by the poor correction of the parameter β which is reduced ZTD to the average sea level. The deviation of the UNB3m model reflects the defect that the model has low accuracy in high altitude areas. The overall accuracy of A_ZTD and GZTD is not much different and the distribution is consistent. And the area where RMS maximum point appears is

the same due to both models are composed of time period fitting and elevation index. It can be seen from Fig. 7b and d that the accuracy of the A_ZTD model on the Australian land is higher than that of the GZTD, but the accuracy is not as good as that of the GZTD in the ocean area around Australia. This may be caused by two aspects. On the one hand, the A_ZTD model is established by the terrestrial CORS stations. Without the data on the ocean, the model extrapolation effect is poor. And the accuracy is not as good as the GZTD model established by the global GGOS_ZTD data. On the other hand, the A_ZTD model is affected by the marine climate. It can be seen from Fig. 7f that the UNB3 model has the largest RMS value at 20°S and 120°E/150°E. The analysis results show that the model is affected by latitude and altitude, and its accuracy is low in high altitude areas, indicating that the UNB3m model has greater restrictions when using it in Australia.

5 Conclusion

In this paper, the spatial and temporal characteristics of the observed tropospheric delay at CORS stations in Australia from 2013 to 2017 are analyzed and a modified polyhedral function based correction model for tropospheric delay, A_ZTD model, is proposed. At the same time, this paper uses different data sources to verify the internal and external coincidence accuracy of A_ZTD model, and compares with UNB3, UNB3m, UNB4 model and GZTD model. Summarized as follows:

(1) The non-meteorological parameter A_ZTD model based on the characteristic parameters of tropospheric delay time variation is expanded by polyhedral function, which not only has simple modeling method and concise parameters, but also reflects the spatial and temporal variation characteristics of tropospheric delay in Australia effectively.
(2) The coincidence accuracy in A_ZTD model is 4.1 cm in average RMS and 0.3 cm in Bias, and 14 CORS stations which are not involved in modeling are used to verify the coincidence accuracy. The statistical results are as follows: the average RMS is 5.24 cm and the average Bias is −0.15 cm. Under different data sources, the accuracy of A_ZTD model is better than that of UNB series model and GZTD model under the same conditions, which verifies the feasibility and stability of the model.
(3) From the analysis of Sect. 4.3, it can be seen that the adoption constant of elevation correction coefficient β has some limitations and needs further optimization. From the comparative analysis of different models under different data, it is found that how to eliminate the systematic deviation between different data, refine the elevation correction coefficient β, and establish a high-precision regional tropospheric delay model need to be further studied.

Acknowledgements. The authors would like to thank GGOS Atmosphere and Geoscience Australia for providing relevant experimental data.

References

1. 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
2. Chen Q, Song S, Heise S, Liou YA, Zhu W, Zhao J (2011) Assessment of ZTD derived from ECMWF/NCEP data with GPS ZTD over China. GPS Solut 15(4):415–425
3. Leandro R, Santos MC, Langley RB (2006) UNB neutral atmosphere models: development and performance. Proc ION NTM 52(1):564–573
4. Li W, Yuan Y, Ou J, Chai Y, Li Z, Liou YA et al (2015) New versions of the BDS/GNSS zenith tropospheric delay model IGGtrop. J Geod 89(1):73–80
5. Yao Y, Peng W, Xu C, Cheng S (2017) Enhancing real-time precise point positioning with zenith troposphere delay products and the determination of corresponding tropospheric stochastic models. Geophys J Int 208(2):1217–1230
6. Penna N, Dodson A, Chen W (2001) Assessment of EGNOS tropospheric correction model. J Navig 54(1):37–55
7. Remy D, Falvey M, Bonvalot S, Chlieh M, Gabalda G, Froger JL et al (2011) Variability of atmospheric precipitable water in northern chile: impacts on interpretation of InSAR data for earthquake modeling. J S Am Earth Sci 31(2):214–226
8. Saastamoinen J (1972) Atmospheric correction for troposphere and stratosphere in radio ranging of satellites. In: The use of artificial satellites for geodesy, vol 15, pp 247–251
9. Yao Y, He C, Zhang B, Xu C (2013) A new global tropospheric zenith delay model GZTD. Chin J Geophys 56(7):2218–2227
10. Yao Y, Zhang R, Yi W, Song W (2012) A new regional troposphere fitting model and its application to PPP. J Wuhan Univ (Inf Sci Ed) 37(09):1024–1027
11. Yao Y, Hu Y, Zhang B (2016) Building a global zenith tropospheric delay model using multi-source data. Sci Bull 61(24):2730
12. Yao Y, Xu X, Hu Y (2017) Accuracy analysis of GGOS tropospheric delay products and its application in ppp. J Surv Mapp 46(3):278–287
13. Zhao J, Song S, Chen Q, Zhou W, Zhu W (2014) Establishment of global tropospheric zenith delay model based on vertical section function formula. Chin J Geophys 57 (10):3140–3153
14. Zhang H, Yuan Y, Li W, Li Y, Chai Y (2016) Assessment of three tropospheric delay models (IGGtrop, EGNOS and UNB3m) based on precise point positioning in the chinese region. Sensors 16(1):122

Water Level Measurements Using Multi-station and Dual-System GNSS-MR – A Case of Shuangwangcheng Reservoir

Minfeng Song, Ruya Xiao, Xiufeng He(✉), and Jie Wang

School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China
xfhe@hhu.edu.cn

Abstract. Water level information is one of the basic parameters to ensure the safety of reservoir dams and optimize the operation of water resources. In recent years, the research on monitoring altitude difference (snow depth, sea level) and surface environment (vegetation, soil moisture, temperature, and so on) using Signal-to-Noise Ratio (SNR) of GNSS observation data has become a new research hotspot. This paper analyzes the characteristics of the GNSS SNR data which is in the case of low satellite elevation angles, and the inversion principle of GNSS-MR technology based on the SNR data of single geodetic-quality Global Navigation Satellite System (GNSS) receiver to detect water level is given in details. The water level inverted by using the deformation monitoring network data of Shuangwangcheng Reservoir, an important reservoir in the East Route of South-to-North Water Transfer Project, was compared to water level data from nearby and in situ sensor. The result indicates that the variation daily time series of water level estimated by GPS L1 band SNR measurements can be achieved to 4.7 cm, 4.9 cm and 3.8 cm, compared with the measured daily water level, the correlation between the estimated value and the daily records can even reach up to 0.99. The accuracy of inversion results based on GPS L2 band data is lower than L1, and the single station RMS using L2 data is about 8.9 cm. What's more, reservoir water level inversion method based on GNSS SNR can be used as an effective supplement for reservoir water level monitoring system. For Beidou system, the SNR data of B1 signal is used for water level inversion. Because of the number of observable satellites, the accuracy is worse than GPS. The RMS, compared with the measured data of water level station, is 6.1 cm. The preliminary study shows that the GNSS-MR water level inversion method can provide real-time water level information and can be an effective supplement to the reservoir water level monitoring system.

Keywords: GNSS-MR · Water level · Multipath · Shuangwangcheng Reservoir

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 163–173, 2019.
https://doi.org/10.1007/978-981-13-7751-8_17

1 Introduction

In recent years, the technology of inversion of surface environment using GNSS reflected signal has been developed gradually. It has become a new research orientation of combining GNSS technology with remote sensing, namely GNSS-R (GNSS-Reflectometry), GNSS-R technology, which uses both GNSS direct and reflected signals at the same time, can regard the receiving antenna as a passive remote sensing capable of self-localization equipment. GNSS direct signal is right-handed signal, and the polarization of reflection signal is mainly left-handed. With the rise of altitude angle, the left-handed signal shows a logarithmic upward trend [1], while the reflection signal at low altitude angle is mainly right-handed circular polarization. Therefore, the general geodetic receiver can receive the reflection signal at low altitude angle. The GNSS remote sensing can be studied by using the general geodetic GNSS receiver. The interference between reflected signal and direct signal is embodied in GNSS SNR data. The technology of GNSS-MR based on SNR has gradually developed and become a new research hotspot. GNSS-MR can retrieve the surface environment (soil moisture, snow depth, snow equivalent, vegetation, tides, etc.) [2]. Anderson [3] first proposed the GNSS water level remote sensing method, which proved the validity of the method, and monitored the sea level in San Diego and California, and compared it with the data of tide gauge stations. The RMS is 12 cm. Larson and Lofgren [4] officially proposed to use ordinary geodetic receivers to monitor sea level for the first time. Experiments were carried out in Friday Harbor, Sweden and the United States. Compared with the measured data of tidal stations, the accuracy of the two methods was 5 cm and 10 cm, respectively. The correlation coefficients were 0.97. Larson [5] conducted a year-long data analysis of GPS stations near Kachemak Bay, Alaska, in 2013, and put forward a dynamic correction model, which corrects the height of the reflective surface of the dynamic change, and obtains that the variation of the daily sea level is more than 7 m. The accuracy of the daily average sea level is about 2.3 cm by comparing the tidal dynamic change inverted by GPS with the data of the traditional tide station at 30 km. Lofgren [6] analyzed the data of five GPS stations, and used elevation angle data below 5°, obtained high accuracy of sea level inversion. RMS ranged from 6.2 cm to 43 cm, and correlation coefficient ranged from 0.89 to 0.99. Santamaría-Gómez [7] uses L1 and L2 bands at the same time, carries out tidal inversion for eight tide stations with GPS joint based on SNR data, combining traditional tide gauge data in the process of processing data. The results obtained are in good agreement with the tide station record, and the RMS reaches 3 cm or even higher. Larson [10] analyzed the data of Friday Harbor's GPS receiver for 10 years, and added dynamic correction of reflector and low elevation angle refraction correction [11]. The accuracy of single GPS water level estimation is 12 cm, the accuracy of daily average water level is 2 cm, and the accuracy of monthly average water level is 1.3 cm.

In recent years, the inversion of water level by GNSS-MR technology in China is in the initial stage. Zhang [8] monitored the tide at SC02 station on the Bank of Friday Harbor in Washington, USA, based on GNSS-MR technology. Compared with the data of tidal station 359 m apart, the inversion accuracy can reach 10 cm, and the correlation coefficient is better than 0.98. Xi [9] has used the GNSS-MR technology to

monitor the water level of the reservoir for one year. Compared with the measured data, the correlation coefficients were above 0.93. Based on the deformation monitoring system of Shuangwangcheng Reservoir Dam, this paper systematically expounds the theoretical method of GNSS-MR inversion technology. The GNSS receiver installed on the dam is used to monitor the water level of the reservoir, which can be used as a reference for the supplement of the water level monitoring system.

2 Principle of Altitude Difference Inversion of GNSS-MR

The generation of multipath is very complicated. Generally, the satellite data of low elevation angle is shielded by setting cutoff angle. This suppression method will inevitably reduce the accuracy of the elevation direction. The GNSS-MR technology can make full use of the discarded low elevation data to the altitude inversion, but the altitude information could not help to improve the accuracy of results of the receiver's elevation direction. It can be well applied to special scenes with height changes with time, such as water level changes and snow depth monitoring. We can basically determine most of the multipath sources, such as monitoring stations located on the shores of wide waters. As shown in Fig. 1, in the waters, the received GNSS signals are synthesized by interference from direct signals and reflected signals which through the water surface.

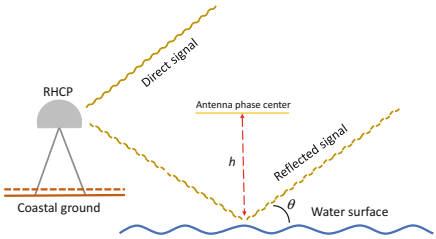
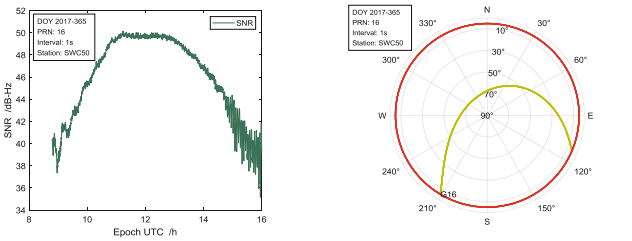


Fig. 1. GNSS-MR water level measurements

In the Fig. 1, h is the distance from the center of the antenna phase to the water surface. In the small range around the station, the signals from the same satellite that are emitted to the water surface and the receiver are parallel to each other. At the same time, the satellite elevation angles in the antenna phase center and the water surface reflection point are equal, θ is the satellite elevation angle at the center of the antenna phase. And the difference between the direct signal and the reflected signal can be easily derived, as shown in Eq. 1.

$$\partial = 2h\sin(\theta) \tag{1}$$

Figure 2(a) is a timing diagram of the PRN16 signal-to-noise ratio of the SWC50 station of Shuangwangcheng Reservoir in DOY365 in 2017.



(a) Signal-to-Noise Ratio Timing Chart **(b) SWC50 Site sky plot of GPS PRN 16**

Fig. 2. SNR and sky plot of GPS PRN 16

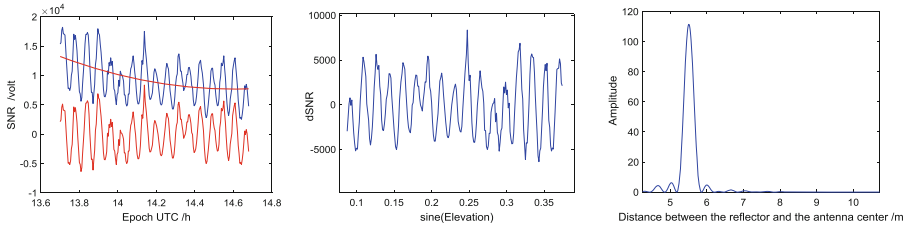
Signal-to-noise ratio is an indicator to evaluate the quality of signals received by the receiver, the larger the SNR value, the better the signal quality, and the positioning accuracy could be improved. The signals of low satellite elevation angle are affected by complex multipath and the SNR is low, showing the variation trend of change over time as shown in Fig. 2(a). In since 9 o 'clock in the morning, receiving the signals with the low angle of satellite elevation angle, as the satellite gradually rises, the signal-to-noise ratio becomes higher and higher, and the SNR fluctuation is more stable relative to the low elevation angle. As the satellite glides down over the sky, the SNR begin reducing and the fluctuations expand accordingly, compared with the SNR value of the low satellite elevation angle at the rising time, and it fluctuates faster and more regularly during the falling time, because the satellite rises in the southwest direction of the antenna. As can be seen from Fig. 2(b), at the rising time of the satellite, the signal received by the antenna is an interference signal composed of reflected and direct signals from the complex surrounding environment on the ground, so the signal-to-noise ratio is irregular and not suitable for inverting as a data source for GNSS - MR technology. During the descent of the satellite, the satellite direction is in the southeast direction of the antenna, which is the reservoir. Due to the low altitude angle, it can be considered that the multipath components in the composite signal are basically signals reflected through the water surface, and the received composite signal is formed by interference between the direct signal and the reflected signal. The principle of GNSS-MR is precisely based on this characteristic, and the sequence of signal-to-noise ratio de-trend items can be expressed as Eq. 2 [12].

$$dSNR = A\cos(2\pi ft + \varphi) \tag{2}$$

where $f = 2h/\lambda_L, t = \sin(\theta)$, φ is the phase value of dSNR. It can be seen from Eq. 2 that dSNR hides the altitude difference information between the reflecting surface and the antenna phase center. The main trend of SNR timing chart is an arch, close to a quadratic parabola. After the SNR value is linearly transformed by using $SNR' = 10^{SNR/10}$, the arc data needs to be extracted according to azimuth angle and satellite elevation angle. Then the useful arc could be fitted by simple quadratic terms, which can remove the trend items better, as shown in Fig. 3(a).

The obtained dSNR sequence needs to be matched with the elevation angle and resampled. The elevation angle value can be obtained through navigation documents or

precise ephemeris. As shown in Fig. 3(b), the horizontal axis is the sine value of the elevation angle at the corresponding epoch time, and vertical axis is dSNR. Then the L-S spectrum analysis method is used to analyze the spectrum of the waveform. The results are shown in Fig. 3(c). The altitude difference of the antenna from the reflecting surface at different times is obtained, so that the altitude change of the reflecting surface can be monitored.



(a) Trend item removal from SNR (b) Resampling dSNR (c) L-S spectrum analysis

Fig. 3. dSNR resampling and spectrum analysis

GNSS - MR technology takes advantage of the characteristics that the changes of dSNR with low elevation angle is similar with cosine curve to carry out further research. It can extract the parameters of the earth's surface in Eq. 2, such as vegetation, snow depth, tide, soil humidity, etc.

3 Example Analysis of Monitoring Water Level by GNSS-MR

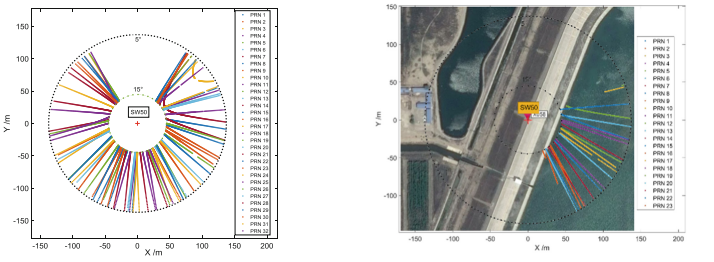
In order to verify the validity of the principle of altitude difference inversion by GNSS-MR, the dam stability monitoring system of Shuangwangcheng Reservoir is selected as an example, as shown in Fig. 4.

(a) Distribution of Shuangwangcheng Reservoir Station (b) Site Map of Monitoring Station

Fig. 4. Data of Shuangwangcheng Reservoir

Shuangwangcheng Reservoir is located in ShouGuang City, WeiFang City, ShanDong Province, with a total length of 9.636 km along the dam axis. The height of dam is 12.5 m, and the maximum reservoir capacity of 61.5 million cubic meters. It is an important reservoir for the Main Line Project of the East Route of the South-to-North Water Transfer Project. In order to monitor the stability of the reservoir dam, three monitoring stations and one reference station have been set up around the dam, as shown in Fig. 4(a), of which SWC43 is the reference station, close to 300 m away from the dam, and the other three stations are all monitoring stations, located on the dam crest respectively, with protective covers installed on the antennas, as shown in Fig. 4 (b). Since the establishment of the dam settlement monitoring system in April 2017. By analyzing data for up to one year, the results indicate that three points are in different degrees of instability. Especially, the SWC53 point is in a continuous linear settlement period, because the dam is in its initial stage, the settlement reaches 20 mm per year. The other two stations have different degrees of settlement and rebound, but they gradually become stable.

There is a water level monitoring station near the SWC50 station. During the dry season, the water level is about 5 m, the GNSS antenna is about 45 m from the water-land boundary, and the water level can reach 11.5 m when there is plenty of rain. This paper compares the water level estimated by GNSS-MR inversion technique with the records of the water level station to evaluate the accuracy. This study collected the observation data of three monitoring stations from May 1, 2017 to December 31, 2017 and the water level data of the water level monitoring stations. The sampling rate of observation data of GNSS receiver is 15 s, and the sampling rate of data of water level monitoring station is once a day.



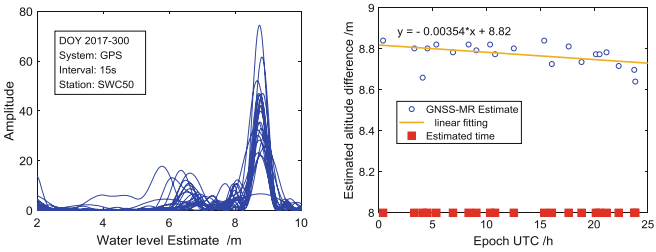
(a) GPS signal reflection point trajectory (b) Trajectory of reservoir directional reflection point

Fig. 5. Reflection point trajectory

Figure 5(a) shows the trajectory of the GPS satellite ground reflection point received by the receiver at day NO. 300 of year 2017. Figure 5(b) shows the result of arc screening according to the vertical height of the GNSS antenna from the water surface and the azimuth of the antenna where the water area is located. At this time, the height of the antenna from the water surface is 8.7 m, the azimuth range is 70°–155°, and the altitude range is from 5 to 15°, as shown in Fig. 5(b), emerging two concentric circles. There are 25 GPS satellite arcs in this azimuth angle and altitude angle range all

day. After matching satellite altitude angles, resampling to obtain dSNR waveforms, and performing spectrum analysis, as shown in Fig. 6, 25 estimated results can be obtained.

Figure 6(a) shows the results of the L-S spectrum analysis of all valid GPS satellite arcs at day NO. 300 of year 2017. The horizontal axis directly converts the frequency values into altitude differences, and the altitude values corresponding to the peak amplitudes in the figure are concentrated around 8.7 m, of which about 6.5 m are concentrated with some smaller secondary amplitude peaks, which can be considered as noise peaks.



(a) L-S spectrum analysis **(b) Inversion time series diagram of water level**

Fig. 6. 300-day inversion results of SWC50

Figure 6(b) shows 25 inversion results for the whole day. the solid line is a curve fitted according to the inversion results. It is obvious that the trend is descending and the rate is about 9 cm/d. Combined with the results of water level recorded by water level station, the water level was rising during this period, and the rising rate of water level is 7 cm/d, both of them have good consistency in qualitative and quantitative aspects. From the red rectangle mark on the horizontal axis of Fig. 6(b), only GPS satellites are used as the data source for inversion, the time resolution can be close to one result per hour. If other constellations are considered, the time resolution will be higher and the accuracy will be improved.

Table 1 is the result of all valid GPS satellite arc estimated at day NO.300 of year 2017, with a total of 25 inversion results, and the standard deviation is 0.056 m. The second column is the vertical height between the water surface and the antenna phase center. It is known that the geodetic height of the antenna of SCW50 station is 14.204 m, solved by direct signal, and the geodetic height of the water surface at the same time can be obtained by subtracting the inversion altitude difference from the geodetic height of the antenna, as shown in the third column. The fourth column is the record of the reservoir water level monitoring station. Since the elevation reference of the water level station is different from the reference of elevation calculated by the direct signal, there will be a systematic deviation. The accuracy can be evaluated by the difference between the two water level data.

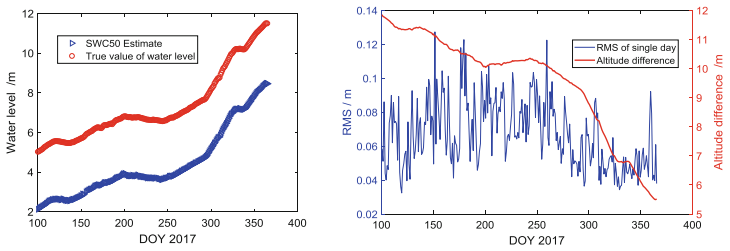
Figure 7 shows the daily average water level inversion result of SWC50 station for day No. 100 to 365 of year 2017. The records of the water level station are marked by red circle, which can be taken as the true value, and the blue triangle marks the daily

Table 1. 300 day of 2017 GPS inversion results

UTC/h SWC50 14.2038m	Altitude difference	Inversion of water level	Water level station data
0.392	8.840	5.363	**8.18**
3.298	8.802	5.402	
4.073	8.660	5.544	
4.517	8.802	5.402	
5.340	8.821	5.382	
6.890	8.783	5.421	
8.343	8.821	5.382	
9.063	8.793	5.411	
10.342	8.821	5.382	
10.802	8.774	5.430	
12.544	8.802	5.402	
15.369	8.840	5.363	
16.079	8.726	5.478	
17.640	8.812	5.392	
18.838	8.736	5.468	
20.173	8.774	5.430	
20.517	8.774	5.430	
21.158	8.783	5.421	
22.292	8.717	5.487	
23.738	8.698	5.506	
23.833	8.641	5.563	

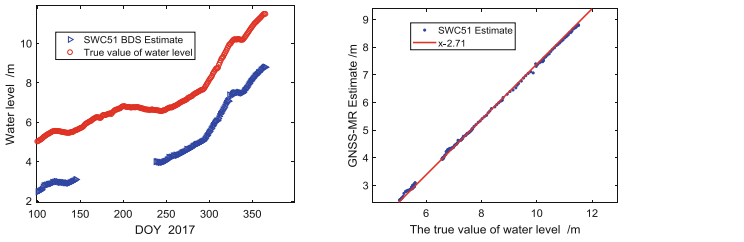
average result estimated by GPS-MR. It is obvious that the general trend of the two curves in Fig. 7(a) is completely the same, with a systematic deviation between them. By analyzing the difference, the deviation between the two elevation is about 2.72 m, and the RMS of the difference on the same day is 0.046 m, and the correlation coefficient is 0.9975. Figure 8(b) shows the relationship between the standard deviation of the daily inverted altitude difference and the vertical altitude difference, with the horizontal axis being the day of year, the left vertical axis being the standard deviation, and the right vertical axis being the vertical altitude difference. As the vertical altitude difference decreases, the standard deviation decreases gradually. This is because when the water surface is higher, the vertical distance between the antenna and the water surface decreases, the water level rises, the horizontal distance between the monitoring station and the boundary line of the water level will decrease, the area of the waters will expand, so that reflected signal received by the antenna will be stronger, and the interference effect will be more obvious. Therefore, the vertical altitude difference accuracy of the high water level inversion will be higher and more stable.

In the same calculation step, the Beidou signal can also be used for reservoir water level inversion, as shown in Fig. 9.



(a) Estimated and True Water Level of SWC50 (b) The relationship between accuracy of inversion water level and height difference

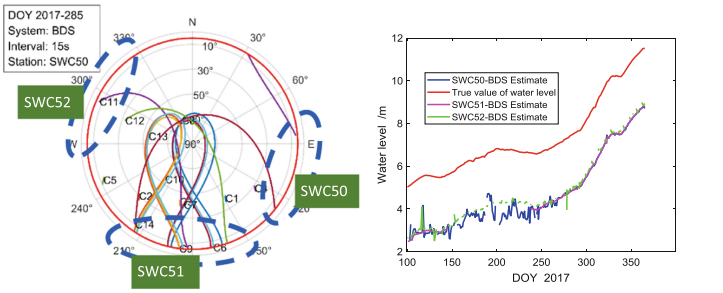
Fig. 7. Monitoring and analysis of water level at SWC50



(a) Estimated and True Water Level of SWC51 (b) Analysis of correlation between inverted water level and true value

Fig. 8. Water level estimation of BeiDou at SWC51

Figure 8(a) is the result of the water level inversion of monitoring station sw51 using the signal-to-noise ratio of Beidou signal B1C, with the blue triangle marked as the daily average water level and the red marked as the water level of the water level station. Since the data of day of year from 146 to 236 are missing from the middle of SWC51, there is no inversion value. Due to the total number of Beidou satellites, the daily average inversion results would be relatively poor, with the average difference between the two being 2.71 m, the standard deviation of the difference being 0.061 m and the correlation coefficient being 0.999, as shown in Fig. 8(b). Compared with the SWC51 monitoring point, the accuracy of SWC50 and SWC52 is much worse because of the special design of Beidou constellation, as shown in Fig. 9(b), Since the second generation of Beidou satellites are basically IGSO, the north-south direction of the chart in the Asia - Pacific region is in the shape of a figure of '8', there will be more Beidou satellites available when the water area is located in the south or north of the monitoring station, such as SWC51 station, which is located in the north dam of the reservoir, and more satellites moving in the north-south direction can be received at this time. On the contrary, the SWC50 and SWC52 monitoring points can only receive a small part of the signals reflected from the water surface to the antenna at a low altitude angle, especially SWC52. Due to the high latitude of the reservoir, there are very few effective satellite arcs every day in the whole process. The blue dotted ellipse marked in the figure indicate the valid arc that can be received by the station, and it is obvious that the SWC51 is more dominant.



(a) Analysis of BeiDou satellite sky plot (b) Inversion results of three stations

Fig. 9. Water level estimation of SWC50, SW5C1, SWC52 by BeiDou

Figure 9(b) shows the results of water level inversion using BeiDou system. The red line is the result of water level station measurement, the blue one is the result of water level inversion at SWC50, and the pink one and the green one are the results of SWC51 and SWC52, respectively. Obviously, at low water level, the inversion results are poor, and the accuracy of SWC51 station is the most stable and the highest.

4 Conclusion

By introducing the principle and inversion process of GNSS-MR technology and combining with the dam deformation monitoring system of Shuangwangcheng Reservoir. The water level of Shuangwangcheng Reservoir is inverted by using GPS and BeiDou system signals, which is compared with the measured data of the water level station. For the L1 band of GPS system, the standard deviation between the daily average water level retrieved from SWC50 monitoring station and the records obtained from water level station data is 0.046 m, and the average of the two is 2.71 m which due to the inconsistency between the measured data and the elevation results estimated by GNSS-MR. The standard deviation between the result of SWC51 monitoring point inversion and the measured data is 0.041 m, the mean value of the difference is 2.70 m, the standard deviation between the result of SWC52 monitoring point inversion and the measured data is 0.038 m, and the mean value of the difference is 2.72 m; Due to objective reasons, the S2 data integrity is poor in the data received this time. In the process of spectrum analysis, the data results show relatively large errors. The accuracy of SWC51 is 0.089 m. In comparison, the accuracy of BeiDou system inversion is low precision, the effective arc obtained every day is less than that of GPS because of the small number satellites, the inversion precision of SWC51 monitoring station is the highest, the standard deviation of the inversion result compared with the true value is 0.061 m, and the inversion precision of SWC50 and SWC52 points is very poor, with the RMS of 0.253 m and 0.179 m, respectively, what's more, the obtained data are very unstable and the time resolution is very low.

The research on GNSS-MR technology is more and more extensive. It can estimate the soil moisture, vegetation cover and growth, snowfall, ocean tide change and so on.

Especially, it provides a new method for the global water cycle research combined with the high-density GNSS monitoring network. In this paper, the inversion application of GNSS-MR on altitude difference is further expanded. The data of GPS and BeiDou are used to monitor the water level of inland reservoirs, which can be an effective supplement to the water level monitoring system of reservoirs.

Acknowledgements. This study is supported by National Natural Science Foundation of China (Nos. 41830110, 41804005, 41474001, 41604018), Natural Science Foundation of Jiangsu Province, China (No. BK20170869), Fundamental Research Funds for the Central Universities (Nos. 2019B17414, 2019B17514) and Major scientific and technological issues of Jiangxi Water Resources Department (No. KT201322).

References

1. Jin YQ, Zhang JR, Zhao RY (1992) Remote sensing of sea ice by multi-frequency microwave radiometers and numerical modeling of radiative transfer. Acta Ocean Sin 7 (1):33–38
2. Bilich A, Larson KM (2007) Mapping the GPS multipath environment using the signal-to-noise ratio (SNR). Radio Sci 42:RS6003. https://doi.org/10.1029/2007rs003652
3. Anderson KD (2000) Determination of water level and tides using interferometric observations of GPS signals. J Atmos Ocean Technol 17(8):1118–1127
4. Larson KM, et al (2012) Coastal sea level measurements using a single geodetic GPS receiver. J Adv Space Res. http://dx.doi.org/10.1016/j.asr.2012.04.017
5. 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
6. Lofgren JS, Haas R, Scherneck HG (2014) Sea level time series and ocean tide analysis from multipath signals at five GPS sites in different parts of the world. J Geodyn 80:66–80
7. Santamaría-Gómez A, Watson C, Gravelle M et al (2015) Levelling co-located GNSS and tide gauge stations using GNSS reflectometry. J Geodyn 89(3):241–258
8. Zhang S, Nan Y, Li Z et al (2016) Analysis of tide variation monitored by GNSS-MR. Acta Geod Cartogr Sin 45:1042–1049
9. Xi R, Zhou X, Jiang W et al (2018) Simultaneous estimation of dam displacements and reservoir level variation from GPS measurements. Measurement S0263224118302124
10. Larson KM, Ray RD, Williams SDP (2017) A 10-year comparison of water levels measured with a geodetic GPS receiver versus a conventional tide gauge. J Atmos Ocean Technol 34 (2):295–307
11. Santamaría-Gómez A, Watson C (2016) Remote leveling of tide gauges using GNSS reflectometry: case study at Spring Bay, Australia. GPS Solut 21(2):1–9
12. Larson KM, Small EE (2017) Estimation of snow depth using L1 GPS signal-to-noise ratio data. IEEE J Sel Top Appl Earth Obs Remote Sens 9(10):4802–4808

Accuracy Analysis of GPT2/GPT2w Models for SLR-Based Satellite Orbits Validation

Tao Li[1], Lei Wang[1(✉)], Ruizhi Chen[1,2], Beizhen Xu[1], and Xinxin Zhang[1]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China
lei.wang@whu.edu.cn
[2] Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan, China

Abstract. The meteorological parameters such as the temperature and the pressure are crucial for tropospheric delay correction of Satellite Laser Ranging (SLR). Considering that the meteorological data of the SLR observations may be missing, this paper selects two common global meteorological models, GPT2 and GPT2w, to evaluate the accuracy of the meteorological model and apply them to SLR-based satellite orbit validation. In this study, we validated the accuracy of the empirical models with the observed meteorological data collected from the SLR stations. The results show that the RMSE of the pressure and temperature generated by GPT2 model is 5.61 hPa and 4.90 K, respectively. The RMSE of the pressure and temperature derived from the GPT2w is 5.58 hPa and 4.83 K, respectively. Then the SLR residuals are compared by applying the empirical meteorological models and the real observations. The results indicate that the RMS of the SLR residuals with the real meteorological data is 3.22 cm, and the RMS of the SLR residuals calculated with the GPT2 and GPT2w models are 3.45 cm and 3.42 cm, respectively. The GPT2w model performs slightly better than the GPT2 model in SLR data processing. It concludes that replacing the real meteorological data by the empirical models is feasible for SLR data processing.

Keywords: Satellite Laser Ranging (SLR) · Tropospheric delay · GPT2 · GPT2w · Satellite orbit validation

1 Introduction

Satellite Laser Ranging (SLR) technology is the most accurate technique for satellite orbit validation, whose accuracy can reach sub-centimeter [1]. Tropospheric propagation delay is one of the major error sources in satellite laser ranging data processing, and its impact reaches up to tens of meters [2]. In order to precisely correct the troposphere delay, most empirical models, such as the Mendes-Pavlis [3] model, requires the meteorological data, including the air pressure, the temperature and the relative humidity as the model input. Therefore, most SLR observatories are equipped with expensive meteorological observing systems. GPT series of meteorological models have been widely recognized and applied in GNSS precise data processing.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 174–183, 2019.
https://doi.org/10.1007/978-981-13-7751-8_18

Yao et al. [4] performed accuracy testing and analysis on GPT2 models in global scale. Hua et al. [5] evaluated the accuracy of GPT2w model's weighted average temperature and water vapor direct reduction rate. The GPT2 and GPT2w models have high accuracy on a global scale, but there are few researchers applying them in SLR data processing. Therefore, this paper validated the accuracy of GPT2 and GPT2w models with the meteorological data collected by SLR observatories, then analyzed the effects of temperature and pressure on the total tropospheric delay in laser propagation. Finally, the two models are applied to the SLR-based satellite orbit validation, and compared with the observed meteorological data.

2 Global Empirical Meteorological Models

Due to the fact that the measured meteorological data is often not available in GNSS observation, Böhm et al. [6] proposed the GPT model that requires only the station coordinates and the day of the year and offers temperature and pressure. Lagler et al. [7] proposed GPT2 on the basis of GPT, increased a semi-annual period term, and replaced the 9th-order spherical harmonic function with 5° grid, which improved the horizontal resolution of the model. GPT2w proposed by Böhm et al. [8] is based on 1° grid meteorological data. The latest GPT3 model [9] has the same strategy for temperature and pressure calculation as GPT2w. Therefore, we focus on the GPT2 and GPT2w model in this study.

2.1 GPT2 Model

The GPT2 model is based on global temperature, pressure and relative humidity data from ERA-Interim for the decade from 2001 to 2010. The empirical meteorological model with average, annual and semi-annual terms has a grid resolution of 5° × 5° [7]. The output of GPT2 includes air pressure, air temperature, water vapor pressure and two projection function parameters. The specific calculation formulas of each output parameter are the same [7]:

$$r(t) = A_0 + A_1 \cos(\frac{doy}{365.25}2\pi) + B_1 \sin(\frac{doy}{365.25}2\pi) + A_2 \cos(\frac{doy}{365.25}4\pi) + B_2 \sin(\frac{doy}{365.25}4\pi) \tag{1}$$

where $r(t)$ represents pressure or temperature, A_0 is the mean value, A_1 and B_1 are the annual amplitudes, A_2 and B_2 are the semi-annual amplitudes of the coefficient, *doy* means the day of the year.

2.2 GPT2w Model

The main calculation formula of the GPT2w model is the same as that of GPT2. It has two grid resolutions: 1° × 1° and 5° × 5°. It provides not only air pressure, air temperature, water vapor pressure and dry delay projection parameters and wet delay projection parameters in VMF, but also temperature lapse rate and water vapor pressure decay index [10]. The 1° × 1° GPT2w model is used in this article.

2.3 Comparison of the Empirical Model and Observed Meteorological Data

In order to study the applicability of GPT2 and GPT2w models in SLR data processing, the differences between the meteorological models and the measured meteorological data of the SLR observatories need to be first analyzed. Most SLR stations around the world are equipped with MET3 meteorological instruments for obtaining atmospheric pressure, temperature and relative humidity on the surface of the station. The accuracy of MET3 is ±0.5 °K for temperature, ±0.08 hPa for pressure, and ±2% for relative humidity [11]. Therefore, the meteorological data obtained by the SLR observatory can be used as reference value for evaluating the meteorological model. This paper uses the meteorological data collected by 16 SLR stations distributed around the world for one year to test the accuracy of the GPT2 and GPT2w models.

The average bias and root mean square error are used as the criteria to evaluate the accuracy and stability of GPT2 and GPT2w. The calculation formula is shown in Eq. (2):

$$\begin{cases} Bias = \frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{r(t)}_i^{GPT} - \mathrm{r(t)}_i^{real}\right) \\ RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(r(t)_i^{GPT} - r(t)_i^{real}\right)^2} \end{cases} \tag{2}$$

where $\mathrm{r(t)}_i^{GPT}$ represents temperature or air pressure offered by GPT2 or GPT2w, $\mathrm{r(t)}_i^{real}$ is the observed meteorology data, N is the number of observations.

Table 1 shows the results of the overall accuracy result of GPT2 and GPT2w. It can be seen that there is certain systematic bias between the pressure provided by the empirical model and the real pressure. The average deviations of GPT2 and GPT2w are −1.38 hPa and −1.32 hPa, respectively. The RMSE values of GPT2 and GPT2w air pressures are 5.61 hPa and 5.58 hPa, respectively. The average temperature deviations provided by GPT2 and GPT2w are near zero, and the RMSE values are 4.90 K and 4.83 K, respectively. Overall, the accuracy of GPT2w is slightly better than GPT2.

Table 1. Accuracy results of pressure and temperature

Model	Pressure/hPa		Temperature/K		Number
	Bias	RMSE	Bias	RMSE	
GPT2	−1.38	5.61	0.08	4.90	15257
GPT2w	−1.32	5.58	0.43	4.83	

2.4 Analysis of Temporal and Spatial Characteristics of Meteorological Model Residuals

2.4.1 Temporal Characteristic Analysis

In order to analyze the temporal characteristics of meteorological models, we select four SLR stations with different latitudes in the northern hemisphere, and calculate the root mean square error of air pressure and temperature in spring, summer, autumn and winter according to formula (2). As Figs. 1 and 2 show, the pressures and temperatures vary widely in different seasons, but the seasonal variations at different latitudes are also different.

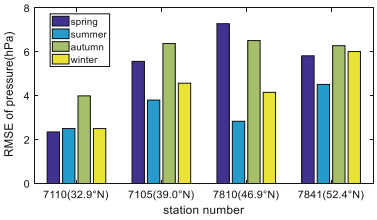


Fig. 1. Seasonal changes for air pressure accuracy of GPT2

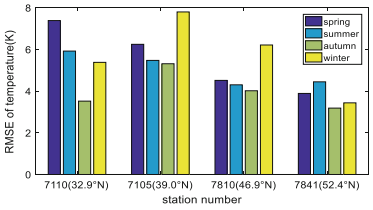


Fig. 2. Seasonal changes for temperature accuracy of GPT2

In order to more intuitively show the difference between the meteorological model and the measured data, as shown in Figs. 3 and 4, this paper selects four stations with different latitudes: 7124 (17.6°S), 7090 (south latitude 29.0°), 7105 (39.0°N) and 7841 (52.4°N), and gives the observed values and model values of air pressure and temperature. It can be seen that GPT2 and GPT2w have some differences in low latitudes, and as the latitude increases, the pressure and temperature data provided by the two models are quite similar. The GPT2 and GPT2w can only provide daily average temperature and daily average pressure, which means the time resolution of models have to be improved.

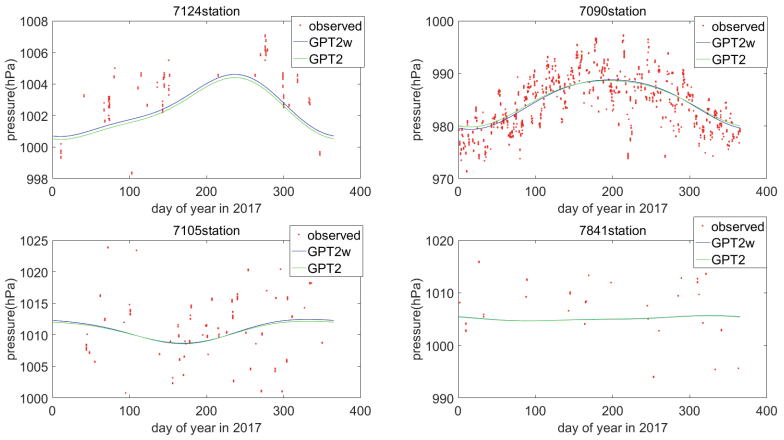


Fig. 3. Pressure at four different latitude SLR stations in 2017, the red dot is the observed value, and the blue and green solid lines represent GPT2w and GPT2, respectively

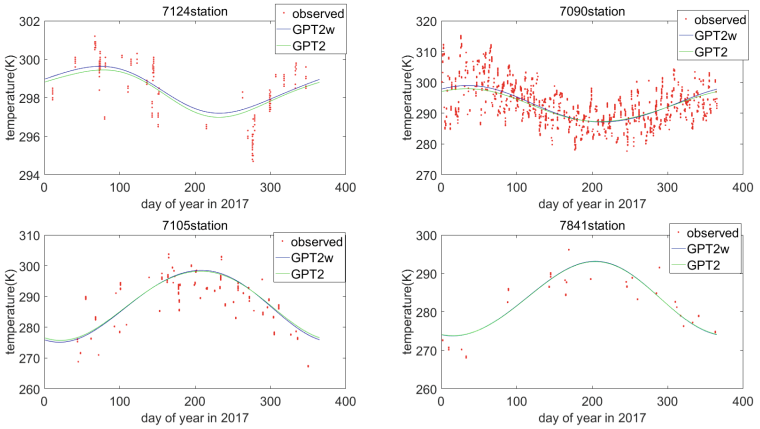


Fig. 4. Temperature at four different latitude SLR stations in 2017, the red dot is the observed value, and the blue and green solid lines represent GPT2w and GPT2, respectively

2.4.2 Spatial Characteristics Analysis

Figures 5 and 6 are the spatial distributions of the root mean square errors of air pressure and temperature of GPT2 and GPT2w, respectively. It can be seen that in terms of air pressure, the RMSE is greater than 5 hPa at those stations such as 7840, 7841, 1879, 1868, whose latitude is higher than 45°, and the RMSE is less than 4 hPa at low latitude stations such as 7124 and 7501. In terms of temperature, the accuracy of the meteorological model is closely related to the climate type of the station. The 7501 station is located in Johannesburg, South Africa, and belongs to the savanna climate, whose temperature RMSE of GPT2 and GPT2w is 7.47 K and 7.81 K, respectively.

The 1868 station and the 7237 station belong to the temperate continental climate, and their temperature RMSE are in the range of 5.5–6.3 K. At other stations belonging to the Mediterranean climate, such as 7840, 7841, 7810, 7839, 7941, and 7825, the root mean square errors are in the range of 3.7–4.8 K.

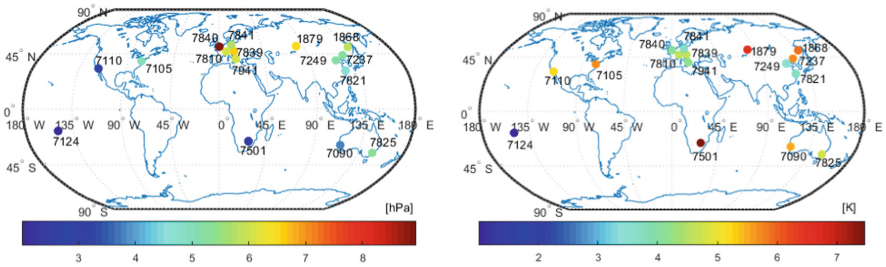


Fig. 5. RMSE of the GPT2 pressure (left), temperature (right) at each SLR station

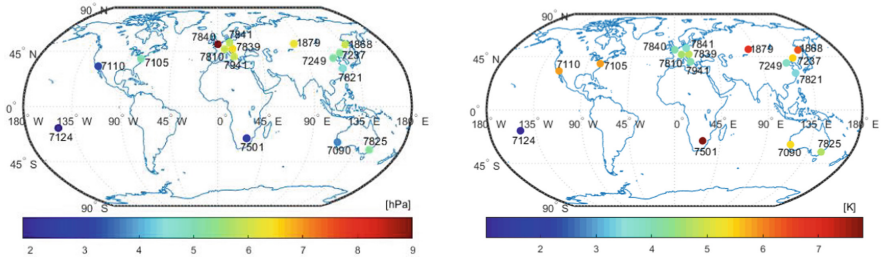


Fig. 6. RMSE of the GPT2w pressure (left), temperature (right) at each SLR station

3 Application of Meteorological Model in SLR

3.1 Influence of Pressure and Temperature on SLR

The SLR is an optical ranging technique. The influence of the troposphere refraction on the SLR reaches several meters, while the influence of the ionosphere is basically negligible [12]. The current tropospheric delay model recommended by IERS Conventions (2010) is the Mendes-Pavlis model, which determines the zenith direction delay from the station's air pressure, temperature, and relative humidity, and then projects it by satellite elevation angle [13].

In this paper, we have studied the variation of total tropospheric delay calculated by Mendes-Pavlis model in different air pressure, different temperature and different humidity by controlling variable. As shown in Fig. 7, at the 7090 station (latitude: 29.05°S, geodetic height: 241.34 m), the air pressure and relative humidity are proportional to the total tropospheric delay. To be specific, the total tropospheric delay varies by 1 mm as the air pressure changes by 0.25 hPa, whereas 10% variation of the

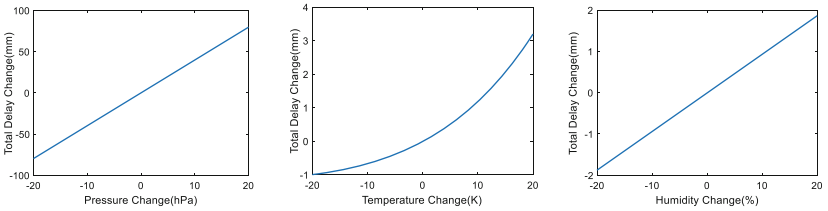


Fig. 7. Relationship between air pressure, temperature, relative humidity and total tropospheric delay, respectively

relative humidity leads to 1 mm change of the total tropospheric delay. Unlike pressure and relative humidity, the effect of temperature on the total tropospheric delay is nonlinear. The temperature drops by 20 K, meanwhile the total tropospheric delay is reduced by 1 mm. However, if the temperature rises 20 K, the total tropospheric delay will increase 3 mm.

3.2 Meteorological Model for SLR Validation of Orbit Products

In order to verify the feasibility of applying the GPT2 and GPT2w models to the SLR validation of satellite orbits, we select the SLR standard point data of GLONASS128 satellite (PRN number R03) provided by ILRS from January 1, 2017 to December 31, 2017, and the precise orbit products of corresponding satellites provided by IGS in corresponding time period. The precision orbit accuracy of GLONASS published by IGS official website is 3 cm. Considering the influence of various errors, according to the 3σ principle, the observations of the ranging residual greater than 10 cm are eliminated. In this paper, the temperature and pressure data in the SLR standard point data file are replaced by the temperature and pressure data provided by the two meteorological models, and the water vapor pressure provided by the models is converted into relative humidity.

We calculate the laser ranging residual for one year based on Bernese 5.2 software. Multiple experiments proved that the relative humidity error has little effect on the laser ranging residual. Figure 5 shows the influence of the deviation between air pressure, temperature obtained in the GPT2 model and measured data of the 7090 station on the laser ranging residual. The color value of every point represents the difference between the real laser ranging residual and the GPT2 model's laser ranging residual. It can be seen from Fig. 8 that the influence of the air pressure on the ranging residual is dominant. When the air pressure deviation maintains static, the change of the temperature almost has little effect on the laser ranging residual. The pressure deviation of 10 hPa has an effect of 20–40 mm on the laser ranging residual, while the temperature deviation of 10 K has an effect of 1–2 mm on the laser ranging residual.

SLR satellite orbit validation is to compare the difference between the distance from the station to the satellite obtained by laser and microwave measurements [14]. The method for evaluating the accuracy of SLR data check satellite orbit products is to measure the root mean square value of the laser ranging residuals of the same satellite at different times and different stations, and use it as the accuracy index [15]. In this

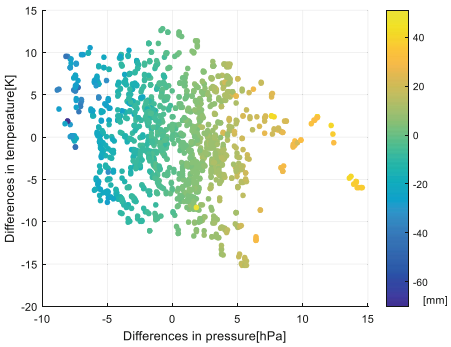


Fig. 8. Influence of pressure and temperature deviation(GPT2-observed) on laser ranging residual at 7090 station

paper, the meteorological data provided by the measured meteorological data, GPT2 and GPT2w are used to calculate the laser residual results of 16 SLR stations around the world to observe the GLONASS128 satellite in 2017. Figure 9 presents the accuracy of the satellite orbit products checked by meteorological models for most stations is comparable to the accuracy of the measured data, and the accuracy of most stations is around 3 cm. The accuracy of the 7237 station and the 1868 station tracking satellite is worse than other stations, and the effective values of the laser ranging residual are 5.84 cm and 5.79 cm, respectively. The empirical model performed the worst at 1879 station, with an accuracy of 2.14 cm for real data, 3.51 cm and 3.43 cm for GPT2 and GPT2w, respectively. Except for the 1879 station, the accuracy of other stations using meteorological models is no more than 1 cm compared to the measurement accuracy. In this paper, the observations with laser residuals greater than 10 cm are regarded as gross deviations. The RMS of the measured meteorological data of all 16 stations to check the satellite orbit is 3.22 cm, the RMS of GPT2 and GPT2w is 3.45 cm and 3.42 cm, respectively. Therefore, when the meteorological data is missing in the SLR standard point data file, the empirical meteorological model can be used to provide temperature and air pressure. Consequently, the utilization of the laser ranging data can be improved.

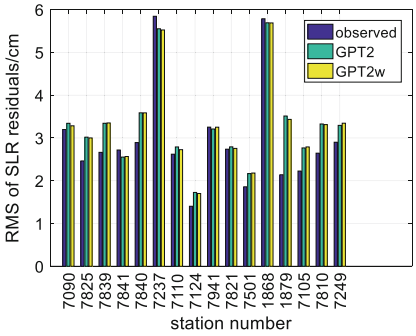


Fig. 9. Statistical results of laser ranging residuals for the GLONASS128 satellite in 2017

4 Conclusion

This paper analyzes the characteristics of the accuracy of GPT2 and GPT2w meteorological models in temporal and spatial distribution, and applies these two meteorological models to SLR for inspecting satellite orbit products.

The RMSE of the pressure and temperature generated by GPT2 model are 5.61 hPa and 4.90 K, respectively. And the RMSE of the pressure and temperature derived from the GPT2w are 5.58 hPa and 4.83 K, respectively. The two meteorological models have obvious seasonal characteristics in terms of time characteristics. In terms of spatial characteristics, the higher the latitude, the lower the accuracy of air pressure. And the accuracy of the temperature and the climate type of the station presents strong correlation.

The meteorological parameters mainly affect the tropospheric delay in the SLR. Pressure changes of 2–3 hpa have an impact on the total delay of 10 mm. 10% humidity variation leads to the 1 mm change of the total tropospheric delay. The effect of temperature changes on the total tropospheric delay is nonlinear. The temperature drops by 20 K, meanwhile the total tropospheric delay is reduced by 1 mm. However, if the temperature rises 20 K, the total tropospheric delay will increase 3 mm.

The accuracy of the GPT2 and GPT2w models applied to the satellite orbit products validation by SLR is comparable to that of measured data. Moreover, the accuracy of 15 stations in the 16 stations is no more than 1 cm compared with the actual measurement. In terms of overall accuracy, the RMS of satellite orbit checked with the measured meteorological data is 3.22 cm, the RMS values of that using GPT2 and GPT2w are 3.45 cm and 3.42 cm, respectively. In the future, we should consider improving the temporal resolution of the meteorological model and the stability of the meteorological model for SLR-based satellite orbits validation.

Acknowledgement. This research is support by the National Natural Science Foundation of China (NSFC 41704002, 91638203), China Postdoc Science Foundation (2017M620337) and the Fundamental Research Funds for the Central Universities. Thanks to IGS, CDDIS, and ILRS for providing data for this article.

References

1. Yang H, Xu T, Sun D (2016) Validation of GPS36 satellite CODE precise orbit with SLR measurements. In: China satellite navigation conference (CSNC) 2016 proceedings, vol 3, Changsha, China
2. Yuan H, Mei H, Huang Y, Rao R (2011) Research on atmospheric refraction correction algorithm and model for satellite laser range-finding. Acta Opt Sin 31(04):31–37
3. Mendes VB (2004) High-accuracy zenith delay prediction at optical wavelengths. Geophys Res Lett 31(14):L14602
4. Yao Y, Cao N, Xu C et al (2015) Accuracy assessment and analysis for GPT2. Acta Geod Cartogr Sin 44(7):726–733
5. Hua Z, Liu L, Liang X (2017) An assessment of GPT2w model and fusion of a troposphere model with in situ data. Geomat Inf Sci Wuhan Univ 42(10):1468–1473

6. Boehm J, Heinkelmann R, Schuh H (2007) Short note: a global model of pressure and temperature for geodetic applications. J Geod 81(10):679–683
7. Lagler K, Schindelegger M, Böhm J et al (2013) GPT2: empirical slant delay model for radio space geodetic techniques. Geophys Res Lett 40(6):1069–1073
8. Johannes B, Gregor M, Schindelegger M et al (2015) Development of an improved empirical model for slant delays in the troposphere (GPT2w). GPS Solut 19(3):433–441
9. Landskron D, BöHm J (2017) VMF3/GPT3: refined discrete and empirical troposphere mapping functions. J Geod 92(4):349–360
10. Zhai S (2018) GNSS tropospheric delay correction and its application. Strategic Support Force Information Engineering University
11. Hulley GC, Pavlis EC (2007) A ray-tracing technique for improving Satellite Laser Ranging atmospheric delay corrections, including the effects of horizontal refractivity gradients. J Geophys Res Solid Earth 112:B06417
12. Petit G, Luzum B (2010) IERS conventions (2010). Iers Techn Note 36:1–95
13. Yang H (2017) GNSS/LEO satellite precision orbit detection based on SLR data. Chang'an University
14. Qin X (2003) Satellite precise orbit determination based on SLR technology. University of Information Engineering, People's Liberation Army
15. Mao Y, Jia X, Sun F (2007) Validation satellite broadcast ephemeris by SLR data. J Geomat Sci Technol 24(4):263–266

Analysis of the Influence of the Superrefraction Effect on the Earth's Neutral Atmospheric Parameters Retrieval

Xian-Sheng Xu[1(✉)] and Zhen-Jie Hong[2]

[1] Ningbo Institute of Technology, Zhejiang University, Ningbo 31500, China
xuxiansheng666@126.com
[2] School of Mathematics and Information Science, Wenzhou University, Zhejiang 325000, China

Abstract. The superrefraction, which can be frequently found in the moist lower troposphere, degrades the accuracy of the retrieved atmospheric parameters from GPS/LEO occultation data. The GPS/LEO occultation data containing superrefraction effect can be simulated by using multiple phase screens technique, with the refractivity model or high resolution ecmPrf profile (echPrf data). Simulation results show that superrefraction effect can lead to refractivity error in the moist lower troposphere, which in some case can be as large as about −9.6%. About 24530 echPrf profiles from DOY(day of year) 151 to DOY 161 in 2007, are divided into A, B, C, D 4 categories. According to the minimum of refractivity gradient of the echPrf data, the minimum of the refractivity gradient of the echPrf data is decreasing from category A to category D. The category C and category D contain the occultation profiles with the superrefraction effect. The refractivity profiles in 4 categories are compared with those of the corresponding COSMIC(Constellation Observing System for Meteorology, Ionosphere and Climate) data. Statistical comparison results show: about 5.6% of the echPrf profiles with superrefraction effect can be found. The mean and standard deviation of the fractional difference of refractivity between COSMIC observation and echPrf data near the ground are about −0.51% and 2.71% for set A, −1.64% and 4.27% for set B, −3.59% and 5.49% for set C, −8.50% and 7.21% for set D. The locations of the superrefraction phenomenon in the tropics are mainly near the following regions: North America (Mexico), South America (the western coast of Peru), North Africa (the Sahara desert), and South Africa on the Atlantic side.

Keywords: COSMIC occultation · Superrefraction phenomenon · Multiple phase screens technique · Open-loop technique

1 Introduction

It is an important issue to use satellite navigation data to explore the Earth's atmospheric parameters. Atmospheric parameters of the Earth can be resulted from inverting GPS (Global Positioning System) signals by employing the GPS/LEO (Low Earth

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 184–192, 2019.
https://doi.org/10.1007/978-981-13-7751-8_19

Orbit) radio occultation (RO) technique. In 2006, a joint Taiwan-U.S. FORMOSAT-3/COSMIC (Constellation Observing System for Meteorology, Ionosphere and Climate) mission was launched and is currently providing 1500–2500 RO daily profiles [1]. With the assumption of spherical symmetry, atmospheric parameters of the Earth, such as bending angles, refractivity, temperature, humidity and pressure can be resulted from inverting GPS phase and amplitude by employing the GPS/LEO RO technique.

The large amount of water vapor in the lower troposphere, especially in the atmospheric boundary layer, results in the refractivity gradient and the retrieval errors. Superrefraction effect and signal tracking error are the main sources of the retrieval error [2]. The GPS signal will bend to the ground, when superrefraction occurs. Under superrefraction condition, Abel integral transform is singular, and thus cannot accurately retrieve the refractivity information. GPS signal, simulated using radiosonde data including superrefraction effect, can be retrieved and analyzed. The results show that superrefraction effect can cause the maximum refractive error about 15% [3]. Xie assumes that the height is a piecewise linear function of the impact parameter in the atmospheric boundary layer [4]. Under certain constraints, the refractivity information inside the atmospheric boundary layer can be obtained. The simulation results show, the bending angle of the same profile corresponds to a continuum of the refractivity. The refractivity obtained by Abel integral transformation inversion, which will cause large negative deviation, is only the minimum value of this continuum [4]. In above method, the height and thickness of the atmospheric boundary layer, which is difficult to obtain accurately, is needed. There are still some defects in the inversion of the GPS/LEO RO data, although the retrieval accuracy is high in the simulation [2]. Superrefraction is a difficult and hot point in the GPS/LEO RO technique.

The atmPrf data, including bending angle, refractivity and dry temperature profiles, is retrieved from the phase data of COSMIC occultation. The superrefraction effect cannot be observed in atmPrf data, because the data are smoothed in the inversion process. The superrefraction effect can be captured by echPrf data [5]. The echPrf is high resolution ECMWF (European Centre for Medium-Range Weather Forecasts) analysis field data. The echPrf field data does not contain COSMIC information. Both atmPrf and echPrf data can be downloaded from website https://www.cosmic.ucar.edu/. The influence of superrefraction effect on neutral atmospheric inversion is discussed, through simulation, inversion, analysis of the GPS/LEO occultations.

2 Superrefraction Effect

In the neutral atmosphere, the refractivity can be expressed as

$$N = (n-1) \times 10^6 = k_1 \frac{P}{T} + k_2 \frac{P_w}{T^2} \tag{1.1}$$

Where $k_1 = 77.6\,\mathrm{K/mb}$ and $k_2 = 3.73 \times 10^5\,\mathrm{K^2/mb}$ are constant. The variables n, P, P_w, T represent the refractive index, total atmospheric pressure (unit mb), atmospheric wet pressure (unit mb) and atmospheric temperature (unit K), respectively.

Take derivative of the refractivity in above equation, with respect to the curvature radius r:

$$\frac{\mathrm{d}N}{\mathrm{d}r} = k_1 \frac{1}{T}\frac{\mathrm{d}P}{\mathrm{d}r} - (k_1 \frac{P}{T^2} + 2k_2 \frac{P_w}{T^3})\frac{\mathrm{d}T}{\mathrm{d}r} + k_2 \frac{1}{T^2}\frac{\mathrm{d}P_w}{\mathrm{d}r} \tag{1.2}$$

The first item on the right side of the equation is caused by atmospheric pressure gradient, in the moist lower troposphere, its value is about $-30\ \mathrm{km}^{-1}$. The second term and the third term are the effect of atmospheric temperature and wet pressure gradient, respectively. The Atlantic trade winds test shows that the impact of the second term is about $-30\ \mathrm{km}^{-1}$, and the third one is about $-120\ \mathrm{km}^{-1}$. Atmospheric wet pressure gradient is the main cause of the superrefraction in the moist lower troposphere.

In the moist troposphere, especially in the subtropical ocean region, the water vapor gradient is inclined to drastic change, leading to a larger refractivity gradient [3]. In an ideal state with the occurrence of superrefraction, the ray travels in a circular orbit around the earth at a fixed height, within and below the superrefractive layer. This means the complete loss of signals within and below the superrefractive layer [4]. As a result, the refractivity profile obtained by inversion of GPS/LEO RO signal is inclined to negative deviation. Subsequent inversion of atmospheric parameters such as temperature, humidity and pressure will also have errors. In order to obtain the atmospheric parameters accurately in the moist lower troposphere, the constraints of background field data are needed. The variational assimilation technique is used to optimize the solution and obtain the atmospheric parameters [6].

3 Simulation

3.1 Multiple Phase Screens Technique

Simulation of RO signals is performed by multiple phase screens technique based on an ideal atmosphere refractivity profile [7]. The atmospheric bending angle profile is obtained by inversion of GPS signals, using the radioholographic (RH) method. The refractivity profile can be obtained from the bending angle profile, using the Abel integral transformation, and compared with the true value. This process can be used to study the phenomenon of the superrefraction, as shown in Fig. 1.

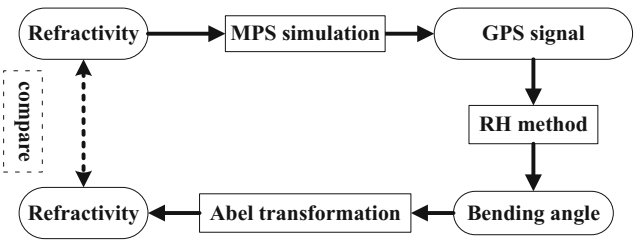


Fig. 1. Simulation with the use of multiple phase screens

Simulation of RO signals is performed by EGOPS (End-to-End Generic Occultation Performance Simulation and Processing System) software based on an ideal atmosphere refractivity profile [8]:

$$N(h) = 315 \exp\left[\frac{-h}{7.35\,\text{km}}\right] + B \exp\left[\frac{-(h - BH)^2}{0.05\,\text{km}^2}\right] \tag{1.3}$$

in which, h denotes height (unit km), constant B and BH determines the strength and location of the multipath or superrefraction. Three different situations: no multipath, multipath, superrefraction, can be simulated when $B = 0$, $0 < B < 37$ and $B \geq 37$, respectively. In simulation, the earth's radius $r_e = 6371\,\text{km}$ and the altitude of the atmosphere $h_{atm} = 130\,\text{km}$.

The total distance traveled by radio waves in the atmosphere is:

$$L_{atm} = 2\sqrt{(r_e + h_{atm})^2 - r_e^2} \approx 2587\,\text{km} \tag{1.4}$$

The distance between the screens is 2 km, which means about 1300 phase screens in the simulations.

3.2 Case Study

In the formula (1.3), $BH = 2\,\text{km}$ and $B = 43$ are taken for simulation. The result is shown in Fig. 2, and its refractivity profiles are shown in Fig. 2(d). The Fig. 2(a) is the corresponding refractivity gradient, in which superrefraction effect can be found at about 2 km. The amplitude of the simulated signal is shown in Fig. 2(b), in which the horizontal axis is occultation time and the vertical axis is signal amplitude. The full spectrum inversion (FSI) method is used to retrieve the bending angle profile, as shown the blue curve in the Fig. 2(c). In Fig. 2(c), the horizontal axis represents the bending angle; the vertical axis represents the impact height, which is defined as the difference between the impact parameter and the curvature radius. The red curve represents the true value, which is obtained from the refractivity model through the Abel integral. As can be seen in Fig. 2(c), the bending angle calculated by the Abel integral tends to infinity at about 3.8 km, which means the singularity appears in the Abel integral. At impact height 3.5 km–4 km, big error can be found between bending angle profile obtained by FSI method and the corresponding true value. Figure 2(d) is the refractivity profile, in which the horizontal and vertical axis are refractivity and height, respectively. Above 2.5 km, there is no superrefraction in the signal, the refractivity obtained by the FSI method is close to the true value. Below about 2.5 km, the superrefraction effect results in the error between the retrieved refractivity and the true value.

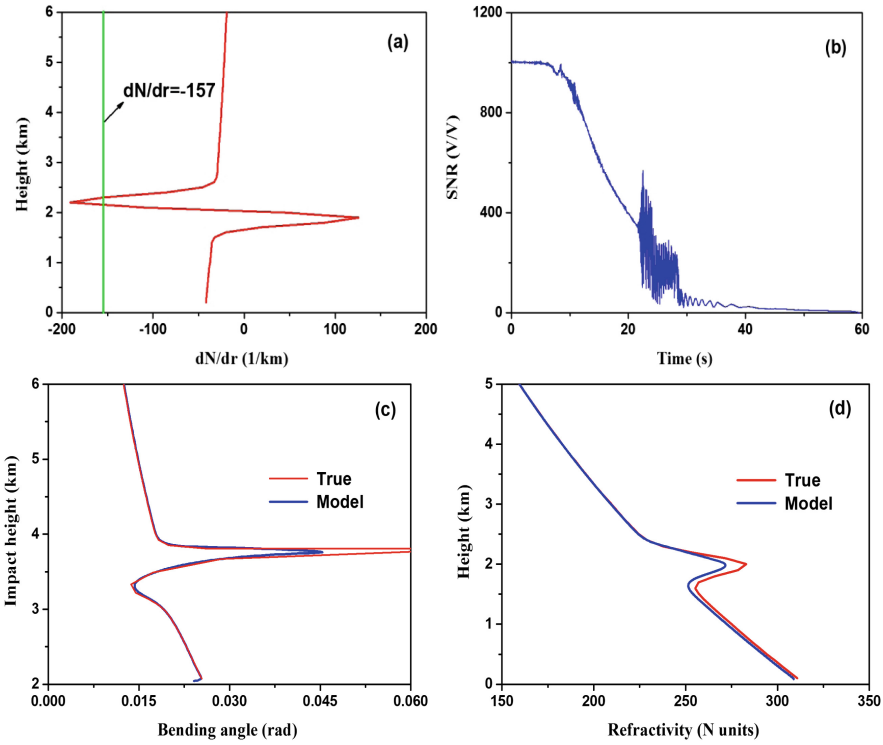


Fig. 2. Simulation of superrefraction phenomenon with refractivity model

Superrefraction effect can also be found in echPrf data, as seen in Fig. 3. The minimum refractivity gradient is about $-300\,\mathrm{km}^{-1}$ at about 1.1 km in Fig. 3(a). Due to the superrefraction effect, the amplitude oscillates strongly at 35 s–45 s, and the average intensity of the amplitude is about 200 V/V. Big difference can be found in both bending angle and refractivity profiles, in Fig. 3(c) and (d). The black dash line denotes bending angle and refractivity profiles obtained from atmPrf data. Big error can be found in the lower troposphere, and the minimum value of the relative error in refractivity is about −9.6% at 0.5 km, in Fig. 3(d). In addition to the superrefraction effect, the influence of other error sources is difficult to estimate, which is not discussed in this paper.

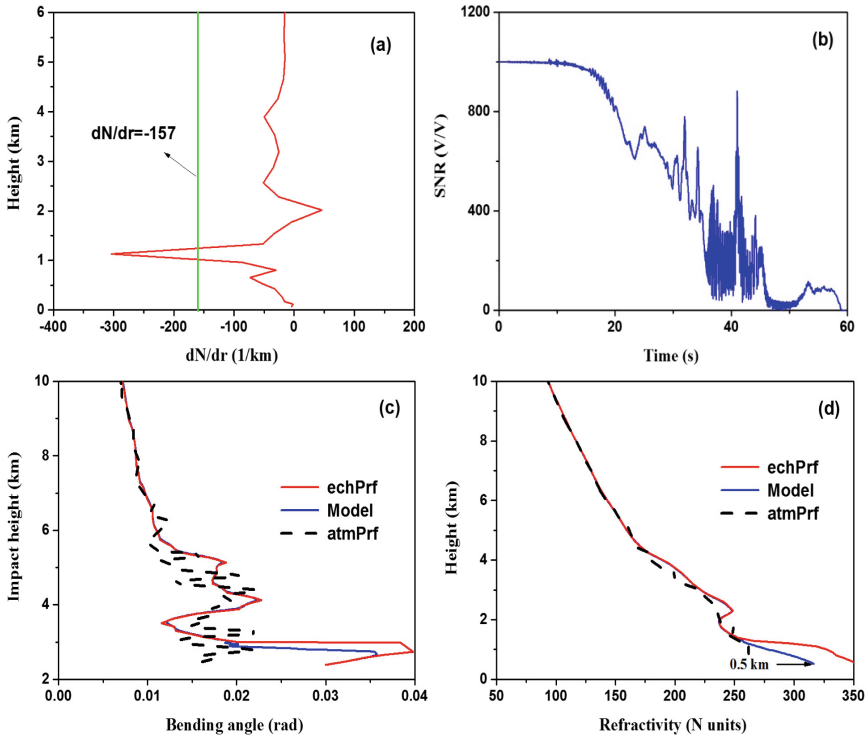


Fig. 3. Simulation of superrefraction phenomenon with echPrf data (C005.2007.152.03.13.G29) (17 N, 63E)

4 Statistical Analysis

4.1 Superrefraction Influence on Inversion

It is difficult to observe the superrefraction phenomenon in atmPrf data, as the refractivity profiles retrieved from COSMIC occultation data are smoothed. In order to study the influence of superrefraction effect on GPS/LEO RO inversion, the minimum refractivity gradient of each echPrf profile, is calculated as $G_{\min} = \min\{\mathrm{d}N_{\mathrm{ECMWF}}/\mathrm{d}r\}$, in which N_{ECMWF} denotes the refractivity profile of echPrf data. According to the minimum value of refractivity gradient, the echPrf data are divided into 4 categories: (A) $-100\,\mathrm{km}^{-1} \le G_{\min} \le 0\,\mathrm{km}^{-1}$; (B) $-157\,\mathrm{km}^{-1} \le G_{\min} < -100\,\mathrm{km}^{-1}$; (C) $-300\,\mathrm{km}^{-1} \le G_{\min} < -157\,\mathrm{km}^{-1}$; (D) $G_{\min} < -300\,\mathrm{km}^{-1}$. A total of 24530 echPrf data, from DOY (day of year) 151 to 161 in 2007, were classified, of which 19 388 (79.0%) were in category A, 3778 (15.4%) in category B, 1238 (5.1%) in category C and 126 (0.5%) in category D.

The mean and standard deviation of the relative refractivity error, between different category and corresponding COSMIC data, are calculated.

The fractional difference of refractivity $\Delta_f N$ between COSMIC occultation and ECMWF analysis (echPrf data) is defined as

$$\Delta_f N = \frac{N_{\mathrm{COSMIC}} - N_{\mathrm{ECMWF}}}{\bar{N}} \times 100\% \tag{1.5}$$

where N_{COSMIC} and N_{ECMWF} represent the refractivity derived from COSMIC occultation and ECMWF analysis, respectively. The symbol $\bar{N}$ denotes the average between N_{COSMIC} and N_{ECMWF}.

The statistical results are shown in Fig. 4. The vertical axis denotes height, and the horizontal axes are mean, standard deviation of the relative error in the refractivity and occultation number of the echPrf data, respectively. In the statistics, data will be truncated when $|\Delta_f N| > \varepsilon$, the truncation parameter $\varepsilon = 0.3$. The red circle, black square, blue triangle and green triangle solid line represent A, B, C and D category, respectively. There is no superrefraction effect in the category A. The minimum mean value of relative error of refractivity is about −0.51%, and the maximum standard deviation is about 2.67%. The minimum relative error of refractivity of the category B is about −1.64%, and the maximum standard deviation is about 4.30%. Category C and D contain superrefraction effect, and there are positive and negative deviations in the results, similar to category B. From category A to category D, the mean and standard deviation of the COSMIC refractivity become worse as a whole, especially below 5 km. The noise and superrefraction effect may be the main causes of the positive and negative deviation, and the unsolved multipath effect and signal tracking error can also have an impact.

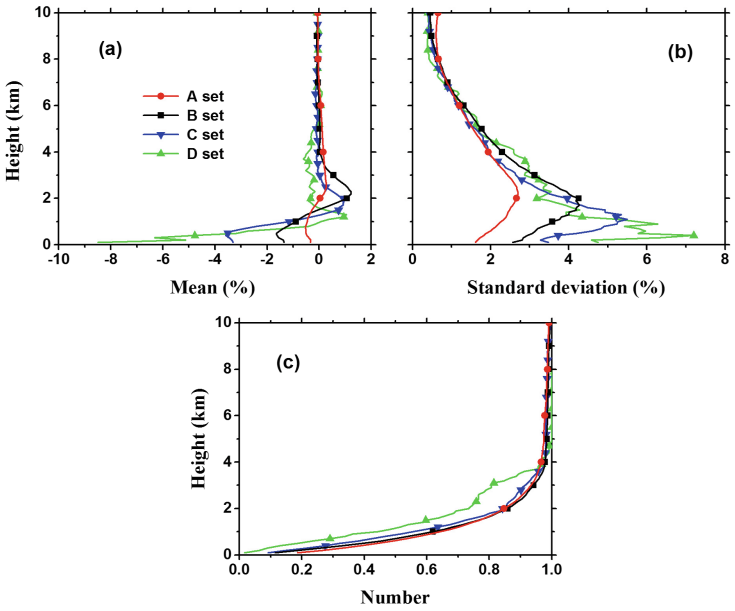


Fig. 4. The means and standard deviations of the fractional difference in refractivity, and the number of the occultation used for comparison (2007.151–2007.161)

4.2 Superrefraction Strength and Depth

Both category C and category D of the echPrf data contain superrefraction effect. Global distribution of minimum refractivity gradient is made, as seen in Fig. 5. The horizontal and vertical axes are longitude and latitude, respectively. In Fig. 5, most of the superrefraction phenomenon occurs in the tropical marine region. The locations of the superrefraction phenomenon in the tropics are mainly near the following regions: North America (Mexico), South America (the western coast of Peru), North Africa (the Sahara desert), and South Africa on the Atlantic side. A large number of cumulus clouds in these regions, results in the occurrence of superrefraction. In addition, there are a few superrefractions in Antarctica, which are related to the inversion layer near the surface of the earth.

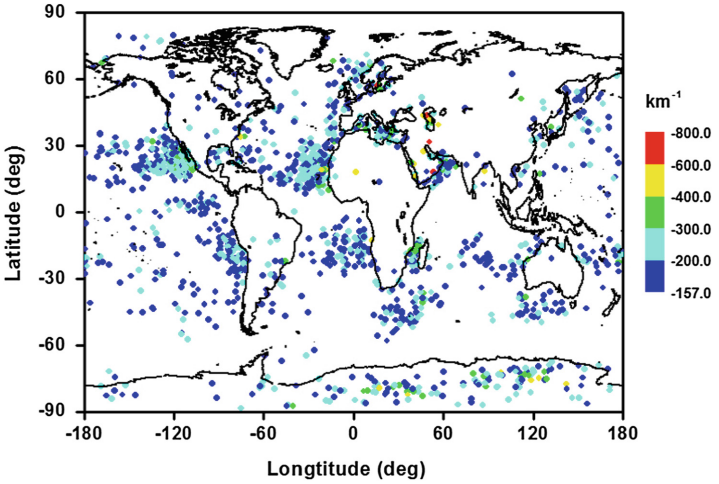


Fig. 5. The minimal vertical refractivity gradient (2007.151–2007.161)

5 Conclusions

The influence of superrefraction on the inversion of earth's neutral atmospheric parameters is discussed, through both numerical simulations and statistical comparisons of the COSMIC occultation data. The results of the simulation show, the superrefraction effect results in big error between the bending angle or refractivity profiles and the corresponding true values. The superrefraction effect may cause the minimum value of the refractivity error of about −9.6%.

In order to discuss the influence of refractivity gradient on inversion of the GPS/LEO RO data, the echPrf data is divided into 4 categories. Categories A and B have not superrefraction effect, while categories C and D contain superrefraction effect. Case studies show, the maximum relative refractivity error between COSMIC occultation and echPrf data changes from 0.65% to −12.91% (near the ground). Statistical comparison results show, the minimum mean and the maximum standard deviation of

the relative refractivity error are −0.51% and 2.71% (category A), −1.64% and 4.27% (category B), −3.59% and 5.49% (category C), −8.50% and 7.21% (category D), near the ground. It still needs a lot of work to effectively and accurately retrieve the atmospheric parameters in the atmospheric boundary layer, so as to improve the retrieval accuracy in the moist lower troposphere.

Acknowledgements. The authors are grateful to University Corporation for Atmospheric Research (UCAR) for access to COSMIC RO data and ECMWF analysis data. Thanks are also due to Wegener Centre and University of Graz for providing the EGOPS software. This research is supported by the National Natural Science Foundation of China under Grant Nos 41305016 and the Natural Science Foundation of Ningbo under Grant Nos 2018A610284, 2017A610130.

Refrences

1. Wu MJ, Guo P, Xu TL et al (2015) Data assimilation of plasmasphere and upper ionosphere using COSMIC/GPS slant TEC measurements. Radio Sci 50:1–10. https://doi.org/10.1002/2015rs005732
2. Xie F, Wu DL, Ao CO et al (2012) Advances and limitations of atmospheric boundary layer observations with GPS occultation over southeast Pacific ocean. Atmos Chem Phys 12:903–918. https://doi.org/10.5194/acp-12-903-2012
3. Xie F, Wu DL, Ao CO et al (2010) Super-refraction effects on GPS radio occultation refractivity in marine boundary layers. Geophys Res Lett 37:L11805. https://doi.org/10.1029/2010gl043299
4. Xie F, Syndergaard S, Kursinski ER et al (2006) An approach for retrieving marine boundary layer refractivity from GPS occultation data in the presence of superrefraction. J Atmos Ocean Tech 23:1629–1644
5. von Engeln A, Teixeira J (2004) A ducting climatology derived from the European Centre for Medium-Range Weather Forecasts global analysis fields. J Geophys Res 109(D18):D18104. https://doi.org/10.1029/2003jd004380
6. Guo P, Kuo YH, Sokolovskiy S et al (2011) Estimating atmospheric boundary layer depth using COSMIC radio occultation data. J Atmos Sci 68:1703–1713. https://doi.org/10.1175/2011jas3612.1
7. Ao CO, Meehan TK, Hajj GA et al (2003) Lower troposphere refractivity bias in GPS occultation retrievals. J Geophys Res 108(D18):4577. https://doi.org/10.1029/2002jd003216
8. Sokolovskiy S (2001) Modeling and inverting radio occultation signals in the moist troposphere. Radio Sci 36(3):441–458

Research on Estimation Method of Surface Soil Moisture Content Based on Multi-star Fusion

Huasheng Huang[1], Yueji Liang[1,2(✉)], Mingli Yuan[1], Yifan Qiu[1], He Huang[1], and Chao Ren[1,2]

[1] College of Geomatics and Geoinformation, Guilin University of Technology, Guilin, China
lyjayq@glut.edu.cn
[2] Research Center of Precise Engineering Surveying, Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin, China

Abstract. Global Positioning System Interferometric Reflectance (GPS-IR) is a new remote sensing technique, which can be used to estimate soil moisture content near the surface. From the view of multi-satellite fusion, an estimating method of surface soil water content based on multi-satellite fusion is proposed. Firstly, the direct and reflected signals of multiple satellites are separated by low order polynomial fitting, and then the sinusoidal fitting model of the reflected signals is established to obtain the relative delay phase. Secondly, the multiple linear regression inversion model of soil moisture is established, and the input variable set of the model is determined by the correlation coefficient of each satellite. Finally, the advantage of multi-satellite fusion is brought into full play to retrieve soil moisture. The feasibility and effectiveness of single satellite and multiple satellite fusion inversion are compared and analyzed through the monitoring data provided by the Plate Boundary Observation (PBO). The experimental results show that the multiple linear regression model can realize the effective fusion of multiple satellites. Compared with the single satellite, the inversion accuracy is higher and the inversion error is more stable.

Keywords: GPS-IR · Soil moisture · Multi-satellite fusion · Multiple linear regression model

1 Introduction

Soil moisture is a key parameter for quantifying land and the exchange of atmospheric energy. Accurate monitoring of soil moisture is an important basis for achieving stable and high-yield agriculture. It is of great significance to study the estimation of surface soil moisture for agricultural, hydrology and meteorology [1].Compared with traditional soil moisture monitoring methods such as drying and weighing method, time domain reflectometry method and frequency domain reflectometry method, Global Positioning System Interferometric Reflectometry (GPS-IR) is used as a new method of remote sensing, and it has the characteristics of high efficiency and high resolution, and can realize non-contact and wide-scale detection of soil moisture. GPS-IR mainly achieves the inversion of soil moisture content by using the interference of direct and

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 193–202, 2019.
https://doi.org/10.1007/978-981-13-7751-8_20

reflected signal superimposed at the single antenna, and it has been recognized in soil moisture monitoring. The GPS reflected signals were used to measure ocean heights in1993. Since 1996, NASA Langley Research Center and the University of Colorado have jointly conducted GPS-R inversion of sea surface wind fields [2]. On this basis, the use of GPS satellite signal reflected by the target to achieve remote sensing detection related research began in many countries and regions. a method for mapping multipath environments was developed by using SNR data, and confirming that the amplitude and frequency of the reflected signal are closely related to the multipath environment [3]. The relative delay phase is the best metric for retrieving soil moisture changes, and there is a theoretical linear relationship with surface soil moisture [4, 5]. In response to the problem of vegetation scattering, many scholars have also carried out research work on vegetation effect correction, vegetation water content estimation and vegetation parameter estimation using GPS-IR [6]. A soil moisture retrieval algorithm was developed for the effects of near-surface soil moisture on GPS multipath, and the effects of near-surface vegetation variation on Signal Noise Ratio was studied [7]. Because the multi- path factors such as soil surface roughness, vegetation coverage information and soil moisture content around the station have different effects on the satellite reflection signal, it is difficult to directly establish an accurate soil moisture estimation model. Many scholars have shown that not all GPS satellite signals are suitable for inversion of soil moisture. At present, there are not much studies involving multi-satellite fusion inversion. The effective combination of multiple satellites is realized through machine learning and neural network inconsideration of the inversion of soil moisture mean in days, and it can improve the accuracy of soil moisture inversion [8, 9]. However, the established model parameters cannot be visualized, which is not conducive to further research.

Based on the above research, a multiple linear regression model for soil moisture based on multi-satellite fusion is proposes. According to the principle of multiple linear regression model, the relative delay phase of each satellite is analyzed by SPSS software, and then the input variables of effective satellite and model are selected, and the multi- satellite fusion inversion model is constructed to realize soil moisture inversion. Based on the monitoring data provided by Plate Boundary Observatory, compare and analyze the feasibility of soil moisture inversion by single satellite or different satellites, and verify the effectiveness of multiple satellite fusion inversion.

2 Principle and Method

2.1 Principle of Satellite Signal Reflection

The signal-to-noise ratio (SNR) is an index to reflect the signal quality of the receiver antenna, which is mainly affected by the antenna gain parameters, the correlator state and the multipath effect in the receiver. The signal received by the GPS antennais a composite signal of direct signal and reflected signal. As shown in Fig. 1, SNR can be described as [7]:

$$SNR = S_d(0) + S_m(\phi) \tag{1.1}$$

Where SNR represents signal-to-noise ratio observation; $S_{\mathrm{d}}(0)$ represents direct signal; '0' represents the initial phase; $S_m(\phi)$ representation reflected signal; 'ϕ' represents the phase value of the reflected signal.

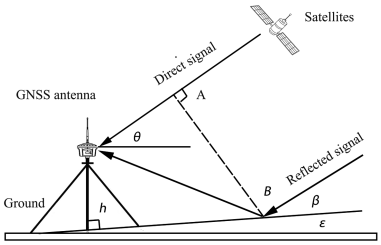


Fig. 1. Geometric model of ground multi-path errors

In Fig. 1, θ represents the incident altitude angle of satellite signal; h represents the vertical height of antenna from the bottom; ε represents the inclination angle of slope; β represents the inclination angle of slope. When ε is small, the formula can be expressed as follows:

$$\theta = \beta + \varepsilon \approx \beta \tag{1.2}$$

Combining with Fig. 1, the phase difference between direct and reflected satellite signals can be deduced as follows:

$$\psi = \frac{4\pi h}{\lambda} sin\theta \tag{1.3}$$

Where ψ is a functional variable associated with the geometric position of a satellite receiver, representing the relative phase difference between the direct and reflected signals (unit: rad). ψ is correlated with the θ, and as with the θ, when t changes, so does ψ. Its angular frequency ω_t is:

$$\omega_t = \frac{d\psi}{dt} = \frac{2\pi}{\lambda} 2h \frac{d(sin\theta)}{dt} \tag{1.4}$$

As can be seen from formula (1.4), And There is a linear relationship between ψ and $sin\theta$.

The relationship between SNR observations and relative delay phase is sinusoidal or cosine, and there exists a sinusoidal (or cosine) function relationship between the multipath reflection component and $sin\theta$ after removing the direct component and a fixed frequency [5]:

$$SNR_{\mathrm{MP2}} = A_{\mathrm{MP2}} cos(\frac{4\pi H}{\lambda} sin\theta + \phi_{MP2}) \tag{1.5}$$

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, ϕ_{MP2} is the relative phase delay. A_{MP2} and ϕ_{MP2} are the feature parameters needed.

2.2 Linear Regression Inversion Model of Soil Moisture Based on Multi-satellite Fusion

Suppose the relative phase delay set of GPS satellites is X:

$$\begin{aligned} X &= [x_1^0, x_2^0, \ldots, x_N^0], (N = 1, 2, \ldots) \\ x_N^0 &= [\phi_{\mathrm{MP2}}^1, \phi_{\mathrm{MP2}}^2, \ldots, \phi_{\mathrm{MP2}}^M], (M = 1, 2, \ldots) \end{aligned} \tag{1.6}$$

Where M is the day of year, and m can be any integer from 1 to M, x_N^0 is the relative delay phase set of each satellite, N is the number of satellites.

Let the soil moisture set corresponding to the satellite's relative delay phase set X be y. Then, X represent the independent variables, y is the dependent variables. The Multiple linear regression model [10]:

$$y = a_1 x_1^0 + a_2 x_2^0 + \ldots. + a_N x_N^0 \tag{1.7}$$

Where Y is soil moisture value retrieved by multiple linear regression equation, a_0 is the Equation constant term, $a_1, a_2, \ldots a_N$ are the coefficients representing the coefficient of dependent variable.

2.3 Multiple Linear Regression Model Estimation Process

(1) Effective separation of satellite reflections. The SNR2 value (L2 carrier) is obtained by calculating GPS monitoring data with TEQC (Translation, Editing and Quality Checking) software, and the satellite direct and reflected signals are separated by low-order polynomial fitting;
(2) Reflected signal resampling. The reflection signal changing with epoch is resampled and transformed into the relationship between the sinusoidal value of the incident altitude angle of the satellite;
(3) Parameter estimation. The least square algorithm is used for sinusoidal fitting of the reflection signal after resampling, and the relative delay phase is obtained;
(4) The estimation model of soil moisture. The correlation between the relative delayed phase and soil moisture is analyzed by linear regression equation, and the effective satellite is reasonably selected, and then the multivariate linear regression model is established;
(5) Set the model input variables to realize the inversion of soil moisture.

3 Experimental Analysis

The experimental data used in this paper are from PBO H2O website. (http://xenon.colorado.edu/portal). Considering account factors such as annual rainfall, terrain fluctuations, vegetation types and density, P041 station is selected for experimental analysis. The station is located at 105.194267° west longitude and 39.949493° north latitude. It can record L2C observations and carry out soil moisture analysis very early. It has a certain representativeness. From Fig. 2, the terrain around the station is relatively flat, open and sparsely populated, and it is conducive to soil moisture monitoring. The site is housed in a steel triangle bracket with a receiver model of TRIMBLENERT9, a radome of SCIT and an antenna model of TRM59800.80. In order to preliminarily verify the feasibility and effectiveness of multi-satellite fusion, the experiment is conducted in a flat terrain and sparse vegetation environment. In the experiment, the inversion model of soil moisture is multi-satellite fusion, and the obtained inversion results are average. The reference value of soil moisture in the experiment is based on the soil moisture products published by PBO. The reference value is equal to the average value of soil moisture in the effective range near the station. For the multi-satellite fusion inversion of soil moisture, the problem of time synchronization is negligible.

Therefore, this experiment selected the monitoring data of 70–290 days in 2011, and the sampling rate was 30 Hz. According to the existing research, the cut-off satellite altitude angle is set to be 5–20. The SNR value (L2 carrier) is obtained by calculating the monitoring data of GPS receiver with TEQC. The direct and reflected signals of each satellite are separated by quadratic polynomial, and the relative delay phases of each satellite are fitted by least square method. Due to space limitation, the satellites PRN06, 11, 14, 18, 21, 22, 26, 31 of P041 station is only given for analysis. The relative delay phase (RDP) and soil moisture changes (SMC) of each satellite are shown in Fig. 3.

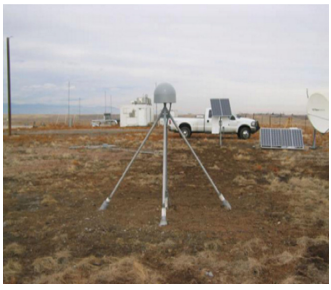


Fig. 2. Surround environment of station

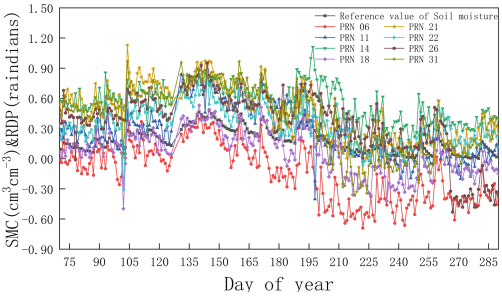


Fig. 3. Relationship between relative delay phase and soil moisture of P041 station

Figure 3 shows that when soil moisture fluctuates upward or downward, the relative delay phases of each satellite can respond. The linear regression equation between the relative delay phase of each satellite and soil moisture is established by using the data of the whole period. The linear regression correlation coefficient (r) is shown in Fig. 4.

It can be seen that the correlation between the relative delay phase of different satellites and soil moisture is different. Comparing with the satellites in Fig. 3, it is found that the response modes of different satellites to soil moisture change are different, the relative delay phase of the same satellite shows abnormal jump in different time periods. The response modes of different satellites to soil moisture are not consistent. The relative delay phase of the same satellite has abnormal hopping at different time periods. This may be due to the geometric trajectory of the satellite relative to the GPS antenna and the performance of the satellite itself during the observation process. Caused by surface multipath environmental impacts. Therefore, it is difficult to eliminate the abnormal jump value of a single satellite directly by some method or means, and it is not conducive to the establishment of continuity estimation model.

Therefore, five satellites (PNR 14, 18, 21, 22 and 31) are selected randomly, and the multivariate linear regression equations of single satellite to five satellites are established by using the whole time period data, and the r of the model is shown in Fig. 5. Inside, five sets of linear regression equations are established for five satellites, three sets of binary linear regression equations are established for two satellites, two sets of ternary linear regression equations are established for three satellites, two sets of quaternary linear regression equations are established for four satellites, and one set of five-element linear regression equations are established for five satellites. By comparing the r, it is found that the r of the multivariate linear regression equation established by the satellite increase with the number of satellites. It is possible to estimate soil moisture by linear fitting through multi-satellite fusion to form complementarity.

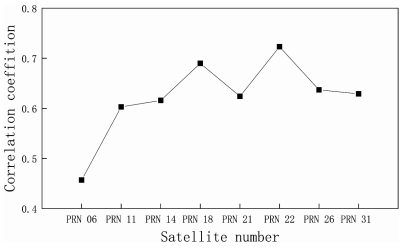


Fig. 4. The correlation coefficient between RDP and soil moisture

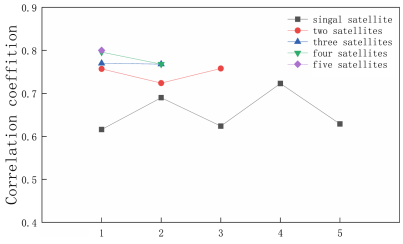
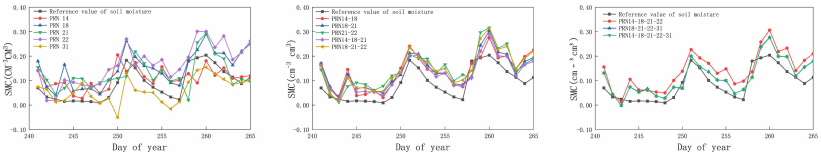


Fig. 5. Comparison of different input variable sets

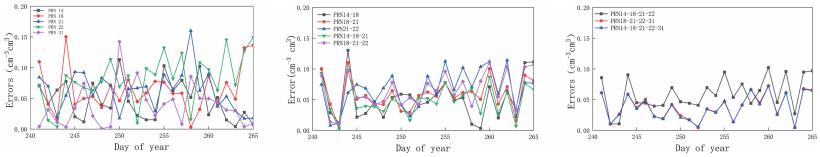
In Fig. 5, the numbers in abscissa represent corresponding the single satellite order number.1 is PRN14, 2 is PRN18, 3 is PRN21, 4 is PRN22 and 5 is PRN31. In legend, understanding with abscissa numbers which means the corresponding satellite fusion

serial number. Like two satellites fused when the abscissa is 1, mean PRN14 fused with PRN18 (Abbreviated as PRN14-18, and so on), at 2 is PRN18 fused with PRN21 (PRN18-21), at 3 is PRN21 fused with PRN22(PRN21-22). Three satellites fused at 1 is PRN14-18-21, at 2 is PRN18-21-22. Four satellites fused at 1 is PRN14-18-21 and at 2 is PRN18-21-31.

In order to verify the feasibility and validity of multi-satellite fusion estimation, the threshold range of r greater than 0.60 is set, and the relative delay phases of PRN14, 18, 21, 22 and 31 stars are selected to establish the multi-linear regression estimation model. Five schemes were set up for comparative analysis: Scheme 1 - Establish a single linear regression equation based on a single family of satellites for estimation (five groups), Scheme 2-Establishing binary linear regression equation based on two satellites fusion for estimation (three groups), Scheme 3-Establishing ternary linear regression equation based on three satellites fusion for estimation (two groups), Scheme 4-Establishing quaternary linear regression equation based on four satellites fusion for estimation (two groups), Scheme 5-Establishing five-element linear regression equation based on five satellites fusion for estimation (one group). In this experiment, the 181th-240th days of data are randomly selected as modeling data, and 241th-265th days of data are used as testing data. The coefficients of the modelling equations for each scheme are shown in Table 1. It can be seen that with the increase of the number of satellites, the parameters of the model will also increase. The estimated results and errors of each scheme are shown in Fig. 6.



(a) Estimates of programme 1 **(c)** Estimates of programme 2 and programme 3 **(e)** Estimates of programme 4 and programme 5



(b) Estimation error for scenario 1 **(d)** Estimation error for scenario 2 and scenario 3 **(f)** Estimation error for scenario 4 and scenario 5

Fig. 6. Comparison and analysis of the estimated results and errors of each model

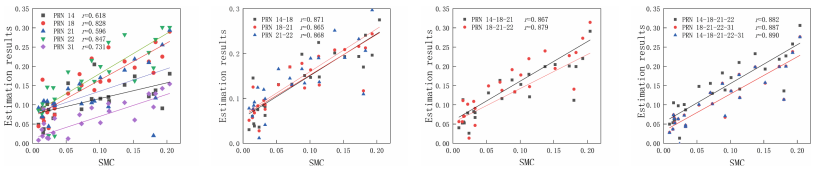
From Fig. 6, seeing the use of a single satellite for soil moisture estimation, it is difficult to accurately grasp the change law of soil moisture; especially when the soil moisture value is small, the estimation results of some satellites show obvious distortion. This may be due to the fact that a single satellite signal is greatly affected by the surface multipath environment in a certain direction at a certain time, so that useful soil moisture reflection information is interfered. Comparing schemes 2 to 5, it is found that

Table 1. Coefficient of each model equation

Coefficient of equation	14	18	21	22	31	14-18	18-21
a_0	−0.031	0.189	0.037	0.031	0.038	0.134	0.128
a_1	0.367	0.449	0.448	0.560	0.371	0.094	0.309
a_2						0.385	0.186
Coefficient of equation	21-22	14-18-21	18-21-22	14-18-21-22	18-21-22-31	14-18-21-22-31	
a_0	0.015	0.092	0.084	0.067	0.062	0.058	
a_1	0.230	0.069	0.205	0.038	0.166	0.011	
a_2	0.351	0.272	0.137	0.192	0.128	0.163	
a_3		0.171	0.219	0.132	0.138	0.127	
a_4				0.204	0.119	0.136	
a_5						0.117	

with the increase of the number of satellites, the inversion result is closer to the reference value of soil moisture, and the inversion error tends to decrease gradually, and the inversion error fluctuation is gradually weakened. It can be seen that different satellites reflect the surface environmental information in different directions. It is difficult to accurately grasp the soil moisture information in a certain area by single satellite estimation. To a certain degree, the Multi-satellite fusion inversion can suppress the impact of surface multi-path environment on soil moisture estimation.

In order to further synthetically evaluate the performance of the schemes, the *r* between the model estimation results and the reference, RMSE, MAE and MAX values of soil moisture are used to evaluate the performance of the schemes. Combined with Fig. 7 and Table 2, it is found that the greater the number of satellite fusions, the closer r is to 1, and the overall RMSE, MAE, and MAX show a downward trend. When the number of fusion satellites reaches 5, the correlation coefficient r between the model estimation result and the soil moisture reference value is 0.890, the RMSE and MAE are 0.131 and 0.028, respectively, and the MAX is less than 0.080. Compared with the single satellite model, each precision index has increased. Further analysis of Schemes 4 and 5 reveals that when the PRN 14 is integrated into the multi-star inversion model, the RMSE and MAE are all increased, and the remaining indicators are better; this may be related to the performance of the PRN 14 satellite itself, for how



(a) Option 1 **(b)** Option 2 **(c)** Option 3 **(d)** Option 4 and Option 5

Fig. 7. Linear regression analysis of estimation results and soil moisture

Table 2. Estimation accuracy of soil moisture in each model

Estimation of each model	14	18	21	22	31	14-18	18-21
r	0.618	0.828	0.596	0.847	0.731	0.871	0.865
RMSE	0.115	0.164	0.139	0.182	0.07S	0.150	0.151
MAE	0.046	0.064	0.063	0.081	0.037	0.052	0.058
MAX	0.113	0.150	0.160	0.149	0.142	0.130	0.110
Estimation of each model	21-22	14-18-21	18-21-22	14-18-21-22	18-21-22-31	14-18-21-22-31	
r	0.868	0.867	0.879	0.882	0.887	0.890	
RMSE	0.161	0.142	0.162	0156	0.086	0131	
MAE	0.066	0.049	0.065	0.059	0.034	0.037	
MAX	0.114	0.098	0.110	0.102	0.128	0.072	

It is a problem that needs to be further explored in combination with the time period in which GPS satellites appear, the satellite elevation angle, and the effective reflection area for satellite selection.

4 Conclusion

In view of the current research on soil moisture, it is more limited to the estimation by single satellite. Based on the idea of multi-star fusion, a linear regression estimation model of soil moisture based on multi-star fusion is proposed, the following conclusions:

(1) During the observation process, the GPS antenna is affected by the satellite geometric motion trajectory and the satellite' s own performance. The relative delay phase of different satellites has different response modes to the surface soil moisture, and there is a certain correlation with soil moisture. The linear regression equation can be used to describe the relationship between the two.
(2) The algorithm fully exploits the advantages of multiple linear regression models in soil moisture, can effectively integrate the performance of each satellite, and to some extent inhibits the impact of surface multi-path environment on soil moisture estimation. At the same time, it can improve the phenomenon of abnormal fluctuation of soil moisture when using a single satellite to some extent.
(3) The algorithm is used to invert soil moisture, and various precision indicators are improved. When the number of satellite fusion increases, the inversion error is more stable, and the *r* between the inversion soil moisture value and the soil moisture reference value is larger.

For different environments, how to better achieve the selection of multiple satellites and the combined inversion of different satellite navigation systems is the focus of the next step.

Acknowledgements. This work was supported by Basic Ability Improvement Project for Young and Middle-Aged Teachers in Universities in Guangxi (2018KY0247), the National Natural Foundation of China (41461089), and College Students' Innovation and Entrepreneurship Project (201810596048).

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 For Meteorol 148(8):1362–1373. https://doi.org/10.1016/j.agrformet.2008.04.003
2. Martin-Neira M (1993) A passive reflectometry and interferometry system (PARIS): application to ocean altimetry. ESA J 17(4):331–355
3. 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
4. Zavorotny VU, Larson KM, Braun JJ, Small EE, Gutman ED, Bilich AL (2011) A physical model for GPS multipath caused by land reflections: toward bare soil moisture retrievals. IEEE J Sel Top Appl Earth Obs Remote Sens 3(1):100–110. https://doi.org/10.1109/JSTARS.2009.2033608
5. 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. https://doi.org/10.1109/TGRS.2014.2364513
6. Wu X, Ying L, Jin X (2012) Theoretical study on GNSS-R vegetation biomass. Geosci Remote Sens Symp. https://doi.org/10.1109/igarss.2012.6352719
7. Chew C, Small EE, Larson KM (2015) An algorithm for soil moisture estimation using GPS-interferometric reflectometry for bare and vegetated soil. GPS Solut 20(3):525–537. https://doi.org/10.1007/s10291-015-0462-4
8. Liang Y, Ren C, Huang Y (2018) Research on multi-satellites fusion inversion model of soil moisture based on sliding window. In: China satellite navigation conference (CSNC) 2018 proceedings. Springer, Singapore. https://doi.org/10.1007/978-981-13-0005-9_7
9. Liang Y-J, Ren C, Wang H-Y, Huang Y-B, Zheng Z-T (2018). Research on soil moisture inversion method based on ga-bp neural network model. Int J Remote Sens 1–17. https://doi.org/10.1080/01431161.2018.1484961
10. Kabacoff R (2015) R in action: data analysis and graphics with R

Analysis of the Effect of Receiver Type on Accurate GPS Data Processing

Shuguang Wu[1(✉)], Guigen Nie[1,2], Changhu Xue[1], Fengyou Peng[1], and Lei Qiu[3]

[1] GNSS Research Center, Wuhan University, Wuhan 430079, Hubei, China
shgwu@whu.edu.cn
[2] Collaborative Innovation Center for Geospatial Information Technology, Wuhan 430079, Hubei, China
[3] Shenzhen Cadastral Surveying and Mapping Office, Shenzhen 518034, Guangdong, China

Abstract. Taking Shenzhen CORS system as an example, this paper performs high-precision calculation on the GPS data of Trimble and Panda receivers from 11 reference stations. By analyzing the difference of baseline accuracy and positioning results between Trimble and Panda receivers, the feasibility in urban CORS system and accuracy of the domestic GPS receiver are then evaluated. The results show that the difference between the NRMS values of the two receivers is small, and the difference of the baseline repetition is also kept in the scale of millimeters, and the maximal difference of the coordinate components after the constraint adjustment is not more than 5 mm. Furtherly, linear least-squares fitting of the baseline repetition shows that both the fixed error and the proportional error coefficient of both receivers meet the specification requirements.

Keywords: GPS receiver · Baseline accuracy · Baseline repetition · Positioning accuracy

1 Introduction

As satellite positioning theory and technology improve, Continuously Operating Reference Stations (CORS) of different levels (national, provincial and municipal) have been established and promoted step by step. At present, dozens of cities across China have established urban CORS systems, with real-time and post-time services provided by them having been widely used in daily production of surveying and mapping departments of all levels [1]. Shenzhen Continuously Operating Reference Stations (SZCORS) is the first practical real-time dynamic CORS system established in China that was completed in 2000. Since its establishment, SZCORS has been providing accurate and real-time positioning services to the society, which plays a vital role in Shenzhen's urban construction and economic development [2].

Nowadays, based on considerations of geographic information safety and better meeting different needs of various CORS users in Shenzhen, SZCORS has been updated in the past two years. System upgrade mainly includes: (1) Domestic software

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 203–210, 2019.
https://doi.org/10.1007/978-981-13-7751-8_21

of data processing in control center, PowerNetwork, and Chinese hardware equipment (Panda receiver, Huaxin antenna) are put into use; (2) Upgrade the original Trimble GPSNet data processing software to a new one, Pivot, with more advanced data processing technology; (3) Original 5 reference stations are expanded to 11 forming a wider and stabler urban CORS network.

After upgrade of SZCORS, two kinds of receivers, Trimble and Panda, both were placed in 11 reference stations. GPS satellite data is received by the same antenna on each reference station, then the data is transmitted to two receivers through feeder and power divider. Through this zero-baseline measurement method, we created a synchronous and equal environment measurement condition, eliminating the effects of other error sources besides internal noise of receiver. In this paper, GPS data collected by the two types of receivers on reference stations are solved separately. By analyzing the accuracy difference of baseline and positioning results, feasibility and accuracy of domestic GPS receiver applying to the urban CORS system are then evaluated.

2 GPS Data Processing

In data processing, we use GAMIT10.5 software to solve baseline, obtaining GPS single-day baseline result, then we use COSA software to perform 3-dimensional constraint adjustment. Spatial distribution of each reference stations is shown in Fig. 1, and receiver/antenna information are listed in Table 1.

(1) Data: ① observation data of 15 reference stations in Fig. 1 from 1st to 7th, August, 2018 (day of year 213–219), with sampling interval 30 s; ② Precise and broadcast ephemeris derived from International GNSS Service (IGS).
(2) Constraint stations: 4 reference stations from Guangdong CORS, including Zhongshan (SHIJ), Dalingshan (DLSH), Humen (HMSC), Huidong (GDHD). Their coordinates were obtained from unified 3-dimensional adjustment of all CORS stations in Guangdong province, and CGCS2000 was employed.
(3) Main strategies in baseline processing: ① Ionospheric refraction is eliminated by LC observation; high-order ionospheric delay model is GMAP; geomagnetic field reference model is IGRF11; ② Tropospheric delay correction: mapping function adopts GMF, which every 2 h estimates a zenith tropospheric delay parameter and every 12 h estimates the atmospheric horizontal gradient parameter; ③ Tidal correction: the tidal corrections used include the Earth solid tide (IERS2003), pole tide, ocean tide (FES2004), and atmospheric tide (global grid model ANU100826);④ Non-tidal correction: non-tidal atmospheric model correction (atmfilt_cm.year); non-tidal oceanic loading and surface water are not corrected at the observation level; ⑤ Satellite and receiver antenna phase center correction (PCO). Correction model of PCO is AZEL where antenna phase center is regarded as a function of satellite height angle and azimuth angle, using bilinear interpolation; using igs08.atx absolute antenna phase center correction model, including PCO, PCV correction.

Table 1. Information about antenna and receiver types in CORS reference stations

Station	Receiver	Antenna	Affiliation
HMSC	TRIMBLE 5700	HXCCGX601A	Guangdong CORS
SHIJ	PD318	TRM59900	Guangdong CORS
DLSH	TRIMBLE 5700	HXCCGX601A	Guangdong CORS
GDHD	TRIMBLE NETRS	TRM55971	Guangdong CORS
SZDP	TRIMBLE NETR9 & PD318	TRM55971	Shenzhen CORS
SZJY	TRIMBLE NETR5 & PD318	TRM55971	Shenzhen CORS
SZLG	TRIMBLE NETR9 & PD318	TRM55971	Shenzhen CORS
SZNS	TRIMBLE NETR9 & PD318	TRM55971	Shenzhen CORS
SZSY	TRIMBLE NETR5 & PD318	TRM55971	Shenzhen CORS
BAOA	TRIMBLE NETR9 & PD318	HXCCGX601A	Shenzhen new CORS
BINH	TRIMBLE NETR9 & PD318	HXCCGX601A	Shenzhen new CORS
DONC	TRIMBLE NETR9 & PD318	HXCCGX601A	Shenzhen new CORS
KENZ	TRIMBLE NETR9 & PD318	HXCCGX601A	Shenzhen new CORS
PINH	TRIMBLE NETR9 & PD318	HXCCGX601A	Shenzhen new CORS
SONG	TRIMBLE NETR9 & PD318	HXCCGX601A	Shenzhen new CORS

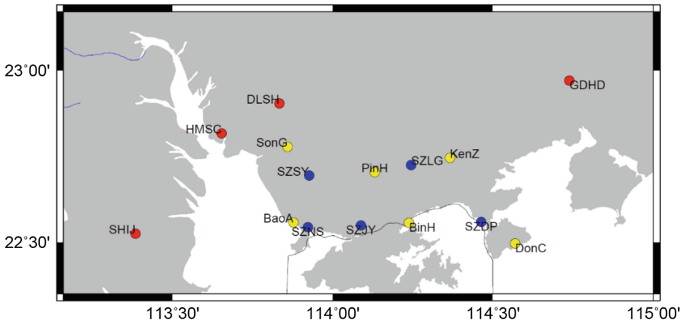


Fig. 1. Spatial distribution of CORS reference stations. 4 red balls are Guangdong CORS stations. 5 blue balls are original Shenzhen CORS stations, and 6 yellow balls are newly-built stations.

3 Analysis of Baseline Results

3.1 NRMS

After GAMIT baseline processing is completed, standard root mean square error (NRMS) of baseline solution can be used as an indicator that reflects quality of GPS synchronization loop, indicating deviation degree that baseline values of the single-period solutions Y_i away from their weighted average Y [3–5], as shown in Eq. (1). NRMS value of less than 0.3 is usually required, and if greater than 0.5, it indicates that the cycle slips may not be fully repaired during baseline processing.

$$\mathrm{NRMS} = \sqrt{\frac{1}{\mathrm{N}}\sum_{i=1}^{n}\frac{(\mathrm{Y_i} - \mathrm{Y})^2}{\sigma_i^2}} \tag{1}$$

Figure 2 shows NRMS values obtained by two solutions. It can be seen that NRMS values of all single-day solutions are less than 0.2, which satisfies the general limit. In comparison, NRMS values of Trimble receiver is slightly smaller than that of Panda, indicating that baseline estimation accuracy of the former is slightly higher, but the difference between them is not more than 0.01.

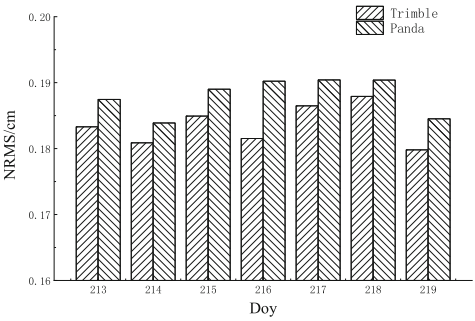


Fig. 2. NRMS values derived from GPS data processing of Trimble and Panda receivers

3.2 Baseline Repeatability

In quality control of baseline solution, repeatability is a higher level of quality control indicators [6]. The definition of repeatability is:

$$\mathrm{R_C} = \left[\frac{\frac{n}{n-1}\cdot\sum_{i=1}^{n}\frac{(C_i - C_m)^2}{\sigma_{c_i}^2}}{\sum_{i=1}^{n}\frac{1}{\sigma_{c_i}^2}}\right]^{\frac{1}{2}};\ C_m = \frac{\sum_{i=1}^{n}\frac{C_i}{\sigma_{c_i}^2}}{\sum_{i=1}^{n}\frac{1}{\sigma_{c_i}^2}} \tag{2}$$

where n is the total number of observation sessions for a baseline. In this experiment, the value of n is 7; C_i is a baseline component or its length obtained in a session; $\sigma_{C_i}^2$ means variance of C_i; C_m is the weighted average of C_i in each session. Baseline repeatability reflects the dispersion of baseline components and lengths. The smaller this value is, the higher the accuracy of baseline alignment, and the better the baseline quality [7, 8].

This paper performs a repeatability check on all 105 baselines to analyze the accuracy difference within the baseline results of the two types of receivers. Figure 3 shows the repeatability difference of the baseline results obtained from Trimble and Panda receivers (the former minus the latter), and it can be seen that this difference is

maintained within the scale of millimeters. Repeatability difference between horizontal directions and length direction is small, within the range of 3 mm. Repeatability difference in vertical component is large, which is basically in the range of 6 mm, but there are also repeatability differences of several baselines close to the degree of centimeters. In addition, in North component, repeatability of the baseline result obtained from Trimble is less than that of Panda, so former internal accuracy is higher, while on the remaining components, the opposite is true. According to statistics, average repeatability of 105 baselines of Trimble on North, East, Up and Length components is 2.6, 3.8, 9.6 and 4.0 mm, and the corresponding results of Panda are 2.9, 3.6, 8.6, and 3.9 mm.

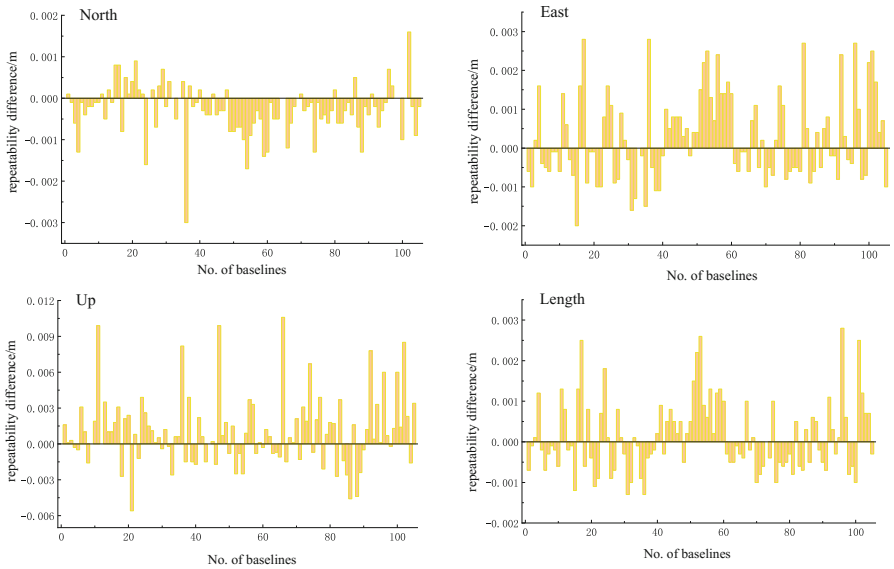


Fig. 3. Repeatability difference of the baseline results obtained from Trimble and Panda receivers (the former minus the latter)

Among all baselines, the shortest and longest baselines are selected, and repeatability of baseline components are listed in Table 2. We see that for short baseline of several kilometers and the long baseline of over 100 km in this experiment, repeatability in baseline horizontal and length components is smaller than that in vertical component. The former is within millimeter level, and as length increases it becomes larger. The latter varied in a range of 2 cm, and its change is less obvious as length increases. This characteristic can also be seen from Fig. 4 in next section. For relative baseline repeatability in this experiment, short baseline is on the order of 10^{-7}, and long baseline is on the order of 10^{-8}, which shows a high baseline resolution accuracy.

Table 2. Repeatability of baseline components of the shortest and longest baselines

Name	Receiver	Length (m)	RN (m)	RE (m)	RU (m)	RL (m)	Relative repeatability
BAOA-SZNS	Trimble	4869.5039	0.0012	0.0016	0.0124	0.0016	3.29E-07
	Panda	4869.5046	0.0010	0.0019	0.0089	0.0018	3.70E-07
GDHD-SHIJ	Trimble	147473.1989	0.0017	0.0089	0.0114	0.0082	5.56E-08
	Panda	147473.1978	0.0034	0.0076	0.0139	0.0073	4.95E-08

3.3 Least Square Linear Fitting

The baseline repeatability can also be expressed as two parts, fixed error and proportional error, that is

$$R_C = a + b \cdot L_c \tag{3}$$

where a and b are fixed error and proportional error coefficients of a certain component of baseline. According to GPS survey specifications of China [9, 10], after the calculation of high-level GPS network, we should,apart from checking baseline repeatability, also ensure fixed error a and proportional error coefficient b within the specified range. According to principle of least squares, the two parameter values are fitted in this paper, as shown in Fig. 4.

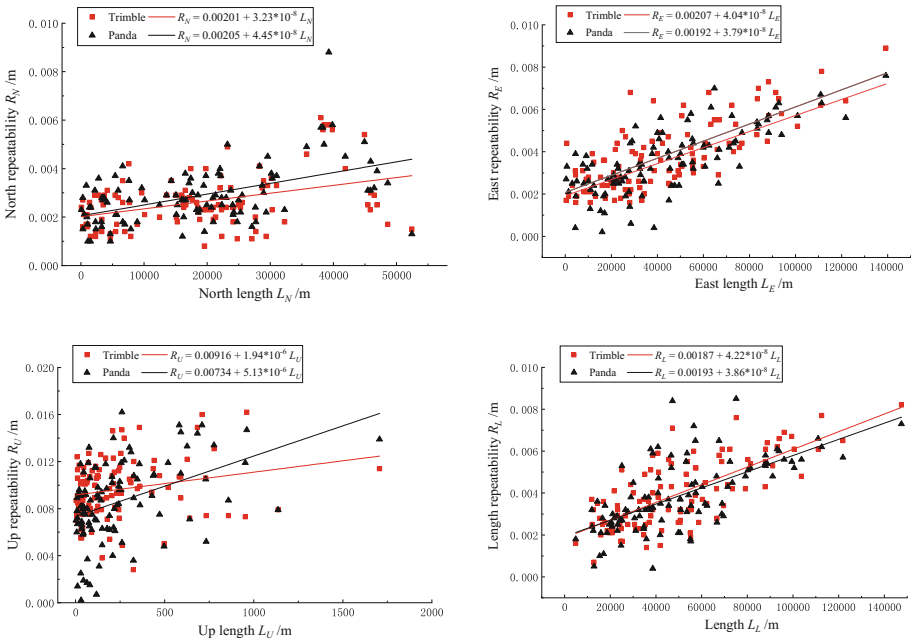


Fig. 4. Least square fittings of baseline results of Trimble and Panda in three components and length direction

According to GPS survey specification, for C-level GPS control network, fixed error a and proportional error coefficient b in length direction of baseline should not exceed 10 mm and 5 ppm, respectively. From linear fitting of baseline length in Fig. 4, fixed errors of Trimble and Panda receivers are 1.87 and 1.93 mm, and proportional error coefficients are 4.22×10^{-2} and 3.86×10^{-2} ppm, far below limits in specifications. Therefore, baseline accuracy of both Shenzhen GPS control networks constructed by two receivers discussed in this paper fully meets the design requirements.

4 Adjustment and Analysis of Positioning Results

After completion of baseline processing, Cosa GPS Data Processing System (Version 5.21) compiled by Wuhan University is then employed to perform 3-dimensional network adjustment under CGCS2000 coordinate framework. The known coordinates of HMSC, SHIJ, DLSH and GDHD are selected as the origin. Figure 5 shows the difference between adjustment results of baselines derived from Trimble and Panda receivers (the former minus the latter). It can be seen that the maximum value of the difference between the two sets of coordinate results is 4.7 mm in Y component of SZDP station, followed by 4.2 mm in Z component of DONC station. Coordinate differences of the remaining station components are not more than 4 mm, indicating that the positioning result is stable and reliable.

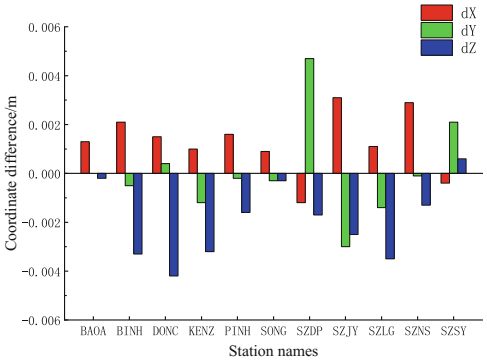


Fig. 5. Difference of 3-dimensional network adjust derived from Trimble and Panda receivers (the former minus the latter)

5 Conclusion

In this paper, GPS data of Trimble and Panda receivers that share the same antenna at the same reference station are solved respectively. NRMS value, baseline repeatability and coordinate difference are then compared. Results show that difference in NRMS values between the two receivers is small. Difference in baseline repeatability is also maintained in the scale of millimeters. Also, the maximum difference of coordinate

components after adjustment is no more than 5 mm. Furthermore, least squares linear fitting was performed on baseline repeatability. Results show that fixed error and proportional error coefficient of both receivers meet the specification requirements.

As China gradually develops and prospers, the issue of geographic information security has received increasing attention from government and people. GPS receivers are important tools for collecting geographic information, and if domestic receivers can gradually reduce or even surpass foreign similar products in terms of performance (including baseline and positioning accuracy) and operate independently, then China's homeland security will be furtherly guaranteed. From the experimental analysis in this paper, domestic receivers on Shenzhen CORS reference stations have almost the same accuracy as foreign receivers perform, and they manage to meet the designing requirements of C-level GPS network. With gradual improvement of Beidou satellite system network in China, more and more satellites are used for data solution. Using GPS+Beidou satellite data stored in domestic receivers for joint data processing will be the following work of this paper.

References

1. Wu S, Nie G, Qiu L et al (2017) Coordinate time series analysis on reference stations of Shenzhen CORS. J Geomat 42(3):117–120. https://doi.org/10.14188/j.2095-6045.2016330
2. Wu S, Nie G, Qiu L, et al (2018) Stability analysis on reference stations of Shenzhen CORS. Sci Surv Mapp 43(6):78–82, 87. https://doi.org/10.16251/j.cnki.1009-2307.2018.06.013
3. Li Z, Zhang X (2009) New techniques and precise data processing methods of satellite navigation and positioning. Wuhan University Press, Wuhan
4. Wang J (2016) Data processing technology with high precision based on GAMIT and its application in Wenchuan earthquake analysis. Master thesis, Wuhan University, Wuhan, China
5. Wu S (2017) Characteristics of coordinate time series from regional CORS stations. Master thesis, Wuhan University, Wuhan, China
6. Li Z, Huang J (2010) GPS surveying and data processing. Wuhan University Press, Wuhan
7. Ma H, He L (2008) The updated GAMIT software's function and characteristics and its application example analysis. Mine Surv 04:35–39+4
8. Huang G, Wang B, Wang Y (2011) Baseline repeatability test and quality analysis of high-precision GPS control network. Bull Surv Mapp 07:9–11
9. State Administration for Quality Supervision and Inspection and Quarantine. GB/T 18314-2001. Specifications for global positioning system (GPS) surveys. China Standards Press, Beijing, China
10. State Administration for Quality Supervision and Inspection and Quarantine. GB/T 18314-2009. Specifications for global positioning system (GPS) surveys. China Standards Press, Beijing, China

GNSS Snow Depth Monitoring Using SNR Observations

Fenfen Li[1], Lilong Liu[1(✉)], Liangke Huang[1], Wei Zhou[1], Junyu Li[1,2], Yunzhen Yang[1], and Donggui Huang[1]

[1] College of Geomatic Engineering and Geoinformatics, Guilin University of Technology, Guilin, China
hn_liulilong@163.com
[2] GNSS Research Center, Wuhan University, Wuhan, China

Abstract. Snow cover plays an important role in global climate regulation and hydrological cycle. However, conventional detection methods cannot achieve accurate detection of snow depth in large-scale. With the development of global positioning system multipath reflection (Global Navigation Satellite Systems Multipath Reflectometry, GNSS-MR) technology, the signal-to-noise ratio (Signal-to-Noise Ratio, SNR) data have been successfully used to detect soil moisture, sea level and other environmental parameters. To verify the difference and accuracy of the thickness of snow cover detected by GNSS using SNR observations at low elevation angles, in this paper, by increasing the elevation angle to obtain the long time series to invert snow depth. Using the GPS observations of the KIRU station in Sweden from January to May 2016 as a data source, the SNR data of L1 and L2 bands at this station are extracted. The snow depth inversion experiment is carried out at three sets of different low elevation angles, and the snow surface is extracted to the receiver antenna by Lomb-Scargle spectrum method. The inverted snow depth is compared with the situ measured snow depth. The results show that at the elevation angle range 0°–20°, 0°–25° and 0°–30°, the inverted snow depth of L1 and L2 bands are associated with the variations in situ measured values. By increasing the elevation angle to obtain the long time series, the inversion values are better at 0°–25°, 0°–30°, and the correlation coefficients are better than 0.93. At the snowless stage, the inverted values of L1 and L2 bands fluctuated around the measured values at different elevation angles.

Keywords: GNSS-MR · SNR · Snow depth · Elevation angle · Lomb-Scargle spectrum analysis

1 Introduction

Snow cover is one of the forms of water resource storage. The monitoring snow depth is an important for the research of snow water content. At present, there are many methods to measure the snow depth. The thickness of snow cover obtained from field measurement is very accurate, but inefficient. It is not suitable for obtaining information about snow depth in a wide range. Relying on the high spatial resolution, satellite remote sensing monitoring technology can provide the coverage of snow cover, but it is

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 211–219, 2019.
https://doi.org/10.1007/978-981-13-7751-8_22

seriously affected by weather factors, such as the cloud layer and atmospheric refraction. In recent years, GNSS-MR technology has been developed and applied to the detection of snow depth. It has the following advantages: rich signal sources, good continuity and wide coverage et al.

Martin-Neira, a scientist in the European Space Agency, is the first to come up with the concept of detecting using reflected GPS signal [1]. French scientist Auber (1994) accidentally received a GPS reflected signal during an airborne flight test, proving that the reflected signal could be received [2]. Based on this, GNSS-MR technology has developed rapidly. Domestic and foreign scholars use conventional measurement GPS receivers to receive data for snow thickness [3–5], sea ice [6–8], soil moisture [9–11] and sea level [12, 13] have achieved a series of application results, which provide reference for the application of GNSS-MR technology. Larson et al. verified that GPS multipath modulated signals can better track changes in surface snow depth by using the 2009 snowstorm in Colorado, USA [14]. Subsequently, Ozeki et al. obtained the time series of snow depth using L4 and SNR respectively at the GPS station in Hokkaido, Japan and verified the feasibility of snow depth inversion using multipath information [15]. Tabibi et al. explored the application of L5-band GPS SNR in the detection of snow depth. With compared variations in L5 signal under different reflector heights, surface materials and surface roughness's, compared them with the L2C and suggested that there was no difference between L5 and L2C in terms of snow depth detection using GPS SNR multipath observations [16]. Yu et al. used a geodetic GPS receiver, multipath reflection measurement, and GPS triple-frequency (L1, L2 and L5) carrier phase measurements to observe the snow depth. Unaffected by the ionosphere, the feasibility of the method is verified by two sets of experimental data recorded in two different environments [17].

To verify the relationship between snow depth inversion and snow depth inversion at different low altitude angles by using SNR, the long time series is obtained by raising the altitude angle. The SNR data of GPS L1 and L2 bands from January to May 2016 at KIRU station in Sweden are used. The snow depth around the station is inversed and compared with the actual values under three groups of different height angles.

2 The Principle of Snow Depth Inversion Using GNSS-MR Technology

The antenna of the GPS receiver can simultaneously receive direct signals and reflected signals caused by surface reflections. Figure 1 shows the schematic diagram for the application of GNSS-MR technology in the detection of the snow depth. The cyan dotted line represented A′r, the reflected signal from the satellite to the surface at the snowless stage. The red dotted line represented Ar, the reflected signal at the snow stage. Set H to the antenna height, that is, the distance from the phase center of the receiver antenna to the earth surface, h is the distance from the phase center of the receiver antenna to the surface of snow cover (collectively referred to is vertical reflection distance in this paper), hsnow is the thickness of snow cover, θ is the angle between the direct signal and snow surface, that is, the elevation angle of the satellite.

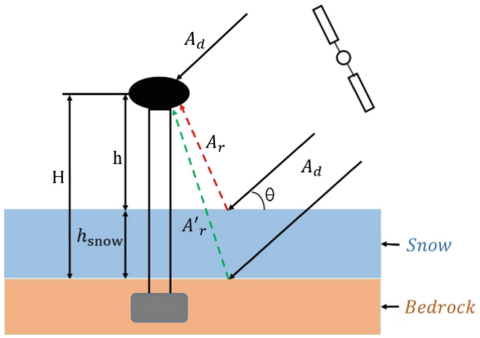


Fig. 1. Schematic of GNSS-MR technology for detecting snow depth

Figure 2 shows the SNR variation of the PRN 19 satellite signal on the 20th day of 2016 at KIRU Station. The horizontal axis represented equal interval observation epoch, while the vertical axis represented the variation of SNR. The unit is dB-Hz and the data sampling rate is 30 s. Since the reflected signal contained multipath information, which would directly affect the SNR data, the corresponding surface environment parameters can be obtained by spectrum analysis of SNR, that is, GNSS-MR inversion technique.

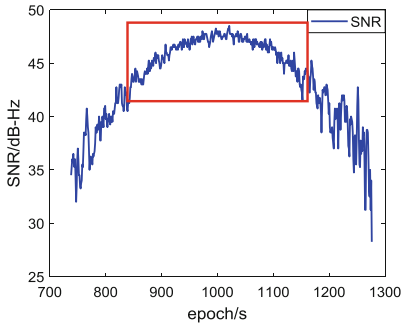


Fig. 2. The variation of SNR of GPS satellite signal

For the antenna of a geodetic surveying GPS receiver, in order to effectively suppress the multipath observation errors caused by the surface reflection, the amplitudes of the direct signal and the reflected signal are the following relationship:

$$A_d \gg A_r \tag{1}$$

From Fig. 1, the relationship between SNR and signal amplitude is as follows [9]:

$$SNR^2 = A_C^2 = A_d^2 + A_r^2 + 2A_dA_r\cos\theta \tag{2}$$

Where Ac is the amplitudes of the composite signal and $cos\theta$ is the cosine value of the angle between the direct signal and the reflected signal. In order to obtain GPS multipath variations information caused by surface reflection in SNR, the multipath effect must be separated from the received SNR observations. Combined with Eq. (1), it can be seen that the influence of multipath effect on SNR mainly in the satellite elevation angle, that is, the value of Ad is much larger than Ar. The trend term Ad (shown in the red box of Fig. 2) is eliminated using the low-order polynomial. The research shows that the SNR residual sequence is mainly caused by low-altitude multipath effects, which provides important data for multipath inversion of surface parameters. Figure 3 shows the relationship between multipath and elevation angle. The amplitudes of reflected multipath signal [18] can be expressed as follows:

$$A_r = A_d \cos\left(\frac{4\pi h}{\lambda}\sin\theta + \varphi\right) \tag{3}$$

Combined with Fig. 1, it can be seen that in Eq. (3), λ is the carrier wavelength, θ is the elevation angle of the satellite and h is the vertical reflection distance. If t = sinθ, f = 2h/λ, then Eq. (3) can be simplified as the standard cosine function expression is:

$$A_r = A_d \cos(2\pi f t + \varphi) \tag{4}$$

The frequency f in Eq. (4) contains the vertical reflection distance parameter h in Fig. 1. If the spectrum analysis is performed on the Eq. (4), the frequency f can be obtained. If the spectrum analysis is performed on the Eq. (4), the frequency f can be obtained. Since Lomb-Scargle algorithm (L-S) transform can effectively extract weak periodic signal from the time-domain sequence, the false signal generated by the non-uniform time domain sequence can be weakened to some extent. Therefore, L-S is adopted for spectral analysis in this paper. Through the L-S spectral analysis of the SNR residual series, the frequency f of the reflected multipath signal can be obtained. The vertical reflection distance h can be calculated by f = 2h/λ. The antenna height H in Fig. 1 is a known quantity that had actually been measured, so the thickness of snow cover can be obtained by hsnow = H-h. Thereby, snow depth measurement using SNR observations is achieved.

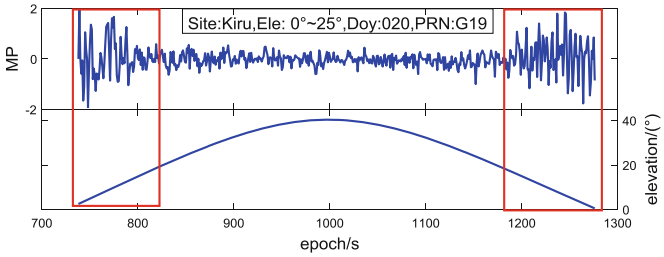


Fig. 3. Relationship between multipath and elevation angle

3 Experiment and Analysis

To verify the relationship between SNR and snow depth inversion at different low elevation angles, the original GPS data of Sweden KIRU station for 2016 are downloaded. The sampling rate is 30 s and the instrument height is 1 m. The station also provides SNR data of L1 and L2 bands. According to the above principle of snow depth principle, the SNR data of L1 and L2 bands of the KIRU Station elevation angles of 0°–20°, 0°–25° and 0°–30°are extracted for experiment. Due to the space limitations, the experiment only use the SNR of 0°–25° L1 and L2 bands as example to illustrate the snow depth. Figure 4 shows the SNR residual sequence of L1 and L2 bands after trend terms eliminated at 0.48 m and 0 m, and the corresponding results of L-S spectral analysis results.

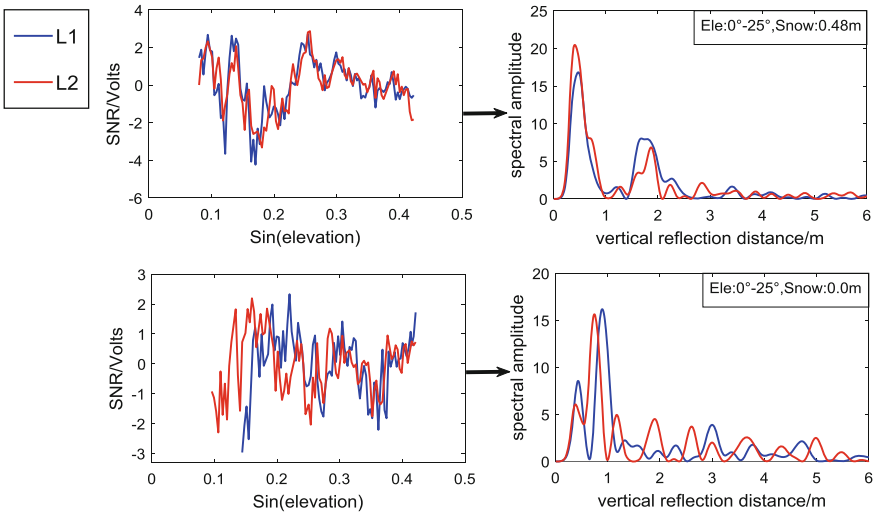


Fig. 4. L2 waveband SNR variations after trend term and L-S results

As can be seen from Fig. 4, it can be seen that both the SNR data of L1 and L2 bands are all related to the variation in snow depth. With the decrease of the snow thickness, the oscillation frequency of the SNR residual sequence of L1 and L2 bands gradually becomes larger, which may be due to the increase of summer temperature and the snow cover gradually melted and made signal received by the instrument contain other surface signal, the amplitude of L-S spectrum also changes, and the vertical distance moves to the right with the decrease of snow thickness.

In the L-S spectral analysis diagram on the right side of Fig. 4 (a) and (b), the horizontal axis represented the distance from the phase center of the receiver antenna to the snow surface. It is the snow depth distance h, which can be calculated by $f = 2h/\lambda$, using the frequency f obtained from spectral analysis. The vertical axis is the amplitudes of L-S spectrums of various frequencies of the SNR reflected signal that are

input. The distance corresponding to the peak of amplitudes of L-S spectrums is the effectively detected vertical reflection height h, which can be calculated by $f = 2h/\lambda$. From Fig. 1, H is a known actual value and finally the inverted snow depth can be obtained by $h_{snow} = H\text{-}h$. The results show that the snow thickness variation is related to the energy spectrum variation of the low-altitude SNR reflection component, which verifies that the SNR spectrum will change with the thickness of the snow, providing a new method for detecting the snow thickness using the SNR reflection signal.

According to the principle of snow depth detection using GNSS-MR technology, the GPS data at KIRU Station from January to May 2016 is processed to obtain, the vertical height from the receiver antenna to the reflecting surface. To further analyze the relationship between the measured snow depth and SNR, the inverted snow depth of L1 and L2 bands at KIRU Station at different elevation angles are compared with the situ measurements. As shown in the Figs. 5, 6 and 7, the horizontal axis represents the days of year, while the vertical axis represents the measured snow depth and the inverted snow depth of L1 and L2 bands.

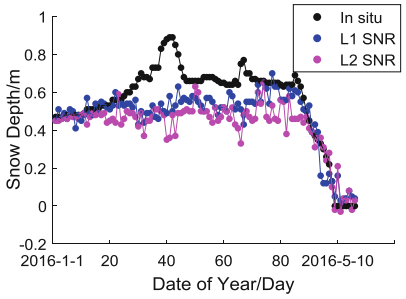


Fig. 5. 0°–20° comparison diagram of Situ measured value and inverse value

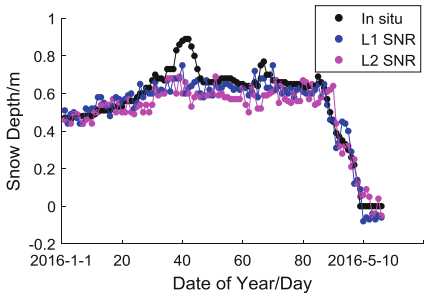


Fig. 6. 0°–25° comparison diagram of situ measured value and inverse value

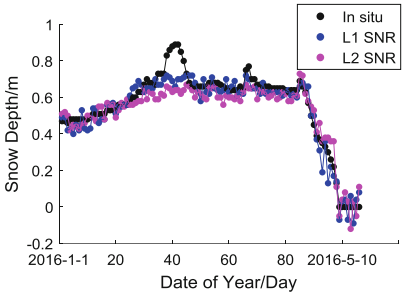


Fig. 7. 0°–30° comparison diagram of Situ measured value and inverse value

By analyzing Figs. 5, 6 and 7, the experiment has the following conclusions:

(1) The elevation angle is range 0°–20°, the correlation between the inverted snow depth and the measured snow depth is analyzed. The correlation coefficient at L1 band is R = 0.86 and the root mean square error is RMSE = 0.08 m. The

correlation coefficient the L2 band is R = 0.79 and the root mean square error is RMSE = 0.08 m. The snow depth is 0.13 m–0.59 m, the inverted values of L1 and L2 bands are in good agreement with the measured values. At the actual snow depth is 0.6 m–0.89 m, the inverted snow depth of L1 and L2 bands are poor. Perhaps due to the continuous snowfall, the thickness of snow cover gradually approached the instrument height and caused the instrument unstable when receiving satellite signal and finally unable to receive the signal.

(2) The elevation angle is range 0°–25°, the correlation analysis is conducted between the inverted snow depth and the measured snow depth. The correlation coefficient at L1 band is R = 0.95 and the root mean square error is RMSE = 0.06 m. The correlation coefficient the L2 band is R = 0.93 and the root mean square error is RMSE = 0.06 m. The elevation angle is range 0°–30°, a correlation analysis is conducted between the inverted snow depth and the measured snow depth. The correlation coefficient the L1 band is R = 0.96 and the root mean square error is RMSE = 0.06 m. The correlation coefficient the L2 band is R = 0.94 and the root mean square error is RMSE = 0.05 m. The inversion snow depth of L1 and L2 bands inversion agrees well with the actual value. When the snow depth is 0.83 m–0.89 m, the inversion values is improved from 0°–20°.

(3) As can be seen from the Figs. 5, 6 and 7, even if the elevation angle is increased and the time series is extended, when the snow thickness is close to the instrument height or even higher than the instrument, the signal would change when it penetrating the snow, making the instrument unstable when receiving signal. The inverted snow depth can't better reflect the variations in the actual snow depth. As a result, the received signal of the instrument is unstable, the inversion of the snow depth value cannot better reflect the actual snow depth changes. In addition, the detection of snow thickness is also affected by season, wind, temperature and other factors. First of all, the KIRU station is seasonal snowfall, and the thickness of snowfall changes every year, and with the arrival of May in summer in May, the air temperature gradually rises and the snow cover melted faster. The snow cover is fluffy, and the wind can blow the snow away. However, the snow depth recorded by the snow depth in situ measurements is only the variation at a certain time, and the disturbance factors such as wind and temperature are not excluded. During the snow-free period, the inverted snow depth of L1 and L2 bands in three groups of different elevation angles fluctuated around the measured values. Although the inversion value is small compared with the measured value, due to different surface roughness, the air temperature and surface temperature are higher in summer, and the data received by the receiver contains various mixed signal. Therefore, the availability of the inverted values at this stage remains to be verified.

4 Conclusion

In this paper, the KIRU station is used as the data source to extract the SNR data of L1 and L2 bands at low altitude angles of 0°–20°, 0°–25°, 0°–30°, and the snow depth data is obtained by L-S spectrum analysis. The accuracy and difference of the snow depth

detected by different low elevation angles L1 and L2 bands are pointed out by comparing the inversion values and measured values detected in the L1 and L2 bands under different thicknesses of snow. The elevation angle of 0°–25° and 0°–30°, the inverted snow depth value is better. The accuracy of inversion can be improved by properly increasing the altitude angle and obtaining the long time series. Since the data used is only the information collected by one station, whether the experimental results are applicable to other stations still needs further study, and the signal received by the receiver are mixed signal. How to extract more "pure" effective data is one of the key work in the future research.

Acknowledgements. This work is 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). The authors 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. Martin-Neira M (1993) A passive reflectometry and interferometry system (PARIS): application to ocean altimetry. ESA J 17(4):331–355
2. Auber JC, Bibaut A, Rigal JM (1994) Characterization of multipath on land and sea at GPS frequencies. In: Proceedings of the 7th international technical meeting of the satellite division of the institute of navigation (ION GPS 1994), pp 1155–1171
3. Qian X, Jin S (2016) Estimation of snow depth from GLONASS SNR and phase-based multipath reflectometry. IEEE J Sel Top Appl Earth Obs Remote Sens 9(10):4817–4823
4. Zhang S, Dai K, Nan Y et al (2018) Preliminary research on GNSS-MR for snow depth. Geomat Inf Sci Wuhan Univ 43(2):234–240
5. Jin S, Qian X (2016) Snow depth variations estimated from GPS L1C/A signal to noise ratio data. In: Geoscience & remote sensing symposium. IEEE, pp 1980–1981
6. Komjathy A, Maslanik J, Zavorotny VU et al (2000) Sea ice remote sensing using surface reflected GPS signal. In: Geoscience and remote sensing symposium. IEEE 2000 International, Honolulu, HI, pp 2855–2857
7. Jacobson MD (2010) Snow-covered lake ice in GPS multipath reception-theory and measurement. Adv Space Res 46(2):221–227
8. Yang M, Cao Y (2014) The following experiment of sea ice observation using GNSS-R signals. GNSS World China 39(4):51–54
9. Larson KM, Small EE, Gutmann E et al (2008) Using GPS multipath to measure soil moisture fluctuations: initial results. GPS Solut 12(3):173–177
10. Rodriguez-Alvarez N, Bosch-Lluis X, Camps A et al (2009) Soil moisture retrieval using GNSS-R techniques: experimental results over a bare soil field. IEEE Trans Geosci Remote Sens 47(11):3616–3624
11. Minsi AO, Jianjun Z, Youjian HU et al (2015) Comparative experiments on soil moisture monitoring with GPS SNR observations. Geomat Inf Sci Wuhan Univ 40(1):117–120 and 127

12. Garrison JL, Katzberg SJ, Hill MI (1998) Effect of sea roughness on bistatically scattered range coded signal from the global positioning system. Geophys Res Lett 25(13):2257–2260
13. Liu L, Feng H, Chen W et al (2017) Inversion of sea level based on signal-to-noise ratio of GPS. J Guilin Univ Technol 37(4):629–634
14. Larson KM, Gutmann ED, Zavorotny VU et al (2009) Can we measure snow depth with GPS receivers? Geophys Res Lett 36(17):L17502
15. Ozeki M, Heki K (2012) GPS snow depth meter with geometry-free linear combinations of carrier phases. J Geod 86(3):209–219
16. Tabibi S, Nievinski FG, Van Dam T et al (2015) Assessment of modernized GPS L5 SNR for ground-based multipath reflectometry applications. Adv Space Res 55(4):1104–1116
17. Yu K, Ban W, Zhang X 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
18. Bilich A, Larson KM, Axelrad P (2004) Observations of Signal-to-Noise Ratios (SNR) at geodetic GPS site CASA: implications for phase multipath. In: The Centre for European Geodynamics and Seismology, Boulder

Dual-Frequency ISB Estimation and Stability Analysis of BDS/GPS

Mingkai Zhang(✉), Zhiping Liu, and Jieqing Yu

School of Environment Science and Spatial Informatics,
China University of Mining and Technology, Xuzhou, China
mkazhang@126.com, zhpnliu@163.com, zhpliu@cumt.edu.cn,
yujieqing@cumt.edu.cn

Abstract. The single-frequency ISB of BDS/GPS is affected by factors such as ionosphere model error and satellite hardware delay. It is difficult to assess accurately and effectively. In this study, the dual-frequency ISB function model of BDS/GPS is derived based on dual-frequency ionosphere-free combination. This method can more accurately determine the time system deviation and hardware delay deviation in receiver. To further analyze the influencing factors and investigate the relation of different receivers (firmware version) and antenna to ISB, 36 MGEX stations' observation data in 2017 are selected. BDS/GPS ISBs series calculated from JAVAD, LEICA, SEPT and TRIMBLE four brand receivers and antennas are analysed in short-term (epoch, day) and long-term (year) time scales. The results show that the ISB variation of diverse brands receiver. In the short-term, the ISB of the above four brand receivers are 40–50 ns, 80–110 ns, 70 ns and 70 ns, respectively. All of above ISBs have periodic time-varying characteristics. The ISB time series of each brand receivers have a consistent but unstable trend in the long-term, and not affected by seasonal factors.

Keywords: GNSS · Inter-system bias · Dual-frequency ISB mode · Timing analysis · Receiver

1 Introduction

With the development of GPS modernization, GLONASS and GALILEO have increased the coincidence of GPS L5 frequency signals. Correspondingly, China's new generation Beidou-3 satellite navigation system has also increased the frequency of B2a signal [1]. The recombination frequency provides technical support for the navigation positioning of multi-system tight combination. The recent dramatic development of multi-GNSS fusion can fully utilize the GNSS multi-constellation and multi-frequency signal, which greatly improve both the positioning accuracy, satellite availability and the positioning reliability.

However, inter-system bias (ISB), which is an inter system deviation between the receiver clock errors of different systems, has to be considered in multiple GNSS fusion positioning. ISB is derived from the difference of time system and the reference frame of different systems, and the hardware delay of the receiver [2–4]. In [5], the

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 220–231, 2019.
https://doi.org/10.1007/978-981-13-7751-8_23

GPS/GIOVE tight combination method is utilized, and the two system receivers' single difference ISB are solved in GPS/GIOVE double difference combination positioning of the GIOVE observation cooperation network(CONGO). In [6], the Galileo/GPS single-frequency DD method is further used, and the differential ISB (DISB) of code and carrier phase observation in two receiver is obtained, which show code DISB reflects the real hardware delay changes between the two systems of the receiver and carrier phase DISB is affected by ambiguity estimation. In [7], the SPP method is used to indicate the difference in the BDS/GPS receiver, and results show the BDS/GPS ISB consistency varied with the type of BDS satellite orbits implemented respectively by the GEO, IGSO, and MEO satellite of SPP. In paper [8], the short-term forecast model of ISB was established, and the PPP convergence time of BDS/GPS was significantly shortened by using the model.

In the research of ISB, the short-term stability of single-frequency ISB is widely studied. However, various receivers and antenna types, and other influencing factors of ISB variation in long-term are not solved perfectly. To accurately estimate the receiver ISB of BDS/GPS, this study investigates BDS/GPS combined PPP method to obtain BDS/GPS dual-frequency ISB using common clock difference, and then analysed ISB time-variation characteristics and influencing factors. Therefore, 36 MGEX monitoring stations data, considering different receiver, firmware and antenna types, are selected to obtain the ISB time series in the year of 2017, and the time series analysis method is employed to verify ISB varying characteristics in long-term time.

2 BDS/GPS ISB Calculation Model

Without considering the observed noise, the code observation equation is expressed as follows.

$$P_k = \rho + c \cdot \left(\left(\delta t_r + D_{r,k} \right) - \left(\delta t^s + D_k^s \right) \right) + \frac{\mathrm{A}}{f_k^2} + d_{\mathrm{trop}} \tag{1.1}$$

Wherein, P_k the code observation value representing the frequency $k(k = 1, 2, 3 \cdots)$, ρ refers to the geometric distance of the satellite, and c is the speed of light, δt_r and δt_s are the receiver and the satellite clock difference without hardware delay, respectively, $D_{r,k}$ and D_k^s the hardware delay of the receiver and the satellite, respectively. f_i the frequency of the point k, $\mathrm{A}(\mathrm{A} = 40.28 \int Ne\mathrm{d}s)$ is the ionospheres' reflection coefficient, which d_{trop} is the troposphere delay.

Since ionosphere-free combination is often used in positioning solution, which has two main advantages for data processing: First, it can make full use of MGEX [9] high-precision satellite clock and orbit products, improve the accuracy of the solution; the second is to avoid the error inherent in the use of ionosphere data (such as GIM) and its mapping function (such as MSLM). The dual-frequency ionosphere-free combined observation equation using two frequencies is expressed as follows.

$$\begin{aligned} P_{\mathrm{LC}} &= \rho + c \cdot (\Delta t_r - \Delta t^s) + d_{\mathrm{trop}} \\ \Delta t_r &= \delta t_r + \frac{f_m^2 D_{r,m} - f_n^2 D_{r,n}}{f_m^2 - f_n^2} \\ \Delta t^s &= \delta t^s + \frac{f_m^2 D_m^s - f_n^2 D_n^s}{f_m^2 - f_n^2} \end{aligned} \tag{1.2}$$

Among them, P_{LC} the combined code observation value, Δt_r is the receiver clock error including the receiver hardware delay combination, and Δt^s the satellite clock error including the satellite hardware delay combination can be directly given by the precision satellite clock products.

In the design of multi-GNSS receiver, GPST is generally used as the reference time, and the integer second deviation of GPST and BDT (14 s) in initial BDS observation is removed, to avoid the BDS code observation value far exceeds the true value [10]. ISB therefore contains the fractional part of inter-system time deviation. The dual-frequency ionosphere-free combined code observation equation and the dual-frequency ISB function model for BDS/GPS are respectively expressed as.

$$\begin{aligned} P_{\mathrm{LC}}^{(G)} &= \rho + c \cdot \left(\Delta t_r^{(G)} - \Delta t^{s,(G)} \right) + d_{\mathrm{trop}} \\ P_{\mathrm{LC}}^{(C)} &= \rho + c \cdot \left(\Delta t_r^{(G)} + ISB - \Delta t^{s,(C)} \right) + d_{\mathrm{trop}} \end{aligned} \tag{1.3}$$

$$\begin{bmatrix} P_{LC}^{(G1)} \\ P_{LC}^{(G2)} \\ \vdots \\ P_{LC}^{(Gs)} \\ P_{LC}^{(C1)} \\ P_{LC}^{(C2)} \\ \vdots \\ P_{LC}^{(Cs)} \end{bmatrix} = \begin{bmatrix} a_X^{G1} & a_Y^{G1} & a_Z^{G1} & 1 & 0 \\ a_X^{G2} & a_Y^{G2} & a_Y^{G2} & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_X^{Gs} & a_Y^{Gs} & a_Y^{Gs} & 1 & 0 \\ a_X^{C1} & a_Y^{C1} & a_Z^{C1} & 1 & 1 \\ a_X^{C2} & a_Y^{C2} & a_Z^{C2} & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_X^{Cs} & a_Y^{Cs} & a_Z^{Cs} & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} dx \\ dy \\ dz \\ c \cdot \Delta t_{rv}^G \\ c \cdot ISB \end{bmatrix} + \boldsymbol{b} \tag{1.4}$$

Where C and G respectively represent BDS and GPS, and then the numbers represent the respective satellite numbers, which a is linearization coefficients, $\boldsymbol{b}$ is the vector of constant terms from the linearized pseudorange equations using initial or updated values of the parameters. According to (1.2) and (1.3), the ISB composition of BDS/GPS in the receiver can be further obtained.

$$ISB = t_{\mathrm{BGTO}} + \delta D_{r,C-G} \tag{1.5}$$

$$\begin{aligned} \delta D_{r,C-G} &= D_{r,C} - D_{r,G} \\ &= \frac{f_m^{2(C)} D_{r,m}^{(C)} - f_n^{2(C)} D_{r,n}^{(C)}}{f_m^{2(C)} - f_n^{2(C)}} - \frac{f_m^{2(G)} D_{r,m}^{(G)} - f_n^{2(G)} D_{r,n}^{(G)}}{f_m^{2(G)} - f_n^{2(G)}} \end{aligned} \tag{1.6}$$

It can be seen from (1.5) and (1.6) that ISB consists of two parts, among t_{BGTO} is the BDS-to-GPS Time Offset (BGTO), which $\delta D_{r,C-G}$ is the hardware delay difference between the two system signals in the receiver. BDS and GPS time systems maintained by their respective master stations using high-precision atomic clocks, respectively named BDT and GPST. The difference is that BDT refers to the National Time Service Centre of the Chinese Academy of Sciences UTC (NTSC) [1, 11], GPST Referring to the U.S. Naval Observatory UTC (UNSO) [12], BDT and GPST establish contact with UTC through UTC (NTSC) and UTC (UNSO) respectively.

$$\begin{aligned} t_{\mathrm{BGTO}} &= t_{\mathrm{BDS}} - t_{\mathrm{GPS}} \\ &= t_c + t_u - t_g \end{aligned} \tag{1.7}$$

$$\begin{aligned} t_c &= t_{\mathrm{BDS}} - t_{\mathrm{UTC(NTSC)}} \\ t_g &= t_{\mathrm{GPS}} - t_{\mathrm{UTC(USNO)}} \\ t_u &= t_{\mathrm{UTC(NTSC)}} - t_{\mathrm{UTC(USNO)}} \end{aligned} \tag{1.8}$$

where: t_c and t_g are time synchronization parameters (modulo 1 s) of BDT and UTC (NTSC), GPST and UTC (UNSO), respectively, and t_u represent time synchronization parameters of UTC (NTSC) and UTC (UNSO). It can be seen that the value of ISB will also change due to the changes of BDT and GPST.

In addition to the differences in time systems, the coordinate framework of both BDS and GPS is different, although existing studies have shown there are only centimetre differences in the system reference frame of CGCS2000 and WGS-84 (G1150) [13], but the difference between the two frameworks cannot be ignored especially the high orbital characteristics of Beidou GEO and IGSO satellites. To reduce the influence of reference frame difference on ISB, several research groups use coordinate transformation method to convert BDS satellite coordinates from CGCS2000 to WGS-84 (G1150) [7], but this method will also bring about the non-self-consistent problem of BDS's own coordinates and clock products. To solve this problem and improve the accuracy of the ISB calculation, it is more reasonable to select the precision orbit and clock products in the same reference frame.

3 BDS/GPS ISB Analysis

This paper uses the 30-second sampling interval observation data of the MGEX monitoring station in 2017, and the GFZ(German research centre for geosciences) precision orbit and clock products [14] to perform PPP calculation to obtain high-precision BDS/GPS ISB value. In order to study the effects of different receivers, firmware versions and antennas on ISB, the observation data from four brands JAVAD, LEICA, SEPT and TRIMBLE were selected, and the distribution of stations of each brand receiver was carried out in Fig. 1, the receiver arrangement information of each station is shown in Table 1. Figure 2 shows the configuration information of the above 36 stations in the long-term, the different colors representing the configuration changes in the receiver. Since the GNSS new satellite constellation and signal generation and

update in recent years, it can be seen that the station receiver and its internal firmware version are frequent changing, which is most active during the period 2016 to 2017. While the station antenna changes are less frequent. Considering the variation of configuration information at station receiver, the experimental design is divided into two parts: (1) BDS/GPS dual-frequency ISB timing changes in short-term (days and weeks); (2) long-term (one-year) BDS/GPS Dual-frequency ISB timing changes. By comparing the sequence of ISB values at different time scales, the influencing factors of ISB are analyzed.

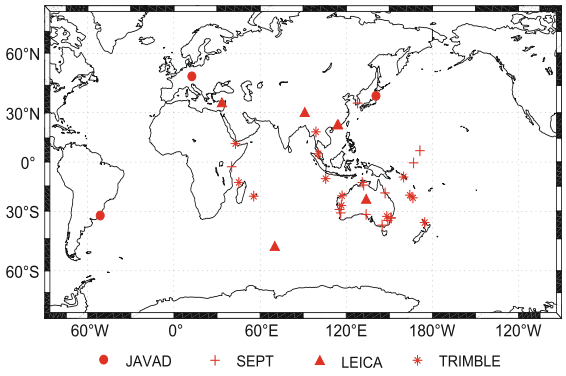
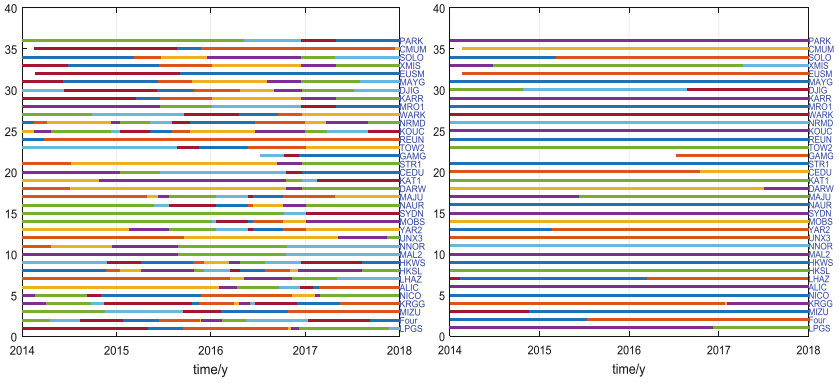


Fig. 1. The distribution of various brands of receivers



(a) Station receiver and firmware version changes **(b) Station antenna changes**

Fig. 2. Station hardware and firmware configuration changes in receivers (different colors represent changes of configuration information)

Table 1. Site receiver configuration information and its ISB result statistics

Band	Site	Receiver	Firmware	Antenna	Epoch STD (ns)	Day STD (ns)	ISB Mean (ns)
JAVAD	LPGS	JAVAD TRE_3 DELTA	3.6.8	JAVRINGANT_G5T	7.21	2.60	72.94
	WTZZ	JAVAD TRE_G3TH DELTA	3.6.9	LEIAR25.R3	6.92	1.93	65.88
	MIZU	JAVAD TRE_G3TH DELTA	3.6.7	JAVRINGANT_G3T	4.29	2.57	73.36
LEICA	KRGG	LEICA GR10	6.523	LEIAR25.R4	3.35	1.95	135.69
	NICO	LEICA GR25	6.523	LEIAR25.R4	3.75	2.56	134.99
	ALIC	LEICA GR25	6.523	LEIAR25.R3	2.95	2.26	130.76
	LHAZ	LEICA GR25	6.523	LEIAR25.R4	3.52	2.72	105.64
	HKSL	LEICA GR50	7.002	LEIAR25.R4	3.06	2.58	70.42
	HKWS	LEICA GR50	7.002	LEIAR25.R4	3.30	2.52	69.79
SEPT	MAL2	SEPT POLARX4	2.9.5	LEIAR25.R4	2.87	1.81	52.72
	NNOR	SEPT POLARX4	2.9.5-extref1	SEPCHOKE_MC	2.74	1.94	53.65
	UNX3	SEPT POLARX4	2.9.3	LEIAR25.R3	4.09	2.34	50.95
	YAR2	SEPT POLARX4TR	2.9.6	AOAD/M_T	2.62	2.04	45.51
	MOBS	SEPT POLARX4TR	2.9.6	ASH701945C_M	3.51	2.07	66.14
	SYDN	SEPT POLARX4TR	2.9.6	ASH701945C_M	3.75	2.28	70.06
	NAUR	SEPT POLARX4TR	2.9.6	LEIAR25.R3	3.64	2.59	52.90
	MAJU	SEPT POLARX4TR	2.9.6	JAVRINGANT_DM	3.70	2.65	53.98
	DARW	SEPT POLARX5	5.10	JAVRINGANT_DM	3.34	2.24	39.70
	KAT1	SEPT POLARX5	5.1.1	LEIAR25.R3	3.01	2.07	39.31
	CEDU	SEPT POLARX5	5.10	AOAD/M_T	2.56	2.15	42.58
	STR1	SEPT POLARX5	5.10	ASH701945C_M	2.81	2.19	60.24
	GAMG	SEPT POLARX5TR	2.9.6	LEIAR25.R4	2.76	2.22	49.81
	TOW2	SEPT POLARXS	2.9.6	LEIAR25.R3	3.28	2.53	53.63

(*continued*)

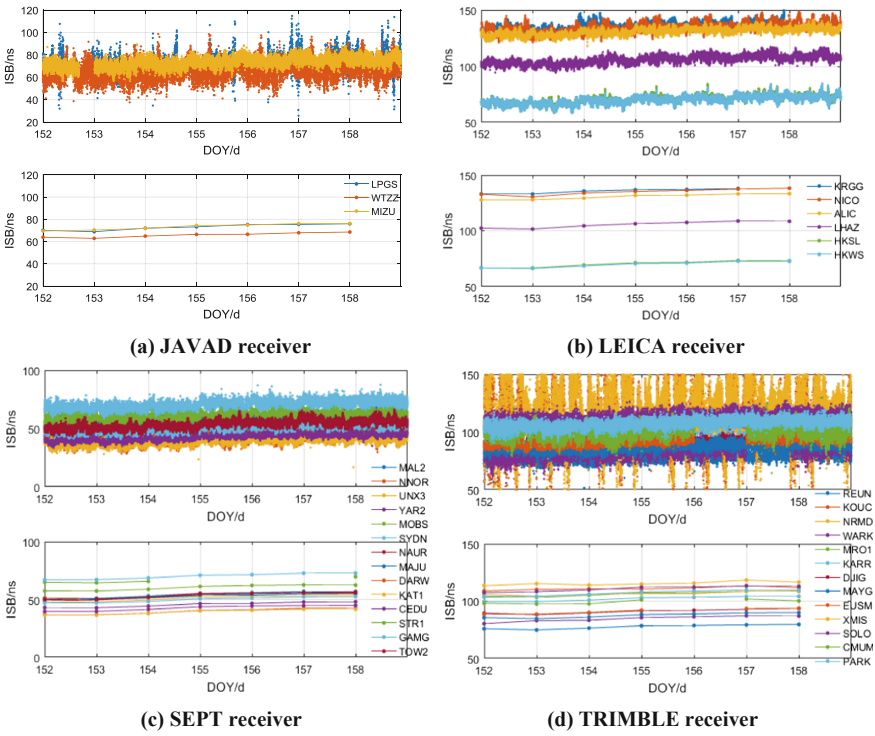
Table 1. (*continued*)

Band	Site	Receiver	Firmware	Antenna	Epoch STD (ns)	Day STD (ns)	ISB Mean (ns)
TRIMBLE	REUN	TRIMBLE NETR9	4.85	TRM55971.00	3.07	1.75	77.70
	KOUC	TRIMBLE NETR9	Nav 5.22	TRM57971.00	18.68	1.36	111.12
	NRMD	TRIMBLE NETR9	Nav 5.22	TRM57971.00	19.24	1.53	115.55
	WARK	TRIMBLE NETR9	5.15	TRM55971.00	4.75	2.37	84.77
	MRO1	TRIMBLE NETR9	5.22	TRM59800.00	3.00	2.09	106.36
	KARR	TRIMBLE NETR9	5.22	TRM59800.00	2.99	1.91	102.17
	DJIG	TRIMBLE NETR9	5.15	TRM59800.00	3.12	1.84	91.17
	MAYG	TRIMBLE NETR9	5.15	TRM59800.00	3.36	1.97	87.67
	EUSM	TRIMBLE NETR9	5.01	JAVRINGANT_DM	2.81	2.16	91.18
	XMIS	TRIMBLE NETR9	5.22	JAVRINGANT_DM	2.96	1.43	106.74
	SOLO	TRIMBLE NETR9	5.22	JAVRINGANT_DM	4.32	2.18	110.97
	CMUM	TRIMBLE NETR9	5.10	JAV_GRANT-G3T	5.16	1.75	99.49
	PARK	TRIMBLE NETR9	5.22	ASH701945C_M	3.45	2.23	107.30

3.1 Short-Term BDS/GPS ISB Timing Changes

According to the composition of type (1.5), the ISB includes the inter-system time deviation and the hardware delay deviation. Therefore, different hardware and firmware effects should be considered in the ISB study, and the receivers and antennas of different brands and types should be analyzed in detail. In the short-term timing change of BDS/GPS ISB, the station observation data of 36 different receivers, firmwares and antennas in the DOY152 to DOY158 period of 2017 are selected, which JAVAD, LEICA, SEPT and TRIMBLE receiver stations' number is 3, 6, 14, and 13, respectively. The time series of ISB results for each receiver is shown in Fig. 3.

Figure 3 shows the ISB time series of JAVAD, LEICA, SEPT, TRIMBLE receivers and different antennas. The upper part of each picture is the epoch solution sequence, and the lower part is the day solution sequence (day epoch solution mean). On the whole, the ISB of each station in the DOY153 to DOY155 period shows an upward trend, and then it is stable. The ISB changes of different receivers are consistent, and both have periodic changes in days. Among them, the periodicity of the



(a) JAVAD receiver **(b) LEICA receiver**

(c) SEPT receiver **(d) TRIMBLE receiver**

Fig. 3. ISB time series for different brands of receivers (the upper part of each subgraph is epoch value, and the lower part is daily mean value)

JAVAD brand receiver ISB is more obvious than that of the other three brands. The WTZZ station has the widest ISB sequence, the fluctuation range is large, and the day period is the strongest. According to the statistical results of Table 1, the JAVAD receivers' standard deviation of epoch calculations(Epoch STD) ISB is above 4 ns, and the variation between epochs is more obvious. The average ISB (Mean, hereinafter referred to as ISB) obtained by different type of JAVAD brands receivers and antennas are not consistent, at about 73 ns, and when using the LEIAR25.R3 antenna, the ISB is reduced to 65.88 ns.

The LEICA brand receivers LEICA GR10 and LEICA GR25 types have ISBs above 100 ns. The difference between receiver and antenna type has little effect on ISB differences, although both NICO and LHAZ stations use the same LEICA GR25 receiver and LEIAR25.R4 antenna, the ISB difference between the two is about 30 ns, which is 134.99 ns and 105.64 ns, respectively. It is not excluded the reason of the individual factors in the receiver. When using the LEICA GR50 receiver, the ISB is around 70 ns, which is a significant reduction compared to the previous types of the brand. Since the inter-system time deviation has the same effect on all stations' ISB, it can be inferred that it is mainly because of the new type LEICA GR50 receiver has improved the BDS/GPS hardware delay difference.

The antenna brands and types used by SEPT brand receivers are more complicated, so there are more sites to choose. When SEPT POLARX4 and SEPT POLARX4TR types receiver use LEIAR25.R4, SEPCHOKE_MC, LEIAR25.R3 and other different brands of antennas, the difference of ISB is not obvious around 50 ns, and the ASH701945C_M antenna is higher, reaching 65 to 70 ns. Similarly, when the SEPT POLARX5 receiver uses other types of antennas, the ISB is about 40 ns, and when using the ASH701945C_M antenna, it is 60.24 ns. It can be seen that the ASH701945C_M antennas in the four types of antennas cause the largest differences in system hardware delay between SEPT receivers. SEPT POLARX5 has a lower ISB than the SEPT POLARX4TR receiver, so the BDS/GPS hardware delays' difference is reduced.

The TRIMBLE brand receiver types are all TRIMBLE NETR9. When using the TRM55971.00 antenna, the ISBs of different receiver firmware versions are different. The ISBs of version 4.85 and 5.15 are not much different, both are about 80 ns. The two stations using version Nav. 5.22 ISB is around 111 ns, and Epoch STD is also about 19 ns, which is much higher than other stations in the same period. As can be seen from Fig. 3(d), the intra-day epoch ISB of the two stations is large, and the main reason is that the firmware version Nav 5.22; when using the same type of antenna, the 5.22 version has an increase compared to the previous version (5.11, 5.01, 4.85), and when the firmware version is constant, its antenna brand and type change have little effect on ISB results.

Finally, the update of the LEICA and SEPT brand receiver types generally reduces the hardware delay of the ISB. In addition to the ASH701945C_M antenna, the SEPT brand receivers use other brands and types of antennas with ISB values of 40 to 50 ns, significantly lower than other brands and types of receiver and antenna combinations. Apart from some types of the LEICA brand receivers, the ISB values of TRIMBLE NETR9 receivers using different brands and types of antennas are generally higher than other brands and types of receivers and antenna combinations. When the antenna type is constant, the ISB value changes when the receiver brand, type or firmware version changes. Therefore, the receiver and its firmware are the main factors determining the size of the ISB, and antenna is the secondary factors.

3.2 Long-Term BDS/GPS ISB Timing Changes

In order to further study the long-term time series changes and influencing factors of ISBs in different brands receivers, the research selects four stations' observation data of MIZU, ALIC, GAMG and MRO1, which from the above four brands of receivers JAVA, LEICA, SEPT and TRIMBLE for the whole year of 2017. The receiver configuration information of each station is shown in Table 2, except that the MRO1 station receiver firmware version changed from 5.15 to 5.22 in DOY121, other station receivers, firmware versions and antennas did not change, the resulting ISB year time series is shown in Fig. 4.

It can be seen from Fig. 4 that the long-term time series of ISBs of the four brands of receivers still have the same trend. In the beginning of 2017, ISB showed a downward trend, the change of ISB in the daytime are more obvious. After the inflection point appeared at DOY068, it rise to the beginning value of the year. The ISB

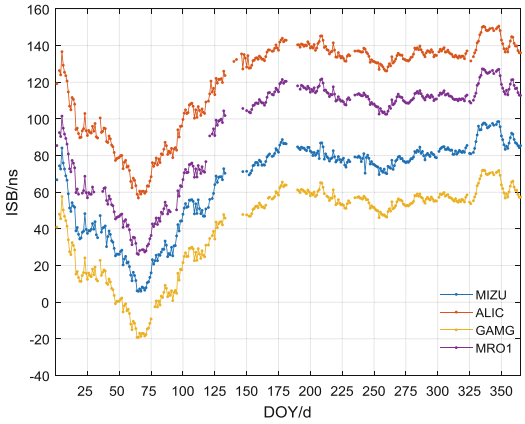


Fig. 4. BDS/GPS ISB one year time series of different receivers

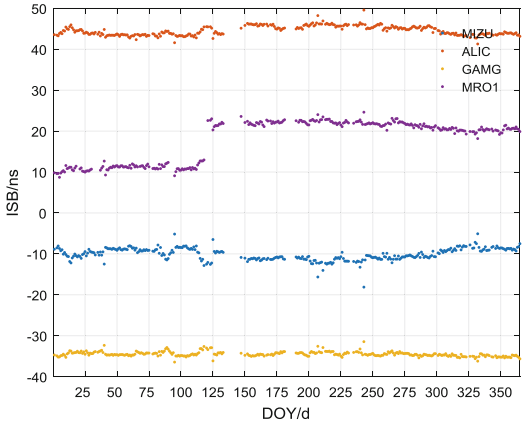


Fig. 5. BDS/GPS ISB detrended year time series for different receivers

Table 2. ISB long-term sequence station information

Site	Lat	Long	Detrended ISB STD (ns)
MIZU	39.14	141.13	1.47
ALIC	−23.67	133.89	1.13
GAMG	35.59	127.92	0.58
MRO1	−26.70	116.64	0.88

is stable in a small range of around 20 ns, and the day-to-day variation of the ISB becomes relatively smooth. The difference between the inflection point and the stabilized ISB is about 80 ns. According to the Table 2, the distribution of the north and south hemispheres of the four stations is uniformity, this phenomenon occurs at the

same time period, which can eliminate the influence of seasonal factors on the receiver. The ISB changes of each station have common factors unrelated to the station receiver, which may be due to the different maintenance of the respective time systems of BDS and GPS. In order to further analyse the ISB relationship between different receivers and remove the influence of public trend items, the four stations ISB are removed from their total day ISB averages (the MRO1 station receiver firmware version changes, so it do not participate in the mean calculation). Four the ISB annual variation sequence of the station without trend item is shown in Fig. 5. It can be seen that except for the influence of firmware factors, the annual ISB of the trend-free items of each station receiver is not changed much. According to the statistical STD of Table 2, the overall change of the ISB of the GAMG station SEPT POLARX5TR receiver is small, and the hardware delay is also relatively stable.

4 Conclusions

A BDS/GPS dual-frequency ISB model augmented by the precise orbit and clock products is presented in detail in this contribution, and the causes and composition of the ISB are also illustrated. Then, the algorithm is evaluated and validated by the comparison of short-term and long-term ISB time-series obtained from different receivers, firmware versions and antennas.

According to the results, we conclude that (1) the ISB of different receivers, firmware versions and antennas has a day-to-day variation in short-term (days). The change of ISB period in JAVAD brand receivers is obvious, and the ISB changes in LEICA, SEPT and TRIMBLE receivers are relatively stable. When the TRIMBLE receiver install with Nav 5.22 firmware version, the ISB fluctuation is significantly larger during the day, and the ISB is extremely unstable; (2) the difference in receiver brand has the greatest impact on the ISB, and different antennas have little effect on ISB of the same receiver. Among the four brands of receivers, the ISB of the SEPT and TRIMBLE NETR9 receivers are about 40 to 50 ns and 80 to 110 ns, respectively. JAVAD and the LEICA GR50 receiver ISB are not much different, both are around 70 ns. The LEICA GR50 ISB are about 130 ns, the ISB value of each brand receiver is sequentially ranked as SEPT < LEICA GR25 (GR10) < JAVAD < TRIMBLE NETR9 < LEICA GR50; (3) in the long term, the ISB time series of the four brand receivers have the same trend, which is basically consistent by the difference of hardware delays, and is not affected by seasonal factors, but the ISB's annual change is about 80 ns. It cannot be considered as a constant amount over a long time scale. Therefore, when performing stability analysis of BDS/GPS ISB, one must take into account not only the influence of hardware delay in the receiver, but also the common factors such as system time deviation. In the short-term, ISB value is consistent for most specific receivers. However, the long-term trend of the ISB has neither stable indication nor forecasting features.

Acknowledgments. The authors gratefully acknowledge IGS Multi-GNSS Experiment (MGEX) for providing GNSS data and products. We appreciate anonymous reviewers for their valuable comments and improvements to this manuscript. Thanks also go to the Natural Science

Foundation of China (41771416), and the Open Foundation of Key Laboratory of Precise Engineering and Industry Surveying of NASMG (PF2017-12).

References

1. China Satellite Navigation Office (2018) BeiDou navigation satellite system signal in space interface control document open service signal B3I (Version 1.0). Beijing
2. Odijk D, Teunissen PJG (2012) Characterization of between-receiver GPS-Galileo inter-system biases and their effect on mixed ambiguity resolution. GPS Solut 17(4):521–533
3. Cai C, Gao Y (2008) Estimation of GPS-GLONASS system time difference with application to PPP. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp 2880–2887
4. Sui X, Shi C, Xu A et al (2018) The Stability of GPS/BDS inter-system biases at the receiver end and its effect on ambiguity resolution. Geomatics and Information Science of Wuhan University, no 2
5. Montenbruck O, Hauschild A, Hessels U (2010) Characterization of GPS/GIOVE sensor stations in the CONGO network. GPS Solut 15(3):193–205
6. Paziewski J, Wielgosz P (2014) Accounting for Galileo–GPS inter-system biases in precise satellite positioning. J Geod 89(1):81–93
7. Zeng A, Yang Y, Ming F, Jing Y (2017) BDS–GPS inter-system bias of code observation and its preliminary analysis. GPS Solut 21(4):1573–1581
8. Jiang N, Xu Y, Xu TH, Xu GC, Sun ZZ, Schuh H (2017) GPS/BDS short-term ISB modelling and prediction. GPS Solut 21(1):163–175
9. Montenbruck O, Steigenberger P, Khachikyan R, Weber G, Langley R, Mervart L et al (2014) IGS-MGEX: preparing the ground for multi-constellation GNSS science. Inside GNSS 9(1):42–49
10. Yu L (2016) Beidou/GPS dual-mode software receiver principle and implementation technology. Electronic Industry Press
11. Han C, Yang Y, Cai Z (2011) BeiDou navigation satellite system and its time scales. Metrologia 48(4):S213–S218
12. Lassiter EM (2006) NAVSTAR GPS Space Segment/Navigation User Interfaces, (Rev. D). ICD-GPS-200 D
13. Montenbruck O, Steigenberger P, Hauschild A (2014) Broadcast versus precise ephemerides: a multi-GNSS perspective. GPS Solut 19(2):321–333
14. Uhlemann M, Gendt G, Ramatschi M, Deng Z (2015) GFZ global multi-GNSS network and data processing results. GPS Solut 143:673–679

Impact of Grid Model on Tropospheric Wet Refractivity Tomography in Multiplicative Algebraic Reconstruction Techniques

Xiaoying Wang, Fuyang Ke(✉), Lianchun Song, and Yunchang Cao

Nanjing University of Information Science and Technology,
Ningliu Road 219, Nanjing 210044, China
kfy_0829@163.com

Abstract. The ground-based GNSS has become a member of global meteorological composite observing system with superiority on the derivation of the atmosphere water vapor compared to traditional observation methods. Grid models decide the spatial distribution of the slant path penetrating the grids and greatly affect the tomography results which is key for further operational water vapor tomography to select optimal grid model in ground-based GNSS. The paper gives the effect of three different grid models on the convergence properties and tomography result in the multiplicative technique to reconstruct the 4D wet refractivity field of troposphere based on simulation method for Shenzhen-Hongkong (SH) GNSS network with data from 15th to 22th August, 2009. The results show that for SH tomography region, the fundamental irregular grid model induces the worst tomography result, for the extended grid model and fundamental regular grid model which one is better depends on the distribution of satellites related to the ground network during the specular time, if there are more slant path signal with low elevation angle coming in, the extended grid model can achieve better results, otherwise the fundamental regular one can get tomography results with higher precision since there are more unknown voxels and no improvement for SWD structure for the extended grid model. In a word, for SH region, the fundamental regular grid model can get best tomography results.

Keywords: Slant wet delay · Algebraic reconstruction techniques (ART) · Water vapor tomography · Ground-based GNSS · Precipitable Water Vapor (PWV)

1 Introduction

Water vapor plays an important role in many of the atmospheric and geophysical phenomena, which include transfer of energy, formation of clouds and weather system [Singh et al. 2014]. Because of its high spatial and temporal variability it is very important to track the status of water vapor in the atmosphere with high temporal-spatial resolution for rainstorm forecasting. There are various techniques for PWV (Precipitable Water Vapor) detection such as radiosonde, water vapor radiometer (WVR), and the GPS (Global Positioning System) is proposed to measure water vapor with features of

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 232–241, 2019.
https://doi.org/10.1007/978-981-13-7751-8_24

high temporal-spatial resolution, high precision and relative economy (Bevis et al. 1992), which is verified by many researches (Rocken et al. 1993; Chen and Liu 2014).

The ground-based GNSS technology can detect the PWV (Precipitable Water Vapor, PWV) (Rocken et al. 1993), SWV (Slant Water Vapor) (Alber et al. 2000) and the 4-dimentional (4D) structure of water vapor in the troposphere (Alber et al. 2000), but for water vapor tomography the result precision is very easy to be unstable because of the weak SWD structure penetrating the grid model.

In a long term, researchers have made a lot of effort to resolve the ill-posedness problem of water vapor tomography equations: (1) improve the SWD model to make its precision better (Eresmaa and Jarvinen 2006); (2) use independent and effective observations to improve the ill-posedness of water vapor tomography equations (Notarpietro et al. 2011; Bender et al. 2011b); (3) use iteration method such as the algebraic reconstruction techniques (ART) to solve the tomography equations which can avoid inversion difficulty and result oscillation in general solving method (Bender et al. 2011a; Wang 2011).

ART has been successfully used to reconstruct the total electron content (TEC) of the ionosphere (Stolle et al. 2006; Jin et al. 2008) and in recent years tried to be used in tropospheric wet refractivity field and water vapour tomography (Bender et al. 2011a; Wang 2011; Notarpietro et al. 2011). The traditional method solving the tomography equations are based on the least square method, there are massive equations and unknown voxels with problems of data management and inversion difficulty, the normal equation is huge sparse matrix, sometimes it is unable to inverse because of the rank deficiency. Later, singular value decomposition (SVD) was used to solve the tomography equations which can accomplish the inversion regardless of rank deficiency (Rohm and Bosy 2011). For the above solving methods it is very difficult to get stable water vapor field because of the weak SWD spatial structure, the small change of input data can induce completely different results. The ART is based on the iteration method to gain water vapor field, which can avoid the inversion problem of huge sparse matrix and the result is stable. The previous research on ART used in tropospheric water vapor tomography focused on the convergence and relaxation parameters and pointed out that multiplicative algebraic reconstruction techniques (MART) iterated fast and got better results, but rarely discussed the impact of grid model on tropospheric wet refractivity tomography. This paper gives the effect of three different grid models on the convergence properties and tomography result in the multiplicative technique to reconstruct the 4D wet refractivity field of troposphere based on simulation method for Shenzhen-Hongkong (SH) GNSS network with data from 15th to 22th August, 2009, which can give advice for the selection of best grid model for operational water vapor tomography based on the ground-based GNSS technique.

2 Data and Method

As shown in Fig. 1, the GNSS network including 4 Shenzhen stations and 11 Hongkong stations (SHG network) is selected to do the simulation experiment. The SHG network covers a latitude range of 22°12′ to 22°36′ and a longitude range of 113° 53′ to 114°17′. The HKNP station has maximum geodetic height equal to 350.672 m, the HKPC station has minimum geodetic height equal to 18.105 m.

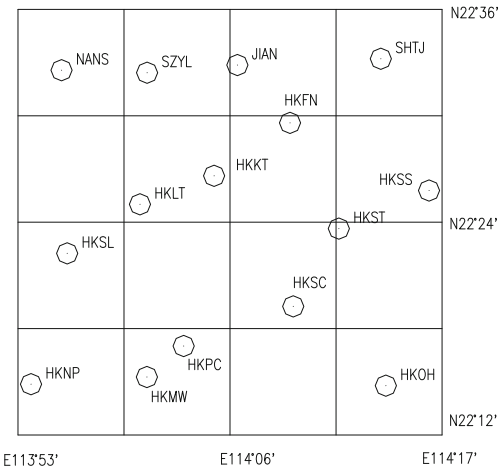


Fig. 1. Distribution diagram for SH GPS network

The precise ephemeris from August 15, 10:00 UTC to August 22, 24:00 UTC are downloaded from the IGS. The simulation test is done based on the ephemeris, the standard atmosphere and the SH network. The wet refractivity field calculated from the standard atmosphere is treated as the true value. The SWD observations are numerically integrated from the true wet refractivity field adding some stochastic error, which is $\sigma_{ZWD} \cdot (\sin \alpha)^{-1}$ for SWD with elevation angle α. The σ_{ZWD} is the error of the zenith wet delay(ZWD) and equal to %3*ZWD.

Three grid models are tested calling basic & regular, extended & regular and basic & irregular respectively. The middle coordinate (B = 22.4, L = 114.1, H = 18.105) of SH network is set to the centre of grid models, the troposphere top is set to 10 km, for basic & regular grid model total voxels are 4*4*20, every voxel is 10 km*10 km*500 m; for extended & regular grid model total voxels are 6*6*20, every voxel is 10 km*10 km*500 m; for basic & irregular grid model total voxels are 4*4*15, its horizontal spatial resolution is same to that of basic & regular grid model, for layers below 5 km every voxel is 10 km*10 km*500 m whereas for layers above that every voxel is 10 km*10 km*1000 m.

The tomography equations are shown as Eq. (1), the principle formula for MART are shown as Eq. (2) (Bender et al. 2011a):

$$A \cdot x = SWD \tag{1}$$

$$x_j^{k+1} = x_j^k \cdot \left(\frac{SWD_i}{\langle A^i, x^k \rangle}\right)^{\frac{\lambda \cdot A_j^i}{\sqrt{\langle A^i, A^i \rangle}}} \tag{2}$$

Where ***SWD*** is the matrix of all SWDs, the ***A*** is the matrix denoting the distance of slant path propagating through the voxel, ***x*** matrix is the wet refractivity field (mm/km) of the troposphere, λ is the relaxation factor. For voxel not penetrated by any slant path

the Gaussian constraint is used to update its value (Bender et al. 2011b), iterations is equal to 100.

Since the true wet refractivity field is known for the whole grid model, four parameters δ, σ, Bias and RMS are used to evaluate the precision of calculated wet refractivity field as shown in Eqs. (3)–(6):

$$\delta = \frac{1}{I}\sum_{i=1}^{I}\left(SWD_i^k - SWD_i^0\right) \tag{3}$$

$$\sigma = \frac{1}{I-1}\sum_{i=1}^{I}\left(\left\{SWD_i^k - SWD_i^0\right\} - \delta\right)^2 \tag{4}$$

$$Bias = \frac{1}{n}\sum_{i=1}^{n}\left(Nw_i^k - Nw_i\right) \tag{5}$$

$$RMS = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\left\{Nw_i^k - Nw_i\right\} - Bias\right)^2} \tag{6}$$

In the above equations, $\mathbf{SWD}_i^k$ denotes the ith back-projection SWD value after the kth iteration, $\mathbf{SWD}_i^0$ denotes the true value for the ith SWD; $\mathbf{Nw}_j^k$ is the solved wet refractivity for jth voxel after the kth iteration,$\mathbf{Nw}_j^0$ is the true wet refractivity for the jth voxel; I denotes the total number of SWDs, n denotes the total number of voxels. The δ and σ can indirectly evaluate the $\mathbf{N_w}$ precision for voxels penetrated by SWDs, the Bias and RMS can directly evaluate the $\mathbf{N_w}$ precision for the total grid, which mainly reflect the $\mathbf{N_w}$ precision for the lower atmosphere since the $\mathbf{N_w}$ value in the lower atmosphere is much larger than that in the upper atmosphere.

2.1 The Impact of Grid Models on Wet Refractivity Tomography in MART

For extended & regular grid model its voxels change from 320 to 720 compared to basic & regular one, but theoretically it can include more SWDs especially with low elevation angle, e.g., for SHG network, the lowest SWD elevation angle decreases from 10° to 7, which is essential for better tomography. For basic & irregular grid model, its voxels decrease from 320 to 240 since for high layer above 5 km the vertical resolution is 1000 m, thus basic & irregular grid model has less unknown voxels. The following tests aim to answer two questions: (1) if extended & regular grid model can achieve better results since generally it includes more SWDs with low elevation angle; (2) can basic & irregular grid model improve the result precision because of decreased unknown voxels?

Two periods are selected to do the simulation test, one is 12:00–12:30 on August 15th, 2009 (hereinafter referred to as Time 1), the other is 0000–0030 on August 16th, 2009 (hereinafter referred to as Time 2). Figures 2 and 3 show the iteration

convergence of statistic results for three different grid models in two periods, Fig. 4 show the elevation angle distribution for extended & regular grid model in two periods, Table 1 shows the statistic results. They tell us that δ and σ parameters iterate fast, Bias and RMS keep stable after initial several convergences.

Table 1. Result statistics for three different grid models (iterations = 100)

Grid model	Period	Effective/Total SWD	Voxels penetrated/total	δ (mm)	σ (mm)	Bias (mm/km)	RMS (mm/km)
4*4*20	Time1	1763/8027	311/320	−0.43	2.08	1.65	5.09
	Time2	2624/7389	312/320	0.01	1.16	1.12	3.37
4*4*15	Time1	1763/8027	231/240	−0.46	2.27	2.49	6.03
	Time2	2624/7389	232/240	−0.00	1.19	1.80	4.15
6*6*20	Time1	2698/8027	528/720	−0.50	2.90	1.54	4.77
	Time2	3351/7389	520/720	−0.11	1.10	1.30	3.62

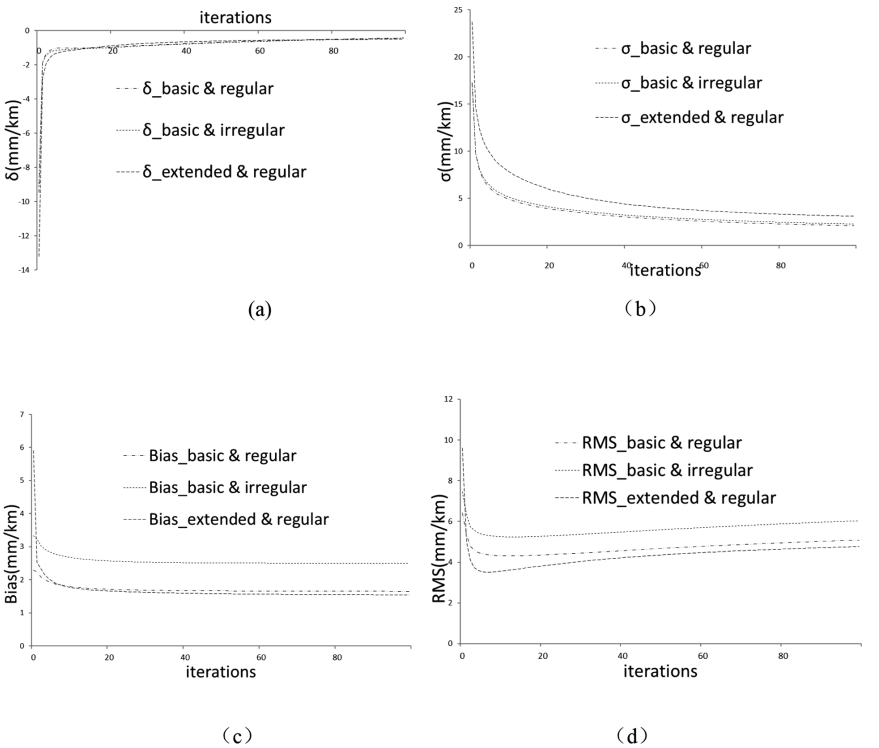


Fig. 2. Statistics for tomography results of different grid model, 12:00–12:30 on 15th August, 2009, (a) δ (b) σ (c) Bias (d) RMS

Table 1, Figs. 2 and 4(a) show that for Time 1, comparing with the basic & regular grid model the extended & regular one includes more SWD observations with low elevation angle, the lowest elevation angle changes from former 14.7° to 11.6°, the number of SWDs with elevation angle between 10°–25° increase from 362 to 721, because the SWDs with low elevation angle have larger error, the δ and σ parameters for extended & regular grid model are higher than that for basic & regular one, but good wet refractivity tomography needs more near horizontal SWDs to improve spatial structure, thus the Bias and RMS of the extended & regular grid model which reflect the whole precision of the wet refractivity field are smaller than that of the basic & regular one.

Table 1, Figs. 3 and 4(b) show that for Time 2, the satellites distribution is relatively concentrated, comparing with the basic & regular grid model the extended & regular one doesn't include more SWD observations with low elevation angle, there is no SWDs with elevation angle between 10°–25°, four parameters of extended & regular grid model show no superiority compared to basic & regular one.

On-orbit satellites distribution relative to the SH network can vary with time very differently, it is not easy to get very even tomography field based on the ground-based GNSS technique and MART method at any time and any area, e.g., the above Time 2 has more effective SWDs than that of Time 1, the tomography field of Time 2 are better

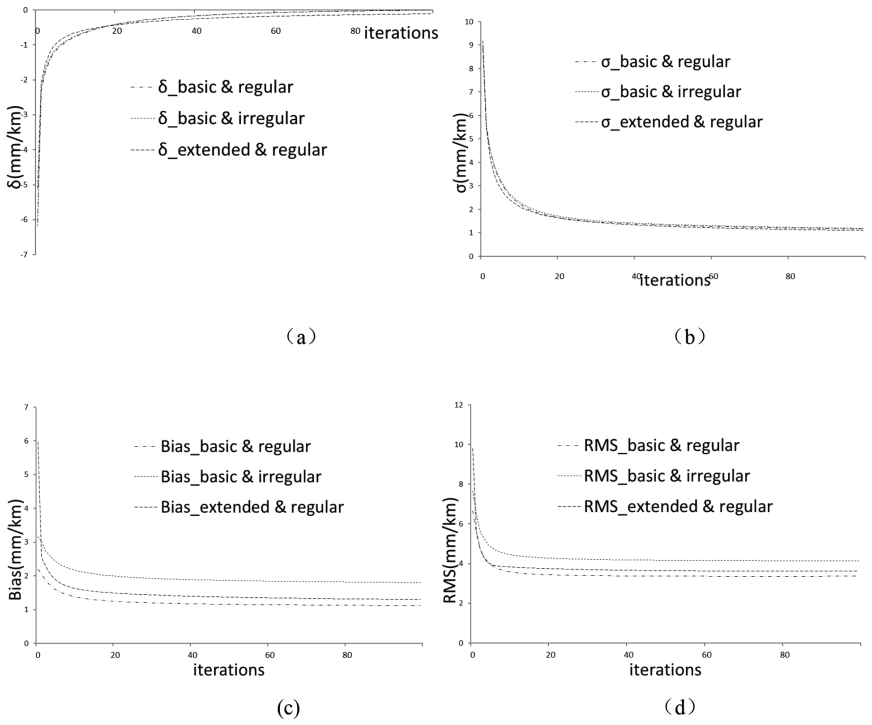


Fig. 3. Statistics for tomography results of different grid model, 0:00–0:30 on 16th August, 2009, (a) δ (b) σ (c) Bias (d) RMS

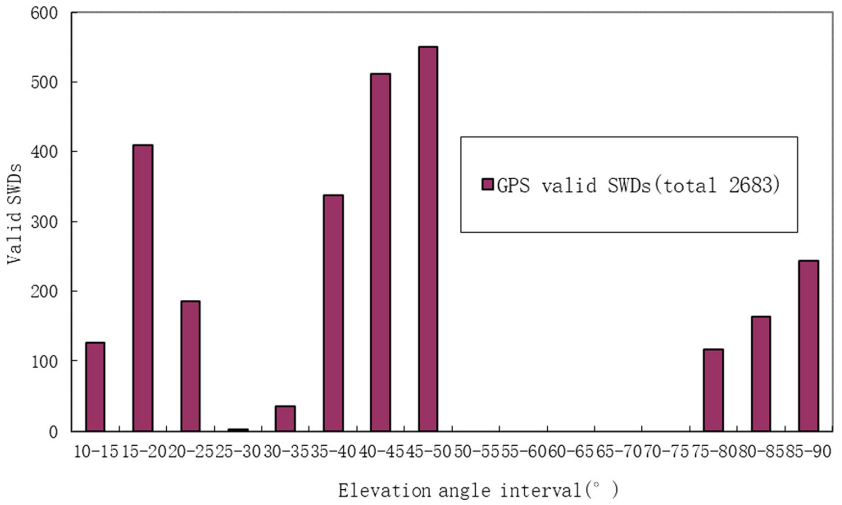


Fig. 4. Elevation angle distribution of SWD in extended model, 12:00–12:30 on 15th August, 2009

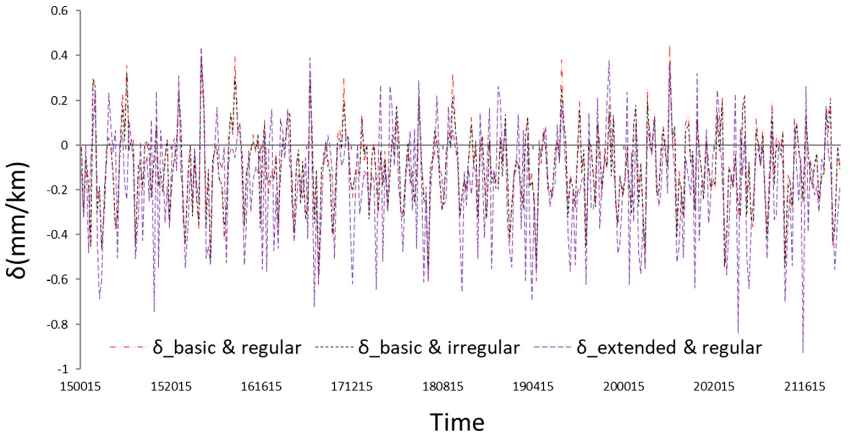


Fig. 5. The δ comparison for three different grid models (iterations = 100)

than that of Time 1 using the same grid model. On the other hand, which grid model can achieve best result for one particular region? It needs more analysis for the on-orbit satellites distribution relative to the ground network, if the extended & regular grid model can include more wide-distributed SWDs, especially those SWDs with lower elevation angle, it can get better tomography field.

For two periods, the basic & irregular grid model gets worst tomography results. Table 1 and Figs. 2 and 3 show that although the basic & irregular grid model has decreased unknown voxels, it doesn't improve the precision of the tomography field. The possible reason may be because that in the basic & irregular grid model the voxels

in high layer above 5 km has low vertical resolution, the wet refractivity in one pretty large voxel are treated as one value, which induces larger sampling error.

Figures 5, 6, 7 and 8 are the δ, σ, Bias and RMS change with time at the 100th iteration from August 15th, 2009 to August 21st, 2009 for three different grid models. Figures 5 and 6 show that the δ and σ for the basic & regular grid model is almost the same to that for the basic & irregular one, but the δ and σ for the extended & regular are obviously larger than those for the other two models, which is because that generally the extended & regular grid model includes more SWDs with low elevation angle with larger error, the δ and σ are calculated from the SWDs and they can only indirectly evaluate the precision of wet refractivity field for voxels penetrated by SWDs. During the seven test days, there are total 336 samples, the σ value for the basic & regular gird model are all smaller than that for the basic & irregular one, whereas for

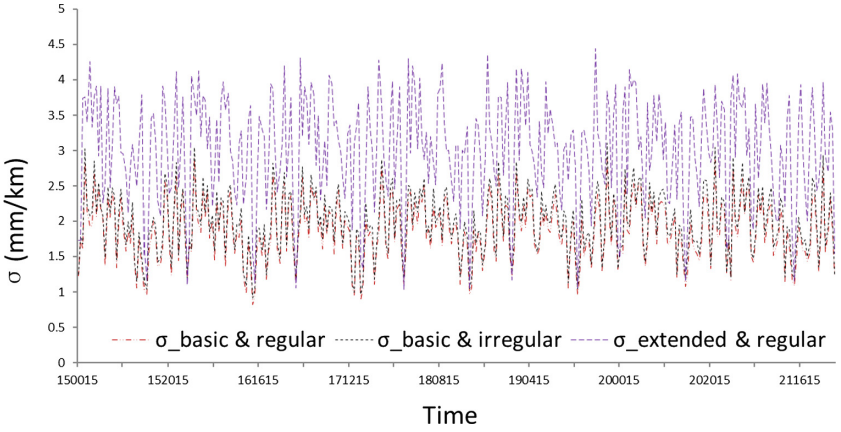


Fig. 6. The σ comparison for three different grid models (iterations = 100)

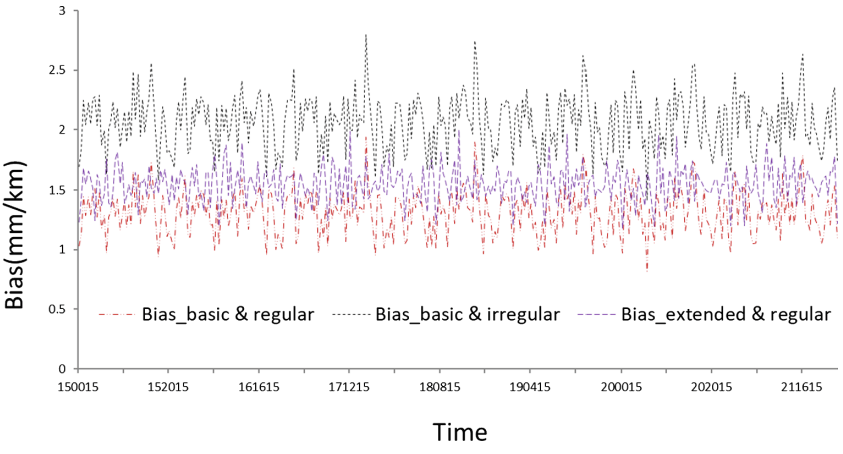


Fig. 7. The Bias comparison for three different grid models (iterations = 100)

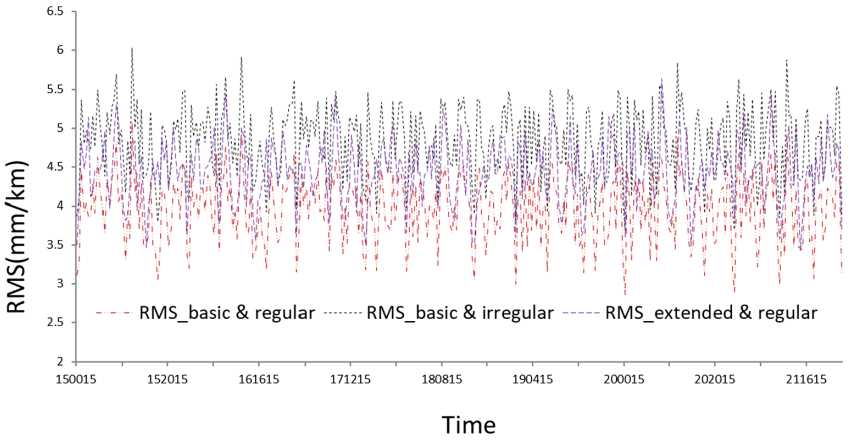


Fig. 8. The RMS comparison for three different grid models (iterations = 100)

δvalue there are only 164 samples in which the basic & regular gird model shows its goodness. So for SH network, when using MART to calculate tomography equations, although the basic & irregular gird model has less unknown voxels, it shows no obvious advantages for improving the tomography result precision, which indicates sparse sampling for high atmospheric layer are not good for wet refractivity tomography.

Figures 7 and 8 show that in most cases for the SH network, the basic & regular grid model can get best tomography result. During the total 336 samples, there are only 52 samples in which the Bias for the extended & regular grid model is smaller than that for the basic & regular, and 48 samples in which the RMS for the extended & regular grid model is smaller than that for the basic & regular, and no one sample in which the Bias and RMS for the basic & irregular grid model is smaller than that for the basic & regular one. On the whole, for SH network it is best to use the basic & regular grid model to do wet refractivity tomography based on the ground-based GNSS technique.

Figures 5, 6, 7 and 8 show the δ, σ, Bias and RMS have obvious characteristics of periodicity. The δ value synchronizes with the σ, and the Bias and RMS also vary with the same trend. For the basic & regular grid model, the δ and σ are smallest on 15:00 UTC per day when the satellites have concentrated distribution and most of the SWDs have high elevation angles; the δ and σ are largest on 02:00 UTC per day when the satellites have dispersed distribution and a lot of SWDs have low elevation angles; the Bias and RMS have maximum value on 09:00 UTC and minimum value on 18:00 UTC per day. The periodicity phenomenon for the above four parameters indicates that the SWDs spatial structure related to the particular SH ground GNSS network repeats by day, when the SWDs have worst spatial structure, the lowest precision for wet refractivity field tomography must be caused. After a detailed analysis, maybe different grid models can be utilized at different tomography time in one day, which can accomplish optimal results for the whole tomography task.

3 Conclusion

For the SH GNSS network, the simulation experiments show that generally the basic & regular grid model get better tomography results for wet refractivity field compared to the basic & irregular and the extended & regular grid model. When the extended & regular grid model includes more SWDs with low elevation angle, it is good for the improvement of SWDs spatial structure, which are expected to increase the precision of the tomography results. Relating to the basic & irregular, the decreased unknown voxels has no goodness for the tomography results, on the contrary because of the sparse sampling for high atmospheric layer above 5 km, the basic & irregular grid model get the worst result. Maybe we can use the basic & regular grid model or the extended & regular one on different tomography time per day after an exact analysis for one particular ground GNSS network, which can lead to optimal tomography results on a whole.

Acknowledgements. This study is funded by a National Natural Science Funds (Grant no. 41674036), NUIST Pre-Research Fund (Grant no. 2014x039), and Key Research & Development Program of Jiangsu Province of China (Grant no. BE2016020).

References

Alber C, Ware RH, Rocken C et al (2000) Inverting GPS double differences to obtain single path phase delays. Geophys Res Lett 27:2661–2664

Bender M, Dick G, Ge M et al (2011a) Development of a GNSS water vapour tomography system using algebraic reconstruction techniques. Adv Space Res 47(10):1704–1720

Bender M, Stosiusa R, Zusa F et al (2011b) GNSS water vapour tomography - Expected improvements by combining GPS, GLONASS and Galileo observations. Adv Space Res 47 (5):886–897

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 97:15787–15801

Eresmaa R, Jarvinen H (2006) An observation operator for ground based GPS slant delays. Tellus A 58(1):131–140

Jin S, Luo OF, Park P (2008) GPS observations of the ionospheric F2-layer behavior during the 20th November 2003 geomagnetic storm over South Korea. J Geod 82(12):883–892. https://doi.org/10.1007/s00190-008-0217-x

Notarpietro R, Cucca M, Gabella M et al (2011) Tomographic reconstruction of wet and total refractivity fields from GNSS receiver networks. Adv Space Res 47:898–912

Rocken C, Ware R, Van Hove T et al (1993) Sensing atmospheric water vapor with the global positioning system. Geophys Res Lett 20:2631–2634

Rohm W, Bosy J (2011) The verification of GNSS tropospheric tomography model in a mountainous area. Adv Space Res 47(10):1721–1730

Stolle C, Schluter S, Heise M et al (2006) A GPS based three-dimensional ionospheric imaging tool: process and assessment. Adv Space Res 38(11):2313–2317. https://doi.org/10.1016/j.asr.2006.05.016

Wang W, Wang J (2011) Tropospheric water vapor tomography based on algebraic reconstruction techniche. J Comput Appl 31(11):3149–3156

Navigation and Location-Based Service

A Self-dependent Camera/Map-Aided Smartphone-Based PNS

Chunyang Yu(✉), Yiran Lou, Haiyu Lan, Minghong Zhu, and Naser El-Sheimy

Department of Geomatics, University of Calgary, 2500 University Drive N.W., Calgary, Canada
chunyang.yu@ucalgary.ca

Abstract. To build a portable and reliable indoor pedestrian navigation system (PNS), signal availability, cost, power consumption, and accuracy are the key issues need to be considered and evaluated. Taken all the above issues into consideration, a totally self-dependent and low-cost camera/map-aided smartphone-based indoor pedestrian navigation system is proposed in this research. Different from the traditional indoor navigation system, in this research, only smartphone-embedded sensors, i.e. Micro-electromechanical System (MEMS) gyroscope, MEMS accelerometers, magnetometer and camera, are needed to establish the positioning system. To effectively use the indoor geospatial information, a wall-cross method is proposed to add indoor map information to the positioning system through a novel cascade structure algorithm. Moreover, a low-cost line segment detection (LSD) - based map acquisition method is utilized to obtain the indoor map information. Nonlinear particle filter is applied to implement the map aiding process to the navigation solutions. The proposed method has been tested and verified through real experiments in different scenarios. This research has important theoretical and practical value for both industry and academic areas.

Keywords: Map-aiding navigation · Pedestrian navigation system · Line segment detection · Inertial navigation system · Smartphones based sensors

1 Introduction

Nowadays, pedestrian location system has a big market requirement, more than 25000 developers in the world are focusing on this market, such as IndoorAtlas™, Streetlight Data™, Focal Point Positioning™, TRX Systems™ etc. Markets and Markets released a market report which indicates that the market of indoor location-based service will grow from $15.04 billion to $77.84 billion by 2021. Precise positioning plays an essential role in helping average person to live a convenient life and the emergency rescuers a powerful aiding tool by shortening their navigation time.

As shown in Table 1, various kinds of techniques, such as Geomagnetism based positioning method, map based method, inertial navigation system (INS) based method, Wi-Fi based positioning, Bluetooth technique and vision based technique

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 245–252, 2019.
https://doi.org/10.1007/978-981-13-7751-8_25

could be used to obtain a pedestrian's position in indoor environment [1]. However, each method has its own drawbacks, there isn't any technique that can satisfy all the requirements and needs of indoor pedestrian navigation system. Therefore, numerous methods have been proposed and integrated for pedestrian navigation by researchers. From Table 1, it is obvious that INS and Vision are two self-dependent positioning techniques, the navigation system based on these two techniques can be operated without installing infrastructures and pre-surveying. However, the estimation error of low-cost MEMS based INS grows quickly with time, and Vision/Camera based positioning method suffers from high computational complexity.

Table 1. Indoor navigation techniques

Technique	Accuracy	Characteristics
INS	Increase with time	Standalone system; Low-cost MEMS sensors; Error grows with time
WLAN	1–10 m	Wi-Fi router is needed (infrastructure-based); Database is needed for fingerprinting method; Access Point's location is needed for trilateration method
Vision/Camera	1–3 m	High computational complexity; Priori database for image processing and matching is needed
Cellular	50–500 m	Time of arrival (TOA): needs synchronization between signal receivers; Time Difference of Arrival (TDOA): needs external information
Map-based	1–5 m	Indoor map is the essential information
Geomagnetism	1–5 m	Vulnerable to environmental interference and magnetic changing

To build a portable and reliable PNS, signal availability, cost, power consumption, and accuracy are the key issues need to be considered and evaluated. For the customers, the major challenges for a pedestrian navigation system is to reduce the cost of the system, including the time-cost and the economic-cost, without decreasing the accuracy of the system. Currently, the MEMS sensors, including MEMS-based inertial sensors, Wi-Fi module, barometer and magnetometer, which is low-cost and convenient have been widely embedded in smart devices (e.g., smart phones, watches, tablets), and the global IMU embedded smartphone adoption rate keeping increasing year by year. Therefore, inertial navigation based method is applied in this research to obtain a primary navigation solutions. However, the navigation errors of standalone low-cost MEMS based inertial navigation system significantly grow with time. Therefore, aiding information/techniques are needed to correct or constrain the sensor-based navigation solutions [9].

Indoor map, which generally do not need additional infrastructures or equipment as well, is also a widely-used and favorable aiding information for indoor pedestrian navigation. Therefore, for this research, as we need to consider about the cost and the indoor operational environment, only free indoor map and the smartphone embedded odometer are integrated with the low-cost INS to achieve a low-cost, continuous, and

accurate indoor position solution. Therefore, the remaining problem of this research is how to obtain and make advantage of the above information to get accurate indoor navigation solution. Map information is essential for map aided INS methods. Currently, only a few public buildings, e.g. airports, museums, universities etc., could offer precise digital indoor maps. Some location-service companies, such as Google Map™ and Openstreet Map™, can also provide digital indoor maps of some public buildings to users. However, when the preceding map sources are unattainable, how can we obtain an indoor map information without increasing the cost of the whole navigation system [2].

Therefore, based on the above background introduction and problem description, the main goal of this research is to design and develop a smartphone based indoor pedestrian navigation system using the existing sensors and information. In this research, a data collecting software is developed to collect the smartphone embedded sensors' output and indoor map information (if needed), an indoor map information acquisition method based on camera is proposed, and a map-aided INS algorithm is designed to obtain the indoor positioning solution.

2 Methodology

The methodology of this research includes indoor map information acquisition using camera, Non-holonomic constraints (NHC) based inertial navigation system and map-aided INS method.

2.1 Indoor Map Information Acquisition

Maps are very informative sources of measurements for indoor navigation as it could be used to plot the indoor navigation results and further correct the estimated position errors. Nowadays, map-based indoor PNS approaches are available and cost-effectiveness, since maps needs no additional infrastructure, e.g. power, appliances, spectrum or processing. Some LBS-based companies, such as Palmap™, NavInfo™, Google Map™ and OpenStreetMap™, can provide several types of indoor maps. Some public buildings could provide indoor floorplan maps, which indicate the location of various offices, rooms on different floors, and show other points of interest such as restrooms, elevators, escalators, shops and so on. However, there are still many old constructions that do not have digital map information, therefore, we cannot always assume that the map information is available for our indoor navigation applications. In this case, the indoor map information is not accessible, a suitable solution is needed to solve the map information acquisition issue.

Currently, various kinds of techniques can be used to obtain digital indoor map information. One of the most common methods is the crowdsourcing approach; however, it is difficult to fuse information from different users, since powerful processors and fusion algorithm are needed for crowdsourcing. Generating indoor map using professional indoor mapping systems is another way to obtain indoor map information. However, the cost of such an indoor mapping system is relatively high, but lower acquisition throughput and processing throughput. The goal of this research is to provide a low-cost, convenient, and accurate indoor position system. The floor

plan maps are used as basic source to generate the digital map information through some image processing techniques. Then, through the edge detection method, such as the Canny algorithm, and the Hough transform method, we can obtain the line segment, as well as coordinates of line segments' pixels through registration. These coordinates, which can be used directly in the MA navigation approach.

2.2 NHC Based Inertial Navigation System

This research is based on the research team's previous publication [1, 10]. Inertial navigation decoding algorithm is used to obtain the pedestrian's positioning solutions. The position, orientation and velocity (POV) of a pedestrian in the navigation frame is given as:

$$\begin{bmatrix} r^n \\ v^n \\ T_b^n \end{bmatrix} = \begin{bmatrix} D^{-1}v^n \\ T_b^n f^b - 2\left(\Omega_{ie}^n + \Omega_{en}^n\right) + g^n \\ T_b^n\left(\Omega_{ib}^b - \Omega_{in}^b\right) \end{bmatrix} \tag{1}$$

in which $D = diag([R_M + h \quad (R_N + h)cos\varphi \quad -1])$, n represents the navigation frame. $r^n = [r^x \quad r^y \quad r^z]^T$ and $v^n = [v^x \quad v^y \quad v^z]^T$ represent the INS-derived position and velocity. T_b^n denotes the transition matrix describing the body frame, b, relative to the navigation frame n. The attitude vector, ε (i.e., pitch, roll and heading) can be derived from T_b^n. f^b is the accelerometer, and ω^b denotes the angular velocity. g^n is the local gravity vector; φ is the local latitude; R_N is the radius of curvature in the prime vertical; R_M is the radius of curvature in the meridian. To make the system state equation easy to calculate, the POV equation will be linearized and discretized as:

$$\delta x_{k+1}^{ins} = \Phi_{k+1}^{ins}\delta x_k^{ins} + G_k^{ins} w_k^{ins} \tag{2}$$

in which Φ_{k+1}^{ins} is the discrete-time error state transition matrix from epoch k to $k + 1$. $x_k^{ins} = [r^n \quad v^n \quad \varepsilon^n]^T$ is the navigation solution at time k. The vector $\delta x^{ins} = [\delta r^n \quad \delta v^n \quad \delta\varepsilon^n]^T$ is the INS error state, which includes three components, position error, velocity error and attitude error.

NHC is a typical motion constraint, as a typical velocity constraint for vehicle navigation that used in the GNSS-denied environments, NHC is based on the assumption that a moving platform (e.g. a land vehicle) cannot move sideways or vertically. Thus, the lateral and vertical velocities of the specific platform along the body frame are zero [4]. Therefore, NHC can be used for pedestrian navigation In this research, the estimated state error could be updated through an EKF using motion constraints update such as ZUPT and NHC, the derivation and linearization process, as well as the detail of the NHC aided INS can refer to the previous publication of the author [2]. The measurement model of the EKF can be represented as,

$$\delta z_k^{INS/NHC} = H_k^{INS/NHC}\delta x_k^{ins} + v_k^{INS/NHC} \tag{3}$$

in which $\delta z_k^{INS/NHC}$ is the measurement vector. Here, it is calculated using the pseudo-velocity of the system minus the INS-derived velocity; $v_k^{INS/NHC}$ is the measurement noise matrix. $H_k^{INS/NHC}$ is the measurement matrix:

$$H_k^{INS/NHC} = \begin{bmatrix} 0_{3\times 3} & T_b^u T_n^b & -T_b^u T_n^b V^n \end{bmatrix} \tag{4}$$

in which V^n is the skew-symmetric matrix of the velocity v_{INS}^n.

2.3 Map-Aided INS Method Through Map Matching

For almost all current indoor navigation systems, indoor map information only acts as a presentation layer on which the position solution or the destination is plotted. Nevertheless, maps should not only serve to display the navigation results, they can be used in the calculation process [6]. Indoor maps can be effectively used as constraints to correct the estimated navigation errors. Therefore, some researchers have started to involve map information in the navigation algorithm [5].

According to the format of the map information, map-based navigation can be classified into 3 types: geometric-based method, probabilistic-based method, and topological-based method. The geometric-based method is easy to implement, as it only uses the geometric relationship between the estimated position and the map information. Besides the geometric map information, the topological-based method needs the topological information, such as the road classification and road restriction information, which may be difficult for some application scenarios. The probabilistic-based method may require the probabilistic map of the user, this map can be either obtained on the offline persevering process or the online learned process [7]. Compared with the last two map aided method, geometric-based method is more popular and efficient as only indoor geometric is need, and much easier to obtained.

In this research, the map information is used to correct the accumulated error of inertial navigation sensors. A wall crossing method is used to constrain the estimated solution through the sequence importance sampling (SIS) approach and pedestrian dead reckoning (PDR) method. The position and heading information of the user can be expressed as,

$$\begin{aligned} E_{k+1}^i &= E_k^i + S_{k,k+1}^i \sin\left(\Psi_k^i\right) \\ N_{k+1}^i &= N_k^i + S_{k,k+1}^i \cos\left(\Psi_k^i\right) \\ \Psi_{k+1}^i &= \Psi_k^i + \delta\Psi_{k,k+1} \end{aligned} \tag{5}$$

in which Ψ is the heading derived from the lower KF. $S_{k,k+1}^i$ is the pseudo-step-length from time step k to time step $k + 1$. $\delta\Psi_{k,k+1}$ represents the heading change between the two headings derived from the lower filter at time step k and time step $k + 1$ respectively. The superscript i denotes the i-th particle of the PF. E_{k+1}^i and N_{k+1}^i are the East and North position of the system at time step $k + 1$ respectively. When we have N particles at time step k, each particle represents a position coordinate $P(E_{k+1}^i, N_{k+1}^i)$, and the weight of the i-th particle is expressed as w_k^i.

If the updated position coordinate can match the indoor digital map information, which means the line segment between the previous position and this coordinate has no intersection with the indoor map information, then this coordinate is an effective estimated position, if not, the new estimated coordinate is regarded ineffective coordinate, and will be deleted by reset its weight to zero.

3 Experiments and Solutions

To test the performance of the proposed map aided INS positioning method, indoor experiments are conducted in the EEEL building at the University of Calgary (U of C). As shown in Table 2, Xiaomi Mi 4 smartphone is used in this research to collect the inertial sensors outputs and indoor digital map information.

Table 2. Sensor performance of Xiaomi Mi 4 smartphone

Sensor	Model	Scale/Rate Range/pixels	Noise Density/Rate
Accelerometer	InvenSense MPU6050	±2/±4/±6/±8/±16 g	400 μg/√Hz
Gyroscope	InvenSense MPU6050	±250, ±500, ±1000, ±2000°/sec (dps)	0.005 mdps/√Hz
Compass	AKM AK8963	±4900 μT	
Camera	Sony IMX214	4224 (H) × 3200 (V) Approx. 13.51 M pixels	

The indoor two-dimensional digital map is download from University of Calgary maps, and can also be obtained from floor plan map on each floor of the building using the smartphone camera. Figure 1 is the testing environment, and Fig. 2 gives the two designed trajectories for test [8].

Fig. 1. Testing environment in the EEEL building at U of C

Fig. 2. Designed test trajectory

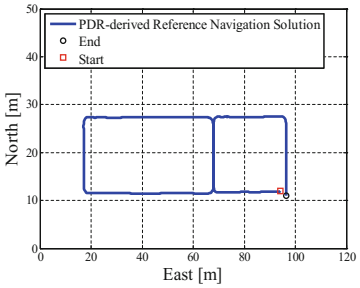


Fig. 3. Reference trajectory

The pure INS method cannot provide reliable position solution to the pedestrian, therefor aiding information or constraint is need (Fig. 3).

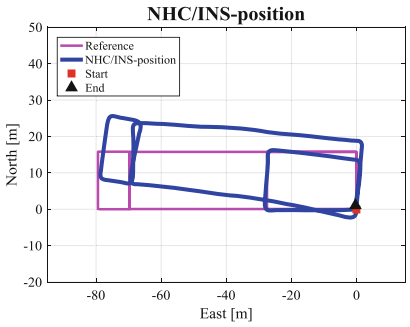


Fig. 4. Estimated position of NHC/INS

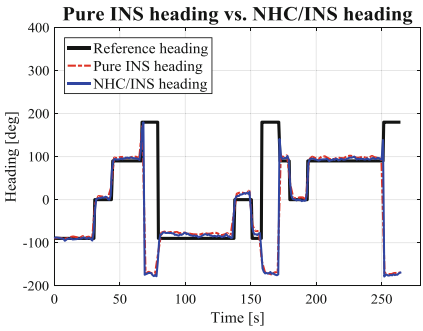


Fig. 5. Estimated position of NHC/INS

Figure 4 shows the estimated position using the NHC/INS method. Figure 5 is the estimated heading using NHC/INS method. In Fig. 4, the pink line is the reference trajectory; the blue solid line is the NHC/INS-derived position. With the NHC, the error drift can be well constrained. However, it still cannot satisfy the requirement of indoor

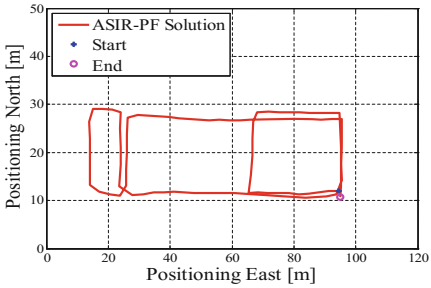


Fig. 6. Estimated position from the map aided INS method

pedestrian navigation. Figure 6 is the estimated position from map aided INS method, it is obvious that the estimated trajectory can well match the reference trajectory, which means the proposed algorithm can provide a reliable indoor navigation solution.

This research proposes a map-aided INS algorithm to provide a low-cost and reliable indoor navigation solution. This algorithm can increase the system efficiency and decrease the system computational burden. Map matching method is applied to the system through wall-crossing method. Moreover, through real experiment we can draw the conclusion that the Map/INS integrated method can dramatically improve INS-derived solution.

References

1. Yu C, Lan H, Gu F, Yu F, El-Sheimy N (2017) A Map/INS/Wi-Fi integrated system for indoor location-based service applications. Sensors 17(6):1272
2. Yu C, El-Sheimy N, Lan H, Liu Z (2017) Map-based indoor pedestrian navigation using an auxiliary particle filter. Micromachines 8(7):225
3. Von Gioi RG, Jakubowicz J, Morel JM, Randall G (2012) LSD: a line segment detector. Image Process Line 2:35–55
4. Leppäkoski H, Collin J, Takala J (2013) Pedestrian navigation based on inertial sensors, indoor map, and WLAN signals. J Signal Process Syst 71(3):287–296
5. Karimi HA (ed) (2015) Indoor wayfinding and navigation. CRC Press, Boca Raton
6. Nilsson JO, Rantakokko J, Handel P, Skog I, Ohlsson M, Hari KVS (2014) Accurate indoor positioning of firefighters using dual foot-mounted inertial sensors and inter-agent ranging. In: 2014 IEEE/ION Position, Location and Navigation Symposium-PLANS 2014, pp 631–636
7. Agrawal R, Vasalya A (2012) Bluetooth navigation system using Wi-Fi access points. Int J Distrib Parallel Syst 3(2)
8. Lan H, Yu C, El-Sheimy N (2015) An integrated PDR/GNSS pedestrian navigation system. In: China satellite navigation conference (CSNC) 2015 proceedings, vol III. Springer, Heidelberg, pp 677–690
9. Gu F, Khoshelham K, Yu C, Shang J (2018) Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders. IEEE Trans Instrum Meas
10. Lan, H (2016) Multiple systems integration for pedestrian indoor navigation. University of Calgary

Linearization Error Analysis of Observation Equations in Pseudo Satellite Positioning System

Yue Zhao, Shuguo Pan(✉), and Yanheng Wang

School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China
psg@seu.edu.cn

Abstract. For the problem that the observation environment is poor and the positioning accuracy is not high, pseudolites can be used for positioning. However, compared with navigation satellites, pseudolites are much closer to user receivers. Using algorithms such as EKF may cause relatively large linearization errors, but UKF algorithm can avoid linearization errors. Through theoretical analysis, it is found that the distance between the satellite and the receiver of the user and the coordinate error of the user receiver will affect the linearization error. In this paper, We used different receiver coordinate errors, and a set of short-baseline pseudo-satellite positioning experimental data was solved by UKF algorithm and EKF algorithm to verify the influencing factors of linearization error. The experimental results show that the linearization error is larger when the receiver coordinate error is larger or the satellite distance from the receiver is smaller. When the receiver coordinate error is relatively large, the convergence speed and overall accuracy of the UKF algorithm are better than EKF algorithm.

Keywords: Pseudo satellite · Linearization error · UKF algorithm

1 Introduction

The reliability and accuracy of satellite navigation systems depend on the geometric distribution of the number of visible satellites [1]. Once there is a poor observation environment such as indoors, canyons and underground, the number of visible satellites will be reduced, which will greatly affect the positioning accuracy [2, 3]. For the problem of low positioning accuracy caused by poor observation environment, pseudolites can be used for positioning. Pseudo-satellite is also called land-based design or land-based satellite. Its function and principle are similar to navigation satellites. It has the characteristics of low cost and flexible setting [4]. Different from navigation satellites, pseudolites are arranged on the ground and are flexible in setting [4, 5], so that the problem of small number of visible satellites and poor geometric distribution can be solved.

For the pseudo-satellite system with independent networking, the pseudo-range positioning equation is nonlinear when pseudo-satellite positioning is solved. The traditional method is to use the extended Kalman filter (EKF) algorithm to perform

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 253–265, 2019.
https://doi.org/10.1007/978-981-13-7751-8_26

first-order Taylor expansion on nonlinear functions, and then skip the second or more items [6, 7]. However, because the distance from the pseudolite to the user receiver is much closer than the distance from the navigation satellite to the user receiver, the linearization error of the pseudolite pseudorange observation equation caused by the EKF algorithm will be relatively large [8], which will affect the positioning accuracy. And when the distance of the satellite from the user receiver is small, the error of the user receiver coordinates also has a greater influence on the linearization error. For the strongly nonlinear equation of the pseudo-satellite pseudo-positioning equation, Julier et al. proposed the unscented Kalman filter (UKF) algorithm [9]. This method directly uses the nonlinear model, using the state equation of the nonlinear system or the observation equation [10], avoiding the influence of nonlinear error on the positioning result. In theory, the effect is better than EKF when the satellite is closer to the receiver.

This paper first analyzes the influencing factors of linearization error of pseudo-satellite pseudo-location equation. Then introduced the UKF algorithm. Finally, in order to verify the influence of the satellite distance from the user receiver and the error of the user receiver coordinates on the linearization error, and to prove that the UKF algorithm is better than the EKF algorithm when the satellite is closer to the receiver, we solved a set of indoor short-base pseudo-satellite positioning experimental data. The observation model is a double-difference pseudorange equation, and the EKF and UKF algorithms are used to calculate the positioning results. The experimental results show that the larger the receiver coordinate error, the larger the linearization error; the smaller the satellite distance from the receiver, the larger the linearization error. And when the receiver coordinate error is relatively large, the convergence speed and overall accuracy of the UKF algorithm are better than the EKF algorithm.

2 Analysis of Linearization Error of Pseudo-satellite Pseudorange Equation

In this section, we will introduce the pseudo-satellite double-difference pseudorange observation model, and then analyze the linearization error of the pseudo-satellite positioning system to theoretically explain the factors affecting the linearization error.

2.1 Pseudo-satellite Double-Difference Pseudorange Observation Model

Because the experiment in this paper is static positioning, the user receiver is fixed, so it is only necessary to analyze the nonlinear error of the observation equation. The pseudo-satellite pseudo-range observation equation is:

$$\rho^n = r^n + \delta t_u - \delta t^n + I^n + T^n + \varepsilon_\rho^n \tag{1}$$

In the formula, ρ^n is the pseudorange between the receiver and the satellite with the PRN number n, r is the actual distance from the receiver to the satellite, δt_u is the receiver clock, δt^n is the satellite clock difference, I is the ionospheric delay, T is the tropospheric delay, and ε_ρ^n is the measured noise parameter.

In order to eliminate receiver clock error, satellite clock error, ionospheric delay and tropospheric delay in pseudo-satellite pseudo-range measurements, a double-difference pseudo-range observation equation is used:

$$\rho_{ur}^{ij} = r_{ur}^{ij} - \varepsilon_{\rho,ur}^{ij} \tag{2}$$

The double difference pseudorange measurement is defined as:

$$\rho_{ur}^{ij} = \left(\rho_u^i - \rho_r^i\right) - \left(\rho_u^j - \rho_r^j\right) \tag{3}$$

The double difference geometric distance is defined as:

$$r_{ur}^{ij} = \left(r_u^i - r_r^i\right) - \left(r_u^j - r_r^j\right) \tag{4}$$

In the formula, i and j refer to different pseudolites, and u and r refer to the user receiver and the reference receiver, respectively.

2.2 Linearization Error Analysis

Use $h(x_u, y_u, z_u)$ to represent the pseudo-satellite pseudorange equation, then expand its Taylor to:

$$h(\hat{x}_u + \Delta x_u, \hat{y}_u + \Delta y_u, \hat{z}_u + \Delta z_u) = h(\hat{x}_u, \hat{y}_u, \hat{z}_u) + \frac{x - \hat{x}_u}{\hat{r}}\Delta x_u + \frac{y - \hat{y}_u}{\hat{r}}\Delta y_u + \frac{z - \hat{z}_u}{\hat{r}}\Delta z_u + \ldots \tag{5}$$

In the formula, ellipsis refers to Taylor's high-order expansion, and satellite to receiver distance $\hat{r} = \sqrt{(x - \hat{x}_u)^2 + (y - \hat{y}_u)^2 + (z - \hat{z}_u)^2}$.

Use X to represent the state vector. X0 as an approximate estimate of the state vector. $\delta X = X - X_0$. A represents the first-order partial derivative of f(X) at X0. εh represents the second-order residual. Then there is:

$$h(X) = h(X_0) + A\delta X + \varepsilon_h \tag{6}$$

$$\varepsilon_h = \frac{1}{2}\delta X^T \partial_{XX}^2 h(X_0 + t \times \delta X)\delta X \quad 0 < t < 1 \tag{7}$$

In the formula, $\partial_{XX}^2 h(X)$ is called Hessian Matrix and consists of second-order partial derivatives. Suppose u refers to the receiver, i and j refer to the satellite. Then the geometric distance from the satellite to the receiver can be expressed as:

$$r_u^i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2 + (z_i - z_u)^2} \tag{8}$$

The Hessian Matrix is:

$$\partial_{XX}^2 h(X) = \frac{1}{\sqrt{\left(r_u^i\right)^3}} \begin{bmatrix} (y_i - y_u)^2 + (z_i - z_u)^2 & -(x_i - x_u)(y_i - y_u) & -(x_i - x_u)(z_i - z_u) \\ -(x_i - x_u)(y_i - y_u) & (x_i - x_u)^2 + (z_i - z_u)^2 & -(y_i - y_u)(z_i - z_u) \\ -(x_i - x_u)(z_i - z_u) & -(y_i - y_u)(z_i - z_u) & (x_i - x_u)^2 + (y_i - y_u)^2 \end{bmatrix} \tag{9}$$

Usually in the EKF method, the second-order residual ε_h is directly rounded off, which is the linearized residual error term. Teunissen gave a boundary estimate for its size in the literature [11]:

$$\frac{1}{2}\lambda_{min} \times \|\delta X\|^2 \leq \varepsilon_h \leq \frac{1}{2}\lambda_{max} \times \|\delta X\|^2 \tag{10}$$

In the formula, λ_{max} is the maximum eigenvalue of Hessian Matrix $\partial_{XX}^2 h(X)$, λ_{min} is its minimum eigenvalue, and δX is the mean square error of $x_i - x_u$, $y_i - y_u$, $z_i - z_u$. Here we make the limit hypothesis that its maximum possible eigenvalue is $\frac{1}{r_u^i}$ and its minimum possible eigenvalue is 0 [12]. Therefore, the following estimation can be further obtained from the formula (10):

$$0 \leq \varepsilon_h \leq \frac{\|\delta X\|^2}{2r_u^i} = \frac{\delta_{(x_i - x_u)}^2 + \delta_{(y_i - y_u)}^2 + \delta_{(z_i - z_u)}^2}{2r_u^i} \tag{11}$$

For the pseudo-satellite double-difference pseudorange experiment conducted in this paper, the pseudo-satellite coordinates are known, so in the above formula:

$$\begin{aligned} \delta_{(x_i - x_u)}^2 &= \delta_{(x_u)}^2 \\ \delta_{(y_i - y_u)}^2 &= \delta_{(y_u)}^2 \\ \delta_{(z_i - z_u)}^2 &= \delta_{(z_u)}^2 \end{aligned} \tag{12}$$

Taylor expansion of the double difference pseudorange Eq. (2), Then, from Eq. (11), the second-order residual ε_{dh} of the double-difference pseudorange equation is estimated as follows:

$$0 \leq \varepsilon_{dh} \leq \frac{\delta_{(x_u)}^2 + \delta_{(y_u)}^2 + \delta_{(z_u)}^2}{2r_u^i} - \frac{\delta_{(x_u)}^2 + \delta_{(y_u)}^2 + \delta_{(z_u)}^2}{2r_u^j} \tag{13}$$

In the formula, $\delta_{(x_u)}^2, \delta_{(y_u)}^2, \delta_{(z_u)}^2$ are the variances of the user receiver coordinates (x_u, y_u, z_u), respectively, r_u^i is pseudo The geometric distance from satellite i to the user receiver, r_u^j is the geometric distance from pseudolite j to the user receiver.

Global navigation satellites are more than 20,000 km away from users. It can be seen from Eq. (11) that if the user's receiver coordinate error is 200 m, the maximum error of the linearization of the observation equation is only 1 mm [13]. Based on this premise and formula (13), the maximum error of linearization of the double-difference pseudo-range observation equation is definitely less than 1 mm. Therefore, when calculating coordinates with the EKF algorithm, the linearization error can be ignored. However, for the pseudo-satellite navigation system, if the receiver coordinate error is 20 m and the two pseudo-satellite distances are 200 m and 150 m respectively, the maximum error of the linearization of the observation equation is 0.33 m. This is already a non-negligible error.

It can be seen from the above analysis of linearization error that since the pseudolite is closer to the user than the global navigation satellite, if the distance between each pseudolite and the receiver is not guaranteed, the linearization error will be relatively large. This error cannot be ignored and is related to the distance of the pseudolite from the receiver and the receiver position error. Therefore, for the pseudo-satellite double-difference pseudorange model, it is not appropriate to use the traditional EKF method for settlement, and the UKF method that does not generate linearization error should be used for settlement.

3 UKF Algorithm

The UKF algorithm is a method that approximates the nonlinear distribution by using the sampling strategy. It is based on the Unscented Transform and uses the Kalman linear filtering framework [14].

Suppose the state equation and the observation equation of the system are:

$$x_k = f(x_{k-1}) + \omega_k \tag{14}$$

$$z_k = h(x_k) + v_k \tag{15}$$

In the formula, x_k and x_{k-1} are state vectors at times k and k-1, respectively, z_k is the observation vector, ω_k is the state noise, v_k is the observed noise, and ω_k and v_k are mutually uncorrelated zero-mean white noise sequences. The variance matrix is QK and Rk respectively. The specific calculation steps of the UKF algorithm are as follows:

Selected initial value of the filter:

$$\hat{X}_0 = EX_0 \tag{16}$$

$$P_0 = E\left[\left(X_0 - \hat{X}_0\right)\left(X_0 - \hat{X}_0\right)^T\right] \tag{17}$$

Calculate the sigma point:

$$\tilde{x}_{k-1}^{(0)} = \hat{X}_{k-1} \tag{18}$$

$$\tilde{x}_{k-1}^{(i)} = \hat{X}_{k-1} + \gamma\left(\sqrt{P_{k-1}}\right)_{(i)} \quad i = 1, 2, \ldots, \mathrm{n} \tag{19}$$

$$\tilde{x}_{k-1}^{(i)} = \hat{X}_{k-1} - \gamma\left(\sqrt{P_{k-1}}\right)_{(i-n)} \quad i = n+1, n+2, \ldots, 2n \tag{20}$$

In the above formula,

$$\gamma = \sqrt{n+\lambda} \tag{21}$$

$$\lambda = a^2(n+k) - n \tag{22}$$

n is the state vector dimension. a is a small positive number used to determine the distribution of sigma points around x, ranging from 10-4 to 1. k is the second scale parameter, k = 3-n [15, 16].

Determine the weight:

$$W_0^{(m)} = \frac{\lambda}{n+\lambda} \tag{23}$$

$$W_0^{(c)} = \frac{\lambda}{n+\lambda} + 1 - a^2 + \beta \tag{24}$$

$$W_i^{(m)} = W_i^{(c)} = \frac{1}{2(n+\lambda)} \qquad i = 1, 2, \ldots, 2n \tag{25}$$

In the formula, β is the state distribution parameter. The β optimal value for the Gaussian distribution is 2, and for the univariate state variable, the β optimal value is 0.

Time update:

$$x_{k/k-1}^{(i)} = f\left(x_{k-1}^{(i)}\right) \quad i = 0, 1, 2, \ldots, 2n \tag{23}$$

$$\hat{\mathrm{X}}_{k/k-1} = \sum_{i=0}^{2n} W_i^{(m)} x_{k/k-1}^{(i)} \tag{24}$$

$$P_{k/k-1} = \sum_{i=0}^{2n} W_i^{(c)} \left[x_{k/k-1}^{(i)} - \hat{\mathrm{X}}_{k/k-1}\right]\left[x_{k/k-1}^{(i)} - \hat{\mathrm{X}}_{k/k-1}\right]^T + Q_{k-1} \tag{25}$$

In the formula, $\hat{\mathrm{X}}_{k/k-1}$ is the prediction state vector, and $P_{k/k-1}$ is the covariance matrix of the prediction state vector.

Measurement update:

$$Z_{k/k-1}^{(i)} = h\left(x_{k/k-1}^{(i)}\right) \tag{26}$$

$$\hat{Z}_{k/k-1} = \sum_{i=0}^{2n} W_i^{(m)} Z_{k/k-1}^{(i)} \tag{27}$$

$$P_{(ZZ)_{k/k-1}} = \sum_{i=0}^{2n} W_i^{(c)} [Z_{k/k-1}^{(i)} - \hat{Z}_{k/k-1}] \left[Z_{k/k-1}^{(i)} - \hat{Z}_{k/k-1}\right]^T + R_k \tag{28}$$

In the formula, $\hat{Z}_{k/k-1}$ is the predicted observation vector, and $P_{(ZZ)_{k/k-1}}$ is the covariance matrix of the predicted observation vector.

Filter update:

$$P_{(XZ)_{k/k-1}} = \sum_{i=0}^{2n} W_i^{(c)} [x_{k/k-1}^{(i)} - \hat{X}_{k/k-1}] \left[Z_{k/k-1}^{(i)} - \hat{Z}_{k/k-1}\right]^T \tag{29}$$

$$K_k = P_{(XZ)_{k/k-1}} P_{(ZZ)_{k/k-1}}^{-1} \tag{30}$$

$$\hat{X}_k = \hat{X}_{k/k-1} + K_k \left(Z_k - \hat{Z}_{k/k-1}\right) \tag{31}$$

$$P_k = P_{k/k-1} - K_k P_{(ZZ)_{k/k-1}} K_k^T \tag{32}$$

In the formula, $\hat{X}_k$ is the estimated state vector, P_k is the covariance matrix of the estimated state vector, and K_k is the gain matrix.

It can be seen from the above steps that the UKF method directly uses the state equation or the observation equation of the nonlinear system to calculate, and avoids the linearization error compared with the EKF algorithm. In addition, the UKF algorithm requires a priori information about the exact known system noise. If the system noise is inaccurate, the filtering accuracy will decrease or even diverge [17, 18]. For the static test in this paper, the receiver is known to be stationary, so the system noise can be set to zero to ensure the accuracy of the UKF algorithm.

4 Experiment and Analysis

The experimental site of this paper is a test site in Shijiazhuang. The experimental device is self-made 8 pseudolites and 2 U-BLOX receivers. The experimental site is indoors, 8 pseudolites, and the distance between the pseudolite to the user receiver and the reference receiver is shown in Table 1.

Table 1. Pseudo satellite to receiver distance

Pseudo satellite PRN number	User receiver	Reference receiver
1	3.979886808	3.138004621
2	3.941818109	3.092444017
3	3.218742146	3.582744618
4	3.198264999	3.561635158
5	3.381152614	2.739856384
6	3.383770973	2.735710877
7	2.985285245	3.629406839
8	3.024775032	3.694472357

EKF and UKF are used respectively for positioning solution. Both methods require a state vector of the initial epoch and a covariance matrix of the state vector, namely X0 and P0. Because this experiment is used to verify the above inference, the given X0 is the deviation of the real coordinates of the user receiver plus a deviation vector $\Delta X = (\Delta x, \Delta y, \Delta z)^T$. ΔX is the coordinate error of the user receiver. Then make the X0 covariance matrix $P_0 = \begin{bmatrix} \Delta x^2 & 0 & 0 \\ 0 & \Delta y^2 & 0 \\ 0 & 0 & \Delta z^2 \end{bmatrix}$. By changing the magnitude of the deviation vector ΔX, the influence of the receiver position error on the linearization error can be verified. According to the requirements of the UKF algorithm and the verification static test of this paper, we can make the state noise matrix $Q_k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Make pseudolite 1 as a reference star. Suppose the M vector represents the distance from different pseudolites to the user receiver, and K is the coefficient matrix. When there are 4 pseudolites, the pseudorange double difference positioning equations can be expressed as:

$$Z = KM \tag{33}$$

In the formula, $K = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$. When making M's covariance matrix $D_{MM} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, Z's covariance matrix $D_{ZZ} = KD_{MM}K^T =$ $\begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

Therefore, the measurement noise matrix R of the pseudorange double difference positioning equation is equal to $D_{ZZ} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. For the UKF algorithm, the set weight parameters are: a = 0.1, k = 0, $\beta = 2$.

The experiment gives X0 different deviation vector ΔX. We made ΔX equal to $(0.5\,\text{m},\ 0.5\,\text{m},\ 0.5\,\text{m})^T$, $(1\,\text{m},\ 1\,\text{m},\ 1\,\text{m})^T$, $(5\,\text{m},\ 5\,\text{m},\ 5\,\text{m})^T$ and $(10\,\text{m},\ 10\,\text{m},\ 10\,\text{m})^T$. Figures 1, 2, 3 and 4 are the error diagrams of the NEU in the three directions of the EFK algorithm and the UKF algorithm under different error vector conditions.

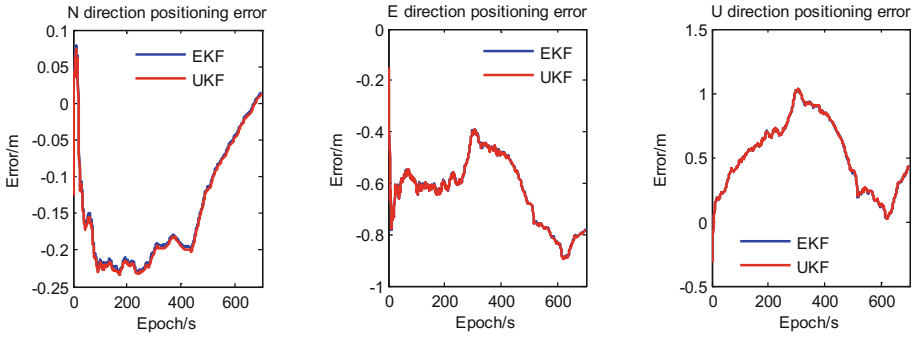


Fig. 1. Positioning result when the user receiver coordinates error 0.5 m

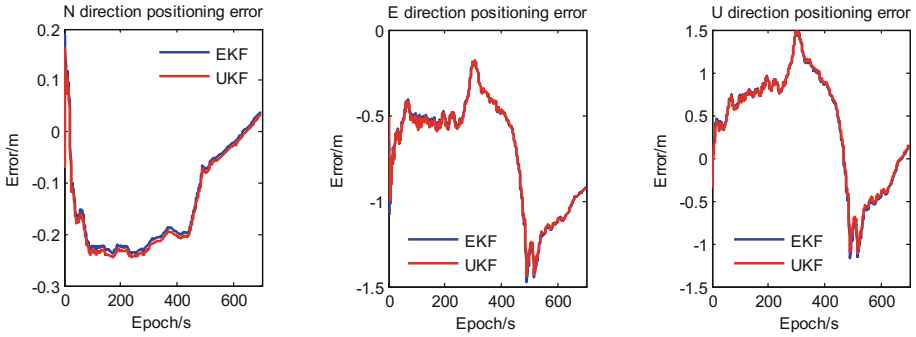


Fig. 2. Positioning result when the user receiver coordinates error 1 m

From Figs. 1 and 2, it can be found that when the receiver error is relatively small, the UKF algorithm and the EKF algorithm have relatively close positioning errors, and the curves almost coincide. Through analysis, the reason may be that the coordinate deviation is small and the linearization error is not obvious, so the UKF algorithm result is close to the EKF algorithm result. From Figs. 3 and 4, it can be found that when the receiver error is relatively large, the convergence result of the UKF algorithm is much faster than the EKF algorithm. This shows that the larger the receiver

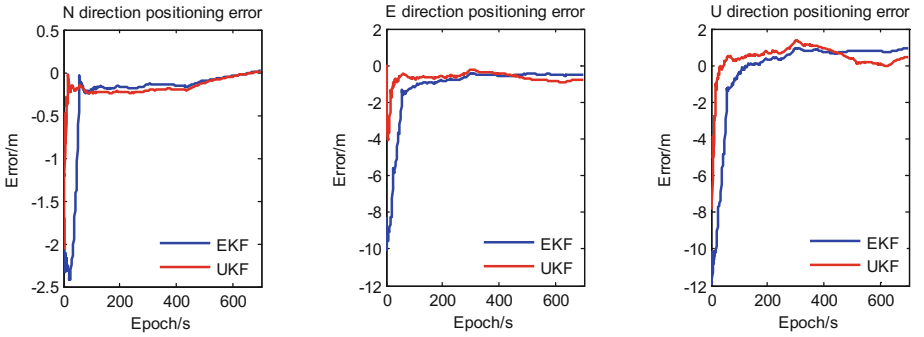


Fig. 3. Positioning result when the user receiver coordinates error 5 m

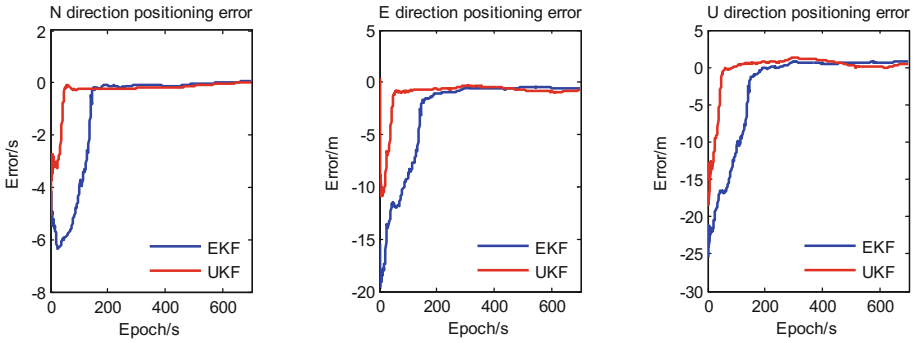


Fig. 4. Positioning result when the user receiver coordinates error 10 m

coordinate error, the larger the linearization error. Moreover, the UKF algorithm is significantly better than the EKF algorithm in the initial stage. Compared with the positioning of the navigation satellite system, it can be explained that the smaller the distance from the satellite to the receiver, the larger the linearization error.

Table 2. Comparison of mean and standard deviation when user receiver coordinate error 0.5 m

Algorithm	Coordinate error/m					
	N		E		U	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
EKF	−0.14489	0.08071	−0.62650	0.13212	0.51969	0.28322
UKF	−0.14837	0.08088	−0.62811	0.13213	0.52100	0.28335

Table 3. Comparison of mean and standard deviation when user receiver coordinate error 1 m

Algorithm	Coordinate error/m					
	N		E		U	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
EKF	−0.13793	0.09699	−0.71354	0.32649	0.38775	0.66190
UKF	−0.14570	0.09732	−0.71387	0.31955	0.39749	0.64963

Table 4. Comparison of mean and standard deviation when user receiver coordinate error 5 m

Algorithm	Coordinate error/m					
	N		E		U	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
EKF	−0.25870	0.09699	−1.08711	0.32649	−0.11446	0.66190
UKF	−0.17180	0.09732	−0.68131	0.31955	0.42670	0.64963

Table 5. Comparison of mean and standard deviation when user receiver coordinate error 10 m

Algorithm	Coordinate error/m					
	N		E		U	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
EKF	−1.02806	1.99124	−2.80897	4.49067	−2.60286	6.40315
UKF	−0.32012	0.66745	−1.10718	1.81457	−0.18189	2.89446

Tables 2, 3, 4, and 5 are the comparisons between the mean and standard deviation of the positioning results for different receiver coordinate errors. It can also be seen from Tables 2 and 3 that the UKF and EKF positioning results are closer when the coordinate deviation is small. As can be seen from Table 4, the standard deviation of the UKF algorithm and the EKF algorithm is closer. But the positioning errors in the N, E, and U directions of the UKF algorithm are slightly smaller than the EKF algorithm. This shows that when the coordinate deviation is 5 meters, the fluctuation of the UKF algorithm and the EKF algorithm is relatively close, but the positioning accuracy of the UKF is relatively high. It can be seen from Table 5 that the mean and standard deviation of the positioning error in the three directions of N, E and U of the UKF algorithm are much better than the EKF algorithm. This shows that the UKF algorithm is much better than the EKF algorithm in terms of accuracy and fluctuation when the coordinate deviation is 10 meters. And from the four tables, it can be seen that the larger the receiver error is, the larger the positioning error of the EKF algorithm is compared with the UKF algorithm. Since the positioning error of the UKF algorithm does not include linearization errors, it can be considered that the linearization of the EKF algorithm increases as the receiver error increases.

Based on the above conclusions, it can be concluded that for the pseudo-satellite positioning system, if the receiver error is larger or the satellite distance from the receiver is smaller, the linearization error is larger, and the UKF algorithm positioning effect is better compared with the EKF algorithm. When the receiver error is large, the convergence speed and overall accuracy of the UKF algorithm are significantly improved compared with the EKF algorithm.

5 Conclusion

This paper firstly analyzes the linearization error of pseudo-satellite double-difference pseudorange observation equation theoretically, and concludes that the linearization error is affected by the distance between the satellite receiver and the user receiver coordinates. Then the specific steps of the UKF algorithm are introduced. Compared with the EKF algorithm, the UKF algorithm can avoid linearization errors, but it needs to ensure the accuracy of the system noise. Finally, we use different receiver coordinate errors, use EKF algorithm and UKF algorithm to calculate a set of experimental data, and find the positioning results. The experimental results show that the linearization error is larger when the receiver coordinate error is larger. When the distance of the satellite from the receiver is smaller, the linearization error is larger. And when the receiver coordinate error is relatively large, the convergence speed and overall accuracy of the UKF algorithm are better than the EKF algorithm. Therefore, the UKF algorithm can avoid the problem of large linearization error, and thus significantly improve the positioning effect of the pseudo-satellite system of the independent network.

Acknowledgments. The authors appreciate anonymous reviewers for their valuable comments and improvements to this manuscript. This work is partially supported by the National Key Technologies R&D Program (Grant No. 2016YFB0502101).

References

1. Zhu Y (2006) Design and implementation of GPS pseudo satellite transmitter. Hohai University
2. Kang G (2006) Research on pseudo-satellite positioning system. Northwestern Polytechnical University
3. Yue X, Yuan J, Wu Q (2005) Overview of the development and application of pseudo satellite technology. Glob Position Syst 30(2):47–51
4. Wang W, Liu Z (2009) Nonlinear compensation of pseudo satellite/inertial integrated navigation. J Beijing Univ Aeronaut Astronaut 35(04):514–518
5. Zhao N, Zhao W, Sun W et al (2014) Pseudo-satellite technology and its application in navigation and positioning. J Navig Position 2(1):82–86
6. Chen H-Q, Wu M (2009) Application of EKF and IUKF in initial alignment of inertial navigation. J Yangtze River Eng Vocat Tech Coll 26(2):40–43
7. Wu L (2011) Research on GPS/DR-based mobile robot combined positioning technology. Nanjing University of Science and Technology

8. Zhao F, Gao S, Wu J, Yang Y (2014) Application of least squares unscented Kalman filter in pseudo satellite positioning. Electron Des Eng 22(19):1–3
9. Julier SJ, Uhlmann JK (1997) A New extension of the Kalman filter to nonlinear systems. In: Proceedings of SPIE - the international society for optical engineering, vol 3068, pp 182–193
10. Zhang W, Sun R (2015) Comparison and application of EKF and UKF performance. J Nanjing Univ Sci Technol 39(05):614–618
11. Teunissen PJG (1990) Nonlinear inversion of geodetic and geophysical data: diagnosing nonlinearity. In: Developments in four-dimensional geodesy. Springer Heidelberg
12. Wang H (2009) Pseudo-satellite enhanced GPS technology and application research. Shanghai Jiaotong University
13. Gao W, Pan N, Zhang X (2008) Application of improved iterative EKF algorithm in pseudo satellite positioning. Sci Mapp Sci (04):73–75
14. Liu W (2009) Research on UKF algorithm and its improved algorithm. Central South University
15. Zhang X, Wang Z, He W et al (2011) Active segment trajectory tracking algorithm based on UKF. Fire Command Control 36(3):95–97
16. Qin Y, Zhang H, Wang S (2012) Kalman filtering and integrated navigation principle. Northwestern Polytechnical University Press
17. Tao Z, Xu X-S, Li P-J (2009) Underwater terrain matching algorithm based on chaotic optimization. J Chin Inert Technol 17(2):156–164
18. Luo ST, Ren ZX (2010) Geomagnetic matching algorithms based on affine model. J Chin Inert Technol

The Status and Challenges of High Precision Map for Automated Driving

Jun Liu[1(✉)], Jihua Xiao[2], HongJie Cao[2], and Jiakai Deng[2]

[1] Beijing BeiDou Navigation and LBS Limited Cooperation, No. 23 Rd. ZhiChun, Haidian District, Beijing 100191, China
jun.liu@chinalbs.org
[2] Beijing Engineering Laboratory of BeiDou Navigation and LBS Technology, No. 23 Rd. ZhiChun, Haidian District, Beijing 100191, China

Abstract. This paper discusses the connotation of high precision map of automated driving in depth, analyzes its important role and compares it with conventional map in terms of expression content, data type and the freshness, then summarizes its data forms. The methods of high precision map building using vehicle-borne mobile mapping system, low-altitude photogrammetry are investigated, and the updating method based on crowdsourcing data is introduced. At last, the problems of the lack of industry standard, the limitation of map accuracy and the prohibition of online service of some map content by law is discussed.

Keywords: Automated driving · High precision map · Navigation digital map · Mobile mapping system · Simultaneous location and mapping · Crowd sourcing

1 Introduction

Map has a long history as a carrier of geospatial information. Since the theory "Pei's six basic principles for cartography" put forward in the Western Jin Dynasty, through research and exploration of domestic and foreign scholars for thousands of years, it has formed a complete cartographic theories and methods. Especially in recent decades, technologies such as satellite geodesy, vehicle-borne mobile mapping system, aerospace photogrammetry and computer-aided mapping have developed rapidly, and the cycle of map production has been significantly shortened. Google, Microsoft, Baidu, AutoNavi and other IT giants rely on the Internet to publish online maps and provide location based services, which is greatly facilitates people's lives. The map also flies from the exclusive use of a few scholars to the homes of ordinary people.

The main function of the current Internet digital map is to facilitate people to travel and provide comfortable location based services. However, with the development of science and technology and the continuous improvement of human requirements for driving safety, driving experience, energy saving, automated and unmanned driving will be the inevitable trend of future development [1]. In the 1980s, with the support of the US Department of Defense Advanced Research Projects Agency (DAPPA), a wave of research on smart car technology was launched. In 1995, Jochem and Pomerleau of

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 266–276, 2019.
https://doi.org/10.1007/978-981-13-7751-8_27

Carnegie Mellon University used computers to drive their NavLab5 minivans, travelled nearly 3,000 miles from the east coast of the United States to the west coast and achieved the task of "driving across the United States without a hand" [2].

At present, the international automated driving research boom is mainly initiated by Google, and its evolution is dominated by the market. By the end of July 2016, Google has experimented with 2.89 million km. GM, Ford, Mercedes-Benz, BMW and other famous car companies, as well as Google, Tesla and other technology companies are developing their own automated driving technology. As the core of automated driving technology, high precision maps have also received extensive attention from domestic and foreign scientific research institutions, but it has not reach a wide consensus at present in various aspect including content, morphology, precision, etc. This paper will discuss and analyze the connotation, production and application of high precision maps for automated driving.

2 Simultaneous Location and Mapping of Mobile Robots

Autonomous vehicle is essentially an autonomous mobile robot. The latter is an important research object in the field of robotics and artificial intelligence. It is necessary to realize its own positioning and autonomous navigation through self-sensing of environmental characteristics in an unknown environment.

"Where I am? Where am I going? How should I get there?" are the three basic issues facing mobile robots' autonomous navigation [3]. They were solved by Simultaneous Location And Mapping (SLAM) technology. SLAM includes two aspects of positioning and mapping, which is one of the key issues to achieve robot autonomy. The specific implementation includes visual SLAM, laser SLAM and other methods. The basic process is that the robot pre-estimates the current position according to the sensors such as the carried vision and laser, and then updates its position by observing the feature points of the previous map. Finally, the environment map is extracted according to the current sensor data to update the environment map. It can be seen that the two processes of positioning and mapping of SLAM are interdependent and are a typical "egg-chicken" problem [4].

Although SLAM has solves the problem of autonomous navigation of mobile robots, it has important defects that accumulate errors. It is often used indoors and at low speeds, the accuracy of the created map is not good enough, and the semantic information such as lane information, height and speed limit, and stop line is absent, which cannot meet the requirements of automated driving on open roads.

As a special case of mobile robots, autonomous vehicles are driven on open outdoor structured roads with a priori information. The static background data of roads can be pre-created by modern surveying and mapping methods, combined with GNSS, inertial navigation and odometry. It enables real-time, high precision positioning, which significantly reduces system complexity and cost. The typical representative is Google's unmanned vehicle. By manually collecting environmental sensing information in advance and establishing high precision digital maps, unmanned vehicle uses map information to improve their positioning accuracy and the reliability to find obstacles. Therefore, building accurate road maps is the key to achieve autonomous driving.

3 The Connotation of Automated Driving High Precision Map

Autonomous vehicles rely on GNSS, radar, inertial measurement units, computer vision, LiDAR and other sensors to sense the surrounding environment, and achieve safe driving without any human active operation. The operating environment of autonomous vehicles is a structured road with prior knowledge, which can reduce the technical complexity of SLAM system, but puts forward higher requirements for road map at the same time.

3.1 The Role of High Precision Maps in Automated Driving

The autonomous mobile intelligent robot realizes the purposeful and safe operation by independently detecting the environmental information. This navigation process requires the cooperation of four systems: environmental sensing, real-time positioning, path planning and control [5].

1. Vehicle positioning. Vehicle positioning technology is a prerequisite for automated driving. Vehicle positioning methods mainly include autonomous positioning, land-based positioning and GNSS, but each has many technical limitations. In order to ensure safety, the most reliable method is based on high precision map, multi-source sensor data fusion to achieve high precision and robust real-time positioning.
2. Environmental sensing. Autonomous vehicles use various types of sensors that can sense the surrounding environment, but due to the limitations of sensor capability only in the vicinity of the vehicle. The high precision map can expand the vehicle's static environment sensing, provide beyond visual range view that other sensors can't provide, and reserve more reaction time for behavioral decision making.
3. Behavioral decision. In the decision-making phase of automated driving, path planning, obstacle avoidance, and intelligent speed control must be based on high precision maps. Similar to current mobile navigation functions, the path planning of autonomous vehicles must also be supported by digital maps. However, due to the very harsh working conditions of automated driving, higher requirements are put forward for navigation digital maps. Avoiding obstacles and ensuring driving safety is the most important consideration for automated driving. Dynamic obstacles such as pedestrians and vehicles can be identified by deep learning, but static obstacles must be contained in high precision maps.

Thus, accurate maps are essential for autonomous vehicles positioning, navigation, control and safe driving. But at the same time, it should be pointed out that the autonomous vehicles have different requirements for maps at different phases of positioning, sensing and decision-making. In the positioning phase, SLAM mainly reverses its position and attitude by registering the point cloud acquired by the vehicle-mounted sensors with the landmark in the map. The map at this time can be a 3D sparse map including only landmarks, or even sparse 3D point cloud. Dynamic path planning is suitable for high precision vector maps with topological relationships. Obstacle avoidance is suitable for dense 3D maps, such as 3D dense point clouds [6] (Fig. 1).

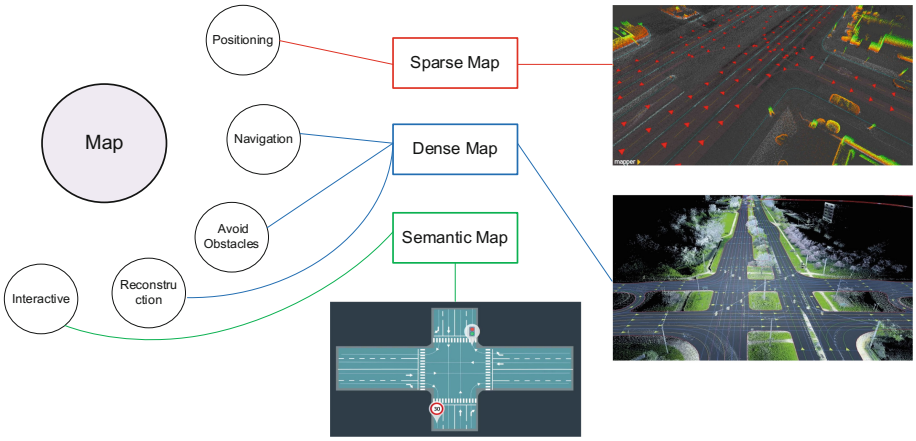


Fig. 1. Differentr kinds of maps needed by SLAM

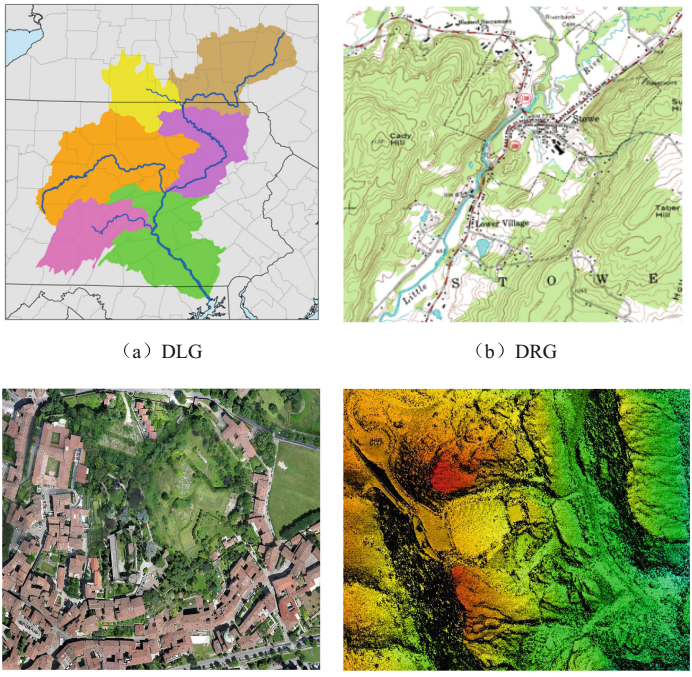
3.2 The Difference Between Automated Driving Map and Regular Map

According to the definition of maps in China's cartography textbooks, "maps are graphs that follow a certain mathematical law, run a symbolic system, and generally reduce the various natural and socio-economic phenomena on the earth on a flat surface" [7]. Obviously, it is specially designed for the purpose of transmitting information. Maps have been the subject of surveying and geography information research for a long time. Their main forms are topographic maps, image maps, thematic maps and navigation digital maps, such as the well-known 4D products: Digital Line Graphic (DLG), Digital Orthophoto Quad (DOQ), Digital Elevation Model (DEM) and Digital Raster Graphic (DRG), illustrate in Fig. 2. To facilitate information transmission, these maps are designed with various intuitive, concise, easy-to-understand and easy-to-remember map symbols to describe various spatial things and phenomena, and their service objects are mainly people. Autopilot technology is essentially a category of robotics. High precision maps mainly serve computers, providing a basis for real-time positioning and control system decision-making of automobiles. Therefore, content and expression methods are inevitably different from current navigation digital maps.

In the expression content, the automated driving of high precision maps not only requires road network information, but also specific to the lane, the accuracy must reach centimeter level, and also requires the road's height limit, speed limit, slope, curvature, bridge bearing and other attributes, as well as steering arrows, traffic lights, signs and other ancillary facilities information [8].

In terms of expression, navigation digital maps are mainly presented in the form of 2D vector maps with topological relationships, However, the high precision map is a combination of 2D and 3D, raster and vector, and point cloud map and semantic map.

On the freshness of map data, the navigation digital map mainly provides reference for people to drive. During the decision-making process, the driver will determine the driving mode by comprehensive maps, road conditions and other factors, so the freshness of the map is not crucial. However, automated driving should plan the driving



（a）DLG （b）DRG

（c）DOQ （d）DEM

Fig. 2. 4D mapping product samples

route independently according to the map and environment. Therefore, road changes must be reflected in the map in real time.

Therefore, compared with conventional maps, autopilot high precision maps have a broader meaning, which is more than just improving the accuracy (or mapping scale) of the map. Instead, it has a richer content and a more diversified expression form, which is more in line with the definition of map by International Cartographic Association (ICA): "map is the representation or abstraction of the real world of geography, and a tool to represent geographic information in a visual, digital or tactile mode" [9].

3.3 The Morphology of the Automated Driving High Precision Map

At present, autonomous vehicles have not yet been mass-produced. How to achieve automated driving has not formed a broad consensus, which indirectly leads to the variety of high precision maps. In summary, the automated driving high precision map mainly has the following forms [10–12]:

1. Raster map: First proposed by Elfes and Moravec, also known as Occupied Grid [13, 14]. For a 2D grid map, the entire environment is divided into uniform cell grids, with each grid assigned a value of 0 or 1 to represent the grid's state: occupied or unoccupied. In this way, the entire environmental space is divided into

occupied space and vacant space. For 3D raster maps, the value of each cell represents the height information of this grid, similar to the Digital Surface Model (DSM) in the surveying field. The resolution of the environment space is related to the size of the grid. Increasing resolution will increase the time and space complexity of the operation.

Due to the complexity of traditional grid representation methods, more efficient quad tree and octree representation methods can be adopted [15].

2. Geometric feature map: It consists of a set of environmental road signs with global coordinate parameters, and each road sign is represented by a geometric prototype, such as points, lines, and areas [16, 17]. This expression can not only provide the spatial reference information needed for positioning, but also have relatively small storage capacity. However, it is relatively difficult to correlate sensor sensing information with map due to the high abstraction of space environment. The 2D geometric feature map is very similar to a digital line graphic or vector map in the field of surveying.
3. Topological map: Initially proposed by Brooks, Mataric and other researchers. Use nodes and connection paths to represent the environment map, and use a map to describe the environment information [18, 19]. A topology map is defined as a graph data structure, graph nodes represent special locations in the environment, and arcs connecting nodes represent path information between specific locations. Obviously, this type of map is very similar to the current navigation digital map. Topological maps are an effective representation for structured environments.

4 High Precision Map of Surveying and Mapping

The information volume and quality of high precision maps directly determine the safety, reliability and efficiency of the automated driving system. High precision maps are more accurate (centimeter-level), faster to update, and contain more information (semantic information) than traditional digital maps [20]. At present, the accuracy requirements of automated driving technology for high precision maps have not been finalized, and the highest demand of current reports is centimeter-level. Considering the efficiency and safety of surveying and mapping operations, the most suitable technical means to achieve such high precision road map production is vehicle-borne mobile measurement and UAV low-altitude photogrammetry.

4.1 Vehicle-Borne Mobile Mapping System

The vehicle-mounted mobile mapping system uses the car as a remote sensing platform, and installs metric CCD digital cameras, high precision LiDAR sensors, and a high precision GNSS/IMU System. After time synchronizing of different sensors, the sequential images and dense laser point clouds can be quickly acquired, and the digital map of the road then can be efficiently and accurately produced after the post-processing and indoor collection.

The vehicle-mounted mobile mapping system operates in Direct Georeferencing (DG) mode enabling direct mapping without ground control. It provides a fast mapping method then reduces the human cost and risk of traditional manual surveying [21].

After more than 30 years of development, the vehicle-mounted mobile mapping system products in the international market are rich and the technologies are relatively mature. The representative ones are the StreetMapper series of the German IGI company, the MX series of Canadian Trimble company, and the LandMark series of Chinese Leador company, etc. However, the high-performance vehicle-mounted mobile measurement system is expensive, and it is reported that the acquisition vehicle of high precision map of AutoNavi company costs 8 million ¥ (Fig. 3).

Fig. 3. StreetMapper and the color point cloud acquired

4.2 UAV Low-Altitude Photogrammetry

Low-altitude photogrammetry of UAV has the advantages of strong maneuverability, simple take-off and landing, convenient transportation, and quick response capability. In recent years, the technology of civilian-grade UAV has developed rapidly, and domestic and foreign companies represented by DJI and eBee have emerged, which has greatly promoted the development of low-altitude photogrammetry technology for UAV. UAV oblique photogrammetry is a new technology developed in recent years. It uses a multi-angle camera (generally no less than 5) to obtain images from multiple view directions. Through aerial triangulation, dense point cloud matching and 3D reconstruction, it constructs a 3D model of the real scene and further produces large-scale maps (Fig. 4).

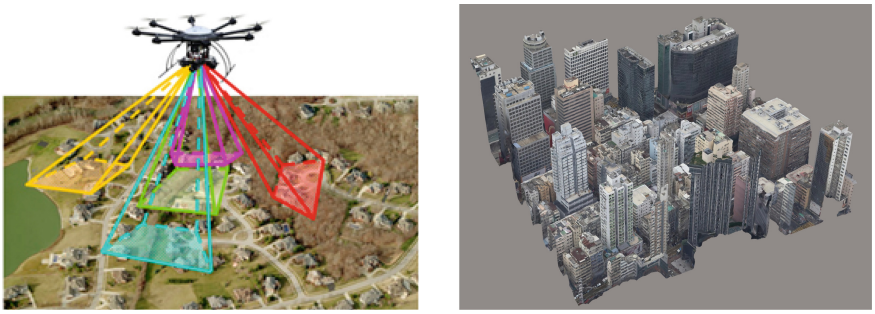


Fig. 4. Oblique photography and reconstruction of true 3D mesh model

The resolution of pixels of UAV ultra-low-altitude oblique photography can reach 1 cm or higher, and the precision of 3D model can reach about 5 cm, which can provide multiple products such as full-area digital elevation model, digital orthophoto quad, 3D point cloud and real scene 3D model. In addition, it can produce high precision road map, which has been successfully applied in highway design and completion surveying. However, the data processing is complex, and the cost of mapping is higher than the vehicle-mounted mobile mapping system.

4.3 GNSS/IMU Positioning and Orientation Technology

One of the issue of road aerial survey mapping is that it is difficult to carry out the ground control point survey, and the operation must be carried out in the way of DG. Therefore, whether it is vehicle-mounted mobile mapping system or low-altitude aerial photography of UAV, the high precision GNSS/IMU system is the core component. It not only ensures time synchronization between different remote sensors, but also provides high precision, high-frequency position and attitude parameters of vehicle or UAV platform during data acquisition, providing accurate external orientation parameters for CCD cameras and LiDAR.

The international mainstream GNSS/IMU suppliers include POS series from Applanix in Canada, AEROControl series from IGI in Germany, and IPAS products from Leica in Switzerland. The post-processing positioning accuracy of Applanix's latest product POS LV 610 is up to 2–3 cm, attitude accuracy is up to 0.0025°. Applanix also offers embedded products AP and APX series for aerial photography (Fig. 5).

（a）POS LV 610　　（b）POS APX 15

Fig. 5. The Vehicle and UAV borned POS products of Applanix

4.4 The Quickly Updating of Maps

Regardless of whether high precision road maps are produced by vehicle-mounted mobile mapping system or UAV low-altitude photogrammetry, all of them have the problems of high cost and long cycle, which cannot meet the current requirements of high precision maps for automated driving. It is only suitable for obtaining the background data of high precision road maps, and other means are needed to update the maps.

Currently, most map manufacturers and autopilot startups around the world use crowdsourcing data to update high precision maps. They cooperate with automakers or travel/logistics to collect road information in real-time using their consumer-grade cameras, laser sensors. By uploading to clouds and processing using big data, deep learning, laser point cloud recognition techniques the high precision maps are updated. According to reports, IT companies of China such as Didi Chuxing, JD, and SF Express are using crowdsourcing technology to lay out their own high precision map production business.

Crowdsourcing mapping is a new technology born with the maturity of surveying and mapping technology and the rapid development of Internet in recent decades. Its technical basis is georeferencing using GNSS devices or smart phone and Web 2.0, and the data source is mainly from GNSS positioning track and geotagged images [22]. One of the most influential projects involved in crowdsourcing geospatial data is OpenStreetMap, which aims to create a world map with free content that can be edited by anyone [23]. By 2015, OpenStreetMap had more than 2 million registered users.

5 Problems with High Precision Map Applications

Due to the variety of forms, complex content and high precision requirements, the high precision map of automated driving faces many challenges before practical use.

5.1 No Industry Standard Has Yet Been Formed

At present, the third working group of the International Organization for Standardization (ISO) Intelligent Transportation System (ITS) Technical Committee (ISO TC204/WG3), which is responsible for map standards, does not introduce standard specifications for automated driving maps. Only some organizations have introduced association specifications, such as OpenDrive, NDS, ADASIS, and SENSORIS, etc.

OpenDRIVE is designed to describe all of the entire road network, but does not express facilities that interact with the road. It is mainly used for simulation and is not suitable for autonomous driving [24]. The high precision map format adopted by Baidu's Apollo open source autopilot platform is also optimized based on OpenDrive. The NDS (Navigation Data Standard) organization consists of automobile manufacturers, map manufacturers, and navigation application developers. The goal is to provide the world's leading autopilot map standard, using a separate mode of map and navigation applications to achieve interoperability of different navigation platforms and support map updates. The ADASIS (ADAS Interface Specification) defines the data model and transmission protocol of the map in ADAS, with CAN as the transmission channel. The SENSORIS (Sensor Interface Specification) defines the data exchange standard between the vehicle-mounted sensor and the cloud.

5.2 Expressed Content and Its Accuracy Are Subject to Policy Restrictions

High precision maps due to high precision, expressed content such as bridge width and bridge bearing may involve military applications, and their data collection, map mapping and Internet publishing are strictly controlled in China. Its production needs to have three qualifications: the use of vehicle mobile mapping or photogrammetry to collect road data, should have map surveying and mapping qualification; navigation digital map data collection, editing, publishing and display, needs to have navigation digital map production qualification. At present, only 12 units in China have this qualification. If it rely on the Internet to provide online map services, it must need to have Internet map service qualifications. At present, China's online map services such as Baidu map, AutoNavi map, etc. have shifted coordinates according to the national requirements, which is obviously contrary to the demand of automated driving. In addition, information necessary for automated driving, such as bridge bearing, width and other policy restrictions, how to transmit through the mobile communication network is also a problem that must be solved in the future.

6 Conclusion

High precision road map is the core of the automated driving technology. From the perspective of surveying and mapping practitioners, this paper makes a comprehensive analysis and discussion of its functions, differences from conventional maps, production methods and problems. It can be seen that the production and application of high precision maps for automated driving involves many fields such as robotics, computer vision, surveying and mapping, big data and cloud computing. It is a new challenge and opportunity for surveying and mapping science, and it will be promising.

Undoubtedly, as the main technical means of GIS data production, surveying and mapping is still an effective and mature way for the production of autonomous high-precision map. However, traditional surveying and mapping must keep up with the progress of current technology, and form a new technological revolution through the cross integration with other disciplines, so as to improve its vitality.

References

1. U.S. Army Engineering Topographic Laboratories (1986) DTIC ADA167472: DAPPA ALV (Autonomous Land Vehicle) Summary
2. Carnegie Mellon University (2017) Look, Ma, No Hands. https://www.cmu.edu/news/stories/archives/2015/july/look-ma-no-hands.html. Accessed 13 Nov 2018
3. Leonard JJ, Durrant Whyte HF (1991) Mobile robot localization by tracking geometric beacons. IEEE Trans Robot Autom 7(3):376–382
4. Durrant Whyte H, Balley T (2006) Simultaneous localization and mapping: part I. IEEE Robot Autom Mag 13(2):99–110
5. Li L, Ye T, Tan M et al (2002) Present state and future development of mobile robot technology research. Robot 24(5):75–480 (ch)

6. Gao X, Zhang T (2017) 14 lectures on visual SLAM: from theory to practice. Electronic Industry Press, Beijing (Ch)
7. Guorui Z (2004) Cartography. Wuhan University Press, Wuhan
8. Kim J, Park B (2017) Defining role, items and accuracy level of high precision road map for autonomous driving. Int J Adv Eng Manag Res 2(1):19–36
9. Yuan K, Zhang R, Wang Y et al (2007) Study on fundamental conceptions and subject system of the modern map and cartography. Geo-Inf Sci 9(04):100–108 (Ch)
10. Xu Z (2004) Simultaneous Localization and Map-building for Mobile Robot. Zhejiang University
11. Chen B (2009) Research on simultaneous localization and mapping of mobile robot in dynamic environments. Zhongnan University (Ch)
12. Deng X (2008) The study on simultaneous localization and mapping algorithms of mobile robots. Nanjing University of Science and Technology (Ch)
13. Elfes (1990) Occupancy grids: a stochastic spatial representation for active robot perception. In: Proceedings of the 6th conference on uncertainty in artificial intelligence
14. Moravec H (1988) Sensor fusion in certainty grids for mobile robots. Sens Devices Syst Robot 9(2):253–276
15. Amir AE, Indyk P et al (1999) Efficient regular data structures and algorithms for location and proximity problems. In: IEEE symposium on foundations of computer science, pp 160–170
16. Chong K, Kleeman L (1999) Mobile robot map building from an advanced sonar array and accurate odometry. Int J Robot Res 18(1):20–36
17. Leonard JJ, Durrant-Whyte HF, Cox IJ (1992) Dynamic map building for an autonomous mobile robot. Int J Robot Res 11(4):286–298
18. Matari MJ (1990) Environment learning using a distributed representation. In: Proceedings of IEEE international conference on robotics and autonomation
19. Francescio S, Benjamin K (2004) Loop-closing and planarity in topological map building. In: IEEE international conference on intelligent robots and systems, no 2, pp 1511–1517
20. Liu S, Tang J, Wu S et al (2017) First book on self-driving. Electronic Industry Press, Beijing (Ch)
21. Deren L (2006) Mobile mapping technology and its applications. In: Annual conference of chinese society of surveying and mapping
22. Heipke C (2010) Crowdsourcing geospatial data. ISPRS J Photogramm Remote Sens 65 (6):550–557
23. WikiPeia (2019) OpenStreetMap. https://en.wikipedia.org/wiki/OpenStreetMap. Accessed 9 Jan 2019
24. WikiPedia (2018) https://en.wikipedia.org/wiki/OpenDRIVE_(specification). Accessed 13 Nov 2018

Validation and Analysis of TPXO over Offshore Areas Using GPS-IR Technology

Huilin Wu[1(✉)], Shuangcheng Zhang[1,2(✉)], Huayi Zhang[3], and Qi Liu[1]

[1] College of Geology Engineering and Geomantic, Chang'an University, Xi'an 710054, China
478875326@qq.com, shuangcheng369@vip.163.com
[2] China Meteorological Administration Urumqi Institute of Desert Meteorology, Urumqi 830002, China
[3] First Institute of Oceanography, SOA, Qingdao 266061, Shandong, China

Abstract. The ocean tide model is an important factor in determining high-precision and high-resolution gravity field models, sea surface topography and ocean circulation. With the rapidly development of GNSS theory and applications, GNSS-IR technology based on multipath effects has been proved by scholars can be used for tidal level monitoring. In order to further expand the application of GNSS-IR technology in the ocean tide field, this paper takes the GNSS measurement station sc02, which located on the shore of the Friday Harbor in Washington, USA, as an example to analyze the tidal wave with 13-year GNSS dataset. Combined with the traditional tide gauge data, the accuracy analysis of the global ocean tide models TPXO6.2 and TPXO09 using GPS-IR technology is carried out. The experimental results show that: the results of amplitude and retardation obtained from the GNSS-IR inversion tide level are different from the results obtained by the corresponding tide station. But the tidal coefficients, which obtained by the inversion of the tidal level of the GNSS station and which calculated by the ocean tide model, are showing a high consistency. The GNSS-IR technique has higher precision for obtaining the tidal coefficient.

Keywords: GPS-IR · Ocean tide model · Tide monitoring · Tidal coefficient

1 Introduction

Coastal and estuary areas are flat and have the geographical advantages of land-sea communication. These areas are densely populated and developed in industry and agriculture. They are the economic center and transportation hub of a region. However, the area is located in the tidal zone, which is affected by the tides all the year round, which will directly or indirectly affect people's normal production and life, and sometimes even threaten the people's lives and property. Therefore, the study of coastal and estuarine tidal essential. The coastal and estuary areas are greatly affected by the tides. For the production and living needs of the people in the coastal areas, there is an urgent need to deeply understand the laws of the tides. The harmonic analysis method

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 277–286, 2019.
https://doi.org/10.1007/978-981-13-7751-8_28

is one of the main methods for tidal analysis and forecasting. The harmonic constant obtained by harmonic analysis is an important characteristic parameter of the tides of the coast and estuary, which plays a key role in the process of forecasting tides and compiling tide tables [1–3].

In 1980, Schwiderski first combined the dynamic equation with the measured data from more than 2,000 tide stations in the world, expanded the Laplace tidal wave equation, and established the world's first ocean tide model, but due to the lack of tide data. The tidal wave coefficient of this model is extremely unstable and the error is large [4]. The emergence of satellite altimetry technology has greatly promoted the study of ocean tides. Many high-precision global tide models have been obtained using tide gauge data and satellite altimetry data, such as NAO.99B, FES2004, SCW80 and TPXO models, et al. [5, 6]. TPXO model belongs to oceanic tidal assimilation model. It is based on tidal wave force equation, uses various assimilation methods, and uses tidal data to improve the accuracy of tidal simulation results, so as to obtain more accurate distribution of global oceanic tides [7].

In recent years, many scholars have carried out a lot of research work on the accuracy assessment of the global tide model, and have achieved fruitful results. Li used the traditional tide gauge data to evaluate the accuracy of five global ocean tide models NAO99b, FES2004, GOT4.7, TPXO7.2 and EOT10a. The results show that EOT10a has the best comprehensive index in the global ocean area and can reflect well. The distribution of global ocean tides [8]; Fu compared the accuracy of three global tide models DTU10, TPXO7.2, NAO99b and one regional tide model NAO.99jb on the coast of China through tide gauge data. The results show that NAO.99Jb has the highest accuracy in China's seas, followed by NAO.99b [9]. Sun analyzed the accuracy of CSR4.0 model and NAO.99b model in China sea based on tide data. The average accuracy of the CSR4.0 model in the offshore China is better than the NAO.99b model [10]. It is not difficult to see from the above that the above scholars are based on the traditional tide gauge data to study and analyze the tide model [11], but the number of tide stations in the world or region is limited, and the space-time resolution is low. With the deepening of GNSS research and application, GNSS-IR technology based on multipath effect has become an emerging tidal level monitoring method [12, 13]. At present, countries have successively deployed GNSS stations in coastal areas. China has also continuously added GNSS monitoring stations along the coast of non-test stations, which will further promote the development of GNSS-IR technology. To some extent, coastal GNSS monitoring stations can be used. An effective supplement to the tide station. The accuracy of GNSS-IR technology for sea level change monitoring is gradually increasing. However, there are relatively few studies on using this technology to obtain tidal wave coefficient and its ocean tide model. This paper uses the measured data of coastal GPS stations to test. And use tide gauge station, the tidal wave coefficient obtained by the data verifies the accuracy of the tidal wave coefficient acquired by GPS-IR, and then compares it with the tidal wave coefficient obtained by the TPXO ocean tide model.

2 GPS-IR Monitoring Tidal Wave and Tidal Coefficient Extraction Principle

GNSS multipath effect is one of the main error sources affecting precision positioning accuracy. The generation of GNSS multipath is mainly related to the structure and dielectric parameters of the reflective surface. The signal received by the GNSS receiver located near the sea surface is actually a direct signal and a composite signal reflected by the sea surface [14] (Fig. 1).

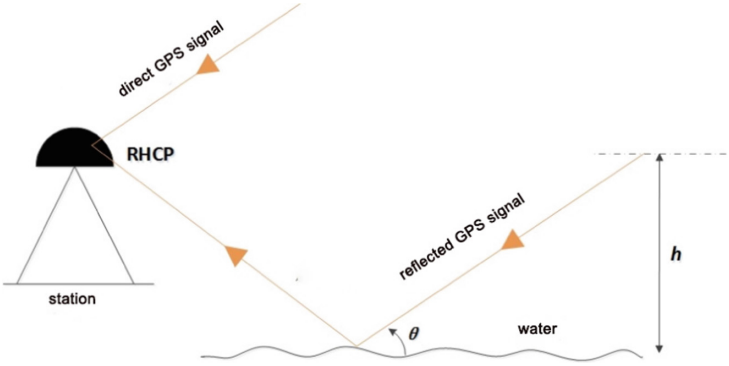


Fig. 1. GPS-IR extracts tidal level change intention

Figure 1 shows the basic principle of GNSS-IR technology for monitoring tidal level changes, where is the angle between the direct signal and the instantaneous tidal plane, which is the distance from the phase center of the receiver antenna to the instantaneous tidal level, that is the vertical reflection distance.

Setting the amplitude of the direct and reflected signals and respectively. For geodetic receivers, the direct signal amplitude can be separated by a quadratic polynomial fitting method. The relationship between signal amplitude and SNR can be expressed as [15]:

$$SNR = A_c^2 = A_d^2 + A_m^2 + 2A_dA_m\cos\theta \tag{1}$$

Where A_c is the amplitude of the actual received signal and $\cos\theta$ is the cosine of the angle between the direct signal and the reflected signal.

The amplitude of the reflected signal obtained after separation can be expressed as:

$$A_m = Acos\left(\frac{4\pi H}{\lambda} sinE + \emptyset\right) \tag{2}$$

Where λ is the carrier wavelength, E is the satellite elevation angle, is the vertical reflection distance. If counted t = sinE, f = $2h/\lambda$, then Eq. (2) can be expressed as:

$$A_r = Acos(2\pi ft + \emptyset) \tag{3}$$

The vertical distance h is a function of the reflection frequency f, $\sin E$ altitude angle is known with varying amounts, but non $\sin E$ sampling intervals, so that the sequence can not guarantee positive SNR residual period truncation, paper uses Lomb-Scargle (LS) spectrum analysis processing.

The actual ocean tide level can be expressed as a superposition of a series of cosine terms on a fixed value [15]:

$$\begin{aligned}\zeta(\mathrm{t}) = h_0 + a(t - t_0) \\ + \sum_i f_i H_i(\lambda, \phi)\cos\left[\sigma_i t + (V_0 + u)_i - \mathrm{g}_i(\lambda, \phi)\right] + r\end{aligned} \tag{4}$$

In Eq. (4), h_0 is the average sea surface; a is the trend term for sea surface change; r is the observed noise; σ is the angular velocity of the divided tide; t_0 is the reference time; f, u are the corrected value of the average amplitude and phase angle introduced by the change of the lunar orbit 18.6 years; H, g are the amplitude of the tide and the Greenwich retardation, which are the harmonic constants of the actual moisture.

In actual harmonic analysis, the tide can be expressed as:

$$\begin{aligned}\zeta(\mathrm{t}) = h_0 + a(t - t_0) \\ + \sum_i f_i C_i(\lambda, \phi)\cos\left[\sigma_i t + (V_0 + u)_i\right] \\ + \sum_i f_i S_i(\lambda, \phi)\sin\left[\sigma_i t + (V_0 + u)_i\right] + r\end{aligned} \tag{5}$$

Which

$$\begin{cases} C_i = H_i cos\mathrm{g}_i \\ S_i = H_i sin\mathrm{g}_i \end{cases} \tag{6}$$

By analyzing the Eq. (5), we can get the tides, and the amplitude and retardation of the tide can be obtained by considering Eq. (6):

$$\begin{cases} H_i = \sqrt{C_i^2 + S_i^2} \\ \mathrm{g}_i = arctan\left(\frac{S_i}{C_i}\right) \end{cases} \tag{7}$$

3 Data Analysis

3.1 Data Sources

In order to verify the accuracy of the tidal wave coefficient obtained by GNSS-IR technology, this paper uses the observation data of the GNSS continuous operation

tracking station SC02 station (Fig. 2) located on the shore of the Harbor in Washington, USA, for experimental analysis, which belongs to the United States. The PBO network is observed at the edge of the plate in the Earth Scope. The SC02 station is built adjacent to the sea and can receive GPS reflection signals from the sea surface in a large space. The station has more than 10 years of continuous observation data, and good data integrity. Figure 2 shows the receiver layout and observation environment of the SC02 station. The receiver used in the SC02 station is the TRIMBLE NETR9 geodetic receiver. The GPS antenna is the TRIMBLE company's choke coil antenna with a fairing (SCIT) (TRM59800.80). At the same time, a comparative analysis was carried out using the measured data of the Friday Harbor tide station at 359 m from the SC02 station. The Friday Harbor Tidal Checkpoint is a continuous operation tide gauge station maintained by the National Oceanic and Atmospheric Administration's Marine Products and Service Center. The tide gauge station was built in 1934 and was equipped with Aquatrak's acoustic test in 1996. The tide gauge can provide tidal level data with a sampling interval of 6 min. Figure 3 shows the position distribution of the SC02 GPS station. The reflection height is 4 m. The color part shows the Fresnel reflection with elevation angles of 5°, 8° and 12° region.

Fig. 2. Surroundings of SC02 station

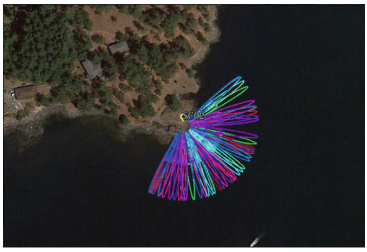


Fig. 3. The Fresnel zones of the SC02

3.2 Experimental Analysis

Figure 4 uses July 2017 (doy 193 to 199) SC02 station L1 signal to noise ratio of the data acquisition process according to the second portion of the tide, and compared with the results of Friday Harbor tidal stations.

It can be clearly seen from Fig. 4 that the GPS-IR inversion results are basically consistent with the observations of the tide gauge, but there are still some deviations in the inversion results. Mainly due to the rapid changes in the sea surface waves and the impact of the gross error elimination method. The poor average of the results obtained by GPS-IR compared with the results of the tide gauge station was 2.93 cm, and the RMS of the difference was 3.57 cm.

In order to further verify the validity and continuity of GPS-IR technology inversion of tidal level changes, the GNSS station (SC02) uses the data length of thirteen

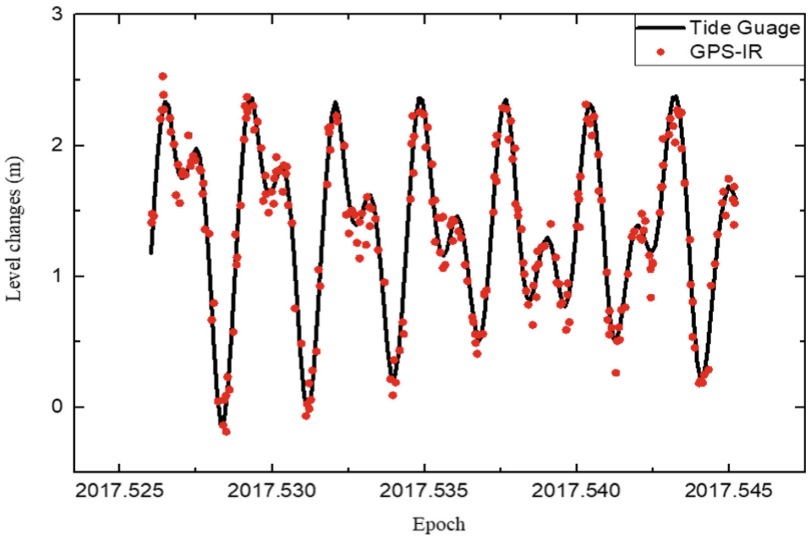


Fig. 4. Comparison of tide level changes in a week

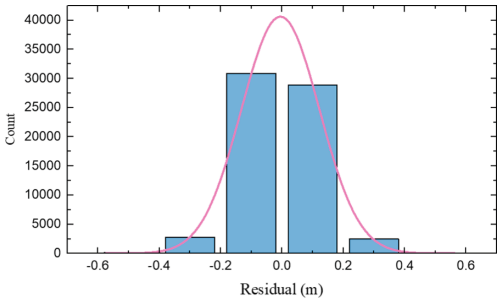


Fig. 5. Histogram of error distribution

years from 2005 to 2017. In order to verify the accuracy of GPS-IR acquisition of tidal wave coefficients, the tide gauge station also uses the same time span data for experiments (shown in Figs. 5 and 6).

Figure 5 shows the error distribution histogram. It can be clearly seen that the error is mostly concentrated around ±15 cm, and the fitting residual distribution is roughly obeying the normal distribution. It can be clearly seen from Fig. 6 that the two have a good consistency in the overall trend.

Accurate and stable tidal wave coefficients generally require tidal data of more than 18.6 years. However, the time span of the SC02 station data is not long enough, and the station is located in a shallow water area, there is a large number of nonlinear composite tides. The results of the harmonic analysis showed that there were 102 tidal components with amplitudes greater than 1 mm in Friday Harbor, and we selected 9 of the larger tidal ranges (Table 1).

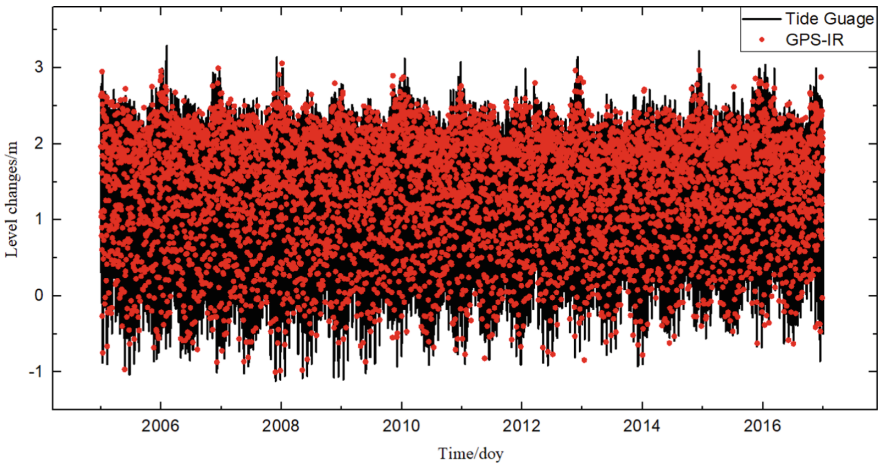


Fig. 6. Thirteen-year sea level changes from tide gauge and GPS-based tide gauge

Table 1. Comparison of tidal wave coefficients obtained by tide gauge station and GPS-IR

Tide	Friday harbor		GNSS (SC02)		Diff
	A (cm)	G (deg)	A (cm)	G (deg)	d (cm)
Mf	2.38	144.19	2.84	146.72	0.47
Q_1	7.27	228.49	7.05	229.05	0.23
O_1	42.78	237.76	41.55	238.00	1.24
P_1	24.93	277.30	23.81	277.63	1.13
K_1	75.32	270.23	71.94	270.56	3.41
N_2	11.98	315.28	11.01	316.13	0.98
M_2	55.92	343.00	51.58	343.90	4.42
S_2	13.25	5.84	11.87	6.60	1.39
K_2	3.61	28.24	3.23	29.00	0.38

The vector difference in the table is calculated as follows:

$$\mathrm{d} = \left| A_1 e^{-iG_1} - A_2 e^{-iG_2} \right| \tag{8}$$

Where, A_1, A_2 are the same tidal wave amplitude corresponding to the two methods, G_1, G_2 are the phase value of each obtained.

It can be seen from Table 1 that the amplitude and phase acquired by GPS-IR agree well with the amplitude and phase acquired by the tide gauge station, and the vector difference between the two is within 5 cm. Among them, the partial tides with amplitudes greater than 20 cm have K_1, M_2, O_1, P_1, and their vector differences are large, while the tidal waves Mf, Q_1, K_2 with smaller amplitudes have smaller vector differences.

According to the literature [16], the Friday Harbor has only one climax and one low tide every day, so it is an obvious sun tide port. The tidal wave is the maximum amplitude of the tide, and its period is a stellar day. It is worth noting that the tide frequency Basically, it is the same as the orbital frequency of the GPS satellite constellation. Since the satellite geometric position is related to the multipath error, the GPS reflection data may be similarly related to the geometric error. But the actual impact was found on GPS-IR obtain tidal wave coefficient is relatively small.

The difference between the two methods may also have a certain relationship with the tide station and the GNSS station, and the difference will also include the difference between the two locations. For the tidal wave analysis itself, only a time series of more than 18.6 years can provide reliable estimates. Even so, many mixed tidal waves have great uncertainty.

In short, the amplitude and phase of each tidal wave obtained under the two methods are not much different. The tidal wave analysis based on GPS-IR can get the same result as the tidal station. Then, using the TPXO9 and TPXO6.2 ocean tide models for comparative analysis, the tidal wave coefficients of the nine tidal waves obtained by the GPS-IR inversion tidal level, TPXO6.2, and TPXO9 tidal model are shown in Table 2:

Table 2. Tidal wave coefficient statistics table obtained by GPS station and TXPO model

Tide	TPXO9		TPXO6.2		GNSS (SC02)	
	A (cm)	G (deg)	A (cm)	G (deg)	A (cm)	G (deg)
Mf	1.59	170.9749	1.55	167.6360	2.84	146.72
Q_1	42.33	271.8331	0.01	347.6236	7.05	229.05
O_1	59.91	204.2858	26.19	226.0167	41.55	238.00
P_1	58.76	210.4402	0.03	357.7957	23.81	277.63
K_1	84.63	237.3035	42.93	240.4076	71.94	270.56
N_2	14.53	325.9247	0.01	122.7984	11.01	316.13
M_2	34.38	46.3724	92.75	235.7758	51.58	343.90
S_2	28.17	234.877	27.44	261.9921	11.87	6.60
K_2	3.01	188.7905	0.1	160.5209	3.23	29.00

It can be seen from the table that for Mf, K1, N2, M2 and K2, the TPXO9 model agrees better with the amplitude calculated by the SC02 station than the TPXO6.2, but for the Q1, O1, P1, and S2, they are equivalent; for the late angle, the TPXO9 model and the TPXO6.2 model are equivalent to the delayed angle of the SC02 station solution, and the difference between the two is large. The reason for this phenomenon may be related to the low accuracy of the TPXO tide model for this sea area [17, 18]. It is clear that GPS-IR and tide station overall good agreement, and compared to TPXO9 TPXO6.2, GPS-IR obtain tidal wave has a higher coefficient of precision in terms of the inversion. To sum up, the tidal wave based GNSS-IR analysis results can be obtained with the corresponding tidal stations.

4 Summary and Outlook

Based on the GPS-IR monitoring of tidal level changes based on SNR observations, the method of acquiring tidal wave coefficients by GPS-IR technique is studied and compared with the ocean tide model. The accuracy of the tidal wave coefficient obtained by the GPS-IR technique is higher than that of the TPXO model.

GPS-IR technology has the advantages of high spatial and temporal resolution, absolute sea surface monitoring, no contact with seawater, and multi-use in the acquisition of regional tidal wave coefficients. At the same time, the high-precision acquisition of the tidal wave coefficient provides a possible integration of satellite altimetry for GNSS-IR technology, providing a possibility for serving fisheries, salt industry, port construction, and seawater power utilization. It is worth noting that this paper only conducts a test analysis on one station, and then needs to verify and solve it through a large number of station data, and combines FES, DTU and NAO.99 models for research.

Acknowledgments. We gratefully acknowledge the provision of data, equipment, and engineering services by the Plate Boundary Observatory operated by UNAVCO for Earth Scope. This work was supported by China Desert Meteorological Science Research Foundation (Sqj2017002), National Science Foundation of China (41731066, 41674001, 41104019) and the Special Fund for Basic Scientific Research of Central Colleges (310826172202).

References

1. Tong Z (2007) Research on methods and applications of tidal harmonic analysis. Hohai University
2. Chen Z (1980) Chaoshanology. Science Press
3. Hu Z, Guo J, Zong G et al (2014) Accuracy analysis of three tidal models using data from tidal station. Ocean Mapp 34(3):13–16
4. Schwiderski EW (1980) Ocean tides, part I: global ocean tidal equations. Mar Geodesy 3(1–4):161–217
5. Lei J, Xu J (2013) Accuracy assessment of tidal harmonic constants in China's coastal tidal stations. Ocean Mapp 33(1):1–4
6. Dong X, Huang W (2000) Monitoring global sea level change using TOPEX/Poseidon satellite altimetry data. J Surv Mapp 29(3):266–272
7. Liu P, Li S, Ai C (2017) Accuracy assessment of different tidal prediction models in offshore waters. Water Resour Plan Des (10):76–79
8. Li D, Li J, Jin T et al (2012) Using the tide gauge data to evaluate the accuracy of the global tide model. Geodesy Geodyn 32(4):106–110
9. Fu Y, Zhou X, Zhou D et al (2017) Accuracy assessment of China's nearshore tide model using tide station data. J Surv Mapp 42(8):28–32
10. Sun J, Guo J, Guo S et al (2013) Accuracy analysis of CSR4.0 model and NAO.99b model based on tide data in Chinese seas. Prog Geophys (5)
11. Cheinway H, Chen S (2000) Fourier and wavelet analyses of TOPEX/Poseidon-derived sea level anomaly over the South China Sea: a contribution to the South China Sea Monsoon Experiment. J Geophys Res: Ocean (1978–2012) 28785–28804

12. Larson KM, Löfgren JS, Haas R (2013) Coastal sea level measurements using a single geodetic GPS receiver. Adv Space Res 51(8):1301–1310
13. Bilich A, Larson KM, Axelrad P (2004) Observations of signal-to-noise ratios (SNR) at geodetic GPS site CASA: implications for phase multipath. Proc Cent Eur Geodyn Seism 23:77–83
14. Zhang S, Nanyang Li Z et al (2016) GNSS-MR technology for monitoring and analyzing tidal level changes. Acta Geodae Sinica 45(9):1042–1049
15. Li J, Liu Y, Xiao F (1995) Error analysis of tidal analysis and forecast. Ocean Mapp (1):31–37
16. Munekane H (2013) Sub-daily noise in horizontal GPS kinematic time series due to thermal tilt of GPS monuments. J Geodesy 87(4):393–401
17. Shum CK, Woodworth PL, Andersen OB et al (1997) Accuracy assessment of recent ocean tide models. J Geophys Res 11(15):125–173
18. Jia D, Sheng S, Zhang L et al (2016) Analysis of deformation characteristics of ocean tide loads in regional GPS stations in China. J Geodesy Geodyn 36(3):206–210

Evaluation and Improvement of BDS Dynamic Positioning Performance in the Smartphone

Zhenqiang Du, Hongzhou Chai(✉), Zongpeng Pan, and Chunhe Liu

Information Engineering University, Zhengzhou, China
gnsserl996@163.com, chaihzl969@163.com

Abstract. Since Google launched the Android 7.0 operating system, smartphone has opened the original GNSS observation reading interface for the first time, enabling developers to further develop precision positioning and location service applications based on the original GNSS observation. Aiming at different brands of Android smartphones, the original BDS data quality of smartphone is analyzed and evaluated in terms of the number of visible satellites, signal-to-noise ratio, pseudorange noise, Doppler velocimetry. Comparing the current measurement receivers, the static and dynamic positioning results of smartphone are analyzed. A dynamic positioning algorithm based on Doppler velocimetry and adaptive Kalman filtering is proposed, which effectively improves the positioning accuracy and reliability.

Keywords: BDS · Smartphone · Data quality · Dynamic positioning

1 Introduction

At present, smart terminal products are developing rapidly. Android smartphones such as HUAWEI, Xiaomi, Samsung, vivo and OPPO occupy about 86% of the Chinese market. With the popular application of GNSS (Global Navigation Satellite System, GNSS) chips, more and more smartphones can support the BDS (BeiDou Navigation Satellite System, BDS). Since Google released the Android 7.0 operating system in May 2016, the smartphone has opened the original GNSS observation reading interface, enabling users to use mobile devices such as smartphones to obtain the data of GPS, GLONASS, and BDS. The original measurement data of satellite navigation systems such as Galileo and QZSS brings new development opportunities for more convenient applications with high-precision location services.

Several studies have attempted to analysis the original GNSS data quality and positioning accuracy of smartphones: Airbus' Moisés Navarro-Gallardo converted the original GNSS data format collected by the smartphone to RINEX format, and evaluated the static positioning accuracy of Galileo. Maninder Gill of York University analyzed the data quality collected by Nexus9 and U-blox chips, and used static precision single point positioning algorithm to evaluate the positioning accuracy. Zhang Xiaohong of Wuhan University analysed the GPS data quality collected by the Nexus9, and proposed a TD Filter algorithm to improve the positioning accuracy.

However, the above research is mainly aimed at GPS observation data in smartphone. The data quality of BDS in smartphone is still unclear. The data quality of

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 287–295, 2019.
https://doi.org/10.1007/978-981-13-7751-8_29

GNSS collected by smartphones such as HUAWEI and Xiaomi, which have a large domestic market share, still needs to be analysed. The accuracy of positioning, especially under dynamic conditions, remains to be further studied.

2 Data Collection and Quality Assessment

2.1 Data Collection Experiment

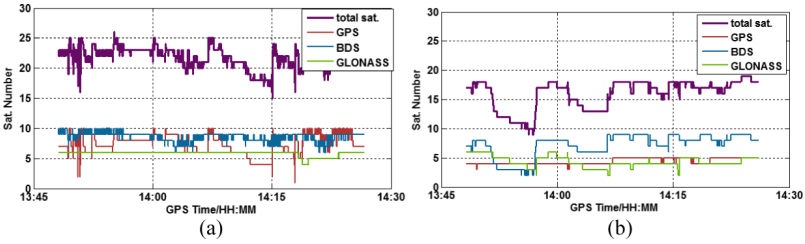
Experiments were conducted at the track and field of Information Engineering University, from 20:00 to 22:20 on October 08, 2018, as shown in Fig. 1. The experimental equipment is two smartphones, two Trimble R10 receivers, and two ComNav boards with zero baseline connection through the power splitter.

(a) (b)

Fig. 1. The static mode is adopted firstly at 20:00–21:30, as shown in left (a). The dynamic mode is adopted at 21:45–22:30, as shown in right (b).

2.2 Satellite Visibility

It can be seen from Fig. 2 that both HUAWEI P20Pro and Xiaomi 6 can support three satellite navigation systems: GPS, BDS and GLONASS. The number of visible satellites in Xiaomi 6 reached about 20 with about 7 BDS. The number of visible satellites of HUAWEI P20Pro reached about 17 with about 4 BDS. For smartphone, the number of visible GNSS satellites has fully meet the positioning requirements.



(a) (b)

Fig. 2. In the dynamic condition, the number of visible satellites in Xiaomi, as shown in left (a). The number of visible satellites in HUAWEI P20Pro, as shown in right (b).

2.3 Signal to Noise Ratio and Satellite Elevation Angle

It can be seen from Fig. 3(a) and (c) that for the same BDS in the same time, the SNR of HUAWEI P20Pro is around 30dBHz, while the ComNav can reach above 40dBHz, which is obviously superior to smartphone. As shown in Fig. 3(a), due to the dynamic situation, the smartphone loses lock on the BDS C22 satellite, and the satellite elevation angle changes significantly, which affects the SNR of C22.

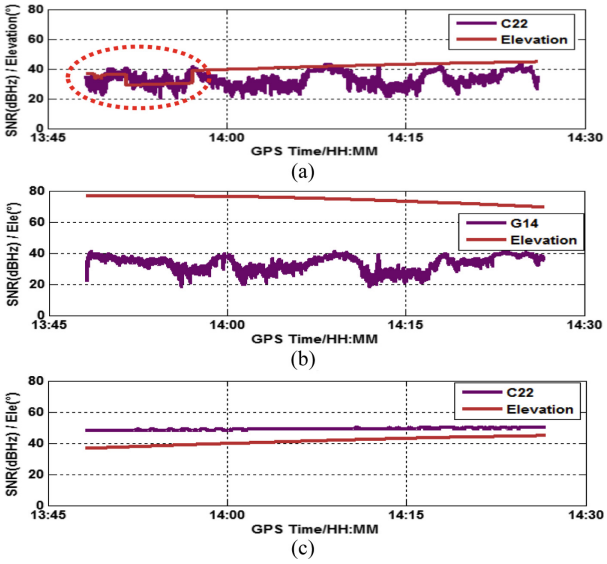


Fig. 3. (a), (b) are the SNR and corresponding satellite elevation angles of BDS C22 and GPS G14 observed by HUAWEIP20 Pro. (c) shows the BDS C22 observed by the ComNav board.

2.4 Pseudorange Observation Noise

In order to evaluate the noise level of the BDS pseudorange observation in smartphone, CC (Code minus Carrier) combined observations were introduced. which is:

$$CC = P - L \tag{1}$$

P is the pseudorange observation. L is the carrier phase observation. Due to the subtraction of pseudorange and carrier observations can eliminate the geometrical bias, such as the satellite clock bias, the receiver clock bias, the receiver-satellite distance, and the tropospheric error. So the CC combined observations after the subtraction between the epochs, only remains pseudorange observation noise (Fig. 4).

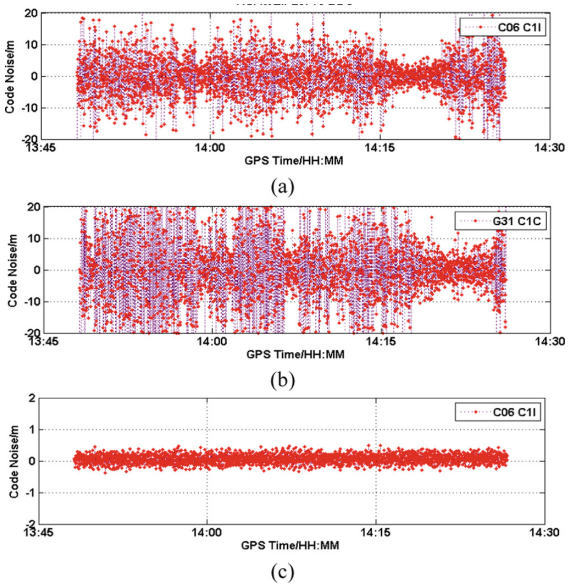


Fig. 4. (a) and (b) show the pseudorange observation noise of HUAWEI P20Pro BDS and GPS respectively. (c) shows the pseudorange noise of BDS observed by ComNav.

Table 1. RMS and STD of Pseudorange noise (m)

Code Bias (m)	HUAWEI BDS	HUAWEI GPS	ComNav BDS
RMS	5.307	5.376	0.152
STD	5.308	5.376	0.138

As can be seen from Table 1, the HUAWEIP20 Pro BDS and GPS pseudorange noise RMS are 5.31 m and 5.38 m respectively. The BDS pseudorange noise of the ComNav board is only 0.15 m. In dynamic condition, the pseudorange noise is still quite different between smartphone and the current measurement receiver.

2.5 Doppler Variation Values and Velocimetry

In order to facilitate the evaluation of the accuracy of Doppler velocimetry, the static mode is used for Doppler velocimetry with the smartphone and ComNav board. As can be seen from Fig. 5(a), the Doppler variation values observed by HUAWEIP20Pro and ComNav boards are relatively flat. From Fig. 5(b), (c), (d), it can be seen that the Doppler velocimetry results of smartphone are worse than ComNav board. But the velocimetry results in X, Y and Z directions are basically no more than 0.5 m/s.

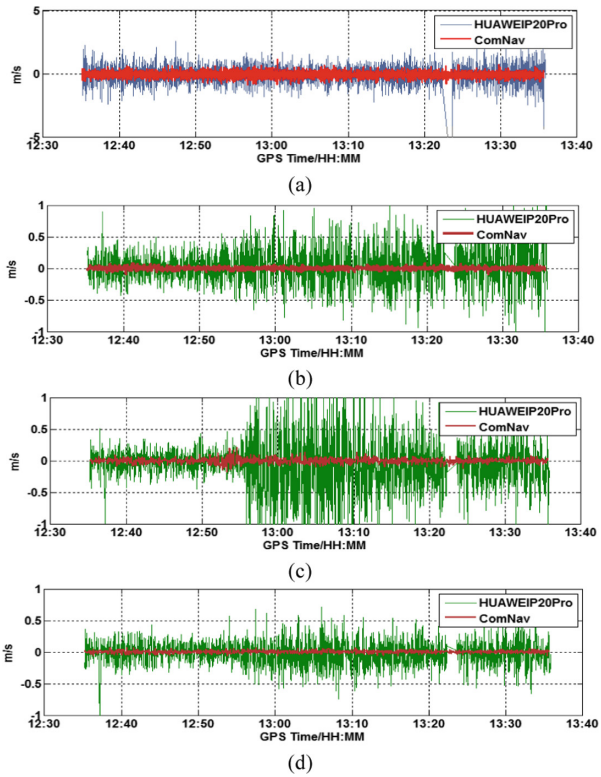


Fig. 5. The Doppler variation values of the HUAWEI P20Pro and ComNav boards are shown in (a), and the Doppler velocimetry results in the X, Y, and Z directions are shown in (b), (c), and (d) respectively.

3 Evaluation and Improvement of Positioning Results

3.1 Evaluation of Positioning Result

It can be seen from Fig. 6 that the BDS, GPS and BDS/GPS combined systems have the accuracy of 50.34%, 63.35% and 78.10% respectively less than 10 m. It indicating BDS/GPS combination system can increase the accuracy and reliability of pseudorange positioning.

Based on the positioning results of the Trimble R10 GPS RTK, the pseudorange dynamic positioning accuracy of the smartphone is evaluated.

It can be seen from Fig. 7(a), (b) and (c) that the accuracy of BDS, GPS and BDS/GPS combined systems have 52.53%, 81.32% and 91.09% within 15 m respectively for the smartphone to adopt the pseudorange positioning mode under dynamic conditions. For Fig. 7(d), with the BDS/GPS combined positioning mode, 80% of the positioning results have an accuracy of 11.69 m. In the dynamic situation, the BDS/GPS combination system of the smartphone can improve the accuracy and reliability of the pseudorange positioning, but its positioning accuracy still needs to be improved.

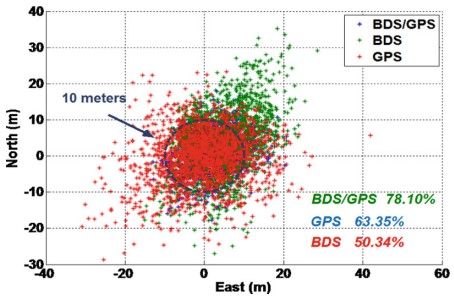


Fig. 6. Taking the static RTK result of TrimbleR10 as the datum, Figure shows the results of positioning of Xiaomi 6 in the N and E directions by BDS, GPS and BDS/GPS combined system.

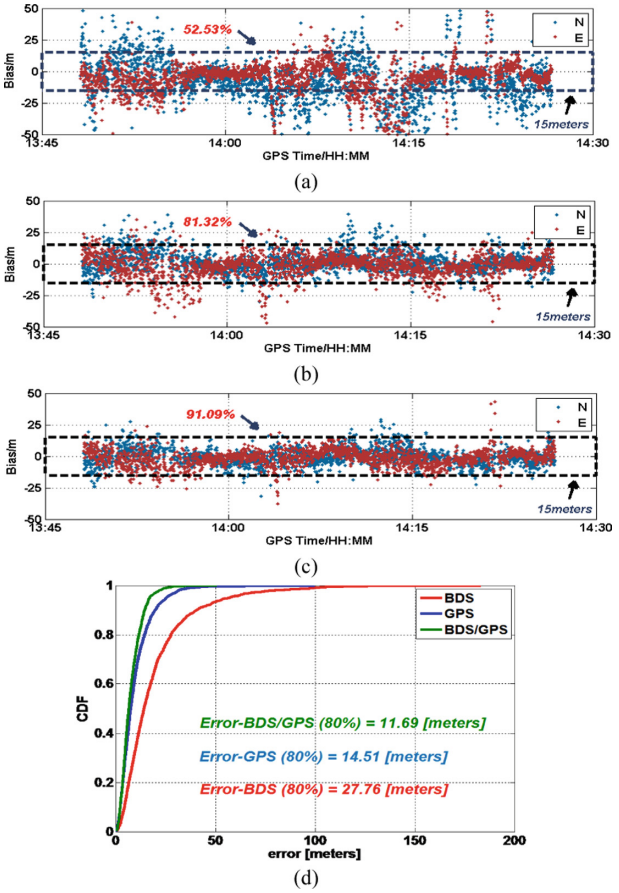


Fig. 7. (a), (b), (c) are the dynamic pseudorange positioning results of the HUAWEI P20Pro BDS, GPS, BDS/GPS combined system. (d) is the statistical results of dynamics positioning accuracy of the HUAWEIP20ProBDS, GPS, BDS/GPS combined system.

3.2 Improved Algorithm for Precise Single Point Positioning

The accuracy of smartphone dynamic pseudorange positioning is not high enough to meet the needs of higher precision applications. Because the smartphone is easy to cause the satellite lose lock. So the Doppler velocimetry is used to improve the state equation of the system, and adaptive Kalman filter is adopted to get the location results of smartphone (Fig. 8).

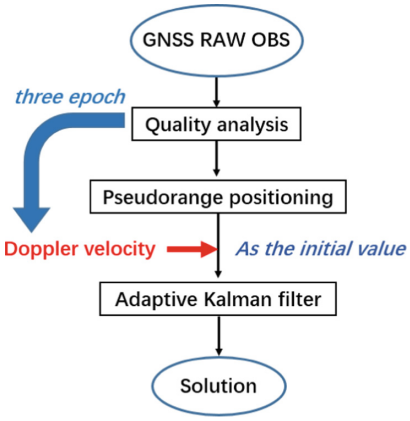


Fig. 8. Algorithm flow chart. A dynamic positioning algorithm based on Doppler velocimetry and adaptive Kalman filtering

Firstly, the original data quality is analyzed, the poor data is eliminated and the Doppler velocimetry is carried out. Then every three epochs are processed once. A position of the current epoch is deduced from the Doppler velocimetry results of previous two epochs. Another position is the pseudorange solution of the current epoch. Weighted them as the initial value of dynamic PPP, the result is obtained by adaptive Kalman filtering finally.

$$\overline{X}_k = \alpha X_k^0 + \beta\left(X_{k-1}^0 + v_{k-1} \times \Delta t_{k-1}\right) + \gamma\left(X_{k-2}^0 + v_{k-2} \times \Delta t_{k-2}\right) \tag{2}$$

$$e_k' = L_k - A_k \bar{X}_k \tag{3}$$

X_k^0 is the pseudorange solution of k epoch. V_{k-1}, V_{k-2} are the speed obtained by Doppler velocimetry results of k − 1, k − 2 epoch. Δt_{k-1}, Δt_{k-2} are the corresponding time difference. α, β, γ are its corresponding weight. So the predicted value $\bar{X}_k$ of k epoch can be deduced synthetically. A_k is the vector of observations. L_k is a matrix of coefficients.

It can be seen from Fig. 9(a) and (b) that the improved PPP positioning result is obviously better than the pseudorange positioning result, especially for some bias. From Table 2, the RMS of dynamic pseudorange results of HUAWEI P20Pro are 8.135 m, 10.004 m, and 25.080 m in the three directions of N, E, and U respectively.

The RMS of improved PPP in N, E, and U directions are 5.844 m, 7.293 m, and 17.704 m respectively. The positioning accuracy is better improved in all three directions, and the N and E directions that converge after 12 min can reach 3.294 m and 6.375 m.

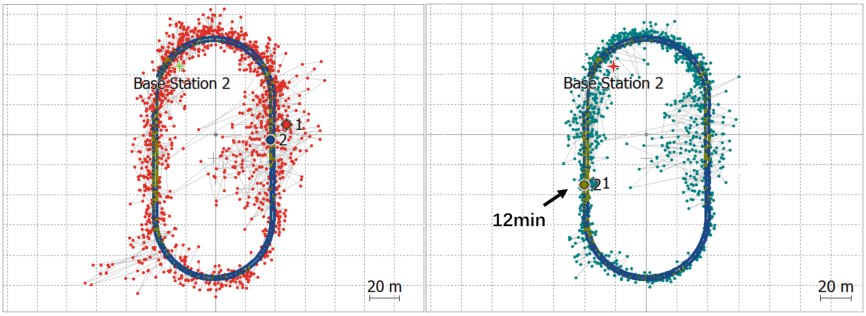


Fig. 9. Taking the positioning result of Trimble R10 GPS RTK as the datum. (a) is the HUAWEI P20Pro GPS pseudorange positioning result, and (b) is the positioning result using the improved PPP algorithm.

Table 2. RMS and STD of dynamic positioning results (m)

	N	E	U
Single RMS	8.135	10.004	25.080
Single-STD	8.133	9.797	24.449
PPP-RMS	5.844	7.293	17.704
PPP-STD	5.842	7.052	16.153

4 Summary

This paper analyses and evaluates the original GNSS data quality of smartphones from the aspects of visible satellite number, signal-to-noise ratio, pseudorange noise and Doppler velocimetry for Android N under dynamic conditions. Taking the Trimble R10 as a comparison, the static and dynamic positioning accuracy of the smartphone are analyzed, and a dynamic positioning algorithm based on Doppler velocity measurement and adaptive Kalman filtering is proposed.

Experiments show that the number of visible satellites of smartphone can meet the demand under dynamic conditions, and the spatial structure is better. However, compared with the current measurement receiver, the signal-to-noise ratio is too low, the pseudorange noise is too large, which mainly affects the accuracy of the positioning result, and the improved algorithm can improve the positioning accuracy and reliability.

Acknowledgement. This paper is funded by the National Natural Science Foundation of China (Fund No. 41574010, 41604013) and the National Key Laboratory of Geographic Information Engineering (ID: SKLGIE2015-Z-1-1).

References

1. Chai H, Pan Z, Cui Y (2016) Research progress on GNSS multi-system combination precision positioning. Ocean Mapp 36(4):21–26
2. Fu F (2016) Current status and development trend of mobile phone chip development. Eng Technol: Cit Version (7):00031
3. He H , Yang Y, Sun Z (2012) Comparative analysis of several GPS velocity measurement methods 31(3):217–221
4. Li X (2013) Research on GNSS precise point positioning and fast determination of non-difference ambiguity. Wuhan University
5. Gill M, Bisnath S, Aggrey J, Seepersad G (2017) Precise point positioning (PPP) using low-cost and ultra-low-cost GNSS receivers. In: Proceedings of ION GNSS 2017, Portland, Oregon, pp 85–100 September 2017
6. Navarro-Gallardo M, Bernhardt N, Kirchner M, Musial JR, Sunkevic M (2017) Assessing Galileo readiness in android devices using raw measurements. In: Proceedings of ION GNSS 2017, Portland, Oregon, pp. 85–100, September 201
7. de Bakker PF, van der Marel H, Tiberius CCJM (2009) Geometry-free undifferenced, single and double differenced analysis of single frequency GPS, EGNOS and GIOVE-A/B measurements. GPS Solutions 13(4):305–314
8. Yang Y (2006) Adaptive Dynamic Navigation and Positioning. Surveying and Mapping Press
9. Zhang X, Tao X, Zhu F et al (2018) Quality assessment of GNSS observations from an Android N smartphone and positioning performance analysis using time-differenced filtering approach. GPS Solutions 22(3):70

Performance Evaluation of Beidou RDSS Position Service

Dongxia Wang[1,2,3(✉)], Rui Guo[1(✉)], Jie Xin[1], Zhiqiao Chang[1], Xin Shi[1], and Tianqiao Zhang[1]

[1] 32021 Troops, Beijing 100094, China
wdx2008abc@163.com, shimbarsalon@163.com
[2] State Key Laboratory of Geodesy and Earth's Dynamics, Wuhan 430037, Hubei, China
[3] Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

Abstract. According to the rare existing researches of RDSS position service of satellite navigation system, this article evaluates RDSS position service performance using the time series analysis method based on Beidou original data. Firstly, we study the RDSS position principle and the evaluation method systematically. Moreover, based on placing the user terminal at different stations such as Beijing, Sanya, Shantou, Kashgar and Urumchi, we evaluated the position performance of different areas on the accuracy analysis of east-west, south-north and plane. The results provide the reference for the position consistency of Beidou system.

Keywords: Beidou · RDSS · Position service · Performance evaluation

1 Introduction

The Global Navigation Satellite Systems (GNSS), like Global Positioning System (GPS), Global Navigation Satellite System (GLONASS) and Galileo, just use Radio Navigation Satellite Service (RNSS) to achieve Position, Navigation and Timing (PNT) service [1, 2]. Therefore, the researches on RDSS position service performance evaluation is the unique direction of our country in the world [3, 4].

Beidou Satellite Navigation System (BDS) combines RDSS and RNSS to realize precise position, real-time navigation, precise timing, short message and position report service. Among them, RDSS service is an important component and distinct feature of Beidou system, and it is also the advantage and competitiveness of systems different from GPS, GLONASS and Galileo [5, 6].

RDSS position is an active position mode, which mainly relies on the pseudo-range measurement data (time difference observation can be transformed into pseudo-range observation) and user elevation data of the user to calculates the three-dimensional position according to the principle of spherical rendezvous [7–9]. At present, there are rare reports about the performance of Beidou RDSS position service. Therefore, this article studies the performance evaluation of RDSS position service, which can enrich

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 296–304, 2019.
https://doi.org/10.1007/978-981-13-7751-8_30

the accuracy evaluation technology of navigation system, monitor the performance timely and then improve the position accuracy of the Beidou system

2 RDSS Position Principle and Evaluation Method

2.1 RDSS Position Principle

RDSS position is an active position mode. Its observation amount is four-way ranging measurement. Firstly, the Main Control Center (MCC) transmits the outbound C-band signal, and then generates the outbound S-band signal through all GEO satellites, which can transmits two beam signals. RDSS user responds to the ranging signal of one beam according to its position, and then transmits inbound L-band signal and generates inbound C-band signal through GEO satellite. Finally, MCC responds to all inbound C-band signals transmitted by satellites, and completes the four-way ranging calculation. RDSS position principle can be shown in Fig. 1.

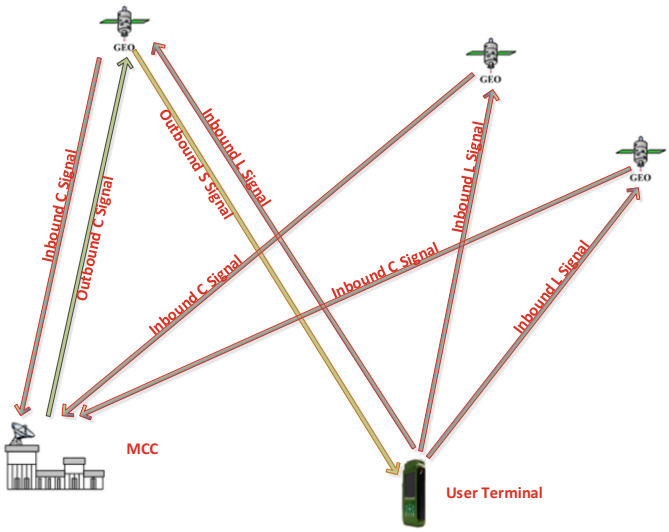


Fig. 1. RDSS position principle

In the process of actual measurement, the central station can calculate the departure time of the corresponding pseudo-range according to the response frame of user terminal. After the time iteration, two observations ρ_{Out_C} and ρ_{In_C} can be calculated accurately, and the following observation equation can be obtained.

$$\begin{cases} \rho' = \rho - \rho_{Out\_C} - \rho_{In\_C} \\ \rho' = \rho_{Out\_S} + \rho_{In\_L} = R_{OutS} + R_{InL} + \Delta D \\ \Delta D = \Delta D_{OutSTrop} + \Delta D_{OutSIono} + \Delta D_{OutSSagnac} + \Delta D_{OutSRel} \\ \quad + \Delta D_{OutSAnt} + \Delta D_{InLTrop} + \Delta D_{InLIono} + \Delta D_{InLSagnac} \\ \quad + \Delta D_{InLRel} + \Delta D_{InLAnt} \end{cases} \tag{1}$$

In Eq. (1), the satellite position is included in the variables R_{OutS} and R_{InL}. When the number of visible satellites is equal or more than 2, the position solution can be carried out under the constraint of known elevation. Thus, Eq. (1) can be linearized to obtain the following error equation:

$$\begin{cases} V_1 = (\frac{X_{OutSat}-X_0}{\rho_{OutSat0}} + \frac{X_{InSat1}-X_0}{\rho_{InSat10}})dX + (\frac{Y_{OutSat}-Y_0}{\rho_{OutSat0}} + \frac{Y_{InSat1}-Y_0}{\rho_{InSat10}})dY \\ \qquad + (\frac{Z_{OutSat}-Z_0}{\rho_{OutSat0}} + \frac{Z_{InSat1}-Z_0}{\rho_{InSat10}})dZ + l_{10} \\ V_2 = (\frac{X_{OutSat}-X_0}{\rho_{OutSat0}} + \frac{X_{InSat2}-X_0}{\rho_{InSat20}})dX + (\frac{Y_{OutSat}-Y_0}{\rho_{OutSat0}} + \frac{Y_{InSat2}-Y_0}{\rho_{InSat20}})dY \\ \qquad + (\frac{Z_{OutSat}-Z_0}{\rho_{OutSat0}} + \frac{Z_{InSat2}-Z_0}{\rho_{InSat20}})dZ + l_{20} \\ \cdots\cdots \quad \cdots\cdots \quad \cdots\cdots \quad \cdots\cdots \quad \cdots\cdots \\ V_m = (\frac{X_{OutSat}-X_0}{\rho_{OutSat0}} + \frac{X_{InSatm}-X_0}{\rho_{InSatm0}})dX + (\frac{Y_{OutSat}-Y_0}{\rho_{OutSat0}} + \frac{Y_{InSatm}-Y_0}{\rho_{InSatm0}})dY \\ \qquad + (\frac{Z_{OutSat}-Z_0}{\rho_{OutSat0}} + \frac{Z_{InSatm}-Z_0}{\rho_{InSatm0}})dZ + l_{m0} \end{cases} \tag{2}$$

$$\begin{cases} l_{10} = \rho'_1 - \rho_{OutSat0} - \rho_{InSat10} - \Delta D_{10} \\ l_{20} = \rho'_2 - \rho_{OutSat0} - \rho_{InSat20} - \Delta D_{20} \\ \cdots\cdots \quad \cdots\cdots \quad \cdots\cdots \\ l_{m0} = \rho'_2 - \rho_{OutSat0} - \rho_{InSatm0} - \Delta D_{m0} \end{cases} \tag{3}$$

where $[X_{OutSat}, Y_{OutSat}, Z_{OutSat}]$ and $[X_{InSatm}, Y_{InSatm}, Z_{InSatm}]$ are the positions of the outbound satellite and the m inbound satellite respectively, $[X_0, Y_0, Z_0]$ is the approximate position of the user terminal, $\rho_{OutSat0}$ is the distance between the outbound satellite and the user terminal, $\rho_{InSatm0}$ is the approximate distance between the m inbound satellite and the user terminal, and $[dX, dY, dZ]$ is the position corrections of the user terminal which need to be estimated.

Furthermore, during to GEO satellites located at the equator, the geometric observation configuration of the satellites is very poor, especially only two available GEO satellites. Therefore, position solution usually needs the support of the ground elevation database to be carried out. The corresponding error equation can be written as follows:

$$V_{m+1} = \cos(\varphi) \cdot \cos(\lambda) \cdot dX + \cos(\varphi) \cdot \sin(\lambda) dY + \sin(\varphi) \cdot dZ \tag{4}$$

2.2 Accuracy Evaluation Method

Time series analysis method [8] is to observe continuously in a certain time period, extract the relevant characteristics with time, analyse the characteristics and trends, and

then find out the rules of time-varying signal sequence to decompose periodic and random changes. In this paper, time series analysis method is used to analyse the position accuracy of different area to evaluate the RDSS position service perfomance.

The position performance analysis is based on the combination of region and position principle. The position accuracy is studied by placing the user terminals at Beijing, Sanya, Shantou, Kashgar and Urumchi. The specific evaluation and calculation method are based on the spherical geometry theory [10].

Assuming that $Lon1$ and $Lat1$ represent the exact value of longitude and latitude of user terminal placement, $Lon2$ and $Lat2$ represent calculate value of longitude and latitude of the user terminal. The difference of longitude can be calculated as $a = Lon1 - Lon2$, and the difference of latitude can be calculated as $b = Lat1 - Lat2$. According to the spherical geometry theory, the position accuracy of east-west direction, north-south direction and the plane can be obtained as:

$$\begin{cases} P_x = a * \mathrm{Re} * \cos(\mathrm{rad}(Lat1)) \\ P_y = b * \mathrm{Re} \\ P_s = 2\mathrm{Re} * \arcsin\left(sqrt\left(\begin{array}{l} \sin^2(b/2) + \cos(rad(Lat1)) \\ * \cos(rad(Lat2)) * \sin^2(a/2) \end{array} \right) \right) \end{cases} \tag{5}$$

where Re is the radius of the earth.

3 Test Results and Analysis

In order to evaluate and analyse the performance of RDSS position service in different areas, the static position test was conducted on 15 April 2017 using user terminals which are placed at the different stations. The point coordinates have been measured by GPS in advance, and then place the RDSS user terminals at the exact GPS points. The precise position point coordinates of RDSS user terminals are determined by using GPS data and IGS super-fast ephemeris. The system differences of user equipment and demodulation unit are allocated to RDSS precise position system respectively, and then the performances of RDSS position service are evaluated and obtained [11].

Combining with the ground station layout scheme of Beidou system, five stations including one main control station Beijing, two injection stations in Sanya and Kashgar, two first-class monitor stations in Shantou and Urumchi are selected. The position accuracy curves of each station are shown in Figs. 2, 3, 4, 5 and 6.

As we can see from Figs. 2, 3, 4, 5 and 6, the position accuracy of Sanya and Shantou is relatively large in the north-south direction, and the plane position accuracy is mainly affected by the north-south direction. Meanwhile, the position accuracy of Kashgar and Urumchi is relatively large in the east-west direction, and the plane position accuracy is mainly affected by the east-west direction. Moreover, the position accuracy of Beijing Station basically maintains in the east-west direction and the north-south direction, whose plane position accuracy is slightly affected by the north-south direction.

Statistical position accuracy on east-west, north-south and plane of Beijing, Sanya, Shantou, Kashgar and Urumchi stations are presented in Table 1.

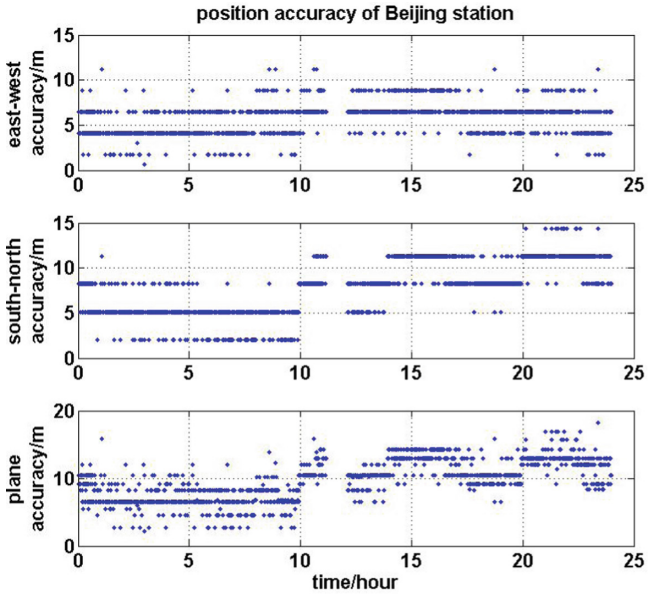


Fig. 2. Curve: position accuracy of Beijing station

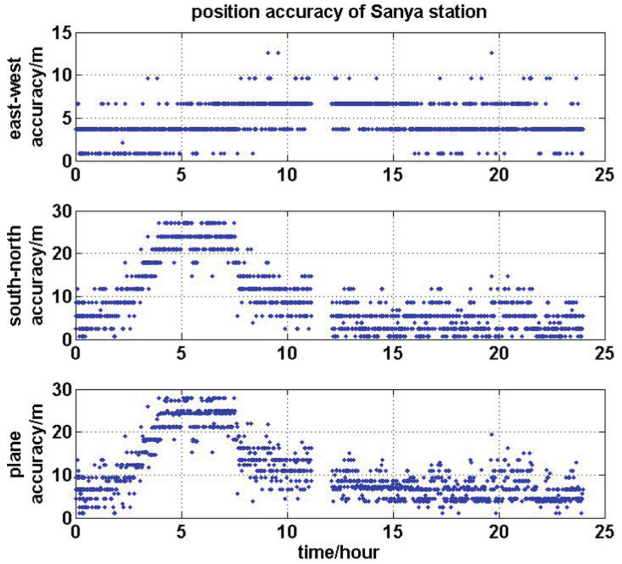


Fig. 3. Curve: position accuracy of Sanya station

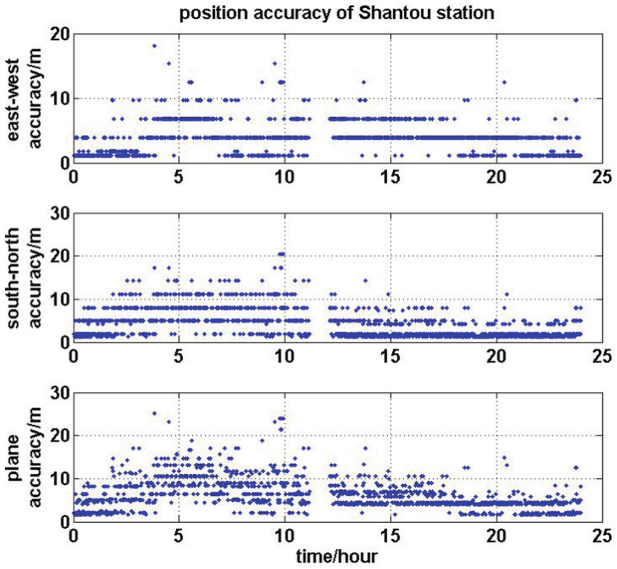


Fig. 4. Curve: position accuracy of Shantou station

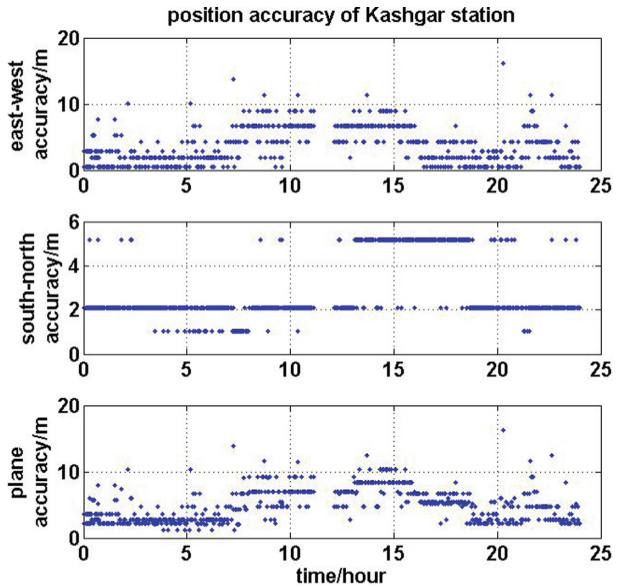


Fig. 5. Curve: position accuracy of Kashgar station

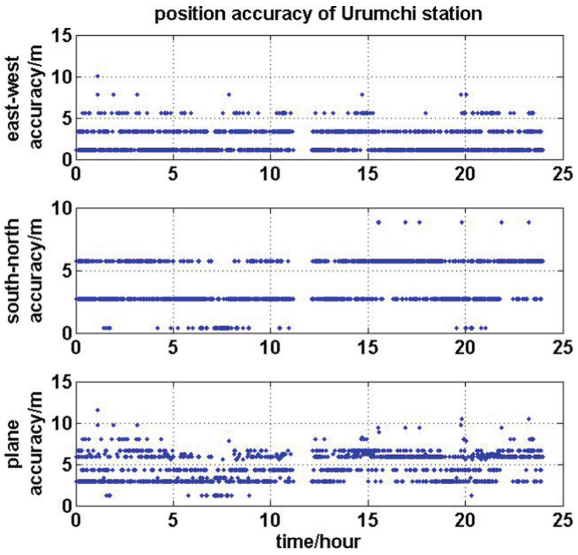


Fig. 6. Curve: position accuracy of Urumchi station

Table 1. Statistical position accuracy of stations (Unit: m)

Ground station	East-West position accuracy			North-South position accuracy			Plane position accuracy		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Beijing	0.67	11.16	5.67	2.03	14.40	7.64	2.14	18.22	9.66
Sanya	0.81	12.56	4.66	0.68	27.15	9.07	1.06	27.96	10.85
Shantou	1.09	18.13	3.84	1.24	20.41	4.37	1.65	25.07	6.20
Kashgar	0.54	16.17	3.60	1.01	5.17	2.78	1.15	16.30	4.89
Urumchi	1.11	10.01	2.22	0.42	8.86	3.85	1.18	11.55	4.71

Analysing the position accuracy of Beijing, Sanya, Shantou, Kashgar and Urumchi stations, the following conclusions can be obtained:

(1) The position accuracy of East-West direction from high to low is as follows: Urumchi-Kashgar-Shantou-Sanya-Beijing. The position accuracy of North-South direction from high to low is as follows: Kashgar-Urumchi-Shantou-Beijing-Sanya. The position accuracy of plane direction from high to low is as follows: Urumchi-Kashgar-Shantou-Beijing-Sanya.

(2) The position accuracy of Urumchi and Kashgar is better, and the ones of Beijing and Sanya is not very high, which is mainly related to the geographic environmental factors. Moreover, Sanya Station is near to the equator where ionospheric activity is intense.

Therefore, Beidou RDSS position service still has space to improve accuracy and stability. As we can see from the position results, there are two problems that must be resolved in future:

(1) The position results of the same users in the same areas have deviations at different times of the day. This situation is usually related to the satellite ephemeris delay and ionospheric error, which is more obvious in the South than in the north.
(2) The position results of the same users in different areas have deviations, which is closely related to the differences of zero-value deviation of inbound and outbound stations, which means position consistency of different areas should be strengthened.

Therefore, the position accuracy of Beidou RDSS can be further improved by RDSS ranging accuracy, the orbit determination accuracy of GEO satellite and the correction accuracy of ionospheric delay.

4 Conclusions

With the development and application of Beidou globalization, RDSS service performance has attracted more and more attention. Therefore, this article studies the RDSS position principle and evaluation method. Combing the operation status of Beidou global system, we evaluate the position service performance of Beijing, Sanya, Shantou, Kashgar and Urumchi based on the position accuracy of East-West direction, north-south direction and plane direction. It provides further theoretical researches in RDSS ranging, GEO satellite orbit determination and ionospheric delay correction.

Acknowledgements. Thanks for the fund support of the National Nature Science Foundation of China (Nos. 61603397, 41874043, 41704037) and the State Key Laboratory of Geodesy and Earth's Dynamics Foundation (Institute of Geodesy and Geophysics, CAS) (No. SKLGED2017-3-3-E).

References

1. Tan S (2017) Innovative development and forecast of BeiDou system. Acta Geodaetica et Cartographica Sinica 46(10):1284–1289
2. Tan S (2009) Theory and application of comprehensive RDSS position and report. Acta Geodaetica et Cartographica Sinica 38(1):1–5
3. Hu Z (2013) Performance evaluation theory and test verification of Beidou satellite navigation system. Wuhan University
4. Han X, Ma X (2015) Positioning performance assessment of BeiDou satellite navigation system. Geotech Investig Surv 4:83–87
5. Zhao J, Qu J, Yuan H (2016) A new ambiguity resolution method using combined RNSS-RDSS of BeiDou. Acta Geodaetica et Cartographica Sinica 45(4):404–410
6. Dou C, Zhang B, Tan S (2010) A new three-satellite high-precision RDSS/RNSS combination positioning method. In: International Workshop on Education Technology & Computer Science, vol 2, pp 288–290

7. Li Q, Liu Z, Xue Y, Bai Z (2006) Modeling and simulation in RDSS system. J Syst Simul 18 (5):1199–1203
8. Liu L, Han C (2004) Two way satellite time transfer and its error analysis. Prog Astron (3):219–226
9. Xu L (2007) BD-2 RDSS subsystem navigation error modification technology study. Chinese Academy of Sciences, Beijing
10. Wu Y (2004) Basic geometry. People's Education Press
11. Tang W, Cui J, Hui M (2016) Analysis of the impact of Beidou regional constellation on relative positioning accuracy. Geomat Inf Sci Wuhan Univ 41(8):1107–1112

The Effect of DCB Correction on Multi-system Combination Precise Point Positioning

Shiming Gu[1,2(✉)], Yamin Dang[2], Hu Wang[2], Jian Wang[1,2], Zhengzhao Ren[1,2], and Jinxu Zhang[1,2]

[1] Shandong University of Science and Technology, Qingdao, Shandong, China
475172140@qq.com
[2] Chinese Academy of Surveying and Mapping, Beijing, China

Abstract. The satellite differential code bias (DCB) characterizes the time it takes for the signal to travel from the reference clock to the transmitting antenna. The internal delay of the signal can be up to ten nanoseconds, and the pseudorange can be up to the meter level. It must be corrected when high-precision point positioning. This paper analyzes the DCB correction formulas of different pseudorange combinations of GPS and BDS. Using GPS, GPS + BDS, GPS + GLAONASS + BDS and iGMAS released DCB products, standard single point positioning (SPP) and precise point positioning (PPP) solutions are performed for multiple stations in China. The results show that after DCB correction, the SPP accuracy is improved in decimeters, the average improvement rate of the three schemes is about 30%. The static PPP accuracy has a millimeter-level improvement and the accuracy is improved by about 10%. DCB correction has the greatest impact on the positioning of single GPS system. However, when multi-system combination positioning is performed, the influence of DCB correction on positioning accuracy is weakened, the improvement rate is greatly reduced, and the coordinate precision improvement of some stations is sub-millimeter.

Keywords: GNSS · DCB · Single point positioning

1 Introduction

Multi-gnss joint positioning can utilize more satellite and multi-frequency signals, effectively reduce the position accuracy factor PDOP value, provide more accurate navigation and positioning services, and ensure the reliability of positioning [1]. The internal delay of the signal characterizes the time it takes for the satellite signal to generate the reference signal from the satellite clock, and then generate the ranging code, the navigation message and the carrier, perform signal modulation, and until modulated signal finally leaves the satellite transit antenna (phase center). The internal delays of signals between different frequency points and even different codes at the same frequency point are different, and their common parts will be automatically absorbed into the satellite clock offsets, so we only need to compare the difference between them, and the differential code bias (DCB) is processed [2]. For the receiver-side DCB, in the single-mode single-point positioning, the internal delay of each

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 305–313, 2019.
https://doi.org/10.1007/978-981-13-7751-8_31

satellite signal at the receiver end is the same, which can be included in the receiver clock offsets to estimate, ignoring its influence on the positioning accuracy [3]. At the same time, because the influence of satellite DCB on the positioning accuracy can not be eliminated, and its value can be up to ten nanoseconds, the influence on the signal propagation distance can reach the meter level [4], so it must be corrected for high-precision navigation positioning. In this paper, the national CORS station data is used to analyze the influence of DCB correction on the positioning accuracy of multi-system combination. The DCB correction equations of BDS and GPS different code combinations are given, and the influence of the bias correction on the single point positioning accuracy is analyzed by using the measured data.

2 Multi-GNSS Joint Single Point Positioning

Precise single-point positioning uses dual-frequency pseudorange and phase observations, as well as precise ephemeris and precision clock offset for positioning processing. The observation equation is

$$\left.\begin{aligned} PC_i^j &= \rho_i^j + c\left(\delta t_i - \delta t_j\right) + \delta_{trop}^j + \varepsilon_{pc} \\ LC_i^j &= \rho_i^j + c\left(\delta t_i - \delta t_j\right) + \delta_{trop}^j + \lambda N_i^j + \varepsilon_{lc} \end{aligned}\right\} \tag{1}$$

where, PC_i^j and LC_i^j represent the observations after combining the dual-frequency pseudorange and the dual-frequency carrier, respectively, $PC = \left(f_1^2 P_1 - f_2^2 P_2\right)/\left(f_1^2 - f_2^2\right)$, $LC = \left(f_1^2 \lambda_1 L_1 - f_2^2 \lambda_2 L_2\right)/\left(f_1^2 - f_2^2\right)$, f and λ represent the carrier frequency and the combined wavelength, $\lambda = \mathrm{c}/\left(f_1^2 - f_2^2\right)$, c represents the speed of light under vacuum, i and j represent the station and the satellite, ρ represents the geometric distance between the station and the satellite, δt_i and δt_j represent receiver clock offset and satellite clock offset, δ_{trop} denotes tropospheric delay, N is float phase ambiguity, $N = f_1 N_1 - f_2 N_2$, after combination, the ambiguity parameter no longer has integer characteristics, ε_{pc} and ε_{lc} represent the residual after the combination of pseudorange and carrier phase.

At the same time, considering the inter-system deviation and the inter-frequency deviation of Multi-GNSS joint positioning, the observation equation is

$$\left.\begin{aligned} PC_i^{j,G} &= \rho_i^j + c\left(\delta t_i - \delta t_j\right) + \delta_{trop}^{j,G} + \varepsilon_{pc}^G \\ LC_i^{j,G} &= \rho_i^j + c\left(\delta t_i - \delta t_j\right) + \delta_{trop}^{j,G} + \lambda N_i^{j,G} + \varepsilon_{lc}^G \\ PC_i^{j,R_k} &= \rho_i^j + c\left(\delta t_i - \delta t_j + ISB_i^{j,R_k}\right) + \delta_{trop}^{j,R_k} + \varepsilon_{pc}^R \\ LC_i^{j,R_k} &= \rho_i^j + c\left(\delta t_i - \delta t_j + ISB_i^{j,R_k}\right) + \delta_{trop}^{j,R_k} + \lambda N_i^{j,R} + \varepsilon_{lc}^R \\ PC_i^{j,C} &= \rho_i^j + c\left(\delta t_i - \delta t_j + ISB_i^C\right) + \delta_{trop}^{j,C} + \varepsilon_{pc}^C \\ LC_i^{j,C} &= \rho_i^j + c\left(\delta t_i - \delta t_j + ISB_i^C\right) + \delta_{trop}^{j,C} + \lambda N_i^{j,C} + \varepsilon_{lc}^C \end{aligned}\right\} \tag{2}$$

Where G, R and C represent GPS, GLONASS and BDS system, ISB_i^{j,R_k}, ISB_i^C indicates the inter-system deviation of GLONASS, BDS relative to GPS. Since GLONASS adopts Frequency Division Multiple Access (FDMA) technology, the delay between the receiver and each satellite signal channel is different, ISB_i^{j,R_k} is related to the station and the satellite. BDS and GPS adopt Code Division Multiple Access (CDMA) technology, that is, the satellite signals of the same satellite system use the same carrier frequency, ISB_i^C is only related to the receiver. Significant deviations are mainly reflected in the firmware differences in receiver signal processing [5]. In addition, CDMA technology will also be adopted in the new generation of GLONASS satellites [6]. The remaining symbols in the formula (2) represent the same meanings as in the formula (1).

3 DCB Correction

The satellite clock offset of the GPS system is obtained by the combination of P1 and P2 dual-frequency ionosphere-free (IF). The time reference is the combined electronic phase center of the L1 and L2 frequency points. The time reference of the clock offset parameter broadcasted by the BDS system is the antenna phase center of B3 frequency point. And the precision satellite clock error is solved by the combination of the B1 and B2 dual-frequency ionosphere-free [7]. Therefore, when positioning with different code combinations, the time must be unified to the same reference point, and the differential code bias (DCB) correction must be performed. In the GPS system and the GLONASS system, the P1 and P2 dual-frequency pseudorange ionosphere-free combination equation is

$$PC = \rho + c(\delta t_r - \delta t_s) + \delta_{trop} + + \frac{c\left(\left(f_1^2\tau_{p1} - f_2^2\tau_{p2}\right)\right)}{\left(f_1^2 - f_2^2\right)} \tag{3}$$

where τ_{P1} and τ_{P2} are signal internal delays of pseudorange P1 and P2.

The satellite clock offset parameter broadcasted by the broadcast ephemeris is

$$\delta t_{sv} = \delta t_s - \left(f_1^2\tau_{p1} - f_2^2\tau_{p2}\right)/\left(f_1^2 - f_2^2\right) = \left(\rho + \delta t_r + \delta_{trop} - PC\right)/c \tag{4}$$

Since the satellite clock offset parameter broadcasted by the broadcast ephemeris includes the device delay of P1 and P2, DCB correction is not considered when positioning with the P1 and P2 dual-frequency ionosphere-free combination. When the combination of C1 and P2 dual-frequency ionosphere-free is used as the observation equation, the satellite clock offset is calculated as

$$\delta t_{sv(c1/p2)} = \delta t_s - \frac{f_1^2\tau_{c1} - f_2^2\tau_{p2}}{f_1^2 - f_2^2} \tag{5}$$

Then, calculate it to the time reference of the P1 and P2 dual-frequency ionosphere-free combination

$$\delta t_{sv(c1/p2)} = \delta t_{sv} + \frac{f_1^2\left(\tau_{p1} - \tau_{c1}\right)}{f_1^2 - f_2^2} \tag{6}$$

Therefore, when using C1 and P2 ionosphere-free combination for positioning, DCB (P1-C1) correction must be performed. Using the same calculation method, when using P1 and C2 ionosphere-free combination for positioning, the clock offset is calculated to the same time reference

$$\delta t_{sv(p1/c2)} = \delta t_{sv} + \frac{f_1^2\left(\tau_{c2} - \tau_{p2}\right)}{f_1^2 - f_2^2} \tag{7}$$

Therefore, when using the P1 and C2 pseudoranges to form the ionosphere-free combination, it is necessary to consider the influence of the DCB (P2-C2) on the positioning result.

For the BDS system, the broadcast ephemeris broadcasts two device delay parameters, TGD1 and TGD2.

$$\left.\begin{array}{l} TGD1 = \tau_{B1} - \tau_{B3} \\ TGD2 = \tau_{B2} - \tau_{B3} \end{array}\right\} \tag{8}$$

Because the time reference of the BDS system is the B3 frequency point, the clock offset parameter of the broadcast ephemeris is

$$\delta t_{sv} = \delta t_s - \tau_{B3} \tag{9}$$

Referring to Eq. (5), when using the B1 and B3 dual-frequency pseudorange ionosphere-free combination, the satellite clock error is

$$\delta t_{sv(B1/B3)} = \delta t_s + \frac{f_1^2(\tau_{B3} - \tau_{B1})}{f_1^2 - f_3^2} \tag{10}$$

Unified calculate it to the point time reference of B3 frequency

$$\delta t_{sv(B1/B3)} = \delta t_{sv} + \frac{f_1^2(\tau_{B3} - \tau_{B1})}{f_1^2 - f_3^2} \tag{11}$$

Therefore, when using the B1 and B3 dual-frequency pseudorange ionosphere-free combination as the observation equation, the TDG1 must be used to correct the pseudorange observations. Similarly, when using B1 and B2 dual-frequency pseudoranges for positioning, the equation for calculating the satellite clock offset to the time reference of B3 frequency point is

$$\delta t_{sv(B1/B2)} = \delta t_{sv} + \frac{f_1^2(\tau_{B3} - \tau_{B1}) + f_2^2(\tau_{B2} - \tau_{B3})}{f_1^2 - f_2^2} \tag{12}$$

Therefore, when using the B1 and B2 dual-frequency pseudorange ionosphere-free combination as the observation equation, the pseudorange observations should be corrected by TGD1 and TGD2. Since the B2 and B3 dual-frequency combination make noise amplification can not be located [3], the calculation model of B2 and B3 dual-frequency combination is no longer considered.

4 Single Point Positioning Experiment

Using the 24 h static observation data collected by the national stations JLPS, JXYS, QHTG and SDCW on January 5, 2018, the sampling interval is 30 s, and the experimental analysis is carried out. The above four national stations all use the TRIMBLE NETR9 receiver. This type of receiver is a non-crossing receiver and can only receive GPS C1 and P2 pseudoranges [8]. According to formula (7), it must be DCB (P1-C1) correction. The effects of DCB (P1-C1) on single point positioning in SPP and PPP modes are compared by using G, G + C, G + R + C system.

4.1 SPP Experiment

Pseudo-range single point positioning is performed using the observation data of the above four national stations. The DCB file is a products advertised by the iGMAS Analysis Center. Since the high precision coordinates of the national station are not published, the coordinates of the GAMIT solved are used as the true values. Compare the coordinate difference in three directions of N, E and U after DCB(P1-C1) correction. Taking the coordinate residuals of the JLPS in the N, E, and U directions as an example, and select first 1000 epochs for analysis, as shown in Figs. 1, 2, and 3. Table 1 shows the RMS values of coordinate residuals of the four stations in the N, E, and U directions.

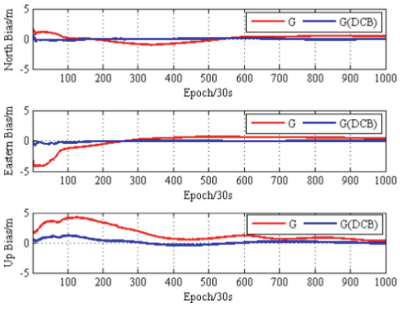


Fig. 1. Coordinate residuals solved by G system

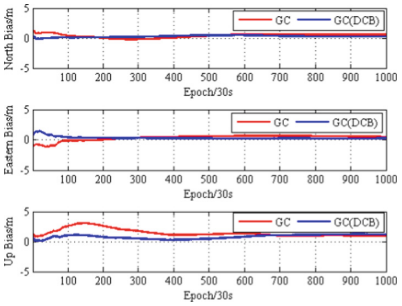


Fig. 2. Coordinate residuals solved by GC system

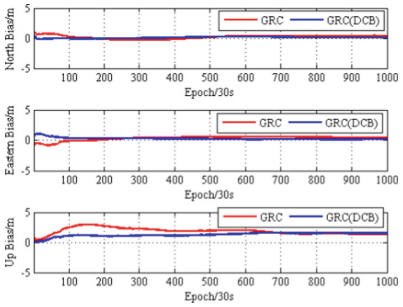


Fig. 3. Coordinate residuals solved by GRC system

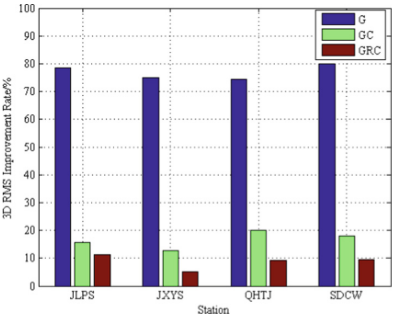


Fig. 4. 3D coordinate improvement rate

(1) Figures 1, 2 and 3 shows that for SPP, after DCB correction, the coordinates of the station change gently in the three directions of N, E, and U, and basically maintained near the 0 value.

(2) Table 1 shows that after DCB correction, the accuracy of SPP has a decimetre level of improvement, the average improvement rate in the N direction is 56.72%, the average improvement rate in the E direction is 44.93%, and the average improvement is 32.29% in the U direction. DCB correction has a greater impact on SPP.

(3) Combined with the analysis of Fig. 4 and Table 1, it can be seen that for single point positioning of single G system, the three-dimensional coordinate improvement rate is 76.92% after DCB correction, the improvement rate for G + C combined system is 16.65%, and the three-dimensional coordinate improvement rate is 8.81% for G + R + C. For multi-gnss joint single-point positioning, the accuracy is improved after DCB correction, but with the increase of the positioning system, the improvement rate of three-dimensional coordinate precision is greatly reduced.

Table 1. The effect of DCB correction for SPP

Station	DCB Correction	G			G + C			G + R + C		
		N/m	E/m	U/m	N/m	E/m	U/m	N/m	E/m	U/m
JLPS	Before	0.3917	0.7147	1.1589	0.3816	0.4622	1.4966	0.2458	0.3993	1.8446
	After	0.0944	0.0981	0.2710	0.2177	0.2797	1.3142	0.0908	0.2342	1.6692
JXYS	Before	0.4726	0.7179	0.8392	0.3337	0.4436	0.6730	0.2194	0.4623	0.9806
	After	0.0784	0.1811	0.2275	0.1790	0.3889	0.6288	0.1345	0.4251	0.9487
QHTG	Before	0.6212	0.4733	1.4052	0.3324	0.5799	1.6566	0.3248	0.7085	2.2711
	After	0.1367	0.0944	0.3788	0.2318	0.3884	1.3533	0.1714	0.5298	2.1053
SDCW	Before	0.4116	0.7399	1.5880	0.3113	0.4098	1.3772	0.2027	0.3965	1.6699
	After	0.0916	0.2476	0.2477	0.1910	0.2556	1.1609	0.0843	0.2600	1.5422

4.2 PPP Experiment

Select the 24 h static observation data of the above four national stations, use the 30 s clock offset products broadcasted by gbm, and the precision ephemeris with sampling interval of 5 min to carry out static precision single point positioning experiment. Take JLPS station as an example, select the first 200 epochs to analysis. Figures 5, 6, and 7 show use G, G + C, G + R + C three combination schemes respectively, and compare the residuals of coordinates after DCB correction in the three directions of N, E, and U.

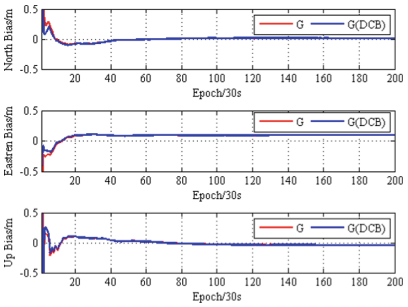


Fig. 5. Coordinate residuals solved by G system

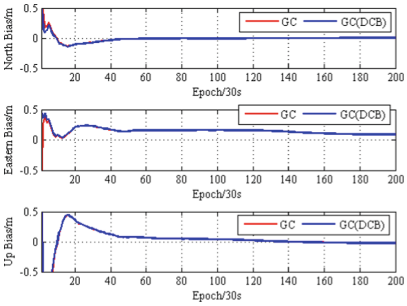


Fig. 6. Coordinate residuals solved by GC system

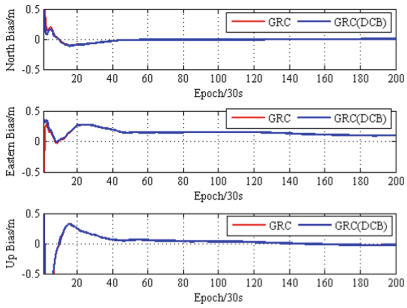


Fig. 7. Coordinate residuals solved by GRC system

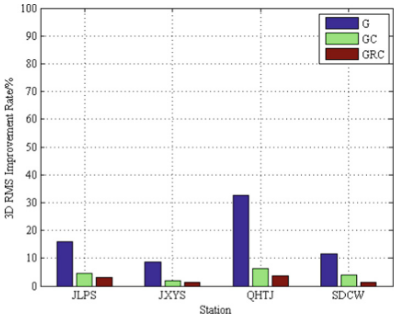


Fig. 8. 3D coordinate improvement rate

(1) Figures 5, 6, and 7 show the DCB correction improves the coordinate accuracy of the precise single point positioning of G, G + C, G + C + R, which can accelerate the coordinate convergence. The improvement of the single G system is more obvious, but the accuracy improvement of the G + C and G + R + C combined systems is limited.

(2) Table 2 shows the accuracy of PPP has a millimetre level improvement in the three directions of N, E, and U after DCB correction, the average improvement

rate in the N direction is 17.67%, the average improvement rate in the E direction is 2.80%, and the average improvement rate in the U direction is 17.65%.

(3) According to Fig. 8 and Table 2, for the single G system, the accuracy improvement of three-dimensional coordinates is centimeter level, and the improvement rate is 17.21%. For multi-gnss joint system G + C, G + R + C, the accuracy improvement is millimeter level. And the improvement rates is 4.19% and 2.31%, respectively.

Table 2. The effect of DCB correction for PPP

Station	DCB Correction	G			G + C			G + R + C		
		N/m	E/m	U/m	N/m	E/m	U/m	N/m	E/m	U/m
JLPS	Before	0.0443	0.1303	0.0417	0.0318	0.1101	0.0799	0.0283	0.1134	0.0658
	After	0.0274	0.1141	0.0285	0.0212	0.1082	0.0752	0.0192	0.1122	0.0628
JXYS	Before	0.0381	0.1508	0.0541	0.0372	0.1334	0.1181	0.0256	0.1360	0.0988
	After	0.0184	0.1432	0.0416	0.0269	0.1322	0.1169	0.0206	0.1356	0.0965
QHTG	Before	0.0264	0.1222	0.1526	0.0411	0.1272	0.0783	0.0342	0.1289	0.0539
	After	0.0264	0.1167	0.0582	0.0404	0.1256	0.0605	0.0321	0.1275	0.0441
SDCW	Before	0.0273	0.1335	0.1069	0.0284	0.1210	0.0839	0.0220	0.1245	0.0685
	After	0.0273	0.1288	0.0783	0.0263	0.1196	0.0754	0.0233	0.1237	0.0659

5 Conclusion

Since the DCB of the satellite can be up to ten nanoseconds, the influence on the pseudorange can reach the meter level. Therefore, the influence of the satellite DCB on the precise single point positioning and the pseudorange single point positioning is analyzed. The three system combinations of G, G + C and G + R + C are designed, and compared the impact of DCB correction on the accuracy improvement of single point positioning. The conclusion is as follows.

(1) For SPP and static PPP, the coordinate accuracy is improved after DCB correction. Among them, the SPP has the greatest improvement and the accuracy improvement is decimeter level. The three-dimensional coordinate accuracy improvement rate of the three schemes is 8% to 77%, the impact on static PPP is relatively small, and the accuracy improvement is millimeter level, and the accuracy is improved by 2% to 18%. Therefore, in the single point positioning, the DCB correction of the satellite side must be considered to improve the positioning accuracy.

(2) DCB correction has the most obvious impact on the positioning of single G system. The three-dimensional coordinate accuracy improvement rates of SPP and PPP are 76.92% and 17.21%, respectively. When multi-gnss joint positioning is performed, Due to the increase in the number of visible satellites and the enhancement of satellite space geometry, the improvement rate of DCB correction on positioning accuracy has been greatly reduced. The accuracy improvement rate of SPP and PPP of G + R + C is 8.81% and 2.31%, and the effect of coordinate

accuracy of some stations is only sub-millimeter level, so multi-gnss joint positioning can reduce the impact of satellite DCB on the accuracy of coordinate calculation.

Acknowledgments. This work is supported by the General Program of National Natural Science Foundation of China (41874042), the National Key R&D Program of China (2016YFB0501405) and the China Postdoctoral Science Foundation (2016M590715).

References

1. Huang G, Lei Z, Wang J, Zhao L (2018) Multi-system combined PPP positioning efficiency analysis under different occlusion scenarios in Asia-Pacific region. J Geodesy Geodyn 38 (06):562–567
2. Li Z, Gong X (2012) New developments in global positioning system (lecture 4) - internal delay of signal and its influence on clock error. J Geomat 37(04):51–54
3. Zeng T, Sui L, Bao Y, Xiao G, Dai Q, Tian Y, Zhang Q (2017) The influence of BDS satellite terminal differential code deviation on positioning and correction model. J Geodesy Geodyn 37(01):53–57
4. Montenbruck O, Hauschild A (2013) Code biases in multi-GNSS point positioning. In: Proceedings of the 2013 International Technical Meeting of the Institute of Navigation, San Diego, California
5. Ren X, Zhang K, Li X, Zhang X (2015) BeiDou, Galileo, GLONASS, GPS multi-system fusion precision single point positioning. Acta Geodaetica et Cartographica Sinica 44 (12):1307–1313+1339
6. Seepersad G, Bisnath S (2015) Reduction of PPP convergence period through pseudorange multipath and noise mitigation. GPS Solutions 19(3):369–379
7. Dai W, Jiao W, Jia X (2009) Application research for compass navigation satellite interfrequency bias correction terms. J Geomat Sci Technol 26(05):367–369+374
8. Zhao L, Zhang S (2016) Analysis of the influence of GPS satellite P1C1 code deviation on dynamic PPP. J Geodesy Geodyn 36(02):143–145

Estimation Method for Position and Posture of Mobile Carrier Based on Multiple Laser Trackers

Haolong Luo(✉), Zhen Yang, Shiyan Wang, and Chenyu Wang

Strategic Support Force Information Engineering University, Zhengzhou, China
1728731316@qq.com

Abstract. Nowadays, the position and posture of moving objects are playing more and more important roles in space flight, navigation, target tracking and so on. It is crucial to obtain high-precision position and posture of the carriers. Using multiple laser trackers, high-precision position and posture of the carriers can be obtained, but there are still some errors due to the influence of noise and other factors. In this paper, a current statistical model is built, the Kalman filter adaptive algorithm is designed to filter the sampling data of the three laser trackers, removing the influence of noise and other factors on the laser tracking system and obtaining more accurate sampling coordinates. Besides, the three-spline interpolation method is used to obtain the position coordinates of the carrier at any time. Finally, the high-precision position and posture of the carrier at any time are obtained through the calculation of position and posture.

Keywords: Kalman filter · Laser tracker · Position and posture · Current statistics

1 Introduction

Nowadays, the position and posture of moving targets are playing more and more important roles in aerospace, navigation, fault diagnosis and target tracking. By acquiring the position and posture of the carrier, we can know the working state and safety status of the carrier, and ensure the normal operation of the carrier. In the field of surveying and mapping navigation, position and posture measurement mainly depends on different kinds of sensors, such as posture measurement based on inertial sensors, GNSS multi-antenna method and the methods of posture measurement of three-axis turntable [1]. Inertial sensors integrate the angular velocity information and acceleration information sensitive to gyroscope and accelerometer based on the initial position and posture information, and then calculate the corresponding position and posture information. However, the speed of initial convergence is slow, existing error accumulation, and the cost of high-precision inertial sensors is high. The measurement of Multi-antenna GNSS posture is based on the carrier phase of the antenna to obtain the

Fund projects: National Key R&D Plan (2017YFF0206000); National Natural Fund projects (No. 41501491).

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 314–325, 2019.
https://doi.org/10.1007/978-981-13-7751-8_32

posture and heading information of the carrier, but the GNSS signal is greatly affected by climate and other factors. The three-axis turntable measures three posture parameters directly through vertical axis, horizontal axis and pitch axis, but the disadvantage is that the accuracy of posture measurement is easily affected by the manufacturing accuracy of the turntable. Laser tracking measurement system has large ranging range, high ranging accuracy and good dynamic performance. Compared with other measurement systems, it has obvious advantages.

At present, the moving speed of maneuvering targets such as airplanes and automobiles is fast, and the position and posture change rapidly. It is necessary to track the target uninterruptedly and grasp the position and posture of the target in real time [2]. Laser tracker can track the target in real time, its measurement range can reach 120 m, the measurement accuracy is um level, and it is less disturbed by the outside. It can meet the requirement of measuring the position and posture of the moving carrier stably and accurately. However, due to the interference of noise factors, there are still some errors in the sampling data of laser tracker. The errors of Point position have great influence on the posture calculation of the carrier when tracking the motion of the carrier over a long distance. However, the original data of laser tracker are directly used at present. How to eliminate the influence of noise on laser tracking system hasn't been studied. In order to eliminate the influence of noise and other factors on the laser tracking system and obtain higher precision, Kalman filter algorithm is used to filter the sampled data.

2 Noise Removal Using Adaptive Kalman Filter

The sampling data of laser tracker is affected by noise and other interference factors. Kalman filter algorithm can remove the influence of noise and other interference factors, and obtain more accurate posture of the carrier.

2.1 Kalman Filtering

Kalman filter is a recursive and optimal estimation theory. It estimates the state of the target based on the current observations and the last moment estimates. It constantly updates the state of objects to correct the state of the target and estimate the optimal solution of the target. Kalman filter is widely used in signal processing, target tracking, petroleum exploration, satellite positioning, aerospace and biomedicine [3].

2.2 Motion Model of the Target

There are many kinds of models describing the motion of the target, such as uniform velocity model and uniform acceleration model, but the moving direction and speed of the target are generally irregular. It is obviously unreasonable to describe the motion state of the target with the above model [4]. Singer model has strong anti-jamming ability and can simulate carrier motion better. However, Singer model is not enough because it is a zero-mean model of acceleration, and there are still large errors in simulating target motion with rapid acceleration change. The current statistical model is a mean model with non-zero acceleration. It considers that there is a certain relationship

between the acceleration changes of the carrier at the adjacent time, and its probability distribution is determined by the modified Rayleigh distribution [5]. The current statistical model is more accurate and reliable in simulating carrier motion.

The equation of state for continuous systems under the "current" statistical model is as follows:

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \\ \dddot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\alpha \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \alpha \end{bmatrix} \bar{\alpha}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \omega(t) \tag{1}$$

Assuming that the sampling period of the laser tracker is T, we can obtain the discrete state equation:

$$X(k+1) = F(k)X(k) + U(k)\bar{\alpha} + W(k) \tag{2}$$

The matrix of state transition is shown below.

$$F = \begin{bmatrix} 1 & T & \frac{1}{\alpha^2}(-1 + \alpha T + e^{-\alpha T}) \\ 0 & 1 & \frac{1}{\alpha}(1 - e^{-\alpha T}) \\ 0 & 0 & e^{-\alpha T} \end{bmatrix} \tag{3}$$

The input matrix is shown below.

$$U(k) = \begin{bmatrix} \frac{1}{\alpha}(-T + \frac{\alpha T^2}{2} + \frac{1-e^{-\alpha T}}{\alpha^2}) \\ T - \frac{1}{\alpha}(1 - e^{-\alpha T}) \\ 1 - e^{-\alpha T} \end{bmatrix} \tag{4}$$

W(k) is a sequence of discrete white noise, and its variance is shown below [6].

$$Q(k) = E\left[W(k)W^T(k)\right] = 2\alpha\sigma_\alpha^2 \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \tag{5}$$

2.3 Adaptive Filtering Algorithm Based on "Current" Statistical Model

The adaptive Kalman filter algorithm under the "current" statistical model is as follows:

$$X(k/k) = \hat{X}(k/k-1) + K(k)[Y(k) - H(k)\hat{X}(k/k-1)] \tag{6}$$

$$\hat{X}(k/k-1) = F(k,k-1)\hat{X}(k-1/k-1) + U(k)\bar{\alpha}(k) \tag{7}$$

$$K(k) = P(k/k-1)H^T[H(k)P(k/k-1)H^T(k) + R(k)]^{-1} \tag{8}$$

$$P(k/k-1) = F(k,k-1)P(k-1/k-1)F^T(k,k-1) + Q(k-1) \tag{9}$$

$$P(k/k) = [1 - K(k)H(k)]P(k/k-1) \tag{10}$$

When the "current" acceleration is positive:

$$\sigma_\alpha^2(k-1) = \frac{4-\pi}{\pi}[\alpha_{\max} - \hat{\ddot{x}}(k-1/k-1)] \tag{11}$$

When the "current" acceleration is negative:

$$\sigma_\alpha^2(k-1) = \frac{4-\pi}{\pi}[\alpha_{-\max} - \hat{\ddot{x}}(k-1/k-1)] \tag{12}$$

2.4 Computer Simulation

The filtering effect of the adaptive Kalman filter algorithm is closely related to the assumed maximum acceleration of the target and the mean square error R of the measured noise [7]. Through relevant experiments, the reasonable selection of R is 1 mm^2, and the frequency of moving target is 0.05 [8]. Sampling data with sampling frequencies of 10 Hz, 50 Hz and 100 Hz are filtered respectively. The three-axis errors of the filtered moving target are as follows (Figs. 1, 2 and 3):

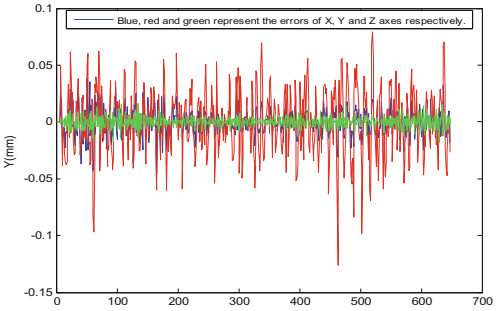


Fig. 1. The errors of three axes when frequency is 10 Hz

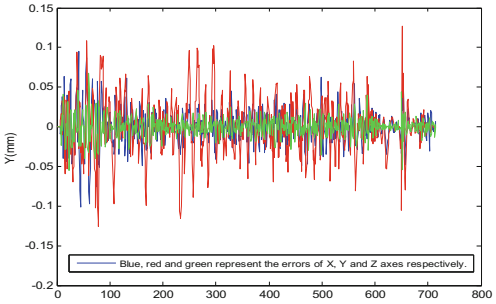


Fig. 2. The errors of three axes when frequency is 50 Hz

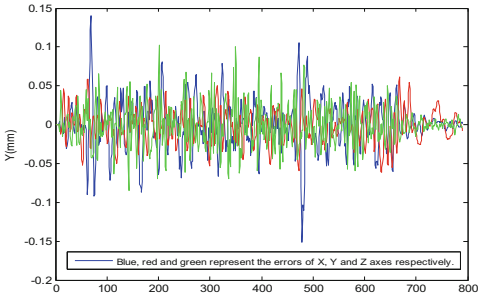


Fig. 3. The errors of three axes when frequency is 100 Hz

It is concluded from the graph that:

1. In 10 Hz and 50 Hz sampled data, the changed speed of Y-axis coordinates is faster, and the red error curve in the corresponding graph is relatively larger; in 100 Hz sampled data, the change speed of X-axis coordinates is faster, and in the corresponding graph, the blue error curve is relatively larger. It shows that within a certain range, the faster the carrier moves, the greater the influence of noise is; the slower the carrier moves, the less the influence of noise is.
2. When the sampling frequency is 100 Hz, the carrier moves faster on the Z axis, and the error is greater than that on the Z axis of 10 Hz and 50 Hz. It shows that within a certain range, the faster the carrier moves, the greater the influence of noise is; the slower the carrier moves, the less the influence of noise is.
3. The maximum errors before and after Kalman filtering are basically within 0.1 mm, which is consistent with the positioning accuracy of laser tracker at um level.

In order to verify the effect of the maximum acceleration of the hypothetical target on Kalman filtering, and to change the maximum acceleration of the hypothetical target, the sampling data of 100 Hz are filtered, and the mean square deviation under different hypothetical maximum acceleration conditions is obtained as follows:

Table 1. The square mean errors of three axis (*mm*)

	3 m/s^2	10 m/s^2	20 m/s^2
X	0.0752	0.0406	0.0589
Y	0.0702	0.0284	0.0465
Z	0.0432	0.0288	0.0362

According to Table 1, different maximum acceleration has different influence on filtering effect, and the assumed maximum acceleration has no linear relationship with filtering error, the filtering effect of adaptive Kalman filter can be improved by choosing the appropriate assumed maximum acceleration.

In order to verify the influence of sampling frequency on filtering effect, the maximum acceleration of 5 m/s^2 is selected. The sampling data with sampling frequency of 10 Hz, 50 Hz and 100 Hz are filtered. The errors after filtering are as follows (Tables 2, 3 and 4):

Table 2. The errors of three axes when frequency is 10 Hz

	Maximum error (*mm*)	Mean square error (*mm*)
X	0.0324	0.0097
Y	0.0634	0.0229
Z	0.0149	0.0044

Table 3. The errors of three axes when frequency is 50 Hz

	Maximum error (*mm*)	Mean square error (*mm*)
X	0.0571	0.0127
Y	0.0777	0.0176
Z	0.0128	0.0026

Table 4. The errors of three axes when frequency is 100 Hz

	Maximum error (*mm*)	Mean square error (*mm*)
X	0.1398	0.0311
Y	0.1023	0.0213
Z	0.0623	0.0146

According to the table,

1. The maximum errors of the three-axis distance of X-axis, Y-axis and Z-axis are generally controlled within 0.1 mm, which is consistent with the positioning accuracy of laser tracker in um level.
2. The errors of Z axis are smaller than that of X axis and Y axis. The reason is that the target moves slowly in Z axis, the influence of noise and other factors is smaller, and the measurement accuracy of laser tracker is better.
3. Sampling frequency has little influence on removing noise. When other factors remain unchanged, the errors before and after Kalman filtering are mainly related to the speed of target.

3 Location Estimation of Direct Interpolation

3.1 Interpolation Method

Interpolation method is a mathematical method to express a continuous function by using discrete known points. Interpolation method is also called "interpolation method" [9].

3.1.1 Linear Interpolation

Given the adjacent points (x_0, y_0) and (x_1, y_1). The result of the function at x is $(y - y_0)(x_1 - x_0) = (y_1 - y_0)(x - x_0)$ within the interval.

3.1.2 Nearest Adjacent Point Interpolation

The principle of the nearest adjacent point interpolation method is that the function value of the nearest known point to the unknown point is regarded as the function value of the unknown point, and its error is often large.

3.1.3 Three Spline Interpolation

There are several points, $x_0, x_1 \ldots x_{n-1}$, in the interval [m, n]. If the function $g(x)$ can satisfy following conditions:

1. In the interval $[x_{k-1}, x_k](k = 1, 2\ldots, \mathrm{n})$, $g(x)$ is multinomial with exponent i;
2. $g^{(m)}(x)(m = 0, 1, \ldots, \mathrm{i} - 1)$ are continuous in the interval [m, n];
3. For a known function $F(x_i) = y_i$, if it satisfies $g(x_i) = y_i$, $g(x)$ is interpolation function whose exponential is i.

When i is 3, $g(x)$ is cubic spline interpolation function [10].

3.2 Interpolation Calculation

In order to compare the advantages and disadvantages of the three interpolation methods, the odd points (n = 1, 3, 5, 7…) of the sample data of the laser tracker are selected. Interpolation results of even points (n = 2, 4, 6, 8…) are obtained by interpolation method. The even point coordinates calculated by interpolation method are compared with the even point coordinates recorded by laser tracker to obtain the errors of the three interpolation methods. The sampling data with sampling frequency of 100 Hz are selected and the odd points are known. The errors of three interpolation methods are obtained by interpolation method as follows (Figs. 4, 5 and 6):

It can be seen from the graph that the errors caused by the three-spline interpolation are the smallest for the sampling data with the sampling frequency of 100 Hz, and the positive and negative distributions are uniform. The error distributions of linear interpolation are uneven and larger than that of three-spline interpolation. The error caused by the interpolation method of the nearest adjacent points is too large. Therefore, the three-spline interpolation method is used to estimate the position of the target at any time.

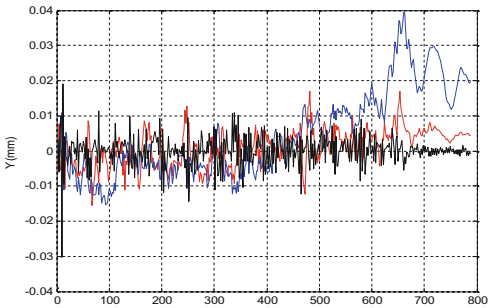


Fig. 4. The errors of linear interpolation

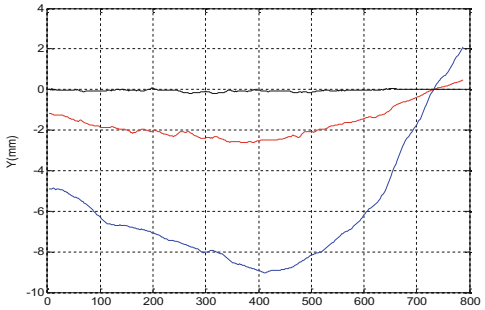


Fig. 5. The errors of nearest neighbor

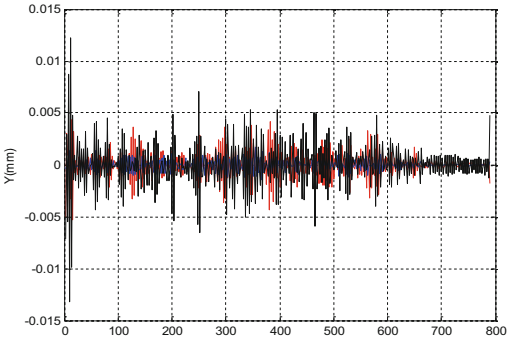


Fig. 6. The errors of three spline

Taking the sample data of 10 Hz as an example, the image and trajectory of three-axis interpolation using the three-spline interpolation vector are as follows (Figs. 7 and 8):

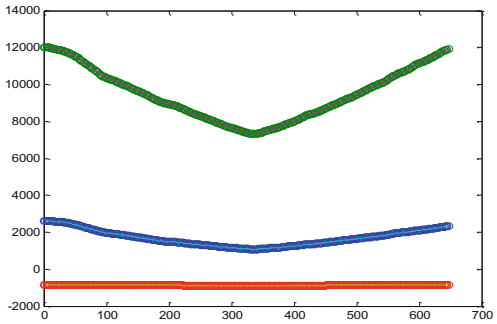


Fig. 7. The errors of three spline

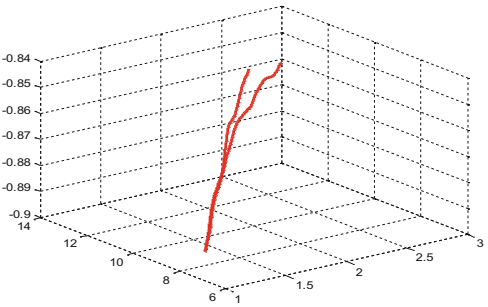


Fig. 8. The trail of carrier

4 Calculation of Carrier Posture

Three target balls are fixed on the carrier, and three laser trackers track three target balls respectively. The same signal source transmits signals to three laser trackers at the same time, obtaining the position data of three target balls at the same sampling time. Assuming that the coordinates of the three target balls in the measuring coordinate system at time t are respectively (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3), six degrees of freedom of object determined by solving equations [11].

4.1 Definition of Various Coordinate Systems

4.1.1 The System of Laser Tracker Coordinate

It is stipulated that the centre of the laser tracker is the origin O, the line from the origin to the rotation axis of the laser tracker is the Z axis, and a line perpendicular to the Z axis is the X axis. According to the right-hand principle, the Y axis is determined to form the coordinate system of the laser tracker [12].

4.1.2 Measuring Coordinate System of Multiple Instruments

The phase centres of three laser trackers are A, B and C respectively. Assuming that A is origin 0, the straight line from point A to point B is the positive direction of X axis,

the straight line perpendicular to the plane where A, B and C are located and pointing to C is the positive direction of Y axis. According to the right-hand criterion, the Z axis is determined to form the measuring coordinate system.

4.1.3 The Coordinate System of the Carrier

Assume that the three target balls tracked by three laser trackers are M, P and Q. It is stipulated that the centre of M is origin O, the line of connection of M and P centre is X axis, and the line passing through origin O and perpendicular to the plane of M, P and Q centre is Z axis. According to the right hand criterion, the Y axis is determined to establish the carrier coordinate system.

4.2 Calculation of Carrier Position

In this paper, the position of the first target ball is defined as the position of the carrier. Assuming that the position coordinates of the first target ball in the measuring coordinate system at time t are (x1, y1, z1), then the position coordinates of the carrier in the measuring coordinate system at time t are (x1, y1, z1) [13].

4.3 Calculation of Carrier Posture

Suppose that the coordinates of the three target ball centers in the measuring coordinate system at time t are (x1, x2, x3), (x2, y2, z2) and (x3, y3, z3). The forward directions of X, Y and Z axes in the carrier coordinate system are expressed respectively in the measurement coordinate system.

$$\vec{X} = \vec{B} - \vec{A} = (x2 - x1, y2 - y1, z2 - z1) \tag{13}$$

$$\vec{Z} = (\vec{B} - \vec{A}) \times (\vec{C} - \vec{A}) \tag{14}$$

$$\vec{Y} = \vec{Z} \times \vec{X} \tag{15}$$

The unit direction vectors of the three axes in the coordinate system of the carrier are obtained through the unit transformation.

$$\vec{\mathrm{x}} = \vec{X} / |X| \tag{16}$$

$$\vec{y} = \vec{Y} / |Y| \tag{17}$$

$$\vec{z} = \vec{Z} / |Z| \tag{18}$$

The cosine matrix from the measuring coordinate system to the carrier coordinate system is shown below.

$$R = \begin{bmatrix} \vec{x} \\ \vec{y} \\ \vec{z} \end{bmatrix} \tag{19}$$

The relation between Euler angle and cosine matrix is shown below.

$$R = \begin{bmatrix} \cos\psi\ \cos\theta & \cos\phi\ \sin\theta\ \sin\psi - \cos\phi\ \sin\psi & \sin\psi\ \sin\phi + \cos\psi\ \cos\phi\ \sin\theta \\ \cos\theta\ \sin\phi & \cos\psi\ \cos\phi + \sin\psi\ \sin\theta\ \sin\phi & \cos\phi\ \sin\psi\ \sin\theta - \cos\psi\ \sin\phi \\ -\sin\theta & \cos\theta\ \sin\phi & \cos\theta\ \cos\phi \end{bmatrix} \tag{20}$$

The posture (ψ, θ and ϕ) angle of the carrier is obtained by solving 19 and 20 equations.

In conclusion, Kalman filtering algorithm is used to filter the data of the moving carrier. The position of the target at any time is calculated by three spline interpolation method. Inputting the time t at will, the position and posture of the carrier are calculated as follows (Table 5):

Table 5. Position and posture of the carrier when the time is t

	0.1	1.3	5.2	12.9	20.5
X	2.620	2.620	2.617	2.596	2.570
Y	12.028	12.028	12.020	11.964	11.893
Z	−0.851	−0.851	−0.851	−0.850	−0.851
θ	0.312	0.312	0.313	0.311	0.309
ϕ	0.119	0.118	0.118	0.122	0.122
ψ	4.629	4.629	4.631	4.642	4.652

5 Summary

In this paper, the "current" statistical model is established; the algorithm of adaptive Kalman filter is used to reduce the influence of noise and other factors on the laser tracking system; three spline interpolation method is used to obtain the position of three targets at any time; finally, through establishing multiple coordinate systems and utilizing the transformation relationship among the coordinate systems, the position and posture of the carrier at any time would be calculated. In this paper, adaptive Kalman filtering algorithm is used to remove noise, improving the accuracy of laser tracker in calculating the position and posture of the carrier and providing reference for obtaining high-precision position and posture of vehicles, satellites and machines.

References

1. Ye Z (2013) Measuring method of relative position and posture based on TOF camera, Hefei Industry University
2. Chou J-Y, Chang C-M (2018) Decentralized damage detection of seismically-excited buildings using multiple banks of Kalman estimators. Adv Eng Inf 38:1–13
3. Mehammer EB, Føre M, Sauder T, Chabaud VB, Parisini T (2018) Kalman estimation of position and velocity for ReaTHM testing applications. J Phys: Conf Ser 1104(1):012008
4. Li J (2018) Research on key technologies of radar maneuvering target tracking, Dalian Maritime University
5. Liu W, Pan H, Li Y (2016) Fuzzily adaptive algorithm for current statistical model of maneuvering target. J Ordnance Eng 37(11):2037–2043
6. Liu N (2016) Adaptive filtering algorithm based on maneuvering target tracking model, Zhejiang University of Technology
7. Peng D (2008) Basic principle and application of Kalman filter, Graduate School, China University of Geosciences
8. Feng K (2015) Research on algorithms and system design in multi-target tracking technology, Jiangsu University of Science and Technology
9. Zhu C (2000) Numeric Computing Method and Its Application, 2nd edn. Science Education Press, pp 23–58
10. (2005) Manual of Modern Applied Mathematics-Computation and Numerical Analysis, 1st edn. Tsinghua University Press, pp 100–130
11. Li X, Xiang M, Fan B, Wang T, Zhou W (2018) Analysis and improvement of calibration accuracy of kinematic parameters of industrial robots. J Surv Mapp Sci Technol 35(03):255–259
12. Xiang M, Fan B, Li X, Long C (2018) Research on coordinate transformation between laser tracker and robot. Aeronaut Manuf Technol 61(Z1):98–101
13. Wang X, Fan B, Wang T, Yang Z, Zhang G (2017) Accuracy analysis of directional solution for dynamic position and posture of laser tracking measurement. Surv Mapp Eng 26(03):56–59+64

PDR/GNSS Fusion Algorithm Based on Joint Heading Estimation

Abdul Rehman, Qiang Liu(✉), Zida Wu, Huiping Zhu, Jiuchao Qian, Yuze Wang, and Peilin Liu

School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
lugialiu@sjtu.edu.cn

Abstract. The performance of Pedestrian Dead Reckoning (PDR) based on a smartphone is limited due to the low-cost MEMS IMU. Heading and stride length estimation tend to accumulate errors over time and those errors lead to the failure of PDR. Considering that Global Navigation Satellite System (GNSS) provides absolute location information in outdoor applications, and the characteristics of estimation errors of GNSS are quite different from those of PDR. In this paper, we analyse the characteristics of two different types of errors first, then we propose a fusion positioning framework that can fuse PDR and GNSS information in real time. The joint estimations of heading and stride length are also given. The experimental results show that the fusion algorithm is superior to PDR or GNSS in heading estimation, anti-multipath and noise performance.

Keywords: PDR · Heading estimation · Fusion positioning

1 Introduction

With the development of smartphones, many navigation sensors are integrated into it, such as GNSS module and IMU. However, the performance of low-cost devices on smartphone differs greatly from that of professional devices. For example, the bias instability of the MEMS accelerometer is generally greater than 0.03 m/s^2, the bias instability of the MEMS gyroscope is generally greater than 100°/h, and the GNSS positioning accuracy of smartphones based on single point positioning is usually about 10 m [1, 5]. For this reason, the traditional Strap-down Inertial Navigation System (SINS) algorithm is almost impossible to implement on the MEMS IMU. Although the low accuracy of MEMS will lead to accumulated errors in SINS, in some scenarios, such as a pedestrian or driving vehicles, the MEMS sensors can be used to assist attitude measurement, and the attitude change has a certain regularity. Therefore, some techniques such as PDR [3] and Vehicle Dead Reckoning (VDR) are developed in recent years. Besides GNSS and IMU, there are many other types of sensors integrated on smartphones, such as magnetometer, Bluetooth, WIFI and camera etc., so the fusion scheme can be varied. Taking the GNSS and INS integrated navigation as an example, this is a highly complementary system, so it has achieved success in many applications. However, MEMS is not good enough to support INS because of the above-mentioned disadvantages. According to the level of coupling, integrated navigation can be divided

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 326–339, 2019.
https://doi.org/10.1007/978-981-13-7751-8_33

into three types: loose coupling at the level of positioning results, tight coupling at the level of observations and deep/ultra-tight coupling at the level of signal tracking. With the open access to the raw GNSS observations of Android smartphones in recent years, the tight coupling between MEMS IMU and low-cost GNSS is theoretically supported, but such research still faces great challenges.

Because the low-cost MEMS IMU and GNSS on smartphones are still unable to achieve GNSS/INS integrated navigation, while the integration of PDR and GNSS based on MEMS is considered to be feasible, so many related research have been published in recent years. As mentioned above, the fusion based on PDR and low-cost GNSS can be either loosely coupled or tightly coupled. Hsu et al. [2] proposed a framework of the fusion algorithm based on 3D map-assisted GNSS and PDR in an urban environment. This algorithm outputs the stride length and heading of PDR as well as the position and accuracy of 3D-GNSS into the Kalman filter to obtain the fusion result, so it is a loosely coupled method. The observations of the Kalman filter come from 3D-GNSS, while the PDR outputs are used to be the control inputs. This is because the trajectory of pedestrians is very uncertain and it is difficult to be described by linear state transition equations. Therefore, the real-time observations of PDR are used to update the state transition equations. However, the accumulated error of PDR still exists in the fusion system, resulting in the final heading estimations. Lan et al. [4] uses the EKF to fuse PDR and GNSS. This algorithm is directly based on the output information of the MEMS sensors, so it can be regarded as a tightly coupled algorithm. This algorithm uses the heading derived by GNSS to assist the heading estimation of the MEMS gyroscope, which improves the overall accuracy of the heading estimation. However, the serious noise problem in the heading of GNSS has not been solved well, so the initial stage of the final trajectory output dependent highly on the heading of GNSS.

Based on the above analysis, this paper focuses on the error analysis of PDR and GNSS, then a fusion framework is proposed, which fully combines the advantages of PDR and GNSS, and effectively eliminates accumulated error and noise. In addition, considering the inconsistency of update frequencies between GNSS and PDR in practice, a synchronization method is also discussed.

The following is organized as follows. The PDR algorithm is briefly introduced in Sect. 2. We discuss the error analysis, data synchronization and fusion framework in Sect. 3. Section 4 shows the experimental setup and results. Section 5 gives the summary and future work.

2 Pedestrian Dead Reckoning

The PDR algorithm typically consists of three steps, they are step detection, stride length estimation and heading estimation, as shown in Fig. 1.

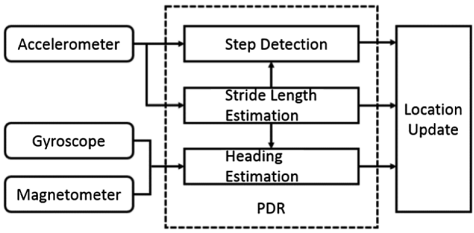


Fig. 1. Block diagram of the PDR algorithm

3-axis accelerometer output is generally used to detect the occurrence of the step and to estimate the stride length, while the 3-axis gyroscope or magnetometer is used to estimate the heading. When PDR detects a step, it will output the heading and stride length estimation {θ, SL} of step k. If the initial position is prior known, we can get the pedestrian position by Eq. 1 as follows.

$$\begin{cases} x_{k+1} = x_k + SL_k \cdot \ sin\theta_k \\ y_{k+1} = y_k + SL_k \cdot \ cos\theta_k \end{cases} \tag{1}$$

2.1 Step Detection

Step detection is mainly based on the periodic signal generated by the accelerometer when a pedestrian is walking, in order to determine whether the pedestrian has stepped or not. Common step detection methods include zero-crossing detection, peak detection, the autocorrelation method and the spectrum analysis method. In this paper, the raw data from accelerometer are pre-processed by setting the low-pass filter with the cut-off frequency is 3 Hz to reduce the device noise. Then the dynamic threshold zero-crossing detection method proposed in [7] is used to control the transition between walking and non-walking states through a finite state machine, it can achieve the 99% accuracy.

2.2 Stride Length Estimation

There are several methods for stride length estimation, they are empirical model, the linear model, the non-linear model and the machine learning model. The stride length estimation model used in this paper is a binary linear regression model with the stride frequency f_s and the acceleration variance σ_a^2 as parameters, and this model was first proposed in [8]. This model is expressed as follows.

$$SL = \alpha f_S + \beta\sigma_a^2 + \gamma \tag{2}$$

In Eq. 2, α and β are weighting factors, while γ is a constant, they need to be determined by off-line training. Once determined, it can be used to estimate the stride length of different pedestrians in real-time applications.

2.3 Heading Estimation

Heading estimation is the most important part of the PDR algorithm because the small estimation errors introduced in the heading estimation will accumulate, which will eventually lead to shape changes between estimated trajectory and real trajectory. The heading estimation usually uses the data from gyroscope, magnetometer or both of them. Since the gyroscope generates a cumulative error during integration, and the magnetometer is susceptible to interference from the surrounding magnetic field, these two sensors have complementary characteristics. The heading estimation used in this paper combines gyroscope and magnetometer, and the optimal heading estimation is able to be obtained by using the redundant information from gyroscope and magnetometer.

3 PDR/GNSS Fusion Framework

This section firstly defines the GNSS heading and stride length, then discusses the synchronization of PDR and GNSS data. After that, the analysis for heading error and stride length error of PDR and GNSS is given and the fusion framework is also proposed.

3.1 GNSS Heading and Stride Length

GNSS single point positioning gives the position of the user at a particular time. It is generally considered that the GNSS trajectory can be obtained by connecting the single point positions in time. The receiver outputs the filtered single point positions in general, but there is still the influence of thermal noise. Especially with low-cost GNSS on smartphone, the noise power is much larger than the professional receiver. Because the observed outputs of the PDR are heading and stride length, in order to unify the outputs of GNSS and PDR, it is necessary to define the GNSS heading and stride length. These two concepts are rarely mentioned in the traditional GNSS field. The definitions are described as follows.

Coordinate transformation: The latitude, longitude and height (LLH) output by GNSS are first converted into ECEF XYZ coordinates, and then the ECEF coordinates can be converted to East-North-Up (ENU) coordinates based on a prior known initial position.

Calculate heading and stride length: Assume (E_k, N_k, U_k) and (E_{k+1}, N_{k+1}, U_{k+1}) are the ENU coordinates of time k and k + 1, then the calculation formula of the GNSS heading θ_{k+1}^{GNSS} and the stride length SL_{k+1}^{GNSS} is shown in Eq. 3. It should be noted that the definition of GNSS heading is the same as that of PDR, that is, the angle between the vector of the two positioning results and the northward direction. The range of the heading in the formula is ($-\pi$, $+\pi$).

$$\begin{cases} \theta_{k+1}^{GNSS} = arctan\left(\frac{N_{k+1}-N_k}{E_{k+1}-E_k}\right) \\ SL_{k+1}^{GNSS} = \left\| \begin{bmatrix} E_{k+1}-E_k \\ N_{k+1}-N_k \end{bmatrix} \right\|_2 \end{cases} \tag{3}$$

3.2 Heading and Stride Length Errors of PDR and GNSS

The Eq. 4 gives an equivalent expression of Eq. 1, where $\{x_1, y_1\}$ is the known initial position. In each iteration process, some errors will be introduced in stride length and heading estimations, which will be reflected in position x and y. Equation 4 can be used to derive the formations of the errors on x and y, but the process is complex, and it is not very helpful to eliminate accumulated errors.

$$\begin{cases} x_{k+1} = x_1 + \sum_{i=1}^{k} (SL_i \cdot \ sin\theta_i) \\ y_{k+1} = y_1 + \sum_{i=1}^{k} (SL_i \cdot \ cos\theta_i) \end{cases} \tag{4}$$

In order to combine PDR and GNSS, one idea is based on the fusion of positions. However, according to above analysis, it is not convenient to analyse the influence of accumulated errors by directly fusing their positions. The following will give the error analysis of heading and stride length for two systems and state the aims of fusion algorithm. The necessary symbolic descriptions are given as follows.

$\varepsilon_k^{\theta-PDR}$: heading estimation error of PDR at step k
ε_k^{SL-PDR}: stride length estimation error of PDR at step k
$\varepsilon_k^{\theta-GNSS}$: heading estimation error of GNSS at time k
$\varepsilon_k^{SL-GNSS}$: stride length estimation error of GNSS at time k

In order to analyze the characteristics of the above errors, we designed the following experiment. This experiment requires some priori control points, for PDR, the tester needs to travel along the road with a straight line and step by step along the ground control points in the line. The stride length can be fixed to a constant, and the real-time data can be recorded and analysed by the post-processing software. The characteristics of the four types of errors can be obtained by carrying out a large number of experiments described above. According to these experimental results, the heading and stride length errors of PDR increase with the increase of k, while the heading and stride length errors of GNSS are more like zero-mean Gaussian noise. In this paper, we define a metric called Mean Cumulative Heading Error (MCHE) as shown in Eq. 5 to measure the heading errors of the two systems.

$$MCHE(k) = \frac{1}{k} \sum_{i=1}^{k} \varepsilon_k^{\theta-SYS} \tag{5}$$

In Eq. 5, SYS can be either GNSS or PDR. MCHE can effectively distinguish the characteristics of Gaussian noise and cumulative error. When k increases, the zero-mean noise MCHE of GNSS will tend to zero, while the cumulative error MCHE of PDR will be divergent. Similarly, we define the Cumulative Stride Length Error (CSLE) as shown in Eq. 6 to measure the stride length errors of the two systems.

$$CSLE(k) = \sum_{i=1}^{k} \varepsilon_k^{SL-SYS} \tag{6}$$

According to the above discussion, one of the objectives of the proposed fusion algorithm in this paper is to make the fused estimations are not only able to avoid the accumulated errors, but them also can reduce the noise power.

3.3 Data Synchronization

In PDR, the output is updated when the step is detected, so the outputs of PDR are unpredictable. However, the GNSS updates its results periodically, which is usually an update rate of 1 Hz. In order to fuse these data in real time, it is necessary to make the output frequency of the two systems be consistent. When a person is walking normally, usually at least one complete step can be completed within 1 s, and if a person cannot complete a step within 1 s, it is more likely that this person does not make a decision for next step. The first case is called a fast step, and the second case is a slow step, as shown in Fig. 2. In general, people take fast steps when they are walking continuously, and the slow steps happen accidentally. If there is no step for a long time, it is called a stop, that is, the pedestrian's feet are closed and stay in place. Figure 2 shows a schematic of the asynchronous output of GNSS and PDR.

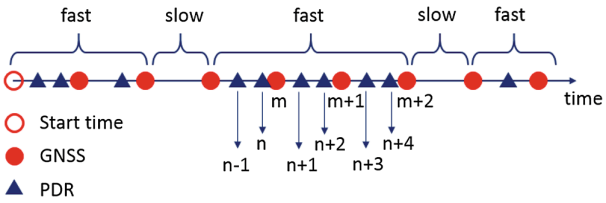


Fig. 2. Data synchronization between PDR and GNSS

The basic idea of the data synchronization method proposed in this paper is to turn the data of PDR into a timing output. In Fig. 2, between time m and m + 1 of GNSS, there are two PDR outputs, they are step n + 1 and n + 2, respectively, corresponding to $\{\theta_{n+1}, SL_{n+1}\}$ and $\{\theta_{n+2}, SL_{n+2}\}$, we define a PDR output $\{\theta_m, SL_m\}$ at time m of GNSS as shown in Eq. 7.

$$\begin{cases} \theta_m = \frac{\theta_{n+1} + \theta_{n+2}}{2} \\ SL_m = \sqrt{SL_{n+1}^2 + SL_{n+2}^2 + 2C} \\ C = SL_{n+1} SL_{n+2} \cdot cos(\theta_{n+1} - \theta_{n+2}) \end{cases} \tag{7}$$

Equation 7 gives a method for calculating the synchronized PDR heading and stride length at a GNSS time. This equation can be extended to the case where there are more than one PDR outputs in one GNSS interval, where the calculation of the heading

is averaging, and the calculation of the stride length is not a simple summation, but the distance between the first PDR output at time m and the first PDR output at time m + 1.

3.4 Fusion Framework

Due to the cumulative errors, the trajectory of PDR will have severe distortion with the increase of time. The trajectory of GNSS appears as a non-smooth curve, which is the effect of thermal noise, but it does not produce distortion. This is a very obvious difference between relative positioning and absolute positioning. The goal of the fusion algorithm is to suppress the cumulative errors and reduce the noise power as much as possible so that the final pedestrian trajectory can truly reflect the real dynamic position of the pedestrian, that is, the trajectory is smooth and not shape changed.

Considering that PDR/GNSS fusion algorithm can directly adopt a loose coupled method, as shown in Fig. 3. The synchronized PDR system outputs position and heading information, then they make difference with the outputs of GNSS. After that, the position and heading residuals will be sent to Kalman filter for corrections. Finally, the corrections will be returned to PDR system to form the estimated position and heading. However, the difficulty of this method is that the observation noise of position still needs to be analysed in the design of Kalman filter. In practical applications, the statistical characteristics of these noises and errors are difficult to obtain. Moreover, as discussed in Sect. 3.2, the error and noise components in the final position observation are difficult to separate.

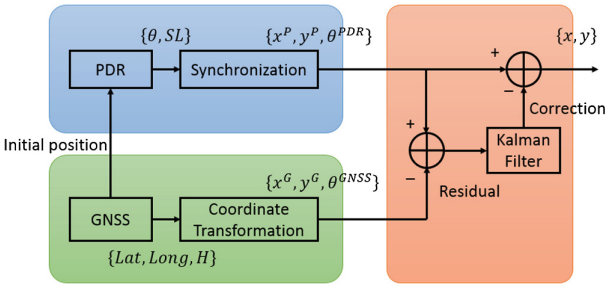


Fig. 3. Traditional loose-coupled framework

In order to reduce the dependence on the selection of Kalman parameters, another fusion idea is presented in this paper. The idea is based on joint heading estimation, which uses PDR heading with low noise power to smooth GNSS heading with high noise. Not only that, because the differential PDR heading is used in the algorithm, the cumulative error is also suppressed. The proposed PDR/GNSS fusion framework is shown in Fig. 4. The initial position is provided by the GNSS at the initial stage.

Then PDR outputs the heading and stride length estimation, and GNSS outputs the latitude and longitude, after coordinate transformation, the GNSS heading and stride length can be obtained according to Eq. 3. At the same time, PDR output the synchronized heading and stride length $\{\theta^{PDR}, SL^{PDR}\}$.

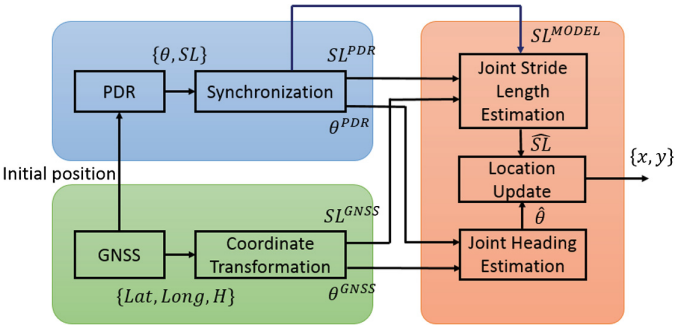


Fig. 4. Joint fusion framework of PDR/GNSS

Then, the heading and stride length estimation of the two systems are respectively sent to the joint heading estimation and joint stride length estimation module, and finally the heading and stride length estimation $\left\{\widehat{\theta}, \widehat{\mathrm{SL}}\right\}$ after the fusion filtering are output by the two modules and sent them to the location update module. The iterative update of the pedestrian location is completed by Eq. 1.

The purpose of the joint heading estimation module is to fuse the PDR heading with cumulative error and the GNSS heading with random noise to achieve the dual effect of reducing the cumulative error and the noise power. The detail is shown in Eq. 8 as follows.

$$\begin{aligned} \widehat{\theta}_{k+1} &= \tfrac{1}{M}\Theta_{k+1} + \tfrac{M-1}{M}\left(\widehat{\theta}_k + \theta^{PDR}_{k+1} - \theta^{PDR}_k\right) \\ \Theta_{k+1} &= p\theta^{GNSS}_{k+1} + (1-p)\theta^{PDR}_{k+1} \end{aligned} \tag{8}$$

The input of the above iterative equation is the heading estimation of two systems θ^{PDR} and θ^{GNSS}, and the output is the joint heading estimation. M and p are the design parameters, where M is the smoothing coefficient, which is used to filter and smooth the GNSS heading with noise. It should not be set too large or too small, if it is too large, the tracking of the heading becomes dull, if it is too small, the filtering effect is not good, and it is usually set to 30 to 50. p is a weighting factor and usually takes a value from 1 to 1.5.

The input of the joint stride length estimation module includes three aspects: the PDR stride length after synchronization, the GNSS stride length, and the output $\mathrm{SL}^{\mathrm{MODEL}}$ of the PDR synchronization module, whereas the synchronization module determines the mode of the pedestrian, that is, the fast step, slow step and stop mentioned in Sect. 3.3. The final output is the weighted output of the three.

4 Experimental Setup and Results

4.1 Experimental Setup

The experimental environment is shown by the red line in the Fig. 9. The selected route is a straight line with ground marks, which length is about 127 m and the off-angle of the north direction is −19.19°. In order to highlight the characteristics of PDR algorithm error accumulation over time, the tester starts from point A in a straight line, a total of three round trips (where A-B-A is a round trip), in a total time of 540 s.

In the experiment, the tester held the Huawei P10 smartphone in his right hand and walked continuously along the straight line. The walking speed was relatively uniform, and occasionally there were fast and slow. When he reached one of the two endpoints, A or B, his feet closed together, took turn and started moving again. During the experiment, there are occasional vehicles on both sides, and the multipath error of the low-cost GNSS chip in smartphone will become larger in this case, so the anti-interference ability of the fusion algorithm can also be tested.

4.2 Experimental Results

4.2.1 Error Analysis

Figure 5(a) shows the step detection results of PDR. The stop time of three round trips can be clearly seen from the figure, as shown by the red dashed box. Figure 5(b) shows the comparison between the heading estimation of PDR and the true value. It can be found that the heading estimation is relatively accurate in the first 200 s, but less accurate in the next 300 s. The heading estimation has a significant deviation during three round trips.

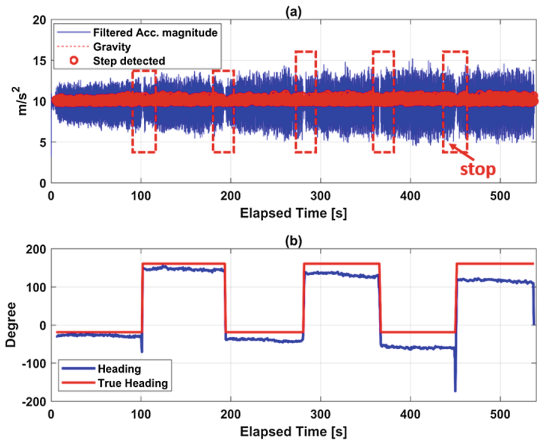


Fig. 5. Step detection and heading estimation

Figure 6 shows the characteristics of the stride length estimation error. Figure 6(a) shows a comparison of the estimated stride length of PDR with the average stride length. It can be seen that the estimated stride length is significantly larger than the average stride length with an increase in time, with an average stride length of approximately 0.73 m. Figure 6(b) shows CSLE of PDR. It can be seen from the figure that since there is an error of about 0.1 m for each PDR stride length estimation, the total error of nearly 50 m at 1000 steps is finally caused. Figure 5(b) and Fig. 6(b) together show that both the stride length estimation and the heading estimation in PDR introduce a large cumulative error.

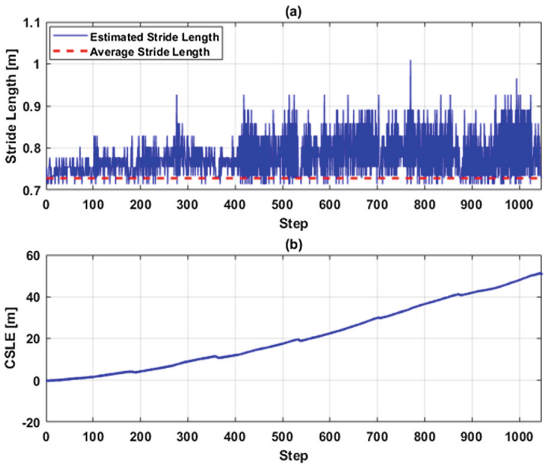


Fig. 6. Stride length error of PDR

Figure 7 shows the error characteristics of PDR heading and GNSS heading. The red line in (a) shows the heading error of GNSS. It can be seen that the GNSS heading is greatly affected by noise, and the maximum error is 20°. The PDR heading indicated by the blue line shows a time accumulation error, but it can be seen that the noise on it, is not very large. The blue line in the figure has several glitches, which is caused by the detection of stop in PDR algorithm, indicating that the pedestrian is stopped at this moment. Figure 7(b) shows the average cumulative heading error of MCHE for GNSS heading and PDR heading. This figure illustrates the completely different characteristics of the two errors, i.e. the GNSS heading error is zero-mean random noise, and the PDR heading error appears the nature of the time function.

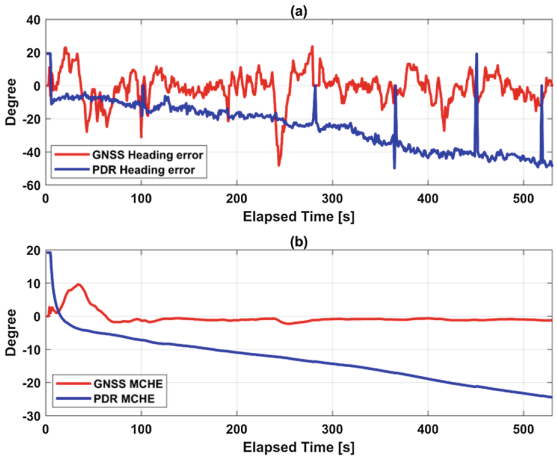


Fig. 7. Heading errors of PDR and GNSS

4.2.2 Fusion Results

The comparison of the GNSS heading, PDR heading and the heading estimation obtained by the fusion algorithm proposed in this paper is shown in Fig. 8. It can be seen from the figure that the GNSS heading does not have time offset, but its noise affects the accuracy of heading estimation, which makes the pedestrian trajectory derived from GNSS not smooth, and GNSS is easily interfered by environmental factors. Where the red circle 1 is marked, there is a large fluctuation in the GNSS data, this is because the occasionally passing of vehicles during the test. This situation will cause the GNSS multipath error at that moment to suddenly increase. At the same time, PDR is not subject to external interference and has high robustness. The combined heading has the respective advantages of the above two kinds of heading estimation, such as no obvious time offset, enhanced anti-interference ability and random noise reduction. The red circle 2 in the figure gives a slow step detected at a certain moment, which shows that the slow step is rarely seen during the normal walking of the pedestrian. Since the PDR heading estimation will detect the slow step, the heading estimation of the fusion algorithm will also show the effect of slow steps, but this will not affect the subsequent heading estimation. That is to say, this phenomenon is harmless.

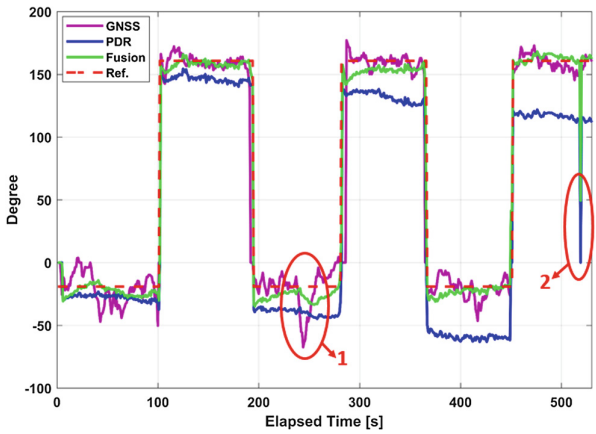


Fig. 8. Comparison of heading estimation by three different methods

Figure 9 shows the final trajectory of the three (GNSS, PDR, Fused) methods, which successively disassembles the three round trips into three graphs (a), (b) and (c). In the above figure, A is the starting point of the path and B is the ending point. As can be seen from the Fig. 9(a), the PDR method has high accuracy in the transient, but it will yaw with an increase in time, while the GNSS trajectory shows greater volatility, and the trajectory of the fusion algorithm shows the high degree of consistency with the real trajectory. Figure 9(b) shows that with the increase in time, the trajectory of PDR produces severe yaw, which can no longer reflect the true trajectory of pedestrians. In this round trip, the distortion of GNSS trajectory is due to the interference of the vehicles during test, even if the GNSS is subject to severe multipath interference and PDR produces severe yaw, in this case, the fusion algorithm can still give relatively smooth, robust and unbiased trajectories. The PDR trajectory in Fig. 9(c) has completely lost its meaning, and the GNSS is less affected by the interference during this round trip. The trajectory of the fusion algorithm can still maintain the non-yaw, indicating that it is affected by the PDR heading estimation at this time, but it also inherits the advantages of PDR low heading noise and the smoother trajectory of PDR. From the above three figures, it can be seen that the fusion algorithm basically eliminates the characteristics of time bias, and obviously enhances the anti-interference ability, and realizes the complementary advantages of PDR and GNSS. However, we have also noticed that compared with the real trajectory, the length of the fusion trajectory is shorter, and it is obvious in the Fig. 9(b) and (c). This is mainly because the stride length estimation in the fusion algorithm still has a small cumulative error. So the performance of the stride length estimation needs to be improved in future work.

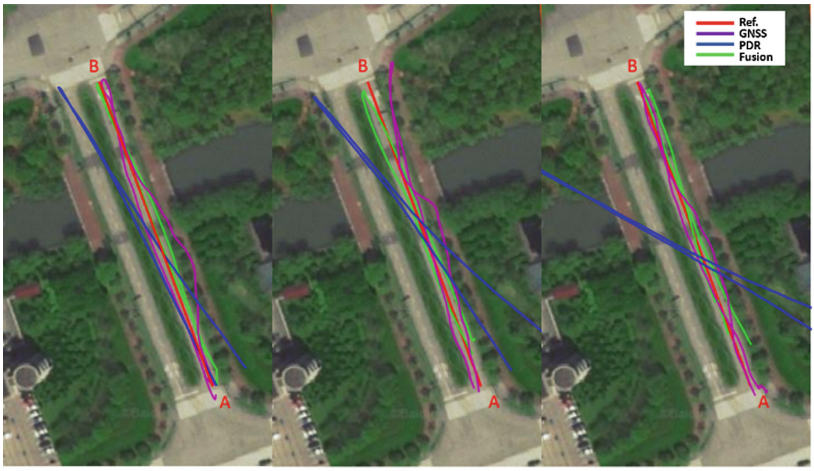


Fig. 9. Comparison of trajectories by three methods

5 Conclusion and Future Work

In this paper, an effective PDR/GNSS fusion framework is proposed by comparing and analyzing the heading and stride length errors of PDR and GNSS. The framework gives a synchronization scheme for PDR and GNSS data and the detailed heading and stride length fusion algorithm. The experimental results show that the fusion algorithm is better than the single GNSS or PDR algorithm in terms of heading estimation, anti-interference and noise performance, which reflects the complementary advantages of absolute positioning and relative positioning. At the same time, however, the accuracy of the stride length estimation needs to be improved. The performance of the fusion algorithm in the urban multipath environment remains to be verified. These are the focus of future work.

Acknowledgements. This work is funded by the China Scholarship Council under grant no. 17B8ECC74F, "The Belt and Road" program.

References

1. Groves PD (2013) Principles of GNSS, inertial, and multi-sensor integrated navigation systems. Ind Robot 67(3):191–192
2. Hsu LT, Gu Y, Huang Y et al (2016) Urban pedestrian navigation using smartphone-based dead reckoning and 3D maps aided GNSS. IEEE Sens J 16(5):1281–1293
3. Jimenez AR, Seco F, Prieto C et al (2009) A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU. In: IEEE International Symposium on Intelligent Signal Processing. IEEE, pp 37–42
4. Lan H, Yu C, El-Sheimy N (2015) An integrated PDR/GNSS pedestrian navigation system, vol 342, pp 677–690

5. Liu Q, Ying R, Wang Y, Qian J, Liu P (2018) Pseudorange double difference algorithm based on duty-cycled carrier phase smoothing on low-power smart devices. In: Sun J, Yang C, Guo S (eds) China Satellite Navigation Conference (CSNC) 2018 Proceedings, CSNC 2018. Lecture Notes in Electrical Engineering, vol 497. Springer, Singapore
6. Malkos S (2016) User location takes center stage in new android OS: Google to provide raw GNSS measurements. GPS World 27:36
7. Qian J, Ma J, Ying R et al (2013) RPNOS: Reliable Pedestrian Navigation on a Smartphone. Geo-Informatics in Resource Management and Sustainable Ecosystem. Springer, Berlin, pp 188–199
8. Shin SH, Park CG, Kim JW et al (2007) Adaptive step length estimation algorithm using low-cost MEMS inertial sensors. In: 2007 IEEE Sensors Applications Symposium, SAS 2007. IEEE, pp 1–5

Error Compensation Algorithm for Dynamic Model Based on Neural Network

Chengxin Li, Jing Peng(✉), Zengjun Liu, Huaming Chen, and Gang Ou

Navigation and Spacetime Technology Engineering Research Center, National University of Defense Technology, Deya Road 109, Changsha, China
Jingpeng_nnc@163.com

Abstract. The noise of GNSS navigation and positioning system is non-priori, while the optimal estimation of standard Kalman filter requires the establishment of accurate system model and observation model, which leads to the low accuracy of Kalman filter. Neural network has strong ability of denoising, learning, self-adapting and complex mapping. In order to improve the filtering accuracy, this paper proposes an algorithm to compensate the error of the dynamic model by using the neural network, and corrects the error of the dynamic model by using the RBF neural network in the filtering estimation part, which inhibits the contribution of the abnormal disturbance of the dynamic model to the navigation solution. The experimental results show that the algorithm can not only eliminate the positioning deviation in all directions, but also reduce the standard deviation in X, Y and Z directions by about 70%, 60% and 60% respectively, compared with the standard Kalman filter.

Keywords: Kalman filter · Neural network · Error compensation

1 Introduction

The Global Navigation Satellite System (GNSS) is constantly developing and its applications have penetrated into every aspect of our daily life. As navigation terminals are applied in more and more fields, the receiving environment they need to face is becoming more and more complex, such as the high-occlusion environment of cities, mountains and canyons, as well as the high-dynamic environment of airborne and projectile loads [1]. The Kalman filter algorithm is a typical algorithm for filtering in satellite navigation information processing. It uses the motion model and the observation model to obtain the state estimation of the carrier [2]. However, it needs linearization of the positioning equation, which may introduce errors. At the same time, the Kalman filter requests the prior information of the state model and the observation model to achieve the optimal estimation. There are many empirical models used in practice, and the model cannot be completely matched [3]. In a complex environment, abnormal observation errors can cause problems such as unstable filtering results and poor precision. In view of the application of Kalman filter in the actual process, algorithms such as extended Kalman filter and unscented Kalman filter have been proposed, but they may not really adapt to navigation and positioning in difficult

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 340–349, 2019.
https://doi.org/10.1007/978-981-13-7751-8_34

environments [4]. The neural network has strong intelligent processing capabilities such as denoising, learning, adaptive, and complex mapping [5]. In order to improve the filtering precision, an algorithm for compensating the dynamic model error by the neural network is proposed. The error of the dynamic model is corrected by the RBF neural network in the filter estimation part, which suppresses the contribution of the dynamic model abnormal disturbance to the navigation solution effectively.

2 Standard Kalman Filter Algorithm

Assuming that the state equations and observation equations of the system are known, and they are as follows

$$\boldsymbol{X}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{X}_{k-1} + \boldsymbol{W}_k \tag{1.1}$$

$$\boldsymbol{L}_k = \boldsymbol{A}_k\boldsymbol{X}_k + \boldsymbol{e}_k \tag{1.2}$$

Wherein, the subscript k represents the time. $\boldsymbol{X}_k$ is the state vector. $\boldsymbol{\Phi}_k$ is the coefficient matrix. $\boldsymbol{W}_k$ is the Gaussian white noise process error vector. $\boldsymbol{L}_k$ is the observation vector. $\boldsymbol{e}_k$ is the observed noise vector.

The standard Kalman filter algorithm is as follows.

The prediction state vector $\bar{\boldsymbol{X}}_k$ and the prediction state covariance matrix $\boldsymbol{\Sigma}_{\bar{\boldsymbol{X}}_k}$ are obtained by using the state estimation vector $\hat{\boldsymbol{X}}_{k-1}$ at time $k-1$, the state estimation vector covariance matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{X}}_{k-1}}$, and the state transition matrix $\boldsymbol{\Phi}_{k,k-1}$.

$$\bar{\boldsymbol{X}}_k = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{X}}_{k-1} \tag{1.3}$$

$$\boldsymbol{\Sigma}_{\bar{\boldsymbol{X}}_k} = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{\Sigma}_{\hat{\boldsymbol{X}}_{k-1}}\boldsymbol{\Phi}_{k,k-1}^{\mathrm{T}} + \boldsymbol{\Sigma}_{\boldsymbol{W}_k} \tag{1.4}$$

The innovation vector $\bar{\boldsymbol{V}}_k$ and the innovation vector covariance matrix $\boldsymbol{\Sigma}_{\bar{\boldsymbol{V}}_k}$ are calculated using the prediction state vector $\bar{\boldsymbol{X}}_k$, the current observation vector $\boldsymbol{L}_k$, and the observation design matrix $\boldsymbol{A}_k$.

$$\bar{\boldsymbol{V}}_k = \boldsymbol{A}_k\bar{\boldsymbol{X}}_k - \boldsymbol{L}_k \tag{1.5}$$

$$\boldsymbol{\Sigma}_{\bar{\boldsymbol{V}}_k} = \boldsymbol{A}_k\boldsymbol{\Sigma}_{\bar{\boldsymbol{X}}_k}\boldsymbol{A}_k^{\mathrm{T}} + \boldsymbol{\Sigma}_k \tag{1.6}$$

Then, the gain matrix is calculated.

$$\boldsymbol{K}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{A}_k\boldsymbol{\Sigma}_{\bar{\boldsymbol{V}}_k}^{-1} \tag{1.7}$$

Finally, the current state estimate is obtained and a new state estimation vector covariance matrix is calculated.

$$\hat{\boldsymbol{X}}_k = \bar{\boldsymbol{X}}_k - \boldsymbol{K}_k \bar{\boldsymbol{V}}_k \tag{1.8}$$

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{X}}_k} = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{A}_k)\boldsymbol{\Sigma}_{\bar{\boldsymbol{X}}_{k-1}}\left(\boldsymbol{I} - \boldsymbol{A}_k^{\mathrm{T}}\boldsymbol{K}_k^{\mathrm{T}}\right) + \boldsymbol{K}_k \boldsymbol{\Sigma}_k \boldsymbol{K}_k^{\mathrm{T}} \tag{1.9}$$

3 Error Compensation Algorithms for Dynamic Models Based on Neural Networks

3.1 Algorithm Principle

According to Eqs. (1.5) and (1.8), the standard Kalman filter solution can be rewritten as

$$\hat{\boldsymbol{X}}_k = \bar{\boldsymbol{X}}_k - \boldsymbol{K}_k(\boldsymbol{A}_k \bar{\boldsymbol{X}}_k - \boldsymbol{L}_k) \tag{1.10}$$

It is assumed that the state prediction vector $\bar{\boldsymbol{X}}_k$ of the dynamic model can obtain the true value $\boldsymbol{X}_k$ of the state vector of the maneuver carrier at time k after error compensation. Define the error compensation vector as $\boldsymbol{\Delta Er}$.

$$\begin{aligned} \boldsymbol{X}_k =& (\bar{\boldsymbol{X}}_k + \boldsymbol{\Delta Er}) - \boldsymbol{K}_k(\boldsymbol{A}_k(\bar{\boldsymbol{X}}_k + \boldsymbol{\Delta Er}) - \boldsymbol{L}_k) \\ =& (\boldsymbol{I} + \boldsymbol{K}_k \boldsymbol{A}_k)\boldsymbol{\Delta Er} - \boldsymbol{K}_k \bar{\boldsymbol{V}}_k + \bar{\boldsymbol{X}}_k \end{aligned} \tag{1.11}$$

By Eq. (1.11), the relationship between the error compensation vector $\boldsymbol{\Delta Er}$ and the product $\boldsymbol{K}_k \bar{\boldsymbol{V}}_k$ between the Kalman filter gain and the innovation vector can be considered as a nonlinear mapping of multidimensional input and output.

$$\boldsymbol{\Delta Er} = F(\boldsymbol{K}_k \bar{\boldsymbol{V}}_k) \tag{1.12}$$

Therefore, the neural network method can be used to approximate the nonlinear mapping F by training, and the relationship between $\boldsymbol{K}_k \bar{\boldsymbol{V}}_k$ and $\boldsymbol{\Delta Er}$ can be learned.

In the network prediction phase, the error compensation vector is predicted and the dynamic model is compensated. As shown in Fig. 1.

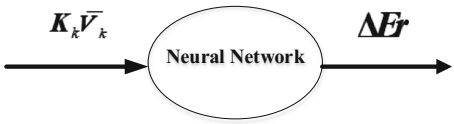


Fig. 1. Compensation flow of neural network dynamics model.

The innovation vector $\bar{\boldsymbol{V}}_k$ can also be called the prediction residual vector, which reflects the dynamic model error mainly. The Kalman filter gain $\boldsymbol{K}_k$ is determined by the state estimation vector covariance matrix, the observation design matrix, and the observation vector weight matrix. When the observation is reliable and the system

dynamics model is abnormal, the state estimation vector covariance matrix may bring the dynamic model error accordingly. Therefore, the Kalman filter gain and the innovation vector are both the source of the dynamic model error and the influencing factors, which explains the rationality of the algorithm in the physical sense.

The algorithm block diagram is shown in Fig. 2.

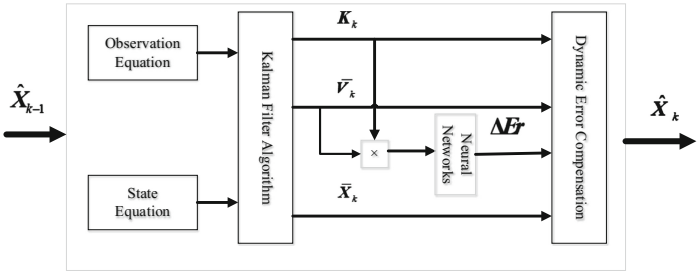


Fig. 2. Compensation algorithms of neural network dynamics model.

In practical applications, the Kalman filter gain $\boldsymbol{K}_k$, the innovation vector $\bar{\boldsymbol{V}}_k$ and the error compensation vector $\boldsymbol{\Delta Er}$ can be collected for offline training according to the carrier's starting motion for a certain length of time. After training, the neural network can be used to correct the dynamic model prediction information $\bar{\boldsymbol{X}}_k$.

3.2 Simulation Analysis

The simulation time lasted for 2000 s. Assume that the carrier only moves in a uniform circular motion on the X-Y plane, and remains stationary in the Z direction($v_z = 0$, $a_z = 0$). Carrier reference track is known. The X-direction track is shown in Fig. 3.

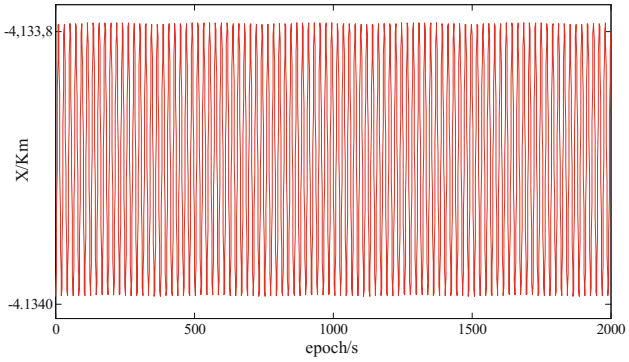


Fig. 3. Reference trajectory of the X-direction.

Assume that the receiver clock error model is a discrete-time first-order Markov model. The GPS precise ephemeris interpolation on April 1, 2007 was used to obtain the position of the satellite, and the distance between the carrier and the satellite was generated. The simulated clock error and the observed pseudorange noise were added. Set the standard deviation of the observed pseudorange noise to the typical value $\sigma_{\mathrm{UERE}} = 5.3$ m of the equivalent distance error of the GPS standard positioning service user.

The network input during training is the corresponding $\boldsymbol{K}_k\bar{\boldsymbol{V}}_k$ of each epoch. The reference output is the dynamic model error $\boldsymbol{\Delta Er}$ calculated by using the true value of the motion state of the carrier. Filter convergence time is 10 s. Train with data from 11 s to 1100 s. The post-990 s data is solved by positioning, and the $\boldsymbol{K}_k\bar{\boldsymbol{V}}_k$ corresponding to the epoch is input by the trained network to obtain the dynamic model error prediction value $\boldsymbol{\Delta Er}$, which is taken into Eq. 1.11, and the motion state prediction of the carrier after the dynamic model compensation is obtained. The value is used as the result of the positioning. RBF neural network has the characteristics of simple structure, simple training, fast learning convergence and unique optimal approximation. Here RBF neural network is used. The maximum number of neurons is set to 1000. The mean square error goal is set to 0.1. The expansion speed of the radial basis function is set to 1. The positioning result is compared with the standard Kalman filter (Figs. 4 and 5).

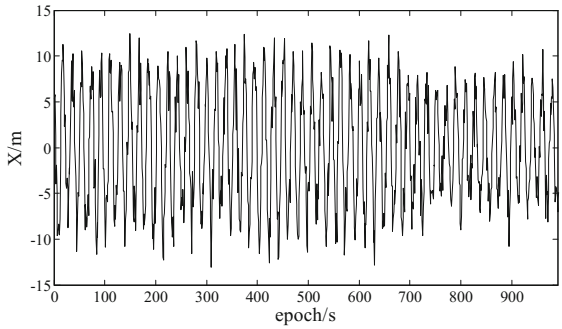


Fig. 4. Standard Kalman filter.

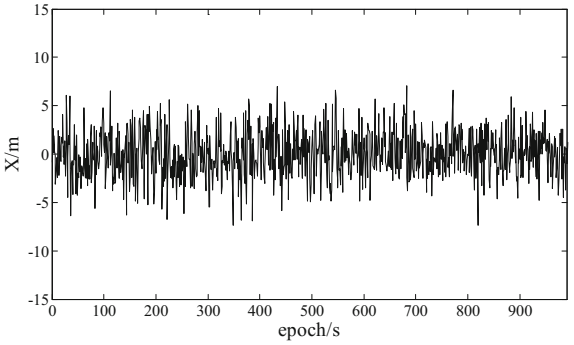


Fig. 5. Error compensation algorithm for dynamic model based on neural network.

Only the coordinate difference of the X-axis component is plotted in the figures of this paper, and the coordinate difference of the Y-axis and Z-axis components is similar to the X-axis component. The root mean square error statistics of each filtered output are listed in Table 1.

Table 1. Statistical results of root mean square error.

Filtering algorithm	RMSE/m			
	X	Y	Z	Radial
Standard Kalman filter	6.142	4.432	3.065	8.171
Error compensation algorithm for dynamic model based on neural network	1.980	1.382	1.219	2.705

It can be seen from the comparison that the neural network-based dynamic model compensation algorithm can effectively suppress the influence of the kinetic model anomaly, and the root mean square error in the horizontal and vertical directions is reduced by about 70%.

4 Measured Data Analysis

4.1 Dynamic Sports Car Measured Data Acquisition

The on-board GPS observation data collected at Songya Lake Park in Changsha City on September 19, 2018 was selected to verify the filtering algorithm proposed in this chapter. The positioning results obtained by the RTK system are used as reference. In Fig. 6, the receiver antenna is placed on the roof of the car during data acquisition. The sports car runs around the Songya Lake Park, and the sports car route is shown in Fig. 7.

Fig. 6. Antenna placement diagram for sports car experiment.

Fig. 7. Trajectory diagram of sports car experiment.

Each epoch is separated by 1 s. The number of visible satellites (cutoff angle) and GDOP values during the acquisition time are shown in Figs. 8 and 9, respectively.

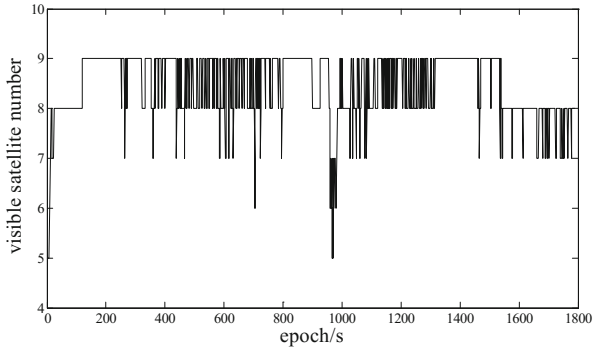


Fig. 8. Visible satellite number diagram.

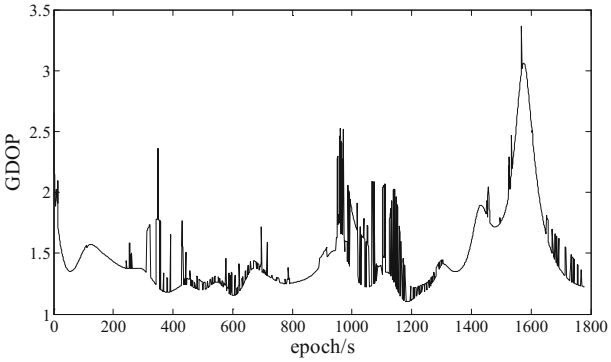


Fig. 9. Error compensation algorithm for dynamic model based on neural network.

4.2 Experimental Results and Analysis

The sports car data has a total of 1787 epochs. In order to fully adapt to the environment and obtain the initial value of each connection weight, the sample of the first 1200 s epoch is taken as the network training sample, and the initial weight of the network is assigned. The data of the last 587 epochs is used as prediction data. The results of the standard Kalman filter and the neural network-based observation error compensation filter algorithm are shown in Figs. 10 and 11.

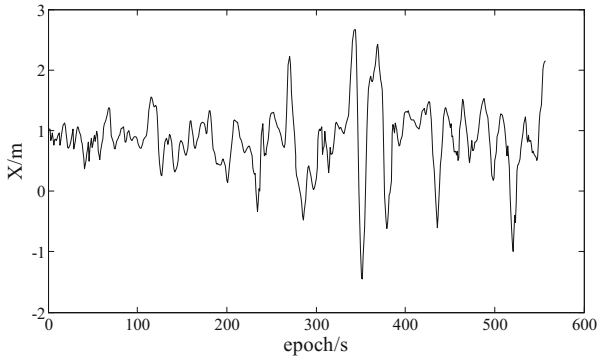


Fig. 10. Standard Kalman filter.

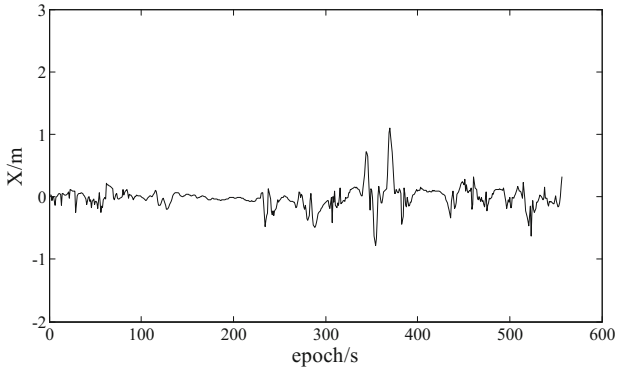


Fig. 11. Standard Kalman filter.

The statistical results of each filtered output are shown in the following Table 2.

Table 2. Statistical results of root mean square error.

Filtering algorithm	RMSE/m			
	X	Y	Z	Radial
Standard Kalman filter	1.177	2.710	1.579	3.350
Error compensation algorithm for dynamic model based on neural network	0.186	0.175	0.172	0.307

Table 3. Mean and standard deviation of errors.

Filtering algorithm	Mean/m			Standard deviation/m		
	X	Y	Z	X	Y	Z
Standard Kalman filter	0.852	−2.551	1.302	0.559	0.408	0.465
Error compensation algorithm for dynamic model based on neural network	−0.001	0.002	0.001	0.175	0.163	0.193

It can be known from the above results that there are some uncertain factors in the observation information and dynamic model information. Standard Kalman filter solution is not ideal. The dynamic model compensation algorithm based on neural network can effectively suppress the model error. Its filtering performance is greatly improved compared with the standard Kalman filter. The root mean square error in X, Y and Z directions is reduced by about 85%, 95% and 90% respectively. Through training and learning, the positioning deviation in all directions is eliminated in the positioning prediction (see Table 3). And in terms of error standard deviation in all directions, the algorithm reduces the X, Y and Z directions by about 70%, 60% and 60% respectively, compared to the standard Kalman filter. However, a burr appears between 300 s and 400 s in the prediction phase, which is due to the large GDOP value of the satellite and the mutation of GDOP value during this period. The learning sample does not contain similar conditions, resulting in no effective elimination of abnormal errors in the prediction phase.

5 Conclusion

Based on the standard Kalman filter algorithm and the RBF neural network, a dynamic model compensation algorithm based on neural network is proposed in this paper. By using the nonlinear mapping approximation ability of the RBF neural network, the relationship between the product of the Kalman filter gain and the new interest and the dynamic model error is learned, and the dynamic model error is compensated in the prediction phase. Simulation and experimental results show that the proposed algorithm can improve the positioning accuracy of standard Kalman filter effectively.

References

1. Wei E, Chai H, Liu J (2005) On the process and key technologies of GPS modernization. Bull Surv Map 12:5–7
2. Ren K, Yang L, Guo J (2012) GPS position in a severe environment. Sci Surv Map 1:5–7
3. Zhou Y (2012) GPS dynamic navigation algorithm based on Bayesian estimation. Blackwell, Tongji Shanghai
4. Teng Y, Shi L, Zheng Z (2011) Research on GPS receiver position algorithm under bad conditions. C J Sci Ins 8:1879–1884
5. Zhao H, Zeng X, He Z (2011) Low-complexity nonlinear adaptive filter based on a pipelined bilinear recurrent neural network. IEEE Tran Neural Netw 9:1494–1507
6. Mu Y (2006) Study on the application of RBF neural network learning algorithm for pattern classification. Dalian University of Technology

Accuracy Analysis of BDS/GPS Navigation Position and Service Performance Based on Non/Double Differential Mode

Yingying Ren[1], Jiexian Wang[1(✉)], Hu Wang[2], Lizhen Lian[3], and Yangfei Hou[1]

[1] College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China
wanjiexian@tongji.edu.cn
[2] Chinese Academy of Surveying and Mapping, Beijing 100039, China
[3] CAS Key Laboratory of Planetary Sciences, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

Abstract. The performance of the Chinese BeiDou navigation satellite system (BDS) for navigation and positioning has been enhanced gradually, which benefits from the improvement of the availability and distribution of monitoring stations. The aim of this paper is to assess positioning precision of GPS and BDS based on the existing high-precision positioning mode. Results showed that the overall accuracy of double differential solution would be better than the un-differential precise point positioning (PPP) solution, i.e., the accuracy of double differential solution could reach about millimetre-level while the un-differential PPP accuracy could be only about centimetre-level. The accuracy of the BDS/GPS combination solution could be roughly equivalent to that of the GPS individual solution, both of which would be superior to that of the single BDS solution with about centimetre-level. After the orbit-normal and yaw-steering correction is adopted, the positioning accuracy of BDS could be better than 5 cm. Moreover, compared with the conventional PPP solution, PPP network solution by double-difference ambiguity fixed and Uncalibrated Phase Delay (UPD) ambiguity fixed would have higher accuracy and converge faster. In contrast with the double-difference precise baseline solution, the UPD ambiguity fixed network solution with faster convergence speed could have the same accuracy and could also ensure the rigor of the whole network solution, providing a new rapid monitoring technology with high precision for the GNSS station network.

Keywords: BDS/GPS · Non-difference/double difference · PPP network solution · UPD · Accuracy analysis

1 Introduction

Since the operation of Beidou-2, both the system construction and the precision positioning and orbiting technology have been greatly developed. The positioning accuracy and the performance of the integrated navigation service of the BDS are also

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 350–359, 2019.
https://doi.org/10.1007/978-981-13-7751-8_35

constantly improved. By the end of 2012, BeiDou-2 has reached regional operational status to provide positioning and navigation services for users over the whole Asian–Pacific region [1, 2]. Compared with GPS, although BDS is still not perfect enough, the comprehensive service performance of BDS will be greatly improved. In addition, multi-system combined positioning has many advantages such as more visible satellites available and better satellite space structure. Therefore, the combined BDS/GPS positioning mode can further improve the reliability and stability of the positioning accuracy.

At present, in terms of precision positioning data processing, there are usually two solving modes: zero-differential PPP and double-difference baseline network solution [3–7]. The former has its own characteristics such as single-station operation, much flexibility of the positioning algorithm, high efficiency of computation, independence between different station solutions, and not affected by the length of the baseline. The advantages of the latter which has been widely used in high-precision GNSS data processing are listed accordingly: good integer ambiguity resolution and high positioning accuracy. In recent years, due to the development of non-difference algorithm, the large network solution scheme based on PPP [8–10], which greatly drives the calculation efficiency without sacrifice of the solution accuracy, have been proposed. They can be classified into PPP with the double-difference ambiguity fixed method and that based on the method of UPD ambiguity fixed.

In view of this, we first compared the accuracy of non-difference PPP, non-difference network solution and double difference network solution based on BDS and GPS, and then analyses the positioning accuracy of BDS in non-difference and double-difference solution mode and the contribution of BDS to BDS/GPS combined positioning.

2 Non-difference Solution

2.1 Data and Processing Strategy

The BDS observations of 20 stations distributed in Australia from the MGEX networks are employed to verify BDS positioning effect in its service range. Data for about 100 days sampled at 30 s from January 1st to April 1st, 2018 are processed.

The non-difference solution, including standard PPP and PPP network solution, are carried out with self-programmed software. According to different ambiguity fixed methods, PPP network solution could be further divided into double-difference ambiguity fixed solution and UPD ambiguity fixed solution. The details of the processing strategy can be found in Table 1.

Table 1. Non-difference processing strategy

Parameter item	Processing strategy
Observational measurement	Ionosphere-free combination observations
Sampling interval	30 s
Satellite cut-off height angle	10°
Satellite orbit and clock difference	GFZ precision products
Phase center deviation	igs14.atx
Tropospheric delay	Tropospheric delay
Ionospheric delay	Combination with no ionospheric
Ocean tide model	FES2004
Ambiguity	Estimate
EOP parameters	IERS C04 model

2.2 Non-difference PPP Solution

Based on the observations of 20 stations in Australia, the daily BDS, GPS and BDS/GPS PPP solutions are generated. Here the coordinates issued by SOPAC are selected as a reference. The root-mean-square error (RMSE) of PPP coordinate residuals at each station are shown in Fig. 1. The RMSE of each station based on BDS, GPS, BDS+GPS PPP solution are shown with red, blue and green bar, respectively. In addition, with these 20 station coordinate residuals covering 100 days, the average RMSE for the three positioning modes are given in Table 2.

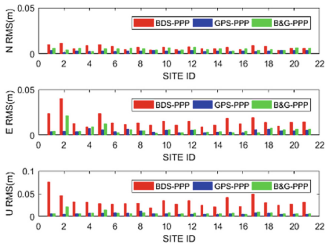


Fig. 1. 100-day average RMS (PPP)

Table 2. Average RMS of the 20 stations (PPP)

Mode	N/mm	E/mm	U/mm
Non-difference PPP-BDS	6.9	15.3	32.6
Non-difference PPP-GPS	3.1	4.5	6.5
Non-difference PPP-B/G	5.0	5.9	8.0

It can be seen from Fig. 1 and Table 2 that in terms of the three positioning modes, their positioning accuracy in the horizontal direction is better than that in the vertical direction, where the accuracy in N direction is better than that in E direction, followed by that in U direction. The accuracy of BDS PPP is worse than GPS. On the one hand, the quality of BDS precise ephemeris and clock error data are worse than that of GPS. On the other hand, no accurate PCO and PCV correction information accounted for the BDS satellite reduce its accuracy inferior to that of GPS. In addition, when solar altitude angle relative to the satellite orbital plane is small, the IGSO/MEO satellites will not track the position of sun, when they are in zero-bias mode, which lead to PPP-based parameter estimation and zenith tropospheric delay inexact and further destroy

the accuracy of satellite orbit determination [11]. After the orbit-normal and yaw-steering correction is adopted, the positioning accuracy of BDS could be better than 5 cm. The accuracy of BDS/GPS combined positioning is roughly equivalent to that of GPS only. On the one hand, the overall quality of BDS observation is worse than that of GPS; on the other hand, some inter-system errors in combined positioning could not be eliminated. The accuracy of long-term PPP static solution based on the three positioning modes could reach cm level or even mm level, where the positioning accuracy of BDS/GPS and GPS in each direction is better than 1 cm, and that of BDS is at cm level.

2.3 Non-difference Precision Network Solution

2.3.1 Double Difference Ambiguity Fixed Technique

(1) Basic principles

In the network-based solution mode, the error by the initial phase of the satellite and the hardware delay of the receiver could be eliminated through the double difference processing, and then the integer characteristic of the ambiguity could be restored [9]. In the actual data processing, a set of independent baselines are first selected. Then, for each baseline, inter-station differential processing is carried out on the two ambiguities of the same satellite. Subsequently, inter-satellite differential processing is carried out on inter-station ambiguities. Thus, the double-differenced ambiguity could be expressed by a wide-lane integer ambiguity plus a narrow-lane integer ambiguity as Eq. (1). At last, the ambiguity can be fixed one by one.

$$\begin{cases} DB_c = \frac{1}{1+g} DN_n + \frac{g}{1-g^2} DN_w \\ DN_w = DN_1 - DN_2 \\ DN_n = DN_1 \end{cases} \tag{1}$$

Where DN_w and DN_n are the ambiguity of wide-lane and narrow-lane, respectively; DN_1 and DN_2 are the double difference ambiguities with different frequencies.

For non-difference data processing, after the double-difference ambiguity is fixed, the fixed double-difference ambiguity as a constraint condition is applied in the normal equation, and the parameter estimation is performed iteratively.

(2) Case study

Based on the constraint of double-difference ambiguity fixing, daily PPP network solution from Australian regional networks could be obtained by use of the BDS-based, GPS-based, BDS/GPS combined positioning modes. The root-mean-square error (RMSE) of each station coordinate residuals are shown in Fig. 2 In addition, with these 20 station coordinate residuals covering 100 days, the average RMSE for the three positioning modes are given in Table 3.

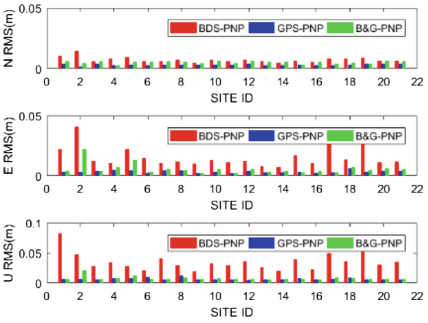


Fig. 2. 100-day average RMS (PNP)

Table 3. Average RMS of the 20 stations (PNP)

Mode	N/mm	E/mm	U/mm
BDS	7.2	15.6	34.9
GPS	3.0	3.3	6.3
B/G	5.0	5.5	7.6

2.3.2 UPD Ambiguity Fixed Technique

(1) Basic principles

The UPD-based ambiguity fixed method is a commonly used method for PPP ambiguity fixed. Some scholars have found that not only the fractional part of the wide-lane ambiguity is stable, but also the narrow-lane ambiguity changes very little in a short period of time [8, 9]. The wide-lane and narrow-lane UPD calibration is generally completed based on the non-difference ambiguity at each station uniformly distributed in the survey area. The wide lane UPD is achieved by the MW combined observations, while the narrow lane UPD is estimated as the random walk parameters.

$$L_W = N_W + f_{wr} + f_W^S \tag{2}$$

Where L_W is the float solution for wide lane ambiguity, which is solved by averaging the MW combination, N_W is the integer solution for wide lane ambiguity, f_{wr} and f_w^s are the fractional parts of the receiver and the satellite UPD, respectively.

$$L_C = \frac{f_1}{f_1 + f_2} L_N + \frac{f_1 f_2}{f_1^2 - f_2^2} L_W \tag{3}$$

Where L_N is the float solution for narrow lane ambiguity, L_w is replaced with N_W and then the formula (3) can be expressed as:

$$L_C = \frac{f_1}{f_1 + f_2} L_N + \frac{f_1 f_2}{f_1^2 - f_2^2} N_W \tag{4}$$

$$L_N = N_N + f_{nr} + f_n^S \tag{5}$$

After the wide lane ambiguity is fixed, the float solution for the narrow lane ambiguity can be determined by formula (4), and the narrow-lane ambiguity fixed is achieved by formula (5).

(2) Example analysis

In this paper, the observations of global IGS stations on August 3, 2015 were used to evaluate the performance of the UPD method. We take the standard PPP solution from SHAO station as an example to compare the positioning accuracy of PPP solution which introduced the constraint of UPD ambiguity fixed. Figure 3 gives the time series of narrow-lane UPD estimation sampled at every 30 s, and Fig. 4 shows the positioning errors of the standard PPP and UPD-based PPP solution.

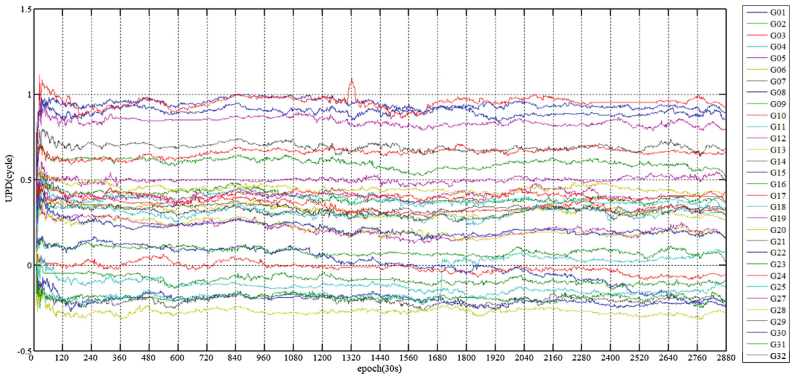


Fig. 3. Time series of narrow-lane UPD estimation

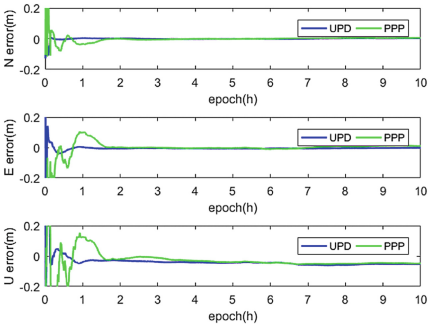


Fig. 4. SHAO station coordinate residuals based on UPD ambiguity fixed constraint and ambiguity float solution

As shown in Fig. 3, the narrow lane UPD is quite stable in a short time, and after stabilization the maximum variation does not exceed 0.15 week, so the ambiguity can be fixed quickly and accurately, and the convergence time is reduced. It can be seen from Fig. 4 that the convergence time of the standard PPP is about 1–2 h, and that of the UPD ambiguity fixed PPP is only tens of minutes or even less, which indicate that

better convergence was achieved with the UPD ambiguity fixed PPP mode and the results of this mode can truly reflect the variation of station position with time. Based on the long-term UPD ambiguity fixed PPP solution, the positioning accuracy in the North, East, Up direction is at mm level.

3 Double Difference Precision Solution

The baseline solutions were first acquired with GAMIT using multi-system observation from the reference network in Australia, and then network adjustment was performed by use of GLOBK. The main strategies in terms of baseline solution are shown in Table 4.

Table 4. GNSS baseline solution strategy

Parameter item	Processing strategy
Satellite cut-off height angle	10°
Solution processing mode orbit relaxation	RELAX
Satellite ephemeris	CODM
Light pressure model	BERNE
Observation mode	LC-AUTCLN
Zenith Delay Correction Model	VMF1
Solid tide model	IERS03
Ocean tide model	FES2004
IGS station prior coordinates	ITRF14

3.1 Comparison of Baseline Solutions

The normalized variance of all daily BDS-based and GPS-based solutions is shown in Fig. 5 The repeat rate of baseline and coordinate repeatability in different modes are counted as shown in Fig. 6.

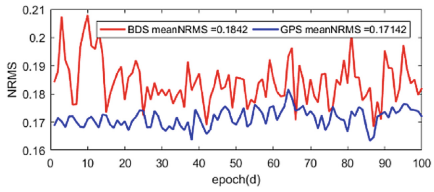


Fig. 5. NRMS of daily BDS-based and GPS-based solutions

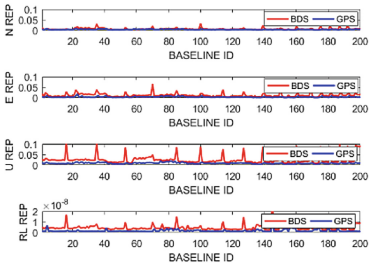


Fig. 6. Repeatability of baseline solutions in different modes

It can be concluded from Fig. 5 that the NRMS of the baseline solutions of BDS and GPS are less than 0.3, which is in good according with the accuracy requirement of single-day data processing. In contrast, the baseline solution of GPS is better than BDS and more stable, the difference between their averages of NRMS is about 0.01.

Table 5. Baseline repeatability in different modes

Mode	N/m	E/m	U/m	RL/10^{-9}
BDS	0.0067	0.0118	0.0260	7.78
GPS	0.0039	0.0039	0.0088	1.62

As presented in Fig. 6, the accuracy of the baseline horizontal component solution of BDS and GPS is significantly better than that of the vertical, where the accuracy of N component is the best, followed by the E component, and that of the U component is the worst. The repeatability accuracy of long baseline is in an order of 10^{-9}, meeting the requirements of the GAMIT baseline solution. The average of the baseline repeatability of 20 stations is counted. As shown in Table 5, it can be concluded that the accuracy of the GPS baseline solution is better than that of the BDS, where the difference of U component is the largest, and that of E component takes second place, followed by that of the N component.

3.2 Adjustment Comparison Analysis

The final solution of 20 stations are obtained by network adjustment using GLOBK. Station coordinates issued in the SOPAC website were used as the reference solution here, and then the coordinate residuals of each station could be generated as shown in Fig. 7. The RMSE of the BDS-based and GPS-based coordinate residuals of these 20 stations is listed in Table 6, respectively.

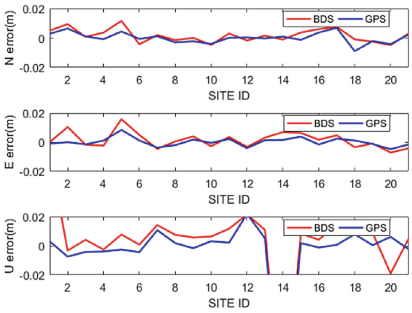


Fig. 7. Coordinate error of the 20 stations

Table 6. RMSE of the 20 stations

Mode	N/mm	E/mm	U/mm
BDS	4.6	5.5	15.4
GPS	3.7	3.0	6.5

We can see from Fig. 7 and Table 6 that the accuracy of BDS and GPS adjustment solutions in horizontal direction is better than that of the vertical. Furthermore, GPS-based positioning accuracy and stability is better than the BDS-based. The positioning

accuracy of the former in horizontal and vertical direction is better than 4 mm and 1 cm, respectively. The positioning accuracy of the latter could reach at the level of 1 cm in horizontal component and 2 cm in vertical component.

4 Comparative Analysis of Different Solution Modes

The time series of the North, East, Up residuals calculated by different solution modes were demonstrated in Fig. 8.

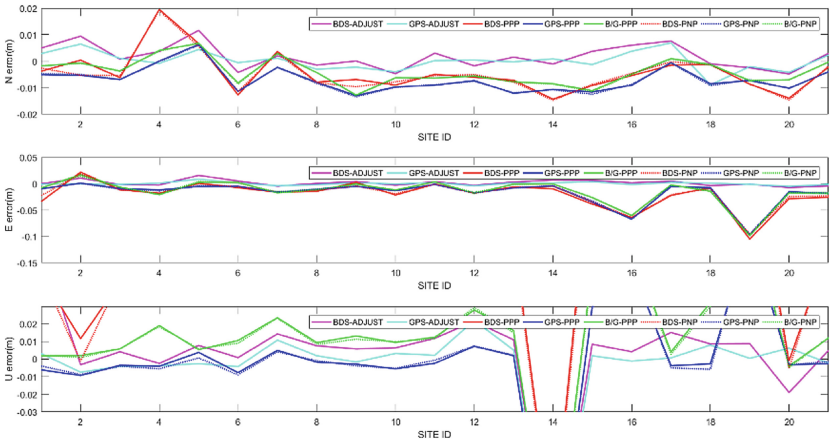


Fig. 8. Errors in different solution modes of 20 stations

It can be concluded from Fig. 8 and Tables 4, 5 and 6 that the positioning accuracy of the BDS/GPS combined positioning mode is roughly the same as that of single GPS, both of which were better than that of the BDS-based. The accuracy of double-difference positioning is better than that of the non-difference positioning, but the positioning accuracy of the double-difference network solution is roughly equivalent to that of the non-difference, where the accuracy could reach at the level of 1 cm. Although BDS-based positioning accuracy is worse than the GPS-based, it could achieve cm-level.

5 Conclusion

With the gradual construction and improvement of Chinese Beidou-3, the BDS/GPS combined positioning will become a development trend in future. Thus, this paper focused on the positioning accuracy evaluation of BDS/GPS navigation position and service performance based on the non/double-difference mode. In order to assess the reliability and stability of the results, we used the 100-day observation of 20 stations (about 2000 observation files). Based on analysis of the experimental results, such following conclusions can be drawn:

(1) With the standard PPP mode, the accuracy of the BDS/GPS combined solution and the GPS-based are roughly equivalent and better than that of the BDS-based. The accuracy of GPS PPP can reach mm level, but that of BDS PPP is at cm level. (2) By the mode of the non-difference network solution, the PPP mode using the double difference ambiguity fixed technique and the UPD ambiguity fixed technique has higher precision and faster convergence than the standard PPP. Compared with double difference baseline solution, although their accuracies were at the same level, the convergence time of the PPP mode using the double difference ambiguity fixed technique and the UPD ambiguity fixed technique were shorter, and the solution using the UPD ambiguity fixed technology also ensures the rigor of the whole network solution. (3) The double-difference precision baseline solution acquired by a conventional high-precision process strategy has strict algorithm and high precision, where the GPS-based positioning accuracy can achieve mm-level and that of the BDS-based is better than 1 cm and 2 cm in horizontal and vertical component, respectively.

At present, due to the imperfect construction of the BDS, its overall positioning accuracy and service performance is slightly inferior to that of the GPS, but the advantages of the BDS have gradually emerged. In addition that BDS positioning accuracy is continuously improved, the BDS/GPS combined positioning accuracy can also meet the requirments. It can be foreseen that when the construction of the Beidou system is completed, the positioning accuracy could be roughly equal to that of GPS. In some areas, the positioning accuracy of Beidou can be even better, and the reliability and stability will be higher.

References

1. Montenbruck O, Hauschild A et al (2013) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. GPS Solutions 17:211–222
2. Zhang X, Wu M, Liu W et al (2017) Initial assessment of the COMPASS/BeiDou-3: new-generation navigation signals. J Geodesy 91(10):1225–1240
3. Li X, Ge M, Douša J et al (2014) Real-time precise point positioning regional augmentation for large GPS reference networks. GPS Solutions 18(1):61–71
4. Chen H, Jiang W, Ge M et al (2014) An enhanced strategy for GNSS data processing of massive networks. J Geodesy 88(9):857–867
5. Ge M, Gendt G, Dick G, Zhang PF, Rothacher M (2006) A new data processing strategy for huge GNSS global networks. J Geodesy 80:199–203. https://doi.org/10.1007/s00190-006-0044-x
6. Blewitt G (2010) Fixed point theorems of GPS carrier phase ambiguity resolution and their application to massive network processing: Ambizap. J Geophys Res Solid Earth 113 (B12):36–44
7. Geng J, Meng X, Dodson AH et al (2010) Integer ambiguity resolution in precise point positioning: method comparison. J Geodesy 84(9):569–581
8. Wang H, Li J, Dang Y et al (2016) PPP network solution UPD ambiguity fixed technique to monitor the earthquake coseismic displacement of the Ms8.1 earthquake in Nepal and the surrounding areas in China. J Surv Mapp 45(S2):147–155
9. Chen H (2015) GNSS unified fast and precise data processing method based on original observations. Wuhan University
10. Wei E, Liu X, Wang L, Liu J (2018) Analysis and evaluation of BDS/GPS combined precision single point positioning accuracy. J Wuhan Univ 43(11):1654–1660
11. Ye S, Xia F, Zhao L et al (2017) Impact analysis of yaw attitude on BDS precision point positioning. Acta Geochim Sinica 46(08):971–977

Trajectory Optimization of LiDAR SLAM Based on Local Pose Graph

Chunxu Chen[1], Ling Pei[1(✉)], Changqing Xu[1], Danping Zou[1], Yuhui Qi[2], Yifan Zhu[1], and Tao Li[1]

[1] Shanghai Key Laboratory of Navigation and Location-Based Services, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
ling.pei@sjtu.edu.cn

[2] School of Information Engineering, Nanchang University, Nanchang 330031, China

Abstract. Simultaneous localization and Mapping (SLAM) is one of the key technologies for autonomous navigation of mobile robots. In recent years, researchers have studied diverse sensors and proposed heterogeneous SLAM algorithms. Owing to the high precision and strong anti-interference ability of LiDAR, SLAM algorithm based on laser sensor has been widely studied and applied for autonomous navigation and 3D reconstruction. This paper studies the three-dimensional localization and mapping based on laser sensor. Aiming at the problem of trajectory accumulated error drift caused by sequential registration in LiDAR odometry, an optimization method based on local pose graph is proposed. Firstly, the initial pose estimation of the robot and corrected point cloud are figured out by a LiDAR odometry algorithm. Then, an omnidirectional local map is constructed by using the corrected point cloud, and the point cloud of current frame is registered with the omnidirectional local map to obtain more precise pose. Finally, the ICP algorithm is used to compute the frame to frame transformation which is used for local pose graph, the pose graph based on the G2O frame is designed to realize the further correction of the robot trajectory. The paper carried out four experiments on the KITTI dataset and the field test dataset, experimental results show that the loop-closure error of the proposed algorithm is reduced to 0.01%, the minimum relative error is reduced to 0.49% and the relative error of all scenarios is about 1/2 of that of LOAM which is known as a state-of-the-art LiDAR SLAM algorithm.

Keywords: LiDAR · SLAM · Pose graph · Pose estimation · Graph optimization

1 Introduction

In recent years, navigation has become one of the important functions of robot. Satellite positioning as a mature positioning technology is able to obtain good results in many conditions. However, it fails or has poor performances while inside or between buildings. Thus, robots need other sensors to realize positioning and navigation in all scenarios.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 360–370, 2019.
https://doi.org/10.1007/978-981-13-7751-8_36

Simultaneous Localization and Mapping (SLAM) [1] can provide real-time positioning and mapping of unknown environments, and then realize path planning and autonomous navigation. At present, the perceptual sensors used in mobile robots are mainly divided into two categories: LiDAR sensors and visual sensors. Therefore, the methods of SLAM can be divided into visual SLAM [2] and LiDAR SLAM [3]. Compared with visual sensors, LiDAR sensors have the advantage of being less subject to environmental interference and directly providing point clouds data describing the geometric information of the environment. LiDAR sensors can be divided into single-line LiDAR and multi-line LiDAR. Generally, the point clouds directly acquired by single-line LiDAR are two-dimensional data. The point clouds in each frame are too sparse to fully describe the environmental scene information. The multi-line LiDAR can acquire 3D spatial point clouds with rich information, which can be used to construct high-quality 3D maps. Therefore, 3D laser SLAM technology based on multi-line LiDAR sensors is widely used in indoor navigation [4], unmanned vehicle and other fields.

The early studies in SLAM realized the pose estimation of robots by using filter approach, such as Extended Kalman filter (EKF) and particle filter (PF). Later, in order to adapt to large-scale and unstructured environments, the method based on filtering is gradually replaced by Graph-based optimization method [5].

The SLAM method based on graph optimization was first proposed by Lu and Milis [6]. This paper describes the basic idea of SLAM technology based on graph optimization. The SLAM problem based on graph optimization can be regarded as a least squares problem in essence. Its modeling methods are commonly used in dynamic Bayesian networks and factor graphs [7, 8]. KartoSLAM [9] and LagoSLAM [10] are 2D laser SLAM algorithms based on graph optimization. The popular ORB-SLAM [11] and ORB-SLAM2 [12] in visual SLAM also adopt graph optimization methods to improve accuracy and robustness.

3D laser SLAM acquires the frame-to-frame transformation matrix by registration of point cloud from two frames, and obtains the trajectory of the robot, and builds the map on this basis. Trajectory estimation and mapping are achieved by sequential registration [13]. Errors accumulate gradually over time and cannot be eliminated, which eventually lead to the failure of the algorithm.

To solve the above problems, this paper constructs a local map for registration and adds a local pose map to reduce the cumulative error. Figure 1 shows the flow of the algorithm. The algorithm consists of three parts: LiDAR odometry, LiDAR mapping and local pose graph.

The main contributions of this paper are as follows:

(1) Build an omnidirectional map for point cloud registration and obtain more precise pose;
(2) Construct local pose graph with keyframes for further trajectory optimization.

Experimental results show that compared the relative accuracy of the proposed algorithm is around 1/2 of that of the state-of-the-art open source algorithm LOAM.

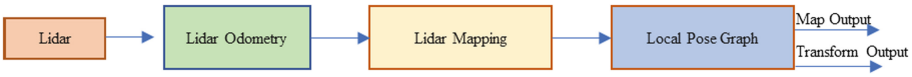


Fig. 1. The block of system overview.

2 Pose Estimation Based on LiDAR Odometry

LiDAR Odometry implements the point cloud sequential registration method to estimate the robot pose. First, the raw point clouds are preprocessed. It is used to extract some points from raw point cloud for point cloud registration, including point clouds denoising, point cloud filtering and feature point extraction.

This paper selects the points used for registration according to the smoothness of the point. The smoothness of the points are evaluated according to Eq. (1) and the edge points and plane points are extracted from the raw point clouds. In Eq. (1), c is the evaluation index of the point clouds' smoothness, s is the number of point clouds selected from the raw points, and r_i is the distance from the i-th point to the laser sensor. This paper sets s to 10. When c is below the threshold c_{th}, the point is marked as a planar point, while c is greater than the threshold, it is marked as an edge point. The default value of c_{th} is 0.1.

$$c = \frac{1}{|s| \cdot \|r_i\|} \left\| \sum_{j \in s, j \neq i} (r_i - r_j) \right\| \tag{1}$$

Then the filter points selected from the adjacent two frames is registered, and the error function is as shown in Eqs. (2) and (3).

$$d_e = \frac{\left|(p_i - q_i) \times (p_i - q_j)\right|}{\left|q_i - q_j\right|} \tag{2}$$

$$d_p = \frac{\left| \begin{matrix} (p_j - q_i) \\ (q_i - q_l) \times (q_i - q_m) \end{matrix} \right|}{\left|(q_i - q_l) \times (q_i - q_m)\right|} \tag{3}$$

Equations (2) and (3) are variants of the ICP algorithm. The distance error function is expressed by the distance d_e from the edge point p_i to the straight line and the distance d_e from the plane point p_j to the plane, where q_i q_j q_m are searched from previous point clouds, which are the nearest neighbors of the point in the current frame. Then the objective function can be written as Eq. (4)

$$\begin{cases} p_i' = Rp_i + t \\ p_j' = Rp_j + t \\ \min \sum_{i=1}^{M} \left\| d_e(p_i') \right\|_2^2 + \sum_{i=1}^{N} \left\| d_P(p_i') \right\|_2^2 \end{cases} \tag{4}$$

Where M is the number of edge points of the current frame, and N is the number of plane points. The objective function can be solved by Levenberg-Marquardt method, and the initial pose estimation of the robot is obtained.

The point clouds data of each frame are acquired at different times, so they cannot be considered to be in the same coordinate systems when the robot moves too fast.

Assuming that the start moment of the scan in the k-th frame is t_k and the transformation between the moment $t(t > t_k)$ and t_k is $T_k^L(t)$. The transformation $T_{(k,i)}^L$ between intermediate moment $t_{(k,i)}$ and t can be calculated as Eq. (5). The point clouds from the k-th scan are transformed to the local coordinate system at the moment t_k, more details can be found in [14].

$$T_{(k,i)}^L = \frac{t_{(k,i)} - t_k}{t - t_k} T_k^L(t) \tag{5}$$

The original point clouds are corrected by the pose of robot utilizing LiDAR odometry, at the same time, the corrected point clouds and the pose estimation of robot are transmitted to the LiDAR Mapping.

3 Omnidirectional Local Map Construction

LiDAR Odometry is coarsely registered using sparse point clouds. The trajectory is further improved by map registration based on LiDAR Odometry.

This paper use point clouds of surrounding around the current frame to construct local map as the horizontal field of view (FOV) of LiDAR sensor is 360° in one scan. As Fig. 2 shows, triangles represent pose of the robot, the red triangle represent current robot, blue triangles represent robots in the searching field with the searching radius R. Yellow regions with different shapes represent point clouds observed at corresponding

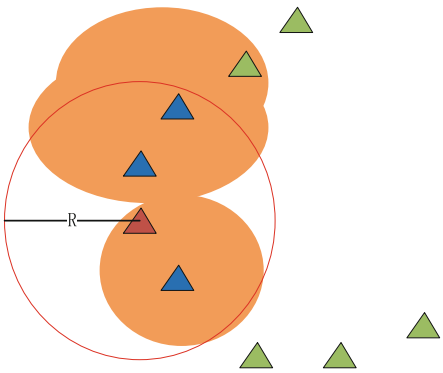


Fig. 2. Omnidirectional Local Map construction.

blue positions. The local map is represented by all yellow regions. In order to speed up searching, the K-D tree algorithm is adopted for neighbor search.

The current frame is register with local map using method same as LiDAR Odometry and then the more accurate poses obtained by registration will be put into pose graph for further correction.

4 Pose Graph Optimization

In the sequential registration SLAM algorithm, the cumulative error cannot be eliminated with distance and time, so in the large scene or complex environment, the positioning result is not accurate. In order to reduce the error, the pose can be further corrected by adding a constraint relationship between poses. The classical graph optimization SLAM problem can be expressed by the factor graph. As shown in Fig. 3, the large node represents the pose of the robot, circles are the nodes of the graph, the lines represent the transformation between the poses, and represent the edges of the graph. The black lines are the constraint generated by the sequential registration in SLAM, and the red line is the constraint provided by other methods such as loop-closure detection and registration between keyframes. Ω_{ij} contains the weight information of the error. The error function of the above pose map can be written into Eq. (6)

$$F(x) = \sum_{<i,j> \in C} e(X_i, X_j, Z_{ij})^T \Omega_{ij} e(X_i, X_j, Z_{ij})$$

$$X^* = \arg\min_X F(X) \tag{6}$$

where Z_{ij} is the edge and e is an error relationship between the node and the edge. According to Eq. (6), graph optimization can be seen as a non-linear least squares problem essentially. Then a first-order Taylor expansion is performed on the objective function and it can be solved by methods such as Gauss-Newton method or Levenberg-Marquardt (L-M). More derivation details can be found in reference [15].

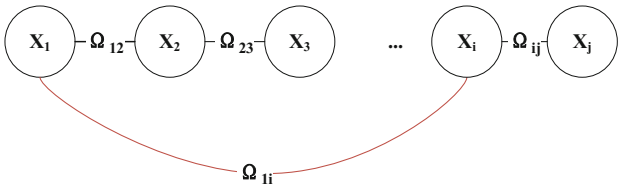


Fig. 3. The model of pose graph in SLAM.

In this paper, the optimized pose of the local map registration is defined as a node. The keyframes are selected from the optimized poses, and the point clouds of the current keyframe C are registered with the point clouds of the previous keyframe P by the ICP algorithm. The registration result is transformation matrix T and the error function e can be written as Eq. (7):

$$e = T_p^{-1} T_c \ominus \mathrm{T} \tag{7}$$

T_p: The six-dimensional pose of robot at the previous moment;
T_c: The six-dimensional pose of robot at the current moment;
$\ominus$: Subtraction of six-dimensional pose;

The information matrix Ω between keyframes is calculated according to the ICP [16] registration score s, and I is the identity matrix.

$$\Omega = S \times I \tag{8}$$

The defined nodes and edges are placed in the pose graph based on the G2O framework for optimization and the optimized robot pose and map are output. Local pose graph is used to optimize the algorithm to ensure the real-time performance. The number of nodes is set 2 and no loop-closure detection algorithm is added. The specific process is shown in Fig. 4.

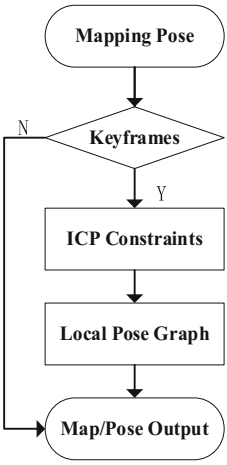


Fig. 4. Flow of Algorithm in pose graph.

5 Test Verification

This paper carried out four experiments on indoor and outdoor datasets and analyzed the performance of two competing methods, LOAM and the proposed method.

5.1 Experimental Platform

In order to evaluate the performance of the proposed method, we designed a portable sensor integration platform. As shown in Fig. 5, it is equipped with Velodyne VLP-16. The measurement range of VLP-16 is up to 100 m with the accuracy of ± 3 cm. It has a vertical FOV of 30°(±15°). The vertical angular resolution is 2° and the horizontal

angular resolution varies from 0.1° to 0.4°. As a multi-sensor platform, it is also equipped with other sensors, such as IMU and camera. In this paper, only VLP-16 data is applied in the proposed approach. In addition, the experimental platform uses Core i7-6500U low-voltage processor.

Fig. 5. Experimental platform.

5.2 Experimental Scenarios

The proposed method and LOAM were tested by four data sets. The experiments 1, 2 and 3 were carried on the field test dataset with the experimental platform. Experiment 4 was carried out on KITTI [17] outdoor data with offline processing at a rate of 0.2 due to the large amount of point clouds from Velodyne HDL-64E.

The experiment 1 was carried out in our laboratory. The experiment 2 was carried out in the same environment as experiment 1. The difference is the number of loops that we walked around the laboratory, which in the experiment 1 is one and in the experiment 2 is three.

The route of experiment 3 was walking one loop from indoor to outdoor.

Figure 6 shows CAD and the picture of the laboratory.

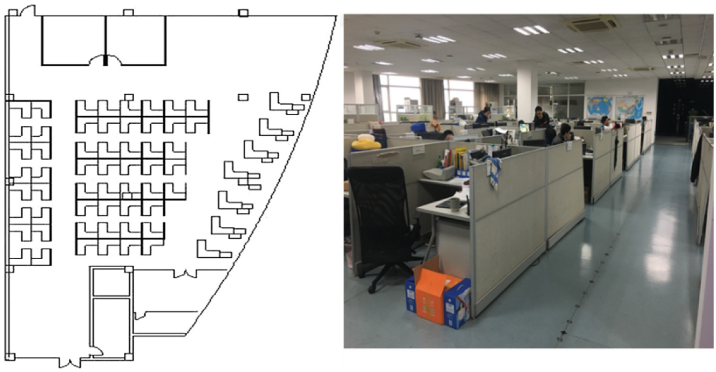


Fig. 6. The CAD and real map of Lab.

5.3 Experimental Results and Analysis

The loop-closure error and relative error are used to evaluate the performance of the algorithm. The relative error is defined as the ratio of the measured drift to the distance.

Figure 7 shows the comparison of trajectory and the ground truth from four experiments for two methods respectively, where red trajectory is the ground truth, the green one is the trajectory of LOAM, and the blue one is the trajectory figured out by the proposed method. Figure 7(a), (b), (c) and (d) correspond to the trajectories from experiments 1, 2, 3, and 4 respectively.

Figure 7(a) shows the performance of the two methods are equivalent in short distance. In long distance, when LOAM has a matching error (highlighted in red circle in Fig. 7(b)), it cannot be corrected for its property of sequential registration. The proposed method utilizes ICP to add constraint between the keyframes and corrects the mismatch. As a result, the drift of LOAM is larger while the pose estimated by the proposed method is more closed to ground truth in Fig. 7(b).

Figure 7(c) shows the closed-loop trajectory from indoor to outdoor. The loop-closure error of LOAM shown in Table 1 is smaller while the relative error is almost six times of the proposed method in experiment 3. Figure 8 shows the outdoor trajectory in experiment 3. Compared with the indoor positioning result, the two methods work worse in outdoor. The main reasons are as follows: First, the outdoor

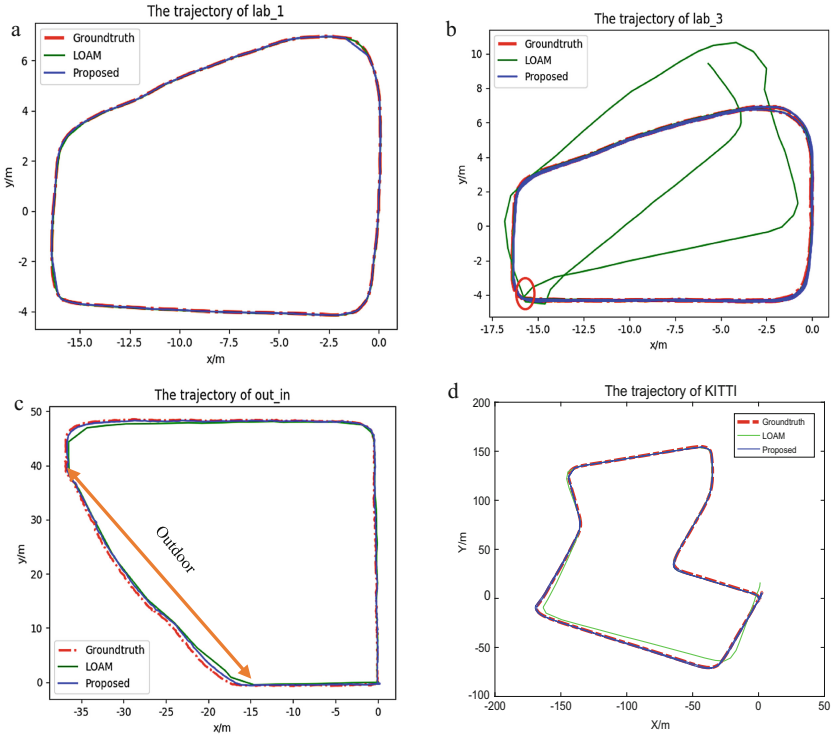


Fig. 7. The comparison of trajectories in different scenarios by two methods.

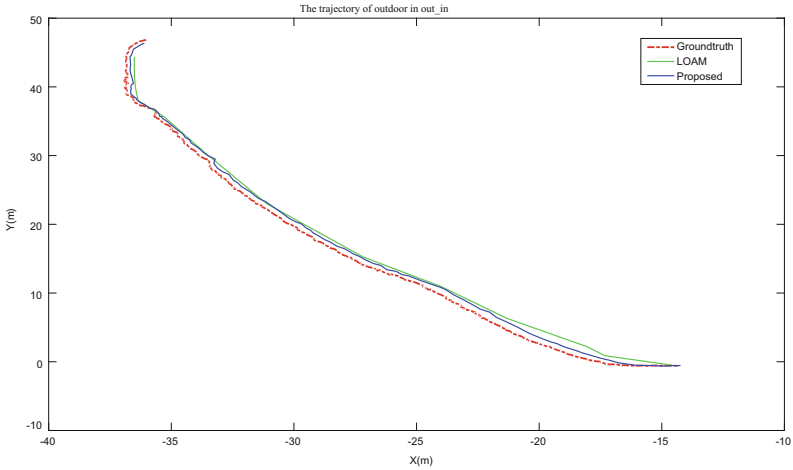


Fig. 8. The trajectory of outdoor in experiment 3 by two methods.

Table 1. Comparison of algorithm in loop-closure error and relative error

Experiments/Mileage		1/47(m)	2/141(m)	3/175(m)	4/750(m)
Loop-closure error	LOAM	0.01%	7.30%	0.01%	1.40%
	Proposed	0.01%	0.01%	0.05%	0.03%
Relative error	LOAM	0.67%	19.46%	6.35%	3.87%
	Proposed	0.49%	0.11%	1.07%	1.60%

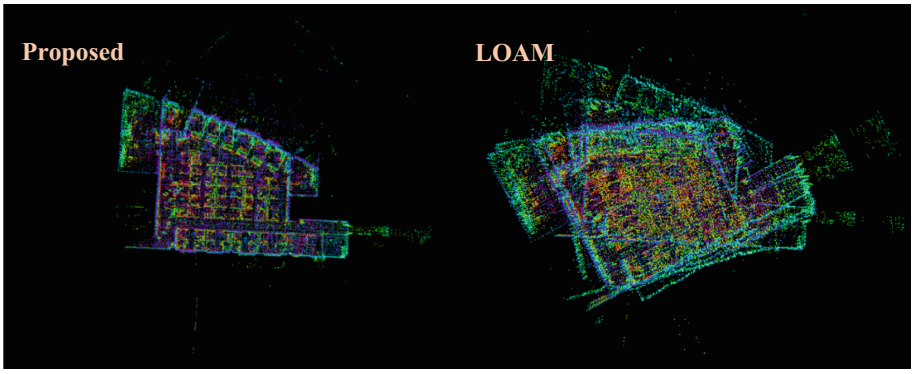


Fig. 9. The mapping built by two methods.

environment is relatively empty compared to the indoor environment, and there are interference factors such as trees or lawn which make the outdoor structural features are not as clear as the indoor structural features. Second, the density of point clouds

decreases with the increase of the obstacle distance, and the information collection for the environment is not comprehensive.

Figure 7(d) shows the results on KITTI outdoor data sets. After half of the trajectory, the cumulative error increases which causes the rest trajectory estimated by LOAM shifting.

This paper analyzed the relative error and loop-closure errors in different scenarios listed in Table 1. As the table shown, the loop error of the proposed algorithm is around 0.03% in each scene, and the minimum relative error can reach 0.49%. The relative error of all scenarios is 50% lower than that of LOAM. These results show that the proposed method performs better and more robust in long-distance scenarios.

The mapping results of SLAM largely depends on the effect of positioning. Figure 9 shows the comparison of the maps built by the two methods in experiment 2. From the mapping results, the performance of proposed method is better than that of LOAM.

6 Conclusion

This paper proposed an optimization method based on local pose graph for the problem of trajectory error accumulation caused by sequential registration in laser SLAM. By constructing an omnidirectional local map and constructing a local pose map for optimization, the trajectory accumulation error is reduced and the performance of the algorithm is improved. The experimental results show that under the same computational performance, the loop-closure error of the proposed algorithm is reduced to 0.01%, and the relative error is about 1/2 of that of LOAM. Although the robustness of the improved algorithm is improved, the performance of the algorithm in outdoor and structured similar scenarios needs to be improved. In the future work, the algorithm's computational efficiency and robustness can be improved by fast point cloud segmentation algorithm [18] and multi-source fusion method.

Acknowledgement. This work was supported in part by the National Nature Science Foundation of China under Grant 61873163 and in part by the Shanghai Science and Technology Committee under Grant 17DZ1100803.

References

1. Chen Y, Tang J, Jiang C et al (2018) The accuracy comparison of three simultaneous localization and mapping (SLAM)-based indoor mapping technologies. Sensors 18(10). https://doi.org/10.3390/s18103228
2. Zhou H, Zou D, Pei L et al (2015) Structslam: visual slam with building structure lines. IEEE Trans Veh Technol 64(4):1364–1375
3. Zhang J, Singh S (2014) LOAM: LiDAR odometry and mapping in realtime. In: Robotics: science and systems conference (RSS)
4. Pei L, Liu J et al (2012) Using LS-SVM based motion recognition for smartphone indoor wireless positioning. Sensors 12(5):6155–6175

5. Thrun S, Montemerlo M (2006) The GRAPH SLAM algorithm with applications to large-scale mapping of urban structures. Int J Robot Res 25(5):403–429
6. Lu F, Milios E (1997) Globally consistent range scan alignment for environment mapping. Auton Robots 4(4):333–349
7. Kschischang FR, Frey BJ, Loeliger HA (2001) Factor graphs and the sum-product algorithm. IEEE Trans Inf Theor 47(2):498–519
8. Gong Z, Pei L, Zou D et al (2016) Graphical approach for MAV sensors fusion. In: International technical meeting of the satellite division of the institute of navigation
9. Konolige K, Grisetti G et al (2010) Efficient sparse pose adjustment for 2D mapping. In: 2010 IEEE/RSJ international conference on intelligent robots and systems. IEEE
10. Carlone L, Aragues R (2011) A linear approximation for graph-based simultaneous localization and mapping. In: Robotics: science & systems VII, University of Southern California, Los Angeles, CA, USA, June 2011. DBLP
11. Mur-Artal R, Montiel JMM et al (2015) ORB-SLAM: a versatile and accurate monocular slam system. IEEE Trans Robot 31(5):1147–1163
12. Mur-Artal R, Tardos JD (2016) ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. IEEE Trans Robot 33(5):1255–1262
13. Liang M, Min H, Luo R (2013) Summary of simultaneous localization and map creation based on graph optimization. Robot 35(04):500–512
14. Zhang J, Singh S (2017) Low-drift and real-time LiDAR odometry and mapping. Auton Robots 41(2):401–416
15. Kummerle R, Grisetti G, Strasdat H et al (2011) g2o: a general frame-work for graph optimization. In: IEEE International Conference on Robotics and Automation
16. Besl PJ, Mckay ND et al (1992) A method for registration of 3-d shapes. IEEE Trans Pattern Anal Mach Intell 14(2):239–256
17. Andreas G, Philip L, Christoph S et al (2013) Vision meets robotics: the KITTI dataset. Int J Robot Res (IJRR). http://www.cvlibs.net/datasets/kitti/eval_object.php
18. Bogoslavskyi I, Stachniss C (2016) Fast range image-based segmentation of sparse 3D laser scans for online operation. In: Proceedings of the IEEE/RSJ international conference on intelligent robots and systems

Analysis of Velocity Estimation Methods Based on BDS PPP

Shunli Duan[1], Wei Sun[1], Junbo Shi[2,3(✉)], and Chenhao Ouyang[2]

[1] School of Geomatics, Liaoning Technical University, Fuxin, Liaoning, China
[2] School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei, China
jbshi@sgg.whu.edu.cn
[3] Key Laboratory of Precise Engineering and Industry Surveying, National Administration of Surveying, Mapping and Geoinformation, Wuhan, Hubei, China

Abstract. Position difference, Doppler observation and time-differenced carrier phase (TDCP) are three commonly used velocity estimation methods. Since the average velocity of two epochs is obtained both in the position difference and TDCP velocity estimation, the accuracy is affected by the carrier's kinematics, large deviation will appear in high dynamic velocity estimation field. While the carrier's real velocity can be obtained in real time utilize the Doppler velocity estimation method, consequently, this method is widely adopted in navigation fields such as vehicle, shipboard and airborne. The accuracy of Doppler velocity estimation will be affected by positioning error through the direction cosine matrix between the receiver and satellite, therefore, high precision position information is the premise to obtain the accurate carrier velocity. The principle of heterogeneous constellations based BDS PPP Doppler velocity estimation is proposed in this paper, and the influence of various errors on the velocity estimation accuracy is analyzed. Based on the high precision position estimated by BDS PPP, the Doppler velocity estimation method is analyzed. The static experimental results show that the velocity estimation accuracy based on BDS PPP position difference is mm/s level, while the Doppler velocity estimation accuracy is cm/s level. In the vehicle dynamic experiment, the velocity estimation results are compared with high precision GNSS/INS integrated navigation results. The results show that the coincidence between the BDS PPP Doppler velocity estimation results and the integrated navigation is cm/s level, while the BDS PPP position difference based velocity estimation is significantly affected by carrier's kinematics, and the coincidence with the integrated navigation results is dm/s level. The Doppler velocity estimation method is recommended in high dynamic applications, the position difference velocity estimation is recommended in low dynamic or low acceleration situations.

Keywords: BDS · PPP · Doppler velocity estimation · Position differential velocity estimation

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 371–379, 2019.
https://doi.org/10.1007/978-981-13-7751-8_37

1 Introduction

BeiDou Navigation Satellite System (BDS) is a self-constructed and independently operated satellite navigation system in China. Its positioning accuracy is better than 10 m, the velocity estimation is better than 0.2 m/s, and the timing is better than 50 ns [1]. High-precision velocity information is especially important in many areas such as autonomous driving [2], seismic monitoring [3], and GNSS/INS integrated navigation [4].

As for the research on BDS positioning method and precision [5], domestic and foreign scholars have done relatively few studies on BDS velocity estimation method and precision research. The satellite velocity estimation methods mainly include position differential velocity estimation, Doppler velocity estimation and TDCP velocity estimation [6, 7]. Position difference method and TDCP velocity estimation method obtain the average velocity between epochs, and the accuracy is subject to carrier motion. The influence of the state is large and cannot meet the accuracy requirements of the high maneuvering velocity estimation field. The Doppler velocity estimation method calculates the instantaneous velocity of the carrier by the original observation of the Doppler shift obtained by the receiver tracking loop in real time, of which the result is not affected by the motion state of the carrier.

In this paper, the principle of position difference and Doppler velocity estimation based on BDS PPP is given. The influence of various errors on the accuracy of velocity estimation and the correction method in Doppler velocity estimation method are analyzed. The two methods were analyzed by taking advantage of two sets of experimental data.

2 Principle and Error Analysis of BDS PPP Velocity Estimation

2.1 Principle of Position Differential Velocity Estimation

Assuming that the receiver sampling interval is Δt, the velocity of the receiver at time t can be expressed as

$$\mathbf{V}_t = \frac{\mathbf{P}_{t+\Delta t} - \mathbf{P}_{t-\Delta t}}{2\Delta t} \tag{1}$$

Where $\mathbf{P}$ is the position vector of the receiver. The velocity obtained by the above formula is the average between two epochs, and if Δt approaches zero, it is the instantaneous velocity.

2.2 Principle of Doppler Velocity Estimation

The linearized Doppler velocity equation can be expressed as:

$$\lambda D_k^j = \mathbf{e}_k^j\left(\mathbf{v}^j - \mathbf{v}_k\right) + c\left(d\dot{\tau}_k - d\dot{\tau}^j\right) + \dot{I} + \dot{T} + \varepsilon \tag{2}$$

where λ is the wavelength and BDS B1 carrier is 0.192 m. D_k^j is the original Doppler shift obtained by the receiver, $\mathbf{v}^j$ stands for the satellite velocity vector, $\mathbf{v}_k$ for the

receiver velocity vector, c is the velocity of light in a vacuum, $d\dot{\tau}_k$ and $d\dot{\tau}^j$ represent the receiver clock drift and the satellite clock drift, respectively. $\dot{I}$ is the ionospheric delay rate, $\dot{T}$ is the tropospheric delay rate, ε is multipath and receiver noise, $\mathbf{e}_k^j$ represents the directional cosine vector between the receiver and satellite and is given by:

$$\mathbf{e}_k^j = \frac{(\mathbf{r}^j - \mathbf{r}_k)}{\|\mathbf{r}^j - \mathbf{r}_k\|} \tag{3}$$

where $\mathbf{r}^j$ and $\mathbf{r}_k$ represent the satellite and receiver positions, respectively.

It can be seen from Eqs. (1) and (2) that the three-dimensional velocity of the receiver and the receiver clock rate variability are four unknowns, and at least four satellites can be solved, and four or more satellites are solved by a least squares algorithm.

2.3 Error Analysis

The accuracy of the position differential velocity estimation method is mainly related to the positioning error and the motion state of the carrier. The Doppler velocity estimation method is mainly affected by satellite position and velocity error, receiver position error, ionospheric and tropospheric delay variation rate, relativistic effect, the satellite clock speed and the observation noise. The satellite clock difference can be corrected by broadcasting ephemeris parameters. Since the velocity estimation is completed in a short time, the influence of the rate of change of the ionosphere and the troposphere on the velocity estimation can be negligible.

2.3.1 Satellite Position and Velocity Error

The satellite position and velocity affect the velocity estimation result by the cosine vector of the station star direction. If the error of the BDS satellite position in the broadcast ephemeris is 36 m, the IGSO and the MEO are 10 m, the influence on the velocity estimation is 1–2 mm/s. Precise ephemeris can be used for calculation in post processing. The velocity of the broadcast ephemeris calculation satellite and the velocity of the precision ephemeris are within 1 mm/s, so the velocity of the satellite can be calculated in real time using the broadcast ephemeris.

Unlike GPS satellites, BDS satellites are heterogeneous constellations, including MEO, GEO, and IGSO. The calculation method of MEO and IGSO satellite velocity is the same as that of GPS satellite. The specific calculation formula can refer to the literature [8], and the GEO satellite orbital inclination is close to 0 degree. It needs to be calculated by coordinate rotation method. The specific calculation method is as follows. The meaning of each symbol can be referred to the literature [9].

(1) Near-point angle change rate

$$\dot{E}_k = n/(1 - e\cos E_k) \tag{4}$$

(2) Rate of change of the pitch angle of the ascending node

$$\dot{\Phi}_k = \frac{\sqrt{1-e^2}\dot{E}_k}{1-e\cos E_k} \tag{5}$$

(3) Corrected latitude change rate, corrected distance change rate, corrected orbital inclination change rate

$$\begin{cases} \dot{u}_k = (1+2(C_{us}\cos 2\Phi_k - C_{uc}\sin 2\Phi_k))\cdot\dot{\Phi}_k \\ \dot{r}_k = ae\ \sin E_k\cdot\dot{E}_k + 2(C_{rs}\cos 2\Phi_k - C_{rc}\sin 2\Phi_k)\dot{\Phi}_k \\ \dot{i}_k = 2(C_{is}\cos 2\Phi_k - C_{ic}\sin 2\Phi_k)\dot{\Phi}_k + I \end{cases} \tag{6}$$

(4) Ascending node longitude change rate

$$\dot{\Omega}_k = \dot{\Omega} \tag{7}$$

(5) The velocity of the satellite in the orbital plane coordinate system

$$\begin{cases} \dot{x}_k = \dot{r}_k\cos u_k - r_k\sin u_k\cdot\dot{u}_k \\ \dot{y}_k = \dot{r}_k\sin u_k + r_k\cos u_k\cdot\dot{u}_k \end{cases} \tag{8}$$

(6) The velocity of the satellite in the geocentric fixed-angle coordinate system

$$\begin{bmatrix} \dot{X}_k \\ \dot{Y}_k \\ \dot{Z}_k \end{bmatrix} = \dot{R}_Z(\omega_e t_k)R_X(-5^\circ)\begin{bmatrix} X_{GK} \\ Y_{GK} \\ Z_{GK} \end{bmatrix} + R_Z(\omega_e t_k)R_X(-5^\circ)\begin{bmatrix} \dot{X}_{GK} \\ \dot{Y}_{GK} \\ \dot{Z}_{GK} \end{bmatrix} \tag{9}$$

Where

$$\begin{bmatrix} X_{GK} \\ Y_{GK} \\ Z_{GK} \end{bmatrix} = \begin{bmatrix} x_k\cos\Omega_k - y_k\cos i_k\sin\Omega_k \\ x_k\sin\Omega_k + y_k\cos i_k\cos\Omega_k \\ y_k\sin i_k \end{bmatrix} \tag{10}$$

$$R_Z(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{11}$$

$$R_X(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix} \tag{12}$$

$\dot{\mathbf{X}}$ is the derivative of $\mathbf{X}$.

2.3.2 Receiver Position Error

The receiver position error further affects the velocity estimation result by affecting the cosine vector of the station star direction. Using static data for simulation calculation, when the positioning error is 50 m, the effect on velocity can reach 1 cm/s. At present, the BDS single point positioning error is within 20 m, and the influence on the velocity estimation is on the order of mm/s. If the BDS PPP position result is used, the influence on the velocity estimation can be ignored.

2.3.3 Relativistic Effects

The relativistic effect is caused by the difference in the velocity of movement of the satellite clock and the receiver clock in the inertial space and the difference in gravity of the Earth [8]. The ranging error caused by relativity is:

$$\Delta\rho = -\frac{2\sqrt{a\mu}}{c}e\ \sin\ E \tag{13}$$

Where a is the long axis of the satellite orbit; μ is the Earth's gravitational constant, and its value is $\mu = 3.986004418 \times 10^{14} m^3/s^2$; e is the satellite orbital eccentricity; E is the near-point angle. The equation for the influence of relativity on the velocity estimation can be obtained from the above equation:

$$\Delta\dot{\rho} = -\frac{2\sqrt{a\mu}}{c}e\ \cos E\frac{dE}{dt} \tag{14}$$

3 Analysis of the Experiment and Results

3.1 Static Experiment

The static experiment is 50 min data collected on June 12, 2018 in an open place in Liaoyang City, Liaoning Province, with a sampling interval of 5 s. The average PDOP value is 2.83, and the number of satellites is 10. The true value of the velocity under static conditions is 0, and the calculated velocity value is the error. Figure 1 shows the error of the BDS PPP positioning result, and the positioning accuracy after convergence can reach cm level. Figures 2 and 3 show the BDS PPP position differential velocity estimation error and PPP-based Doppler velocity estimation error. Table 1 shows the error statistics of the two methods. It can be seen that the BDS PPP position differential velocity estimation method can achieve a velocity estimation accuracy of mm/s after the convergence of the positioning results. The accuracy of the Doppler velocity estimation method based on BDS PPP is cm/s, and the elevation direction is slightly worse than the horizontal direction. The relationship between the spatial positional accuracy factor and the three-dimensional velocity accuracy can be expressed as:

$$M_v = PDOP \cdot M_\rho \tag{14}$$

In the formula, M_v is the velocity estimation error; M_ρ is the sum of the errors of the error to the pseudo-range rate, and M_ρ can be calculated as 1.9 cm/s, indicating that the Doppler observation noise of this type of receiver is about 1.9 cm/s.

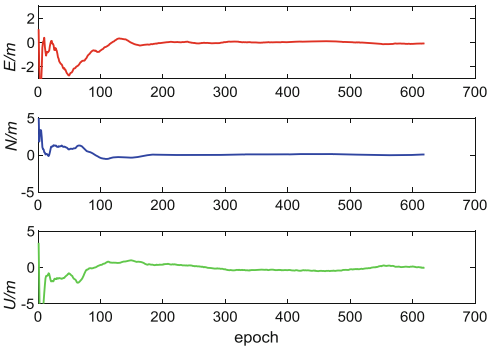


Fig. 1. BDS PPP positioning error

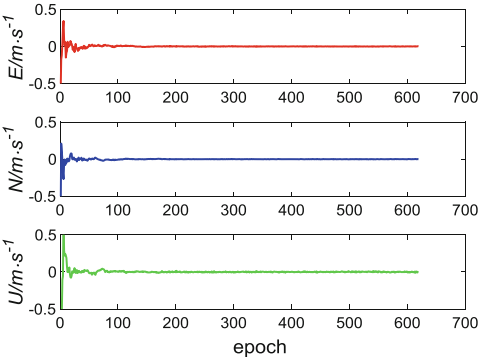


Fig. 2. BDS PPP position differential velocity estimation error

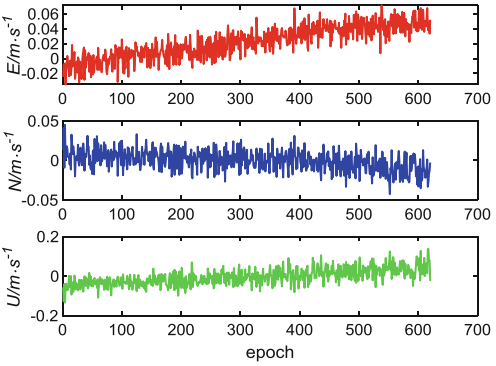


Fig. 3. Doppler velocity estimation error based on BDS PPP

Table 1. Error statistics

	RMS (cm/s)		
	E	*N*	*U*
Position difference method	0.26	0.17	0.48
Doppler method	3.10	1.31	4.19

3.2 Dynamic Experiment

The land vehicle experiment was carried out in Fuxin City on 2017-06-07. The data acquisition time was 1 h 30 min and the data recording frequency was 1 Hz. Using the velocity result of the post-processing of the MP-POS520 integrated navigation device as a reference, the velocity estimation accuracy is 0.02 m/s. The experimental environment and trajectory are shown in Fig. 4. Figures 5 and 6 are the horizontal comparison of BDS PPP position differential velocity estimation (PPPvel), BDS PPP-based Doppler velocity estimation results (DOPvel) and POS520 velocity results. It can be seen that the BDS PPP-based Doppler velocity estimation results. Obviously better than the position differential velocity estimation results, Fig. 5 shows that the difference between the carrier velocity changes, the difference can reach several m/ s, and the Doppler velocity estimation results are not greatly affected. The two methods are different from the POS velocity results. The statistical results are shown in Table 2. It

Fig. 4. Experimental environment and trajectory

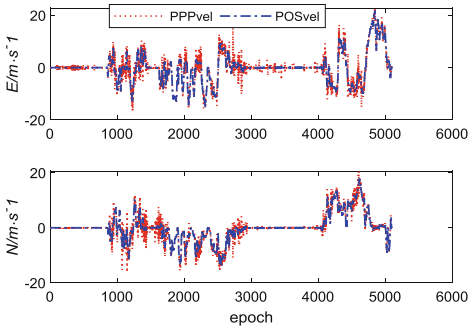


Fig. 5. Position differential velocity comparison

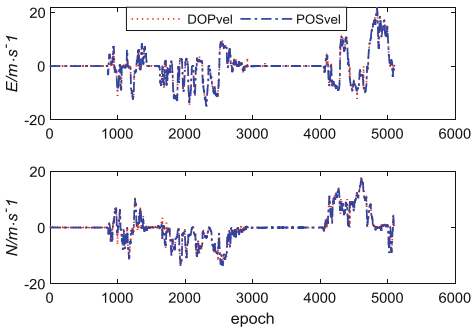


Fig. 6. Doppler velocity comparison

Table 2. Dynamic velocity difference results statistics

	RMS (cm/s)		
	E	*N*	*2D*
Position difference method	30.01	25.53	39.37
Doppler method	4.13	5.82	7.13

can be concluded that the Doppler velocity estimation based on BDS PPP and the high-precision POS velocity result are in the accuracy of cm/s, and the BDS PPP position differential velocity is consistent. The accuracy is at the dm/s level.

4 Conclusions

In this paper, the principle of position difference and Doppler velocity estimation based on BDS PPP is given. The influence of various errors on the accuracy of Doppler velocity estimation and the correction method are analyzed. Through static and sports car experiments, the BDS PPP position differential velocity estimation method and the BDS PPP-based Doppler velocity estimation method are compared and analyzed, turning out that the Doppler velocity estimation method is not affected by the carrier motion state, and the accuracy can reach cm/s level, but the position-difference velocity estimation method is greatly influenced by the motion state of the carrier, with the error even reaching several m/s under high maneuver. It is recommended to use the Doppler velocity estimation method in high dynamic applications, while position-difference velocity estimation method in low dynamic or low acceleration fields. With the continuous improvement of the BDS satellite system, the accuracy of the velocity estimation can be further improved.

References

1. The State Council Information Office of the People's Republic of China: China's BeiDou Navigation Satellite System. People's Publishing House, Beijing
2. Yang F (2014) Development status and prospects of driverless cars. Shanghai Auto 3:35–40
3. Zhang XH, Guo BF (2013) Real-time tracking the instantaneous of crust during earthquake with a stand-alone GPS receiver. Chin J Geophys 56(6):1928–1936
4. Yan K, Zhang T, Niu X et al (2017) INS-aided tracking with FFT frequency discriminator for weak GPS signal under dynamic environments. GPS Solutions 3(21):917–926
5. Shi J, Huang Y, Ouyang C et al BeiDou/GPS relative kinematic positioning in challenging environments including poor satellite visibility and high receiver velocity. Surv Rev. https://doi.org/10.1080/00396265.2018.1537227
6. Sun W, Duan SL, Ding W et al (2017) Comparative analysis on velocity determination by GPS single point. J Navig Positioning 5(1):81–85
7. Freda P, Angrisano A, Gaglione S et al (2015) Time-differenced carrier phases technique for precise GNSS velocity estimation. GPS Solutions 2(19):335–341
8. Sun W, Duan SL, Kong Y et al (2017) Velocity and acceleration of doppler calculation for carrier based on GPS broadcast ephemeris. Chin J Sens Actuators 30(11):1630–1635
9. Li ZH, Huang JS (2012) GPS surveying and date processing. Wuhan University Press, Wuhan

Analysis on Performance of BDS/GPS Fusion Pseudorange Positioning with ISB and Its Influence on DOP

Yifan Jing[1(✉)], Anmin Zeng[2,3], Ang Zhao[1], Yangyin Xu[1], and Yueyuan Ma[1]

[1] Information Engineering University, No. 62 Kexue Street, Zhengzhou, Henan, China
yifan_jing@foxmail.com
[2] Xi'an Research Institute of Surveying and Mapping, No. 1 Yanta Road, Xi'an, Shaanxi, China
[3] State Key Laboratory of Geo-information Engineering, No. 1 Yanta Road, Xi'an, Shaanxi, China

Abstract. The Global Navigation Satellite System (GNSS) user will benefit a lot from the multi-system fusion positioning in availability, accuracy and reliability. However, the Inter-System Bias (ISB) may weaken the performance of multi-system, especially for single point positioning applications. In this paper, the BDS/GPS (Beidou Satellite Navigation System; Global Positioning System) fusion pseudorange positioning model with the ISB estimation is presented. With the principle of Bayesian estimation, the influence of ISB on the function model is evaluated by analyzing the Generalized Dilution of Precision (G-DOP) factor. A set of measured BDS and GPS observations is used to test and verify the aforementioned models. The result demonstrates that the ISB is ubiquitous and the fusion model with ISB estimation performs well since the fusion positioning accuracy is improved. As an additional parameter, the ISB will cause a deterioration in G-DOP. The priori information of ISB can be utilized to reduce its impact, especially when the amount of observed navigation satellites is not enough.

Keywords: BDS/GPS · ISB · Fusion positioning · Generalized Dilution of Precision

1 Introduction

Global Navigation Satellite System (GNSS) users can benefit a lot from the multi-system application, which provides a better PNT (Positioning, Navigation and Timing) accuracy and reliability mainly because of the increase of observed navigation satellites [1]. Since systems are developed independently by different countries or organizations, GNSS differ greatly in system performance and some compatibility issues have to be deliberated. Taking Global Positioning System (GPS) and Beidou Satellite Navigation System (BDS) for an example, the coordinate and time reference applied in GPS are 1984 World Geodetic System (WGS84) and GPS Time (GPST) [2], while 2000 China Geodetic Coordinate System (CGCS2000) and Beidou Time (BDT) are applied in

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 380–388, 2019.
https://doi.org/10.1007/978-981-13-7751-8_38

BDS [3]. In addition, the GNSS observations with different frequencies usually contain different hardware delay errors and they are hard to be calibrated properly [4]. Generally, the hardware delay can be integrated into the receiver clock error term and estimated in single system process, but the inter-system difference of hardware delays in multi-system observations cannot be estimated in fusion positioning [5]. These discrepancies, known as Inter-System Bias (ISB), have a negative impact on the performance of multi-GNSS fusion positioning and a reasonable solution is necessary.

The BDS/GPS fusion relative positioning has been widely studied. The overall error of kinematic BDS/GPS differential positioning is about 3 cm when the baseline is more than hundred kilometers [6, 7]. With the variance component estimation used to adjust random errors from different systems, the fusion positioning performance will be better and the error is at millimeter level [8]. The relative fusion positioning has an advantage in the ISB correction for using the intra-system dual-differential model. The observation errors, such as the hardware delay, will be weakened and make fewer effects in the fusion system.

On the contrary, the non-differential single point positioning is more vulnerable to ISB since the system-depend error cannot be eliminated and persists in the fusion system. To solve this problem, some scholars recommend the positioning method that calculates the actual receiver clock error and ISB as a combined receiver clock term [9, 10]. And the fusion model with the virtual observation parameter for estimating ISB is also useful in some applications [11, 12]. The single point GNSS fusion positioning method still needs to be improved.

In this paper, we propose a BDS/GPS fusion pseudorange positioning model with the ISB estimation and the solution equation is deduced. And we discuss the influence of ISB as an additional parameter by analyzing the variation of the Dilution of Precision (DOP) factor. Finally, some measurement and calculation results are presented for testing the performance of the aforementioned BDS/GPS fusion positioning model.

2 BDS/GPS Fusion Pseudorange Positioning with ISB Estimation

The original BDS and GPS pseudorange observation equation is

$$\begin{cases} P^{\mathrm{B}} = \rho^{\mathrm{B}} + c\delta t_{\mathrm{r}}^{\mathrm{B}} - c\delta t^{\mathrm{s,B}} + I + T + D^{\mathrm{B}} + \varepsilon^{\mathrm{B}} \\ P^{\mathrm{G}} = \rho^{\mathrm{G}} + c\delta t_{\mathrm{r}}^{\mathrm{G}} - c\delta t^{\mathrm{s,G}} + I + T + D^{\mathrm{G}} + \varepsilon^{\mathrm{G}} \end{cases} \tag{1}$$

where B and G stands for BDS and GPS respectively; ρ is the geometric distance from the navigation satellite to the receiver; I and T presents the signal delay caused by ionosphere and troposphere; c is the speed of light; δt_{r} and δt^{s} is the receiver and satellite clock error; D is the hardware delay; ε is the pseudorange observation noise and other errors. Then, the BDS to GPS ISB can be described as [13].

$$\mathrm{ISB} = (c\delta t_{\mathrm{r}}^{\mathrm{B}} + D^{\mathrm{B}}) - (c\delta t_{\mathrm{r}}^{\mathrm{G}} + D^{\mathrm{G}}) \tag{2}$$

Theoretically, the ISB should also contain the coordinate bias. The coordinate reference applied in BDS and GPS, namely CGCS2000 and WGS84, has a similar origin, scale, and orientation, while tiny difference exists between the ellipsoid parameters such as the earth gravitational constant. Current studies indicate that CGCS2000 and WGS84 coordinates are consistent despite the centimeter-level difference [14], so the BDS and GPS observations are treated to be in the same coordinate reference.

The fusion observation equation with the ISB estimation can be written as [13]

$$\begin{cases} P^{\mathrm{B}} = \rho^{\mathrm{B}} + \left(c\delta t_{\mathrm{r}}^{\mathrm{G}} + D^{\mathrm{G}} + \mathrm{ISB}\right) - c\delta t^{s,\mathrm{B}} + I + T + \varepsilon^{\mathrm{B}} \\ P^{\mathrm{G}} = \rho^{\mathrm{G}} + \left(c\delta t_{\mathrm{r}}^{\mathrm{G}} + D^{\mathrm{G}}\right) - c\delta t^{s,\mathrm{G}} + I + T + \varepsilon^{\mathrm{G}} \end{cases} \tag{3}$$

where most observation errors in Eq. 3 can be corrected properly. For instance, the ionosphere and troposphere delay will be weakened by the corresponding correction models and the exact satellite clock error is already given in satellite ephemeris information.

For the adjustment solution, Eq. 3 can be transformed into

$$\begin{cases} \mathbf{L}_{\mathrm{B}} = \mathbf{A}_{\mathrm{B}}\mathbf{X} + \mathbf{B}_{\mathrm{ISB}}\mathbf{S} + \boldsymbol{\varepsilon}_{\mathrm{B}} \ , & \mathbf{P}_{\mathrm{B}} \\ \mathbf{L}_{\mathrm{G}} = \mathbf{A}_{\mathrm{G}}\mathbf{X} + \boldsymbol{\varepsilon}_{\mathrm{G}} \ , & \mathbf{P}_{\mathrm{G}} \end{cases} \tag{4}$$

where $\mathbf{L}_{\mathrm{B}}$ and $\mathbf{L}_{\mathrm{G}}$ is the corrected observation vector; $\boldsymbol{\varepsilon}_{\mathrm{B}}$ and $\boldsymbol{\varepsilon}_{\mathrm{G}}$ is the observation noise vector; $\mathbf{X}$ is the receiver coordinate and clock error parameter vector; $\mathbf{A}_{\mathrm{B}}$ and $\mathbf{A}_{\mathrm{G}}$ is the designed matrix of $\mathbf{X}$; $\mathbf{S}$ is the ISB parameter vector; $\mathbf{B}_{\mathrm{ISB}}$ is the designed matrix of $\mathbf{S}$; $\mathbf{P}_{\mathrm{B}}$ and $\mathbf{P}_{\mathrm{G}}$ is the observation weight matrix. The error equation and least square adjustment solution is

$$\mathbf{V} = \mathbf{A}\hat{\mathbf{X}} + \mathbf{A}\hat{\mathbf{S}} - \boldsymbol{l} \tag{5}$$

$$\begin{pmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{S}} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{B} \\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{B} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\boldsymbol{l} \\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\boldsymbol{l} \end{pmatrix} \tag{6}$$

where $\boldsymbol{l} = \begin{bmatrix} \mathbf{L}_{\mathrm{B}} - \boldsymbol{\varepsilon}_{\mathrm{B}} & \mathbf{L}_{\mathrm{G}} - \boldsymbol{\varepsilon}_{\mathrm{G}} \end{bmatrix}^{\mathrm{T}}$, $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{\mathrm{B}} & \mathbf{A}_{\mathrm{G}} \end{bmatrix}^{\mathrm{T}}$ and $\mathbf{B} = \begin{bmatrix} \mathbf{B}_{\mathrm{ISB}} & 0 \end{bmatrix}^{\mathrm{T}}$; $\hat{\mathbf{X}}$ and $\hat{\mathbf{S}}$ stands for the estimate of $\mathbf{X}$ and $\mathbf{S}$.

The minimum quantity of observed satellites for positioning is increased to five because the ISB is involved and needs to be estimated simultaneously in the observation functional equation. Since navigation satellites can be used cross systems, this requirement is easy to reach in the multi-system condition.

3 Impact of ISB on DOP Factor

The ISB estimation makes fusion model more reasonable. But the participation of ISB as an additional parameter may reduce the performance of the positioning function model. The DOP is the factor that indicates the geometry strength of the measurements,

error magnitude and the estimation risk criteria, and it is typically used as a major index to gauge the contribution of satellites [5].

The original DOP of the BDS/GPS fusion observation equation is

$$\mathrm{DOP}_{(\mathrm{fusion})} = \sqrt{tr\left[\left(\mathbf{A}_{\mathrm{B}}^{\mathrm{T}}\mathbf{P}_{\mathrm{B}}\mathbf{A}_{\mathrm{B}} + \mathbf{A}_{\mathrm{G}}^{\mathrm{T}}\mathbf{P}_{\mathrm{G}}\mathbf{A}_{\mathrm{G}}\right)^{-1}\right]} \tag{7}$$

where letters have the same meaning as Eq. 4. And it is easy to be found that

$$\begin{cases} \mathrm{DOP}_{(\mathrm{fusion})} < \mathrm{DOP}_{(\mathrm{BDS})} = \sqrt{tr\left[\left(\mathbf{A}_{\mathrm{B}}^{\mathrm{T}}\mathbf{P}_{\mathrm{B}}\mathbf{A}_{\mathrm{B}}\right)^{-1}\right]} \\ \mathrm{DOP}_{(\mathrm{fusion})} < \mathrm{DOP}_{(\mathrm{GPS})} = \sqrt{tr\left[\left(\mathbf{A}_{\mathrm{G}}^{\mathrm{T}}\mathbf{P}_{\mathrm{G}}\mathbf{A}_{\mathrm{G}}\right)^{-1}\right]} \end{cases} \tag{8}$$

It is proved that the positioning performance of the fusion system is usually better than the single system in the DOP view. When the ISB becomes an additional parameter in the observation functional equation, the common DOP factor based on the least squares estimation and unified functional and stochastic models is no longer applicable. To describe the relationship between DOP and the influence function, the generalized DOP (G-DOP) is utilized [5].

When additional parameters for the multi-system interoperability are added to the functional model, the second type of G-DOP (G-$\mathrm{DOP}_{\mathrm{II}}$) should be used [5]. The G-$\mathrm{DOP}_{\mathrm{II}}$ of BDS/GPS fusion model is

$$\mathrm{G\text{ - }DOP}_{\mathrm{II}} = \left\{ tr\left[\begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{B} \\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{B} \end{pmatrix}^{-1} \right] \right\}^{1/2} \tag{9}$$

Comparing Eq. 8 with Eq. 9, we can find that $\mathrm{G} - \mathrm{DOP}_{\mathrm{II}} \geq \mathrm{DOP}_{(\mathrm{fusion})}$, which means the estimation of ISB as an additional parameter will deteriorate geometry strength of the measurements and the positioning performance may be decreased.

To weaken effects of the additional parameter, the priori can be applied. The priori ISB can be derived from the solution of former observations or the calibration by receiver manufacturers. Based on Bayesian estimation [15], the positioning solution (Eq. 6) can be transformed into

$$\begin{aligned} \begin{pmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{S}} \end{pmatrix} &= \left(\mathbf{N}\begin{bmatrix} \mathbf{X} \\ \mathbf{S} \end{bmatrix} + \mathbf{N}_{\bar{\mathbf{S}}} \right)^{-1} \begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\boldsymbol{l} \\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\boldsymbol{l} + \mathbf{P}_{\bar{\mathbf{S}}}\bar{\mathbf{S}} \end{pmatrix} \\ \mathbf{N}\begin{bmatrix} \mathbf{X} \\ \mathbf{S} \end{bmatrix} &= \begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{B} \\ \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{P}\mathbf{B} \end{pmatrix}, \mathbf{N}_{\bar{\mathbf{S}}} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{P}_{\bar{\mathbf{S}}} \end{pmatrix} \end{aligned} \tag{10}$$

where $\bar{\mathbf{S}}$ is the priori matrix with a weight matrix $\mathbf{P}_{\bar{\mathbf{S}}}$. And the third type of G-DOP (G-$\mathrm{DOP}_{\mathrm{III}}$) with the priori is [5].

$$\begin{aligned} \mathrm{G}-\mathrm{DOP}_{\mathrm{III}} &= \left\{ tr\left[\begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{PA} & \mathbf{A}^{\mathrm{T}}\mathbf{PB} \\ \mathbf{B}^{\mathrm{T}}\mathbf{PA} & \mathbf{B}^{\mathrm{T}}\mathbf{PB}+\mathbf{P}_{\bar{\mathbf{S}}} \end{pmatrix}^{-1} \right] \right\}^{1/2} \\ &= \left\{ tr\left[\left(\mathbf{N}_{\begin{bmatrix} \mathbf{X} \\ \mathbf{S} \end{bmatrix}} + \mathbf{N}_{\bar{\mathbf{S}}} \right)^{-1} \right] \right\}^{1/2} \end{aligned} \tag{11}$$

It is clearly to be known that

$$\mathrm{G}-\mathrm{DOP}_{\mathrm{III}} = tr\left(\left[\mathbf{N}_{\begin{bmatrix} \mathbf{X} \\ \mathbf{S} \end{bmatrix}} + \mathbf{N}_{\bar{\mathbf{S}}} \right]^{-1} \right) \le tr\left(\mathbf{N}^{-1}_{\begin{bmatrix} \mathbf{X} \\ \mathbf{S} \end{bmatrix}} \right) = \mathrm{G}\text{ - }\mathrm{DOP}_{\mathrm{II}} \tag{12}$$

The priori can be used to reduce the influence of the additional parameter. If the priori of ISB is accurate enough, the ISB parameter can be corrected completely and need not be estimated again.

4 Calculation and Analysis

We carried out several BDS/GPS observation experiments at different locations in China. The observing period lasted for the entire day of DOY 228, 2014, and the sampling rate was 30 s. The minimum satellite elevation was set to 15° and the weight ratio between BDS and GPS observations was decided by the satellite elevation. The ionosphere and troposphere delay was corrected by Klobuchar and Saastamoinen model [16, 17].

The difference between BDS and GPS positioning results and estimated receiver clock errors is summarized in Table 1.

As we can see from Table 1, the mean deviations of BDS and GPS positioning results are at meter level. It proves that the accuracy of BDS and GPS observations is different, which is mainly related to the system capabilities. The difference of "receiver clock errors" calculated by BDS and GPS is varied. It is an opposition to the fact that the receiver clock error is unique for the single receiver, and the calculated "receiver clock error" is actually a combination that contains the real clock error and other systemic deviations between BDS and GPS. These systemic deviations are known as ISB and it is distinct in different receivers.

To make the analysis exhaustive, the result of Guilin is illustrated in detail as an example. The estimated ISB from former calculations is used as the priori.

The receiver clock errors calculated by BDS and GPS are shown in Fig. 1. As it can be seen, both BDS-derived and GPS-derived estimated clock errors are varied with times and the differences of two series seem similar on all epochs. Knowing from Table 1 that the standard deviation of the clock error difference in Guilin is 15 ns, we believe that the

Table 1. The comparison of BDS and GPS positioning results and estimated receiver clock errors

Stations	Clock error difference \ ns		Position difference on B \ m		Position difference on L \ m		Position difference on H \ m	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
Sanya	930	12	1.6092	2.2258	−1.9001	2.0146	−1.9981	6.3518
Guilin	1078	15	1.5818	2.6261	−1.7598	2.1767	−0.0961	5.6904
Yichang	498	13	2.6145	3.1552	−2.1853	2.9482	−2.1087	6.5848
Xi'an	943	14	2.0437	2.0917	−2.0594	2.1754	−1.7397	4.0624
Togtoh	114	11	2.3513	2.8615	−1.9068	2.3786	−2.8246	4.4339

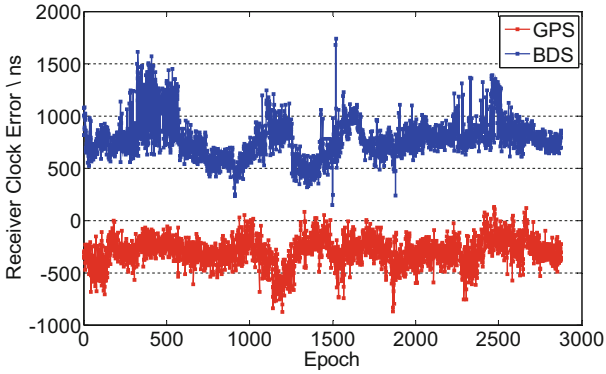


Fig. 1. The receiver clock errors calculated by BDS and GPS in Guilin

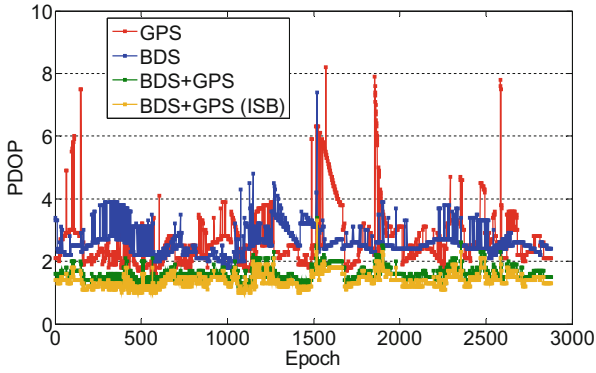


Fig. 2. The result of PDOP factors in Guilin

difference between calculated clock errors is an approximate constant for a day. The same results can be obtained by observations in other stations. This conclusion agrees with the result in reference [18], which indicates that the ISB is stable and repeatable over a day and can be used as a priori for positioning applications [18].

The PDOP results are presented in Fig. 2. As we can see, the PDOPs calculated by BDS and GPS are in the same magnitude and the value of BDS-derived PDOP is a little larger overall. The PDOP performance is improved a lot in BDS/GPS fusion system. Compared to GPS results, the fusion system PDOP is reduced by 40%, and the reduction comparing to BDS is 37%. It proves that the fusion of BDS and GPS observations makes an obvious improvement in the DOP factor and it will bring a better positioning performance. On the other hand, the participation of ISB estimations makes fusion system DOP increase slightly, about 12%. It is acceptable because the DOP improvement is much greater. When the interoperability requirements are achieved, which means the effects of ISB and other observation errors are eliminated or weakened, the BDS/GPS fusion positioning accuracy can be significantly promoted.

The single epoch positioning results calculated by BDS, GPS and BDS/GPS are shown in Figs. 3 and 4, and some statistics are listed in Table 2.

Table 2. The statistic of BDS, GPS and BDS/GPS positioning results in Guilin \ m

	Positioning Error	BDS	GPS	BDS/GPS	
				Without ISB estimation	With ISB estimation
B	Avg.	1.2237	−0.3580	11.8551	0.2683
	Std.	1.8732	1.6053	17.9517	1.0706
L	Avg.	−4.1456	0.6141	−12.5764	−1.9136
	Std.	2.4243	1.9791	15.5495	1.4627
H	Avg.	−0.9862	−0.8901	−17.4019	−0.5413
	Std.	6.5218	2.5832	22.4747	1.8747

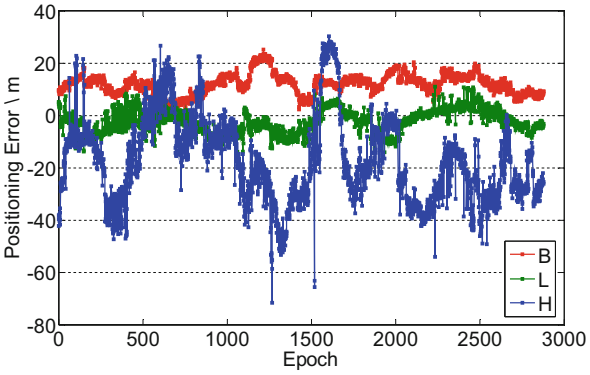


Fig. 3. The result of BDS/GPS fusion positioning without ISB estimation in Guilin

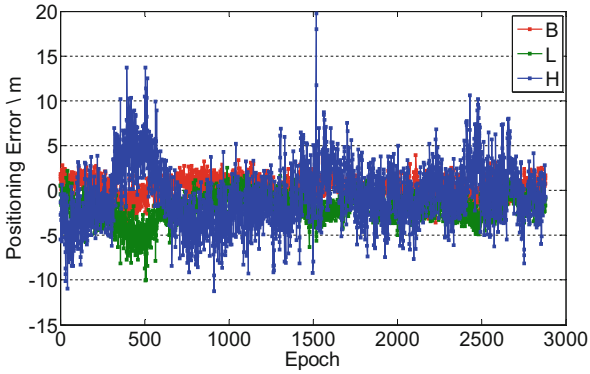


Fig. 4. The result of BDS/GPS fusion positioning result with ISB estimation in Guilin

As we can see from Fig. 3, the result of BDS/GPS fusion positioning without ISB estimation is awful, and it can be known from Table 2 that the standard deviations of the result are 17.9517 m, 15.5495 m and 22.4747 m. It is a disaster since the standard deviation result for the single system is even better, which are 1.8732 m, 2.4243 m, 6.5218 m for BDS, and 1.6053 m, 1.9791 m, 2.5832 m for GPS. The systematic deviation ISB between different systems is ubiquitous. As an additional error, the ISB will make the function model faulty, and the positioning performance will become worse.

As it is shown in Fig. 4, the improvement by adding the ISB estimation to BDS/GPS fusion positioning is obvious. Knowing from Table 2, the standard deviations of the positioning results are 1.0706 m, 1.4627 m and 1.8747 m. By setting an additional parameter to estimate ISB along with positions and the receiver clock errors, the fusion function model becomes reasonable and the positioning performance is promoted significantly.

5 Conclusion

The positioning performance will be improved greatly when the observations collected by multi-GNSS can be combined properly. The proper fusion model should contain the ISB estimation since its influence may deteriorate the fusion positioning performance. For BDS and GPS, the ISB is mainly consisted of the deviation of time reference systems and hardware delays while the deviation of the coordinate reference system, CGCS2000 and WGS84, can be neglected. To correct BDS to GPS ISB in pseudorange observations, we use a fusion model with the additional parameter and estimate ISB along with positions and receiver clock errors. The additional parameter may affect the positioning performance and cause an increase in G-DOP factor, which is an upgrade for common DOP factor by considering the additional parameter. Fortunately, the ISB has been proved stable for a day and the former solution can be utilized as the priori information. By using the priori, the influence of the additional parameter is reduced and the G-DOP factor can be improved. It is effective especially in the condition that observed navigation satellites is insufficient [18].

Acknowledgement. This work is supported by the China National Key Research and Development Program (2016YFB0501701) and the National Natural Science Foundation of China (41474015, 41604013 and 41874016).

References

1. Yang Y (2010) Progress, contribution and challenges of compass/beidou satellite navigation system. Acta Geodaetica Cartogr Sinica 39(1):1–6
2. National Imagery and Mapping Agency (1997) Department of Defense World Geodetic System 1984, 3rd edn. http://earth-info.nga.mil/GandG/publications/tr8350.2/wgs84fin.pdf
3. China Satellite Navigation Office (2012) BD-SIS-ICD: BeiDou Navigation Satellite System Signal in Space Interface Control Document. http://en.beidou.gov.cn/SYSTEMS/ICD/201806/P020180608523308843290.pdf
4. Cai C, Li Z, Zhang X (2002) A study of calibration method of GPS satellite and receiver instrumental biases. Bull Surv Mapp 4:15–16
5. Yang Y, Li J, Xu J et al (2011) Generalized DOPs with consideration of the influence function of signal-in-space errors. J Navig 64:3–18
6. Gao X, Guo J, Chen P (2012) Fusion positioning of BeiDou/GPS based on spatio-temporal system unification. Acta Geodaetica Cartogr Sinica 41(5):743–758+755
7. He H, Li J, Yang Y et al (2014) Performance assessment of single-and dual-frequency BeiDou/GPS single-epoch kinematic positioning. GPS Solution 18(3):393–403
8. Jing Y, Zeng A, Xu T (2014) Fusion positioning of BDS/GPS based on variance component estimation and its application for geodetic control network. In: Sun J, Jiao W, Wu H, Lu M (eds) China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume I, vol 303. LNEE. Springer, Heidelberg, pp 115–123. https://doi.org/10.1007/978-3-642-54737-9_12
9. Zhao C, Ou J, Yun Y (2005) The simulation analysis of positioning accuracy and reliability of Galileo and GPS-Galileo based on single point positioning model. Chin Sci Bull 50:811–819
10. Li H, Dang Y, Bei J et al (2014) Research on model and performance of BDS/GPS/GLONASS multi- mode fusion positioning. Bull Surv Mapp 9:1–5
11. Cai C, Gao Y (2009) A combined GPS/GLONASS navigation algorithm for use with limited satellite visibility. J Navig 62:671–685
12. Angrisano A, Gaglione S, Gioia C (2013) Performance assessment of GPS/GLONASS single point positioning in an urban environment. Acta Geod Geophys 48:149–161
13. Zeng A, Yang Y, Ming F et al (2017) BDS–GPS inter-system bias of code observation and its preliminary analysis. GPS Solutions 21(4):1573–1581
14. Wei Z (2008) China geodetic coordinate system 2000 and its comparison with WGS84. J Geodesy Geodyn 28(5):1–5
15. Koch KR (1990) Bayesian inference with geodetic applications. Springer, Heidelberg
16. Klobuchar JA (1996) Ionospheric effects on GPS. In: Parkinson BW, Spilker JJ (eds) Global positioning system: theory and applications, vol I, Chap 12, pp 485–515
17. Saastamoinen J (1946–1975) Contribution to the theory of atmospheric refraction. Bull Géodésique 105(1): 279–298
18. Zeng A, Yang Y, Jing Y et al (2017) Systematic bias compensation model of inter-system bias and its performance analysis for BDS/GPS fusion positioning. Geomatics Inform Sci Wuhan Univ 42(10):1423–1430

Study on Search and Rescue System Based on BeiDou Global Navigation Satellites

Yuan Guojing(✉), Liu Jiang, and Liu Han

China Academy of Space Technology (Xi'an), Xi'an, China
yunqil013@163.com

Abstract. BeiDou satellite navigation system (BDS), which is expected to offer PNT service around the world in 2020, is currently in the stage of continuous launching and improving. In order to develop service capabilities, BDS carries the medium-altitude Earth orbit navigation satellite search and rescue (MEOSAR) payloads and global short message communication (GMSC) payloads on a part of MEO satellites. These payloads provide global MEOSAR service and GMSC service, while the GMSC payloads can also offer SAR service. The characteristics of SAR system based on BeiDou global navigation satellites, which is abbreviated as BDS SAR system can be analyzed by studying and comparing MEOSAR system with the SAR system based on GMSC, which is defined as GMSC SAR system. Benefitting from using the dedicated frequency allocated by ITU, the MEOSAR system can avoid the possible interference from other systems. Moreover, the completion of MEOSAR system construction will improve to establish an international image of China's responsibility for fulfilling humanitarian relief. While the GMSC SAR system also has some helpful merits to carry out rescue work smoothly, such as implementing ID identification which enable to obtain the identity information of the user directly, and realizing the short message communication which can be used to contact the distress users to report the location information, physical condition. Furthermore, owing to complete intellectual property and characterizing with independent and confidential, the GMSC SAR system is especially suitable for military use. Here come two development proposals of the BDS SAR system through analyzing the characteristics and current situation. First of all, the return link should be established as soon as possible. The other advice is to develop an integrated user terminal for both MEOSAR system and GMSC system.

Keywords: MEOSAR system · GMSC system · RLS

1 Introduction

The rapid search and rescue for people in aviation distress, navigation distress and major natural disasters has always been an important research subject around the world. Since the beginning of new century, the comprehensive national strength of China has rapidly increased. Accompanying, the development and national rejuvenation requires our country to integrate into the world, while the international community also needs China to assume more international affairs and humanitarian obligations. At the same time, with the expansion of the country's overseas strategic interests, it is necessary for

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 389–395, 2019.
https://doi.org/10.1007/978-981-13-7751-8_39

military forces to radiate the world, and it is of great significance to construct a complete SAR system with independent intellectual property rights.

2 Status of Satellite SAR System

2.1 Cospas-Sarsat

Cospas-Sarsat is an international satellite system for SAR distress alerting which was established in 1979 by Canada, France, the USA and the Former USSR [1]. Including low-altitude Earth orbit SAR (LEOSAR) system and geostationary Earth orbit SAR (GEOSAR) system, it has been successfully applied to a large number of distress search and rescue operations worldwide.

Due to the characteristics of the orbits of LEO and GEO, there are some shortcomings of the system such as long response time and blind areas. In 2000, the USA, the European Commission (EC) and Russia began consultations with the International Search and Rescue Satellite Organization regarding the feasibility of installing 406 MHz SAR instruments on their respective medium-altitude Earth orbit navigation satellite systems, including the USA distress alerting satellite system (DASS), European Galileo MEOSAR, and the Russian GLONASS MEOSAR [1, 2]. Many investigations show that the disadvantages of LEOSAR and GEOSAR could be overcome by MEOSAR system. Therefore, the council of the International Search and Rescue Satellite Organization decided to develop the MEOSAR system in 2004.

The MEOSAR system should be implemented in six phases to clearly delineate development and implementation activities [1]:

(1) Definition and Development Phase;
(2) Proof of Concept (POC)/In-orbit Validation Phase;
(3) Demonstration and Evaluation Phase (D&E);
(4) Early Operational Capability (EOC);
(5) Initial Operational Capability (IOC);
(6) Full Operational Capability (FOC).

From 2013, MEOSAR system entered global D&E phase, which was in line with the expectation on reliability and precision. Then it entered EOC phase by the end of 2016 and is expected to enter FOC phase in 2020, when enough MEOSAR satellites and commissioned ground stations are available to provide worldwide service.

2.2 Domestic SAR System Development

China joined Cospas-Sarsat programme and became a user since 1985. Then became a provider of ground system devices after MCC and LUT stations were completed and commissioned in 1998. After September 19, 2018, when the 13th and 14th satellites of Beidou global satellite navigation system, which equipped MEOSAR payloads, were launched, China became a potential supplier of space equipment for the Cospas-Sarsat system. According to the Beidou global system development plan, SAR payloads will also be deployed on other four MEO satellites.

On the other hand, the regional BeiDou Radio Determination Satellite Service (RDSS) system was completed on December 27, 2012, and began to offer service for Asian-Pacific region formally. This system which can achieve services such as location reporting and short message communication, has played a important role in the earthquake relief work of the Wenchuan earthquake in 2008 and the return of manned space flight search and rescue [3–5]. Moreover, the GMSC system, which provides the location report and short message communication services could be used for global search and rescue service, is further extended to the whole world in the BeiDou global system.

3 BDS SAR System

3.1 BDS SAR System Scheme

A part of BeiDou global system MEO satellites have been equipped with international MEOSAR payloads to provide SAR services. At the same time, MEO satellites are also equipped with GMSC payloads, which can be used to global search and rescue in distress events. The BDS SAR system is made up of three segments: the user segment including Cospas-Sarsat 406 MHz distress beacons and the user terminal of BDS GMSC, the space segment consisting of BeiDou global system MEO satellites, which are equipped MEOSAR payloads, GMSC payloads and the cross links, and the ground segment comprising of Cospas-Sarsat part which includes MEOLUT, MCC, RCC, and GMSC part, which includes BeiDou system operation control center and GMSC SAR control center, as shown in Fig. 1.

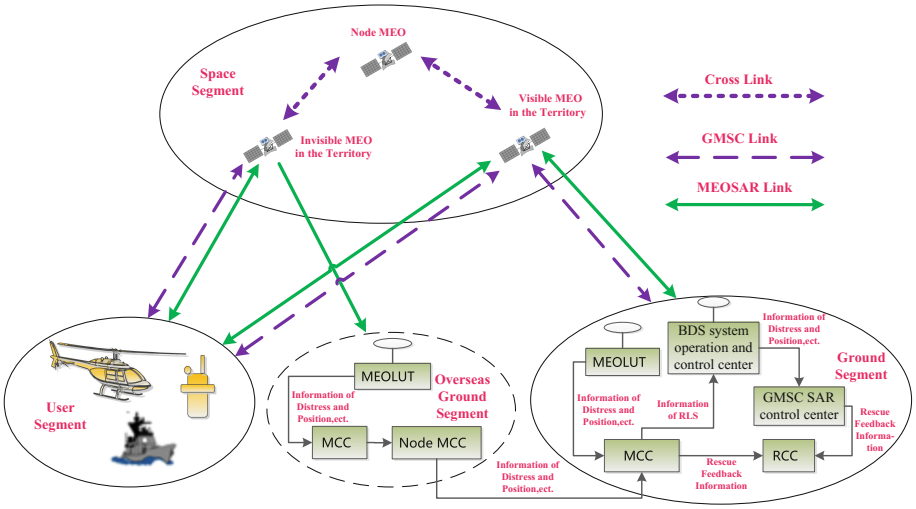


Fig. 1. The schematic of BDS SAR system MEOLUT represents the local user terminal of MEOSAR system, while the MCC represents search and rescue mission control center, and RCC represents rescue coordination center. The overseas ground segment in the virtual box include the MEOLUTs and MCCs of Cospas-Sarsat system which are located in foreign countries.

3.1.1 BDS MEOSAR System

BeiDou MEOSAR system is a part of Cospas-Sarsat system, which is composed of the distress beacons, BeiDou MEO satellites that equipped with MEOSAR payloads, ground MEOLUT, MCC, and RCC. The distress beacons can trigger a 406 MHz distress signal transmission in an active or passive manner when a distress event occurs. Then the MEO satellite receives the beacon signal and transmits it to the ground through the antenna at a downlink frequency of 1544.21 MHz after being frequency converted and amplified. If this satellite is visible in the territory, the downlink distress signal will be received by the domestic MEOLUT, and the location of the user in combination with the distress information can be calculated and transmitted to MCC. Otherwise, when the distress signal is received by the overseas MEOLUT, the user position and distress information will be sent to the corresponding MCC and transmitted to the MCC in China via the relay MCC. After receiving the distress information from domestic MEOLUT or abroad relay MCC, the domestic MCC will contact to the nearest RCC through the communication network to organize the search and rescue operation.

As the forward link, MCC receives the distress information from the user and transmits it to RCC. Meanwhile, the return link service (RLS) information, including the confirm information generated by MCC and rescue feedback information from RCC, will be sent to BeiDou operational control center, transmitted to satellites, and finally received by the users in distress through return links of the satellites. The RLS of BeiDou MEOSAR is still in research.

3.1.2 BDS GMSC SAR System

The BDS GMSC SAR system composes of user terminals, BeiDou MEO satellites which equipped with GMSC payloads and cross links between MEO satellites and the ground segment including BDS system operation and control center, the GMSC SAR control center and the RCC.

When a distress event occurs, the L-band signal modulated distress information and position report is sent to visible MEO satellites by the user terminal, then received and demodulated by the GMSC devices on satellites, transmitted to BDS system operation and control center by the MEO satellites themselves, or through cross links, finally send to RCC to implement search and rescue.

Subsequently, the BeiDou operation and control center transmits the return link information which includes the rescue feedback information from RCC to the user terminals through the satellites up and down link, sometimes the cross link is needed.

3.2 Comparative Analysis

Comparative analysis of the BDS MEOSAR system and the BDS GMSC SAR system are shown in Table 1.

Table 1. A comparison of BDS MEOSAR system and the GMSC SAR system

Item	BDS MEOSAR system	BDS GMSC SAR system
Positioning principle	TDOA/FDOA positioning, or GNSS passive positioning	GNSS passive positioning
Positioning accuracy	TDOA/FDOA positioning: about 5 km; GNSS passive positioning: about 5 m	GNSS passive positioning: about 5 m
Service capabilities	Expected to be FOC phase in 2020	Plan to have global service capabilities by 2020
Service object	Registered or unregistered users, ships, aircraft, individuals, etc.	registered users
User frequency	Frequency of SAR allocated by ITU	GMSC frequency, there are possible interference signals from other system abroad
Rescue feedback	RLS standard is not proposed by Cospas-Sarsat; the RSL of Beidou MEOSAR is researched and not implemented	Reliable feedback is achieved
Overseas user information link	Invisible MEO in the Territory— > overseas MCC — > node MCC— > domestic MCC	Invisible MEO in the Territory - > Cross link - > Visible MEO in the Territory - > Ground Station
International interoperability	Beidou MEOSAR is developed in accordance with international interoperability standards and will be integrated in Cospas-Sarsat system after commissioning	BDS GMSC system is unique, which is not suitable for international promotion and cannot achieve international interoperability
Communication function	No	Short message communication

3.3 Characteristics and Development Suggestions of BDS SAR System

The BDS SAR system involves the MEOSAR system conforming to international interoperability standards and the BDS GMSC SAR system having independent intellectual property rights. While the frequency of MEOSAR is allocated by ITU, which can avoid the possible interference from other systems, moreover, the construction of the MEOSAR system is conducive to improving the ability of life and property rescue on a global scale, enhancing the international influence of the BDS system in the field of satellite navigation, and establishing an international image of China's responsibility for fulfilling humanitarian relief, the GMSC SAR system also has the following characteristics:

1. ID management

Benefiting from the unique ID number is settled for each terminal, and the location information of the distress person can be obtained through receiving and demodulating the user uplink signal, the identity information of the user can be directly confirmed by

BDS operation control center, which is contributing the search and rescue activity successfully.

2. Short message communication function

In search and rescue process, the short message communication function can be used to contact the distress user in case of the ground communication network is out of work, thereby improving the successful rate of rescue. Furthermore, it can timely report the physical condition, surrounding environment and threat situation of the people in distress, and receive short message instructions from the rescue center to help rescuers carry out rescue work smoothly.

3. Accurate positioning and communication functions based on a single satellite navigation system

The navigation and communication integration is completed in the navigation system, which not only expands the navigation function, but also avoids communication obstacles caused by different communication systems and departmental preparations.

4. Independent intellectual property

The GMSC SAR system, whose location reporting link of the GMSC SAR system is combined with the navigation link in the same navigation satellite system, makes the information transmission link to be secure, independent and confidential, suitable for military use especially.

In the BeiDou satellite system, the MEOSAR system and GMSC SAR system can complement each other and offer the SAR services of various users. However, the RLS of BDS MEOSAR system has not been realized, and there is no unified user terminal for the two SAR systems, which brings great inconvenience to users. Therefore, the following two recommendations are proposed for the BDS SAR system:

1. Suggest accelerating the implementation of the RLS

Studies show that if the distress users can receive feedback, the successful rate of rescue will greatly improve. Therefore, recommendation of accelerating BDS MEOSAR system RLS is necessary. The RLS information can be sent through existing basic navigation downlink signals, such as B2b signals [6].

2. Suggest developing an integrated user terminal for both the MEOSAR system and GMSC SAR system

From the comparative analysis in Sect. 3.2, both the MEOSAR system and the GMSC SAR system have their own advantages, such as: The MEOSAR system can avoid the possible interference from other systems benefiting from the dedicated frequency, while the GMSC SAR system has the communication function which is great helpful for search and rescue events [7]. In order to enable users to experience the advantages of the BDS SAR system fully through employing two kinds of SAR systems in the meantime, best idea is to develop an integrated user terminal for both the MEOSAR system and the GMSC SAR system.

4 Conclusions

The BDS SAR system includes both the international MEOSAR system and the GMSC SAR system. Through the research on the BDS SAR system, the development proposals are put forward. The completion of the BDS SAR system construction will greatly enhance the international influence of the BDS system in the field of satellite navigation, and provide a strong guarantee for human life safety worldwide.

References

1. Cospas-Sarsat 406 MHz MEOSAR implementation plan (2018). C/S R.012, Issue 1 - Revision 13, International Satellite System for Search and Rescue, Montreal
2. Zheng Q, Wang B, Cheng X et al (2015) Global satellite search and rescue system and its development status. Satell Appl 10:29–31
3. Youjun P (2013) Principle and structure of SAR system based on BDS. Avionics Technol 44 (4):7–13
4. Qingyun Z, Yunbin L, Yutao S (2014) Research on maritime search and rescue system based on Beidou system. Tianjin Navig 3:64–65
5. Liu S, Yu L, Wang S (2012) Application exploration of Beidou system search and rescue system. Silicon Valley 2012(6):128
6. Zhao Y, Liu C, Yuan G (2017) Beidou satellite navigation SAR system design. In: The 8th China satellite navigation conference, Nanjing
7. Xingang F, Zhiqiang S, Lixin Z (2014) Study between satellite search and rescue and Beidou radio determination satellite system. Space Electron Technol 11(2):63–66

Satellite Navigation Signal and Signal Processing

Research on Low-Power Short Burst Signal Synchronization Segment Design Based on MEO Satellites and the Fast Acquisition Algorithm

Tianqiao Zhang(✉), Henglin Chu, Aiyong Zhang, Hui Ren, and Jinglei Guo

Beijing Satellite Navigation Center, Beijing 100094, China
zhangtqee@aliyun.com

Abstract. The message service is the most distinctive part of the BeiDou System (BDS), whose coverage area will be expanded from China & surroundings to the global surface space at BDS-3 stage. This is an innovative integration of a global mobile communications service in a Global Navigation Satellite System (GNSS), with great practical value and significant strategy. However, BeiDou Global Message service relies on the global Medium Earth Orbit (MEO) satellite constellation to achieve the global beam-forming, the space-borne antenna has smaller size and lower receiving gain than regional beam-forming, so a key issue in designing Global Message is the uplink signal synchronization segment to ensure the space-borne message receiver achieves fast and reliable capture of short burst signals at very low carrier-to-noise ratio (C/N0). In this paper comparing with the uplink channel in Radio Determination Satellite Service (RDSS) based on the BDS Geostationary Earth Orbit (GEO) satellites, the signal condition parameters of the optional BeiDou Global Message are calculated. Then the design scheme of the shortest length synchronization segment based on the constant false alarm and specified detection probability is given. According to this, a fast acquisition algorithm for the uplink message receiver based on MEO satellite is proposed. Finally, through software simulation analysis and ground equipment testing the proposed signal design scheme is proved, as well as the fast acquisition algorithm with high-reliability reception of short burst signals in the condition of signal-to-noise ratio lower than 34 dBHz. The research results of this paper provide important technical support for the overall design of Global Message service in BDS-3, and will be further tested in the subsequent actual applications. It is also hoped to provide theoretical basis for the continuous optimization of Global Messages in the future BDS.

Keywords: BDS-3 · Global Navigation Satellite System · Global Message · MEO constellation

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 399–409, 2019.
https://doi.org/10.1007/978-981-13-7751-8_40

1 Introduction

Since the establishment of BeiDou System (BDS), it has been more than just a navigation satellite system, but a multi-service radio satellite system [1]. Message service is one of the earliest functions in the continuously developing BDS, and it is also the main feature of BDS that distinguished from other Global Navigation Satellite System (GNSS) systems. In the BDS-1 and BDS-2, the BeiDou message payloads were only deployed on Geostationary Earth Orbit (GEO) satellites, covering China & surrounding areas [2], and played an important role in many application fields, such as navigation monitoring, emergency communication, search & rescue, forest fire prevention, etc. According to the literatures [3, 4], the message payloads in BDS-3 are first carried on the Medium Earth Orbit (MEO) satellite and will achieve continuous global coverage ability. BDS-3 coordinates the global constellation and L-band downlink signal and Ka-band inter-satellite link to realize the integration of navigation and communication, which has great technical innovation and practical value.

However, we should note significant differences between the Global Message based on MEOs and the regional message based on GEOs in technical approaches and radio channel conditions [5]:

(1) the geostationary orbit satellites become mobile orbit satellites;
(2) the payload is changed from transparent transponder to on-board processing and inter-satellite relay;
(3) the satellite beam is changed from spot beamforming to global beamforming;
(4) the carrier-to-noise ratio (C/N0) of the user arrival signal becomes weaker.

Although the radio transmission loss from the Earth's surface to the BeiDou MEO satellite is about 3.7 dB lower than to the GEO satellite, but the receiving gain of the MEO satellite antenna is nearly 15.5 dB lower than that of the GEO satellite antenna [5]. The above information indicates that under the same ground user terminal transmit power conditions, the arrival C/N0 of the MEO satellite message burst signal is about 11.8 dB lower than that of the GEO satellite, also can be expressed as 1/15.

The rapid acquisition of short burst signals under extremely low carrier-to-noise ratio condition is a difficult issue in the design and implementation of the BDS-3 Global Message service. In the following sections, Sect. 2 compares the significant differences of signal parameters between the BDS-3 MEO Global Message and the GEO Regional Message, and points out the key aspects of the Beidou Global Message in signal design. From the perspective of signal design, Sect. 3 analyzes the synchronization segment design method to meet the fast acquisition requirement of short burst signals in low carrier-to-noise ratio, and proposes. The shortest synchronization segment length in constraints of constant false alarm, detection probability and implement loss. Section 4 optimizes the parameters of the spaceborne acquisition algorithm and the synchronization segment length by simulation calculation. Section 5 gives the results of test verification on the ground.

2 Characteristics of BeiDou Global Message Signal

As the extension of the regional RDSS Message, BDS-3 Global Message has certain inheritance in common with the past RDSS Message signal.

(1) The user uplink signals are all short burst signals;
(2) The transmission power of the user terminal are all at the level of 10 W;
(3) Signal frames can be divided into three segments: synchronization segment, control segment, and data segment;

The control segment includes information about user identification (ID) and the message data format or the message length. The data segment carried data content according to the specified format.

Since the Beidou Global Message service is based on MEO satellites constellation to achieve continuous coverage of the Earth's surface, according to the MEO height 21,500 km [2], the global beam angle of the MEO satellite-to-earth can be calculated that is about 26.5° [1], and the receiving gain of satellite antenna is only about 10 dB. According to the channel parameters given in [5], the Comparison of Uplink Calculations between Beidou Global Message and Regional Message are shown as below (Table 1).

Table 1. Comparison of Uplink Calculations between Beidou Global Message and Regional Message

Parameter	MEO	GEO
Direction of transmission	Earth-to-Space	Earth-to-Space
Coverage area	Global	Regional
Frequency bands (GHz)	1.6	1.6
Nominal user EIRP (dBW)	8 to 12 (select 8)	3 to 10 (select 8)
Max trans distance (km)	26600	41000
Space trans loss (dB)	−185.0	−188.7
Polarization loss and others (dB)	−1	−1
Satellite G/T (dB/K)	–16.5 to –14.5 (select −16.5)	−1
Typical C/N0 (dBHz)	34.0	45.8

According to the above table, the typical C/N0 of the BeiDou Global Message arrival signal at satellite is about 34.0 dBHz, which is about 11.8 dB lower than the typical C/N0 of the BeiDou Regional Message. This has caused the biggest difficulty in the design of the Beidou Global Message signal and the implementation of satellite message receiver payload. In order to achieve reliable reception of the short burst signals for non-stationary satellites in very low C/N0 conditions, we need to solve these key problems that include how long the signal synchronization segment is enough, and how to quickly acquire the short burst signal on the satellite.

3 Design of Synchronization Segment Length in BeiDou Global Message Signal

Due to the high dynamic, short burst and low C/N0 of the BeiDou Global Message Signal, sufficient synchronization segment length should be reserved in the signal design to ensure that the MEO satellite payload can accurately and reliably completed acquisition in the signal receiving process. The design of the synchronization segment length is essentially an analysis process of the fast acquisition performance of the short burst signal. In this paper, an acquisition performance evaluation method is used based on the constraint of constant false alarm, detection probability and process loss, and the signal synchronization segment length is optimized.

According to the principle of matched filtering, the quick acquisition processing of spaceborne short burst signals can utilize the method of pre-detection coherent integration and post-detection noncoherent accumulation (referred as "coherent integration-segment accumulation" method), which is also a common method in GNSS continuous signal acquisition. The difference just is that the short burst signal needs to be captured in real-time and full time domain, but the continuous signal only needs the fragmentation acquisition with quasi-offline processing. The implementation structure of the above acquisition method is illustrated in Fig. 1.

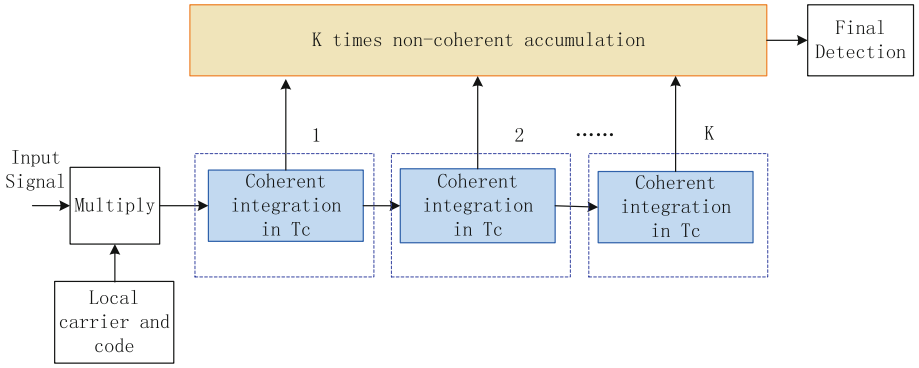


Fig. 1. Real-time matched filter capture structure

The above acquisition calculation process is mainly composed of two parts: coherent integration and segmentation accumulation. The result of a single coherent integration can be expressed as:

$$SNR_B = \frac{CT_c}{N_0} \cdot R^2(\tau) \cdot sinc^2(\pi f_d T_c) \tag{1}$$

Where, C is the signal arrival power, N_0 is the noise spectral density, T_c is the coherent integration duration, $R(\tau)$ is the autocorrelation function between the local code and the received spreading code, and f_d is the Doppler estimation error. Therefore,

$R^2(\tau)$ in the above equation represents the correlation loss caused by the spread code deviation, and $sinc^2(\pi f_d T_c)$ represents the correlation loss due to the Doppler shift.

During the acquisition process, the coherent integration results are first subjected to envelope detection and then segmentation accumulation. Considering the envelope detection loss [6], the final output signal-to-noise ratio (SNR) is as follows:

$$SNR_{out} = \frac{(SNR_B)^2}{SNR_B + 2.3} \cdot K \tag{2}$$

Where, K is the number of non-coherent accumulations. Then, $T = KT_c$ is the total duration of signal acquisition.K

According to the signal detection theory [7], under the condition of specifying the false alarm probability P_{fa} and the acquisition probability P_d, the relationship between the SNR and the acquisition performance obtained by ideal matched filtering is as follows:

$$SNR_{thd} = \frac{1}{2}\left(Q^{-1}(P_{fa}) - Q^{-1}(P_d)\right)^2 \tag{3}$$

The following figure illustrates the relationship between the detection probability and the required SNR under the condition of constant false alarm probability $P_{fa} = 10^{-6}$ (Fig. 2).

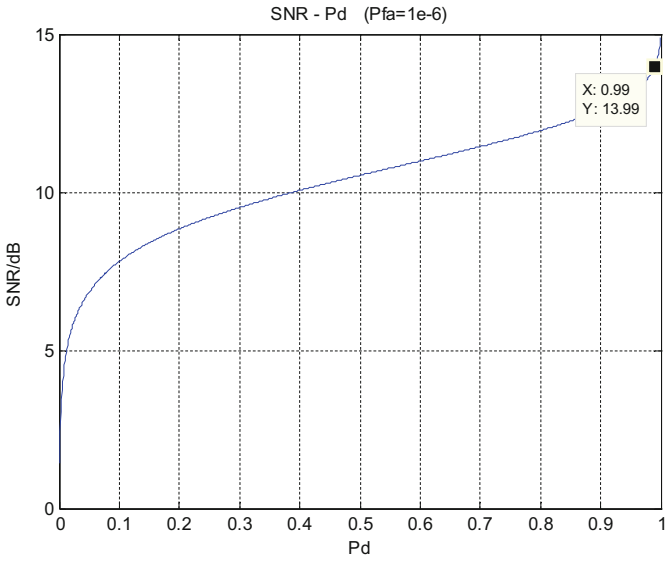


Fig. 2. Relationship between detection probability and required signal-to-noise ratio

According to the above curve, when the detection probability requirement is $P_d = 99\%$, the minimum required detection SNR threshold is about 14 dB. Therefore, the ideal length of the signal sync segment can be expressed as Eq. (4), without considering any acquisition implementation loss. Bring the SNR threshold and the arrival C/N0 parameters into this equation, T_{min} is only 10 ms.

$$T_{min} = SNR_{thd} - \frac{C}{N_0} \tag{4}$$

However, as a practical engineering system, it is impossible for the signal without power loss during processing. Therefore, the main task of the design is to optimize the acquisition algorithm parameters, minimize the Doppler deviation loss, code deviation loss, envelope detection loss and other processing implementation losses, and then design a synchronization segment length scheme that meets the actual requirements.

Based on the above analysis, this paper proposes a message synchronization segment length formula under multiple constraints with constant false alarm, detection probability and realizing loss, as shown in the following Eq. (5). This formula also can be used as a general method for the design of the synchronization segment length of the short burst signal and the acquisition duration for signal processing.

$$\begin{aligned} &T = KT_c \left| SNR_{out} \geq SNR_{thd} \right. \quad \text{or wrote as} \\ &T = KT_c \left| \frac{(SNR_B)^2}{SNR_B + 2.3} \cdot K \geq SNR_{thd} \right. \end{aligned} \tag{5}$$

4 Acquisition Algorithm Optimization

In order to perform the quick acquisition processing in the satellite message receiver, the above-mentioned "coherent integration-segment accumulation" acquisition algorithm needs to be further optimized, and the optimization goal is to minimize the acquisition integrated processing loss. The parameters optimized in the acquisition algorithm include two parts: the coherent integration duration and the post accumulation count. The optimized total acquisition duration is also the length of the synchronization segment in the burst message signal.

Before the simulation calculation, we need to pay attention to the optimization condition limit:

(1) satisfies $T_c < 1/f_d$, otherwise the Doppler loss in the coherent integration may cause the signal energy to be completely lost under certain circumstances.
(2) satisfies $T_c < T_{symbol}$, because the coherent integration is better not to cross the symbol bit, avoid causing positive and negative cancellation, and reduce the coherent integration value.

Firstly, assume that the maximum code deviation phase is less than 1/12 chip, and the code coherent loss can be expressed as:

$$L_\tau = 1/R^2(\tau) = -10log_{10}(1 - 1/12)^2 \approx 0.76\,dB$$

Then, the optimized parameters coherent integration time and number of accumulation times of are numerically simulated respectively, and the optimization results are given in follow.

(1) Optimization of the coherent integration time Tc

Under the condition of arrival signal C/N0 34 dBHz, with different Doppler deviation such as 500 Hz, 1000 Hz, etc., the code deviation loss is L_τ, the relationship between the detection output power loss and the coherent integration time obtained is shown in Fig. 3 by the simulation calculation (Fig. 4).

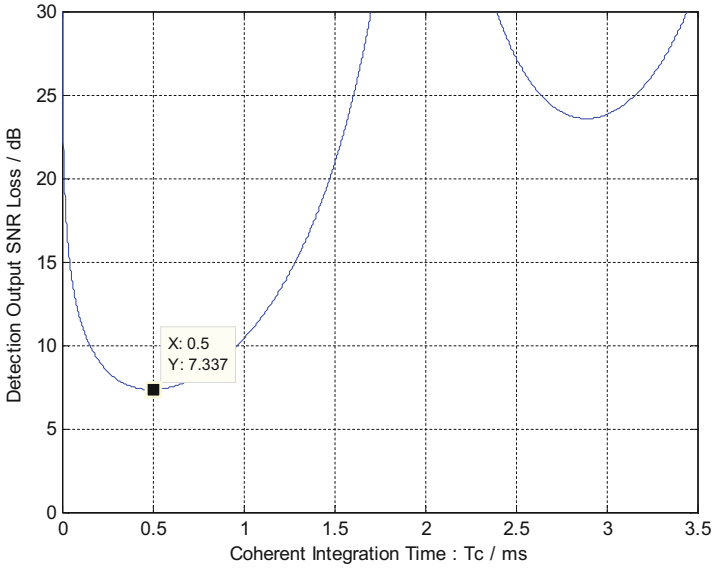


Fig. 3. Relationship between detection output loss and coherent integration time (f_d = 500 Hz)

The above curves illustrate that when the Doppler deviation is 500 Hz, the detection output loss of the acquisition process is minimum at 0.5 ms coherent integration, and when Doppler deviation is 1000 Hz, the detection output loss is minimal at the coherent integration of 0.256 ms. It can be seen from Eqs. (2) and (3) that the main factors affecting the detection output loss result are the Doppler deviation and the received C/N0. Therefore, the influence of Doppler deviation and received C/N0 on the loss results are analyzed by optimizing the coherent integration time under different Doppler and different C/N0 conditions.

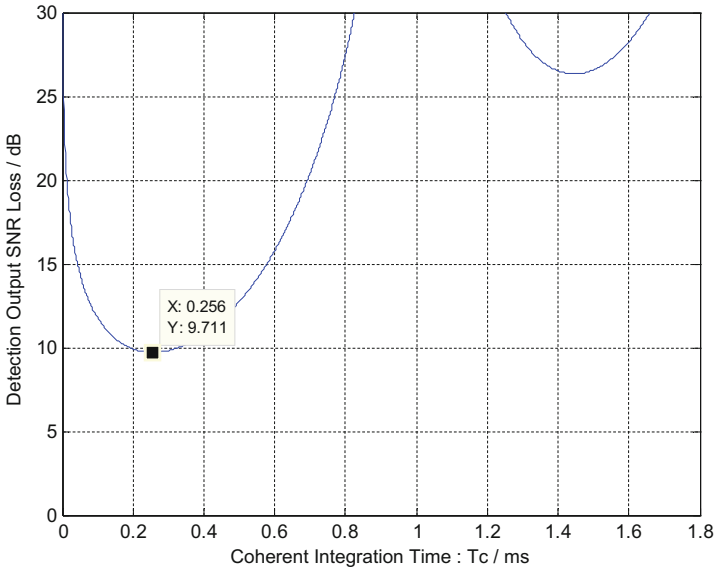


Fig. 4. Relationship between detection output loss and coherent integration time (f_d = 1000 Hz)

(1) Different Doppler deviation conditions

When the Doppler deviation range is 100 Hz to 1000 kHz, and other simulation conditions are consistent with the above, the final coherent integration optimization results are shown in the below Table 2.

Table 2. Coherent integration optimization results with matched filters under different Doppler conditions

Doppler deviation	Optimal Coherent integration time
100 Hz	2.0 ms
200 Hz	1.125 ms
300 Hz	0.8 ms
400 Hz	0.61 ms
500 Hz	0.5 ms
1000 Hz	0.256 ms

(2) Different C/N0 conditions

When the received C/N0 range is 32 dBHz to 40 dBHz, and other simulation conditions are consistent with the above, the final coherent integration optimization results are shown in the below Table 3.

Table 3. Coherent integration optimization results with matched filters under different C/N0 conditions

Received C/N0	Optimal Coherent integration time
32 dBHz	2.17 ms
34 dBHz	2.0 ms
36 dBHz	1.84 ms
38 dBHz	1.65 ms
40 dBHz	1.50 ms

It can be seen from the above two comparison tables that as the Doppler deviation increases, the coherent integration time optimization result decreases inversely; when the received C/n0 increases, the coherent integration time optimization result increases.

(2) Optimization of accumulation times K

After obtaining the optimized coherent integration time, the number of accumulation times is evaluated so that the detection output SNR can exceed the threshold. When the Doppler deviation is 100 Hz, the code deviation loss is L_τ, and the C/N0 is 34 dBHz. The SNR detection threshold is about 14 dB from the curve of Fig. 3. According to the above conditions and the optimized coherent integration time Tc, the detection output SNR obtained by different accumulation times are shown in the below Fig. 5.

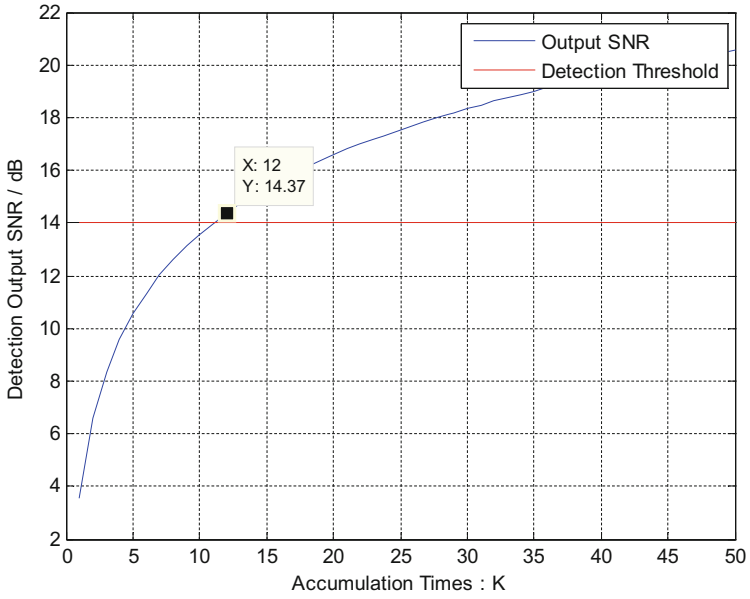


Fig. 5. Accumulation times and detection output C/N0

According to the above, under the condition of receiving C/N0 34 dBHz and maximum Doppler deviation 100 Hz, the optimized coherent integration time is 2 ms. As can be seen from Fig. 5, when the number of accumulations is at least 12, as well as the output SNR is 14.37 dB, that exceed the ideal threshold with the matched filter and meet false alarm rate and detection probability. Therefore, in order to meet the BeiDou Global Message signal acquisition requirements, to achieve high dynamic, low C/N0, high-reliability detection of short burst signals, the user signal synchronization segment length is at least 24 ms.

5 Test and Verification

Based on the message signal synchronization segment length and the on-board acquisition method proposed in this paper, we conducted a ground single-satellite device test & verification, as shown in Fig. 6. The single satellite device test & verification environment includes multi-user simulation equipment and satellite message receiver. We perform the test & evaluation under the receiving carrier-to-noise ratio range of 40 dBHz–33 dBHz, and test 1000 samples under each power condition. It can be seen from the results that under the constraint of false alarm rate of 10^{-6} and capture success rate of 99%, the sensitivity of the satellite message receiver reaches 34 dBHz, which confirms the theoretical analysis in this paper and validates the BeiDou Global Message sensitivity performance based on MEOs is 11.8 dB better than the Regional Message based on GEOs (Table 4).

Fig. 6. Ground test & verification equipments

Table 4. Ground test & verification results

No.	C/N0 of received signal	CRC check in data segment	Success rate
1	40 dBHz	√	100%
2	39 dBHz	√	100%
3	38 dBHz	√	99.90%
4	37 dBHz	√	99.90%
5	36 dBHz	√	99.80%
6	35 dBHz	√	99.75%
7	34 dBHz	√	99.75%
8	33 dBHz	×	93.69%

6 Conclusion

Based on the analysis of radio channel conditions about BeiDou Global Message and the requirements of message signal process, this paper proposes a synchronization segment length calculation method in short burst signal design constrained by constant false alarm and detection probability, and a set of optimized "coherent integration-segment accumulation" acquisition parameters. The acquisition simulation optimization provides feedback for the signal synchronization segment design. Finally, the implementation performance of the proposed method are fully verified by ground test.

The optimized synchronization segment design method proposed in this paper provides important theoretical support for BDS-3 Global Message signal design. The short message signal acquisition algorithm also provided a viable solution for the short message receiver on MEO satellites.

References

1. Tan S, Zhang T BeiDou multi-service satellite system and application. J Nat Univ Defense Technol 39(5):1–5
2. The State Council Information Office of the People's Republic of China (2016) China's BeiDou navigation satellite system, Foreign Languages Press, Beijing, June 2016
3. Beidou adds two satellites and surpasses GPS to become the world's first navigation system that provides Global Message functions. https://baijiahao.baidu.com/s?id=1612323620536294199&wfr=spider&for=pc Sept 2018
4. Shanghai Beidou Navigation Innovation Research Institute (2018) Beidou Satellite Navigation System Construction Progress and Application—BD3. http://www.sohu.com/a/274146940_99924008 November 2018
5. ITU-R (2018) Recommendation ITU-R M.1184-3 January 2018
6. Brown A, May M, Tanju B (2000) Benefis of software GPS receivers for enhanced signal processing. GPS Solutions 4(1):56–66
7. Kay SM (1993) Fundamentals of statistical and adaptive signal processing. Prentice Hall, pp 108–109

Quality Assessment of Galileo E1A Signal

Dejin Yang[1,2], Yongnan Rao[1,2(✉)], Huihui Shi[1,2], Xiaochun Lu[1,2], Li Kang[1,3], Xue Wang[1,2], Chengyan He[1,2], and Meng Wang[1,3]

[1] National Time Service Center, CAS, Xi'an, China
raoyongnan@163.com
[2] Key Laboratory of Precision Navigation Positioning and Timing, CAS, Xi'an, China
[3] University of Chinese Academy of Sciences, Beijing, China

Abstract. The E1A signal of Galileo system provides high-precision location information service for authorized users. The confidentiality of codes and the auto-correlation multi-peaks of high-order Binary Offset Carrier lead to the lack of relevant research on the quality assessment of E1A signal in the real environment. In this paper, two kinds of high-order BOC tracking algorithms including Bump-Jump and Double Estimator (DET) Technical are studied, then their principles how to eliminate the correlation curve multi-peaks are introduced and compared about advantages and disadvantages. Based on DET, the E1A tracking loop is initialized by using the E1B signal's carrier information, and then the spreading code modulated message symbol is accurately recovered by waveform-matching method, so the tracking problem of the E1A actual signal is solved. Based on the tracking information of E1A signal, we perform the evaluation of S-Curve Bias (SCB) and correlation loss, and compared those with E1B signal. The results show that the E1A signal correlation loss is greater than the E1B signal, and within 0.15 chip correlator spacing, both SCB are within 0.1 ns. The results have verified the effectiveness of the assessment method of E1A signal quality, as an important reference for improving signal analysis and assessment of GNSS.

Keywords: Galileo system · E1A signal · S-Curve Bias · Correlation loss · DET

1 Introduction

The E1A signal of Galileo system as a authorized service signal with the authorized spreading code and high-order BOC modulation, are used to provide high-precision location information service for specified users. Centered at 1575.42 MHz in the Aeronautical Radio-navigation Service band, the E1A signal adopt the high-order BOC modulation [1]. It is well known that BOC modulation can expand signal spectrum, improve the accuracy of code tracking and the ability of multipath suppression, and resist narrowband interference. The E1A signal is modulated by BOC (15, 2.5) with the aim to achieve better ranging accuracy and anti-jamming performance. However, there exists multiple peaks in the auto-correlation of high-order BOC signal, which lead to

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 410–420, 2019.
https://doi.org/10.1007/978-981-13-7751-8_41

false lock for the tracking loop. The research on eliminating the multiple correlation peaks of BOC signal has been a hot topic at home and abroad.

With the ability of monitoring the satellite's working status and the propagation environment's damage, signal-in-space quality assessment can find and solve problems in time, to provide important guarantee for the continuity and reliability of navigation signal service [2]. At present, there has been a lack of analysis and evaluation of E1A signal in real environment at home and abroad. Frank validates the effectiveness of the tracking algorithm by simulating BOC (15, 2.5) modulation data [3]. Wendel used 3 m antenna to collect E1 signals to verify the effectiveness of the high-order BOC tracking algorithm, however, it did not accomplish the signal quality evaluation [4]. In the paper, Galileo E1 signal is collected with 40 m high-gain antenna. The E1A signal tracking loop is initialized by demodulating carrier frequency and phase of E1B signal. The E1A spreading code modulated by navigation message symbol is recovered by using chip waveform-matching method. In order to solve the ambiguity of high-order BOC tracking, P. Fine proposed a Bump-Jump method which adds additional correlators and compares the correlation magnitude [5]. Hodgart proposed a more stable DET [6] which adds an independent Subcarrier Locked Loop to the DLL and PLL tracking loops, and treats subcarrier as a specific spreading code. It has high robustness but increases the design complexity of hardware receiver. Comparing two tracking algorithms, this paper chooses DET to get baseband signals for signal quality evaluation, and measures signal power loss and ranging accuracy from correlation loss and S curve zero-crossing bias, respectively. The achievements of this paper, the analysis and evaluation of the authorized E1A signal of Galileo system, can provide a worthy reference for authorized signal evaluation of Global Navigation Satellite System (GNSS).

2 E1A Signal Structure

Galileo E1 signal uses Interplex to modulate E1A, E1B and E1C signals on the same carrier frequency, as well as E1A modulated on the orthogonal branch, E1B and E1C on the In-phase branch. Baseband signal is given by following expression [7].

$$s_{E1}(t) = \sqrt{2P}\left(\begin{array}{l}(\cos(m)s_{E1B}(t) - \sin(m)s_{E1C}(t)) + \\ j(\cos(m)s_{E1A}(t) + \sin(m)s_{E1A}(t)s_{E1B}(t)s_{E1C}(t))\end{array}\right) \quad (1)$$

hereby, P denotes signal total power, m is the modulation coefficient, $s_{E1A}(t)$, $s_{E1B}(t)$ and $s_{E1C}(t)$ are the three baseband signal.

The E1A signal uses BOC (15, 2.5) modulation to move the signal spectrum to ±15.345 MHz. Due to its cosine subcarrier modulation, the subcarrier chip rate is twice faster compared with the sinusoidal subcarrier and the chip rate reaches 2*(2 * 15 * 1.023 MHz) = 61.38 MHz. Without considering the band limitation and channel distortion of the front-end filter, the E1A signal structure is simplified as:

$$X(t) = d * c(t) * b(t) * cos(2\pi f_c t + \theta) + n(t) \quad (2)$$

Where d is the message symbol, c(t) denotes the spreading code, b(t) is the bipolar rectangular subcarrier, and n(t) is additive white Gaussian noise, f_c is the carrier frequency, θ is the initial phase.

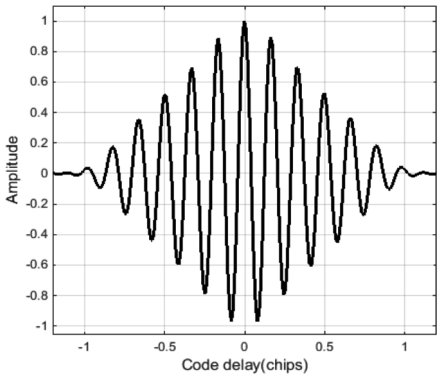


Fig. 1. BOC (15, 2.5) modulated signal correlation curve

Compared with Galileo E1B/C signal modulated by CBOC (6, 1, 1/11), the number of E1A signal auto-correlation peaks is much more than open service E1B/C signal, and the main peak width of correlation curve is 1/6 chip. As can be seen from Fig. 1, the correlation curve of BOC (15, 2.5) signals has the characteristics of multiple correlation peaks and narrow spacing among the peaks.

3 The Solution of E1A Signal Spreading Code

In order to carry out the analysis and evaluation of E1A signal, the spreading code of E1A must be solved firstly, because the spreading code of Galileo E1A signal is confidential to the civil users. According to the multiplex characteristics of E1 signal, the information of open service E1B signal is used to assist the demodulation of E1A signal. E1B and E1A signals are orthogonal phase and have the same doppler frequency, which facilitate the separation of signals. Suppose the stable tracking of E1B signal is t_k, the corresponding doppler frequency is f_k, carrier phase estimation is θ_k, These parameters are applied to E1A carrier loop initialization. E1A baseband signal as following:

$$B(t) = imag\left(s(t) * e^{j(2\pi(f_c + f_k)t + \theta_k)}\right) \tag{3}$$

where *imag*(.) denotes imaginary operation, Start time is t_k, $s(t)$ is intermediate frequency signal.

The baseband signal B (t) is composed of navigation message, spreading code and subcarrier. One spreading code chip modulated by cosine BOC (15, 2.5) has 24 subcarrier chips. Ideal subcarrier cycle sequence is [1 −1 −1 1], and the specific steps on solving the spreading code as follows:

1. The digital baseband signal is multiplied by 24 subcarrier chips to get the sampled code.
2. Summing on each chip and taking symbolic operation to obtain bipolar code.

$$C(k) = sign\left(\sum_{n=1}^{24} B((k-1)*24+n)*b(n)\right),\ k = 1, 2, 3\ldots10230 \tag{4}$$

where $sign(.)$ denotes symbolic operation, $b(n)$ is the subcarrier sequence.

In the process of analyzing E1A spreading code, coherent integration time with 4 ms is used to recover 10230 codes each time. Although the recovered spreading codes contain navigation messages, it does not affect the subsequent analysis and evaluation.

4 High-Rate BOC Tracking

Navigation signal is transformed into digital IF signal after RF front-end amplification, filtering and analog-to-digital conversion. The carrier frequency and code phase of IF signal are roughly estimated after acquisition. The satellite number, carrier frequency and code phase are used as the local carrier and spreading code of the tracking loop.

The traditional structure of tracking loop is PLL + DLL. PLL and DLL are composed of phase detector, loop filter and Numerically Controlled Oscillator (NCO). Carrier discriminator uses four quadrant arc-tangent to detect the phase error between the real signal and the local carrier, removes the high frequency component through low-pass filter, adjusts the carrier NCO and updates the local carrier frequency to converge to the real signal frequency. Using the symmetry of correlation curve, the code discriminator detects the difference of correlation magnitude between local advance code and lagging code. After loop filter, the difference makes NCO update the local code frequency, adjusting the local code delay to align the code delay of the real signal (Fig. 2).

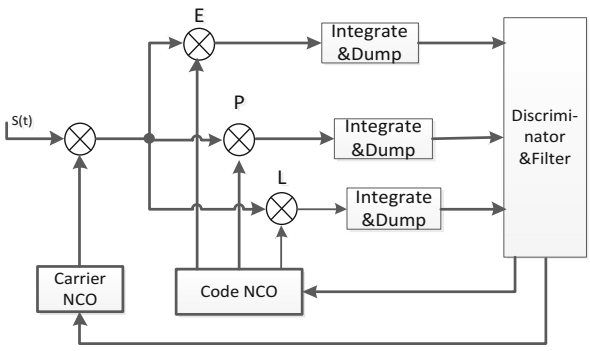


Fig. 2. Traditional tracking structure

4.1 Bump-Jump Algorithm

Multiple auto-correlation peaks of BOC modulation signal easily lead to false lock for the traditional tracking loop and the ranging results deviating from the true values. The Bump-Jump technology adds two additional correlators named VE and VL for the traditional tracking loop structure to monitor whether the main peak is locked correctly. In the Bump-Jump detection process, when the correlation values of early and late branch are equal, we can determine the loop locking state according to the VE and VL correlation values and adjust the local code delay. If $R_{VE} > R_{VL}$, the loop locks the side peak on the right side of the main peak, and the local code phase is increased; If $R_{VE} < R_{VL}$, the loop locks the side peak on the left side of the main peak, and the local code phase is decreased. In order to avoid a false lock caused by noise jitter, a threshold value is usually set to adjust the local code phase when the number of times $R_{VE} > R_{VL}$ *or* $R_{VE} < R_{VL}$ exceeds the threshold value. The Bump-Jump achieves BOC signal tracking by adding two additional correlators. The algorithm is simple in principle. Due to the difficulty in distinguishing the main and secondary peaks of signal correlation curve as in low signal-to-noise ratio, the detection technology based on the amplitude difference of side peaks cannot adjust the local code phase in the right direction.

4.2 Tri-Loop Tracking Algorithm

Tri-Loop tracking technology treats subcarrier as special spreading codes. Subcarrier Lock Loop (SLL) is added to traditional loop, that is, subcarrier discriminator, filter and NCO are added correspondingly, independent of PLL and DLL (Fig. 3).

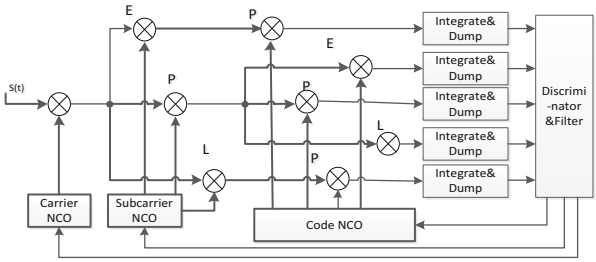


Fig. 3. DET tracking structure

Similarly to the spreading code, SLL's local early, prompt, and late subcarrier are expressed as follows:

$$\begin{aligned} S_e(t) &= b\left(t - \tilde{\tau} + \tfrac{T_s}{2}\right) \\ S_p(t) &= b(t - \tilde{\tau}) \\ S_l(t) &= b\left(t - \tilde{\tau} - \tfrac{T_s}{2}\right) \end{aligned} \tag{5}$$

hereby, $\tilde{\tau}$ is the subcarrier delay estimate, T_s is the subcarrier period.

The In-phase branch expression of correlator output:

$$\begin{aligned} p_{IPE} &= A * d * r(\tau^* - \tau) * a(\tilde{\tau} - \tau - T_s/2) * \cos(\Delta\theta) \\ p_{IPL} &= A * d * r(\tau^* - \tau) * a(\tilde{\tau} - \tau + T_s/2) * \cos(\Delta\theta) \\ p_{IEP} &= A * d * r(\tau^* - \tau - T_c/2) * a(\tilde{\tau} - \tau) * \cos(\Delta\theta) \\ p_{ILP} &= A * d * r(\tau^* - \tau + T_c/2) * a(\tilde{\tau} - \tau) * \cos(\Delta\theta) \end{aligned} \tag{6}$$

A is an amplitude of signal, T_c is code chip duration, $r(.)$ is the correlation function of spreading code, $a(.)$ denotes the correlation function of subcarrier. $\Delta\theta$ is the phase error between the received signal and local carrier.

In Tri-Loop tracking mode, DLL and SLL estimate the code phase separately. The spreading code has only one correlation peak in the code period, as well as the correlation main peak is wider and the estimate of code phase τ^* is an unambiguous but less accurate delay; The subcarriers have multiple correlation peaks in the code period, as well as the main peak width is narrow and estimate $\tilde{\tau}$ is ambiguous but more accurate.

$$\tilde{\tau} = \tau + p * T_s \tag{7}$$

Where τ is the actual subcarrier phase, p is the number of subcarrier period.

In order to make full use of the phase estimation accuracy of SLL loop and eliminate ambiguity, the phase of subcarrier is corrected by using DLL unambiguous code phase estimation.

$$\tau_{out} = \tilde{\tau} + round((\tau^* - \tilde{\tau})/(T_s/2)) * \frac{T_s}{2} \tag{8}$$

where *round* operation guarantees that the output result is not affected by the error when it is less than half of the subcarrier chip width.

5 Correlation Performance

Correlation performance is an important part of navigation signal quality evaluation. The external interference caused by digital distortion, error code and multipath can be reflected by correlation performance. Calculating correlation function is the basic work of evaluating correlation performance. After carrier wiping off, the received signal is correlated with the local spreading code, and the normalized correlation function is obtained as follow [8]:

$$CCF(\varepsilon) = \frac{\int_0^{T_p} S_{BB-\mathrm{Pr}e\,\mathrm{Pr}oc}(t) * S^*_{\mathrm{Ref}}(t-\varepsilon)dt}{\sqrt{\left(\int_0^{T_p} \left|S_{BB-\mathrm{Pr}e\,\mathrm{Pr}oc}(t)\right|^2 dt\right) * \left(\int_0^{T_p} \left|S_{\mathrm{Ref}}(t)\right|^2 dt\right)}} \tag{9}$$

where $S_{\mathrm{BB-PreProc}}(t)$ is the received satellite's baseband signal, $S_{\mathrm{Ref}}(t)$ is the local code. T_{p} is code period. Because of the correlation peak of a single period affected by noise jitter, the correlation peaks are usually calculated by means of accumulating data in multiple milliseconds.

5.1 Correlation Loss

The correlation loss describes the difference between the actual correlation power and the ideal correlation power of the navigation signal within the effective bandwidth. The correlation power of the signal is defined as:

$$P_{CCF}[dB] = \max_{all\,\varepsilon}(20 * \log_{10}(|CCF(\varepsilon)|)) \tag{10}$$

The correlation loss defined as:

$$CL[dB] = P_{\underset{ideal}{CCF}}[dB] - P_{\underset{actual}{CCF}}[dB] \tag{11}$$

where $P_{\underset{ideal}{CCF}}$ denotes correlation power of ideal signal, $P_{\underset{actual}{CCF}}$ denotes correlation power of actual signal.

In the progress of the navigation signal is generated, propagated and received, signal's amplitude may be reduced and the correlation loss of the signal will increase, such as power reduction caused by limited signal bandwidth of satellite load, unbalanced signal amplitude due to non-linearity of high power amplifier, multiple signal components modulated on the same carrier frequency by multiplexing technology, and correlation loss caused by signal interaction [9].

5.2 S-Curve Biases

The S-curve is the function of the output of the code phase discriminator about the local code delay. Ideally, the receiver is locked at the zero value of the S-curve stably, and the code phase bias is zero. In real environment,, when the signal arrives at the ground, there will be various degree of distortion due to the influence of band-limited filter, noise and ground multipath of transmission channel. SCB varies with the change of the correlator spacing [10]. The difference of S-curve among different discriminator algorithms is mainly reflected in S-curve gain and linear interval width. Taking the typical incoherent Early-Minus-Late Power discriminator as an example. The S-curve can be calculated from the following equations [11]:

$$\mathrm{S}(\varepsilon,\delta) = \left|\mathrm{CCF}\left(\varepsilon - \frac{\delta}{2}\right)\right|^2 - \left|\mathrm{CCF}\left(\varepsilon + \frac{\delta}{2}\right)\right|^2 \tag{12}$$

Where δ is the correlator spacing, ε is the code phase bias.

The lock point error $\varepsilon_{bias}(\delta)$ meets the following equation:

$$S(\varepsilon_{bias}(\delta), \delta) = 0 \tag{13}$$

Different modulations have different characteristics of S-curve, and BOC signal's S-Curve has multiple zero-crossings, so S-Curve Biases are defined as:

$$SCB = \max_{over\ all\ \delta}(\varepsilon_{bias}(\delta)) - \min_{over\ all\ \delta}(\varepsilon_{bias}(\delta)) \tag{14}$$

6 Results

In this paper, Galileo GSAT 0204 satellite signal is collected from the 40-m parabolic antenna located at HaoPing Radio observatory(HRO), National Time Service Center of Chinese Academy of Sciences. The antenna gain is about 51 dBi, the sampling rate is 250 MHz, the Quantification is 14 bits, and the intermediate frequency of the signal is 62.5 MHz. The power spectrum of the signal is shown in the Fig. 4.

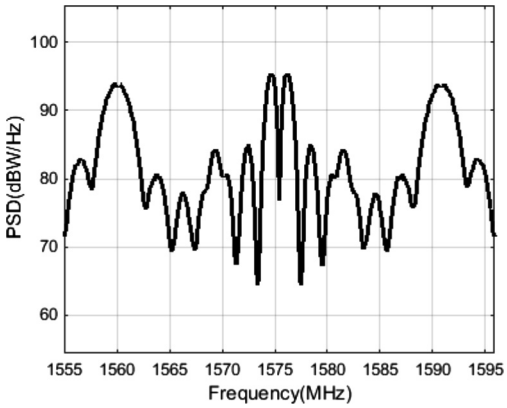


Fig. 4. Power spectral density of Galileo E1 signal

6.1 Tracking Results

Tri-Loop tracking algorithm is used to process Galileo E1A signal. After the E1B signal is tracked and stabilized, the carrier information is provided to the E1A tracking loop to ensure accurate information.

The tracking time of E1A signal is 1 s, the code correlator spacing is 78.2 ns, and the subcarrier correlator spacing is 4.88 ns, noise bandwidth of PLL is 10 Hz, noise bandwidth of DLL and SLL is 5 Hz. In the receiver parameter settings, the loop tracking time is determined by the length of the collected data. The correlator spacing and the loop noise bandwidth are both set with the common receiver parameter setting as a reference. 78.2 ns corresponds to 0.2 times the spreading code width, and 4.88 ns corresponds to 0.3 times the subcarrier chip width. Because the collected data based on

high gain antenna is characterized by low noise, the selected noise bandwidth of DLL is a litter large than that of the common receiver, in order to accelerate the convergence speed of the loop.

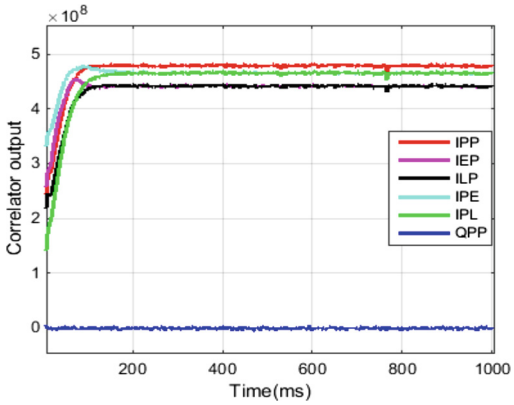


Fig. 5. I/Q branch correlator output

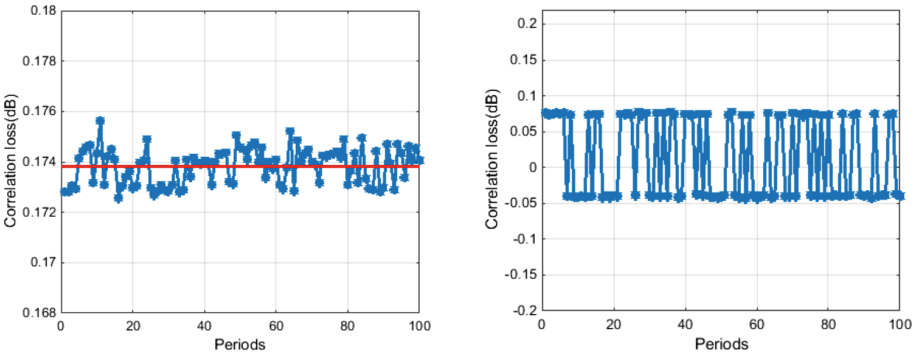
In Fig. 5, IPP and QPP denote I/Q prompt branch outputs respectively, IEP and ILP represent the early and late correlation outputs of spreading code, and IPE and IPL represent the early and late correlation outputs of subcarrier, respectively. The receiver tracks the E1A signal steadily, and the correlation output has unique polarity, because the local code has navigation message and is not affected by symbol flipping. The correlator spacing of subcarrier is less than that of spreading code, and the amplitude is more than that of spreading code, correspondingly

6.2 Signal Assessment Results

From Fig. 6 (a), it shows the correlation loss of 100 integration cycles of E1A signal, whose value fluctuates from 0.172 dB to 0.176 dB. The red line represents the average value of correlation loss and the value is 0.1738 dB. For comparison, it shows the E1B signal correlation loss, whose value is less than 0.1 dB, and the up-and-down jitter of the correlation loss is caused by mutual interference with the same-frequency in-phase E1C signal in Fig. 6 (b).

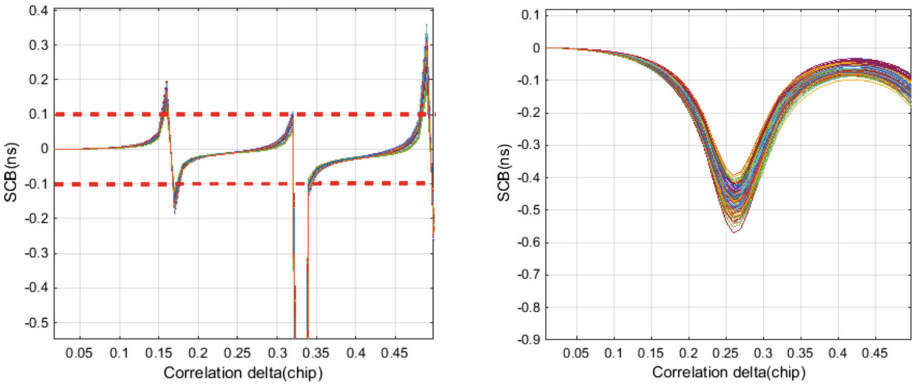
The S-Curve Biases directly reflects the receiver's ranging accuracy. It shows the 100 SCB curves of E1A and E1B signal in Fig. 7, respectively.

As shown in Fig. 7 (a), the trend of the 100 SCB curves are basically the same, indicating that the ranging results are stable at each time interval. Because there are many nonlinear intervals in the correlation curve, SCB has multiple hops. In the linear range of the correlation curve from 0.01 to 0.15 chips, the absolute value of the E1A signal's SCB is less than 0.1 ns, indicating that the E1A signal is better than 0.03 m without interference from the ground environment. For comparison, the SCB of the



(a) Correlation loss of E1A signal **(b)** Correlation loss of E1B signal

Fig. 6. Correlation loss of Galileo E1A and E1B signal



(a) S-Curve Biases of E1A signal **(b)** S-Curve Biases of E1B signal

Fig. 7. S-Curve Biases of Galileo E1A and E1B signal

E1B signal is shown in Fig. 7 (b). The absolute value of SCB is less than 0.1 ns in the 0.15 chip correlator spacing, which is basically consistent with the E1A signal.

7 Conclusion

In this paper, the Galileo satellite signal in real environment is collected by 40 m antenna, and the spreading code of E1A signal is analyzed by using the carrier information of E1B. The tracking ambiguity problem caused by the multiple correlation peaks of E1A signal is successfully solved by using DET algorithm. In this paper, the correlation loss and SCB of E1A signal are analyzed in details. The average value of correlation loss is 0.1738 dB, and the absolute value of SCB in linear interval of 0.15 chips is no more than 0.1 ns, Compared with the E1B signal, the correlation loss of the

E1A signal is greater, indicating that the inter-modulated signal influence of the branch of the E1A signal is greater than the interference between the E1B and E1C signals. Both SCB are less than 0.1 ns in the 0.15 chip correlator spacing. This paper fills in the blank of E1A signal quality assessment in the real environment. On this basis, the other navigation systems' authorized signals with high-order BOC modulation can be studied, subsequently.

References

1. Betz JW (2001) Binary offset carrier modulations for radio-navigation. Navigation 48:227–246
2. Lu X, Zhou H (2010) GNSS signal-in-space quality analysis and research. China Sci Phys Mech Astron 40(5):528–533
3. Schubert FM, et al (2014) The astrium correlator: unambiguous tracking of high-rate BOC signals. In: IEEE/ION position location and navigation symposium (PLANS), Monterey, CA, USA, May 2014
4. Wendel J, Schubert FM et al (2014) Validation of PRS tracking algorithms using real life signals. In: Proceedings of the 27th international technical meeting of the ION satellite division, ION GNSS+ 2014, Tampa, Florida, 8–12 September 2014
5. Fine P, Wilson W (1999) Tracking algorithm for GPS offset carrier signals. San Diego, CA
6. Hodgart MS, Blunt PD, Unwin M (2007) The optimal dual estimate solution for robust tracking of binary offset carrier (BOC) modulations. In: Proceedings of the 21th international technical meeting of the satellite division, Fort Worth, TX, 25–28 September 2007, pp 1017–1027
7. Xie G (2013) Principles of GNSS: GPS, GLONASS and Galileo. Publishing House of Electronics Industry, Beijing, p 111
8. Thoelert S, Furthner J (2013) GNSS survey - signal quality assessment of the latest GNSS satellites. In: Proceedings of the 2013 international technical meeting of the institute of navigation, San Diego, California, January 2013, pp 608–615
9. Yang D, Lu X, Wang X (2017) Analysis and evaluation of Galileo E1 signal. J Time Freq. 40(3):178–192. https://doi.org/10.13875/j.issn.1674-0637.2017-03-0178-15
10. Kang L, Lu X, Wang X, He C, Rao Y, Yang D (2018) Authorized signals quality assessment on GPS L1. J Electron Inf Technol 40(4)
11. Kakos DM, Esterhuizen S, Mitelman A (2004) High gain antenna measurements and signal characterization of the GPS satellites. In: Proceeding of the institute of navigation GNSS meeting, Long Beach, CA, pp 1724–1731

Analysis of Interaction Between Navigation Payload and Constant Envelope Design of Navigation Signal

Ying Wang(✉), Tao Yan, Guoyong Wang, Lang Bian, and Yansong Meng

China Academy of Space Technology (Xi'an), Xi'an 710000, China
eaglesoars@l26.com

Abstract. Navigation signal design considers many factors, such as service requirements, in-heritance, compatibility and interoperability, and patent avoidance. The only consideration for the realization of navigation load is constant envelope. However, the constant envelope in current signal design is based on the assumption of unlimited bandwidth. In practice, satellite load cannot be realized. From the load point of view, there are two main types of loads on navigation satellites: one is linear devices, such as filters, which are mainly used to prevent aliasing, suppress mirror and extra-band radiation, prevent interference with other systems and achieve maximum useful power; the other is devices with nonlinear effects, such as power amplifiers. In principle, linear devices before power amplifier can compensate in-band characteristics by analog and digital means, but power amplifier introduces more complex effects, which are related to multicarrier, power ratio of signal components and constant envelope characteristics of signals. It is the common responsibility of load design and signal design to avoid or reduce the adverse effects caused by these practical devices. In this paper, the deterioration of Beidou navigation signal in navigation load is analyzed, and the influence of constant envelope deterioration of navigation signal near the saturation point of amplifier on signal quality is verified by simulation. By generating a navigation signal, it passes through a real L-band fixed amplifier. The test results show that the specific signal constant envelope design has little effect under the specific trans-mitting bandwidth condition.

Keywords: Payload distortion · Navigation signal · Constant envelope

1 Introduction

At present, taking GPS modernization, Galileo and BDS3 system construction as the representative, the design of navigation signals of multiple systems has absorbed the successful experience of GPS. On this basis, the different needs of military and civil businesses are considered. At the same time, a great deal of energy is spent on intellectual property rights, compatibility and interoperability. However, when designing signals, the realization of satellite loads is considered. Less consideration makes the final signal state not reach the desired state of the designer, and limits the further improvement of system performance in some aspects.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 421–431, 2019.
https://doi.org/10.1007/978-981-13-7751-8_42

An actual device component is always unsatisfactory. Always various errors and parts cannot be modelled accurately. Navigation loads need to generate and transmit signals with time-frequency information, and the channel part is not ideal mainly because of filter and amplifier components.

The filtering in navigation payload mainly distributes in three places: the first is digital baseband filter before IF modulation, mainly to prevent aliasing during modulation; the second is before power amplifier, mainly to filter the mirror after frequency conversion, while suppressing local oscillation leakage; the third is after power amplifier, the purpose is to suppress aliasing. Extra-band radiation can prevent interference to satellite receiving equipment and other adjacent services.

At present, there are two main technical routes of power amplifier on navigation satellite: one is traveling wave tube amplifier (TWTA) based on microwave field technology, which has wide bandwidth, high gain, high efficiency and is a mature and reliable amplification technology; the other is solid-state amplifier (SSPA) based on semiconductor technology [1]. Solid-state amplifiers have gradually overcome the shortcomings of low output power in recent years, and have begun to be applied in the field of navigation loads with high reliability and long life.

Most of the satellite navigation loads of GPS and BDS in China are equipped with traveling wave tube amplifiers. In the development stage, Galileo satellite of the European Union carried out a lot of demonstrations and tests on solid-state amplifiers, confirming that they meet the navigation requirements, and finally equipped with solid-state amplifiers [2–4]. Because of the small size and lightweight of solid-state amplifier, the output power and efficiency are gradually improved, which has become the development direction of space-borne power amplifier.

For multi-carrier signals, traveling wave tube amplifiers generally need to return 4–6 dB due to the non-linear effect of amplifier superposition, while solid-state amplifiers generally only return 3 dB [1]. However, for navigation loads, platform constraints and landing power requirements, efficiency must be considered, so the general navigation load power amplifier works near the saturation point. On the one hand, the nonlinearity is related to the characteristics of power amplifier itself. On the other hand, it is closely related to the constant envelope characteristics and multi-carrier characteristics of signals. Different signal systems have different additional effects.

Navigation signal structure design is an important aspect of satellite navigation system. Navigation signal is the only channel for interaction between satellite navigation system and users. Therefore, how to ensure the quality of space signal should be an important aspect of signal design. The real signal design should be completed after the comprehensive evaluation of the load on-orbit verification. On the one hand, the organizer of the system construction should stimulate the innovation and progress of the load design unit to improve the level of load development. In addition, based on the actual device capacity limitation and technical level, the organizer should put forward requirements for signal design to achieve the optimal performance of the current system and future performance improvement.

2 Signal Model

2.1 Signal Model

The model of navigation signal generation and transmission process in navigation payload is shown in Fig. 1.

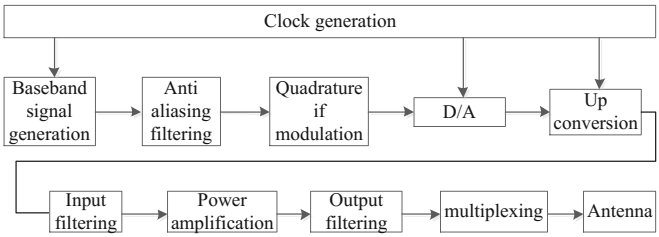


Fig. 1. Functional block diagram of navigation payload signal generation

Baseband signal generation, anti-aliasing filtering and orthogonal IF modulation are all processed in digital domain. Up-conversion, input filtering, power amplification, output filtering, multiplexing and antenna are all processed in analog domain.

For the signal, from the receiver's point of view, the user's main concern is the measurement performance of the signal. The performance of code can be measured by describing various characteristics of correlation function, including correlation loss and S curve deviation describing symmetry.

By analyzing the signal generation and transmission process mentioned above, the signal generation process in the digital domain can be accurately simulated, and the Sinc fading of digital-to-analog conversion can be compensated by the digital domain, because the homology can be considered that the up-conversion does not deteriorate the code and carrier of the signal. Input filtering can compensate for the unsatisfactory characteristics of in-band amplitude and phase by predistortion in digital domain before power amplification. In addition, output filtering, multiplexing and antenna can be considered to exert only linear distortion effects. Thus, the navigation load model can be simplified as follows (Fig. 2):

Fig. 2. Simplified block diagram of navigation payload signal generation

The pre-filtering includes all the band-limiting effects in digital and analog domains. The filtering attributes destroy the constant envelope attributes of the signal and cause the fluctuation of the signal envelope, which makes the power amplifier exert a non-linear influence on the signal. Therefore, it needs to be taken into account.

Post-filtering is mainly used to suppress extra-band radiation, to prevent interference to other satellite equipment and other adjacent band systems, and to influence the performance of navigation signal mainly by distorting the in-band amplitude-frequency and phase-frequency. If the predistortion operation is used to compensate in the digital domain, the predistortion will inevitably result in the loss of the constant envelope performance of the signal. This effect makes the power amplifier exert more non-linear characteristics, thus causing a vicious circle. Therefore, it is difficult to eliminate this part thoroughly by the predistortion method in the digital domain, and it needs the combination of signal systems. Design, optimize the performance of the device itself and control the way combined with predistortion.

In fact, the non-constant envelope signal, including the signal passing through the filter after the constant envelope, has fluctuations in the power envelope. On the AM-AM transfer characteristic curve, the working point is broadened to the working area (Fig. 3).

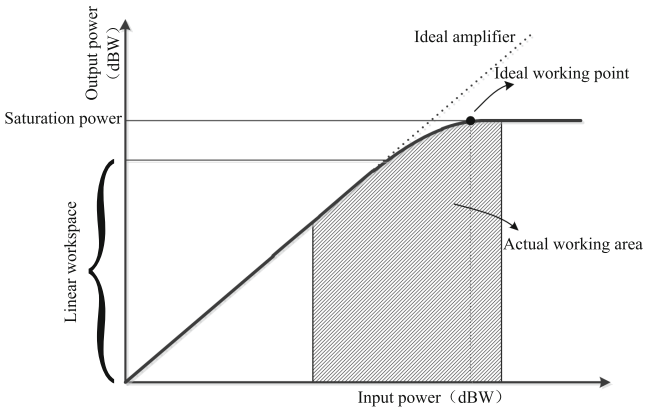


Fig. 3. The diagram of amplifier working point

The non-linear load model is established in reference [1], and the measured power characteristics are converted to the amplitude domain (Fig. 4).

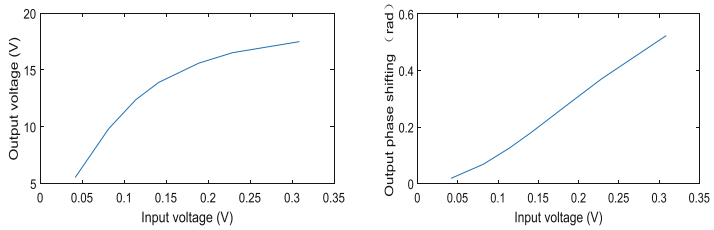


Fig. 4. Transfer curve of power amplifier voltage-voltage and voltage-phase

Pre-filter and power amplifier are investigated. Since the power amplifier can use the equivalent baseband model [2], the baseband signal is written as

$$s_0(t) \tag{1.1}$$

After the pre-filter

$$s_1(t) = s_0(t) * h(t) = m(t)e^{j\varphi(t)} \tag{1.2}$$

After power amplification

$$s_2(t) = U[m(t)]e^{j(\varphi(t)+\phi[m(t)])} \tag{1.3}$$

Fig. 5. Equivalent simplified model of power amplifier

The voltage domain power amplifier model and parameters [1]

$$U[m(t)] = \frac{\alpha_a m(t)}{1+\beta_a m^2(t)}, \quad \phi[m(t)] = \frac{\alpha_b m^2(t)}{1+\beta_b m^2(t)} \tag{1.4}$$

Where, the parameters α_a, β_a, α_b, β_b are obtained by fitting the graph and curve (Fig. 5).

2.2 Signal Evaluation

To measure the symmetry of correlation peaks, the S-Curve Bias (SCB) is used here. The S-Curve is defined as [5].

In order to evaluate the power variation of the constant envelope synthesis signal and the non-constant envelope synthesis signal through the power amplifier, the correlation method as shown in Fig. 6 is used here.

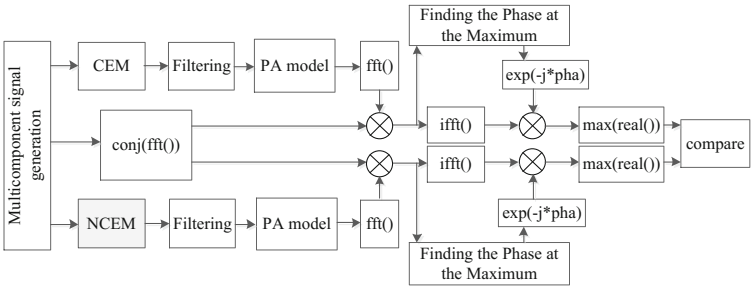


Fig. 6. Power evaluation scheme

The baseband component signals are generated, and then the constant envelope modulation (CEM) and non-constant envelope modulation (NCEM) are synthesized respectively. Then using the same filter and power amplification (PA) model, the correlation power is finally calculated and compared, and the power ratio change is obtained.

3 Simulation Analysis

BDS B1 navigation signal is a typical multi-carrier signal, which is more complex than ordinary OFDM signal, including MBOC modulation signal [6, 7, 8]. In order to achieve constant envelope, more intermodulation components are added, which leads to the addition of a series of different functions to the power amplifier. Therefore, B1 signal is used for simulation analysis.

From ICD, it can be concluded that B1 signal has at least three open signals: B1I, B1Cd and B1Cp. Among them, B1Cp is a QMBOC signal, which contains two orthogonal components of BOC(6,1) and BOC(6,1) [9]. From the analysis of 6 m antenna data, we can conclude that there is another BOC (14,2) signal on B1 signal (Fig. 7).

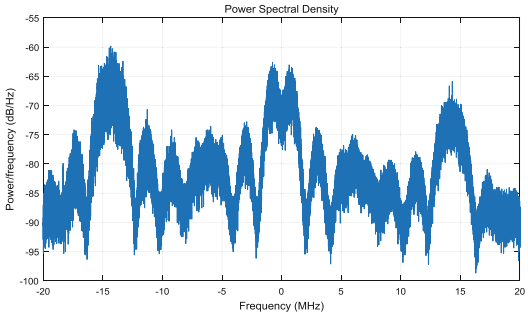


Fig. 7. Signal power spectrum (CAST Xi'an 6 m aperture dish antenna data)

Referring to POCET method and the design of BDS B1 constant envelope method in reference [7–11], the constant envelope mode of signal for simulation is obtained, and the simulation signal is obtained. Sections 3.1 to 3.4 are based on this analysis. Section 3.4 designs a non-constant envelope synthesis signal, and compares its partial performance with that of constant envelope synthesis signal.

3.1 Power Spectrum

As shown in Fig. 8, the baseband signal is generated at 500 MHz sampling frequency. When the baseband signal is filtered directly into the power amplifier model without power amplifier pre-filtering, it can be seen from the power spectrum that there is almost no loss, and B1 signal is a typical complex multi-carrier signal.

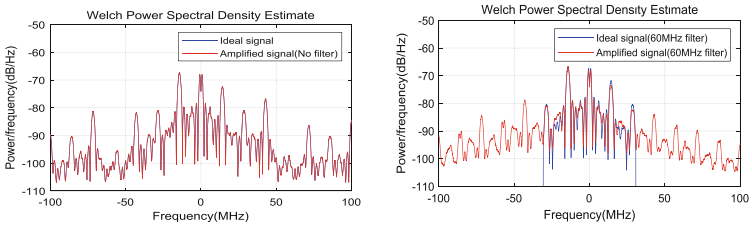


Fig. 8. Signal power spectrum

The ideal signal is filtered with a 100 MHz bandwidth ideal filter. The power spectrum before and after the power amplifier are compared with that before the power amplifier. The nonlinearity of power amplifier makes out-of-band spectrum regeneration.

The ideal signal is filtered with a 60 MHz bandwidth ideal filter. The power spectrum before and after the power amplifier are compared with that before the power amplifier as shown in Fig. 8(b). The regeneration of out-band spectrum is different from that of 100 MHz filtering. The attenuation trend is more obvious and the regeneration spectrum is more regular, which is closely related to the 14*1.023 MHz spectrum.

A 40 MHz bandwidth ideal filter is used to filter the ideal signal. The power spectrum before and after the amplifier are compared with that before the amplifier. Out-band spectrum regeneration is similar to that of 60 MHz filtering.

According to the actual received signal power spectrum, the actual transmitted signal bandwidth is between 40–60 MHz, so the 40–60 MHz filtering is closer to the real situation.

3.2 Peak-to-Average Ratio

Peak-to-average ratio (PAR) can be used to measure the difference between peak power and mean power. If it is an ideal constant envelope signal, the ratio is 1 (0 dB). Peak-to-average ratio (PAR) mainly represents the range from mean to peak power, which corresponds to the most non-linear part of the power amplification transfer curve of power amplifier. Therefore, the smaller the value, the smaller the non-linear effect.

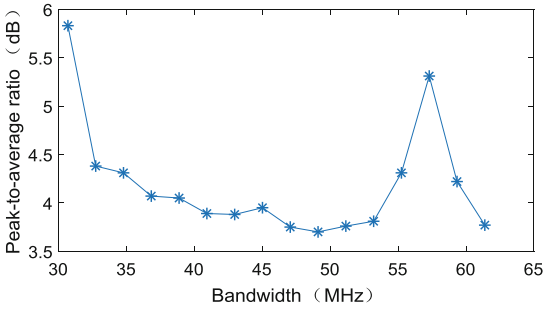


Fig. 9. Signal peak-to-average ratio

From the general trend (Fig. 9), the larger the bandwidth, the smaller the peak-to-average ratio and the less the non-linear effect. However, there is a higher energy of the cross-modulation component where the center frequency is (+28*1.023 MHz). The influence of multi-carrier makes it possible for the cross-modulation component to have a higher energy at the center frequency (+28*1.023 MHz).

It is not that the larger the bandwidth, the smaller the peak-to-peak ratio. Therefore, for this constant envelope signal, the optimal bandwidth before power amplifier should be 48*1.023 = 49.104 MHz, and the transmission bandwidth should be less than that.

3.3 Correlation Performance

As shown in Figs. 10 and 11, for B1Cd and B1Cp signals, compared with ideal unlimited bandwidth signal, the narrower the bandwidth, the larger the SCB.

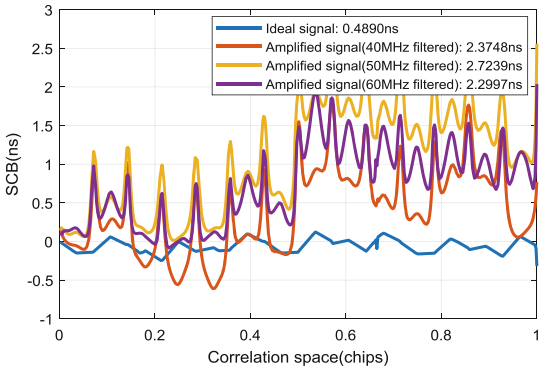


Fig. 10. B1Cd signal SCB

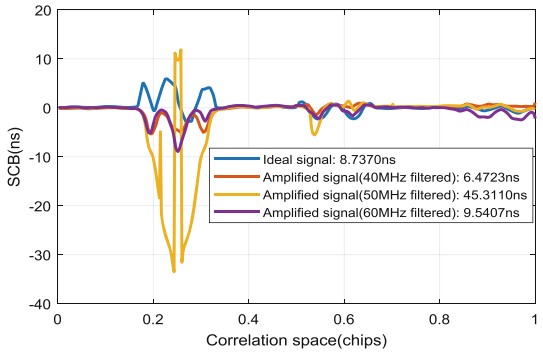


Fig. 11. B1Cp signal SCB

In addition, we can see that the symmetry of B1Cp signal at the first shoulder is worse under 50 MHz bandwidth filtering than other filtering conditions, and there is a very serious negative slope, where the receiver is not easy to lock correctly.

3.4 Performance of Non-constant Envelope Signal

Non-constant envelope signal is designed with reference [6, 7], the correlation power is shown in Table 1. Here, B1Cp signal SCB is shown in Fig. 12.

Table 1. Correlation power

Signal	B1Cd			B1Cp		
Bandwidth (MHz)	Unlimited bandwidth	40	50	Unlimited bandwidth	40	50
CEM	1.64	1.00	0.99	4.87	4.44	4.59
NCEM	1.66	1.00	0.99	3.07	5.17	5.13

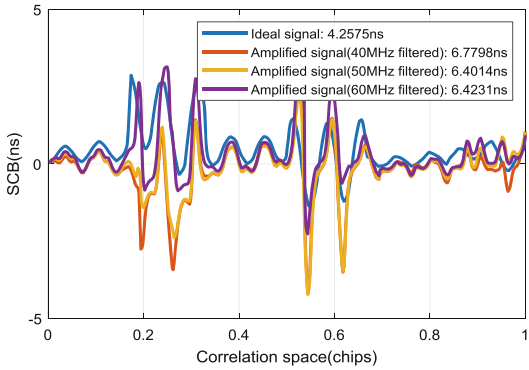


Fig. 12. B1Cp signal SCB

The performance of SCB under constant envelope and non-constant envelope is summarized, as shown in Table 2.

Table 2. SCB performance

Signal	B1I		B1Cd		B1Cp	
Bandwidth(MHz)	40	50	40	50	40	50
CEM SCB(ns)	1.19	1.16	2.37	2.72	6.47	45.3
NCEM SCB(ns)	1.38	0.96	3.38	2.76	6.78	6.40

It can be seen that under the two most possible transmission bandwidth conditions and the specific constant envelope conditions, the 40 MHz band-limited SCB is better than the non-constant envelope SCB as a whole. When the 50 MHz band-limited SCB

is used, each signal has its own advantages and disadvantages. Overall, SCBs with constant envelope at 50 MHz have little difference between non-constant envelope and constant envelope, except for SCBs with constant envelope at 50 MHz.

4 Conclusions

Although only one constant envelope method and one non-constant envelope method are listed in the analysis and simulation, the following conclusions can be drawn:

(1) Constant envelope design of signal is only useful for power amplifier, so ensuring constant envelope performance before power amplifier is the direction of signal design and load design. It is not necessary to study constant envelope performance of real space transmitting signal.
(2) The constant envelope design of signal is helpful to alleviate the non-linear influence of power amplifier, but the channel of engineering realization is not infinite bandwidth, so the constant envelope design of signal should not take infinite bandwidth as input condition.
(3) When the constant envelope cannot be guaranteed due to the principle of signal generation, the peak-to-average ratio (PAPR) should be used as one constraint condition for signal and band-limited design.
(4) Because of the influence of filtering, the performance of constant envelope signal is not necessarily better than that of non-constant envelope signal.

In addition, the results of space signal quality assessment of navigation satellites should be used as common input for signal design and load design. The final space signal quality of navigation should be guaranteed by both signal design and load design. Signal design should be carried out iteratively and interactively with load design. Any independent design that is not determined by engineering test is not optimal.

Acknowledgment. This work is supported by National Natural Science Foundation of China (Grant 11803023).

References

1. Yao Z, Lu M (2016) ACE-BOC: dual-frequency constant envelope multiplexing for satellite navigation. IEEE Trans Aerosp Electron Syst 52:466–485
2. Kazimierczuk MK (2015) RF power amplifiers, vol 2. Wiley, Hoboken
3. Giofre R, Colantonio P, Cipriani E, Danieli R, Giannini F, Gonzalez L, De Arriba F, Cabria L, Baglieri D (2015) A first step towards a 230 W SSPA for Galileo system exploiting European GaN technology. In: Proceedings of the 10th European microwave integrated circuites conference, Paris, France
4. Giofré R, Colantonio P, Giannini F (2015) A GaN high power and efficient amplifier for L-band Galileo system. In: 2015 integrated nonlinear microwave and millimetre-wave circuits workshop (INMMiC), Taormina, Italy

5. Saleh AAM (1981) Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers. IEEE Trans Commun 29(11):1715–1720
6. Yao Z, Minguan L (2016) Signal design principle and realization technology of new generation satellite navigation system. Electronic Industry Press, Beijing, p 4
7. (2016) China Satellite Navigation Office. Beidou Navigation Satellite System Signal in Space Interface Control Document Open Service Signal (Version 2.1), November 2016
8. (2017) China Satellite Navigation Office. Beidou Navigation Satellite System Signal In Space Interface Control Document Open Service Signal B1C (Version 1.0), December 2017
9. Yao Z, Lu M, Feng ZM (2010) Quadrature multiplexed BOC modulation for interoperable GNSS signals. Electron Lett 46(17):1234–1236
10. Dafesh PA, Cahn CR (2009) Phase-optimized constant-envelope transmission (POCET) modulation method for GNSS signals. Proce Int Tech Meetings Satell Div Inst Navig 5538 (1):2860–2866
11. Zhang K, Zhou HW, Wang FX (2011) Multiplexing performance assessment of POCET method for compass B1/B3 signals. J Navig 64(S1):S41–S54

Carrier Phase Measurement Technique Based on Non-cooperative GNSS Signals

Ying Yuan(✉), Feng Yu, Hua Zong, and Run Wang

College of Aeronautics, Nanjing University of Aeronautics and Astronautics, Nanjing, China
2254940674@qq.com

Abstract. The "Yinhe incident" and the chemical weapon crisis happened in Syria both showed that, the US army may have the power to shut down GPS civil signals in some specific areas. In this situation, traditional technologies of signal processing are no longer available, new methods must be developed to make use of GPS signal resources. A new system named Navigation via Signals of Opportunity (NAVSOP) is presented. This system is able to calculate users' position by making use of hundreds of random signals that are all around us. Then a specific method to realize NAVSOP is proposed, in which the non-cooperative GNSS signals are used as the input signal. The word "non-cooperative" means that the pseudo-code loaded on the carrier is unknown, so the carrier phase measurement technique based on non-cooperative GNSS signals will be the key problem to deal with. GNSS receivers can receive navigation signals with high SNR (signal-to-noise ratio) by using CPRA (controlled radiation pattern antenna), thus users can obtain superior signal source to carry out research. The main research contents include signal acquisition and tracking based on non-cooperative GNSS signals. In this paper the whole process of signal acquisition is expressed mathematically. Effect of noise with different intensities on the acquisition results are researched. And in theory it is proved that, when the SNR of input signals is above −15 dB, the acquisition algorithm proposed can successfully capture the frequency. A threshold is set to decide if the acquisition results are corresponded to Doppler frequency shifts. The optimal value of parameters are determined based on their impact on the performance of the acquisition algorithm. The result of signal acquisition is the estimation of carrier frequency. A phase locked loop (PLL) is utilized to lock the carrier and track its variation. The major error sources of PLL in GNSS receiver are phase jitter and dynamic stress error. Optimal parameters are set to improve the performance of tracking loop and reduce the error. Besides, the order of PLL are determined according to the dynamic environment of user's satellite, and the isolation and isolation frequency are proposed to reduce the influence of the close carrier frequency on the tracking results. The simulation results showed that the error of signal acquisition algorithm is within 3 Hz, and the number of satellites which is correctly acquired is more than 6. When PLL stabilized to track the signal, the tracking error (difference between the frequency of output signals and input signals) is within 15 Hz. The method of signal acquisition and tracking conducted in this paper realized carrier recovery of non-cooperative GNSS signals, it also laid the technical foundation of implement of NAVSOP.

Supported by National Natural Science Foundation of China (Grant No. 61673212).

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 432–441, 2019.
https://doi.org/10.1007/978-981-13-7751-8_43

Keywords: Non-cooperative GNSS signal · NAVSOP · Signal acquisition and tracking · PLL

1 Introduction

Although the GNSS (Global Navigation Satellite System) possess strong ability in navigation location service, these systems are also highly vulnerable to all types of attacks. Therefore, in the 1980s, the US military carried out a series of studies on the vulnerability of navigation system, and pioneered the modern NAVSOP (navigation warfare) theory. No matter the "Yinhe incident" which happened 25 years ago or the chemical weapon crisis in the Syria all showed that the US army may have the power to shut down the GPS civil signals in some specific areas. So traditional technologies of signal processing are no longer available.

In this situation, new methods must be developed to make use of GPS signal resources. Known as Navigation via Signals of Opportunity (NAVSOP), BAE Systems' new system is able to calculate its position by making use of the hundreds of random signals that are all around us. Then a specific method to realize NAVSOP is proposed, in which the non-cooperative GNSS signals are used as the input signal. The word "non-cooperative" means that the pseudo-code loaded on the carrier is unknown, so the carrier phase measurement technique based on non-cooperative GNSS signals will be the key problem to deal with.

It is very difficult to capture the signal when the SNR is very low. Now by employing the controlled radiation pattern antenna (CRPA), GNSS receivers can raise the capability of anti-jamming and increase the signal power by 30 dB, which can meet the requirements of signal acquisition of non-cooperative GNSS signals.

Traditional receivers generally capture the C/A codes first, and then P codes. But when the PRN (Pseudo-Random Noise) code of the navigation satellite is unknown, the traditional methods are no longer available.

In this paper a new method is conducted to acquire and track the non-cooperative GNSS signals, that is a special way to realize NAVSOP. The primary principle is: the square method are used to remove the PRN code of GNSS signals, the estimation of frequency of GNSS signals are obtained by serial search. The result of signal acquisition is the estimation of carrier frequency, and a phase locked loop (PLL) is utilized to lock the carrier and track its variation.

2 The Method of Signal Acquisition

In order to get the initial frequency of the tracking loop, at first, the square method are used to remove the PRN-code of input signals. Local signals with a series of doppler shifts are also generated. Then a square operation is performed for both the input signals and local signals. The correlation function is the product of the results.

The correlation function over a period of time is integrated. When the doppler frequency shift of the local carrier is consistent with the input signal's, the integral

results will generate a peak value. The doppler frequency shift of input signals is estimated corresponding to the peak value, and the estimation of carrier frequency is obtained after conversion.

The functional block diagram of the whole process is shown in Fig. 1.

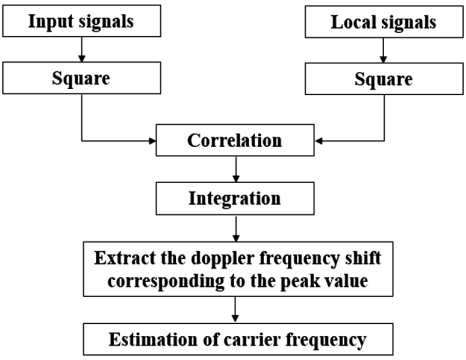


Fig. 1. Functional block diagram of signal acquisition

In general, the power of GPS signal of L1 band is about −160 dBW, and the SNR is −19 dB. Since the receiving bandwidth of signals with P-code is 10 times of the signals with C/A code, the SNR of P-code signal will be lower. In this situation, it is hard to carry out signal acquisition operation. By employing the controlled radiation pattern antenna (CRPA), the SNR of received P-code signal is about −5 dB, which can meet the requirements of signal acquisition algorithm presented in this paper.

2.1 Mathematical Model and Analysis

Navigation satellite signals are converted into intermediate frequency (IF) signals after down-conversion operation. The mathematical model of IF signals can be expressed as:

$$y_I(t) = AC(t)D(t)\cos[2\pi(f_{IF} + f_d)t] + n_1(t) \tag{1.1}$$

$$y_Q(t) = AC(t)D(t)\sin[2\pi(f_{IF} + f_d)t] + n_2(t) \tag{1.2}$$

where A represents the signal amplitude; C is PRN code; D is the data code; f_{IF} is the intermediate frequency; f_d is the doppler frequency shift; $n_1(t)$, $n_2(t)$ is the noise. The process of signal treatment are all the same, the following analyses are based on I brunch.

Local signals can be expressed as:

$$y_{IL}(t) = \cos[2\pi(f_{IF} + f_l)t] \tag{1.3}$$

where f_l is the doppler shift of the local signal.

The square of input signals is multiplied by the square of the local signal to get the correlation function:

$$y_{IA}(t) = y_I^2(t) \cdot y_{IL}^2(t) = S(t) + N(t) \tag{1.4}$$

The correlation function can be divided into two parts: the signal part $S(t)$ and the noise part $N(t)$. The acquisition result of I branch is the integration of the correlation function over a period of time. Integral time and the doppler shift values are the most important factors that can affect the acquisition result. The following analyses present the specific mathematics expression about the integral results of the signal part $S(t)$ and the noise part $N(t)$.

The specific mathematical expression of the signal part $S(t)$ is

$$\begin{aligned} S(t) = &\tfrac{1}{4} + \cos[4\pi(f_{IF} + f_d)t] + \tfrac{1}{4}\cos[4\pi(f_{IF} + f_l)t] \\ &+ \tfrac{1}{8}\cos[4\pi(2f_{IF} + f_d + f_l)t] + \tfrac{1}{8}\cos[4\pi(f_d - f_l)t] \end{aligned} \tag{1.5}$$

The integral result of Cosine function will change periodically over time. If the integral time is long enough, compare to the integral result of constant, the amplitude of the integral result of Cosine function is much smaller and can be ignored. Set N as the total integral points of the system, and the integral result of the signal part will be:

$$Y_{IS} = \sum_{t=1}^{N} S(t) = \frac{N}{4} + \frac{1}{8}\sum_{t=1}^{N} \cos[4\pi(f_d - f_l)t] \tag{1.6}$$

When the doppler frequency shift of the local signal is completely the same with the input signal's, $\cos[4\pi(f_d - f_l)t]$ will equal to 1. In this case, the integral result of signal part will be $3N/8$. But In other cases, the integral result of $\cos[4\pi(f_d - f_l)t]$ can be ignored, which means the integral result of signal part will reduce to $N/4$. The difference between the peak value of integral results and the others is $N/8$. The doppler frequency shift of the input signal are obtained by detecting the peak value among the integral results of signal part.

The specific mathematical expression of the noise part $N(t)$ is

$$\begin{aligned} N(t) = &\tfrac{1}{2}n^2(t) + C(t)\cos[2\pi(f_{IF} + f_d)t] \cdot n(t) \\ &+ \tfrac{1}{2}n^2(t)\cos[4\pi(f_{IF} + f_l)t] \\ &+ C(t)\cos[2\pi(f_{IF} + f_d)t]\cos[4\pi(f_{IF} + f_l)t] \cdot n(t) \end{aligned} \tag{1.7}$$

The integral result of the noise part will add a fluctuation for the integral result of the signal part. If the fluctuation is big enough, it will lead error or even mistake to the acquisition algorithm. Obviously, in order to implement the acquisition algorithm, the fluctuation value must be much smaller than $N/8$. Simulation experiments show that, the value of the fluctuation is directly related to the SNR of the input signal. The higher the SNR is, the smaller the fluctuation is. When the SNR is -15 dB, the magnitude of the fluctuation is about $N/100$. Therefore, when the SNR is -15 dB or higher, the acquisition algorithm proposed in this paper can correctly capture the signal and get the initial value of the signal frequency.

2.2 Parameters Selection

The result of acquisition algorithm is the sum of squares of the integral results of IQ branches.

The length of integral time directly affects the performance of the acquisition algorithm. If the integral time is too short, the peak value of the integral result is not obvious. If the integral time is too long, the calculation amount for each search will increase, eventually lead to a great extend to the search time. But within the allowable calculation amount, the longer the integral time is, the better the integral results are. In order to meet the precision requirement of signal acquisition, the searching interval of frequency is set as 1 Hz. Plenty of simulation experiments showed that when the integral time is 80 ms, the signal acquisition algorithm can successfully capture the signal within the allowable calculation amount.

3 Signal Tracking

When GNSS receivers are on work, the system keeps the tracking of satellite signals, and estimates the precise doppler frequency shift and carrier phase. But the doppler frequency shift of the GNSS signals is changing all the time because of the relative motion between the navigation satellites and receivers, so a PLL (phase locked loop) is used to lock and track the signals' phase.

3.1 Mathematical Model of PLL

PLL is a typical closed-loop feedback control system, which could use the reference signal imported from outside to control the frequency and phase of the signal generated within the loop. PLL is consisted of three components: PD (Phase Detector), LF (Loop Filter) and VCO (Voltage Controlled Oscillator). Figure 2 shows the model of PLL.

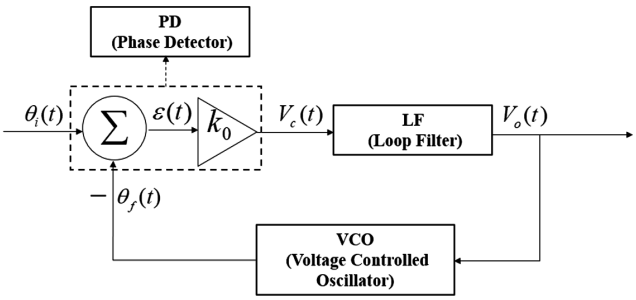


Fig. 2. Phase locked loop model

The order of loop filter determines the order of PLL. In general, GNSS receivers use second-order PLL, that is to say, the loop filter is first-order. The transfer function of second-order PLL is:

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{1.8}$$

where $\omega_n = \sqrt{k_0 k_1/\tau_1}$ is the natural frequency of the loop, and $\zeta = \tau_2\omega_n/2$ is the damping factor. By then we can get the NBW (Noise Bandwidth) of second-order PLL, that is:

$$B_n = \int_0^{+\infty} |H(\omega)| d\omega = \frac{\omega_n}{2}(\zeta + \frac{1}{4\zeta}) \tag{1.9}$$

The NBW of the loop determines the allowable noise magnitude of the system. With wide noise bandwidth, the loop can track the signal quickly, but the tracking error is also large. If noise bandwidth is narrow, the loop is difficult to track the signal quickly, but once the signal is tracked, the tracking error will be smaller.

3.2 Parameters Selection of Tracking Loop

The selection of loop parameters will directly affect the performance of carrier tracking loop. The main phase error sources of PLL in GNSS receiver are phase jitter and dynamic stress error, which can be expressed as the following equation:

$$3\sigma_{PLL} = 3\sigma_j + \theta_e \leq 45^\circ \tag{1.9}$$

Where σ_j is the phase jitter caused by all the other error sources except the dynamic stress error θ_e.

Phase jitter is the square root of the quadratic sum of each irrelevant phase error source, including thermal noise and oscillator noise. The influence of thermal noise is always accompanied by the receiver when it works, while other sources of oscillation are instantaneous or ignorable. Therefore, thermal noise is often regarded as the only error source of phase jitter. The phase jitter error of tracking loop caused by thermal noise can be expressed as:

$$\sigma_j = \sqrt{\frac{B_n}{C/N_0}(1 + \frac{1}{2TC/N_0})} \tag{1.10}$$

where C/N_0 represent the carrier to noise ratio, T is the integral time.

In order to reduce the influence of thermal noise and improve the tracking accuracy, it is necessary to extend the integral time. but in this way, the robustness of the tracking loop to dynamic environment is lower. In the premise that the tracking loop can complete the tracking to carrier, when the integral time is set as 5 ms, the tracking loop has the best performance.

Assume that a satellite with a radial velocity v_n broadcasts signals whose frequency is f and carrier wave length is λ. The radial velocity of the receiver is v, then the doppler frequency shift of the navigation satellite signals received by the user satellite can be expressed as:

$$f_d(t) = (v - v_n)/\lambda = at/\lambda \tag{1.11}$$

where a is the radial acceleration. The integral result of above equation is the variation of carrier phase $\theta_i(t) = \pi at^2/\lambda$, and the complex domain of it is $\theta_i(s) = 2\pi a/\lambda s^3$.

The phase error function of PLL under the frequency domain is:

$$E(s) = \theta_i(s)\frac{s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{1.12}$$

Then the dynamic stress error of second order PLL is:

$$\theta_e = \lim_{s\to 0} s \cdot E(s) = 2\pi a/\lambda\omega_n^2 \tag{1.13}$$

Associated with the phase jitter caused by thermal noise, we can get:

$$\sigma_{PLL} = \sqrt{\frac{B_n}{C/N_0}\left(1 + \frac{1}{2TC/N_0}\right)} + \frac{2\pi a}{3\lambda\omega_n^2} \tag{1.14}$$

The relationship between noise bandwidth and phase error under different carrier to noise ratio is shown in Fig. 3. We can see that phase error is always under 25°, combining with the tracking threshold mentioned above, the second order PLL can meet the requirements of signal tracking. The smaller the phase error is, the higher accuracy the tracking loop can get. So the optimal selection range of noise bandwidth is 25–40 Hz.

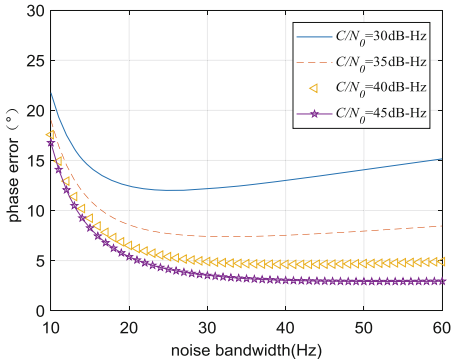


Fig. 3. Relationship between noise bandwidth and phase error

3.3 The Isolation of the Tracking Loop

PLL should have a sufficient suppression on the signals whose frequency is close to the target carrier (difference of the frequency is within 300 Hz). The amplitude - frequency characteristics of the system under different noise bandwidth are shown in Fig. 4.

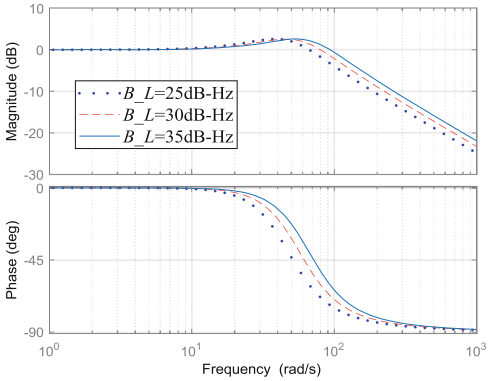


Fig. 4. Bode Diagram at different noise bandwidth

We can see from Fig. 4, the smaller the noise bandwidth is, the greater the inhibitory effect is. So the noise bandwidth is set as 25 Hz. And the amplitude characteristic corresponding to 300 Hz is −15 dB, therefore, the isolation of the system is selected as −15 dB, that is $20\lg|H(j\omega_b)| = -15\text{dB}$.

4 Simulation Verification and Analysis

In order to verify the signal acquisition and tracking algorithm based on non-cooperative GNSS signals, a simulation platform for navigation satellite signals is established.

The parameters of the signals are set as follows: the intermediate frequency of the signal is 0.2046 MHz, the sampling frequency is 1.023 MHz, and the SNR of the signal is −5 dB, integral time is 80 ms (the number of integral points is $N = 1.023 \times 80 \times 10^3 = 81840$). The acquisition results showed that 8 satellites were captured, one of them is shown in Fig. 5. The analysis in Sect. 2.1 shows that, when the doppler frequency shift loaded on the local signal is consistent with the simulated signal's, the theoretical result of the acquisition algorithm is $(N/4 \times 8 + 8/N)^2 \times 2 = 6.05 \times 10^{10}$. In other cases, the theoretical results are: $(N/4 \times 8)^2 \times 2 = 5.36 \times 10^{10}$.

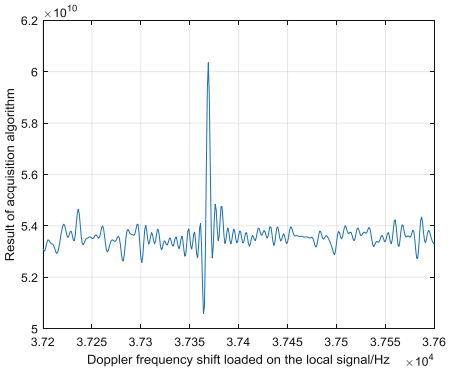


Fig. 5. Result of signal acquisition

Fluctuations in Fig. 5 are mainly caused by the noise. The reason why the peak value is slightly lower than the theoretical value is that, the doppler frequency shift of the signal changes slightly during the integral time of 80 ms, and the doppler frequency shift cannot completely correspond to the signal's. The simulation results verify the feasibility of acquisition algorithm.

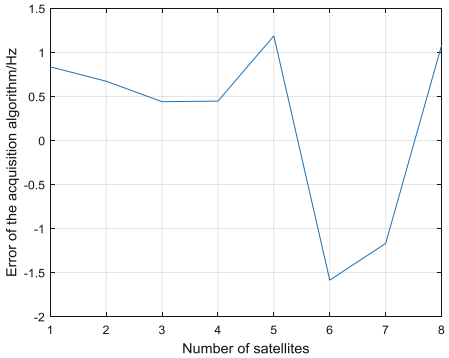


Fig. 6. Error of signal acquisition algorithm

The truth value of the signal doppler shift can be calculated by the relative position of the navigation satellite and the user satellite. By comparing the mean value of the captured value of doppler frequency shift and the theoretical value in the integral time, the error of acquisition algorithm is obtained, as shown in Fig. 6. The error of signal acquisition algorithm is within 2 Hz, and the number of satellites correctly acquired is more than 6.

The parameters of the tracking loop are set as follows: the damping factor of PLL is 0.637, and the Noise Bandwidth is 25 Hz.

Compare the observed frequency of output signals of tracking loop and the truth value. The error of the tracking loop is shown in Fig. 7.

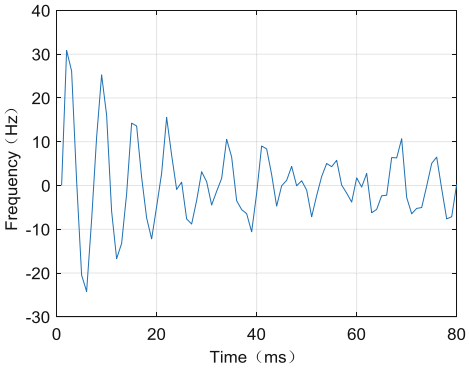


Fig. 7. Error of the tracking loop

As Fig. 7 shows, in the early stage of tracking, there is a relatively big difference between observed frequency and the truth value. When PLL stabilized to track the signal, the tracking error (difference between the frequency of output signals and input signals) is within 15 Hz.

References

1. Liu Y, Zhen S, Han Y (2002) Research of GPS countermeasures technique in navigation warfare. In: Electronic information warfare technology, vol 17, no 4
2. Wang F, Liu X, Zhou H (2014) Challenge and countermeasure for navigation warfare. In: Command information system and technology, vol 04, no 001
3. Lu H, Yu F, Liu Y, Zhu H (2016) Codeless carrier frequency initial value estimation and navigation satellite recognition. Syst Eng Electron
4. Woo KT (1999) Optimum semi-codeless carrier phase tracking of L2. In: The 12th international technical meeting of the satellite division of the institute of navigation, Nashville, Tennessee, 14–17 September 1999
5. Liu D (2003) Brief analysis of receive techniques based on L2 codeless or semi-codeless carrier. China Electron Technol Group Corporation
6. Xurong D (2008) GNSS single frequency software receiver development and applications. National Defense Industry Press, Beijing

INS-Assisted GNSS Loop Tracking Hardware Implementation Algorithm Design

Xiaoming Hao(✉), Chunying Li, and Jinshan Liu

Space Star Technology CO., LTD., Beijing 100086, China
794804174@qq.com

Abstract. In GNSS/INS integrated navigation, Due to the deep processing to the baseband level of GNSS, the hardware integration is difficult, so most of them are still in the stage of simulation and principle verification. In this paper, for the deep combination technology, at the baseband processing level, for the problem of different output frequency of GNSS system, INS system and loop, the data re-engineering method is designed for data alignment. For the data transmission delay problem of the GNSS system and the INS system, the time synchronization of the GNSS system, the INS system and the loop part is performed by means of hardware timing. For the INS solution and combined filtering takes a long time, affecting the real-time output of the INS data, the dual INS core is designed to ensure the real-time and accuracy of the INS result. The experimental results show that the INS-assisted GNSS loop can complete the loop tracking task, and the smaller loop bandwidth can be obtained when using the auxiliary information. The designed loop-assisted algorithm can realize the deep combination method and lay a good foundation for the next step in high dynamic application.

Keywords: INS-assisted GNSS loop · Hardware implementation · Time synchronization

1 Introduction

In GNSS/INS integrated navigation, the combination method has three combinations of loose combination, tight combination and deep combination. At present, the loose combination technology has matured, reducing the divergence problem of inertial navigation results; the tight combination is also almost mature, and the continuity and stability of the navigation results are improved by the assistance of the navigation data processing level. Data processing at the baseband level of GNSS, hardware integration is difficult, so most of them are still in the stage of simulation or principle verification.

Deep combination technology has important research value and practical significance: deep combination is the deepest level of integration, is the development trend of integrated navigation technology, and carries out deep combined technology research, which is conducive to timely grasping the future development direction of integrated

Our thanks to Beijing Municipal Science and Technology Commission for supporting us to develop the paper, Subject number: Z171100005017001.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 442–451, 2019.
https://doi.org/10.1007/978-981-13-7751-8_44

navigation technology and mastering cutting-edge core technologies. The existing deep combination related technologies are listed as key military technologies, subject to the strict export restrictions of the Western countries led by the United States, and the implementation of deep combined technology research is the only way to enhance the independent intellectual property rights of China's advanced receiver technology and integrated navigation technology. The deep integrated navigation system based on the micro-electromechanical system inertial measurement unit has the advantages of low cost, small size, light weight, low power consumption, etc., and has very important significance in military and commercial applications. The development of integrated real-time deep combined system is At present, China's Beidou satellite system has entered the regional service stage, and the development of deep combined technology research can not only provide technical reserves for China's integrated navigation system, but also effectively promote the industrialization of Beidou.

This paper mainly studies the deep combination method. When data processing and hardware integration are performed at the GNSS baseband processing level, the GNSS information and INS information are derived from different systems, so that the output frequencies of GNSS and INS are different, and the tracking loop runs in the baseband. The frequency varies greatly, data recursion and alignment are required; the delay of information transmission of different systems, and the delay on the hardware, so that there is a certain delay error in time synchronization through software; and because of the INS solution and combination in the program. The filtering operation takes a long time and affects the real-time output and correction of the INS data. By studying these issues, a deep combination approach is implemented on the hardware.

2 INS Auxiliary Loop Tracking

In the deep combination technology, the most important thing is to use the INS positioning speed measurement result to assist the satellite navigation tracking loop. In high dynamic and ultra-high dynamic environments, the satellite navigation loop is easy to lose lock, loss of satellite signals, and can improve receiver capture and tracking status when using INS assist [5, 6], INS can be in a short time. The accuracy of the positioning and speed measurement is ensured, and the tracking loop of the satellite navigation is assisted by the accurate positioning speed measurement result, which can ensure the locking state of the satellite navigation loop (Fig. 1).

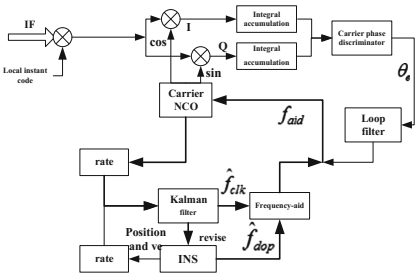


Fig. 1. INS-assisted GNSS loop tracking

The upper part of the figure is the tracking loop branch. Since the GNSS receiver antenna measures the dynamic change of the carrier by receiving the satellite signal, the direct measurement result is the motion information of the carrier and the satellite in their line of sight (LOS), so the motion. There is an LOS projection relationship between the dynamic information of the carrier and the dynamic information contained in the signal processed by the receiver radio frequency. The satellite signal containing the motion information of the carrier is processed by the receiver radio frequency, and then sent to the carrier tracking loop for carrier tracking and processing, which is the same as the common carrier tracking loop, including carrier stripping, code stripping, coherent accumulation, phase discrimination, frequency discrimination, filtering and Adjust the local carrier NCO.

The lower part of the figure is the feed-forward branch of the combined navigation result. The attitude and position information measured by the combined navigation result is not directly related to the loop assistance, and can be given autonomously. The position and velocity information of the carrier obtained by the combined navigation result is combined with the position and velocity information of the satellite to calculate the relative speed of the LOS direction. Finally, the speed is converted into Doppler and combined with the computer clock to obtain Doppler auxiliary information. For the problem of output frequency, the obtained auxiliary information needs interpolation extrapolation, and the auxiliary information is sent as a feed-forward link to the carrier tracking loop of the receiver.

If the estimation error is not considered, the Doppler auxiliary information including satellite motion information, carrier motion information, and receiver clock drift is equal to the input frequency of the carrier loop to remove the intermediate frequency effect. Therefore, the Doppler auxiliary information, like the output value of the tracking loop filter, is an estimate of the frequency variation of the input signal, so the Doppler auxiliary information can introduce a tracking loop at the output of the filter. Intuitively, the external Doppler auxiliary information estimates the frequency variation of the input signal, and the tracking loop only needs to track the residual error of the auxiliary information estimation. If external Doppler assistance can improve quasi-static operating conditions for the tracking loop, the tracking loop can greatly compress the bandwidth and thus reduce the effects of noise on the loop.

In the hardware implementation of loop assist technology, the focus is on the auxiliary Doppler and the influence of various information errors on the final auxiliary Doppler accuracy. Therefore, the paper firstly assists the Doppler in loop assist. Accuracy analysis; for the information transmission delay and hardware delay of different systems, the hardware timing method is used to calculate the time to ensure the accuracy of the data; the INS solution and the combined filtering operation in the program take a long time. The problem is that the design uses the dual INS core to separately output and correct the INS to ensure the real-time and accuracy of the data required for the loop operation.

2.1 Auxiliary Doppler Calculation Principle

The difficulty in the deep combination is to use the INS information to assist the receiver loop. In this case, the auxiliary Doppler shift needs to be calculated according

to the position and velocity information of the satellite and the position and velocity information of the receiver, and the local NCO is adjusted according to the auxiliary Doppler shift. Perform local carrier generation, reduce tracking error caused by receiver motion, reduce loop bandwidth, reduce noise intensity entering the loop, enhance loop anti-interference ability, improve observation accuracy, and realize INS-assisted GNSS receiver quickly and accurately Track satellite signals.

For the problem that the output frequency of GNSS and INS is different, and the frequency of tracking loop operation in the baseband is different, the model recursive method is used to extrapolate the data interpolation of GNSS data and INS data, and the operation of each loop operation is required. The data is then obtained according to the vector of the satellite position and the position of the receiver, and then the projection of the velocity vector in the line of sight direction is obtained according to the vector of the satellite speed and the receiver speed, and then the receiver clock drift information calculated according to the time is calculated. And the signal frequency and speed of light of the satellite signal that needs assistance, converting the velocity information into Doppler information.

2.2 Time Synchronization

In the traditional combined filtering algorithm, the combined filtering period is 1 s, and it needs to wait for the GNSS output positioning result to be solved. The combined filter dimension reaches ten dimensions, and the solution in the hardware platform takes more than milliseconds. There is a large time delay in calculating the loop auxiliary information from the combined filtering result. In the GNSS loop processing process, the loop is updated and adjusted every millisecond, and the loop auxiliary information is solved by the combined filtering result once per second. Road demand. In order to meet the real-time requirements of the loop for auxiliary information, the time interval is calculated by the crystal oscillator of the hardware platform, and the double INS core solution is used for improvement. The loop auxiliary information is provided by the INS high rate solution, and the INS solution is solved by the combined filtering result. The correction is performed to provide not only accurate loop auxiliary information in a high dynamic scene, but also to correct the INS solution by using the combined filtering result to prevent divergence.

The design principle of the dual INS core is:

Set the master and slave INS solution functions. The main INS function updates and outputs the IMU data transmitted by the serial port in real time; at the same time, it starts to buffer the received IMU data in the whole second; it is in the waiting state from the beginning of the INS, and will be filtered after the program completes the combined filtering. The navigation information is used as the initial value of the whole second, and the IMU data buffered from the whole second time is extracted, and the solution is chased to the current solution time of the main INS function, and then the navigation information of the main INS function is updated. The improvement principle is shown in the figure below (Fig. 2).

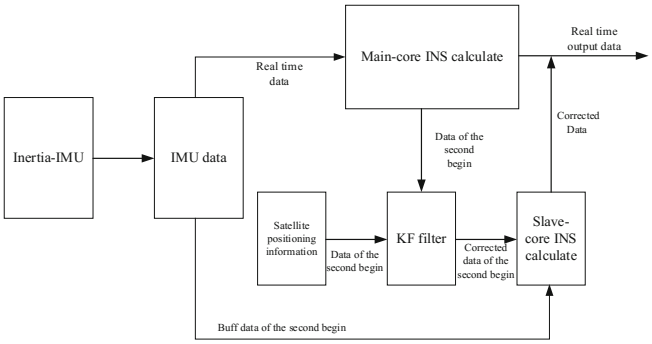


Fig. 2. Double INS core schematic

3 Error Analysis

There are multiple sources of error in the auxiliary Doppler calculation process. Therefore, the accuracy of the auxiliary Doppler depends mainly on the error of the information used in the calculation of the auxiliary Doppler.

Auxiliary Doppler calculation mainly uses satellite data (satellite position, speed, time, clock drift) and receiver data (receiver position, speed, time, clock drift). The error is mainly divided into original data error and recursive model error., delay error.

3.1 Raw Data Error

The original data error mainly includes satellite error (including position error, speed error, satellite clock drift error), combined receiver output error (including position error, speed error, receiver clock drift error).

Satellite error: The satellite data is calculated by the satellite ephemeris, and this part of the error is also included in the receiver resolution calculation, so it is not considered;

Receiver error: Satellite and receiver position information is mainly used to calculate the line-of-sight direction unit vector. Since the satellite is far away from the receiver, the position vector error of the line-of-sight direction caused by the position error is small and negligible; the receiver speed is combined by filtering. Output, there is a speed error, need to consider the auxiliary Doppler error caused by the receiver speed error; the receiver clock drift error includes the clock drift error of the crystal oscillator in the receiver and the thermal noise error caused by the hardware operation, which will be obtained during satellite positioning estimated value.

3.2 Model Extrapolation Error

Due to the hysteresis of satellite navigation and the delay of the combined filtering algorithm, the position and velocity information of the satellite and receiver at the current time cannot be obtained in real time. Therefore, the extrapolation model design needs to be based on the information known at the previous moment.

3.2.1 Extrapolation of Satellite Data

There are four consecutive whole second moments t0, t1, t2, t3, t1 is the current time, data recursion needs to use the data of t0, t1 two times, but because the receiver positioning solution takes time, when t1 time When the positioning data is obtained, this time t1 + τ has been between two times t1 and t2. At this time, the loop has been running for τ seconds with respect to t1, so the time from t1 + τ to the next time the positioning result is acquired. The auxiliary information during this time needs to be recursed by the positioning result at time t1, and so on. For the convenience of simulation, assuming τ = 1 s, the data from t2 to t3 is recursed during simulation.

The recursive method is to perform linear interpolation according to the satellite speed at time t0 and t1, calculate the satellite speed information every 1 ms from t2 to t3, and then obtain the average speed according to the time t1 and the calculated current time. The position of the current time is calculated based on the time intervals of the two moments, thereby calculating the satellite position information every 1 ms from the time t2 to the time t3.

3.2.2 Receiver Data Recursion

The receiver positioning result data recursive is similar to the satellite data recursive method. Firstly, linear interpolation is performed according to the acceleration of the INS output, and then the speed is obtained according to the average acceleration, and the position is obtained according to the average speed.

3.3 Delay Error

The delay error is mainly the error between the satellite information and the receiver information caused by the time synchronization of the Doppler calculation time.

4 Simulation Analysis

Since the high dynamic environment needs to be realized in environments such as cruise navigation and supersonic aircraft, the simulation data is first used for verification.

Data simulation was performed using SimGEN software. The simulation setting scene is linear acceleration motion. The motion setting is: static 30 s, linear acceleration motion 5 s, speed from 0 to 2000 m/s, uniform motion for 10 s, linear deceleration for 5 s, speed from 2000 m/s to 0, and stationary for 10 s.

4.1 Error Simulation Analysis

The Doppler error caused by each error source was tested separately to test its influence on the Doppler error (Figs. 3, 4, 5, 6 and 7).

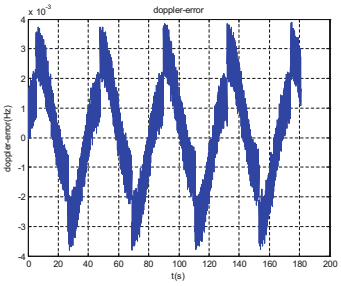


Fig. 3. Model recursive error

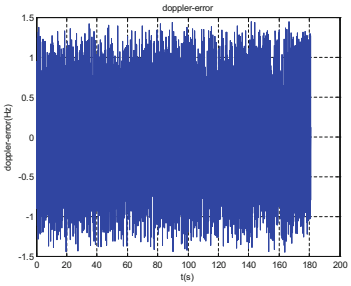


Fig. 4. Receiver speed error

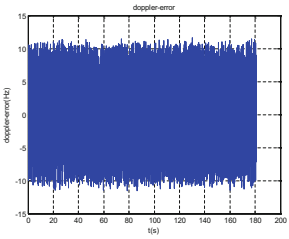


Fig. 5. Receiver drift error

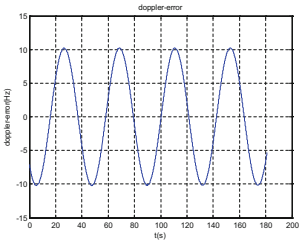


Fig. 6. Time delay error

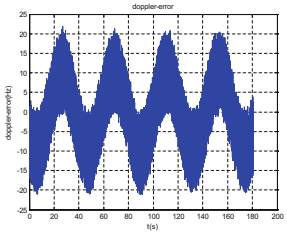


Fig. 7. Comprehensive error

It can be seen from the above results:

(1) Since linear interpolation is used in the extrapolation model, the linear motion of the uniform shift is simulated, so the model error will be caused when the receiver performs the variable acceleration motion, and the Doppler change rate caused by the acceleration is larger, the model. The larger the error, the greater the calculation error caused by the delay;

(2) The Doppler error caused by the receiver speed error is roughly in accordance with the Doppler change caused by the line-of-sight direction speed, that is, the 1 m/s speed error causes a Doppler change of 5 Hz;

(3) The error caused by the receiver drift of 2 m/s is about 10 Hz; the receiver drift error is set to the maximum error. After the general positioning is completed, the receiver clock drift error is equal to the receiver speed error, which is about 0.2 m/s.
(4) The error caused by the delay error is the largest, and this error is a fixed direction deviation, which is superimposed on other errors.
(5) The Doppler error needs to strictly control the time synchronization of all information in the calculation, and minimize the error and speed error of the receiver clock drift.

4.2 Time Synchronization Simulation

When the simulation is performed, the filter lag is set to 100 ms, and whether the data update is completed after the filter lag is compared has an effect on the filter result (Figs. 8, 9 and 10).

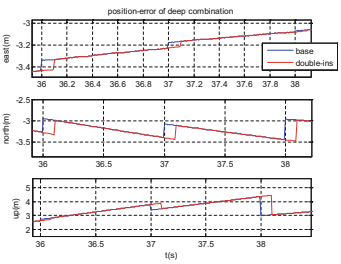


Fig. 8. Position error comparison

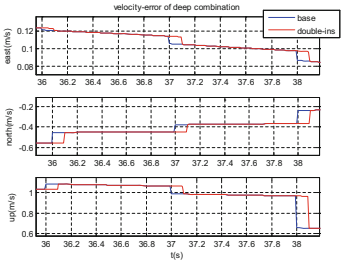


Fig. 9. Speed error comparison

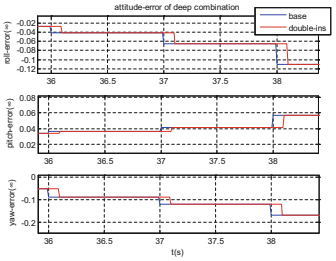


Fig. 10. Attitude error comparison

It can be seen from the above figure that after the double INS core is improved, due to the delay of the positioning solution and the combined filtering (100 ms delay set by the simulation), the position information is corrected after the combined filtering is completed, at the time of the whole second. Until the combined filtering is completed,

the inertia error accumulation will continue to be generated. The corrected error is the same as before the improvement, indicating that the algorithm is correct, and the information update is correctly completed after the combined filtering.

4.3 High Dynamic Simulation

The high-dynamic simulation is performed by using the integrated navigation signal source, and the satellite signal and IMU data are received by the hardware receiver for solution and positioning, and the designed INS auxiliary loop algorithm is tested. The results are as follows (Figs. 11 and 12):

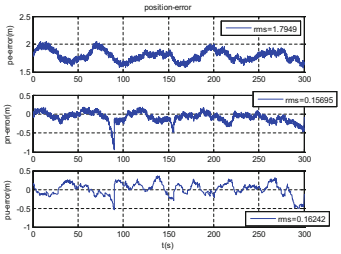


Fig. 11. Position error

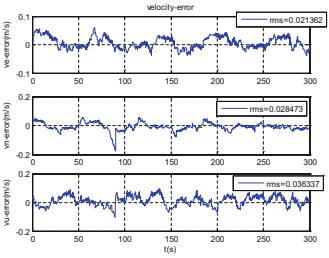


Fig. 12. Speed error

It can be seen from the above figure that under high dynamics, the three-axis positional accuracy is less than 2 m, and the three-axis velocity accuracy is less than 0.1 m/s, and the precision error is small.

5 Summary

In the deep integrated navigation technology, the calculation method and accuracy analysis of Doppler in loop assist are studied for the INS-assisted GNSS loop tracking. The influence of each data precision on the auxiliary results in loop assist; the loop assist in high dynamic environment The accuracy and real-time performance of the required INS data were improved by the algorithm of the double INS core. Finally, the auxiliary loop algorithm was designed. The results of high dynamic simulation show that:

In the loop assist, the extrapolation model simulates the uniform linear motion, which causes the model error to be small; the velocity error in the line of sight direction and the Doppler error caused by the receiver error depend on the magnitude of the GNSS positioning and INS solution error. The error caused by the delay error is the largest, and this error is a fixed direction deviation, which is superimposed on other errors.

After the double INS core is improved, due to the delay of the positioning solution and the combined filtering (100 ms delay artificially set during simulation), the

navigation information is corrected after the combined filtering is completed, and will continue until the combined filtering is completed until the combined filtering is completed. The inertial conduction error is accumulated, and the corrected error is the same as before the improvement. The algorithm is correct. After the combined filtering, the information update is completed correctly, and the double INS core is successfully improved.

The loop assist algorithm designed by the high dynamic simulation test shows that the positioning speed measurement accuracy of the algorithm can be similar to that of the general low dynamic, indicating that the designed loop assist algorithm can make the tracking loop work normally.

References

1. Liu K, Liu H (2005) Application and research trend of GPS/INS combined guidance technology in modern warfare. China Aerosp (9)
2. Abbott AS, Lillo WE (2003) Global positioning system and inertial measuring unit ultra-tight coupling method. United States 6516021, 4 Feb 2003
3. Tan S (2008) Development and thinking of Beidou satellite navigation system. J Astronaut 29 (2):391–396
4. Jing Z, Chong C (2008) Research on the development of GLASS system and its technology. Glob Positioning Syst 33(5):27–31
5. Garcia G (2010) A rapid acquisition GPS receiver based on ultra tightly coupled IMU and GPS. Chalmers University of Technology, Goteborg, pp 22–40
6. Xie G (2009) GPS principle and receiver design. Publishing House of Electronics Industry, Beijing, pp 2–6

Algorithm Research on High-Precision Tracking of Beidou-3 B1C Signal

Yao Guo[1,2], Xue Wang[1,2(✉)], Xiaochun Lu[1,3], Yongnan Rao[1], and Chengyan He[1]

[1] National Time Service Center, Chinese Academy of Sciences, Xi'an 710600, China
wangxeu@ntsc.ac.cn
[2] School of Electronic, Chinese Academy of Sciences University, Beijing 1014082, China
[3] School of Astronomy and Space Sciences, Chinese Academy of Sciences University, Beijing 1014082, China

Abstract. The B1C signal of Beidou-3 satellite navigation system adopts data and pilot separation to realize the transmission of message information and ranging functions. The traditional tracking method uses separate data to demodulate the message information, which doesn't fully utilize the signal energy and data-pilot phase relationship, resulting in low tracking accuracy. In order to demodulate the navigation message information more accurately, based on the phase relationship between the data channel and the pilot channel, this paper introduces the tracking strategy of "data + pilot" joints in loop tracking, and proposes three joint methods with different complexity, including "correlator output amplitude superposition", "discriminator output linear union" and "loop filter output linear union". The tracking accuracy of pseudo-code tracking loop and carrier tracking loop under different joint modes is theoretically analyzed, and the applicability of joint tracking in different scenarios is given. Finally, the measured data based on different satellites shows that the higher the inter-code consistency between the data and the pilot, the more obvious the improvement of joint tracking is. The 'correlator output amplitude superposition' method performs the high tracking accuracy and the simplest structure under higher SNR conditions. Under lower SNR conditions, the combination of "code loop filter output linear union + carrier loop discriminator output linear union" shows the highest tracking accuracy and the lowest tracking threshold. Although the loop filter combination increases the loop complexity, it adapts to a wider scene.

Keywords: B1C signal · Joint tracking · Tracking accuracy · Tracking threshold

1 Introduction

In order to meet the higher requirements of users for global satellite navigation system (GNSS) in the performance of positioning, navigation and timing, new navigation signal modulation methods and multiplexing methods are widely used. Compared with

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 452–464, 2019.
https://doi.org/10.1007/978-981-13-7751-8_45

traditional navigation signals, the new satellite navigation signal isn't a single signal anymore, but a design scheme of "data channel + pilot channel", that is, navigation message is broadcast only on the data component, and the pilot channel is modulated the known secondary code information stream. For example, the GPS L5 signal is composed of two components of equal amplitude and orthogonal carrier phase. Similarly, the MBOC signal, as a civilian interoperable signal for GPS, Galileo, and BDS, also uses a composite signal with pilot and data separation. This scheme improves the overall tracking performance of the signal and provides a more flexible algorithm for combinatorial tracking.

On November 5, 2017, the successful launch of the first group of Beidou-3 satellite navigation system marked the full launch of Beidou's global construction. It is expected to build a satellite navigation system from the original regional navigation capability to global coverage by 2020. "BeiDou Navigation Satellite System Signal In Space Interface Control Document Open Service Signal B1C (Version 1.0)" pointed out that B1C is a new civil signal, its data channel modulates BOC (1, 1) signal, pilot channel uses low frequency component BOC (1, 1) and high frequency Component BOC (6,1) orthogonal multiplexed QMBOC (6, 1, 4/33) signals. In the power distribution ratio of pilot and data, the B1C signal adopts a 25:75 allocation scheme. Traditional methods only use single-channel signal tracking for navigation demodulation, which causes to waste of useful power. A single data channel demodulation message will cause a 6 dB signal power loss, consequently, the tracking sensitivity and tracking accuracy will be greatly deteriorated. In response to this problem, some scholars have carried out related research, but haven't achieved the ideal joint tracking performance. Reference [1] proposes a joint tracking method for GPS L5 signal at the code tracking loop discriminator level and the filter layer, and theoretically analyzes the applicability of the filter joint tracking under the condition of carrier loop-assisted pseudo-code tracking. However, joint tracking at all levels is not implemented in the carrier loop. Reference [2] achieves a linear union at the discriminator level for the Galileo E1 signal, but does not validate the method in the real satellite data. Reference [3] introduces the GPS L5 joint tracking method, but doesn't compare the accuracy improvement performance of different joint tracking methods for navigation message demodulation and applicability in different SNR signals.

Based on the design scheme of B1C signal, this paper proposes the different joint tracking strategies for improving the demodulation accuracy of navigation messages. First, the tracking algorithm under the three different joint modes of "data + pilot" is introduced. Then, theoretical analysis of its tracking threshold and tracking accuracy in pseudo-code tracking and carrier tracking, and the applicable scenarios of each method are given. Finally, based on satellite measured data, the influence of inter-code consistency on joint tracking is verified, and the optimal joint tracking combination method under different CNR signals is given.

2 B1C Signal Structure and Tracking Model

2.1 B1C Signal Structure

The Beidou new civilian B1C signal consists of the data component S_{B1Cd} and the pilot component S_{B1Cp}. The power ratio is 1:3. Both have a code length of 10230 and a code cycle of 10 ms. The signal structure is shown in Table 1 [4].

Table 1. B1C signal structure

Signal	Signal component		Phase	Power ratio
B1C	S_{B1Cd}		0	1/4
	S_{B1Cp}	S_{B1Cpa}	90	29/44
		S_{B1Cpb}	0	1/11

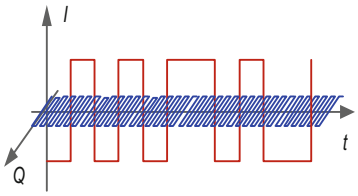


Fig. 1. QMBOC (6, 1, 4/33) time domain waveform

The S_{B1Cd} signal component is combined with the navigation data stream D_{B1Cd} and the spreading code C_{B1Cd} by subcarrier SC_{B1Cd} component is obtained by spreading the modulation. The S_{B1Cp} signal component is modulated by QMBOC (6, 1, 4/33), which is further composed of a narrowband component S_{B1Cpa} and a wideband component S_{B1Cpb}, wherein the S_{B1Cpa} component is obtained by spreading the code C_{B1Cp} modulated subcarrier SC_{B1Cpa}, the S_{B1Cpb} component is obtained by spreading the code C_{B1Cp} modulated subcarrier SC_{B1Cpb}. The B1C signal baseband expression is as follows [5]:

$$\begin{aligned} S_{B1C}(t) = \tfrac{1}{2}S_{B1Cd} + j\tfrac{\sqrt{3}}{2}S_{B1Cp} = \tfrac{1}{2}D_{B1Cd}\cdot C_{B1Cd}\cdot sign(\sin(2\pi f_{SCB1Cd}t)) \\ + \sqrt{\tfrac{1}{11}}C_{B1Cp}\cdot sign(\sin(2\pi f_{SCB1Cpb}t)) + j\sqrt{\tfrac{29}{44}}C_{B1Cp}\cdot sign(\sin(2\pi f_{SCB1Cpa}t)) \end{aligned} \tag{1}$$

As can be seen from Eq. (1), the pilot component uses an orthogonally multiplexed binary frequency-shifted carrier modulated signal QMBOC (6, 1, 4/33) composed of BOC (1, 1) and BOC (6, 1). QMBOC (6, 1, 4/33) is a new time domain implementation of MBOC interoperability signals. Its baseband signal expression is as follows:

$$S_{QMBOC}(t) = \sqrt{\frac{29}{33}}S_{BOC(1,1)}(t) - j\sqrt{\frac{4}{33}}S_{BOC(6,1)}(t) \tag{2}$$

Different from the time division multiplexing of the GPS TMBOC signal and the amplitude superposition of the Galileo CBOC signals, the BOC (1, 1) and BOC (6, 1) components are respectively modulated on two orthogonal phases of the carrier. QMBOC (6, 1, 4/33) time domain waveform is shown in Fig. 1.

The signal time domain waveform correlation operation can be obtained as follows:

$$R_{QMBOC}(\tau) = \mathrm{E}\left\{S_{QMBOC}(t)S^{*}_{QMBOC}(t)\right\} = \sqrt{\frac{29}{33}}R_{BOC(1,1)}(\tau) + \sqrt{\frac{4}{33}}R_{BOC(6,1)}(\tau) \quad (3)$$

The BOC (1, 1) component and the BOC (6, 1) component are located on two orthogonal phases, and the autocorrelation function of the QMBOC signal doesn't have a cross-correlation term. So the implementation of QMBOC is more flexible [5].

2.2 B1C Signal Loop Tracking Model

The data/pilot component tracking loop structure is shown in Fig. 2. The Costas is used in the data component carrier tracking loop. After the pilot component is stripped of the secondary code, the loop design may not consider the 'carrier phase half-cycle ambiguity' problem, which is caused by the navigation message bit flip. Therefore, it is possible to abandon the traditional Costas phase-locked loop that is insensitive to the navigation message flipping, and instead use a pure phase-locked loop with better performance in pilot tracking loop. A delay locked loop (DLL) is used in the pseudo code tracking loop. After the digital intermediate frequency signal is stripped by the carrier, it is coherently integrated with the BOC (1, 1)/QMBOC (6, 4/33) local replica code. The promote branch P is used for carrier tracking. The early branch E and the latter branch L are used for Code tracking.

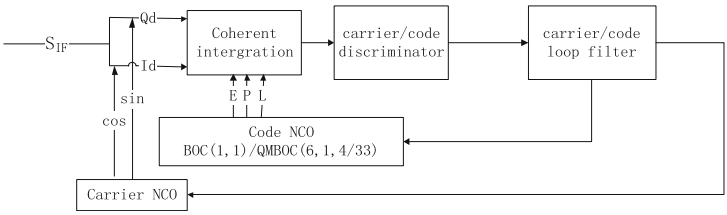


Fig. 2. Block diagram of a single data/pilot tracking loop

Taking the data channel as an example, the receiver coherent integration output: the early branch, promote branch and latter branch can be expressed as:

$$E_d = IE_d + jQE_d = I_0R(\delta_{\tau,d} + \frac{d}{2}) + n_{IE,d} + jQ_0R(\delta_{\tau,d} + \frac{d}{2}) + n_{QE,d} \quad (4)$$

$$P_d = IP_d + jQP_d = I_0R(\delta_{\tau,d}) + n_{IP,d} + Q_0jR(\delta_{\tau,d}) + n_{QP,d} \quad (5)$$

$$L_d = IL_d + jL_d = I_0R(\delta_{\tau,d} - \frac{d}{2}) + n_{IL,d} + jQ_0R(\delta_{\tau,d} - \frac{d}{2}) + n_{QL,d} \quad (6)$$

Where I_0 and Q_0 is:

$$I_0 = A_d D \sin c(\pi \delta_f T) \cos(\delta_{\varphi,d}) \tag{7}$$

$$Q_0 = A_d D \sin c(\pi \delta_f T) \sin(\delta_{\varphi,d}) \tag{8}$$

Where A_d is the signal data component amplitude, D is the navigation message symbol, δ_f is the frequency error (Hz) between the local carrier and the received signal carrier, and T is the coherent integration time. $\delta_{\tau,d}$ is the code phase error between the local pseudo code and the actual received signal. $\delta_{\varphi,d}$ is the carrier tracking phase error.

The output of pseudo-code discriminator and the carrier discriminator are respectively:

$$\hat{\Delta\tau_d} = \frac{\sqrt{IE_d^2 + QL_d^2} - \sqrt{IL_d^2 + QL_d^2}}{\sqrt{IE_d^2 + QL_d^2} + \sqrt{IL_d^2 + QL_d^2}} \tag{9}$$

$$\hat{\Delta\varphi_p} = \text{atan}\left(\frac{imag(P_d)}{real(P_d)}\right) \tag{10}$$

This data pilot uses the 'normalized non-coherent early minus late' pseudo-code phase discriminator and the 'two-quadrant arctangent' carrier phase discriminator.

Loop filter output:

$$\hat{\Delta f_{code\_d}} = k_{1code\_d}\,\hat{\tau_d}\,T + k_{2code\_d}\,\hat{\tau_d} \tag{11}$$

$$\hat{\Delta f_{\text{carr}\_d}} = k_{1\text{carr}\_\text{d}}\,\hat{\varphi_d}\,T + k_{2carr\_d}\,\hat{\varphi_d} \tag{12}$$

Where k_{1code_d}, k_{2code_d}, $k_{1\text{carr}\_\text{d}}$, and $k_{1\text{carr}\_\text{d}}$ are the coefficients of the pseudo code loop filter and the carrier loop filter, respectively.

The special design of the B1C signal data and the pilot channel makes the pseudo-code of B1Cd and B1Cp completely synchronized when the satellite transmits the signal. The B1Cp and B1Cd are modulated at the same frequency point, wherein the B1Cpa phase leads the B1Cd signal by 90°, and the B1Cpb signal is in phase with the B1Cd signal. Based on the specific relationship between data and pilot, the receiver can select three methods to demodulate the navigation message: (1) use B1Cd single component to achieve signal tracking; (2) use B1Cp to achieve signal tracking, and use tracking result of pilot components assists B1Cd in implementing navigation message demodulation. (3) B1Cd and B1Cp are jointly tracked, and the original observations are extracted to complete the coherent integration and message extraction of the data channel [6].

The third method is undoubtedly the optimal tracking method, which makes full use of the energy of the data and the pilot, and the total signal power is increased to obtain higher tracking accuracy. The "data + pilot" joint tracking method is applicable to both pseudo-code tracking and carrier tracking.

3 Two-Component Joint Tracking Algorithm and Performance Analysis

The tracking loop can combine B1Cd and B1Cp signals from three levels: the correlator level, the discriminator level, the and loop filter level. When dual-component joint tracking is used, the phase of the spreading code of the two components for the received signal can be considered synchronous, and the B1Cd carrier phase is delayed by 90 degrees from the B1Cpa signal [7], which is:

$$\delta_{\tau,d}=\delta_{\tau,p} \quad \delta_{\varphi,d}=\delta_{\varphi,p}-\frac{\pi}{2} \tag{13}$$

Therefore, when the carrier loop is jointly tracked, the pilot channel coherent integration result needs to be rotated by $-\pi/2$ radians to achieve carrier tracking with the data components:

$$P_p=(IP_p+jQP_p)\mathrm{e}^{(-j\pi/2)} \tag{14}$$

3.1 Correlator Output Amplitude Superposition

Figure 3 shows the tracking loop structure of the correlator output amplitude superposition. It can be seen from Eqs. (4), (5), and (6) that, besides the respective noise portions, the coherent integration amplitude of the B1Cd signal and the B1CPa signal is determined by the power of each component, and the symbol is determined by the message modulated by each component and the secondary code. Therefore, when combining data and pilot coherent integration results, it is necessary to consider the error association caused by the bit symbols cancels each other out. Equations (16), (17), and (18) are the results of the coherent integration after the union.

$$\lambda=\begin{cases} 1 & |P_d+P_p|\geq|P_d-P_p| \\ -1 & |P_d+P_p|<|P_d-P_p| \end{cases} \tag{15}$$

$$P_{cmb}=P_\mathrm{d}+\lambda P_p \tag{16}$$

$$E_{cmb}=E_\mathrm{d}+\lambda E_p \tag{17}$$

$$L_{cmb}=L_\mathrm{d}+\lambda L_p \tag{18}$$

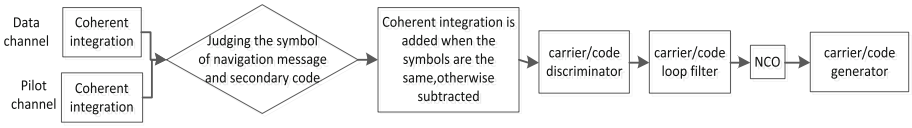


Fig. 3. Structure of Correlator output amplitude superposition

The correlator output amplitude superposition method makes full use of the energy of the entire signal, which can greatly improve the tracking accuracy. The structure of correlator joint is the simplest, and the loop requires only one discriminator and filter. However, in the case where the carrier noise is relatively low, the step of determining the bit symbol is affected, and the noise effect of the data channel is directly added and amplified, so that the structure of the correlator joint in the weak signal tracking is inferior.

3.2 Discriminator Output Linear Union

Figure 4 shows the structure of the discriminator output linear joint of discriminators. The combined weight coefficients "a" and "b" of the data component and the pilot component discriminator output are determined by the signal transmission power allocation ratio. As can be seen from Sect. 2.1, the power ratio of the B1Cd signal to the B1Cpa signal is 11:29. Therefore, the jointed phase estimation error of the pseudo-code tracking loop and carrier tracking loop is:

$$\Delta\hat{\tau}_{cmb} = \frac{11}{40}\Delta\hat{\tau}_d + \frac{29}{40}\Delta\hat{\tau}_p \tag{19}$$

$$\Delta\hat{\varphi}_{cmb} = \frac{11}{40}\Delta\hat{\varphi}_d + \frac{29}{40}\Delta\hat{\varphi}_p \tag{20}$$

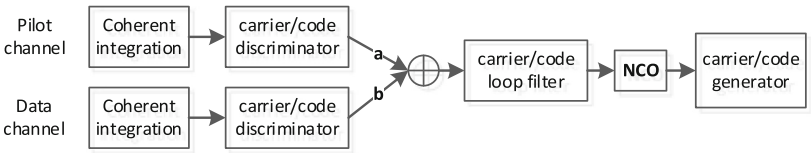


Fig. 4. Structure of Discriminator output linear joint

The pilot channel doesn't broadcast the navigation message, so the pilot channel uses the four-quadrant arctan phase detector, 'ATan2' (pure phase-locked loop). The data component uses a two-quadrant arctan phase detector, 'ATan' (Costas) that is insensitive to the change of the message. The discriminator curves of the two discrimination algorithms are shown in Fig. 5. It can be seen that the linear range of 'ATan' is $(-\pi/2, \pi/2)$, and the linear range of 'ATan2' is $(-\pi, \pi)$. As the carrier-to-noise ratio decreases, the discrimination curve linear range of the 'ATan' is not as wide as 'ATan2'.

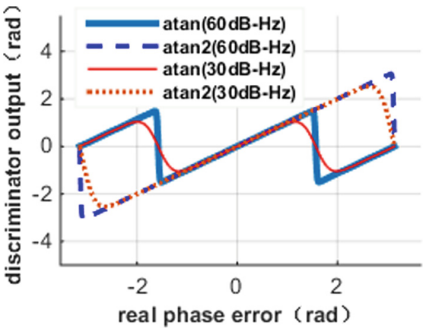


Fig. 5. 'ATan' and 'ATan2' curve characteristics

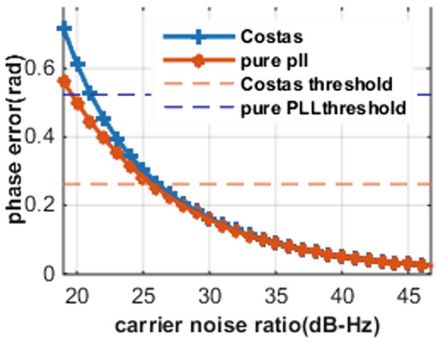


Fig. 6. Costas and pure PLL thermal noise error

The empirical threshold of tracking error is that three times the mean square error of the phase measurement error does not exceed one quarter of the traction range of the discriminator, so the two discriminator tracking thresholds can be expressed as follows [8]:

$$3\sigma_{a\ \tan} + \theta_e \leq \pi/4 \tag{21}$$

$$3\sigma_{a\ \tan 2} + \theta_e \leq \pi/2 \tag{22}$$

The main error sources of the carrier tracking loop are phase jitter error σ and dynamic stress error θ_e. The data used in this paper are collected in the static environment, so the dynamic stress error is very small and will not be considered here. The error caused by phase jitter is mainly the thermal noise error. For the 'ATan' discriminant algorithm and the 'ATan2' discriminant algorithm, the thermal noise error calculation formulas are:

$$\sigma_{a\ \tan} = \sqrt{\frac{B_n}{C/N_0}\left[1 + \frac{1}{2TC/N_0}\right]}[rad] \tag{23}$$

$$\sigma_{a\ \tan 2} = \sqrt{\frac{B_n}{C/N_0}}[rad] \tag{24}$$

Where B_n is the carrier noise loop thermal noise bandwidth and T is the coherent integration time. Figure 6 shows the variation of the standard deviation of the carrier loop thermal noise error with the carrier-to-noise ratio. In this paper, the carrier loop bandwidth and the coherent integration time are set to 25 Hz and 0.01 s. It can be seen that when the carrier noise is low, the tracking accuracy of the pure PLL is higher than that of the Costas PLL; the theoretical tracking threshold of the Costas PLL is 26 dB-Hz, and the pure PLL is 20 dB-Hz; the pure PLL can improve the tracking sensitivity by 6 dB.

Based on the above analysis, it can be judged that the discriminator output linear joint has a better tracking accuracy and a lower tracking threshold under the lower SNR than the correlator output amplitude superposition. The discriminator combination is more complex than the correlation output amplitude of the correlator, but the tracking accuracy can be improved at a lower carrier-to-noise ratio.

Therefore, it can be judged that the discriminator joint has better tracking accuracy under the weak SNR than the correlator joint and has a lower tracking threshold. The structure of the discriminator joint is more complex than the correlator joint, but the tracking accuracy can be improved at a lower carrier-to-noise ratio.

3.3 Loop Filter Output Linear Union

Figure 7 shows the structure of the linear combination of filter outputs. Under the condition that the structure and coefficient of the data filter are exactly the same as the pilot, the result of the union of the discriminator and filter is exactly the same. The selection principle of the weight coefficient is the same as the discriminator joint. The loop filter joint result of the pseudo code tracking loop and the carrier tracking loop are respectively (25) and (26).

$$\Delta \hat{f}_{code\_cmb} = \frac{11}{40}\Delta \hat{f}_{code\_d} + \frac{29}{40}\Delta \hat{f}_{code\_p} \tag{25}$$

$$\Delta \hat{f}_{carr\_cmb} = \frac{11}{40}\Delta \hat{f}_{carr\_d} + \frac{29}{40}\Delta \hat{f}_{carr\_p} \tag{26}$$

The advantage of the loop filter combination is that the loop filter bandwidth of the two components can be set separately. Under weak signals, increasing the coherent integration time can improve loop tracking accuracy. An important role in setting the pilot component is to provide a prerequisite for pseudo code tracking to carry out longer coherent integration (more than one pseudo-code period). In the joint mode of the pseudo code filter, the data channel is subjected to coherent integration (not exceeding the length of the telegram bit) and then non-coherently integrated so that the total integration time is the same as the pilot component. However, the increase of the coherent integration time causes the loop dynamic stress error to increase. Therefore, in order to balance the correlation loss caused by the non-coherent integration of the data channel and the dynamic stress error caused by the excessive coherent integration time

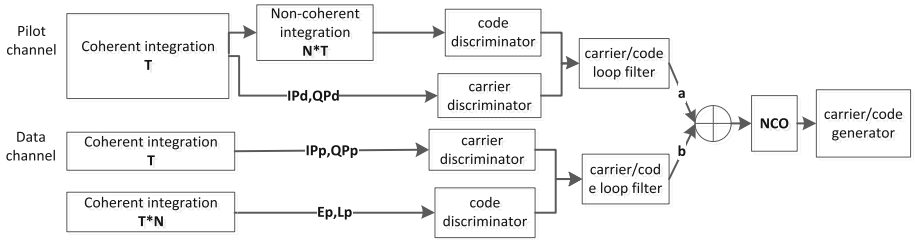


Fig. 7. Structure of Loop filter output linear joint

of the pilot channel, different filter coefficients in the data and pilot components can be adopted in the code loop filter joint. For data components, a narrower filter bandwidth can be set to compensate for the squared loss due to non-coherent integration; for the pilot component, to absorb most of the dynamic stress error, the filter bandwidth can be set wider.

The loop filter combination requires two filters with the highest complexity, but the wide application range is suitable for signal tracking with high dynamics and low signal to noise ratio.

4 Satellite Measured Data Analysis Results

The data used in this paper is collected by high-speed acquisition device, the sampling rate is 250 MHz, the center frequency is 62.5 MHz; The satellite used in this paper is the Beidou-3 satellite navigation system MEO-1, PRN = 19 (the inter-code consistency error of data and pilot is estimated to be −2.67e−03 ns through large-diameter antenna) and MEO-7, PRN = 29 (The inter-code consistency error of data and pilot is evaluated as −1.30e−06 ns by large-diameter antenna); this paper uses the same filter structure and parameters in each tracking strategy in order to compare the tracking accuracy of the three joint methods. The carrier loop bandwidth is 20 Hz, the code tracking loop bandwidth is 4 Hz, the carrier tracking coherent integration time is 10 ms, the pilot channel pseudo-code tracking coherent integration time is 20 ms, the data channel pseudo-code tracking coherent integration time and non-coherent integration time is 10 ms and 20 ms respectively.

Figure 8 shows the DLL phase-detection standard deviation of the B1Cd component of MEO-1 and MEO-7 satellite under different strategies. The carrier tracking uses single data component tracking. The tracking results are as follows: Under different tracking strategies, the phase-detection standard deviation of the DLL loop becomes larger as the carrier-to-noise ratio decreases. Under the condition of the high carrier-to-noise ratio, the three joint methods in the pseudo-code tracking loop can effectively reduce the tracking error, and the improvement effect is equivalent. When the carrier-

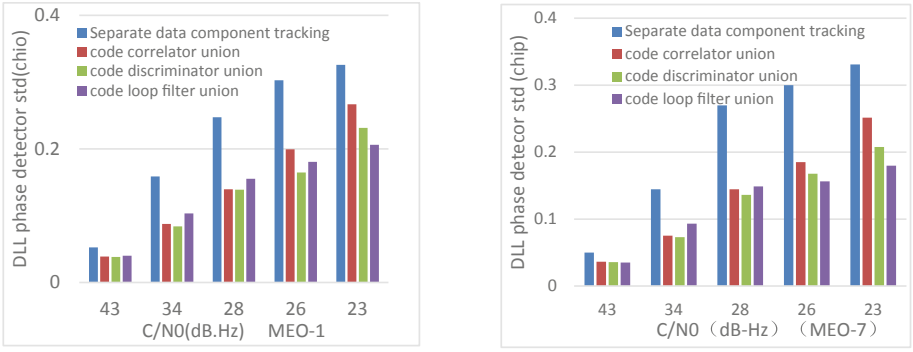


Fig. 8. B1Cd phase-detection standard deviation of pseudo-code tracking under each tracking strategy

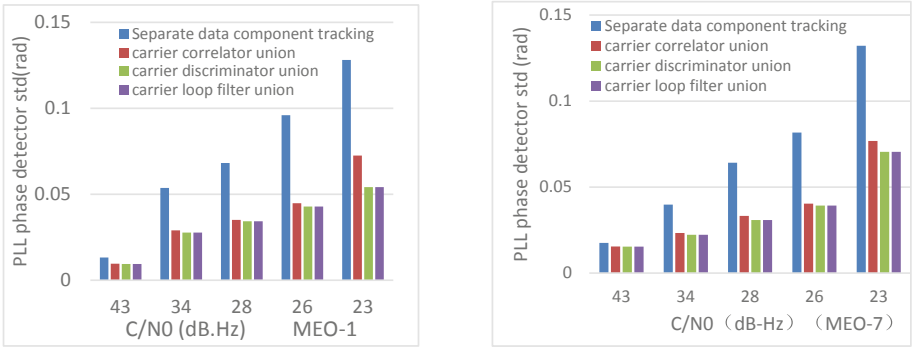


Fig. 9. B1Cd phase-detection standard deviation of carrier tracking under each tracking strategy

to-noise ratio is lower than 34 dBHz, the discriminator joint precision is better than the correlator joint mode. This is because under the lower signal-to-noise ratio condition, the correlator joint misjudges the bit symbol. As the carrier-to-noise ratio continues to decrease, the pseudo-code loop filter combination shows the best tracking accuracy. In addition, comparing the MEO-1 and MEO-7 analysis results, the higher the inter-code consistency between the data and the pilot, the more obvious the joint tracking accuracy is improved.

Figure 9 shows the PLL phase-detection standard deviation of the B1Cd component for MEO-1, MEO-7 satellite under different tracking strategies, and the code tracking uses single data component tracking. The tracking results are shown as follows: The smaller the carrier-to-noise ratio, the greater the advantage that the discriminator joint can demonstrate, and the loop filter output combination is equivalent to the discriminator output, which is consistent with the theoretical analysis results.

In view of the above statistical results, the following three combinations of joint tracking are used in the tracking loop. Combination 1 represents 'code correlator joint + carrier correlator joint', combination 2 represents 'code discriminator joint + carrier discriminator joint', and combination 3 represents 'code filter joint + carrier discriminator joint'.

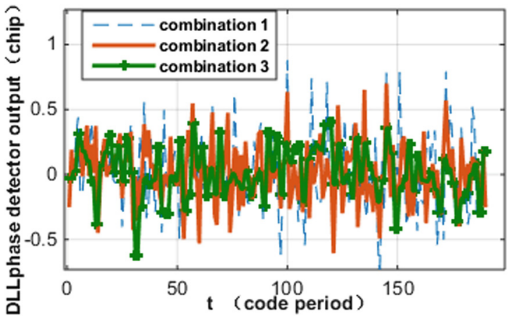


Fig. 10. DLL phase detector output under the three combined strategies (MEO-1)

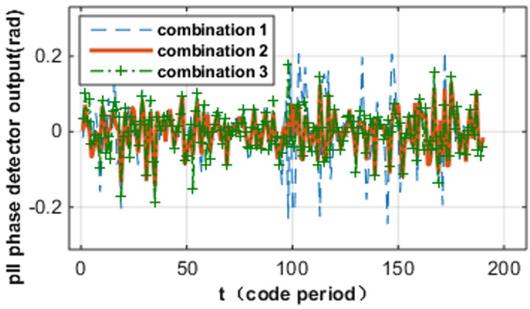


Fig. 11. PLL phase detector output under the three combined strategies (MEO-1)

Figures 10 and 11 show the comparison of the three combinations of B1Cd component code phase detector output and carrier phase detector output at lower carrier-to-noise ratio signals. The results show that the loop tracking accuracy is best under the combination of 'code filter joint + carrier discriminator joint'.

The B1Cd component's coherent integral absolute value of the tracking loop output can measure the tracking accuracy. The same data is jointly tracked in the loop using the following combinations. The mean values of the B1Cd component coherent integration amplitude are as follows:

It can be seen from the results in Table 2 that the performance improvement in joint tracking is better at low SNR. When the carrier-to-noise ratio is higher than 43 dBHz, it is recommended to use the combination of 'code correlator joint + carrier correlator joint' for tracking, whose structure is the simplest. When the carrier-to-noise ratio is less than 23 dBHz, it is recommended to use the combination of 'code filter joint + carrier discriminator joint' for tracking. In other cases, the 'code discriminator joint + carrier discriminator joint' can be selected.

Table 2. Mean value of B1Cd coherent integration amplitude under different combined strategies (unit 10^6)

Carrier noise ratio (dBHz)	43	34	28	26	23
Separate data component tracking	65.7	18.94	11.32	8.15	5.18
Combination 1	65.9	19.23	12.03	9.08	5.98
Combination 2	65.9	19.41	12.21	9.24	6.34
Combination 3	65.8	19.21	12.06	9.16	6.56
Maximum improvement	*+0.3*	*+4.9*	*+7.86*	*+13.3*	*+26.6*

5 Conclusion

To make full use of the total energy of the signal, and improve the tracking accuracy of the data channel. According to the design characteristics of B1C signal, this paper proposes three joint tracking algorithms to demodulate the messages of B1Cd signal

components. The tracking accuracy and applicable scenarios of each method in the code tracking loop and carrier tracking loop are analyzed. The measured results of satellite signals show that the tracking accuracy of the joint tracking method is more obvious under weak signal conditions. Under the condition of higher SNR, the accuracy of the three-way joint tracking method is equivalent. It is suggested to use the simplest structure of the correlator output amplitude superposition method to realize pseudo-code and carrier tracking. Under the condition of lower SNR, the carrier discriminator output linear joint mode shows the best tracking precision, and the code loop filter output joint mode shows the best tracking accuracy. Therefore, under the condition of low SNR, it is recommended to use the combination of 'Code loop filter output Joint + Carrier Discriminator output Joint'. But two sets of filters are required in loop filter output combination mode, so the loop is more complicated.

Acknowledgments. The author is grateful to the National Time Service Center for providing high-gain antenna data and high-speed acquisition equipment. This paper is supported by the National Natural Science Foundation of China (6501430) and the State Key Laboratory of Geo-information Engineering (SKLGIE2017-M-2-2).

References

1. Ries L, Macabiau C, Nouvel O (2002) A software receiver for GPS-IIF L5 signal. In: Proceeding of ION GPS, 24–27 September 2002
2. Zhongying Z, Lixin Z, Yansong M (2014) Research on MBOC signal joint tracking technology. China Acad Space Technol 1:92–96
3. Dongyang X, Yanhong K (2011) Dual-component combined tracking of GPS L5 signals. SCIENTIA SINICA Phys Mech Astron 5:653–662
4. China Satellite Navigation Office (2017) BeiDou Navigation Satellite System Signal in Space Interface Control Document Open Service Signal B1C (Version 1.0)
5. Yao Z, Lu M (2010) QMBOC modulation and its multiplexing techniques. Sci Sinica 40 (5):575–580
6. Yang C, Hegarty C, Tran M (2004) Acquisition of the GPS L5 signal using coherent combining of I5 and Q5. In: Proceedings of international technical meeting of the satellite division of the institute of navigation
7. Li Y, Nagaraj C, Dennis M (2018) An open source BDS-3 B1C/B2a SDR receiver. In: Proceedings of the 2018 international technical meeting
8. Dorf R, Simon M, Milstein L (2016) Digital communications 32(1):3–5

Research and Implementation of Acquisition Algorithm Based on Multi-correlator Parallel Frequency Search for Long Code

Qing Wang(✉), Wenquan Feng, Chen Zhuang, Qi Zhao, and Zebin Sun

School of Electronic and Information Engineering, Beihang University, Beijing 100191, China
2218061372@qq.com

Abstract. The long pseudo-noise codes have the advantages of strong anti-interference capability and high ranging accuracy, which has been used in GPS, Compass and other navigation satellite systems. However, the general acquisition algorithm is no longer applicable in the long codes because the long period property leads to the problem of large computation and high consumption of hardware resource. In order to solve the problems mentioned above, firstly, a long code acquisition algorithm based on parallel phase search which has already been proposed currently is improved by the method of partial averaging, and the simulation results show that the calculated amount is reduced. Then, the employ of multi-correlator and FFT transform in parallel frequency search can not only reduce the computation of acquisition algorithm, but also increase the acquisition speed. In order to receive the long code on the receiver, this paper concentrates on the hardware implementation of long code acquisition algorithm based on multi-correlation parallel frequency search in FPGA. Besides, the operating principle of each module is explained in detail. Through the test on the FPGA board, it is proved that the designed acquisition module can effectively and timely acquire the navigation signal. Moreover, the method of hardware implementation proposed in this paper for long code acquisition has the advantages of small hardware resource consumption and simple implementation structure, which is able to provide a reference for the engineering application of long code acquisition.

Keywords: Long code acquisition · Parallel phase search · Parallel frequency search · Multi-correlator · FPGA

1 Introduction

Pseudo random code is used to realize direct spread spectrum in communication system, which can be divided into long code and short code according to the length. Compared with short code, the long code has advantages in anti-interference, confidentiality and measurement accuracy, which has been applied in GPS, Compass and other navigation satellite systems. Acquisition is one of the key technologies for satellite navigation signal reception and a prerequisite for precise synchronization of

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 465–481, 2019.
https://doi.org/10.1007/978-981-13-7751-8_46

communication systems. The long code period is long and has a wide range of phase to be searched, which acquired by serial searches cause many problems such as large computation and long acquisition time. While using the parallel search algorithm can improve the acquisition speed and reduce the occupation of hardware resources. Van Nee [5] proposed a direct acquisition algorithm which used FFT transformation to search multiple code phases in time domain. Wolfert [6] used a large number of parallel correlator and FFT technology to search 511 code phases and 64 frequency points at a time, which are able to directly acquire the long code. In addition, the extended replica Folding Acquisition Search Technique [8] and the matching filter with FFT [2] search methods are also included in the long code acquisition algorithm.

The method proposed by Van Nee [5] has the problem of large FFT computation, and the hardware implementation structure proposed by Wolfert [6] is complex and difficult to be implemented in the receiver. In order to solve the computation problem caused by long code, this paper studies the long code acquisition algorithm, and proposes a simple acquisition implementation structure with small hardware resource consumption. In this paper, the research content is arranged as follows: firstly, a long code acquisition algorithm based on parallel code phase search is introduced in the second section. Secondly, the third section describes a long code acquisition algorithm based on parallel frequency search. After that, the two parallel acquisition algorithms above are compared in section four. Finally, this paper emphatically introduces the FPGA hardware realization of long code acquisition based on multi-correlation parallel frequency search method, and the details of how each module works are presented in turn.

2 Long Code Acquisition Algorithm Based on Parallel Phase Search

2.1 Algorithm Principle

Parallel code phase search algorithm seeks multiple code phases at one time by FFT transformation of code phase, which reduces two-dimensional search to one-dimensional search in frequency. The basic principle of this algorithm is that the Fourier transform of two signal correlation values is equal to the conjugate of the Fourier transforms of the product of two signals.

The correlation values of two sequences $x(n)$ and $y(n)$ with period length of N can be calculated as follows:

$$z(\mathrm{n}) = \frac{1}{\mathrm{N}} \sum_{\mathrm{m}=0}^{\mathrm{N}-1} \mathrm{x}(\mathrm{m})\mathrm{y}(\mathrm{m}-\mathrm{n}) \tag{1}$$

The Fourier transform of the above equation on both sides is given by:

$$Z(k) = \frac{1}{N} X(k)\overline{Y(k)} \tag{2}$$

Where X(k) and Y(k) are the discrete Fourier transform of $x(n)$ and $y(n)$ respectively, and $\overline{Y(k)}$ is the complex conjugate of Y(k) [7].

2.2 Correlation Method

Because the calculation of Fourier transform of long-period pseudo-code is very large, in order to reduce the calculation and improve the acquisition speed, the local pseudo-code is often segmented in practical engineering, and then the parallel phase search is carried out for each pseudo-code successively until the correlation peak is found.

In the process of long code acquisition, a piece of received data with phase offset may overlap with two local pseudo codes in different degrees, generating two different correlation peaks and resulting in the fuzzy problem of phase acquisition as can be seen in Fig. 1.

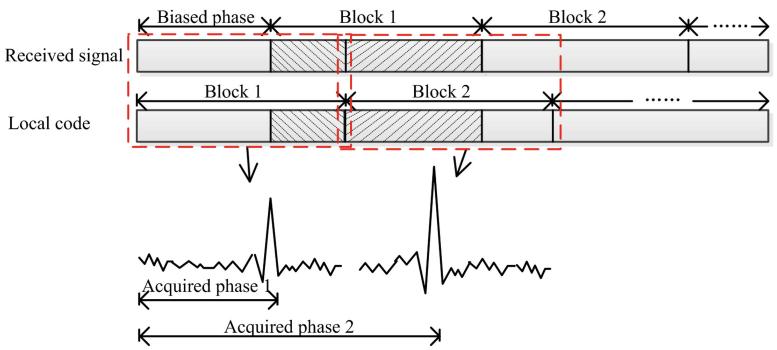


Fig. 1. Two correlated peaks of different values

In order to make parallel code phase search can be used to acquire long codes, Jing Pang used direct averaging method to reduce the amount of FFT computation in reference [3], and proposed the pseudo-code segmentation approach and related operation methods different from short code.

Assuming that the sampling rate is f_s, FFT points are N_{FFT}, and the coherent integration time is T_{coh}, the specific steps of the direct average method are shown as follows:

1. Sample the received signal $s(t)$ at a sampling rate of f_s within $T_{coh}/2$, and then averaged at every other $n = T_{coh} \bullet f_s/N_{FFT}$ sampling points to obtain $N_{FFT}/2$ sampling points, and filled with $N_{FFT}/2$ zeros to obtain N_{FFT} sampling points of the received signal which denoted as block 1;
2. Sample the local pseudo code of length T_{coh} at equal intervals to obtain N_{FFT} sampling points and denote them as block 2;
3. Transform the block 1 by FFT in N_{FFT} points;
4. Transform the block 2 by FFT in N_{FFT} points and take its conjugate;
5. Take the FFT inverse transformation of the product of (3) and (4), and abandon the former $N_{FFT}/2$ relevant results.

6. If the peak value is less than the threshold value, move the local pseudo code by the length of $T_{coh}/2$ sampling points and repeat (2)–(5) until the relevant peak is found.

In step (1), if only the first $m(m<n)$ points of n sampling points are partially averaged, the calculation amount can be m/n of that of the direct average.

The partial direct average formula:

$$s_i = \sum_{p=1+(i-1)n}^{m+(i-1)n} s(p/f_s) \tag{3}$$

Among them, $i = 1, 2, \ldots, N_{FFT}/2$.

The physical significance of the partial average method is equivalent to use a low-pass filter to calculate the weighted average of n sampling points. The first m points are passed by given a weight of 1, and the last n−m points are filtered by given a weight of 0. In this way, the partial averaging method can also prevent the generation of aliasing interference in the descending sampling process.

2.3 Simulation Result

The simulation is carried out for the long code signal acquisition by using the improved averaging method based on parallel code phase search, and the acquisition result of partial averaging is given in Fig. 2.

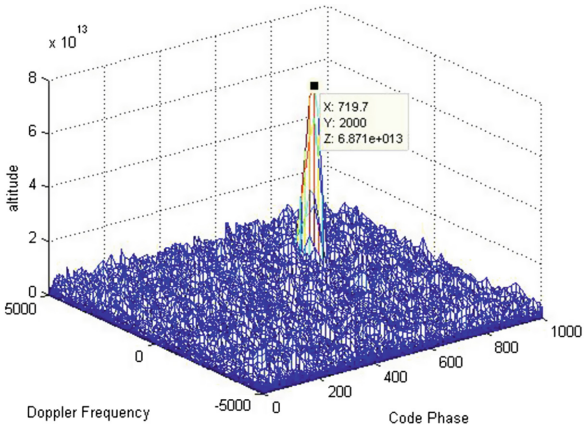


Fig. 2. Acquired result using partial average

Signal parameters in the simulation: pseudo code rate f_c is 1.023 MHz, pseudo code period is 1 s, data rate is 50 bps, modulation mode is BPSK, phase offset is 720 chips, carrier frequency is 13.5 MHz, Doppler frequency offset is 2100 Hz, and carrier noise ratio is −23 db.

Acquisition parameters: sampling rate is 124 MHz, FFT points are 8192, coherent integration time is 2 ms, frequency search interval is 200 Hz, frequency search range is

−5 kHz −5 kHz. The sampling interval $n = T_{coh} \cdot f_s / N_{FFT}$ is calculated to be about 15, and partial average points are selected as m = 8.

As can be seen from the simulation result, the acquisition peak value of the partial average is about 60 times of the noise. The acquired code phase is 719.7 chips, carrier Doppler is 2000 Hz, code phase error is 0.3 chip, and Doppler error is 100 Hz.

Therefore, the acquisition results of the partial average method are consistent with the actual situation, and the correlation peaks are still obvious in the −23 db noise environment. Therefore, the partial average method can be used to directly acquire the long codes. Furthermore, compared with the direct average method, it reduces the calculation amount by 1−8/15 = 46.7% in the relevant parts.

3 Long Code Acquisition Algorithm Based on Parallel Frequency Search

3.1 Algorithm Principle

Parallel frequency search is able to reduce the two-dimensional search to one-dimensional search. The intermediate frequency input signals is correlated with local pseudocode after being stripped the carrier, and then the Fourier transform of correlation results can reflect the correlated peaks at different frequency points.

The acquisition processes based on partially correlation are introduced in the Fig. 3. In the correlation integral time of T_{coh}, the signal and code are partially correlated at the integrated rate of f_{int} to obtain $D(D = T_{coh} f_{\text{int}})$ partial correlation values. The FFT transformation is going to be done in $K(K > D)$ points after the number of partial correlation results is zero padding to K. This partial correlation and zero padding method can not only effectively overcome the scallop loss caused by FFT transformation [1], but also improve the accuracy of frequency acquisition.

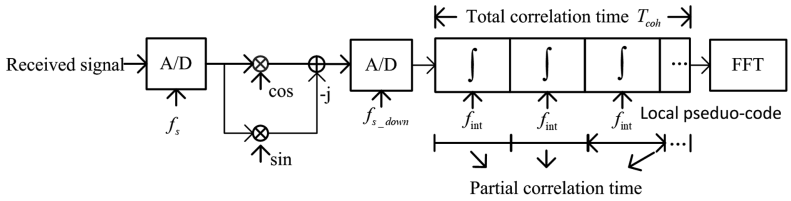


Fig. 3. Acquisition processes based on partially correlation

When the phase of received signal is aligned with that of local pseudo-code, the output amplitude of FFT transform at the $k(k = 1, 2, \ldots K)$ point is [4]:

$$X(k) = \left| \frac{sin(P\pi f_d T_c)}{P sin(\pi f_d T_c)} \right| \cdot \left| \frac{sin(N\pi f_d T_c - \pi \frac{D}{K} k)}{sin(\frac{N}{D} \pi f_d T_c - \pi \frac{1}{K} k)} \right| \tag{4}$$

Where, $N = f_{s\_down} T_{coh}$ is the number of sampling points within the coherent integration time T_{coh}, $P = N/D$ is the number of sampling points within the partial

coherent integration time, f_d is the carrier Doppler, and $T_c = 1/f_{s\_down}$ is the sampling interval.

The formula of Doppler acquisition resolution using partially correlation is shown as:

$$\Delta f = \frac{f_{int}}{K} \tag{5}$$

Therefore, it can be seen that the Doppler resolution Δf is related to the number of FFT points K and the partial coherent integration rate. The acquired Doppler resolution can be improved by decreasing the partial coherent integration rate when the number of FFT points is constant.

3.2 Performance Analysis

When the number of FFT is certain and the Doppler offset is large, the ratio of FFT points to the number of partial correlator will affect the acquired peak. The sampling rate is 8.192 MHz, FFT point K is 128, and the coherent integration time is 1 ms. As the number of partial correlation integrators is 32, 64, and 128 respectively, the variation of FFT output peak in the carrier Doppler ranging from 0 kHz to 10 kHz range is shown in the Fig. 4.

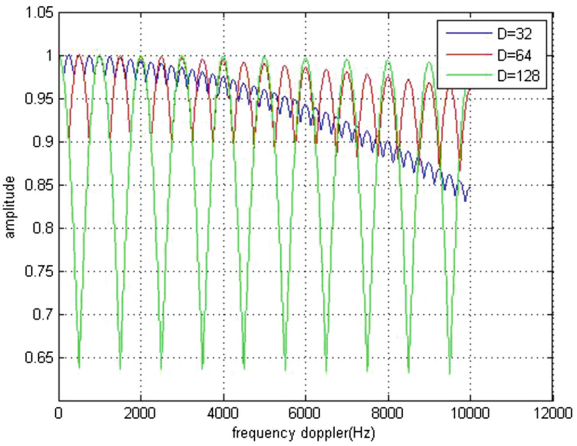


Fig. 4. Curve of FFT peak values changing with Doppler at different partial correlation integrators

According to the figure above, the magnitude of FFT peak is affected by the different number of coherent integrators. As D = 128, that is, FFT transformation is carried out directly without padding zero, it can be seen from the figure above that the acquisition peak is greatly affected by Doppler especially when the Doppler frequency offset is a multiple of 500 Hz. The attenuation of the acquired peak reaches a maximum

of 35%. When half of FFT data is zero, that is D = 64, the max attenuation of the acquired peak value caused by Doppler offset is 10% and the curve envelope change is not obvious with the increase of Doppler. When D = 32, in other words, the number of zero padding is less than half of FFT number, the peak attenuation can be less than 10% as the carrier Doppler is between 0–8000 Hz and the peak attenuation is more than 10% when the carrier Doppler is between 8000–10000 Hz.

It can be seen that the peak value loss of FFT can be effectively overcome by padding the relevant data with 0. When the Doppler frequency offset is small, the frequency resolution can be raised if the number of zero padding is less than half of FFT points. When the Doppler frequency offset is large, the peak fluctuation range can be reduced with the number of zero padding exceeding half of FFT points.

3.3 Parallel Correlator

The acquisition algorithm above can only search one code phase at a time. In this paper, a parallel multiple correlator is used for searching multiple code phases simultaneously to reduce the acquisition time and improve the acquisition speed as shown in Fig. 5.

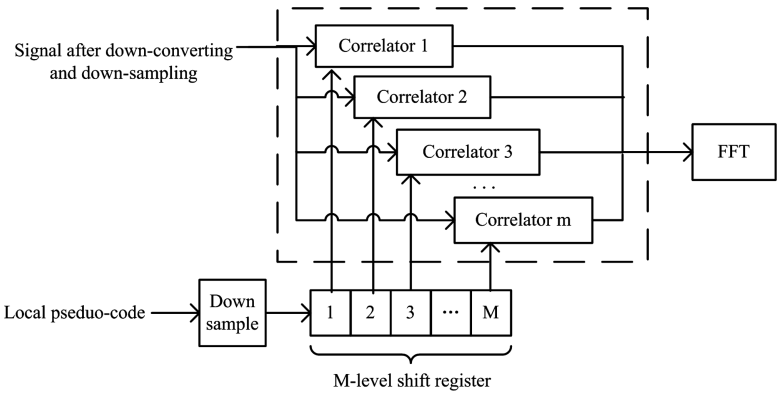


Fig. 5. The structure of parallel correlator

After the local pseudo going through M -level shift register, pseudo codes with M different phases are generated, and they are then input into M parallel correlator respectively together with the received signal. After that the correlator outputs are padded by zero for K point FFT transformation, which can simultaneously search M*K search units and improve the acquisition speed by M times.

3.4 Simulation Result

The signal parameters adopted in the simulation are consistent with Sect. 2.3. Some important parameters of acquisition are shown as follows: coherent integration time

T_{coh} is 2 ms, sampling rate is f_{s_down} 8192 MHz, partial correlation integration rate f_{int} is 20 kHz, FFT point K is 128, and the number of parallel correlator M is 64.

It can be calculated that the captured Doppler resolution is $\frac{f_{int}}{K} = 156.25$ Hz and the acquired code phase resolution is $f_c/f_{s\_down} \approx 0.125$ chips under the current simulation parameters.

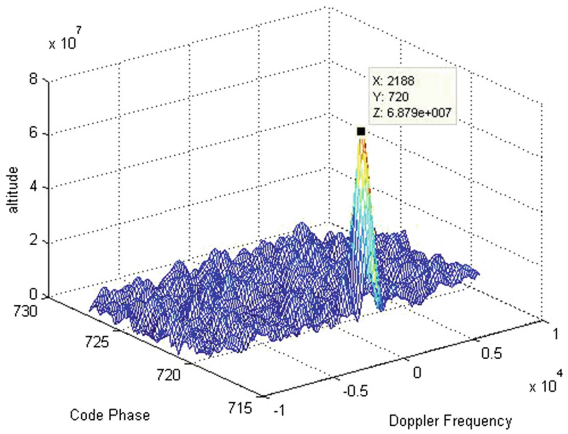


Fig. 6. The acquisition result of multi-correlator frequency search

As can be seen from the acquired results shown in Fig. 6, the acquired Doppler error is 87.5 Hz and code phase error is 0.232 chips. In the noise environment of −23 DB, the acquisition accuracy can still reach the theoretical error range. The simulation results show that the algorithm can acquire long code directly.

4 Parallel Acquisition Algorithm Comparison

The long code has a longer cycle and a wider code phase range than short code. Hence, when use the parallel search methods in long code acquisition, not only the relevant methods need to be improved, but also the algorithms require to be optimized in order to accelerate the acquisition. The compared results of different acquisition methods are listed in the Table 1.

Table 1. Comparison of acquisition methods

Acquisition methods	Correlation methods	Optimization
Parallel phase search	Fill the received signal with zero and segment the local pseudo codes with 50% overlap	Partial average for reduction in computation
Parallel frequency search	Correlate partially and fill with zero	Multiple correlator to improve capture speed

The two parallel algorithms were used to acquire the 1s-period long code respectively, and the average acquisition time of each 500 tests was recorded under their corresponding conditions. The quantitative test results about long code acquisition performance are shown in Tables 2 and 3 respectively.

Table 2. The mean acquisition time of parallel frequency search

Number of partial correlation integrator	Number of parallel correlator	Mean acquisition time (s)
16	512	682.5
32	512	764.3
48	512	795.1
64	512	858.5
20	32	3755.7
20	128	1207.2
20	512	689.2
20	2048	572.1
20	8192	480.6

Table 3. The mean acquisition time of parallel code phase search

Average method	Percentage	Mean acquisition time (s)
Direct average	100%	628.3
Partial average	53.3%	618.9

The simulations continue to use the parameters mentioned in Sects. 2 and 3, except for special instructions.

As can be seen from Table 2, the increased number of partial correlation integrators will increase the average acquisition time. When the number of parallel correlator is increased from 32 to 512, the average acquisition time decreases greatly and the acquisition speed increases significantly. When it is increased from 512 to 8192, the average acquisition time decreases little and the enhanced of capture speed is not obvious.

The search range, number of frequency search points and search of parallel code phase search are changed to 20 kHz, 128 and 156.25 Hz respectively to be consistent with that of parallel frequency search, so that average acquisition time of the two search methods can be compared fairly.

The percentage in the Table 3 represents the proportion of the number used for average calculation in the total sample points at the adjacent sampling time.

When 8192 parallel correlator are used in parallel frequency search, 8192 sampling phases can be simultaneously searched within a coherent integral time just like parallel code phase search. What's more, the average acquisition time of the former (480.6) is less than that of the latter (618.9).

However, the number of parallel correlator cannot be too large due to the constraints of hardware resources and timing. Therefore, in order to make the acquisition speed of parallel frequency search almost as fast as that of the parallel code phase search, this article use 64 parallel correlator to run 8 times continuously in a coherent integral time, which are equivalent to using 512 parallel correlator.

The method of parallel phase acquisition can search multiple code phases at one time and has a fast acquisition speed, but the FFT transform requires high number of points and it requires two FFT transformations and one FFT inverse transformation, which takes up high hardware resources. This method is more suitable for software receiver and can acquire long code quickly.

The hardware is fast in processing speed and limited by FFT conversion points, thus it is more suitable to use parallel frequency search method. In the next section, this article will introduce an implementation method of parallel frequency search in hardware FPGA.

5 Hardware Implementation of Acquisition Algorithm

According to the long code acquisition algorithm described in the third section, the designed implementation construction on the hardware platform is shown in Fig. 7.

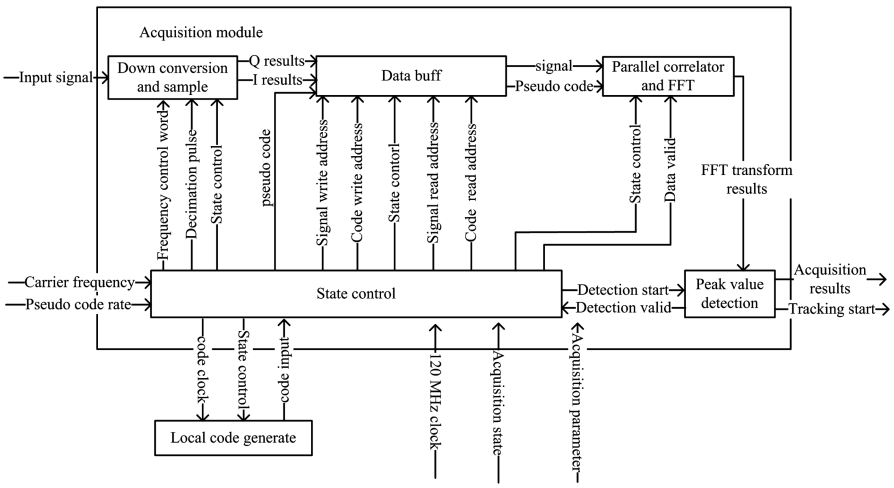


Fig. 7. The block diagram of the overall acquisition structure

Firstly, the down-conversion and extraction module converts down the received signal with the search frequency and then it is extracted at a certain down-sampling rate. Afterwards, the output sampled signal from the down-conversion and the pseudo-code are cached in the data buffer module, waiting for being read. Then, the corresponding data and pseudo-code are read from the cache module according to the read-write address from the state control module and the correlation values of the two are

calculated in the parallel correlation FFT module. Finally, the detection module detects the FFT transform results. If the peak value is greater than the threshold value, the detected code phase and carrier Doppler are output; otherwise the searching process is continued.

The input and output from the down converting and sample, data buff, parallel correlator and peak detection modules are all transmitted via the state control module. On the one hand this design is to coordinate and control the running state of each module, and on the other hand it can achieve the function of searching in different code phase and frequency range.

Parameters of the acquisition module can be adjusted according to the actual engineering requirements, but they must meet the timing constraints to avoid the timing conflicts between modules. The following part is a detailed description of how each module works with the capture parameters provided in Table 4.

Table 4. The important parameters of acquisition module

Parameter	Values
Coherent integration time T_{coh}	1 ms
Down sampling rate	8192 kHz
Number of parallel correlator M	64
Times of continuous correlation C	8
FFT points K	128
Processing clock frequency	120 MHz
Rate of correlated integration	20 kHz

5.1 Down Conversion and Extraction

Once the acquisition module is started, it wills continuously down-convert the received signal. If no peak value is found after searching all the code phase, the state control module will adjust the frequency control word of down conversion module and switch the frequency search range to search again.

As shown in Fig. 8, signal resample_en is on the behalf of the clock with a frequency of 8192 kHz which is used to down sample the signal sampled by 120 MHz. Thus the signal needs to be calculated every 15 sampling points on average. This paper adopts the partial averaging method described in Sect. 2.2, only the first 8 useful signals are averaged. The sampling results after down converting of I channel and Q channel are shown as result_i and result_q respectively.

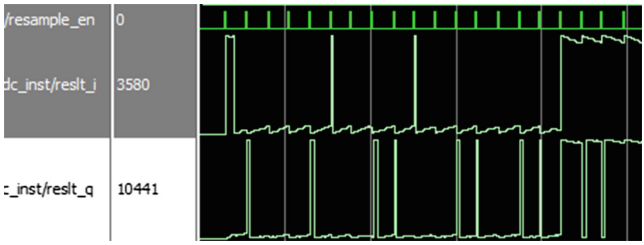


Fig. 8. The sampling results of down-conversion

5.2 Data Buff

The sampled signal and pseudo-code are both stored in three segments of ring RAM, each length of which is equal to the number of down sampling points in a coherent integration time.

The read-write operations of data buffer module can be seen in Fig. 9. Three segments of RAM are read and written in turn while the signal and pseudocode are cached. After the acquisition is started, it begins to read from the first segment when data is written to the third segment. When the third segment finishes writing the data within a coherent integration time, it begins to write the data to the first segment and reads from the second segment, and the rest can be done in the same manner.

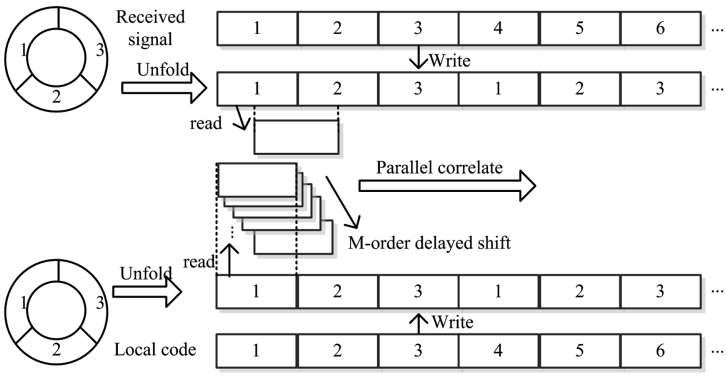


Fig. 9. The read-write operations of data buffer module

On the first search, both the signal and pseudocode are read from the start of their first segment of RAM. The read pseudo codes are delayed by the m-level shift register, which are then correlated with the received signal to search M phases. When the second search is carried out, the reading position of pseudo-code remains unchanged, but the signal is moved to read from the M-th position to be correlated with the m local pseudo-code in parallel. In the same way, keep moving the receiving signal to be

associated with the local pseudo code until the third segment of RAM finishes being written.

Signals are read across a block of RAM, so signals need three blocks of RAM to be stored to ensure that two blocks of RAM are ready for reading while one of them is being written. Pseudo code just need two pieces of RAM to store in order to achieve the function of read and write in circular. However, three pieces of RAM are also used to cache the code in order to share the same write address with the signal.

According to the timing analysis of the parallel correlation FFT in the following section, only the 512/8192 = 1/16 signal phase search is completed during the signal cache period about 120 k clocks. In order to complete the search for the remaining code phase of the same 1 ms signal, both the pseudo-code and the signal need to be switched to the next segment at the same time. Therefore, this paper designs the storage structure of three segments of ring RAM to realize the cyclic operation of signal reading and processing.

5.3 Parallel Correlator and FFT

The working time process of Parallel and FFT is shown in Fig. 10. The parallel correlation and FFT module reads the sampling down-conversion signal and pseudo-code from the data buff module, which is correlated and sampled at the correlation integration rate. Then the results are stored in two independent RAM to buffer alternately. After completing a correlation operation, the Streaming I/O type of IP core is used to start the FFT calculation, and the FFT transform result is transmitted to the peak detection module.

All 64 parallel correlator read 8,192 sampled data in each channel, which last for 8256 clocks in total. Values from 64 channels are output every 410 clocks in turn and are written to the RAM to be cached. Every time a search is completed, the two buffers swap caches. Therefore, there are a total of (8192 + 64) * 8 = 66048 clocks required to complete the correlation operation in 1 ms.

The first FFT operation starts after completing a correlation operation, and the output results are delayed after about 365 cycles. Each time 20 correlations of one correlator are read from RAM, and then 128 FFT operations are performed by zero padding the 20 correlations. A total of 64 * 8 = 512 FFT operations are performed within 1 ms. Therefore, 8192 + 64 + 365 + 128 * 512 = 74,157 clocks are needed to complete the 512 FFT operations.

Parallel correlation and FFT are running simultaneously in timing, thus 512 phases and 64 frequency points in 8192 sampling points are searched, occupying a total of 74,157 clocks.

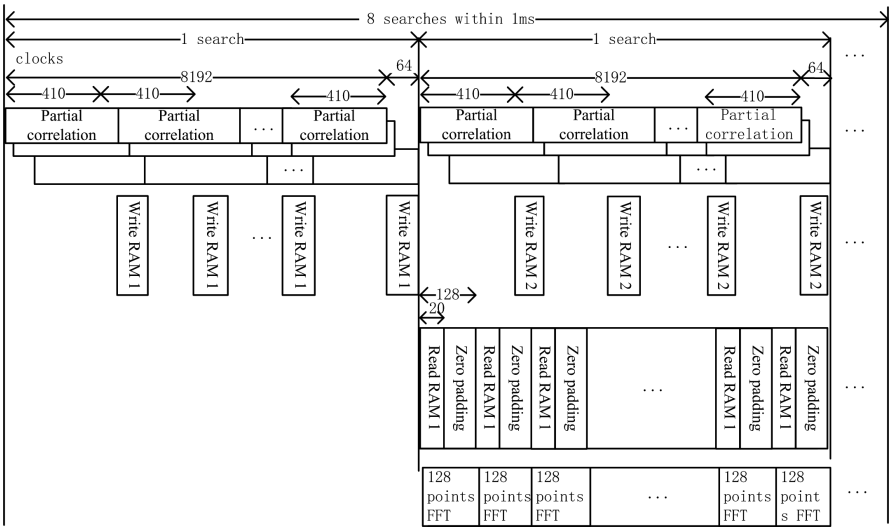


Fig. 10. The working time process of Parallel and FFT

5.4 State Control

Because range of the cached signal cannot be searched completely in 1 ms and the long code acquisition also needs to switch the searched code segment, so the state control

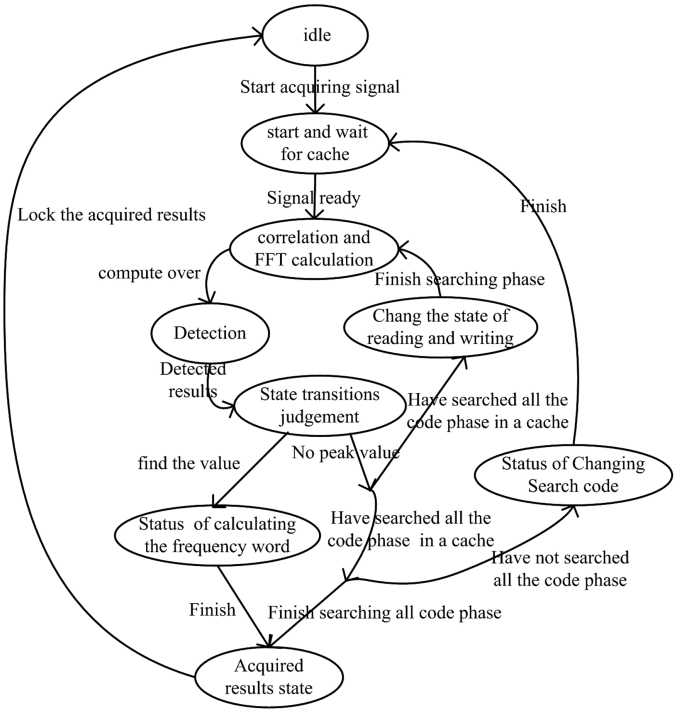


Fig. 11. The state transition of a finite state machine

module needs to control the running state of each module to realize the search of all code phase. The state control module designed in this paper realizes the timing control of the entire acquisition module through the finite state machine, and its state transition diagram is shown in Fig. 11.

When receiving the flag of acquisition starting, the idle state machine enters the state of starting and waiting for cache, in which it will down convert and decimating the received signal. The parallel relevant FFT module will start working after the cache is ready. After finishing the correlation and FFT calculation, the next transition state is judged according to the detection result: when the peak value is detected, calculate the acquired frequency word result and then enter the acquisition result state; When there is no peak, if all the signal phase in a cache has not been searched the state of read will be changed to continue to search the rest of the phase, if there are still residual signals need to be searched further after finishing the search in a cache the status of code segment searching need to be changed to store the next segment of signal, if all code phases have been searched a failed acquisition result will be returned and the state machine go to the state of acquisition results.

In order not to affect the normal running state and timing of each module when switching data segments, the processes including parallel correlation FFT, peak detection, and all group delays, a total of about 74157 + 1000 + 375 = 75532 clocks, all need to be completed within 120,000 clocks of a signal cache.

5.5 Peak Value Detection

The functions of peak detection module are as follows:

1. Square the real and imaginary parts of FFT results and cache them;
2. Compare all FFT amplitudes to find the maximum value, and record the amplitude, phase and frequency of the maximum value;
3. According to the location of the peak, the data locating at a certain range away from the peak (such as 1000 points) is selected for statistical average. The averaged results are as the estimation of noise power, which is then multiply by the input threshold coefficient as the signal judgment threshold;
4. When the signal power is greater than the threshold, the effective detection mark is given and the value of phase, frequency, maximum peak and threshold are locked at the same time.

As shown in Fig. 12, the Verilog code of the acquisition model was flashed to FPGA board of XC7K325T type from Xilinx company for long code acquisition test. The signal parameters adopted are consistent with Sect. 2.3 except for the Doppler and phase bias. The acquisition starts at a random phase in the signal each time, and the acquired results observed timely by Chipscope are shown in Fig. 13.

The flag acq_success is high level which illustrates that the signal has been successfully acquired. The acquired carrier Doppler expressed by carr_dopler_freq_fcw is 0 Hz which is agreed with the real conditions, and it can be known that the phase offset is occurred at 5378 sampling points in paragraph 848.

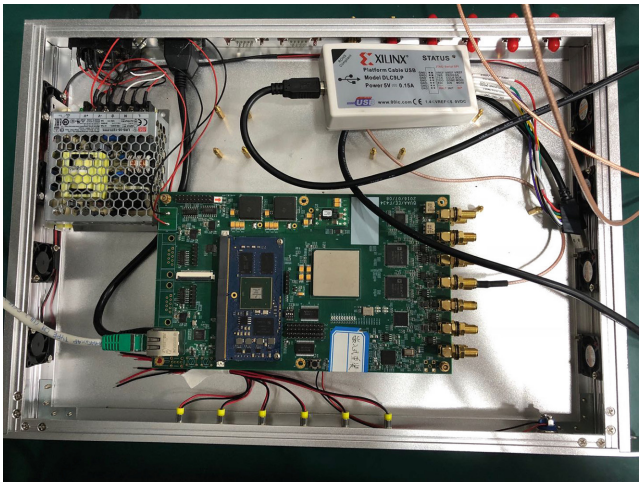


Fig. 12. Acquisition test on the FPGA board

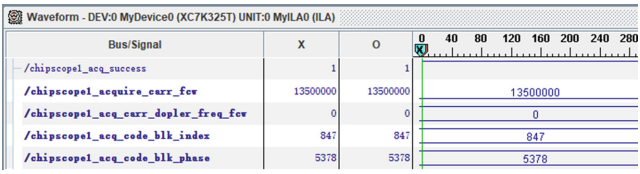


Fig. 13. The acquisition results observed by Chipscope

The usage of chip logic is shown in Fig. 14, in which the registers are used 4%, the lookup table is used 7%, and the memory is used 2%. It can be seen that the hardware acquisition structure designed in this paper consumes little hardware resources and has a simple implementation structure.

```
Slice Logic Utilization:
  Number of Slice Registers:              18,569 out of 407,600    4%
    Number used as Flip Flops:            18,542
    Number used as Latches:                   21
    Number used as Latch-thrus:                0
    Number used as AND/OR logics:              6
  Number of Slice LUTs:                   15,800 out of 203,800    7%
    Number used as logic:                 12,983 out of 203,800    6%
      Number using O6 output only:         5,958
      Number using O5 output only:         1,083
      Number using O5 and O6:              5,942
      Number used as ROM:                      0
    Number used as Memory:                 1,668 out of  64,000    2%
      Number used as Dual Port RAM:            0
      Number used as Single Port RAM:          0
      Number used as Shift Register:       1,668
        Number using O6 output only:         626
        Number using O5 output only:          67
        Number using O5 and O6:              975
    Number used exclusively as route-thrus:  1,149
      Number with same-slice register load:    898
      Number with same-slice carry load:       251
      Number with other load:                    0
```

Fig. 14. The hardware source assumption in FPGA

6 Conclusion

Due to the wide phase range and long period of long codes, when parallel code phase search and parallel frequency search are applied to long code acquisition respectively, it is necessary to improve relevant methods and optimize the acquisition algorithm to reduce the computation and acquisition time. In this paper, on the one hand, the computation of the parallel code phase acquisition algorithm is reduced by 46.7% by using the partial average method; on the other hand, the acquisition time of parallel code phase acquisition algorithm is greatly reduced by running 64 parallel correlator for 8 times in a row. Software simulation verifies that both of these two acquisition algorithms can directly capture long code, and parallel code phase search is more suitable for software receiver. When the parallel frequency search algorithm is implemented in FPGA hardware, the modules of down-conversion extraction, data cache, parallel correlation FFT, peak detection and state control are designed. The functions of searching code phase and frequency are realized by reasonably coordinating the running state and effectively controlling the timing behavior of each module. The tested result on FPGA board proves that this module can acquire the received long code signal in time. By comparing the resource consumption diagram in the FPGA, it can be concluded that the acquisition module designed in this paper based on multi-correlation parallel frequency search has the advantages of small resource occupation and simple implementation. The research results of this paper can provide reference for the engineering implementation of long code acquisition.

References

1. Suli G, Yunfei L (2003) A fast pseudo code acquisition method based on FFT. Commun Technol 29(1):11–13
2. Chang L, Jun Z, Zhu Y, et al (2011) Analysis and optimization of PMF-FFT acquisition algorithm for high-dynamic GPS signal. In: Cybernetics and intelligent systems (CIS), Conference Publications, Qingdao, pp 185–189
3. Pang J, Starzyk J (2003) Fast direct GPS signal acquisition using FPGA. In: The 16th European conference on circuits theory and design, ECCTD 2003, Krakow, September 2003
4. Spangenberg SM, Scott I, Mclaughlin S et al (2000) An FFT-based approach for fast acquisition in spread spectrum communication systems. Wirel Pers Commun 13(1–2):27–55
5. Nee DJRV, Coenen AJRM (1991) New fast GPS code-acquisition technique using FFT. Electron Lett 27(2):158–160
6. Wolfert R, Chen S, Kohli S, et al (1998) Direct P(Y)-code acquisition under a jamming environment. In: IEEE position location and navigation symposium, IEEE
7. Xie G (2011) Principles of GPS and receiver design. Electronic Industry Press, Beijing
8. Yang C, Vasquez J, Chaffee J (1999) Fast direct P(Y)-code acquisition using XFAST. In: Proceedings of international technical meeting of the satellite division of the institute of navigation, pp 317–324

Time Primary Standard and Precision Time Service

A New Apparatus with Constant Voltage to Measure the Flux of Atomic Hydrogen Beam in Hydrogen Maser

Xiumei Wang(✉), Tiezhong Zhou, Chengyuan Zhang, Qiong Wu, Mengzhi Wang, Yaxuan Liu, and Lianshan Gao

Beijing Institute of Radio Metrology and Measurements, Beijing, China
wangxiumei@siom.ac.cn

Abstract. A new apparatus with constant voltage to measure the flux of atomic hydrogen beam in hydrogen maser is developed and verified in this paper, which consists of series circuit with thermistor and the temperature-controlled cylinder etc. During the measurement, the flux of atomic hydrogen beam can be transferred in the flux voltage signal on the measurement thermistor through self-heat and the temperature on this apparatus is controlled to reduce the influence of environment temperature. This apparatus is installed on the physical package of hydrogen maser and then the supply voltage of series circuit is optimized. The test result shows that the flux voltage signal on the measurement thermistor is correlated positively to the flux of atomic hydrogen beam induced by the Nickel purifier voltage under the optimal working parameter. The variation of Nickel purifier voltage 0.1 V can change the flux voltage signal about 23.7 mV when the Nickel purifier voltage is 1 V. Compared with the ion current characterizing the flux of atomic hydrogen beam traditionally, the noise of the flux of atomic hydrogen beam measured by the new apparatus is 3 orders of magnitude lower than the ion-pimp current. So this new apparatus can measure the flux of atomic hydrogen and the measured flux voltage signal can characterize the flux of atomic hydrogen more precisely which can be the error to stabilize the flux of atomic hydrogen beam entering the storage bulb and to fulfill the requirement of the long term frequency stability of Hydrogen Maser in the future.

Keywords: Hydrogen maser · Nickel purifier · The flux of atomic hydrogen beam · Flux meter

1 Introduction

The output frequency of oscillator in Hydrogen Maser is locked to the radiated transition in the hyperfine structure of atomic hydrogen in the ground state, whose transition frequency is about 1420 MHz [1]. During the frequency locking process, the density of hydrogen atoms depends on the total flux of hydrogen atoms and the spin-exchange collisions between hydrogen atoms in $|F = 0, m_F = 0>$ and $|F = 1, m_F = 0>$ can lead directly to a residual frequency shift [4]. In addition to this direct shift, there is a broadening of atomic resonance due to the associated relaxation [2], which can lead to a frequency shift via Q factor of atomic line [5]. Then the Q factor of atomic line will

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 485–495, 2019.
https://doi.org/10.1007/978-981-13-7751-8_47

affect the long-term frequency stability through cavity pulling shift [6]. So the flux of atomic hydrogen beam is an important factor to limit the performance of Hydrogen Maser, and it is necessary to detect the flux of atomic hydrogen beam to improve the long-term frequency stability.

In present, the hydrogen molecules beam is purified by the Nickel cube or palladium-silver alloy which is convolved on ceramics cube [8]. When the Nickel purifier is supplied by the constant voltage or current to high temperature, the large gaps between Nickels can allow hydrogen molecules pass through the wall of cubes to the vacuum chamber and the other gas molecules are forbidden. There is no physical character in this technology to characterize the flux of atomic hydrogen beam and this technology only can ensure the purity of hydrogen molecules beam. Although the ion-pump current related to the pressure in physical package can present the flux of atomic hydrogen beam, the excessive suction will cause the high temperature on the ion-pump and the unstable ion current. So there is no apparatus to measure the flux of atomic hydrogen beam and no physical character to characterize it to fulfill the requirement of long term frequency stability of Hydrogen Maser.

The conductivity of hydrogen molecule is 0.17 W/(m • K), which is higher than other common gases. This paper is aimed to measure the flux of atomic hydrogen beam and the measurement thermistor in the new apparatus with constant voltage to measure the flux of atomic hydrogen beam is as a heat source. The dependence of the flux voltage signal on the measurement thermistor on the flux of atomic hydrogen beam induced by the Nickel purifier voltage is measured and the noise of the flux voltage signal is analyzed. This new apparatus is also called flux meter. This paper is divided into four sections. The second section presents the set-up and theory of this apparatus. The third section shows the optimal working condition, the measurement and noise analysis of flux voltage signal. The last section is conclusions.

2 Experimental Set-Up and Theory

The structure of the new apparatus with constant voltage to measure the flux of atomic hydrogen beam is shown in Fig. 1, which consists of measurement resistance R_1, Platinum-Iridium lead wire, rods plated with gold, pins, rubber seal band, main body, circular electrical connector, the heating coils on measurement body, temperature-controlled cylinder, aerogel, fastener, lead wire on resistance, input pipe, socket, socket bracket and measurement circuit. The measurement resistance R_1 is the key components and the sensitivity of this apparatus to the flux of atomic hydrogen beam is mainly determined by its characteristics of resistance and conductivity. Here, one negative temperature coefficient (NTC) glass bead thermistor is selected. The resistance and temperature characteristics of measurement thermistor R_1 and the specification performance of R_1 are shown in Fig. 2 and Table 1, respectively. The negative temperature coefficient of this resistance can avoid the large surge current and it has wide temperature and resistance range, small dissipation constant, short response time and small size. The right upper corner of Fig. 1 shows the connection method of this thermistor to the circular electrical connector, in which the measurement thermistor R_1 is welded on the two rods plated with gold using its platinum-iridium lead wires and

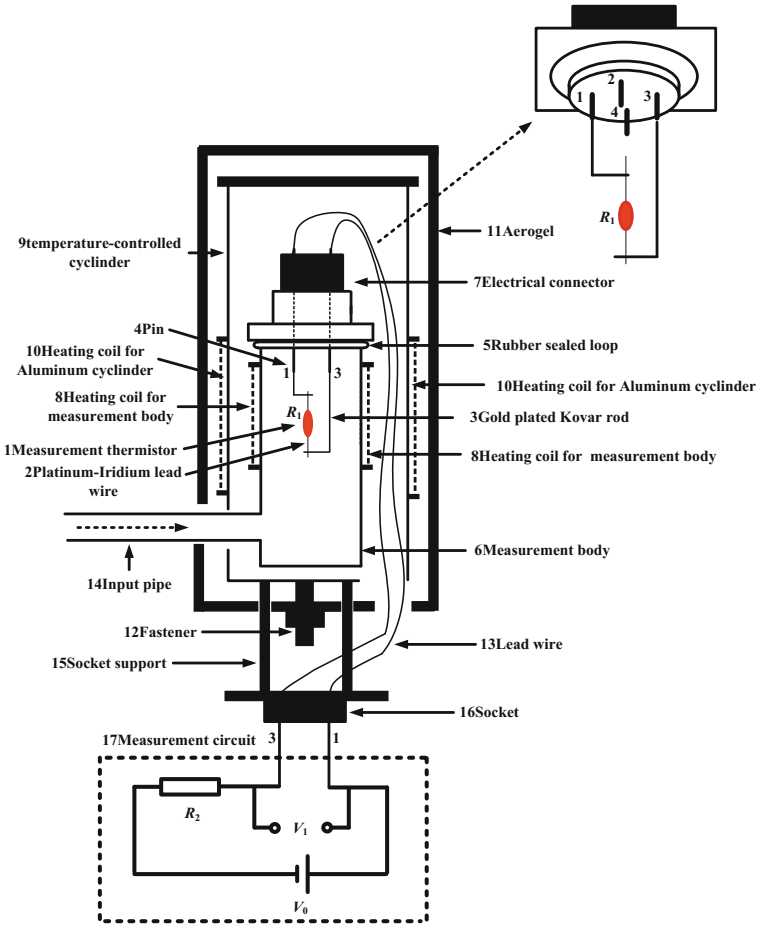


Fig. 1. The structure of the new apparatus with constant voltage to measure the flux of atomic hydrogen beam (flux meter)

two rods plated with gold is installed on the connector with two pins. Not only the rods plated with gold can support, fix and connect the measurement thermistor R_1, but the low heat coefficient can reduce the impact of environment temperature on the measurement thermistor R_1. The circular electrical connector and measurement body which are made by stainless steel are sealed by the rubber seal band and the main body is installed on the bottom of temperature-controlled aluminum cylinder. To reduce the impact of environment temperature further, there are two heat coils on the external of measurement body and aluminum cylinder, respectively. Besides, the aerogel is packed at the outside of the temperature-controlled aluminum cylinder. There is a series circuit in the measurement circuit which is made of a measurement thermistor R_1, a precise thermistor R_2 and the supply voltage of series circuit V_0.

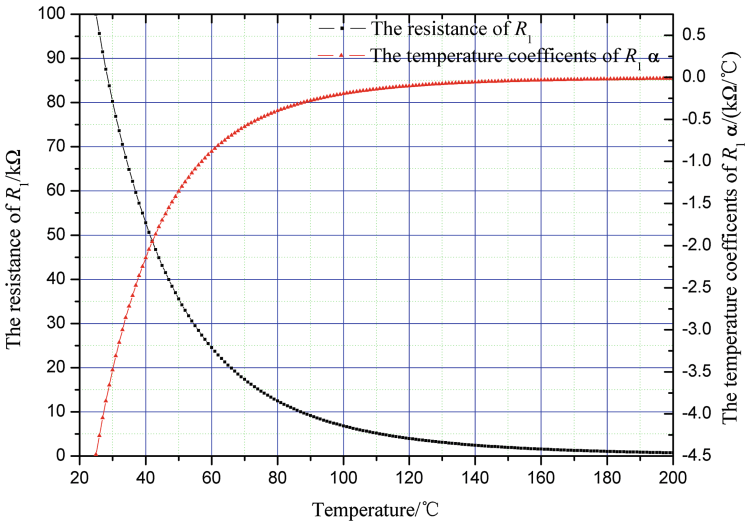


Fig. 2. The resistance and temperature characteristics of measurement thermistor R_1

Table 1. The specification performance of R_1

Parameters	Value
Resistance	100 kΩ@25 °C
Working temperature	−60 °C to 300 °C
Diameter	0.36 mm
B value	3984
Dissipation constant	0.1 mW/°C
Response time	0.5 s
Package	Glass bead
Lead wire	Platinum-Iridium lead wires

The inner vacuum of physical package is a self-balanced system. When the system is in balance, the input flux of atomic hydrogen beam from Nickel purifier W_1 is equal to the output flux of atomic hydrogen beam absorbed by the ion-pump W_2 and the pressure in physical package P is constant. During this process, the flux of atomic hydrogen beam entering the storage bulb W_s is the half of the output flux of atomic hydrogen beam absorbed by the ion-pump W_2 thanks to the magnet state selection,

$$W_s = W_2/2. \tag{1}$$

At the same time, W_2 is proportional to the pressure in physical package P due to the feedback of ion-pump,

$$W_2 = K_B P, \tag{2}$$

where K_B is the sensitivity of absorption of ion-pumps. Considering Eqs. (1) and (2), we have

$$W_s = 1/2K_B P. \tag{3}$$

That means the flux of atomic hydrogen beam entering the storage bulb W_s is proportion to the pressure in physical package P. Therefore the measurement of W_s can be transferred into the measurement of pressure in physical package P.

When there is no hydrogen beam, the temperature on the measurement thermistor R_1 becomes higher and then tends into balance through self-heat supplied by voltage V_0. The initial temperature, resistance and voltage in balance on the measurement thermistor R_1 are T_{01}, R_{01} and V_{01}, respectively. When there are hydrogen molecules beam, there are three ways to dissipate the heat on R_1, the heat conduction of hydrogen molecules Q_c, the heat conduction of lead wires Q_l and the heat radiation Q_r. In low pressure (10^{-1} to 10^0Pa), there is no relation between the heat conduction of lead wires Q_l and the heat radiation Q_r and the pressure P, but the heat conduction of hydrogen molecules Q_c is proportional to the pressure P [10],

$$Q_c = \mathrm{APK}(T_1 - T_0)\mathrm{S}, \tag{4}$$

where A is the adaption coefficient, K is conductivity of hydrogen molecules, T_1 is the self-heat temperature of measurement thermistor R_1, T_0 is the temperature of hydrogen molecules and S is the surface of measurement thermistor R_1. The heat conduction of hydrogen molecules Q_c will break the balance of self-heat in measurement thermistor R_1 and the variation of temperature, resistance and voltage are ΔT_1, ΔR_1, and ΔV_1, respectively. They are written as

$$\Delta T_1 = Q_c/C, \tag{5}$$

$$\Delta R_1 = \alpha \Delta T_1, \tag{6}$$

$$\Delta V_1 = V_0 R_2 \left(\frac{1}{R_{01} + R_2} - \frac{1}{R_{01} + \Delta R_1 + R_2}\right) \tag{7}$$

where C is dissipation constant and α is the temperature coefficients. From Eqs. (3)–(7), it is obtained that the voltage variation ΔV_1 is proportional to the pressure P which is proportional to the flux atomic hydrogen beam entering the storage bulb W_s. So the voltage variation ΔV_1 is proportional W_s and it can characterize the flux of atomic hydrogen beam entering the storage bulb. Here, the voltage variation ΔV_1 is called flux voltage signal.

The flux meter is applied on the hydrogen maser and the experimental set-up is shown in Fig. 3. There are two ports (A and B) in Nickel purifier and they are connected to discharge bulb and flux meter, respectively. The hydrogen molecules from hydrogen source (temperature-controlled on T_H) purified the Nickel purifier enter the

discharge bulb and this flux meter (temperature-controlled on T_M) through the port A and B, respectively. The ion-pump maintains the pressure of the set-up and the ion current is written as I. To stimulate the actual temperature environment of the flux meter, the hydrogen source, Nickel purifier, flux meter, discharge bulb and so on are placed in a temperature-controlled cabinet whose temperature is $T_0 = (30 \pm 1)$ °C. After the measurement thermistor R_1 supplied by V_0 is in balance, the Nickel purifier is heated by the voltage power V_N and the flux of atomic hydrogen beam entering the storage bulb W_s is converted into the measured flux voltage signal ΔV_1. All physical parameters V_0, V_1, V_N and so on are collected by acquisition module and then deal with and saved by computers.

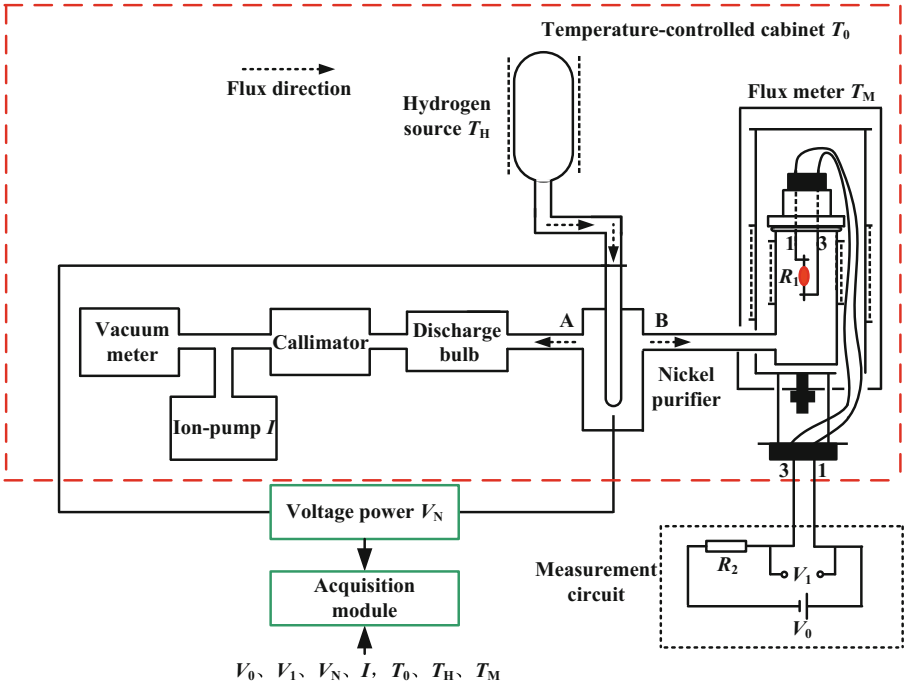


Fig. 3. The experimental set-up

3 Results and Discussions

The temperature-controlled of flux meter is set as $T_M = 57$ °C $\pm$ 0.02 °C and the Nickel purifier voltage V_N is tuned to stimulate the variation of flux of atomic hydrogen beam entering the storage bulb ΔW_s. The supply voltage of series circuits V_0 is mainly determined by the sensitivity of flux voltage signal ΔV_1 to the variation of the flux of atomic hydrogen beam entering the storage bulb ΔW_s. Then at this optimal working condition, the flux voltage signal ΔV_1 is measured and analyzed under the different Nickel purifier voltage V_N.

3.1 The Supply Voltage of Series Circuit V_0

The self-heat of the measurement thermistor R_1 is determined by the supply voltage of series circuit V_0. Figure 4 shows the dependence of flux voltage signal ΔV_1 on the supply voltage of series circuit V_0 when the Nickel purifier voltage V_N is increased from 0.5 V to 1 V. It is obtained that the flux voltage signal ΔV_1 can characterize the variation of flux of atomic hydrogen beam entering the storage bulb ΔW_s and there is an optimal value $V_0 = 2.2$ V which can maximize ΔW_s. Considering the temperature characteristic of thermistor R_1,the reasons are (1) when $V_0 < 2.2$ V, the low sensitivity of flux voltage signal ΔV_1 to ΔW_s is caused by the low initial temperature of R_1, although the temperature coefficient α is great; (2) when $V_0 > 2.2$ V, the low sensitivity of flux voltage signal ΔV_1 to ΔW_s is caused by low temperature coefficient α, although the initial temperature of R_1 is great. Only when $V_0 = 2.2$ V, both the initial temperature of R_1 and the temperature coefficient α are optimal and the flux voltage signal ΔV_1 is the most sensitive to the variation of flux of atomic hydrogen beam ΔW_s.

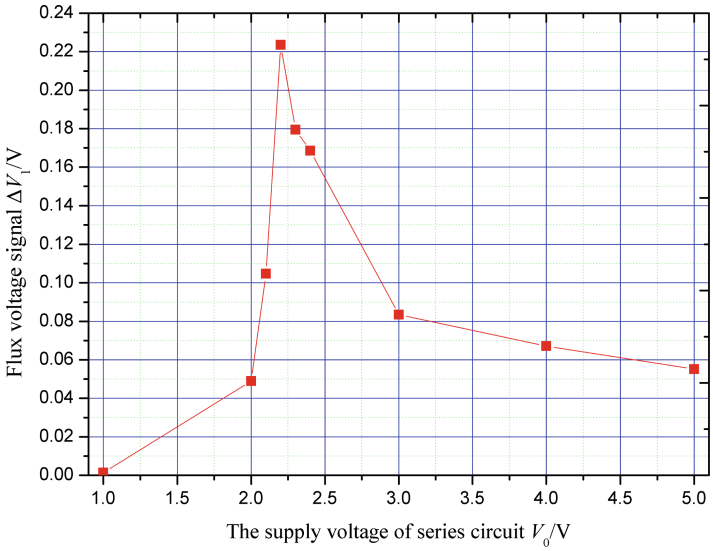
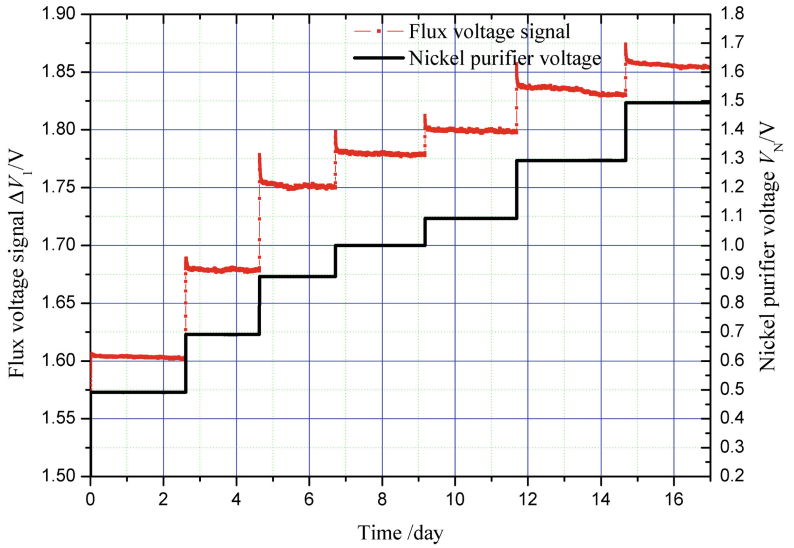


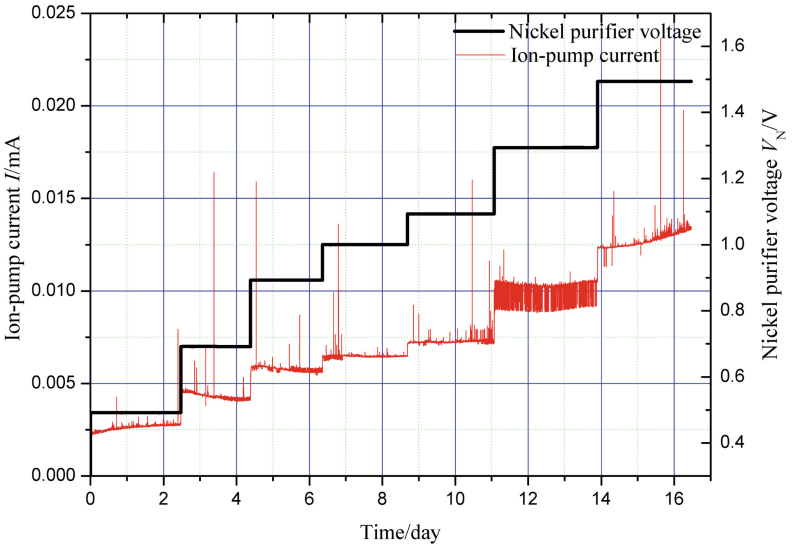
Fig. 4. The dependence of flux voltage signal ΔV_1 on the supply voltage of series circuit V_0 when the Nickel purifier voltage V_N is increased from 0.5 V to 1 V

3.2 The Flux Voltage Signal ΔV_1

The Nickel purifier voltage V_N is set as 0.5 V, 0.7 V, 0.9 V, 1.0 V, 1.1 V, 1.3 V and 1.5 V, respectively. The flux voltage signal ΔV_1 and the ion-pump current I are measured at the same time and the sampling interval is 16 s. Figure 5(a) and (b) are the respective long-term dependence of flux voltage signal ΔV_1 and Ion-pump current I on Nickel purifier voltage V_N. In Fig. 5(a), the higher the Nickel purifier voltage V_N and the greater ion-pump current I, but there are obvious eurosis noises when the Nickel purifier voltage V_N is constant. In Fig. 5(b), the flux voltage signal will increase and then tend into balance without any eurosis noise when the Nickel purifier voltage V_N is



(a)



(b)

Fig. 5. The respective long-term dependence of flux voltage signal ΔV_1 and Ion-pump current I on Nickel purifier voltage V_N, (a) flux voltage signal ΔV_1, (b) Ion-pump current I

increased. It is speculated that the overcastting phenomenon is caused by that the input flux of atomic hydrogen beam from Nickel purifier W_1 is too great compared with output flux of atomic hydrogen beam absorbed by the ion-pump W_2.

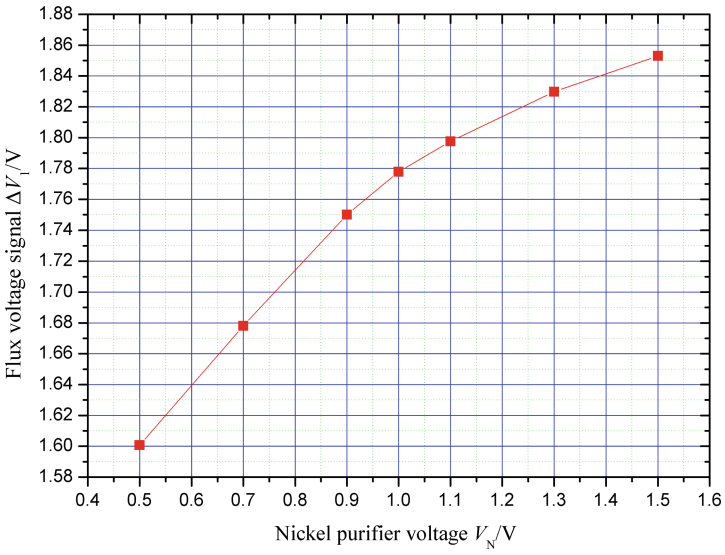


Fig. 6. The dependence of flux voltage signal ΔV_1 on the Nickel purifier voltage V_N

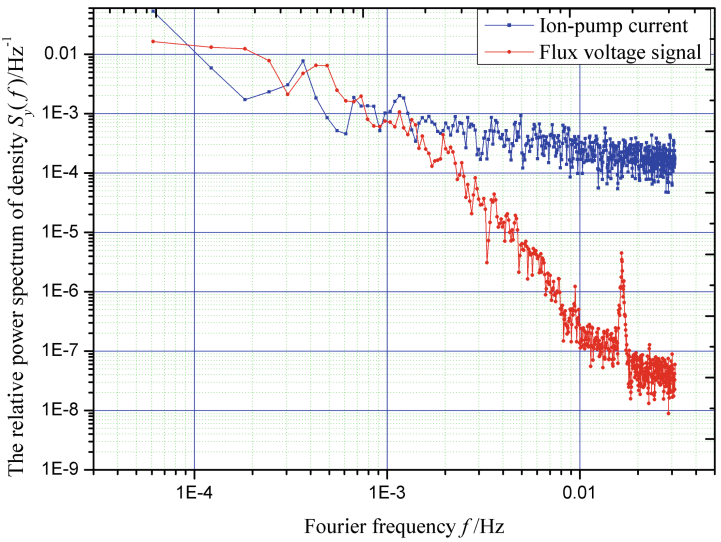


Fig. 7. The relative power spectrum of density $S_{y1}(f)$ of flux voltage signal ΔV_1 and ion-pump current I

Figure 6 shows the dependence of flux voltage signal ΔV_1 on the Nickel purifier voltage V_N and the higher the Nickel purifier voltage V_N and the balanced flux voltage signal ΔV_1. At the present V_N = 2.2 V applied, the balanced flux voltage signal ΔV_1 is linear to the Nickel purifier voltage V_N and after fitting we obtain that

$$V_1 = 0.237V_N + 1.537. \tag{8}$$

The variation of flux voltage signal ΔV_1 is 23.7 mV, when the Nickel purifier voltage V_N changes 0.1 V.

Figure 7 shows the relative power spectrum of density $S_{y1}(f)$ of flux voltage signal ΔV_1 and ion-pump current I, and they are written as $S_{y1}(f)$ and $S_{y2}(f)$ respectively. It is obtained that (1) when the frequency $f > 0.01$ Hz (integrating time $\tau < 100$ s), the relative power spectrum of density $S_{y1}(f)$ of the flux voltage signal ΔV_1 is about −7dBc/Hz but the relative power spectrum of density $S_{y2}(f)$ of the ion-pump current I is about −4dBc/Hz, which means $S_{y1}(f)$ is about 3 orders of magnitude lower than $S_{y2}(f)$; (2) when 0.001 Hz $< f <$ 0.01 Hz(100 s $< \tau <$ 1000 s), $S_{y1}(f)$ is closed to $S_{y2}(f)$ from $f = 0.001$ Hz until $f = 0.01$ Hz; (3) when $f < 0.001$ Hz($\tau > 1000$ s), $S_{y1}(f)$ and $S_{y2}(f)$ are almost the same. This analysis result shows the detection noise of this new apparatus is lower than the ion-pump current for integrating time $\tau < 1000$ s and even about 3 orders of magnitude lower than the ion-pump current for integrating time $\tau < 100$ s. So the flux voltage signal ΔV_1 measured by this new apparatus can characterize the flux of atomic hydrogen beam entering the storage bulb W_s more precisely than the ion-pump current.

4 Conclusion

The new apparatus with constant voltage to measure the flux of atomic hydrogen beam in hydrogen maser is developed and verified in this paper. The flux voltage signal on the measurement thermistor is correlated positively to the pressure in the physical package which is proportional to the flux of atomic hydrogen beam entering the storage bulb, so the flux voltage signal on the measurement thermistor can characterize the flux of atomic hydrogen beam entering the storage bulb. Firstly, according to the sensitivity of flux voltage signal to the variation of flux of atomic hydrogen beam entering the storage bulb mainly, the supply voltage of series circuit is determined. Secondly, the positive correlation between the flux voltage signal and Nickel purifier are measured at the optimal working condition. The variation of Nickel purifier voltage 0.1 V can change the flux voltage signal about 23.7 mV when the Nickel purifier voltage is 1 V. At last, the relative power spectrum of density of the flux voltage signal and ion-pump current are compared at the Fourier frequency domain, and it is obtained that the detection noise of this new apparatus is lower than the ion-pump current for integrating time $\tau < 1000$ s and about 3 orders of magnitude lower than the ion-pump current for integrating time $\tau < 100$ s. So the flux voltage signal measured by this new apparatus can characterize the flux of atomic hydrogen beam entering the storage bulb more precisely than the ion-pump current, which can be the error to stabilize the flux of atomic hydrogen beam entering the storage bulb and to fulfill the requirement of long term frequency stability of Hydrogen Maser.

References

1. Golden HM, Kleppner D, Ramsey F (1960) Atomic hydrogen maser. Phys Rev Lett 5:361–365
2. Kleppner D, Goldenberg HM, Ramsey NF (1962) Theory of the hydrogen maser. Phys Rev 126:603–615
3. Kleppner D, Berg HC, Crampton SB et al (1965) Hydrogen-maser principles and techniques. Phys Rev A 138:972–983
4. Crampton SB, Wang HTM (1975) Duration of hydrogen-atom spin-exchange collisions. Phys Rev A 12:1305
5. Koelman JMVA, Crampton SB, Stoof HTC et al (1988) Spin-exchange frequency shifts in cryogenic and room-temperature hydrogen masers. Phys Rev 38:353–3547
6. Walsworth RL, Silvera IF, Mattison EM et al (1992) Measurement of a hyperfine-induced spin-exchange frequency shift in atomic hydrogen. Phys Rev A 46:2495–2512
7. Riehle F (2006) Frequency standards basics and applications. Wiley-VCH, Weinheim, pp 230–234
8. Chen Q, Zhang JH, Zou F et al (2008) Analysis of the characteristic of the nickel purifier in hydrogen maser. In: 2008 national frequency control technology conference, pp 155–157
9. Mcleod LS (2008) Hydrogen permeation through microfabricated palladium-silver alloy membranes. Georgia Institute of Technology, Georgia
10. Gan Y (2005) The development of the new wide range vacuum gauge. Huazhong University of Science and Technology

Analysis of BDS-2+BDS-3 Combination Real-Time Time Transfer Based on iGMAS Station

Peipei Dai[1,2,4], Xuhai Yang[1,2,3(✉)], Weijin Qin[1,2,4], Ran Wang[1,2,4], and Zhe Zhang[1,2,4]

[1] National Time Service Center, Chinese Academy of Sciences, Xi'an, China
yyang@ntsc.ac.cn
[2] Key Laboratory of Precise Positioning and Timing Technology, Chinese Academy of Sciences, Xi'an 710600, China
[3] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China
[4] University of Chinese Academy of Sciences, Beijing 100049, China

Abstract. Since December 27, 2012, BDS-2 has provided navigation, positioning and timing (PNT) services in the Asia Pacific region. From November 2017 to the end of August 2018, the BeiDou Global System (BDS-3) has successfully launched 12 satellites. With the development of BDS, time transfer based on BDS has become a hot topic. In this work, 3 tracking stations, BRCH (PTB), XIA5 (NTSC) and WUH1, of the international GNSS Monitoring and Assessment System (iGMAS), 31 consecutive days of observation data, are selected. The real-time time transfer based on BDS-2, BDS-3, BDS-2+BDS-3 three schemes are preliminarily analyzed. After that, compared with the results based on GPS precise point positioning (PPP) method, the preliminary results shown that the standard deviation (STD) value of BRCH-XIA5 link based on BDS-2 is 25.1 ns, and the STD value of BDS-2+BDS-3 combination is 20.5 ns, which is increased by 18.33% in the case of unfixed station coordinates. The STD value of WUH1-XIA5 link with BDS-2, BDS-2+BDS-3 combination is less than 7 ns, and the improvement degree of BDS-2+BDS-3 is relatively small, which is about 1.47%. For frequency stability, the stability of the BRCH-XIA5 link BDS-2+BDS-3 combination is better than that of BDS-2, while WUH1-XIA5 link show a similar level with two schemes. In the case of fixed station coordinates, compared with the STD values of BRCH-XIA5 link with BDS-2 and BDS-3, BDS-2+BDS-3 combination is increased by 23.08% and 45.01% respectively. The values of the WUH1-XIA5 link with BDS-2 and BDS-2+BDS-3 combinations are all less than 1.5 ns. The STD of BDS-2+BDS-3 is reduced approximately about 7.14%, as compared to BDS-2; the STD value of BDS-3 scheme is 3.2 ns, which is slightly larger than BDS-2 and BDS-2+BDS-3, due to the fact that BDS-3 satellites are less observed at this stage. For frequency stability, the stability of two link with BDS-2 and BDS-2+BDS-3 combination is basically equal, while BDS-3 is slightly worse. In addition, and the stability of WUH1-XIA5 link clock offset is better than that of BRCH-XIA5 link.

Keywords: BDS-3 · BDS-2 · iGMAS · Time transfer · Real-time

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 496–508, 2019.
https://doi.org/10.1007/978-981-13-7751-8_48

1 Introduction

With the rapid development of Global Navigation Satellite Systems (GNSS), Test and Assessment Research Center of China Satellite Navigation Office has launched the international GNSS Monitoring and Assessment System (iGMAS) to monitor and evaluate the performance and operational status of the Chinese BeiDou Navigation Satellite System (BDS) [1, 2]. BDS-3 is expected to have three GEO, three IGSO and 24 MEO satellites to provide PNT services to users worldwide by 2020 [3]. Until the end of August 2018, 12 BDS-3satellites had been launched successfully. The old signal B1I/B3I and new signal B1C/B2a/B2b had been broadcasted. With the development of BDS system, the time transfer based on BDS has become a hot topic in BDS satellite application.

Since December 2012 BDS-2 has been provided PNT services in the Asia Pacific region [4]. Yang et al. [5] studied the time transfer ability of BDS-1 through experiments, obtained 6.5 ns timing accuracy. By using the BDS carrier phase (CP) time transfer method, Zhang et al. [6] obtained that the accuracy of the BDS CP method and GPS CP method can be reached an equivalent accuracy in the nanosecond magnitude. Zhang et al. [7] used the time transfer technology methods of common-view (CV) and CP to obtain that the accuracy of BDS CV time transfer can reach within 6 ns and the accuracy of BDS CP time transfer can reach nanosecond.

Currently, the CV method and the All-in-view (AV) method are widely used for remote time comparison. CV method has a time interval of 16 min, of which 13 min for data acquisition, 2 min for data processing, 1 min waiting for the next common viewing moment, so the method has a 16-min lag [8]. The AV method allows a direct comparison of any two stations about the system time or the International GNSS Service Time (IGST) in the case of using the precise products. The depend of precise products is a drawback of AV technology, because the update period of the precise products is about two weeks to be publicly available [9]. The latency period of the AV method is the update period of the precise products. Therefore, CV and AV methods do not satisfy the real-time application. Hence, the author used the single point positioning (SPP) mode to further investigate real-time time transfer. However, real-time time transfer based on BDS-2 and BDS-3 combination have not been investigated. Hence, based on the observation data of BRCH (PTB), XIA5 (NTSC) and WUH1 stations provided by iGMAS, the real-time time transfer performance based on BDS-2 and BDS-3 combination is evaluated at the present stage. Due to the relatively small number of satellites observed by BDS-3 satellites at the same time, the dual-frequency time transfer accuracy of BDS-2 and BDS-3 B1I/B3I was mainly analysed in this paper under the two conditions, including fixed station coordinates and unfixed station coordinates. Finally, the results were compared with the results of GPS precise point positioning (PPP).

2 Mathematical Models

The pseudorange observation equation is expressed as [10–12]:

$$P = \rho_0 + c(\delta t_{rcv} - \delta t_{sv}) + D_{trop} + \frac{I_{on}}{f^2} \tag{1}$$

where δt_{sv}, δt_{rcv} represent the satellite and the receiver clock offset, respectively; D_{trop} is the tropospheric delay; $\frac{I_{on}}{f^2}$ is the ionospheric delay corresponding to frequency; c refer to the speed of light; ρ_0 denotes the geometric distance from satellite to observatory [13].

The pseudorange observation error equation can be written as:

$$V = \begin{bmatrix} l_1 & m_1 & n_1 & -1 \\ l_2 & m_2 & n_2 & -1 \\ & \cdots & \cdots & \\ l_n & m_n & n_n & -1 \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \\ \delta_t \end{bmatrix} - L \tag{2}$$

V is the residual; L is the pseudorange observation [14]. There is no significant systematic bias between the B1I/B3I frequency of BDS-3 and BDS-2 [15]. Hence, there is only one clock offset parameter in the equation.

In this work, the dual-frequency ionosphere-free combination is used, which expressed as follows:

$$P_{IF} = \frac{1}{f_1^2 - f_3^2}(f_1^2 P_1 - f_3^2 P_3) \tag{3}$$

where f_1 and f_3 are the signal frequency of B1I and B3I, P_1 and P_3 are pseudorange observations to frequency.

Combined with Eqs. (1) and (3):

$$\frac{1}{f_1^2 - f_3^2}(f_1^2 P_1 - f_3^2 P_3) = \rho_0 + c\delta t_{rcv} - c\frac{f_1^2(\delta t_{sv})_{B1I} - f_3^2(\delta t_{sv})_{B3I}}{f_1^2 - f_3^2} + D_{trop} + \varepsilon \tag{4}$$

$(\delta t_{sv})_{B1I}$ is the satellite clock offset corresponding to the B1I frequency; $(\delta t_{sv})_{B3I}$ is the satellite clock offset corresponding to the B3I frequency; ε is pesudorange noise; Importantly, that the BDS satellite's broadcast ephemeris provides a B3I reference for the clock offset [16, 17]. The combined model of B1I+B3I ionospheric-free is employed in this work, so the observations of B1I need to be modified by time group delay (TGD), which can be written as:

$$(\delta t_{sv})_{B1I} = \delta t_{sv} - T_{GD1} \tag{5}$$

δt_{sv} is the phase time offset of satellite pseudorange; T_{GD1} is the difference between the device delay of the B1I signal and the delay of the reference signal (B3I) device, its uncertainty is less than 1 ns [18].

3 Data Description and Processing Strategies

In this experiment, the observation data were selected from BRCH (PTB), XIA5 (NTSC) and WUH1 provided by iGMAS. The data set is from day of year (DOY) 220 to 250, 2018. The sampling rate is 30 s. All iGMAS stations (BRCH, XIA5 and WUH1) can track all BDS-3 open signals. The BDS-3 old signals B1I/B3I broadcast ephemeris is provided by Multi-GNSS experiment (MGEX), and the data processing strategy is detailed in Table 1. All station information and external atomic clocks are shown in Table 2.

Table 1. Data processing strategy

Project	Model and parameter estimation
Positioning model	Standard single point positioning
Ionospheric delay	Ionosphere-free
Tropospheric delay	Saastamoinen
Station coordinates	PPP (Fixed station coordinates)
Signal frequency	B1I (1561.098 MHz)/B3I (1268.52 MHz)
Elevation cutoff angle	10°
Parameter estimation	Least squares method
Sampling interval	30 s

Table 2. Station information

Station	Receiver	Antenna	Note
BRCH	CETC-54-GMR-4016	NOV750.R4	PTB
XIA5	CETC-54-GMR-4011	TRM59900.00	NTSC
WUH1	CETC-54-GMR-4016	LEIAR25.R4	–

4 Validation and Analysis

In order to analyze the time transfer performance of BDS-2, BDS-3 and BDS-2+BDS-3, in this work, two links are designed with XIA5 (NTSC) as the central node: (1) BRCH (PTB)-XIA5 (NTSC) link, BRCH station located at the Physikalisch-Technische Bundesanstalt (PTB), XIA5 station is located National Time Service Center Punctuality Laboratory, Chinese Academy of Sciences Center; (2) WUH1-XIA5 link, WUH1 station is in Wuhan, China. Since BDS-3 currently has relatively few satellites, the single BDS-3 experiment is only conducted with fixed coordinates. Analysis using the following two processing schemes.

Scheme 1: In the case of unfixed station coordinates, the solution of receiver clock offset based on single station BDS-2 and BDS-2+BDS-3 combination is carried out on two stations respectively. Then the difference of receiver clock offset of two stations is calculated.

Scheme 2: In the case of fixed station coordinates, the solution of receiver clock offset based on single station BDS-2, BDS-3 and BDS-2+BDS-3 combination is carried out on two stations respectively. After that, comparing the receiver clock offset of the two stations is presented.

It should be explained that the experiment used the GPS PPP results based on the IGS final products are regarded as a reference to evaluate the real-time time transfer accuracy of the BDS-2, BDS-3 and BDS-2+BDS-3 combination.

4.1 Unfixed Station Coordinates

Figures 1 and 2 are the clock difference time series between solutions of BDS-2 and BDS-2+BDS-3 and GPS PPP without fixed station coordinates. The left side of Fig. 1 shows the clock difference time series between solutions of BDS-2 and GPS PPP on BRCH-XIA5 link. The right side shows the clock difference time series between solutions of BDS-2+BDS-3 combination and GPS PPP on BRCH-XIA5 link. By comparing the left and right two figures, it can be concluded that the noise of the clock difference time series of the BRCH-XIA5 link is obviously smaller after adding the

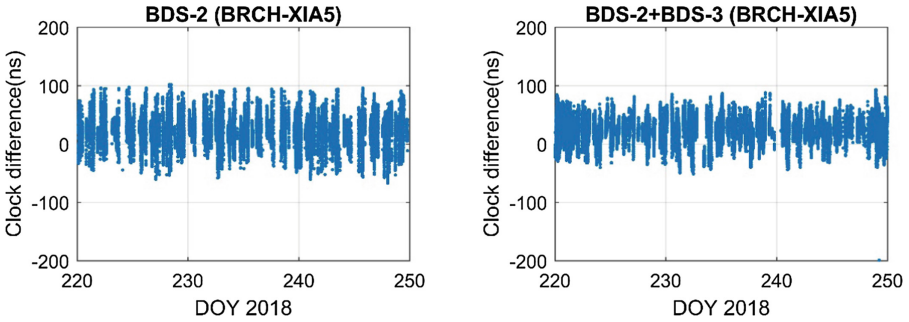


Fig. 1. Clock difference time series between the solutions of BDS-2 and BDS-2+BDS-3 and GPS PPP on BRCH-XIA5

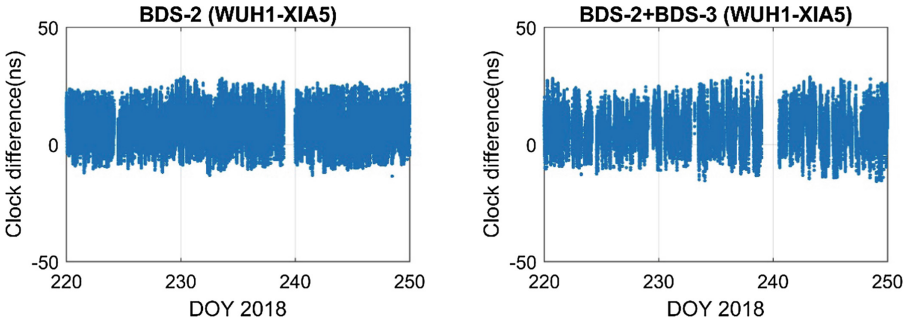


Fig. 2. Clock difference time series between the solutions of BDS-2 and BDS-2+BDS-3 and GPS PPP on WUH1-XIA5

BDS-3 satellites. The reason of these phenomena is that BRCH is located in the European region, so the number of observed BDS-2 satellites is relatively small. In addition, some BDS-2 satellites have high degree of noise at lower elevation angel in the European region. The number of satellites is increased significantly, when BDS-3 joins, which improves the equation structure. On the other hand, it can be seen from Fig. 1 that the BDS-2+BDS-3 time series is more intensive than the BDS-2. The main reason is that the number of BDS-2 satellites is less than 4 at some time, the equation is not solvable, while after adding BDS-3, the number of satellites increase and the equation can be solved. The above results can be further demonstrated in Fig. 3, it shows the number of observable satellites available in the BRCH and XIA5 stations, in the schemes of BDS-2, BDS-3, and BDS-2+BDS-3. In Fig. 2, the left side presents clock difference time series between the solutions of BDS-2 and GPS PPP on WUH1-XIA5, and the right side presents the clock difference time series between the solutions of BDS-2+BDS-3 and GPS PPP on WUH1-XIA5. It should be explained that, the data discontinuity because of the lack of observational data at DOY 239, 2018 in the figure. Combined with Figs. 1 and 2, we obtain the conclusions that the noise of BDS-2 is greater than that of BDS-2+BDS-3. After adding BDS-3 satellites, the noise of two links results are all decreased, moreover, the noise of BRCH-XIA5 link is decreased more obviously. The results of WUH1-XIA5 link are less noisy than BRCH-XIA5 link both BDS-2 and BDS-2+BDS-3 combination. The main reason is the fact that WUH1 and XIA5 located at China, therefor more BDS satellites that can be observed two station.

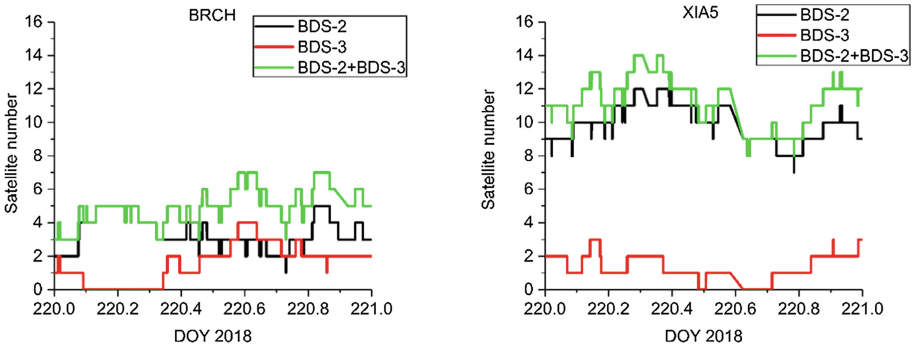


Fig. 3. Observable satellite numbers on BRCH, XIA5 at DOY 220, 2018

Table 3. The STD and mean values of the clock difference between scheme1 and GPS PPP

	STD			Mean	
	BDS-2	BDS-2+BDS-3	Improvement (%)	BDS-2	BDS-2+BDS-3
BRCH-XIA5	25.1	20.5	18.33%	23.4	23.0
WUH1-XIA5	6.8	6.7	1.47%	7.2	6.3

In order to further evaluate the time transfer performance of the BDS-2 and BDS-2 +BDS-3 combination, Table 3 shows the Standard deviation (STD) and mean values of the comparison results between scheme 1 and GPS PPP. The STD value of BRCH-XIA5 link based on BDS-2 in Table 2 is 25.1 ns, and the STD value of BDS-2+BDS-3 combination is 20.5 ns, which is increased by 18.33% compared with BDS-2. The value of WUH1-XIA5 link based on BDS-2 and BDS-2+BDS-3 combination are less than 7 ns, and the improvement degree of BDS-2+BDS-3 is relatively small for BDS-2, about 1.47%. Combined with Figs. 1 and 2, it can be seen that the BRCH-XIA5 link noise is also much larger than the WUH1-XIA5 link, which can be explained that the BRCH station is located in Europe, and there are fewer observable BDS satellites. In addition, the mean difference of different links has distinction, due to the different hardware delay of different stations. Meanwhile, we look forward to achieving better results after the construction of the BDS-3 Global System is completed.

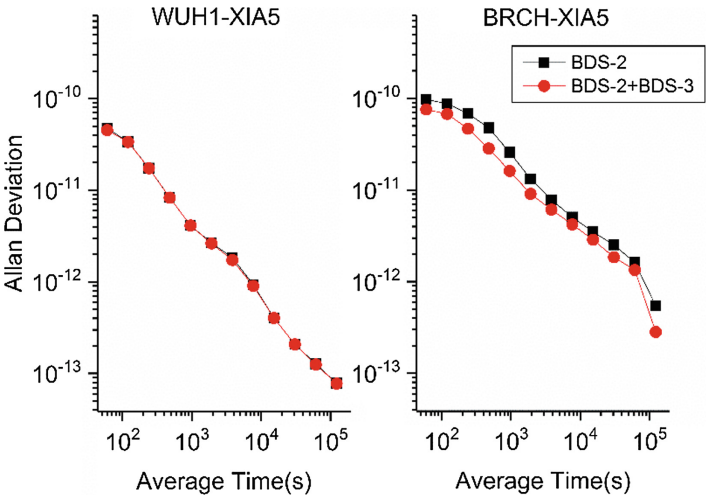


Fig. 4. Comparison of frequency stability between BDS-2 and BDS-2+BDS-3

Although we have presented the results compared to GPS PPP at previous study, the external atomic clocks of different link are different. Hence, they are not sufficient to evaluate BDS-2+BDS-3 time transfer performance because an absolute reference is lack [14]. Therefore, we calculated the Allan deviation of all the schemes to further evaluate the results of our experiment [19]. Figure 4 shows the Allan deviation of the two links with BDS-2 and BDS-2+BDS-3 combination in the case of unfixed station coordinates. The stability of BRCH-XIA5 link based on BDS-2+BDS-3 combination is better than BDS-2, and the stability of WUH1-XIA5 link with two schemes are basically consistent. Figure 5 display the improvement percentage of BDS-2+BDS-3 combination compared to the BDS-2. We can obtain the conclusion that the Allan deviation of the BRCH-XIA5 link based on BDS-2+BDS-3 combination can be increased by 50% and the WUH1-XIA5 link can increased to 5% which compared to

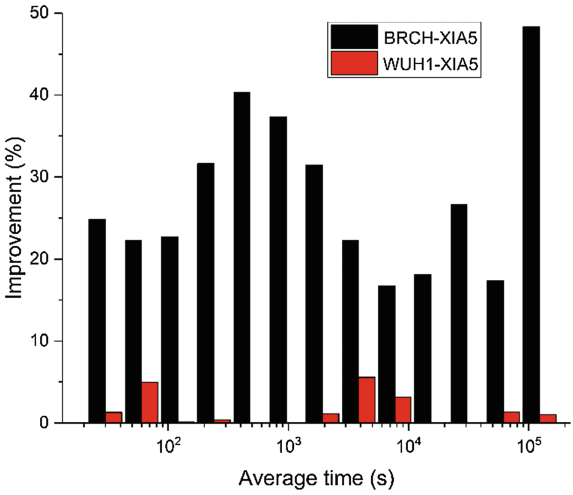


Fig. 5. Improvement in frequency stability of BDS-2+BDS-3 compared to BDS-2

the BDS-2. Therefore, it can be further proved that it can obviously improve the accuracy of BDS time transfer when the number of BDS-3 satellites has increased in Europe region. However, in the Asia Pacific region, the improvement of accuracy after adding BDS-3 satellites is relatively small due to better observation conditions.

4.2 Fixed Station Coordinates

In time community, the receiver station coordinates are usually known, so the station coordinates are fixed and better results are expected, especially in the European region. Clock difference time series between three links based on BDS-2 and GPS PPP in the case of fixed station coordinates are plot in Figs. 6, 7, and 8. By comparing the left and right two figures in Fig. 6, we can see that the noise of the clock difference time series in the BRCH-XIA5 is obviously decreased after adding the BDS-3 satellites. At the same time, compared with Fig. 1, fixed station coordinates can obviously improve the accuracy of time transfer. The left side of Fig. 7 shows the clock difference time series between BDS-2 and GPS PPP on WUH1-XIA5, and the right side shows the clock difference time series between BDS-2+BDS-3 and GPS PPP on WUH1-XIA5, and the reason for data interruption are is the same as before. The noise of time link results based on BDS-2 is larger than that of BDS-2+BDS-3 combination. With increased BDS-3 satellites, that the noise of two links is decreased, while the noise reduction of BRCH-XIA5 link is more obvious. In Fig. 8, the left side presents clock difference time series between BDS-3 and GPS PPP on BRCH-XIA5, and the left side presents clock difference time series between BDS-3 and GPS PPP on WUH1-XIA5. Compared with the results of BDS-2 and BDS-2+BDS-3 combination results, the noise of BDS-3 is relatively large, the main reason can be explained from Fig. 3, because BDS-3 currently has a small number of satellites that can be observed in one epoch.

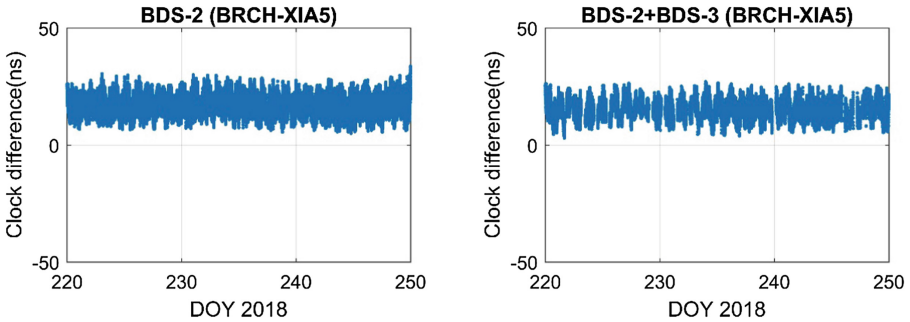


Fig. 6. Clock difference time series between BDS-2 and BDS-2+BDS-3 and GPS PPP on BRCH-XIA5

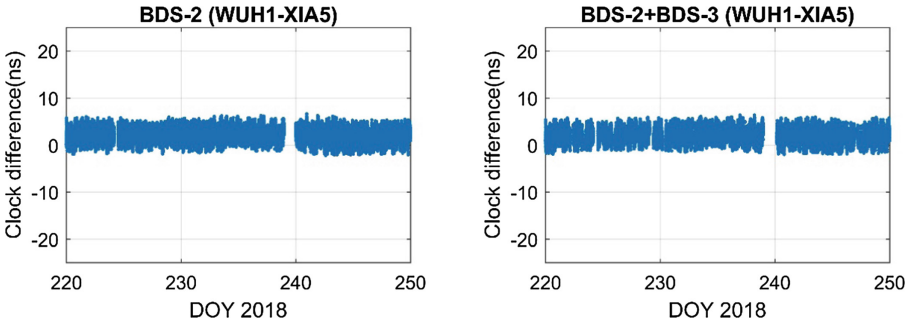


Fig. 7. Clock difference time series between BDS-2 and BDS-2+BDS-3 and GPS PPP on WUH1-XIA5

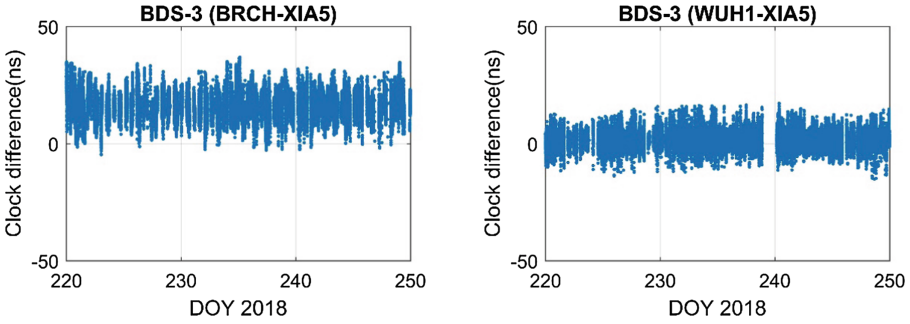


Fig. 8. Clock difference time series between BDS-3 and GPS PPP on BRCH-XIA5 and WUH1-XIA5

Table 4 shows the STD and mean values of the clock difference between scheme 2 and GPS PPP. In Table 4, the STD values of BRCH-XIA5 link based on BDS-2 and BDS-3 are 6.5 ns and 9.1 ns, respectively, while BDS-2+BDS-3 combination is 5.0 ns, which is increased by 23.08% and 45.01%, respectively, compared to BDS-2 and BDS-3. The STD of the WUH1-XIA5 link based on BDS-2 and BDS-2+BDS-3 combination is less than 1.5 ns, and the reduction of STD based on BDS-2+BDS-3 combination is about 7.14% compared with BDS-2; the STD value of WUH1-XIA5 link based on BDS-3 is 3.2 ns, slightly larger than BDS-2 and BDS-2+BDS-3. Combined with Figs. 5, 6 and 7, the results show that the noise of BRCH-XIA5 link is also larger than the WUH1-XIA5 link. That may be explained by the fact that BRCH station is located in Europe, there are less BDS satellites can be observed and the satellite elevation angle is lower.

Table 4. The STD and mean values of the clock difference between scheme2 and GPS PPP

	STD			Improvement		Mean		
	BDS2	BDS2 +BDS3	BDS3	BDS2 (BDS2 +BDS3)	BDS3 (BDS2 +BDS3)	BDS2	BDS2 +BDS3	BDS3
BRCH-XIA5	6.5	5.0	9.1	23.08%	45.01%	16.8	15.1	15.9
WUH1-XIA5	1.4	1.3	3.2	7.14%	59.38%	2.0	1.9	1.3

As we mentioned before, there is a lack of an absolute reference, so we calculate the Allan variance after the fixed station coordinates. Figure 9 shows the Allan deviation of two links BDS-2, BDS-3, and BDS-2+BDS-3 in the case of fixed station coordinates. The stability of two links shows that BDS-2 and BDS-2+BDS-3 combination is basically equal, while the solutions of BDS-3 is slightly worse. In addition, the stability of WUH1-XIA5 performs better than that BRCH-XIA5. Figure 10 present that percentage improvement in frequency stability of BDS-2+BDS-3 compared to BDS-2 and BDS-3. It can be concluded that the Allan deviation of BRCH-XIA5 link BDS-2+BDS-3 combination is better than that of BDS-2 and BDS-3 at different time intervals. The maximum improvement BDS-2+BDS-3 combination compared to BDS-2 is approximately 25%, and is about 72% compared to BDS-3. The Allan deviation of WUH1-XIA5 link based on BDS-2+BDS-3 combination is relatively small as compared to BDS-2, while the maximum improvement is about 95%, as compared to BDS-3.

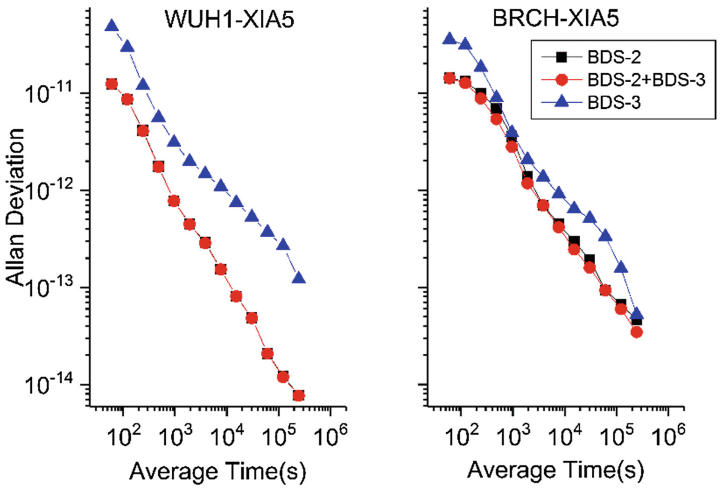


Fig. 9. The Allan deviation of two time-links based on BDS-2, BDS-3 and BDS-2+BDS-3

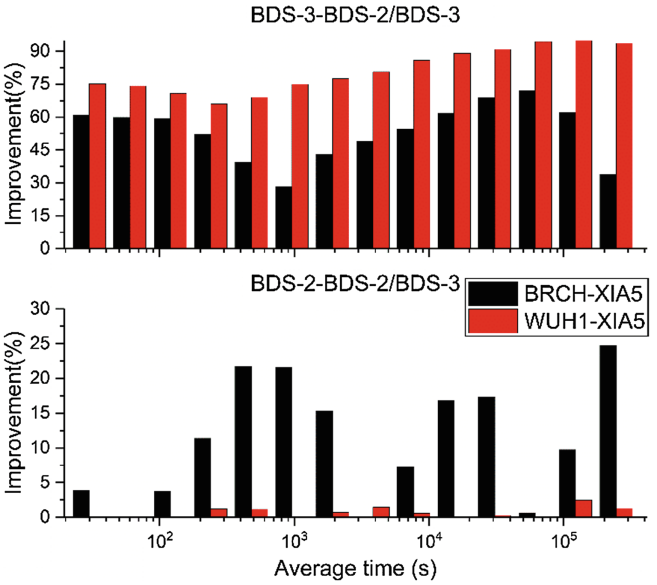


Fig. 10. Percentage improvement in frequency stability of BDS-2+BDS-3 compared to BDS-2 and BDS-3

5 Conclusion

With the rapidly development of China's BDS-3 system, 12 BDS-3 MEO satellites have been successfully launched by the end of August 2018. Furthermore, BDS-3 satellites signals have been monitored by the iGMAS to assess their performance and operational status.

This work mainly evaluates the real-time time transfer accuracy of BDS-2 and BDS-3. Two links, BRCH-XIA5 and WUH1-XIA5 from iGMAS, were selected. And 31 days observation data were collected for our test. Note that the GPS PPP results using IGS final products were regarded. Finally, the following conclusions were obtained:

(1) In the case of unfixed station coordinates, the STD of BRCH-XIA5 link based on BDS-2 is 25.1 ns, and the STD of BDS-2+BDS-3 combination is 20.5 ns, which is increased by 18.33% compared to BDS-2. The STD of WUH1-XIA5 link based on BDS-2 and BDS-2+BDS-3 combination are all less than 7 ns. The improvement of BDS-2+BDS-3 are relatively small, which is about 1.47%. For frequency stability, the stability of the BRCH-XIA5 link based on BDS-2+BDS-3 combination is better than that of BDS-2, while the stability of two schemes for WUH1-XIA5 link show similar level.

(2) In the case of fixed station coordinates, the STD of BRCH-XIA5 link BDS-2 and BDS-3 are 6.5 ns and 9.1 ns respectively, while BDS-2+BDS-3 combination is 5.0 ns, which is increased by 23.08% and 45.01% respectively for BDS-2 and BDS-3. The STD of the WUH1-XIA5 link based on BDS-2 and BDS-2+BDS-3 combination are less than 1.5 ns, the reduction of STD is about 7.14%; the WUH1-XIA5 link based on BDS-3 STD is 3.2 ns, which is slightly larger than BDS-2 and BDS-2+BDS-3, due to less BDS satellites can be observed. Besides, the frequency stability of two links based on BDS-2 and BDS-2+BDS-3 combination is basically equal, the stability of BDS-3 is slightly worse, and the stability of WUH1-XIA5 link is better than BRCH-XIA5 link.

(3) Under the two schemes, the BDS-2+BDS-3 combination time transfer perform better than that of BDS-2 and BDS-3, especially for the BRCH-XIA5 link. Its support that that BDS-3 will be able to better serve the whole world. We also expect better results when the construction of the BDS-3 is completed.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (No. 11703033). The authors gratefully acknowledge iGMAS for funding. Many thanks go to the iGMAS for providing data.

References

1. Xie X, Fang R, Geng T, Wang G, Zhao Q, Liu J (2018) Characterization of GNSS signals tracked by the iGMAS network considering recent BDS-3 satellites. Remote Sens 10 (11):1736. https://doi.org/10.3390/rs10111736
2. Jiao W, Ding Q, Li J, Lu X, Feng L, Ma J, Chen G (2011) Monitoring and assessment of GNSS open services. Sci Sinica 64(S1):S19–S29

3. Hu C, Wang Q, Wang Z, Hernández AM (2018) New-generation BeiDou (BDS-3) experimental satellite precise orbit determination with an improved cycle-slip detection and repair algorithm. Sensors 18(5):1402
4. Xie X, Geng T, Zhao Q, Liu J, Wang B (2017) Performance of BDS-3: measurement quality analysis, precise orbit and clock determination. Sensors (Basel) 17(6):1233. https://doi.org/10.3390/s17061233
5. Yang Z, Zhang J, Yang J (2009) Experimental research on the time transfer capability of compass I. Electron Measur Technol 32(08):16–19
6. Zhang R, Liu H, Shu B, Qian W (2017) Research on time transfer: based on BDS precise point positioning and accuracy comparison. J Geodesy Geodyn 37(10):1070–1073
7. Zhang P, Tu R, Gao Y, Guang W, Cai H (2017) BeiDou time transfer method and its accuracy analysis. Chin J Sci Instrum 38(11):2700–2706
8. Huang W, Defraigne P (2016) BeiDou time transfer with the standard CGGTTS. IEEE Trans Ultrason Ferroelectr Freq Control 63(7):1005–1012. https://doi.org/10.1109/TUFFC.2016.2517818
9. Lee SW, Kim J (2009) Performance evaluation of real-time carrier-phase GPS time transfer. Meas Sci Technol 20(6):065105. https://doi.org/10.1088/0957-0233/20/6/065105
10. Chen L (2017) Research on pseudo-distance single point location algorithm of BDS/GPS combination system [硕士]. Nanjing University of Science and Technology
11. Hao J, Liu Z (2016) Pseudorange single-point based on BeiDou navigation satellite system. Hydrogr Surv Charting 36(06):59–61+69
12. Li X, Zheng K, Li X, Liu G, Ge M, Wickert J, Schuh H (2018) Real-time capturing of seismic waveforms using high-rate BDS, GPS and GLONASS observations: the 2017 Mw 6.5 Jiuzhaigou earthquake in China. GPS Solutions 23(1). https://doi.org/10.1007/s10291-018-0808-9
13. Guang W, Dong S, Wu W, Zhang J, Yuan H, Zhang S (2018) Progress of BeiDou time transfer at NTSC. Metrologia 55(2):175–187. https://doi.org/10.1088/1681-7575/aaa673
14. Zhang P, Tu R, Zhang R, Gao Y, Cai H (2018) Combining GPS, BeiDou, and Galileo satellite systems for time and frequency transfer based on carrier phase observations. Remote Sens 10(2):324. https://doi.org/10.3390/rs10020324
15. Li X, Yuan Y, Zhu Y, Huang J, Wu J, Xiong Y, Zhang X, et al (2018) Precise orbit determination for BDS3 experimental satellites using iGMAS and MGEX tracking networks. J Geodesy. https://doi.org/10.1007/s00190-018-1144-0
16. Ge Y, Feng Z, Sun B, Wang S, Bo S (2017) The impact of satellite time group delay and inter-frequency differential code bias corrections on Multi-GNSS combined positioning. Sensors 17(3):602
17. Guo F, Zhang X, Wang J (2015) Timing group delay and differential code bias corrections for BeiDou positioning. J Geodesy 89(5):427–445. https://doi.org/10.1007/s00190-015-0788-2
18. SCIO (2016) China's BeiDou Navigation Satellite System by the State Council Information Office of the People's Republic of China (SCIO). http://www.beidou.gov.cn/xt/gfxz/201712/P020171221333863515306. Accessed 10 June 2017
19. Ge Y, Sun B, Qin W, Yang H, Wang S (2017) Time and frequency transfer analysis of BDS PPP based on the observations with multipath mitigation using a wave absorbing shield. In: VIII China satellite navigation conference, vol 6

The Engineering Applications of Laser Cooling-Cesium Atomic Fountain Frequency-Standard

Fengfeng Shi(✉), Xiaobin Wang, Shiyi Xu, and Xianglei Wang

Beijing Satellite Navigation Center, Beijing, China
fengfeng_shi@163.com

Abstract. As a SI time unit second realization equipment, laser cooling-cesium atomic fountain frequency-standard has high accuracy and no drift. NIM5-M is a new generation of laser cooling-cesium atomic fountain frequency-standard developed by NIM (National Institute of Metrology, China). In this paper, thorough research on the engineering applications of NIM5-M is conducted. The continuous automatic operation of NIM5-M is achieved. And a software is implemented to realize automatic evaluation of its frequency shift as well as the performance of atomic clocks and TA(BSNC), which can also detect the atomic clock frequency hopping. Firstly, the main factors affecting the continuous operation of NIM5-M are analyzed. It is found that temperature variation is the main factor affecting the uncertainty of NIM5-M, based on which the basic strategy of laser cooling-cesium atomic fountain frequency-standard continuous operation is proposed. Nine systematic frequency shifts of NIM5-M were evaluated, which indicates that the overall uncertainty of NIM5-M is 1.77×10^{-15}. NIM5-M is used to evaluate the frequency drift of the atomic clock directly. With the interference of the reference source drift avoided, the true frequency offset of the atomic clock is reflected. Similarly, the evaluation result of NIM5-M on TA(BSNC) can also be considered as a reference for TA (BSNC) steering and atomic clock frequency hopping detection.

Keywords: Time & frequency standard · Atomic clock · Cooling atomic fountain clock · Engineering application

1 Introduction

In 1967, the General Conference on Weights and Measures (CGPM) defined atomic seconds (s) as: 133Cs hyperfine structure transition radiation for 9, 192, 631, 770 cycles. In 2010, a laser cooling-cesium atomic fountain frequency-standard named NIM5 was successfully developed by NIM, which can realize the SI time unit second [1]. Firstly, the cesium atomic vapor is put into the magneto-optical trap. Under the action of cooling laser and magnetic field, the cesium atoms are "trapped" into an atomic cloud, which is then vertically ejected upwards. During its ascension and drop, it interacts with the microwave twice, thus a Ramsey signal is obtained and a standard clock signal is outputted by the servo-controlled crystal oscillator. Since the cold atom thermal motion velocity is about 8.8 cm/s, which is much lower than the cesium atomic

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 509–517, 2019.
https://doi.org/10.1007/978-981-13-7751-8_49

motion velocity of 300 m/s in the cesium atomic beam clock, it has a longer atomic flight time and the line width of the Ramsey frequency-detection signal is smaller, thus the short-term stability is improved [2]. In terms of accuracy, the cold atomic fountain frequency-standard eliminates the main error source that limits the accuracy of the cesium atomic clock, thereby increasing the uncertainty index by two orders of magnitude and without frequency drift [3].

The cesium atomic fountain frequency-standard combines the drift-free characteristics of the cesium atomic clock with less uncertainty, so the performance of time-keeping system can be significantly improved when it is applied and operated continuously. However, due to its large volume and complex structure as well as the aging and instability of the laser, it is difficult to continuously and stably operate for a long time. In this paper, the application of NIM5-M in time-keeping system is analyzed in detail. Data analysis on the application is carried out and it is shown that the continuous operation rate of NIM5-M is more than 89%, and frequency jump and phase jump of atomic clocks can be monitored in real time. And frequency drift of the atomic clock can be calculated. Meanwhile, the evaluation result of the fountain frequency-standard on TA can be used as a direct and reliable reference for frequency steering, thus the accuracy of TA can be significantly improved.

2 Cesium Fountain Continuous Automatic Operation Strategy

The precondition of using the cesium atomic fountain frequency-standard data to evaluate TA and atomic clocks is that the cesium atomic fountain frequency-standard is continuously and stably operated. In order to ensure the continuous operation of the cesium atomic fountain frequency-standard, it is necessary to control the temperature and humidity environment of the equipment room. When controlled precisely, the temperature of the equipment room can be maintained within the range of 23 ± 0.2 °C. As is shown in Fig. 1, after deducting the temperature change caused by personnel in and out, the room temperature is always maintained between 23.05 °C and 23.15 °C, and the temperature variation is no more than 0.1 °C.

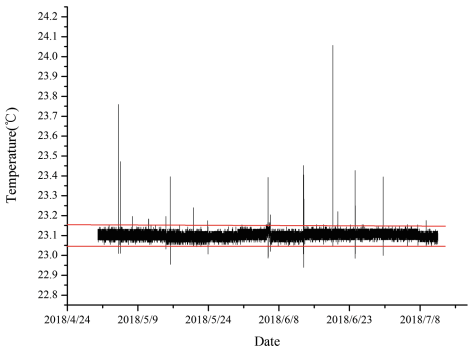


Fig. 1. Temperature of cesium atomic fountain frequency-standard room

The ac stark frequency shift caused by the ambient heat radiation from the atomic cloud on the atomic motion is calculated according to the formula given in [4]:

$$(\frac{\Delta f}{f_0})_{B.R} = -(169 \pm 4) \times 10^{-16}(T/300)^4 \quad (1)$$

The environmental temperature of the cesium atomic fountain frequency-standard room is maintained within the range of 23 °C ± 0.1 °C, and the corresponding black body radiation frequency shift is -16×10^{-15}. Differentiating formula (1):

$$\partial(\frac{\Delta f}{f_0})_{B.R}/\partial T = -(169 \pm 4) \times 10^{-16} \times 4/300(T/300)^3 \quad (2)$$

According to formula (2), the 0.1 °C temperature fluctuation corresponds to an uncertainty of 0.02×10^{-15}. This uncertainty is much smaller than the uncertainty caused by cold atom collisions. Therefore, the precise temperature control system can improve the stability of the optical system of the cesium atom fountain frequency-standard and reduce its uncertainty.

Since an optical systems is included in the cesium atomic fountain frequency-standard, the most common operational failure is the sudden frequency change of the cold collision caused by the atom number change of the cold atomic cloud, which in turn increases the uncertainty of the cesium atomic fountain frequency-standard. Therefore, it is necessary to monitor the number of cold atoms in real time. When the number of cold atoms is drastically reduced or the probability of transition is abrupt, the optical system needs to be maintained as soon as possible. However, as can be seen from Fig. 1, personnel in and out of the equipment room will cause the temperature to change by nearly 1 °C, and the corresponding uncertainty is 0.22×10^{-15}. Therefore, the influence of external factors on the operation of the cesium atomic fountain frequency-standard needs to be minimized and the personnel in and out of the equipment room needs to be controlled. In order to solve these problems, the state of the cesium atomic fountain frequency-standard is remotely monitored in another room in real time, reducing the frequency of personnel entering and leaving the cesium atomic fountain frequency-standard room, thus the factors affecting the environment of the room is minimized and the uncertainty caused by temperature fluctuation is reduced. By taking corresponding measures to ensure the continuous operation of the cesium atomic fountain frequency-standard, the number of days of its normal operation (uncertainty $\leq 5 \times 10^{-15}$) from April 25, 2018 to September 27, 2018 was 137 and the normal operation rate was 89.1%.

3 Products Cesium Fountain Frequency-Standard Frequency Shift and Uncertainty Evaluation

Although the main error source that limits the accuracy of the cesium atomic clock is eliminated in the atomic fountain frequency-standard, which is atomic thermal motion, systematic frequency shifts stills exists. In accordance with the specifications and recommendations on the evaluation of time-frequency reference uncertainty of BIPM [5], the following nine systematic frequency shifts are evaluated: 2nd order Zeeman,

Collisional shift, Microwave interferometric Switch, Distribution cavity phase, Light shift, Blackbody radiation, Gravitational red shift, Majorana, Cavity pulling, etc. According to the frequency shift and uncertainty calculation method in [6], the NIM5-M frequency deviation and uncertainty evaluation results are calculated, as shown in Table 1. It can be seen from Table 1 that the overall uncertainty of NIM5-M is 1.77×10^{-15}. By precision control of laboratory temperature, the evaluation of daily frequency shift only remaining cold atom collision frequency shift [7].

Table 1. Cesium atomic fountain frequency-standard frequency deviation and uncertainty

Physical effect	Frequency deviation (10^{-15})	Uncertainty (10^{-15})
2nd order Zeeman	74	0.2
Collisional shift	−12.5	1.3
Microwave interferometric Switch	0	0.6
Distribution cavity phase	0	1.0
Light shift	0	0.1
Blackbody radiation	−16.2	0.4
Gravitational red shift	4.3	0.1
Majorana	0	<0.1
Cavity pulling	0	<0.1
Total	49.6	1.8

4 Applications of Cesium Atomic Frequency-Standard

In order to apply the cesium fountain frequency-standard (NIM5-M) to the time keeping system, a frequency shift automatic evaluation software is developed which can automatically evaluate the frequency shift of the cesium fountain frequency-standard on a daily basis and can calculate the frequency accuracy of one hydrogen maser. Furthermore, the clock difference data between the hydrogen maser and other masers in the atomic clock assembly are extracted automatically, thus the frequency accuracy of each atomic clock in the time keeping laboratory is obtained indirectly by the cesium fountain frequency-standard. Further, based on the frequency difference between TA(BSNC) and the hydrogen maser, the accuracy evaluation of TA(BSNC) by the cesium fountain frequency-standard can be calculated. Based on the frequency evaluation data of cesium fountain frequency-standard, the calculation of atomic clock drift, the reference of the TA(BSNC) frequency steering and the detection of atomic clock frequency hopping can be realized.

4.1 The Drift of Atomic Clock

Quadratic clock difference model [8] is a common way to calculate atomic clock drift:

$$x(t) = x_0 + y_0 t + \frac{1}{2} D t^2 + \varepsilon_x(t) \tag{3}$$

In the formula, x_0, y_0, D are the initial phase difference, the initial frequency difference and the frequency drift, respectively. $\varepsilon_x(t)$ is the stochastic variation component of the time difference of the atomic frequency standard. The drift is two times the quadratic coefficient, which means that the drift rate cannot be viewed intuitively in the clock difference curve. The evaluation results of cesium fountain frequency-standard are frequency accuracy, that is, using the expression of relative frequency difference $y(t)$ to calculate the drift rate:

$$y(t) = y_0 + Dt + \varepsilon_y(t) \tag{4}$$

The drift is converted to primary coefficient, which can be used to observe the change of drift rate more intuitively. The evaluation data of hydrogen maser A and B by cesium fountain frequency-standard are shown in Fig. 2.

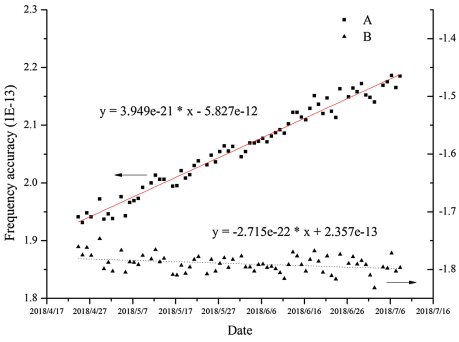


Fig. 2. The drift of hydrogen maser A and B calculated by cesium fountain frequency-standard

The accuracy evaluation results of hydrogen maser A and B are given in Fig. 2, the red curve gives the fitting result of A's accuracy, and the black dotted line gives the result of B's, so that the frequency difference model of A can be obtained as:

$$y = 3.949 \times 10^{-21}x - 5.827 \times 10^{-12} \tag{5}$$

where B's unit is seconds. Therefore, the drift of hydrogen maser A, assessed by cesium fountain frequency-standard, is 3.41E−16.

Similarly, the frequency difference model of hydrogen maser B is:

$$y = -2.715 \times 10^{-22}x + 2.357 \times 10^{-13} \tag{6}$$

which means the drift of hydrogen maser B is −2.35E−17.

When using traditional time difference fitting method to calculate the drift, it is necessary to fit the time difference of the two atomic clocks, and it is assumed that one of them has a very small drift or no drift at all. The drift of hydrogen maser A is calculated, using hydrogen maser B as a reference, and the results are shown in Fig. 3.

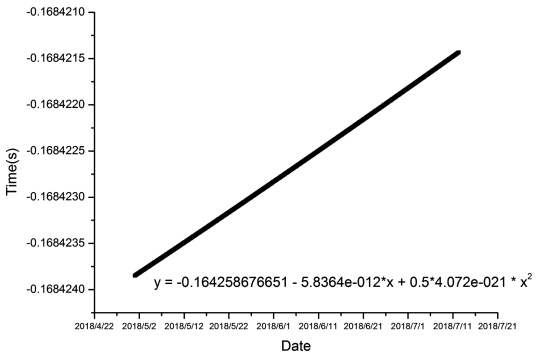


Fig. 3. Time difference between hydrogen maser A and B and its fitting result

The black curve in Fig. 3 is time difference data, and the frequency drift of hydrogen maser A relative to hydrogen maser B is 3.69E−16 using quadratic fitting method.

Table 2. Frequency drift evaluation results

A - NIM5 M	B - NIM5 M	A - B
3.4E−16	−2.4E−17	3.7E−16

The results of frequency drift using cesium fountain frequency-standard and frequency drift using traditional time difference data are given in Table 2. As can be seen from the results in the table, for the same maser, the drift using the traditional time difference data calculation is larger. The main reason is that the data of time difference contains the drift of the reference maser. Therefore, the frequency drift calculated by traditional time difference data is actually the overall drift of the two atomic clocks, rather than the drift of a single atomic clock. The drift of cesium fountain can be ignored, so the drift of atomic clock evaluated by cesium fountain frequency-standard is more accurate, which provides a direct and reliable method for the calculation of TA (BSNC).

4.2 Performance Evaluation of TA(BSNC)

International time keeping laboratories use atomic clock assembly of hydrogen masers and cesium atomic clocks to calculate TA(BSNC), taking advantages of the short-term stability of hydrogen masers and long-term stability of cesium atomic clocks. By using TA(BSNC), the long-term stability of the time keeping system is improved but its accuracy is not. In order to improve the accuracy of system time, international time keeping laboratories often use atomic frequency standard devices, such as the United States rubidium atomic frequency standard, Russia's cesium atomic frequency standard, China's cesium atomic frequency standard and so on.

The output frequency of atomic frequency standard is used to correct the time and frequency of the system, so as to improve the accuracy of the system time, that is, the system time is evaluated continuously by atomic frequency standard, and then the system time is steered by using the evaluation results [9]. Figure 4 shows the accuracy results of system time steered according to NIM5-M.

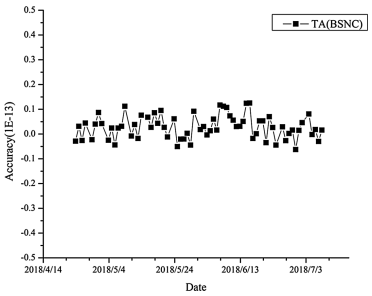


Fig. 4. The frequency accuracy of TA(BSNC)

The accuracy result of 75 days of system time is given in Fig. 4, which indicates that the accuracy is better than 1E−14, the average value is 2.8E−15, and the accuracy of system time steered is higher than that it's designed.

In the absence of external time comparison, the cesium fountain frequency-standard's evaluation on TA(BSNC) is the only reference for frequency steering. Figure 5 shows the frequency evaluation results of TA(BSNC) by the cesium fountain frequency-standard (a) and the time difference curve between GPST and the TA (BSNC) (b).

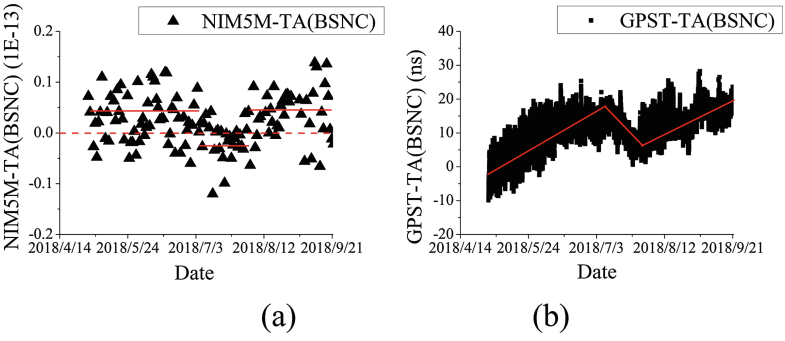


Fig. 5. The frequency comparison of TA(BSNC) with NIM5-M (a) and GPST (b)

The black triangle point in Fig. 5(a) represents the distribution of the frequency difference between the cesium atom fountain clock and the synthetic atom. Where the red dotted line represents Time frequency difference of the cesium atom fountain clock

and the synthetic atom is zero; the place above 0-wire represents the cesium atom fountain clock frequency is greater than the synthetic atom, and the place under 0-wire represents the cesium Atom fountain clock frequency is smaller than the synthetic atom. The red solid line represents the average of the frequency difference of the cesium atom fountain clock and the synthetic atom during corresponding time periods. In Fig. 5(b), the black square point represents the time difference between the GPST and the synthetic atom, and the red solid line represents a fitting of the data over the corresponding time period, and the slope of which is the frequency difference between the GPST and the synthetic atom. As can be seen in Fig. 5(a), from May 1, 2018 to July 8 and from July 31 to September 21, the results of NIM5 M-TA (BSNC) were generally positive, indicating that the TA (BSNC) frequency was relative small, so that the time difference of GPST should gradually increase, and Fig. 5(b) validates this phenomenon. Similarly, from July 9 to July 30, the results of NIM5 M-TA (BSNC) were generally negative, indicating that the TA (BSNC) frequency was relative large, so the time difference of GPST should gradually reduce, and Fig. 5(b) validates this phenomenon. Therefore, the evaluation results of cesium atom fountain clock on synthetic atoms can be used as the basis for the control of atoms.

4.3 Detection of Frequency Hopping

The frequency jump of atomic clock will worsen the index of synthetic atoms, and the elimination of the jumping atomic clock in the clock difference measurement data of atomic clock group is the work that must be done before the calculation of synthetic atomic clock, and the frequency hopping of atomic clock can be detected directly by using cesium atomic fountain device, which is more intuitive and simple.

The frequency of hydrogen atom clock jumped −6E−15 on June 1, 2018, and the frequency change is the jump step when shown on the jet lag curve, which means the slope between the front and back points occurred sudden change; the device of the atomic clock frequency by the cesium atomic fountain is directly manifested as a mutation of the data, which is more intuitive and easier to be understood. The black curve in Fig. 6 gives the time difference curve, and the frequency jump of the hydrogen atom clock can be seen on June 1, 2018 from the time difference curve, and the red curve in Fig. 6 is the assessment results of the frequency accuracy of the hydrogen

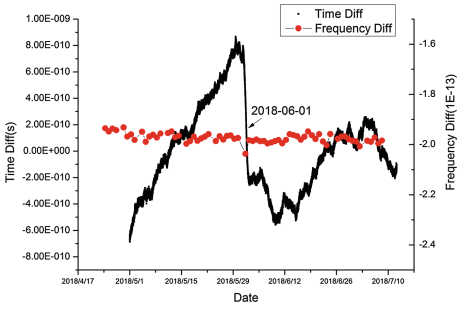


Fig. 6. The results of frequency difference and time difference

atomic clock made by the cesium atomic fountain device, and it can be found that the results of the evaluation on June 1, 2018 mutated, indicating that the frequency of the moment has jumped.

5 Conclusion

Cesium atomic fountain device is a second signal reappearance device. On one hand, the system time can be monitored and evaluated directly by using output frequency signal of cesium fountain device, and the evaluation results can be used as a reference and basis for the system time control, which can effectively improve the accuracy of the system time; on the other hand, the output signal of the cesium atomic fountain device can be used to monitor the output signals of the atomic clock, and to detect the frequency jump of the atomic clock, so that the problem atomic clock in the clock group can be found timely and effectively. Therefore, making full use of the data of cesium atom fountain clock plays an important role in improving the performance of time frequency signal in punctual laboratory.

Acknowledgments. Thanks for the maintenance and technical support provided by National Institute of Metrology.

References

1. Liu X (2011) Not exceeding one second after 15 million years China National Institute of Metrology successfully developed NIM5 cesium atomic fountain Clock. China Metrol 2011 (1):11
2. Bingying H (2004) New generation time frequency standard-laser cooling, manipulation of cesium atom fountain clock. China Metrol 2:47–48
3. Wynands R, Weyers S (2005) Atomic fountain clocks. Metrologia 42:S64
4. Itano WM, Lewis LL, Wineland DJ (1982) Shift of hyperfine splittings due to blackbody radiation. Phys Rev A 25(2):1233–1235
5. Ramsey NF (1956) Molecular Beams. Clarendon Press, Oxford, p 80
6. Lang Z et al (2011) Analysis of influencing factors of frequency shift of NIM5-M Cs cooling atom fountain clock的. In: China satellite navigation conference
7. Lin R, Wang XL, Liu DD et al (2017) Research on differential measurement method of the cold atomic collision shift in the Cesium fountain clock. J Time Freq
8. Yang S, et al (2007) Analysis of frequency drift characteristics of navigation satellite atomic clock. GNSS world China 32(6):5–10
9. Lang Z et al Research on the external comparison method of the frequency datum of cesium fountain. China satellite navigation conference

Analysis on BDS Satellite-Induced Multipath and Its Impact on Time Transfer

Guojun Li(✉), Fengfeng Shi, Yuting Lin, and Zhiping Wang

Beijing Satellite Navigation Center, Beijing, China
lolo551750@qq.com

Abstract. There are systematic multipath errors on BeiDou code measurements, ranging from several decimeters to more than 1 m. The impact of BDS satellite-induced multipath on time transfer was comprehensively analyzed. Firstly, the piecewise liner correction model was constructed using the observation data collected from 17 MGEX stations over a time span of 30 days. Secondly, the time transfer results of 4 time labs located in Europe-Asia were analyzed in different time transfer mode, including single-station BDS CV (common view), two-station BDS CV, single-station BDS PPP (precise point positioning) and two-station BDS PPP. Experiments show that the systematic error originated from the satellite-induced multipath is about 0.5 ns–1.0 ns when time transferring in Asia or Europe. The maximum error are 3 ns–4 ns for single satellite BDS CV. The error will be obviously canceled when time transferring in short baseline, but not for long baseline, which is up to 1.5 ns and must be corrected.

Keywords: Multipath · Time transfer · BDS CV · BDS PPP

1 Introduction

The BDS-2 has been providing position, navigation, and timing (PNT) services since 27 December 2012, covering the whole Asia-Pacific region. Currently, the system consists of 15 satellites, including 5 geostationary earth orbit (GEO) satellites, 7 inclined geosynchronous orbit (IGSO) satellites, and 3 medium earth orbit (MEO) satellites. The constellation will be fully operational by 2020. With the development of the BDS, the system performance has been widely monitored by many agencies and scholars at home and abroad.

Presently, the time transfer methods of the time laboratories participating in TAI (International Atom Time) mainly relay on GPS and GLONASS [1]. GPS PPP has great advantages in high-precision time and frequency transfer over a long distance. It can reach 1E-15–1E-16 over an averaging period of 1 day when frequency transferring. Moreover, its type A uncertainty is less than 0.3 ns [2, 3]. However, there exists high risks relaying on only one time transfer method. It's more recommended that multi-GNSS combined time transfer should be applied. CGGTTS V2E, an extension of the CGGTTS format including GPS, GLONASS, Galileo, BDS, and QZSS, was proposed by the Working Group on GNSS Time Transfer of the Consultative Committee for Time and Frequency (CCTG) in 2015 [4]. Preliminary experiments showed that the

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 518–530, 2019.
https://doi.org/10.1007/978-981-13-7751-8_50

BDS CV in Eurasia could achieve an accuracy of 2 ns–3 ns with 2–3 common-view satellites [5]. The BDS PPP time transfer currently was realized only in the Asia-Pacific region with an accuracy of better than 0.5 ns [6].

In recent years, it was found that there exists systematic errors on BDS code measurements [7, 8], which were closely related to the signal frequencies and the satellite elevations, and be independent of the type of receiver antenna. Therefore, it was preliminarily concluded that the pseudorange deviation was originated from the satellite-induced multipath. In order to eliminate the errors, Wanninger [8] estimated the correction parameters of IGSO/MEO with an elevation interval of 10° by using observation data collected from the MGEX stations. Guo [9] estimated the correction parameters covering 5°–85° by the total least squares method, and verified that the pseudorange deviation had a significant influence on fixing the ambiguity. Li [10] analyzed the influences on the correction parameters estimated, such as the number of the observation stations, the length of the observation period. It was proved that the accuracy of PPP could be improved after the pseudo-range corrected. In order to eliminate the ambiguities, Zou [11] estimated the correction parameters by fitting a third order polynomial. Zhao [12] analyzed the correlations between the multipath and the signal frequencies or the satellite elevations. The results showed that the single-station multipath was closely related to the elevations. The multipath can be significantly reduced by performing difference between two stations. Pan [13] constructed a piecewise liner correction model with an elevation interval of 1° by averaging the 1-order difference of the multipath series. For the stationary GEO satellites, Lou [14] estimated the correction of the GEO with observation data collected from many stations. Ruan [15] concluded that the multipath of the GEO are similar to that of the IGSO for that the difference between them was less than 0.06 m.

In summary, the above researches constructed the satellite-induced multipath correction model with the observation data collected from many stations, and mainly focused on the impact of the satellite-induced multipath on positioning. Few studies have been conducted to analyze the impact of this systematic error on BDS time transfer. Firstly, the concepts and characteristics of the satellite-induced multipath were introduced. Secondly, a piecewise linear correction model was constructed for the IGSO/MEO with the data collected from 17 MGEX stations over a time span of 30 days. Finally, the impact of the satellite-induced multipath on time transfer was analyzed, with the experiments of single-station BDS CV, two-station BDS CV, single-station BDS PPP and two-station BDS PPP among 4 time-labs in Eurasia.

2 BDS Satellite-Induced Multipath

2.1 The Multipath Combination

The code multipath (MP) can be estimated through a combination of code range and carrier phase observables. The multipath combination on frequency i can be expressed as [8]:

$$\begin{aligned} MP_i &= P_i - \frac{f_i^2 + f_j^2}{f_i^2 - f_j^2} L_i + \frac{2f_i^2}{f_i^2 - f_j^2} L_j \\ &= M_i - \frac{f_i^2 + f_j^2}{f_i^2 - f_j^2} m_i + \frac{2f_i^2}{f_i^2 - f_j^2} m_j + B_i + \varepsilon_i \end{aligned} \tag{1.1}$$

where i and $j\,(i, j = 1, 2, 3, i \neq j)$ represent the signal. P and L are, respectively, the frequency and the wavelength of the signal. f is the signal frequency. M and m is the code and phase multipath respectively. B consists of the phase ambiguity and the signal delay. ε is the noise. The signal delay is constant in a short time. Compared with the code multipath, the phase noise and phase multipath are neglected. Therefore, if there are no cycle slip, the fluctuation of the MP combination should be the same as that of the code multipath. Actually, the multipath were obtained by subtraction of the mean value of the MP time series in each continuous observation arc.

Figure 1 shows the code multipath related to the elevations for each GNSS satellites. Obviously, the MP series was much noisier when the elevation was lower. Except for the BDS, the mean of the multipath was almost 0. The MP series of the IGSO and MEO contained systematic bias which was related to the elevations, so as for the other IGSO/MEO satellites.

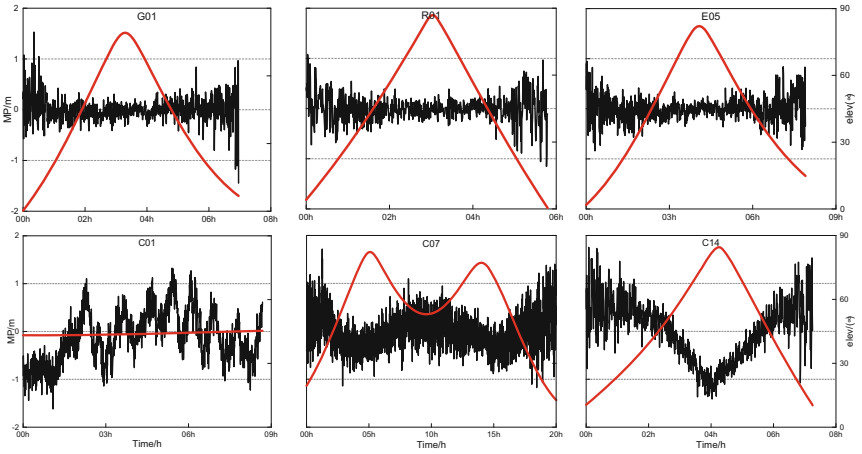


Fig. 1. The multipath of different GNSS satellites

2.2 MEO/IGSO Pseudorange Correction Model

We used the data collected from the widespread distributed MGEX stations to construct the correction model for the satellite-induced multipath. It was proved that the selected stations should be distinguished for IGSO and MEO [11]. The pseudorange RMS of the IGSO was very large outside of the Asia-Pacific. It may pollute the whole observation data and lead to the reduction of the modeling accuracy. Therefore, for the IGSO, the selected stations mainly located in the Asia-Pacific. For the MEO, the selected stations

should be distributed as widely as possible. Figure 2 shows a total of 17 MGEX stations we selected for the MEO correction modeling, including 5 stations (read) for the IGSO correction modeling.

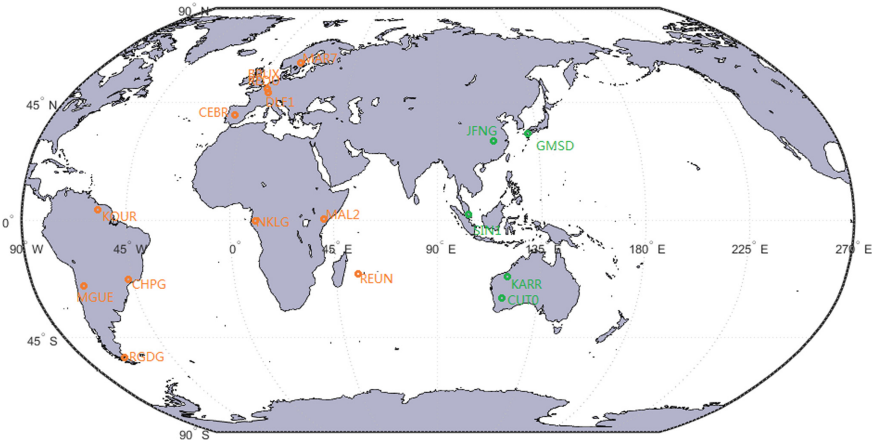


Fig. 2. The distribution of 17 MGEX stations

We adjusted best fitting models for each MEO and IGSO satellites, for each of the frequencies. Assuming that the MP series is piecewise continuous, the object function can be expressed as [9]:

$$\begin{aligned} S &= \sum_{j=1}^{m-1} \sum_{i=1}^{n_j} (f_{j,i} - MP_{j,i})^2 = \min \\ s.t. \quad & f_j(e_j) - f_{j+1}(e_j) = 0 \end{aligned} \tag{1.2}$$

where $j = 1, 2, \cdots, m-1$, m is the number of nodes. $i = 1, 2 \cdots, n$ is the length of the MP series. f is the piecewise linear function. MP is the multipath combination. We discarded the observations below 5° for large RMS. Besides, the observations higher than 85° were also discarded for not enough observations. Given an observation $[e, \mathrm{MP}]$, the equation can be expressed as:

$$\mathrm{MP} = (1 - \frac{e - e_j}{5})x_j + \frac{e - e_j}{5} x_{j+1} \tag{1.3}$$

where x is the node value. j is the segment index. Accounting for the large amount of observations, the equation is solved by sequential adjustment.

Since the dual-frequency CGGTTS for BDS is based on the ionosphere-free combinations of pseudorange on B1 and B2. We only estimated the MP correction parameters on B1 and B2, not including B3. Figure 3 shows the piecewise linear functions on B1/B2 for IGSO/MEO, which was estimated with the MP series of 17 MGEX stations, covering 30 days from 2018/10/01 to 2018/10/30. The grey denotes

the piecewise linear function constructed by the MP series for each satellite. The blue denotes the averaged of all the IGSO/MEO linear functions. The red denotes the piecewise linear function estimated with the total least square method. It was shown that the piecewise linear functions were very similar for the same type of satellites. The results obtained by the averaged method were similar to that of the total least square method. The corrections on B1 were larger than that on B2. The corrections of the MEO were obviously larger than that of the IGSO, which was up to 1 m when the elevation was very high.

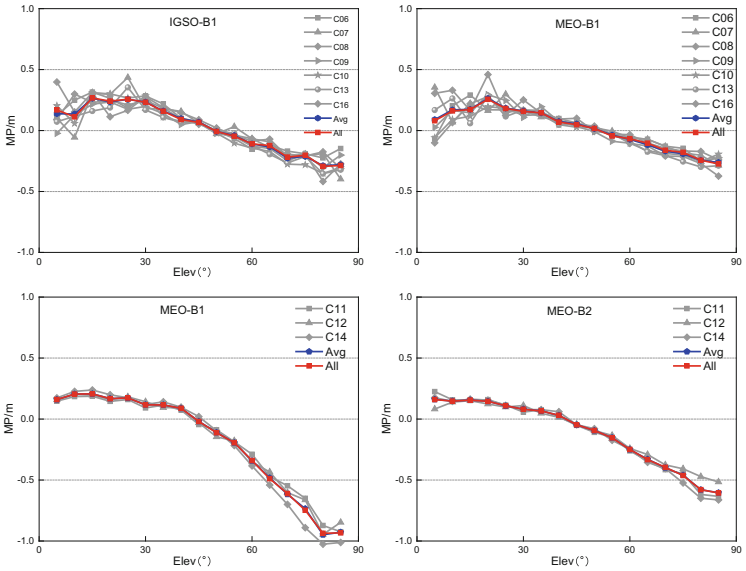


Fig. 3. The piecewise liner functions of the IGSO/MEO on B1/B2

Table 1. The pseudo-range corrections of the IGSO/MEO on B1/B2

Elevation/(°)	Corrections/m			
	IGSO		MEO	
	B1	B2	B1	B2
5	−0.169	−0.08	−0.16	−0.16
10	−0.112	−0.161	−0.205	−0.147
15	−0.267	−0.172	−0.206	−0.155
20	−0.241	−0.256	−0.167	−0.145
25	−0.257	−0.181	−0.173	−0.109
30	−0.232	−0.157	−0.116	−0.08
35	−0.158	−0.145	−0.116	−0.066
40	−0.09	−0.069	−0.088	−0.031
45	−0.066	−0.048	0.02	0.047

(*continued*)

Table 1. (*continued*)

Elevation/(°)	Corrections/m			
	IGSO		MEO	
	B1	B2	B1	B2
50	0.008	−0.017	0.112	0.094
55	0.047	0.045	0.194	0.154
60	0.109	0.067	0.34	0.249
65	0.124	0.104	0.489	0.333
70	0.218	0.162	0.611	0.396
75	0.203	0.179	0.748	0.463
80	0.294	0.244	0.937	0.579
85	0.287	0.274	0.933	0.606

To verify the validity of the piecewise linear correction function, the parameters estimated by the total least square method in Table 1 were applied to correct the pseudorange. Figure 4 shows the multipath of IGSO/MEO on B1 with or without corrections. Figure 5 shows the statistic of the multipath. It can be seen that: (1) the RMS was not improved at low elevation for the large observation noises; (2) the RMS was unchanged when the elevation was nearby 45° for that the corresponding correction was almost 0; (3) the RMS became smaller when the correction was applied at high elevation, especially for the MEO satellite. Therefore, the piecewise linear model can effectively reduce the systematic error resulting from the satellite-induced multipath, especially when the elevation is high.

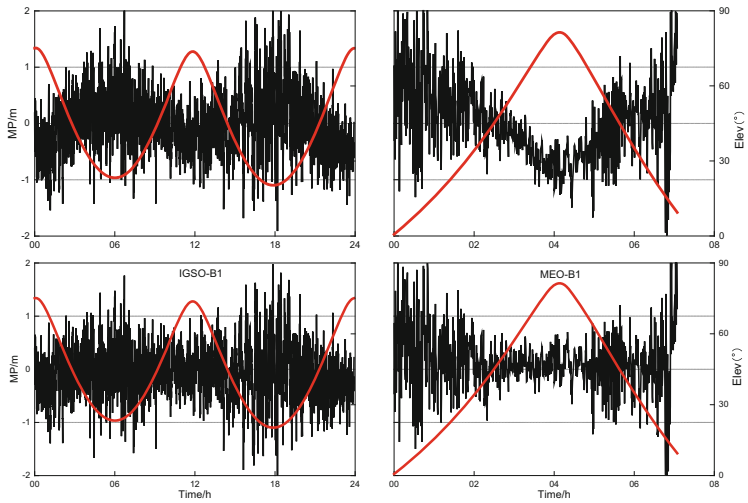


Fig. 4. The multipath of IGSO/MEO on B1 with or without corrections (upper: uncorrected; lower: corrected)

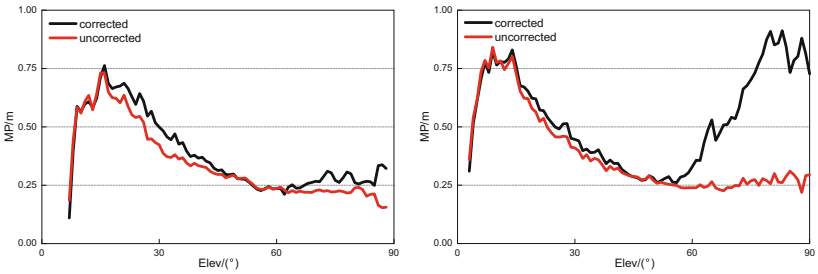


Fig. 5. The multipath RMS of the IGSO/MEO on B1 with or without corrections

3 The Impact of BDS Satellite-Induced Multipath on BDS CV Time Transfer

It should be noted that the multipath estimated by formulas (1.1) and (1.2) reflects the relative multipath, not the absolute multipath. It was proved that the position accuracy can be improved when the satellite-induced multipath was corrected. If not, the error was absorbed in the receiver clock. Therefore, it is necessary to further study the impact of the satellite-induced multipath on time transfer.

As defined in CGGTTS V2E, the satellite common-view equation can be expressed as [4]:

$$t_{clock} - t_{sat} = \frac{1}{c}\left[\overline{P_{IF}} - \left\|\overrightarrow{x_{sat}} - \overrightarrow{x_{rec,IF}}\right\| - S\right] + \Delta t_{rel} - \Delta t_{trop} - GD \tag{1.4}$$

where $\overline{P_{IF}}$ is the ionosphere-free combination. $\overrightarrow{x_{sat}}$ is the satellite position. $\overrightarrow{x_{rec,IF}}$ is the antenna phase center corresponding to the ionosphere-free combination. S is the Sagnac correction. Δt_{rel} is the relativistic clock correction. Δt_{trop} is the ionospheric delay. GD is the broadcast group delay. The BDS TDG is referenced to B3, the group delay must be considered:

$$GD = \frac{T_{GD1} f_1^2 - T_{GD2} f_2^2}{f_1^2 - f_2^2} \tag{1.5}$$

As the satellite-induced multipath was small, the impact of the corrections was obvious on precise point positioning, but not on single-point positioning. Therefore, when analyzing the impact of the satellite-induced multipath on time transferring, we mainly focused on the receiver data in the time labs to avoid the distortion of the receiver clock. 3 domestic time labs (NIM, NTSC and BSNC, BIRM has no BDS data) and BRUX were selected for analysis, as shown in Table 2. Presently, the international time comparison system in BSNC is in pilot. We only got a continuous data over a time span of 5 days from 2018/10/20 to 2018/10/24.

Table 2. The time lab receivers' information

Time-lab	Receiver	Distance from BSNC
BRUX	SEPT POLARX4TR	7941 km
NTSC	SEPT POLARX4TR	890 km
NIM	TF-GNSS-200B	19 km
BSNC	SEPT POLARX5TR	–

The GEO were excluded when time transferring for that the correction model of the GEO has not been constructed. In addition, all the time transfer results were uncalibrated for that the calibration method for BDS has not been clearly defined. Figure 6 shows the results of BDS CV in 4 time-labs. The black and the red in the left, respectively, were the results of BDS CV uncorrected and corrected. The black in the right was the difference between the uncorrected and the corrected. The blue was the number of visible satellites above 10°. It can be seen that: (1) the fluctuation of BDS CV in BRUX was obviously greater than that of the other three time labs; (2) the fluctuation of the time difference between corrected and uncorrected of BDS CV is the greatest; (3) the difference between the corrected and the uncorrected in three domestic time labs is negative, while positive in BRUX. The above results were mainly due to the fact that the elevation of the IGSO/MEO observed in BRUX was obviously lower than that of the domestic time labs, as shown in Fig. 7. The elevations of the observed satellites in BRUX were almost 30°, and the corresponding corrections were about −0.2 m for the MEO, and −0.1 m for that of the IGSO. Therefore, comparing with the uncorrected, the corrected for each satellites decreased. Sometimes the elevations of the IGSO could reach 80° in BRUX, such as C14 at 2018/10/21 13:10:00, the elevation was 83.5°. At this moment, the pseudorange contained obvious systematic error of about +0.9 m. The final common-view results calculated by weighted average algorithm based on the elevation may be seriously affected by the satellite-induced multipath.

In order to further illustrate the effects of satellite-induced multipath on BDS CV between stations. Figure 8 shows the time transfer results between BSNC and BRUX, NTSC or NIM. It can be seen that the time difference of the time transfer results between the corrected and the uncorrected in time-link BRUX-BSNC is the greatest, with maximum value of 1.5 ns, much greater than that of the time-link NIM-BSNC. It's due to that the elevations of the common-view satellites are nearly the same in time-link NIM-BSNC of distance 19 km. The satellite-induced multipath was greatly reduced by come-view time transfer between stations.

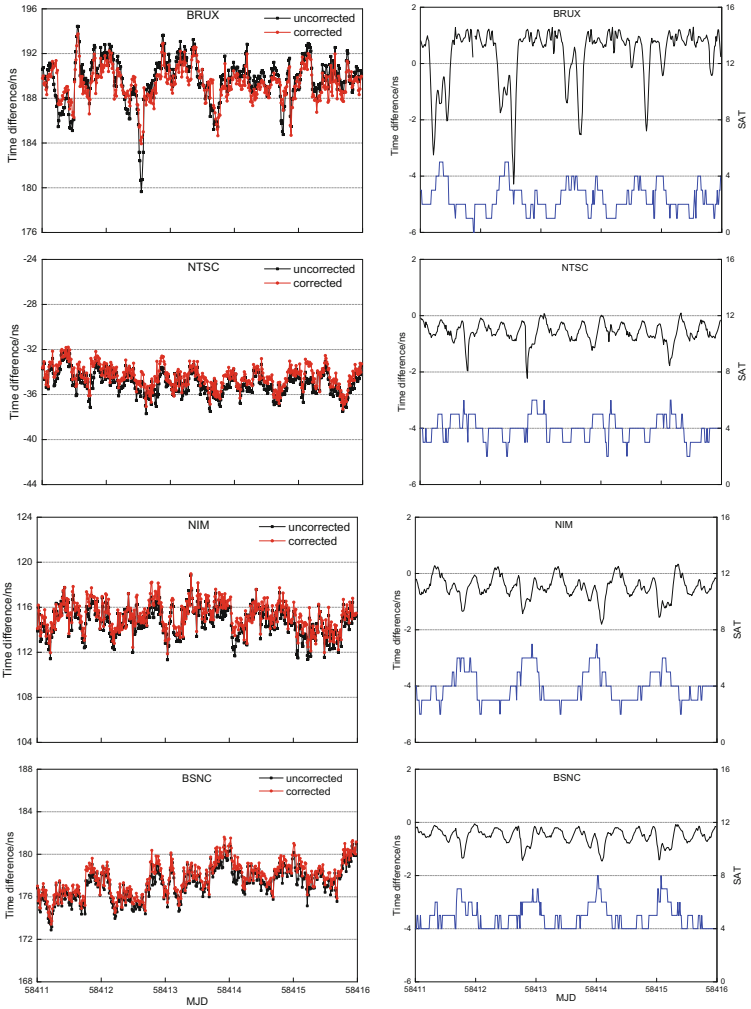


Fig. 6. The results of single-station BDS CV

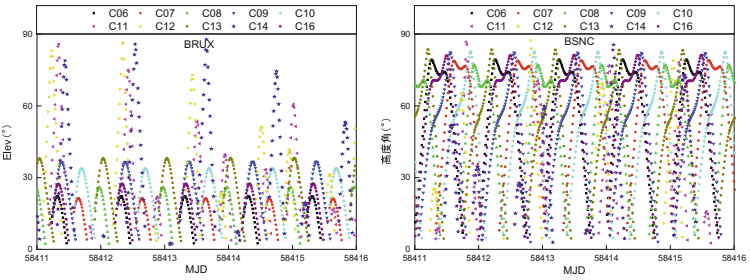


Fig. 7. The elevations of the single-station common-view satellites

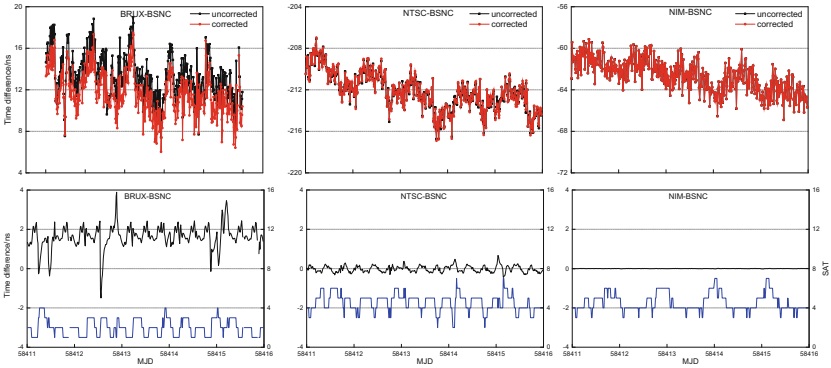


Fig. 8. The results of two-station BDS CV

4 The Impact of BDS Satellite-Induced Multipath on BDS PPP Time Transfer

According to Table 1, for a single BDS satellite, it is predicted that the error originated from the satellite-induced multipath is about 1 ns–3 ns when time transferring by BDS CV. The error may be reduced by averaging the time transfer results of all the common-view satellites. Generally, the type A uncertainty of BDS CV is about 2 ns. Hence the satellite-induced multipath maybe not obvious when the baseline is short or with many common-view satellites. To further analyze the impact of the satellite-induced multipath on time transfer, the BDS PPP time transfer based on RTKLIB [16] was developed with the multipath corrected or uncorrected.

Different form BDS CV, the observations of the GEO, though without corrections, were used with down-weights to ensure enough satellites for BDS PPP time transfer. The precise ephemeris we adopted were provided by GFZ. Because of the lack of visible satellites in Europe, only three time labs in Asia are analyzed here.

Figure 9 shows the single-station BDS PPP time transfer results of NTSC, NIM and BSNC. It can be seen that there existed similar day-boundary jumps in the time transfer results for all stations. This mainly due to that the precise ephemeris provided by GDZ was not continuous. The differences between the corrected and the uncorrected were about −0.5 ns, which were very similar in every day. Comparing Fig. 9 with Fig. 6, we can found that the time difference between the corrected and the uncorrected for the single station BDS CV were in good agreement with that of BDS PPP. Therefore, it can be preliminarily concluded that the satellite-induced multipath will produce systematic errors in the single-station time transfer, which are closely related to the number of observed satellites and the elevation.

The time transfer results of BDS CV has shown that the error originated from the satellite-induced multipath is related to the length of the baseline. Next, we illustrate the impact of the satellite-induced multipath on two-stations BDS PPP time transfer. Figure 10 shows the BDS PPP time transfer results in the time-link NTSC-BSNC and NIM-BSNC. The time difference between the corrected and the uncorrected for

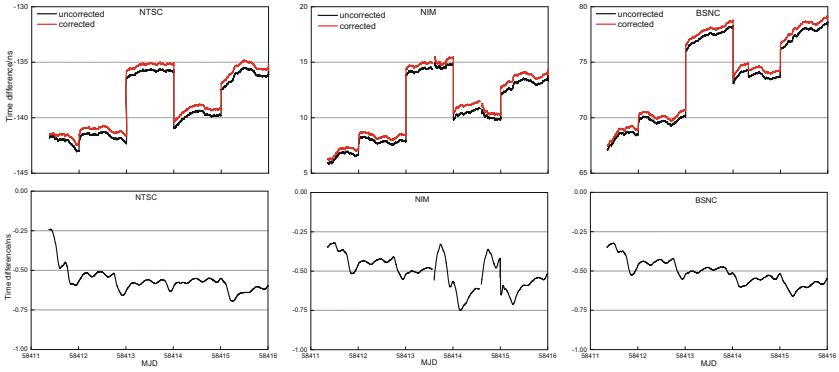


Fig. 9. The results of single-station BDS PPP

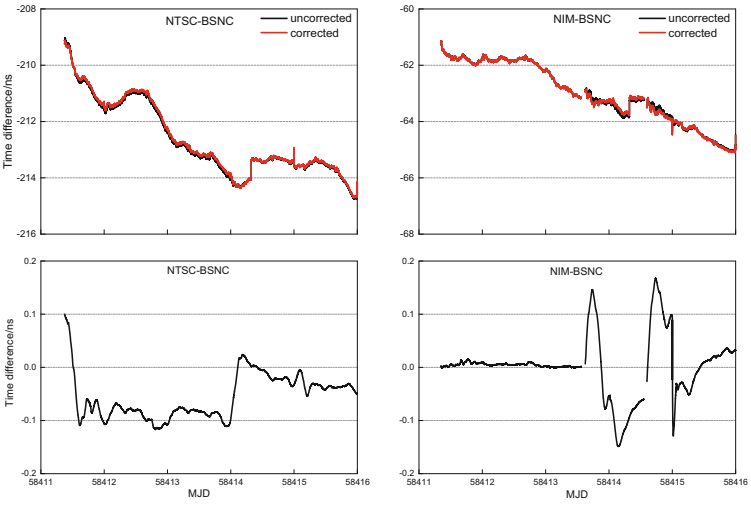


Fig. 10. The results of two-station BDS PPP

BDS PPP time transfer is about 0.1 ns, which is almost 0 in the time-link NIM-BSNC in the beginning. There appeared an obvious fluctuation between during the third day and the fourth day in the time-link NIM-BSNC. It was mainly due to the improper data preprocessing strategy for the receiver observations in NIM. In addition, the time transfer results of the two time-links occurred a jump of about 0.5 ns at the same time during the fourth day, which was confirmed that the connected cables were accidentally touched by the technicians.

5 Conclusion

(1) Using the observation data collected from the 17 MGEX stations over a time span of 30 days, a piecewise linear correction model with an elevation interval of 5° was constructed, which was feasible to correct the pseudorange ranging from 5° to 85°. The corrections on B1 are greater than that on B2, whose maximum value is about 1 m. The constructed correction model could effectively reduce the RMS of the pseudorange at high elevation.
(2) The impact of the satellite-induced multipath on BDS CV was analyzed based on the time transfer results in BRUX, NSTC,NIM and BSNC. It was shown that the errors originated from the satellite-induced multipath were closely related to the number of common-view satellites and the elevations. The corrections on BDS CV in Europe is almost positive, whose maximum value is up to 3 ns–4 ns, while negative in Asia. The error will be canceled when time transferring in short baseline, while it must to be corrected in long baseline.
(3) The impact of the satellite-induced multipath on BDS PPP was analyzed based on the time transfer results in NSTC,NIM and BSNC. The errors originated from the satellite-induced multipath was up to 0.5 ns, which is neglected when time transferring in short baseline by BDS PPP.
(4) This is our first time to verified the status of the time comparison system in BSNC. It was proved that the observed data was available and the system worked well.

Accounting for that the modifications on the existed software was accessible and fixed, we strongly recommend that the satellite-induced multipath should be considered. It's unnecessary for the BDS-3 satellites, for that this problem has been solved [17].

Acknowledgments. Thanks to the observation data provided by NIM and NTSC.

References

1. Arias EF (2018) BIPM annual report on time activities. BIPM, Sevres Cedex
2. Ray J, Senior K (2005) Geodetic techniques for time and frequency comparisons using GPS phase and code measurements. Metrologia 42(4):215–232
3. 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
4. Defraigne P, Petit G (2015) CGGTTS-version 2E: an extended standard for GNSS time transfer. Metrologia 52(6):G1–G1
5. Guang W, Dong S, Wu W et al (2018) Progress of BeiDou time transfer at NTSC. Metrologia 55(2):175–187
6. Zhang J, Guang W, Wu W et al (2018) BDS PPP time transfer at NTSC. In: Frequency control symposium. IEEE (2018)
7. Hauschild A, Montenbruck O, Sleewaegen JM et al (2012) Characterization of compass M-1 signals. GPS Solutions 16(1):117–126
8. Wanninger L, Beer S (2015) BeiDou satellite-induced code pseudorange variations: diagnosis and therapy. GPS Solutions 19(4):639–648

9. Guo F, Li X, Liu W (2016) Mitigating BeiDou satellite-induced code bias: taking into account the stochastic model of corrections. Sensors 16(6):909
10. Li X, Zhang X, Zeng Q et al (2017) The estimation of BeiDou satellite-induced code bias and its impact on the precise positioning. Geomat Inf Sci Wuhan Univ 42(10):1461–1467
11. Zou X, Li Z, Li M et al (2017) Modeling BDS pseudorange variations and models assessment. GPS Solutions 21(3):1–8
12. Zhao Q, Wang G, Liu Z et al (2016) Analysis of BeiDou satellite measurements with code multipath and geometry-free ionosphere-free combinations. Sensors 16(1):123
13. Pan L, Guo F, Ma F (2018) An improved BDS satellite-induced code bias correction model considering the consistency of multipath combinations. Remote Sens 10:1189
14. Lou Y, Gong X, Gu S et al (2017) Assessment of code bias variations of BDS triple-frequency signals and their impacts on ambiguity resolution for long baselines. GPS Solutions 21(1):177–186
15. Ruan R, Jia X, Feng L (2017) Analysis on BDS satellite internal multipath and its impact on wide-lane FCB estimation. Acta Geod Artographica Sin 46(8):961–970
16. RTKLIB (2018) RTKLIB: an open source program package for GNSS positioning [EB/OL]. http://www.rtklib.com/
17. Yang Y, Xu Y, Li J et al (2018) Progress and performance evaluation of BeiDou global navigation satellite system: data analysis based on BDS-3 demonstration system. Sci China Earth Sci 61(5):614–624

Analysis of GNSS Multi-constellation Time Transfer Accuracy

Guang Liu(✉), Changsheng Cai, and Zhipeng Li

Central South University, Changsha, China
liuguang95l0l6@l63.com

Abstract. Timing service is an indispensable part of national economy, and timing technology directly restricts the accuracy of time obtained by users. High precision time transfer technology is an indispensable technology for time synchronization in all time laboratories. In this paper, based on GNSS multi-constellation precise single point positioning (PPP) technology, a time transfer method based on GNSS quad-constellation was proposed, and the precision satellite orbit and clock difference products provided by IGS were used for time transfer accuracy and frequency stability analysis. The experimental results show that the GNSS multiple constellation PPP can complete the time transfer of the accuracy of nanosecond. The PPP time transfer accuracy of GPS/BDS/GLONASS three-system combination is improved to a certain extent compared with GPS single system and GPS/BDS two-system, while the precision of GPS/BDS/GLONASS/Galileo four-system is not improved obviously compared with three-system. Through the analysis of the satellite data for 4 consecutive days, it is found that the inter-station clock difference solution calculated by GNSS multi-constellation PPP can reach the frequency stability of 5×10^{-13}–6×10^{-13}, which is in good frequency consistency with the clock difference product released by IGS.

Keywords: PPP · Transfer time · GNSS multi-constellation · Clockerror

1 Introduction

High precision time transfer is the basis of establishing and maintaining standard time scale and keeping time synchronization in time lab. With the development of space technology, the accuracy of time transfer is getting higher and higher. Recision single point positioning (PPP) technology has great application potential in high-precision time transfer due to its advantages of high accuracy, simple data acquisition and easy data processing. Since 2000, Kouba has first evaluated the potential of the PPP method for time transfer [1]. Then in 2003, the international GNSS service (IGS) and international bureau of standard measurement (BIPM) in order to better use of advantages in the field of GPS in time, conducted a survey time, determine the IGS time scales, and coordinated universal time (UTC), for the PPP calculating more stable clock difference product, promoted the growth of the PPP time transfer technology [2]. Next, Costa et al. [3] studied the use of PPP method in time transfer, and the results showed that there was good consistency between PPP receiver bell difference solution and IGS

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 531–543, 2019.
https://doi.org/10.1007/978-981-13-7751-8_51

clock difference product, which reached the level of yanna-second, and the short-term and medium-term stability was better than traditional GPS common vision method and total vision method. Domestic scholars started late in applying PPP technology to timing. Zhang et al. [4] carried out PPP time transfer accuracy analysis in 2010, showing that static PPP can carry out time transfer of yannas and analyzing the accuracy of GPS PPP for remote time comparison. With the development of beidou navigation system, Chinese scholars began to use beidou PPP technology to carry out time transfer research. For example, Zhang et al. [5] used beidou PPP technology to conduct time transfer experiments, and compared the accuracy of GPS PPP time transfer. In terms of studies on GNSS multi-constellation PPP time transfer, Defraigne et al. [6] used combined GPS/GLONASS PPP technology to make time comparison. The results showed that the addition of GLONASS data could significantly reduce the timing synchronization error and improve the short-term stability of time transfer. Liang et al. [7] showed that the time transfer accuracy based on GPS/BDS carrier phase observation was improved relative to single accuracy. At present, most researches are based on the time transfer of a single system or a dual system, while GNSS has shown the coexistence of multiple constellations. This paper analyzes the accuracy of time transfer based on multi-module and multi-constellation PPP technology and proposes a time transfer method based on quad-constellations. And the frequency stability between multi - system and single system is compared.

2 Multi-constellation PPP Time Transfer Method

The PPP time transfer principle is similar to the single point positioning principle. PPP requires precision track data and precision clock difference data to overcome the error of broadcasting track and clock difference. Unlike the difference method, PPP method can not eliminate or reduce the influence of the misalignment between satellite antenna phase center and satellite center of mass, and the deviation of satellite and receiver phase center provided or recommended by IGS is used to correct it. At the same time, more sophisticated error correction models were considered, including antenna phase center correction, solid tide correction, ocean tide correction, multipath effect, etc. Using PPP technology and high precision time, every time laboratory needs to maintain consistent with IGS time scales (IGST), it is using a dual-frequency receiver for external frequency target ranging pseudo code and carrier phase observation value, calculate the time difference between local time and IGST time, difference between stations, the results can be the difference between two stations.

For a GNSS satellite *j*, the ranging pseudocode and carrier phase observations on the *ith* frequency can be expressed as [8]:

$$P_i = \rho + cdt - cdT + d_{orb} + d_{trop} + d_{ion/L_i} + b_{p_i} + \varepsilon_{P_i} \tag{1}$$

$$\Phi_i = \rho + cdt - cdT + d_{orb} + d_{trop} - d_{ion/L_i} + B_i + \varepsilon_{\Phi_i} \tag{2}$$

Where P_i is the pseudorange observation at the *ith* frequency, m; Φ_i is the carrier phase observation at i frequencies, m; ρ is the geometric distance between the satellite

and the station, m; c is the speed of light, m/s; dt is the receiver clock difference, s; d_{orb} is the satellite orbit error, m; d_{trop} is the tropospheric delay, m; d_{ion/L_i} is the ionospheric delay at the *ith* frequency, m, b_{P_i} is the hardware delay deviation in the code pseudorange on the *ith* frequency, m; B_i is the phase ambiguity term on the ith frequency, including the receiver and satellite initial phase offset and phase hardware delay, m; ε_{P_i} and ε_{Φ_i} contain multipath error and measurement noise, m.

Taking GPS as the system time, using the traditional ionospheric combined model, the GNSS multi-constellation combined PPP observation equation is:

$$P_{IF}^{g} = \rho^{g} + cdt + d_{trop}^{g} + \varepsilon_{P_{IF}}^{g} \tag{3}$$

$$\Phi_{IF}^{g} = \rho^{g} + cdt + d_{trop}^{g} + B_{IF}^{g} + \varepsilon_{\Phi_{IF}}^{g} \tag{4}$$

$$P_{IF}^{r} = \rho^{r} + cdt + cdt_{sys}^{r,g} + d_{trop}^{r} + \varepsilon_{P_{IF}}^{r} \tag{5}$$

$$\Phi_{IF}^{r} = \rho^{r} + cdt + cdt_{sys}^{r,g} + d_{trop}^{r} + B_{IF}^{r} + \varepsilon_{\Phi_{IF}}^{r} \tag{6}$$

$$P_{IF}^{e} = \rho^{e} + cdt + cdt_{sys}^{e,g} + d_{trop}^{e} + \varepsilon_{P_{IF}}^{e} \tag{7}$$

$$\Phi_{IF}^{e} = \rho^{e} + cdt + cdt_{sys}^{e,g} + d_{trop}^{e} + B_{IF}^{e} + \varepsilon_{\Phi_{IF}}^{e} \tag{8}$$

$$P_{IF}^{c} = \rho^{c} + cdt + cdt_{sys}^{c,g} + d_{trop}^{c} + \varepsilon_{P_{IF}}^{c} \tag{9}$$

$$\Phi_{IF}^{c} = \rho^{c} + cdt + cdt_{sys}^{c,g} + d_{trop}^{c} + B_{IF}^{c} + \varepsilon_{\Phi_{IF}}^{c} \tag{10}$$

Where, g, r, e, c represent GPS, GLONASS, BDS and Galileo satellites respectively. $dt_{sys}^{r,g}$, $dt_{sys}^{e,g}$, $dt_{sys}^{c,g}$ are the system time differences of GPS-GLO, GPS-GAL and GPS-BDS respectively. B_{IF} is the phase combination *IF* fuzzy term; $\varepsilon_{P_{IF}}$ and Φ_{IF} contain multiple path errors and measurement noise.

In this experiment, mips-ppp software developed independently by central south university is adopted for processing, which can process the current observation data of GPS, BDS, GLOANSS and Galileo multiple systems separately or jointly, and provide the clock difference results of the receiver. The PPP of multiple constellations firstly preprocesses the data, including gross difference elimination of observation data, phase cycle slip detection, initial value calculation of parameters, etc. The kalman filter is used for parameter estimation, and various error models are used for correction. The specific calculation process is shown in Fig. 1:

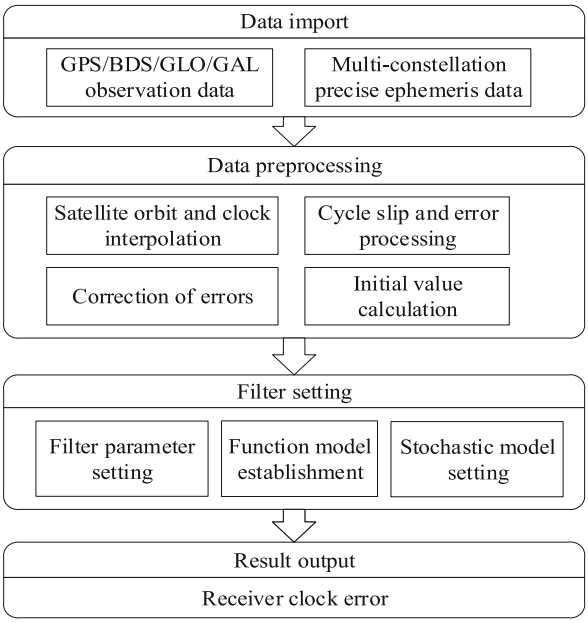


Fig. 1. Multi-constellation PPP time transfer algorithm process

3 Time Transfer Accuracy and Frequency Stability Evaluation Method

Strictly speaking, the precision of PPP time transfer is essentially determined by the stability of satellite clock, the precision of PPP clock difference solution and the frequency stability embodied in many of them [4]. In order to evaluate the precision and frequency stability of PPP clock differential solution, we introduced RMS in internal compliance, RMS in external compliance and Allan variance as evaluation criteria.

Inner coincidence RMS means that we use multiple constellation PPP to calculate the clock difference of the receiver in the experimental station, and use the least square fitting to remove the linear trend term from the result of the clock difference, obtain the result of the bell difference residual after deducting the trend term, and analyze the residual standard deviation.

External RMS is to take the inter-station clock difference of the experimental link provided by IGS GBM center as the reference, and the difference between the link clock and the two measuring stations is relatively poor.

$$RMS = \sqrt{\Delta^T P \Delta / n} \tag{11}$$

Where, Δ is the sequence of clock difference; n is the total number of samples of clock difference sequence; P is the weight matrix (in this case, P is the identity matrix).

Allan variance is a commonly used characterization method of frequency stability. Overlapping Allan variance increases the number of equivalent degrees of freedom, thus improving the confidence of estimation:

$$\sigma(\tau) = \sqrt{\frac{1}{2\tau^2(N-2m)}\sum_{i=1}^{N-2m}(X_{i+2m} - 2X_{i+m} + X_i)^2} \tag{12}$$

In the formula, X_i is the clock difference solution of the element i, and N is the number of samples of the clock difference solution sequence (the number of elements); τ is the sampling interval and m is the smoothing factor. In this paper, a clock difference file of 5 min is adopted, so τ is set as 300 s for better comparison.

4 Analysis of Experimental Results

Data of two stations, GMSD and JFNG, were obtained from MEGX station from August 25, 2017 to August 28, 2017. All observation files contain four system data, the time system is GPS time, the precision ephemeris file is 15 min, and the precision clock difference file is 5 min.

4.1 Internal Accuracy Results

Based on the four constellations, four combinations of GPS single system, GPS/BDS dual system, GPS/BDS/GLO three system and GPS/BDS/GLO/GAL four system combined PPP were used to calculate the clock difference between JFNG and GMSD station, perform the least squares fitting, and analyze the residual results to show its internal consistency accuracy. The four-day results showed no significant differences. Due to space limitations, only the results of August 25, 2017 are shown here (Figs. 2, 3, 4 and 5).

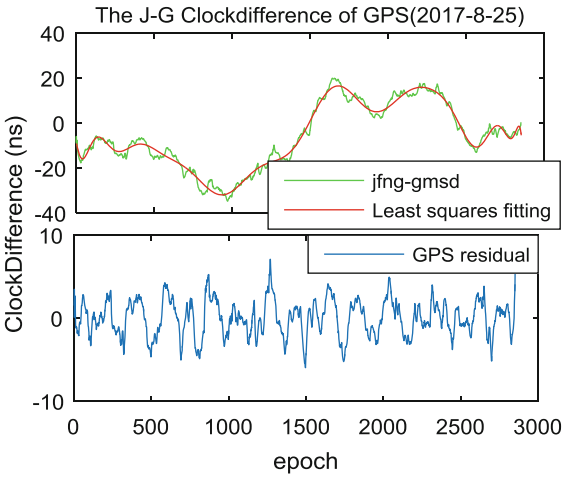


Fig. 2. Residuals of Clock error between JFNG and GMSD based on GPS PPP

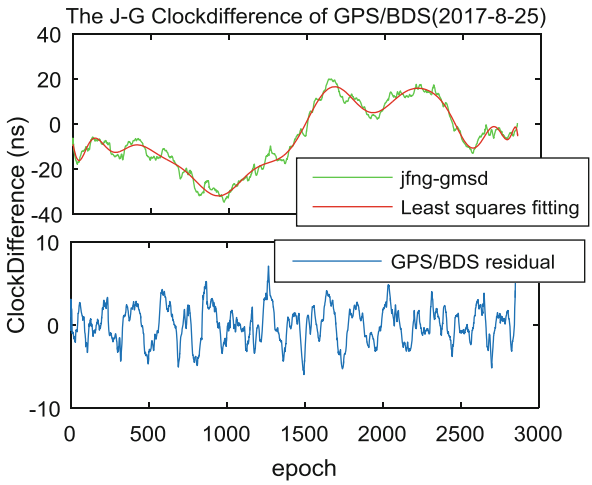


Fig. 3. Residuals of Clock error between JFNG and GMSD based on GPS/BDS PPP

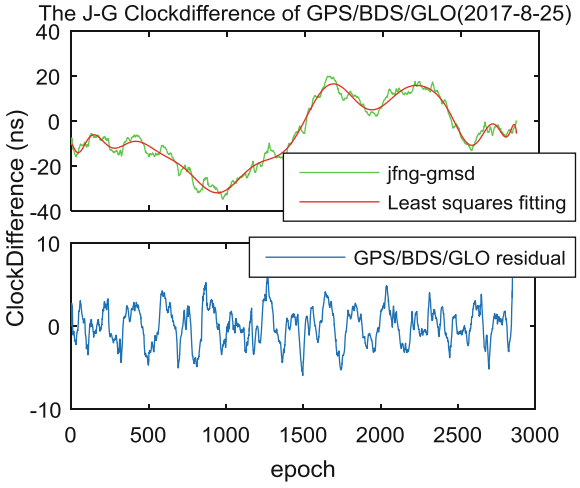


Fig. 4. Residuals of Clock error between JFNG and GMSD based on GPS/BDS/GLO PPP

The waveforms of inter-station time transfer and residual waveforms based on the above multi-system PPP calculation are shown in Table 1. The following conclusions can be drawn:

The precision of PPP time comparison of multiple system combinations is equivalent, and the accuracy of all four combinations is better than 0.22 ns, and GPS/BDS and GPS/BDS/GLO three systems are better than GPS single system. Four systems have no obvious improvement over three systems. This is because there are few observable GAL data.

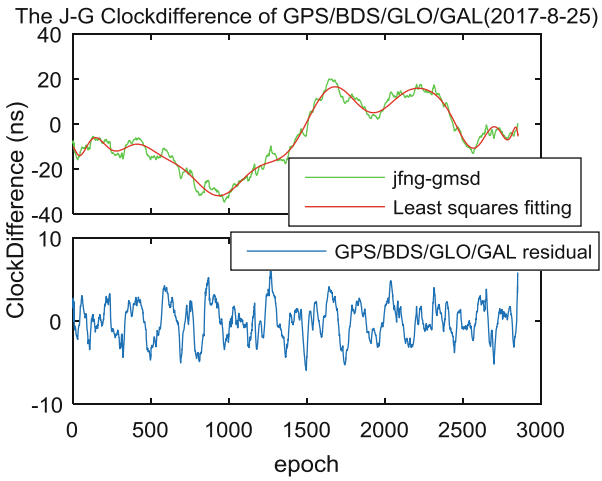


Fig. 5. Residuals of Clock error between JFNG and GMSD based on GPS/BDS/GLO/GAL PPP

Table 1. Internal accuracy analysis results of quad-constellation GNSS PPP

TYPE	Max (ns)	Min (ns)	Mean (ns)	Std (ns)
GPS	7.0993	−5.9674	5.43E−5	0.2134
GPS/BDS	7.0991	−5.9586	5.22E−5	0.2133
GPS/BDS/GLO	7.0999	−5.9397	4.93E−5	0.1287
GPS/BDS/GLO/GAL	7.0999	−5.9402	5.02E−5	0.1284

4.2 External Accuracy Results

To evaluate the effectiveness of the self-research algorithm, the most effective method to measure the external conformance is to compare with the result of the clock difference between stations provided by IGS. The result provided by IGS is the weighted average of the results of other processing centers [9], which is the most accurate and reliable GPS data product at present, and can be compared as truth value. In this paper, the clock difference of measuring station provided by GBM center is taken as the true value, and the clock difference between jfng-gmsd stations calculated by ourselves is relatively poor, and the clock difference residual sequence is obtained, and the standard deviation of the residual sequence is calculated, as shown in Table 2.

It can be seen from the above figures and tables that the standard deviation of IGS clock residuals calculated based on the combined PPP of the four systems is on the order of sub-nanoseconds, and the small standard deviations indicate that the comparison results are stable and reliable algorithm software. According to the above four consecutive days of experiments, it can be found that the maximum difference between the inter-station clock difference calculated by the combination of the four systems and the inter-station clock difference provided by IGS is no more than 2 ns (Figs. 6 and 7).

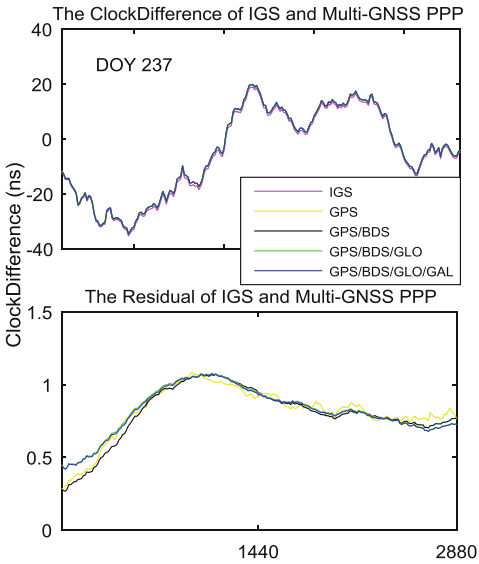


Fig. 6. Waveform and difference between Multi-GNSS PPP clock error and IGS clockerror (2017-8-25)

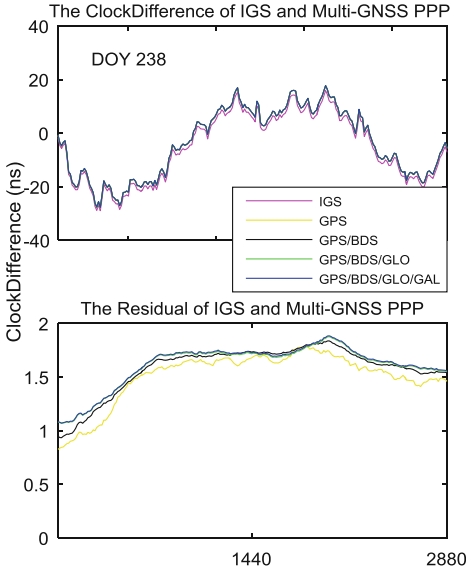


Fig. 7. Waveform and difference between Multi-GNSS PPP clock error and IGS clockerror (2017-8-26)

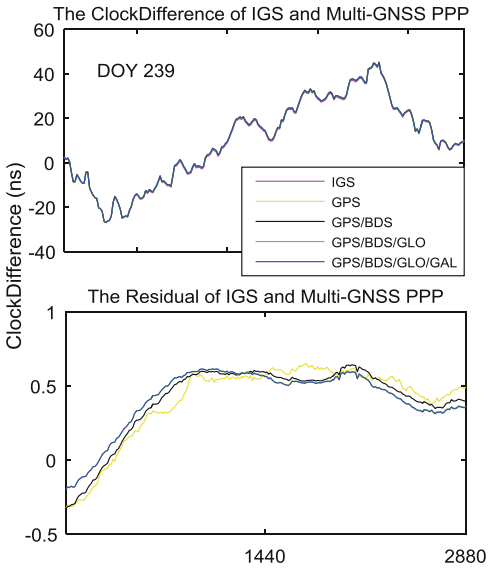


Fig. 8. Waveform and difference between Multi-GNSS PPP clock error and IGS clockerror (2017-8-27)

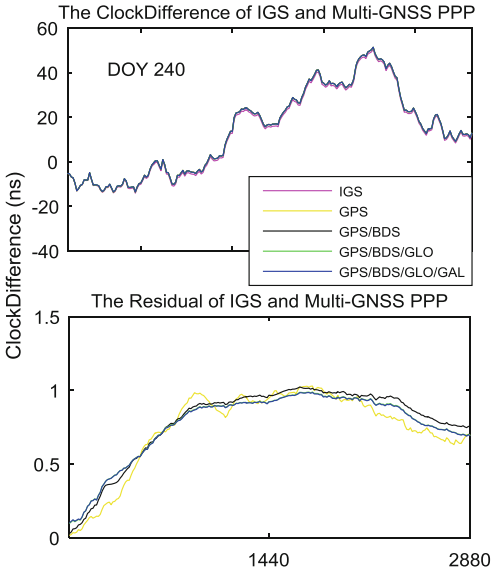


Fig. 9. The waveform and difference between Multi-GNSS PPP clockerror and IGS clockerror (2017-8-28)

Table 2. External accuracy analysis of JFNG-GMSD multi-system PPP time transfer(ns)

JFNG-GMSD	GPS/BDS	GPS/BDS/GLO	GPS/BDS/GLO/GAL
DOY 237	0.194	0.165(7.3%)	0.165
DOY 238	0.223	0.199(17.1%)	0.196
DOY 239	0.246	0.210(18.9%)	0.209
DOY 240	0.256	0.231(15.4%)	0.229

Except for the decrease of compliance accuracy after DOY added BDS in 237 days, the accuracy was improved after adding BDS data in the other three days. It is believed that DOY 237 May be an optimization of the network structure after the addition of BDS system, but the GPS data quality of this day is better than that of BDS data, so the accuracy decreases after the addition of BDS. GPS/BDS/GLO showed a significant improvement in precision compared with GPS single system and GPS/BDS dual system, and the maximum accuracy was up to 18.9% compared with single system. The accuracy improvement of four systems is not obvious compared with three systems (Figs. 8 and 9).

4.3 Stability Analysis

The stability of time transfer is also an important index for the evaluation of time transfer. In this paper, Allan variance formula was used to analyze the bell difference stability between PPP and IGS stations, as shown in Figs. 10, 11, 12 and 13. The results of the bell difference stability were calculated with 2017-8-25 as an example, as shown in Table 3.

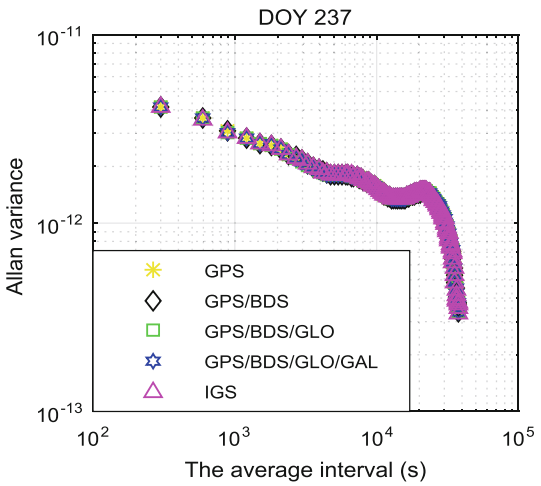


Fig. 10. Time transfer stability of Multi-GNSS PPP clock error and IGS clockerror (2017-8-25)

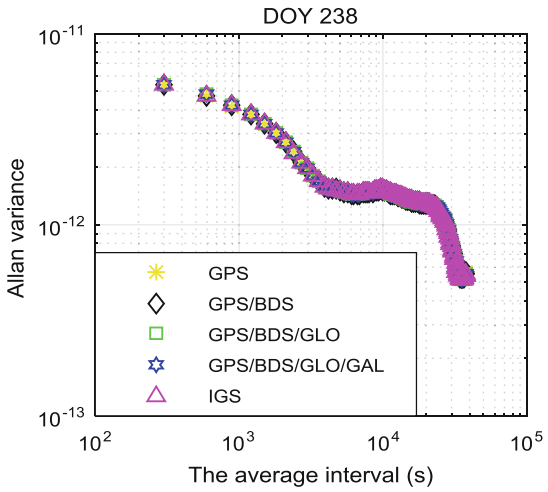


Fig. 11. Time transfer stability of Multi-GNSS PPP clock error and IGS clockerror (2017-8-26)

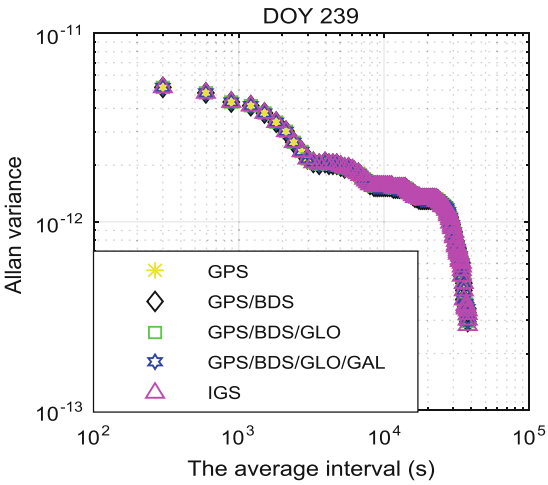


Fig. 12. Time transfer stability of Multi-GNSS PPP clock error and IGS clockerror (2017-8-27)

As can be seen from the above graph, the clock stability trend of the multi-system combination calculation is completely consistent with the IGS clock stability. The value notes that the calculation accuracy of the combination of GPS and GPS/BDS/GLO is slightly better than the clock difference stability of IGS between 4500 s and 18900 s, which may be because the clock difference is smoothed to improve the stability of the clock difference. In half a day, the frequency stability can reach 1×10^{-13}–4×10^{-13}. Within one day, the frequency stability can reach 5×10^{-13}–6×10^{-13}.

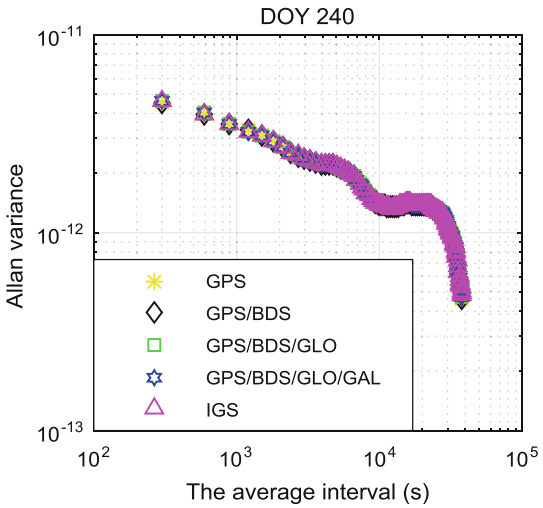


Fig. 13. Time transfer stability of Multi-GNSS PPP clock error and IGS clockerror (2017-8-28)

Table 3. Result of clockerror stability base on IGS and Multi-GNSS PPP

Time (s)	JFNG-GMSD				
	IGS	GPS	GPS/BDS	GPS/BDS/GLO	GPS/BDS/GLO/GAL
300	4.1597E−12	4.1713E−12	4.1733E−12	4.1785E−12	4.1784E−12
900	3.0442E−12	3.0611E−12	3.0652E−12	3.0512E−12	3.0512E−12
2100	2.4552E−12	2.4556E−12	2.4593E−12	2.4493E−12	2.4493E−12
4500	1.8318E−12	1.8312E−12	1.8313E−12	1.8241E−12	1.8240E−12
9300	1.5913E−12	1.5898E−12	1.5929E−12	1.5910E−12	1.5910E−12
18900	1.4394E−12	1.4347E−12	1.4367E−12	1.4415E−12	1.4417E−12
38100	3.3418E−13	3.3563E−13	3.3481E−13	3.3454E−13	3.3449E−13
76200	6.2323E−13	6.2432E−13	6.2422E−13	6.2369E−13	6.2355E−13

5 Conclusion

Based on the inter-station clock difference between IGS JFNG and GMSD, this paper calculates the inter-station clock difference by PPP combination of multiple constellations, and the results show that:

1. Multi-mode GNSS precision single point positioning can complete the time transfer of the accuracy of nanosecond. Considering from the external and internal accuracy, the accuracy of three-system system is improved greatly compared with single-system and dual-system, while the time transfer accuracy of four-system system is not improved much or even reduced on the basis of three-system.
2. In terms of time transfer frequency stability, the clock difference solution calculated by multimode GNSS PPP can maintain a certain consistency with IGS clock

difference. At some time, the frequency stability of PPP difference solution is slightly better than that of IGS difference, but with the passage of time in the day, the frequency stability of PPP difference solution is gradually consistent with that of IGS difference product.

As BDS satellite navigation system begins to be put into global use, the development of GNSS multi-constellation PPP timing research may, on the one hand, promote the application of PPP time transfer technology in the comparison of long-distance and high-precision time in China, and on the other hand, lay a foundation for participating in the international PPP remote time comparison.

Acknowledgement. Funding from National Key R&D Program of China (No. 2016YFB050 1803), National Natural Science Foundation of China (No. 41674039) and Joint teacher-student innovation and entrepreneurship project at Central South University (No. 2018gczd005) are appreciated.

References

1. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solutions 5(2):12–28
2. Ray J, Senior K (2003) IGS/BIPM pilot project: GPS carrier phase for time/frequency transfer and timescale formation. Metrologia 40(3):270–288
3. Costa R, Orgiazzi D, Pettiti V, Sesia I (2004) Performance comparison and stability characterization of timing and geodetic GPS receivers at IEN. In: European frequency and time forum, EFTF 2004, pp 279–286
4. Zhang X, Cai S, Li X (2010) Accuracy analysis of time and frequency transfer based on precise point positioning. Geomat Inf Sci Wuhan Univ 35(3):274–278
5. Zhang R, Liu H, Shu B (2017) Research on time transfer: based on BDS precise point positioning and accuracy comparison. Geod Geodyn 37(10):1070–1073
6. Defraigne P, Harmegnies A, Petit G (2010) Time and frequency transfer combining GLONASS and GPS data. In: Joint conference of the IEEE international frequency control and the european frequency and time forum, vol 47. IEEE, pp 1–5
7. Liang K, Zhang A, Pei C, Zuo F, Jin Z (2014) Preliminary implementation of time and frequency transfer by BDS. In: European frequency and time forum and international frequency control symposium. IEEE, pp 517–520
8. Cai C, Gao Y, Pan L, Zhu J (2015) Precise point positioning with quad-constellations: GPS, BeiDou, GLONASS and Galileo. Adv Space Res 56(1):133–143
9. Guo M, Lu H, Xiao Y (2016) Accuracy analysis of SPP and PPP time transfer methods based on BDS and GPS. In: China satellite navigation academic annual meeting

Evaluation of High Precision Time Transfer in the Relativity Framework

Keliang He[1], Yuling He[1(✉)], Ran Cheng[2,3], and Guoyong Wang[1]

[1] China Academy of Space Technology (Xi'an), Xi'an 710100, China
kewenli@yeah.net, heyuling_cast@163.com
[2] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China
[3] University of Chinese Academy of Sciences, No. 19 Yuquan Road, Beijing 100049, China

Abstract. In this paper, we systematically discussed and analyzed the uncertainties of four typical time transfer links, including the satellite-ground two-way link, the satellite-ground lambda link, the ground-ground TWSTFT link and the ground-ground double-lambda link. Combining with the estimations of the time delays that predicted by relativistic theory for one-way transfer and the derivations of the clock differences of four typical links, we found that all the maximum uncertainty terms of the four links are the order of c^{-2} if we do not consider the time delays between receiving and emitting signals on the satellite itself. Specifically, for the satellite-ground lambda link, the clock difference of order c^{-2} is uncorrelated with the satellite velocities, so that the clock difference uncertainties is superior to the satellite-ground two-way transfer link. Similarly, the ground-ground double-lambda link is better than the ground-ground TWSTFT transfer link. Meantime, the accuracy of the links arising from the relativistic time delays are also discussed.

Keywords: Time transfer · Relativistic time delay · Lambda configuration link · Two-way link

1 Introduction

The purpose of time or frequency transfer between two distant stations is to realize the time and frequency measurements, comparisons and time synchronizations over a long distance. When considering time and frequency transfer in the framework of general relativity, it is necessary firstly to derive the time delay and frequency shift of the one-way transfer of electromagnetic wave to high order terms (series expansion of c^{-1}). Then, these terms should be substituted into the calculation model of clock difference or frequency difference of different links. And the terms affecting the accuracy and stability specifications should be retained. For one-way transfer in an isolated, axisymmetic rotating system, the time-dilation effects to order 10^{-18} in the vicinity of earth was already estimated [1]. The time delay and frequency shift terms to order c^{-3} were derived when considering only the Sagnac effect and the earth's mass potential [2]. Moreover, the time delay and frequency shift terms to order c^{-4} were also derived

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 544–555, 2019.
https://doi.org/10.1007/978-981-13-7751-8_52

based on the world-function method [3–5], which make kinds of time delays and frequency shifts to order 10^{-18} can be calculated. These time delays and frequency shifts are attributed to the earth's mass and spin multipole moments, the tidal potential and the inertial potential. These works provide theoretical basis for considering the complex links of time and frequency transfer. The calculation models of clock difference depend on the link configurations, which affect the link uncertainty and accuracy. The two-way satellite time and frequency transfer (TWSTFT) link is widely used for time synchronizations between time-keeping laboratories all over the world. The clock difference of the TWSTFT link was already derived to order c^{-2} [6], and the transfer accuracy can reach sub-ns [7, 8]. The ACES project projected by ESA adopted the lambda link firstly used by the gravity probe A (GP-A) experiment [9]. The satellite-ground clock difference was also derived to order c^{-2} and the time stability was expected to reach picosecond level. Specifically, the requirement of orbit determination was also discussed [10, 11]. The clock difference uncertainty of two ground stations in the T2L2 (time transfer by laser link) experiment developed by CNES and OCA was also assessed to order c^{-2} [12], the uncertainty can drop to hundreds of picoseconds and the time stability can reach picosecond level at present [13, 14].

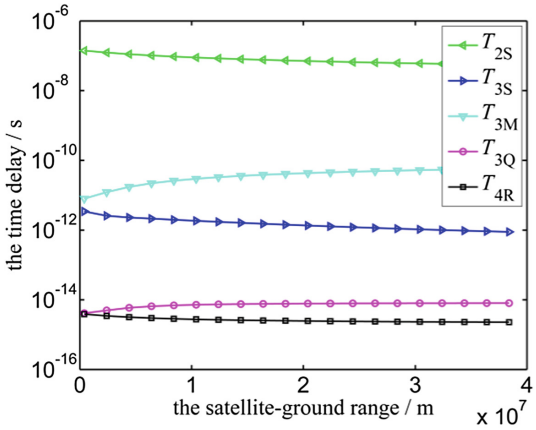
In this paper, we attempts to calculate systematically the clock differences and evaluate the uncertainties of the four typical links of time transfer. In order to determine the dominant uncertainty term, various time delays and frequency shifts are estimated firstly. Then, we calculate respectively the clock difference of the four links, and estimate the uncertainty limits. It is also discussed that which terms of the relativistic time delays should be included for assessing link accuracy.

2 The Time Delays and Frequency Shifts of One-Way Transfer and the Maximum Uncertainty Term

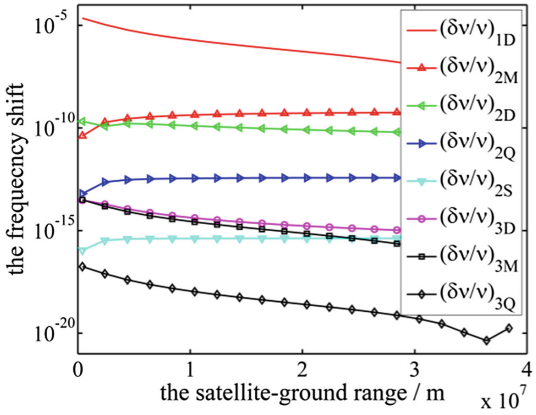
According to general relativity, time and space are not independent. Four-dimensional coordinates are defined as $x^{\alpha} = (ct + x^{i}) = (ct, \boldsymbol{x})$, and four-dimensional space-time interval is defined as $ds^2 = -c^2 d\tau^2 = g_{\alpha\beta} dx^{\alpha} dx^{\beta}$, where $g_{\alpha\beta}$ is the metric tensor, τ is the proper time in the local inertial reference system and t is the coordinate time in classical space-time. Metric tensor is a function of gravitational potentials, which indicates the gravitational effect on space-time. It can be derived by approximately solving Einstein field equations with symmetry and boundary conditions. As the space-time curvature is determined by the distribution of matter, the propagation path of electromagnetic wave is the solution of the null geodesic equation, which can be approximated as a curve with a slight deviation with respect to the straight line in Euclidean space under weak field approximation [15]. Blanchet et al. derived the time delays and frequency shifts of one-way transfer to order c^{-4} term but spared the trouble of integrating of the null geodesic by using the world function method. His model is still the highest precision processing method at present. We adopt the multipolar expansions of earth's gravitational potential and Blanchet's analytical form of time delay and frequency shifts [3] to calculate numerically the variations of the satellite-ground one-way transfer delays and frequency shifts with given satellite orbit altitude,

as shown in Fig. 1. For the convenience of analysis, it is supposed that both satellite and ground station are in the equatorial plane, the angle between satellite and ground station is 15° relative to earth's barycentre, and the orbit of satellite is circular. The Arabic numerals m in the subscript represents that the time delay or frequency shift term is linear form of c^{-m}; the S, M, Q and R in the subscript denote respectively the Sagnac delay caused by the motion of station, the monopolar potential delay (Shapiro delay), the quadrupole moment delay and the earth rotation delay; D, M, Q and S in the frequency shift subscript represent respectively the frequency shift caused by doppler effect, monopolar potential, quadrupole moment and hexadecapole moment.

As can be seen from Fig. 1, the time delay or frequency shift term with lower order will be more sensitive to the variation of satellite-ground range (except for the second order multipole moment frequency shift). Similarly, the term with lower order will be



(a) time delay



(b) frequency shift

Fig. 1. The absolute values of time delay and frequency shift of electromagnetic wave one-way transfer between satellite and ground station

more sensitive to the variation of velocity. In turn, it indicates that when the accuracy of satellite orbit determination and ground station position are invariable, the contribution to uncertainty from the lower order terms is larger than that of higher terms. Therefore, when calculating the constraints on station kinematic parameters according to the link uncertainty requirement, only the lowest order terms of clock difference or frequency difference need to be considered. The lowest order term of the lock difference or frequency difference of one-way transfer is the order c^{-1} term and that in other configuration transfer links is the order c^{-2} term without considering the time delays between the reception and emission of signals on the satellite. Therefore, when assessing the uncertainty of the actual time and frequency transfer links, the order c^{-2} terms need to be considered primarily.

3 The Satellite-Ground Time Transfer Links

Considering the one-way time transfer from A to B to order c^{-3}, the transfer time is given by

$$\begin{aligned} T_{AB} = {} & \frac{D_{AB}(t_A)}{c} + \frac{\boldsymbol{D}_{AB}(t_A)\cdot\boldsymbol{v}_B(t_A)}{c^2} + \frac{D_{AB}(t_A)}{2c^3}\left[|\boldsymbol{v}_B(t_A)|^2 + \boldsymbol{D}_{AB}(t_A)\cdot\boldsymbol{a}_B(t_A)\right. \\ & \left. + \left(\frac{\boldsymbol{D}_{AB}(t_A)\cdot\boldsymbol{v}_B(t_A)}{D_{AB}(t_A)}\right)^2\right] + \frac{2GM}{c^3}\ln\left(\frac{r_A(t_A)+r_B(t_A)+D_{AB}(t_A)}{r_A(t_A)+r_B(t_A)-D_{AB}(t_A)}\right) \\ & + (T_{J2})_{AB} + \Delta T^{iono}_{AB} + \Delta T^{tropo}_{AB} \end{aligned} \tag{1}$$

where $\boldsymbol{D}_{AB}(t_A) = \boldsymbol{X}_B(t_A) - \boldsymbol{X}_A(t_A)$, $\boldsymbol{v}$ and $\boldsymbol{a}$ are the velocity and acceleration respectively, t_A is the emitting time, ΔT^{iono}_{AB} and ΔT^{tropo}_{AB} are the ionosphere delay and troposphere delay respectively. The second and third terms are respectively the order c^{-2} term and order c^{-3} term of the Sagnac delay, the fourth term is the order c^{-3} term of the Shapiro time delay; $(T_{J2})_{AB}$ is the time delay caused by the earth's quadrupole moment and it is given by Eq. (61) in [3].

3.1 The Satellite-Ground Two-Way Time Transfer Link

The satellite-ground two-way transfer link is as shown in Fig. 2. τ^S and τ^A denote respectively the proper time of satellite and earth station, ideally, $t_1 = t_3$, $\tau^S(t_3) = \tau^A(t_1)$, but actually there is a proper time difference with respect to the same coordinate time and it is defined by $\Delta\tau = \tau^S(t_1) - \tau^A(t_1)$. The propagation times observed in the local reference system of the satellite have the expressions

$$\begin{cases} \tau^{\mathrm{S}}(t_2)-\tau^{\mathrm{S}}(t_1)=\tau^{\mathrm{S}}(t_2)-\tau^{\mathrm{S}}(t_3)+\tau^{\mathrm{S}}(t_3)-\tau^{\mathrm{S}}(t_1)=\tau^{\mathrm{S}}(t_2)-\tau^{\mathrm{S}}(t_3)+\tau^{\mathrm{A}}(t_1)-\tau^{\mathrm{S}}(t_1) \\ \qquad\qquad =\Delta\tau^{\mathrm{S}}-\Delta\tau=\left(\Delta T_1^{\mathrm{A}}+T_{12}+\Delta T_2^{\mathrm{S}}\right)^{\mathrm{S}} \\ \tau^{\mathrm{S}}(t_4)-\tau^{\mathrm{S}}(t_3)=\tau^{\mathrm{S}}(t_4)-\tau^{\mathrm{S}}(t_1)+\tau^{\mathrm{S}}(t_1)-\tau^{\mathrm{S}}(t_3)=\tau^{\mathrm{S}}(t_4)-\tau^{\mathrm{S}}(t_1)+\tau^{\mathrm{S}}(t_1)-\tau^{\mathrm{A}}(t_1) \\ \qquad\qquad =\left(\Delta\tau^{\mathrm{A}}\right)^{\mathrm{S}}+\Delta\tau=\left(\Delta T_4^{\mathrm{A}}+T_{34}+\Delta T_3^{\mathrm{S}}\right)^{\mathrm{S}} \end{cases} \tag{2}$$

where ΔT_1^{A} and ΔT_3^{S} are respectively the instrument time delays on the emitting segment, ΔT_2^{S} and ΔT_4^{A} are respectively the instrument time delays on the receiving segment, $(\Delta\tau^{\mathrm{A}})^{\mathrm{S}}$ and $(\Delta T^{\mathrm{A}})^{\mathrm{S}}$ represent time interval on the satellite corresponding to the time interval observed in the local reference system of the ground station and vice versa. The time interval between the reception and emission of signals on the satellite is represented by $\Delta\tau^{\mathrm{S}}=\tau^{\mathrm{S}}(t_2)-\tau^{\mathrm{S}}(t_3)$, corresponding observation in the ground station is represented by $\Delta\tau^{\mathrm{A}}=\tau^{\mathrm{A}}(t_4)-\tau^{\mathrm{A}}(t_1)$. From Eq. (2), the satellite-ground clock difference can be given by

$$\Delta\tau=\frac{1}{2}[\Delta\tau^{\mathrm{S}}-(\Delta\tau^{\mathrm{A}})^{\mathrm{S}}+(\Delta\tau_4^{\mathrm{A}}+T_{34}+\Delta\tau_3^{\mathrm{S}})^{\mathrm{S}}-(\Delta T_1^{\mathrm{A}}+T_{12}+\Delta T_2^{\mathrm{S}})^{\mathrm{S}}] \tag{3}$$

Considering T_{12} and T_{34} can be expressed using Eq. (1), and expressing $\boldsymbol{D}_{\mathrm{SA}}(t_3)$ and $\boldsymbol{v}_{\mathrm{A}}(t_3)$ in terms of $\boldsymbol{D}_{\mathrm{AS}}(t_1)$, $\boldsymbol{v}_{\mathrm{A}}(t_1)$, $\boldsymbol{v}_{\mathrm{S}}(t_1)$, $\boldsymbol{a}_{\mathrm{A}}(t_1)$ and $\boldsymbol{a}_{\mathrm{S}}(t_1)$, $T_{34}-T_{12}$ in Eq. (3) can be given by

$$\begin{aligned} T_{34}-T_{12}\approx{}& \frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{v}_{\mathrm{SA}}(t_1)\Delta t}{D(t_1)\mathrm{c}}+\frac{\boldsymbol{v}_{\mathrm{SA}}(t_1)^2\Delta t^2}{2D(t_1)\mathrm{c}}+\frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{a}_{\mathrm{SA}}(t_1)\Delta t^2}{2D(t_1)\mathrm{c}} \\ &-\frac{[\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{v}_{\mathrm{SA}}(t_1)]^2\Delta t^2}{D_{\mathrm{AS}}(t_1)^3\mathrm{c}}-\frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot[\boldsymbol{v}_{\mathrm{S}}(t_1)+\boldsymbol{v}_{\mathrm{A}}(t_1)]}{\mathrm{c}^2} \\ &+\frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{a}_{\mathrm{A}}(t_1)\Delta t+\boldsymbol{v}_{\mathrm{SA}}(t_1)\cdot\boldsymbol{v}_{\mathrm{A}}(t_1)\Delta t+\boldsymbol{v}_{\mathrm{SA}}(t_1)\cdot\boldsymbol{a}_{\mathrm{A}}(t_1)\Delta t^2}{\mathrm{c}^2} \\ &+O(c^{-3})+\Delta T_{34}^{iono}+\Delta T_{34}^{\mathrm{tropo}}-\Delta T_{12}^{iono}-\Delta T_{12}^{\mathrm{tropo}} \end{aligned} \tag{4}$$

where $\Delta t=t_3-t_1$, the subscript SA denotes the same vector subtraction of the satellite and the ground station. When the initial synchronization makes the clock difference between the satellite and the ground station less than 10 ns, the second-order terms of Δt in the c^{-1} term and the first-order terms of Δt in the c^{-2} term can be ignored. In addition,$\Delta\tau=\tau^{\mathrm{S}}(t_1)-\tau^{\mathrm{A}}(t_1)=-\gamma\Delta t$, where γ is the time dilation factor. Inserting Eq. (4) into Eq. (3) and using iterative calculation, the satellite-ground clock difference is given by

$$\Delta\tau \approx -\frac{1}{2}[\Delta\tau^{\mathrm{S}} - (\Delta\tau^{\mathrm{A}})^{\mathrm{S}} + (\Delta T_4^{\mathrm{A}} + \Delta T_3^{\mathrm{S}})^{\mathrm{S}} - (\Delta T_1^{\mathrm{A}} + \Delta T_2^{\mathrm{S}})^{\mathrm{S}} + \Delta T_{34}^{iono} + \Delta T_{34}^{\mathrm{tropo}} - \Delta T_{12}^{iono} - \Delta T_{12}^{\mathrm{tropo}}] \times [1 - \frac{1}{2\gamma}\frac{\boldsymbol{N}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{v}_{\mathrm{S}}(t_1)}{\mathrm{c}}] - \frac{1}{2}\frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot[\boldsymbol{v}_{\mathrm{s}}(t_1)+\boldsymbol{v}_{\mathrm{A}}(t_1)]}{\mathrm{c}^2} + O(\mathrm{c}^{-3}) \tag{5}$$

Without considering the ionosphere delay and troposphere delay, the uncertainty of the satellite-ground clock difference is written as

$$-\frac{1}{2}\delta\frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot[\boldsymbol{v}_{\mathrm{S}}(t_1)-\boldsymbol{v}_{\mathrm{A}}(t_1)]}{\mathrm{c}^2} \tag{6}$$

The clock difference uncertainty is correlated with the uncertainty of satellite-ground distance vector $\boldsymbol{D}_{\mathrm{AS}}(t_1)$, the speed vectors $\boldsymbol{v}_{\mathrm{A}}(t_1)$ and $\boldsymbol{v}_{\mathrm{S}}(t_1)$ of the ground station and satellite. For instance, considering the time transfer between a geostationary satellite and a ground station and supposing that the orbit determination accuracy is 1 m, the velocity uncertainty is 0.01 m/s and the ground station velocity uncertainty is 1×10^{-5} m/s, the maximum time uncertainty is about 0.4 ps.

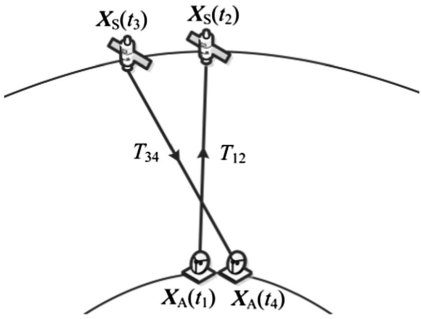


Fig. 2. The satellite-ground two-way time transfer

In order to investigate which time delay terms should be included for assessing link accuracy, the differences of the Shapiro time delay, the third-order Sagnac delay and the third-order quadrupole delay between the up-link and the down-link are calculated respectively. Where the coordinate time difference between the emission times of the satellite and the ground station is set to be 0.1 s, the result is as shown in Fig. 3. It can be seen that the second and third order terms of the Sagnac delay and the Shapiro delay should be taken into account, while the quadrupole moment delay and earth rotation delay do not if assessing the accuracy of the two-way time transfer at femtosecond level (the orbit altitude is less than 40 000 km).

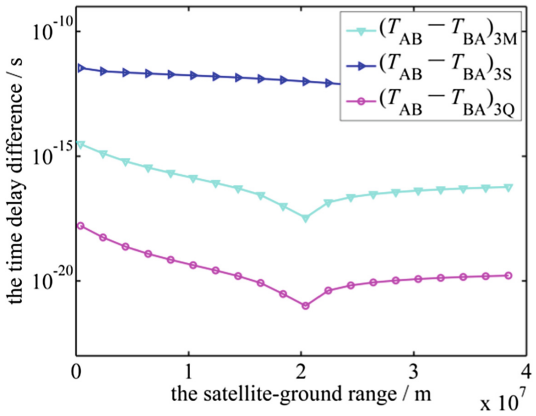


Fig. 3. The third-order time delay difference between the up-link and the down-link in the satellite-ground two-way transfer link

3.2 The Satellite-Ground Lambda (Λ) Configuration Time Transfer Link

The time transfer link of lambda configuration is shown in Fig. 4. In contrast to the satellite-ground two-way transfer link, the difference is that the satellite firstly receives the signal at the time t_2 and then emits a signal at the time t_3. Adopting the similar calculation method used in Sect. 3.1, we get the clock difference

$$\begin{aligned}\Delta\tau = \frac{1}{2}\big[&-(\tau^{\mathrm{A}}(t_4)+\tau^{\mathrm{A}}(t_1))+\tau^{\mathrm{S}}(t_2)+\tau^{\mathrm{S}}(t_3)+(\Delta T_4^{\mathrm{A}}+\Delta T_3^{\mathrm{S}})^{\mathrm{A}}\\ &-(\Delta T_1^{\mathrm{A}}+\Delta T_2^{\mathrm{S}})^{\mathrm{A}}+(T_{34}-T_{12})^{\mathrm{A}}\big]\end{aligned} \tag{7}$$

where

$$\begin{aligned}T_{34}-T_{12}\approx{}&\frac{(\boldsymbol{v}_{\mathrm{S}}-\boldsymbol{v}_{\mathrm{A}})\cdot\boldsymbol{N}_{\mathrm{AS}}T_{23}}{\mathrm{c}}-\frac{2\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{v}_{\mathrm{A}}(t_1)}{\mathrm{c}^2}+O(\mathrm{c}^{-2})\\ &+O(T_{23}^2)+\Delta T_{34}^{iono}+\Delta T_{34}^{\mathrm{tropo}}-\Delta T_{12}^{iono}-\Delta T_{12}^{\mathrm{tropo}}\end{aligned} \tag{8}$$

where $\boldsymbol{N}_{\mathrm{AS}}=[\boldsymbol{X}_{\mathrm{S}}(t_2)-\boldsymbol{X}_{\mathrm{A}}(t_1)]/|\boldsymbol{X}_{\mathrm{S}}(t_2)-\boldsymbol{X}_{\mathrm{A}}(t_1)|$, T_{23} is the time delay between t_2 and t_3. Without considering the ionosphere, troposphere and equipment delays and supposing that T_{23} is very small, the uncertainty of the satellite-ground clock difference is given by

$$-\delta\frac{\boldsymbol{D}_{\mathrm{AS}}(t_1)\cdot\boldsymbol{v}_{\mathrm{A}}(t_1)}{\mathrm{c}^2} \tag{9}$$

Similarity, the uncertainty of clock difference is correlated with the satellite-ground distance vector $\boldsymbol{D}_{\mathrm{AS}}(t_1)$ and the speed vector $\boldsymbol{v}_{\mathrm{A}}(t_1)$ of the ground station. For instance, considering the time transfer between the geostationary satellite and the ground station and supposing that the uncertainty of satellite orbit determination and the ground

station position are the same as that in Sect. 3.1, the maximum uncertainty of clock difference is about 6.8 fs. The time delay terms needed to be taken into account for assessing the link accuracy is similar to that of the two-way transfer link.

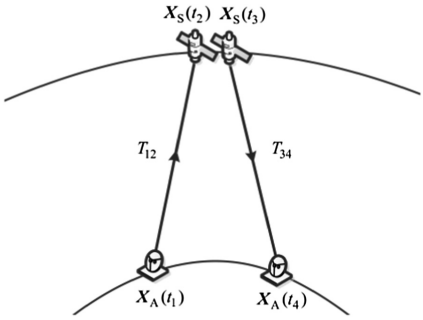


Fig. 4. The satellite-ground time transfer link of lambda configuration

In the two kinds of satellite-ground time transfer schemes mentioned above, the order c^{-2} term of clock difference contribute to the greatest time uncertainty. The difference is that the clock difference of the lambda configuration scheme is not relative with the satellite speed, while the clock difference of the two-way satellite-ground scheme is relative with the satellite speed. Because the velocity uncertainty of satellite is much greater than that of ground station, so the clock difference uncertainty of lambda configuration transfer link is smaller.

4 The Ground-Ground Time Transfer Links Transmitted by Satellite

The TWSTFT link and the double lambda link are two kinds of high accuracy transfer links between two ground stations. Both need a satellite to transmit the time signals. The difference is that the satellite is just as a signal transponder and do not need a clock to record the signal arriving times and emitting times in the TWSTFT link, while there is a clock to record the signal arriving times and emitting times on the satellite in the double lambda link.

4.1 The Two-Way Satellite Time and Frequency Transfer (TWSTFT) Link

The TWSTFT link is shown in Fig. 5. Adopting a similar calculation method used in Sect. 3.1, we can get the clock difference of the two ground stations

$$\Delta\tau = \frac{1}{2}(\Delta\tau^{\mathrm{C}} - \Delta\tau^{\mathrm{D}}) + \frac{1}{2}(T_{01}^{\mathrm{C}} + T_{12}^{\mathrm{C}} + \Delta T_{\mathrm{r}} - T_{\Delta 3}^{\mathrm{D}} - T_{34}^{\mathrm{D}} - \Delta T_{\mathrm{l}}) \tag{10}$$

where ΔT_r and ΔT_l denote respectively the equipment delays from C to D and from D to C. Without considering the ionospheric, tropospheric and equipment delays, the uncertainty of the clock difference between the two ground stations is given by

$$\begin{aligned}\frac{1}{2}\delta\Bigg\{&\frac{[\boldsymbol{D}_{\mathrm{CS}}(t_0) - \boldsymbol{D}_{\mathrm{DS}}(t_0)]\cdot\boldsymbol{v}_{\mathrm{S}}(t_0)}{\mathrm{c}^2} + \frac{\boldsymbol{D}_{\mathrm{SD}}(t_0)\cdot\boldsymbol{v}_{\mathrm{D}}(t_0)}{\mathrm{c}^2} - \frac{\boldsymbol{D}_{\mathrm{SC}}(t_0)\cdot\boldsymbol{v}_{\mathrm{C}}(t_0)}{\mathrm{c}^2}\\ &- \frac{[\boldsymbol{N}_{\mathrm{SD}}(t_0)\cdot\boldsymbol{v}_{\mathrm{SD}}(t_0)]D_{\mathrm{CS}}(t_0)}{\mathrm{c}^2} + \frac{[\boldsymbol{N}_{\mathrm{SC}}(t_0)\cdot\boldsymbol{v}_{\mathrm{SC}}(t_0)]D_{\mathrm{DS}}(t_0)}{\mathrm{c}^2}\Bigg\}\end{aligned} \tag{11}$$

Similarly, the clock difference uncertainty is correlated with the satellite-ground distance vector, the speed vectors of the ground station and the satellite. For the time transfer between geostationary satellites and the ground, supposing that the uncertainty of satellite orbit determination and the ground station position are the same as that in Sect. 3.1, the maximum uncertainty of clock difference is about 1.6 ps.

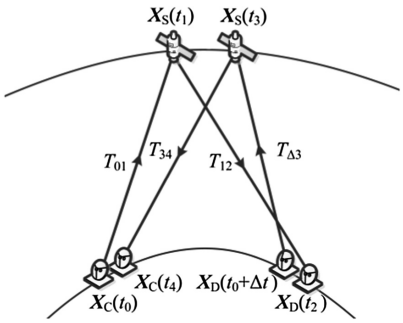


Fig. 5. The time transfer link of TWSTFT

4.2 The Double Lambda Configuration Time Transfer Link

The time transfer link of double lambda configuration is shown in Fig. 6. Adopting a similar calculation method used in Sect. 3.1, we get the clock difference of two ground stations

$$\begin{aligned}\Delta\tau = &-\frac{1}{2}[-(\tau^{\mathrm{C}}(t_0) + \tau^{\mathrm{C}}(t_2)) + 2\tau^{\mathrm{S}}(t_1) + T_{12}^{\mathrm{C}} - T_{01}^{\mathrm{C}} + \Delta T_{\mathrm{l}}]\\ &+ \frac{1}{2}[-(\tau^{\mathrm{D}}(t_0 + \Delta t) + \tau^{\mathrm{D}}(t_4)) + 2\tau^{\mathrm{S}}(t_3) + T_{34}^{\mathrm{D}} - T_{\Delta 3}^{\mathrm{D}} + \Delta T_{\mathrm{r}}]\\ &+ [\tau^{\mathrm{C}}(t_3) - \tau^{\mathrm{C}}(t_1)] - [\tau^{\mathrm{S}}(t_3) - \tau^{\mathrm{S}}(t_1)]\end{aligned} \tag{12}$$

Without considering ionospheric, tropospheric and equipment delays, the uncertainty of clock difference between the two ground stations is given by

$$\delta\left[\frac{\boldsymbol{D}_{\mathrm{CS}}(t_0)\cdot\boldsymbol{v}_{\mathrm{C}}(t_0)}{\mathrm{c}^2}-\frac{\boldsymbol{D}_{\mathrm{DS}}(t_0+\Delta t)\cdot\boldsymbol{v}_{\mathrm{D}}(t_0+\Delta t)}{\mathrm{c}^2}\right] \tag{13}$$

Similarly, for the time transfer between the geostationary satellite and the ground, supposing that the uncertainty of satellite orbit determination and the ground station position are the same as that in Sect. 3.1, the maximum uncertainty of clock difference is about 14 fs.

In the above two kinds of ground-ground time transfer schemes, the clock difference of order c^{-2} in the double lambda configuration link is uncorrelated with the satellite speed, otherwise the clock difference of order c^{-2} in the TWSTFT link is related with the velocity of satellite. Therefore the uncertainty of the double lambda configuration transfer link is smaller than that of the TWSTFT link.

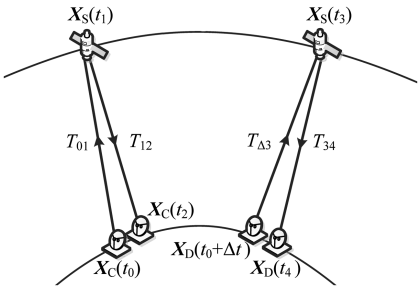


Fig. 6. The time transfer link of double lambda configuration

In order to compare conveniently the uncertainty minimums of these time transfer links, the uncertainty estimations of these links are summarized in Table 1 with the suppose that the orbit determination accuracy is 1 m, the velocity uncertainty is 0.01 m/s and the ground station velocity uncertainty is 1×10^{-5} m/s.

Table 1. Uncertainty estimations of several time transfer links

Link configuration	Order	Uncertainties
One-way	1/c	3.3 ns
Two-way	$(1/c)^2$	0.4 ps
Lambda	$(1/c)^2$	6.8 fs
TWSTFT	$(1/c)^2$	1.6 ps
Double lambda	$(1/c)^2$	14 fs

5 Conclusions

After above systematic analyses and evaluations for the four typical transfer links, we can obtain some conclusions. Firstly, the maximum uncertainty terms for all the links discussed here are at the level of order c^{-2} if we do not consider the time delays between the reception and emission of signals on the satellite itself. Secondly, for the satellite-ground lambda link and the ground-ground double-lambda link, the clock differences at the level of order c^{-2} both are uncorrelated with the satellite velocities, so that the clock difference uncertainties are superior to the satellite-ground two-way transfer link and the ground-ground TWSTFT transfer link, respectively. Finally, the third order terms of the Sagnac and Shapiro delay but not the quadrupole moment and earth rotation delay should be included if assessing the accuracy of time transfer at the level of femtosecond magnitude when the orbit altitude is less than 40 000 km.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (NO. 11803023).

References

1. Wolf P, Petit G (1995) Relativistic theory for clocks syntonization and the realization of geocentric coordinate times. Astron Astrophys 304:653–661
2. Blanchet L, Salomon C, Teyssandier P et al (2001) Relativistic theory for time and frequency transfer to order c^{-3}. Astron Astrophys 370(1):320–329
3. Blanchet L, Teyssandier P (2002) Time transfer and frequency shift to the order $1/c^4$ in the field of an axisymmetric rotating body. Phys Rev D 66:024045
4. Soffel MH, Han WB (2015) The gravitational time delay in the field of a slowly moving body with arbitrary multipoles. Phys Lett A 379:233–236
5. Han WB, Cheng R, Tao JH et al (2015) A relativistic time-delay model at the micrometer level for satellite laser ranging. Astrophys Space Sci 359:43–51
6. Liu X, Wu X, Zhang C (2009) Computational model and its precision evaluation for time comparison by two-way satellite commom-view. Acta Geod Graph Sin 38:415–419
7. Chistian S, Franco C, Luca L et al (2009) Results of the 2008 TWSTFT calibration of seven european stations. In: IEEE international frequency and time forum, 2009 joint with the 22nd European frequency and time forum, pp 1209–1215
8. Huang YJ, Fujieda M, Takiguchi H et al (2016) Stability improvement of an operational two-way satellite time and frequency transfer syatem. Metrologia 53:881–890
9. Vessot RFC, Levine MW, Mattison EM et al (1980) Test of relativistic gravitation with a space-borne hydrogen maser. Phys Rev Lett 45:2081–2084
10. Duchayne L, Mercier F, Wolf P (2009) Orbit determination for next generation space clocks. Astron Astrophys 504:653–661
11. Delva P, Meynadier F, Wolf P et al (2012) Time and frequency transfer with a microwave link in the ACES/PHARAO mission. In: 2012 European frequency and time forum, Gothenburg, Sweden
12. Thomas C, Wolf P, Uhrich PJM et al (1995) Anticipated uncertainty budgets of PRARETIME and T2L2 techniques as applied to ExTRAS

13. Samain E, Vrancken P, Guillemot P et al (2014) Time transfer by laser link (T2L2): characterization and calibration of the flight instrument. Metrologia 51:503–515
14. Samain E, Exertier P, Courde C et al (2015) Time transfer by laser link: a complete analysis of the uncertainty budget. Metrologia 52:423–432
15. Weinberg S (1980) Gravitation and cosmology principles and applications of the general theory of relativity. Science Press, Beijing

A Lamp-Pumped Rubidium Atomic Frequency Standard with a Short-Term Stability at the Level of $2 \times 10^{-13}\tau^{-1/2}$

Shuai Nie[1,2], Pengfei Wang[1(✉)], Zijing Qiu[1], Feng Zhao[1], Feng Qi[1], Gang Ming[1], Fang Wang[1], Xiumei Niu[1], Songbai Kang[1], and Ganghua Mei[1(✉)]

[1] Key Laboratory of Atomic Frequency Standards, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, China
{wpengfei,mei}@wipm.ac.cn
[2] University of Chinese Academy of Sciences, Beijing 100049, China

Abstract. A high-performance lamp-pumped rubidium atomic frequency standard (RAFS) prototype is developed. In the physics package of the RAFS, a rubidium spectral lamp with Xe as the starting gas and the isotope filtering technique are used to depress the optical shot noise of atomic signal. 30 mm-diameter absorption and filter cells located in a slotted-tube microwave cavity are utilized to increase the amplitude of clock transition signal. For the electronic system, a new low phase noise microwave chain is developed. The frequency stability measurement of the prototype shows a short-term frequency stability of $2.1 \times 10^{-13}\tau^{-1/2}$ within 1 s–100 s. This is the best short-term stability result for rubidium clock obtained so far, suggesting that the stability of the traditional RAFS could be comparable to that of the laser-pumped RAFS.

Keywords: Rubidium atomic frequency standard · Physics package · Low noise microwave chain · Frequency stability

1 Introduction

To meet the high-precision requirement for the navigation application, the performance of spaceborne rubidium atomic frequency standard (RAFS) has been increasingly improved in the last two decades. In the 1990s, the short-term and long-term frequency stabilities of the RAFS used in the GPS-IIR satellites have reached $3 \times 10^{-12}\tau^{-1/2}$ and 1.5×10^{-14} [1] at one-day time scale, respectively. At the beginning of this century, the stability of spaceborne RAFS is able to reach $1 \times 10^{-12}\tau^{-1/2}$ with a one-day stability at the level of 10^{-15} [2]. In 2016, we have demonstrated the RAFS with a short-term stability of $7 \times 10^{-13}\tau^{-1/2}$ and a one-day stability of 3×10^{-15} [3]. Recently, Q. Hao et al. designed a proof-of-principle desk-top system of RAFS, which shows an unprecedented frequency stability of $2.4 \times 10^{-13}\tau^{-1/2}$ [4]. In this proceeding, we will report our work on an integrated-phase prototype and its stability performance.

The RAFS can be understood as a frequency-locked loop, which is composed of a physics package acted as frequency discriminator and an electronic system. The

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 556–563, 2019.
https://doi.org/10.1007/978-981-13-7751-8_53

frequency stability of RAFS is determined by the signal-noise ratio (S/N) of the frequency discrimination curve, which can be expressed as [5]:

$$\sigma = \frac{\sqrt{S_N}}{\nu_0 K_D} \tau^{-1/2} \tag{1}$$

where ν_0 is the clock transition frequency, S_N is the noise power spectral density of the discriminator signal, K_D is the frequency discrimination slope. The S/N can be expressed by $\sqrt{S_N}/K_D$. According to Eq. (1), we need to decrease the noise and increase the frequency discrimination slope in order to improve the frequency stability. The noise of the frequency discrimination signal is dominated by the shot noise of the physics package and the intermodulation noise of the electronic circuits. In addition, K_D is mainly determined by the amplitude of the atomic transition signal at a certain linewidth level of $\sim$100 Hz.

2 Physics Package

2.1 Structure

The physics package of RAFS has a rubidium spectral lamp as the pumping source and a cavity-cell assembly where the clock interrogation happens, which is shown in Fig. 1.

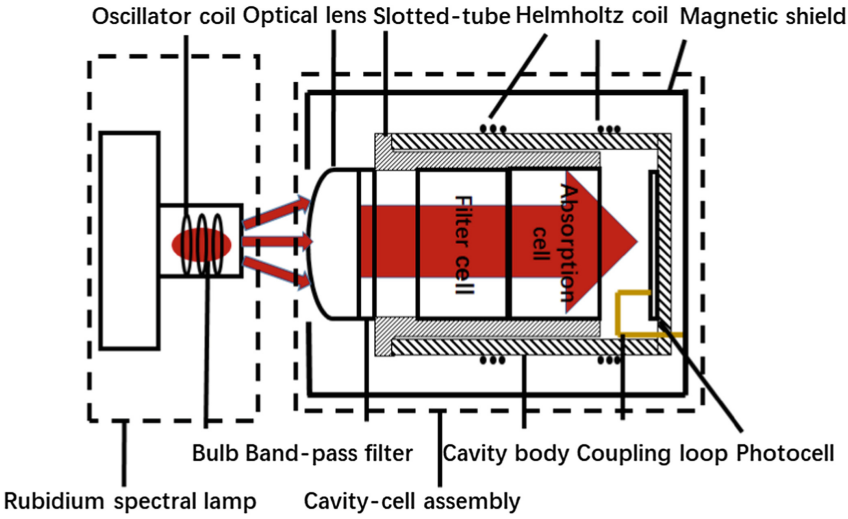


Fig. 1. Schematic diagram of the physics package

The pumping beam, emitted from the spectral lamp, is collimated by the optical lens and then filtered by a band-pass filter and a ^{85}Rb filter cell, eventually realize ^{87}Rb optical pumping. The 6.8 GHz microwave signal is fed into the microwave cavity by a

coupling loop which generates the standing wave filed with specific strength distribution to excite the clock transition of ^{87}Rb. The Helmholtz coil provides a weak static magnetic field (C field) which defines the quantization axis. The magnetic shield is used to eliminate the residual magnetic field inside the physics package. The optical-microwave double resonance signal is monitored by the photocell. Because the amplitude of the optical signal contains the frequency information of the ^{87}Rb atomic clock transition, it is eventually used to stabilize the output frequency of the oscillator.

2.2 Rubidium Spectral Lamp

The rubidium spectral lamp is an electrodeless lamp, mainly composed of a radio frequency (RF) driver circuit and a glass bulb. The Clapp oscillating circuit is utilized as the driver circuit. The glass bulb is filled with isotope ^{87}Rb and noble gas Xe. In our experiment, we use $\sim$100 MHz RF signal to ignite the lamp radiation.

The rich light spectrum emitted from the spectral lamp are from the radiations of Rb atoms and Xe. The components from Xe are useless for Rb optical pumping and only act as a noise floor. In order to reduce the noise, we use a band-pass filter to remove the spectrum of Xe gas. The pumping light spectrum, with and without the band-pass filter, are shown in Fig. 2(a) and (b). We see that the light components from Xe have been completely filtered out and only 780 nm and 795 nm from ^{87}Rb atoms are reserved. A high-resolution spectral measurement of ^{87}Rb D_1 line is shown in Fig. 3. We see that the line distortion due to the self-absorption effect is very week [6].

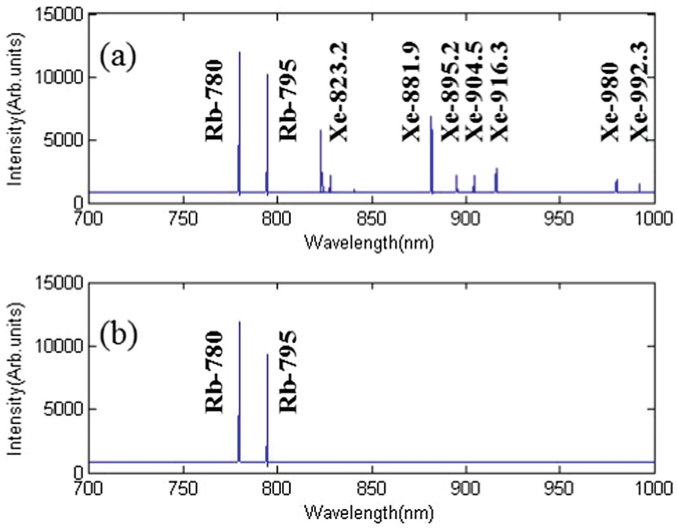


Fig. 2. (a) Spectrum of Xe lamp without band-pass filter; (b) Spectrum of Xe lamp with band-pass filter

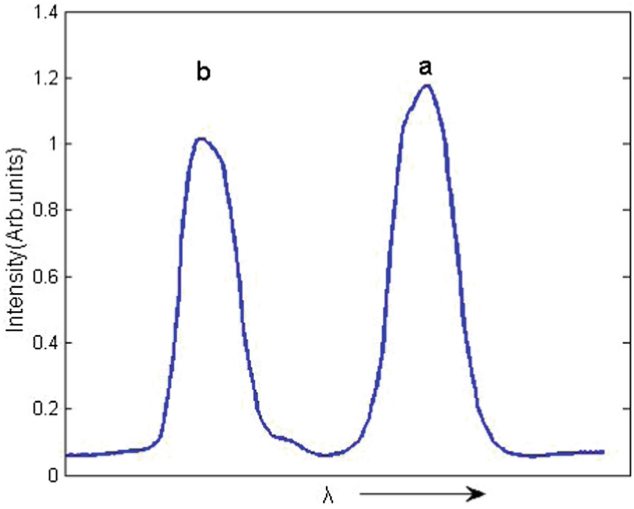


Fig. 3. D_1 line spectrum of ^{87}Rb atom

2.3 Cavity-Cell Assembly

The most crucial component of the cavity-cell assembly is the microwave cavity. Because the clock transition is a magnetic dipole transition, the clock transition only happens when the microwave magnetic field is parallel to quantization axil. Consequently, the amplitude of the clock transition directly depends on the microwave filed distribution in the cavity. A high-performance microwave cavity means its magnetic component is parallel to quantization axil (normally the cavity vertical axial) over the whole atom-microwave interaction area. The slotted-tube cavity [7] developed in our lab is shown in Fig. 4. The cavity has an inner diameter of 30 mm and the tube has six slots symmetrically distributed. Compared to the previous design [4], this new cavity-cell assembly has two changes. One is the absence of the dielectric ring wrapping outside the slotted tube, which can avoid the frequency shift induced by the intrinsic instability of the dielectric ring; the other is a longer length of the slotted tube containing both of the absorption and filter cells, which is good for the physics package's miniaturization and temperature coefficient (TC) suppression. The narrow slots in Fig. 4(a) can be understood as the space where the magnetic flux exists and excites the magnetic field parallel to the quantum axis in the tube. The simulation of the microwave magnetic field distribution is shown in Fig. 4(b). The absorption cell is located in the blue-dashed space where the magnetic lines are tightly distributed; the filter cell is located in the red-dashed space. The magnetic lines are highly parallel to the cavity axial around the centre ($\sim$20 mm, the black-dashed space). The calculated orientation factor is about 0.9, which can guarantee a large amplitude of the atomic transition signal.

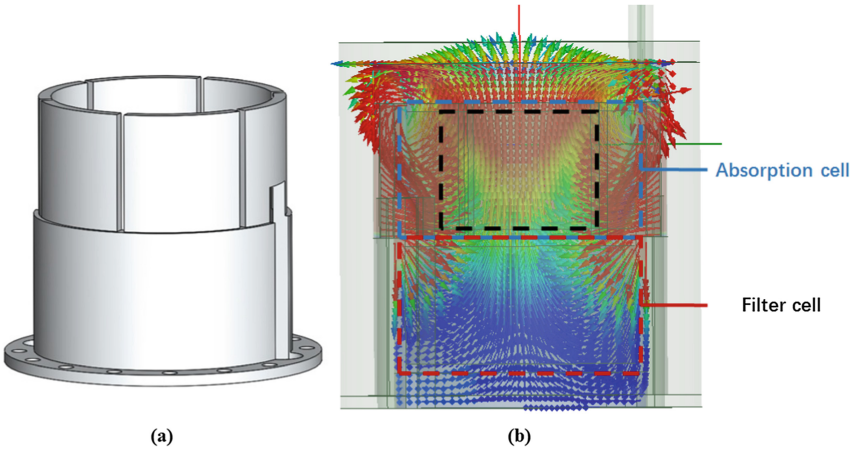


Fig. 4. (a) Structure of slotted tube; (b) Simulated magnetic field distribution in the cavity

3 Electronic System

The electronic system of RAFS consists of microwave chain, temperature controlling circuits and servo loop. The microwave chain generates the clock-interrogation signal. Due to the intermodulation effect, the phase noise of the microwave signal at even modulation frequency harmonics ($2nf_m$) will be eventually converted to the noise of the frequency discrimination signal.

The microwave chain is used to produce the 6834 MHz signal to excite the Rb atomic transition. It is composed of several components, such as a 10 MHz oscillator, a synthesizer, a frequency modulator, a 9-time frequency multiplier and a 76-time frequency multiplier. Normally, the 9-time frequency multiplier gives the main contribution to the microwave signal phase noise. Here, a low-noise Schottky barrier diode (SBD) is used as the nonlinear component instead of the transistor. In addition, a low-noise oscillator and a new modulation method at 30 MHz are utilized to further reduce the noise created during the frequency multiplication process [8]. The measured phase noise of 90 MHz signal is shown in Fig. 5. The phase noise at 2 fm is −143.6 dBc/Hz. Theoretical calculation shows that the phase noise limits the RAFS frequency stability at $1.26 \times 10^{-13}/\sqrt{\tau}$.

The key component of the 76-time frequency multiplier is a step recovery dioxide (SRD) whose frequency multiplying efficiency is sensitive to the ambient temperature. We use an intracavity multiplying design in which the SRD is integrated into the microwave cavity. This design can minimize the microwave power variation due to the environment temperature fluctuation.

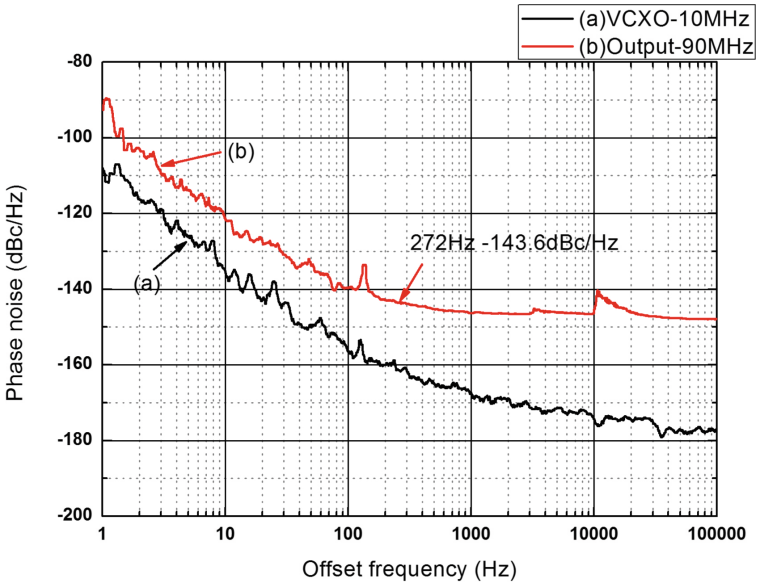


Fig. 5. Measured phase noise of the 9-time frequency multiplier

4 Frequency Stability

4.1 Theoretical Limitation of the Physics Package

The shot noise of optical detection is a main noise source for the lamp-pumped RAFS. The relationship between the detected current I_0 and its noise power spectral density $S_I(f)$ is [5].

$$S_I(f) = 2eI_0. \tag{2}$$

Substituting Eq. (2) into Eq. (1), the short-term frequency stability limited by the physics package can be calculated.

$$\sigma = \frac{\sqrt{2eI_0}}{K_D v_0} \tau^{-1/2}. \tag{3}$$

In our experiment, the measured I_0 is $\sim$100 μA. The measured discrimination curve is shown in Fig. 6, where K_D is 4.3 nA/Hz. Substituting the measured I_0 and K_D into the Eq. (3), the short-term frequency stability limited by the physics package is about $1.93 \times 10^{-13}/\sqrt{\tau}$.

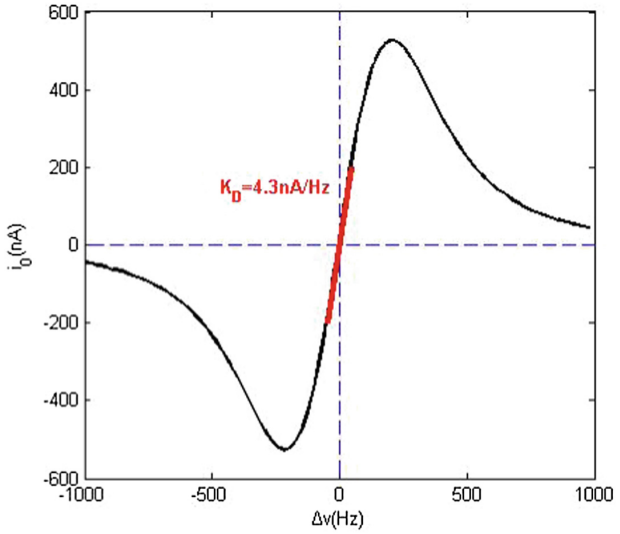


Fig. 6. Measured discrimination cure of the physics package

4.2 Experimental Result

The measured frequency stability of RAFS prototype in the atmospheric environment is shown in Fig. 7. We use the A7-MX frequency stability test instrument referenced on a hydrogen maser. The result shows that the prototype's short-term frequency stability is about $2.1 \times 10^{-13}/\sqrt{\tau}$ within 100 s. To our knowledge, it is the highest stability

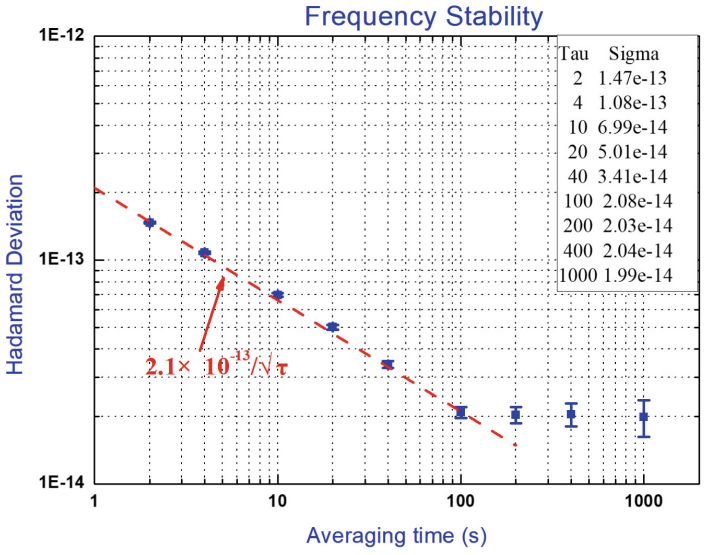


Fig. 7. Measured frequency stability of RAFS prototype

reported so far, which is comparable with the best results obtained by the laser-pumped RAFS [9]. The long-term stability (>100 s) is likely limited by the frequency shifts induced by the light and the temperature drifts in the physics package and will be investigated in the further research.

Both of the physics package and the electronics system contribute to the frequency stability of the RAFS. As we known from the above sections, the limitation of physics package and electric circuits are $1.93 \times 10^{-13}/\sqrt{\tau}$ and $1.26 \times 10^{-13}/\sqrt{\tau}$, respectively. The estimated short-term frequency stability of the RAFS from the result is $2.3 \times 10^{-13}/\sqrt{\tau}$, which is consistent with the experimental result.

5 Summary

An integrated-phase prototype of RAFS is developed in our lab. A low self-absorption Xe-spectral lamp and a 30 mm-diameter slotted-tube microwave cavity are utilized in the physics package. A new low phase noise microwave multiplier is used in the electronics system. Under these efforts, the measured short-term frequency stability reaches a level of $2 \times 10^{-13}/\sqrt{\tau}$ comparable with the stability achieved by the current best laser-pumped RAFS. The long-term stability is likely limited by the frequency shifts induced by the environment temperature fluctuation. We expect a better long-term stability performance after our research on the parameter optimizations in future.

References

1. Riley WJ (1990) Rubidium atomic frequency standards for GPS block IIR. 22nd PTTI Meeting, Vienna, pp. 221–230
2. Vannicola F, Beard R, White J et al (2010) GPS block IIF atomic frequency standard analysis. In: proceedings of the 42nd annual precise time and time interval (PTTI) applications and planning meeting, pp. 181–196
3. Mei G, Zhong D, An S et al (2016) Main features of space rubidium atomic frequency standard for BeiDou satellites. In: Proceedings of the 30th European frequency time forum (EFTF). https://doi.org/10.1109/eftf.2016.7477803:1-4
4. Hao Q, Li W, He S et al (2016) A physics package for rubidium atomic frequency stability of $2.4 \times 10^{-13}\tau^{-1/2}$. Rev Sci Instr 87:123111. https://doi.org/10.1063/1.4972567
5. Vanier J (1981) On the signal-to-noise ratio and short-term stability of passive rubidium frequency standards. IEEE Trans Instr Meas 30(4):277–282
6. Hao Q, He S, Xu F, et al (2015) Influence of the lamp spectral profile on short-term stability and light shift of a rubidium atomic clock. In: 2015 Proceedings China satellite navigation conference (CSNC), vol III, pp. 387–397
7. Mei G, Zhong D, An S et al (2001) Miniaturized microwave cavity for atomic frequency standard, US Patent: US6225870B1
8. Qiu Z, Qi F, Ming G et al (2017) Design a low noise ×9 frequency multiplication chain for rubidium atomic frequency standard. Acta Electron Sin 45(12):3046–3050
9. Bandi T, Affolderbach C, Stefanucci C et al (2014) Compact high-performance continuous-wave double-resonance rubidium standard with $1.4 \times 10^{-13}\tau^{-1/2}$ stability. IEEE Trans Ultrason Ferroelectr Freq Control 61(11):1769–1778

A Novel Method Based on Kalman Filter for Phase Noise Rejection in Two-Way Optic-Fiber Time Transfer System

Shenhui Xue[1(✉)], Bo Li[1], Yupu Wang[1], Xin Zhao[2], Xinming Huang[2], and Xiangwei Zhu[2]

[1] Beijing Satellite Navigation Center, Beijing 100094, China
xsh219281@163.com
[2] Engineering Research Center for Position, Navigation and Time, National University of Defense Technology, Changsha 410073, China

Abstract. Optic-fiber time transfer is the main method for long-distance high-precision time synchronization, and is the supporting technology of national ground time synchronization system. To overcome the degradation of phase noise induced by long-distance optical fiber, optic relay amplifier and local frequency references in two-way optic-fiber time transfer based on spread spectrum, and to improve the accuracy of carrier phase measurement and time synchronization, the characteristic of phase noise in optic-fiber time transfer is analyzed. The high-precision carrier tracking model is proposed based on kalman filter and the joint optimization is analyzed in the presence of white noise and phase noise. This model has good performance on tracking carrier phase at the transmitting end, and improving carrier measurement accuracy. In all, it is useful for phase noise rejection and improving the accuracy and stability of time synchronization in long-distance high-precision time transfer.

Keywords: Optic-fiber time transfer · Carrier phase measurement · Kalman filter · Phase noise · Joint optimization

1 Introduction

Remote high-precision optic-fiber time and frequency transfer technology is the important foundation of state time and frequency system construction. High-precision two-way optic-fiber time transfer method based on spread spectrum system can use light to realize two-way time difference measurement with carrier modulation and code division multiple access (CDMA). It can be more convenient to transfer absolute time in optic-fiber and realize time synchronization without ambiguity [1].

The distance of optic-fiber time transfer is restricted with time synchronization accuracy, which is a problem that still needs to be solved in the remote high-precision optic-fiber time transfer system. In the long distance transmission, the optical fiber amplifier amplifies phase noise while relaying, which further deteriorates at the receiving end. Thus it reduces the carrier phase measurement accuracy. After all, two-way optic-fiber time synchronization system need to achieve high precision under the

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 564–576, 2019.
https://doi.org/10.1007/978-981-13-7751-8_54

condition of long-distance time transfer. So that the influence of phase noise on carrier phase measurement needs to be analysed under sub-nanosecond condition, and the robust tracking method needs to be designed, which can improve the accuracy and stability of time synchronization.

So far, there has been a lot of researches on remote high-precision optic-fiber time transfer at home and abroad. In [2], two-way optic-fiber time transfer based on pseudo-random noise code was proposed and implemented as early as 2009, but it was limited by the measurement performance of modem at that time so that the experiment on long-distance fiber was not performed. In [3], it has realized the time synchronization accuracy at the order of sub-nanosecond based on the pseudo-range measurement on the long-distance optical fiber, but the carrier phase measurement method is not adopted in the experiment. In [4], a kalman filter based on multi-channel observation values is proposed to reduce phase noise in satellite navigation receiver. It provides new idea for phase noise rejection in optic-fiber time transfer. In [5], the tracking performance of the traditional second-order phase-locked loop is analysed in the presence of phase noise in GNSS receivers and the optimal loop parameters are calculated. But the tracking performance of traditional carrier phase lock loop could not meet the requirements of remote high-precision time transfer.

The work of this paper is to introduce the characteristic of phase noise, and explain the influence on carrier tracking in two-way optic-fiber time transfer system, and give the high-precision carrier tracking model based on kalman filter. Further the optimization of the model in the presence of phase noise and thermal noise based on kalman filter is designed to improve the carrier phase measurement accuracy.

2 Algorithm and Principle

2.1 Two-Way Optic-Fiber Time Transfer Based on Carrier Phase

The basic idea of optic-fiber time transfer is to transmit time and frequency signals between two ground stations through optical fiber, so as to achieve the consistency of phase and frequency variations between one ground station and the other. Two-way optic-fiber time transfer system includes two fiber links, which are shown in Fig. 1. Link A is used for one-way frequency transfer, and link B is used for two-way time synchronization and data transmission which still can be achieved only through link B in the presence of frequency standards at the end node. So this paper focuses on link B. Due to two-way transmission and reception, both ends of the link are identical in physical structure. The difference is that time synchronization is to adjust time and frequency of the end node to be consistent with the central node.

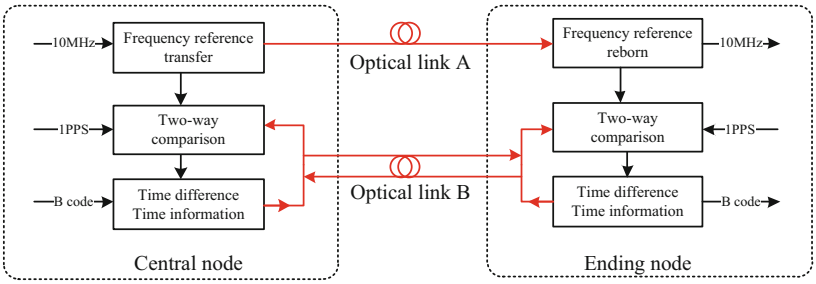


Fig. 1. Diagram: optical two-way time transfer system

The basic working principle of two-way optic-fiber time transfer system based on carrier phase measurement is shown in Fig. 2. Different from satellite navigation signal transmission, photoelectric conversion equipment needs to be added at both ends of the system. In the long-distance optic-fiber time transfer of hundreds of kilometers, relay amplifiers need to be added to ensure the intensity of bidirectional optical signals.

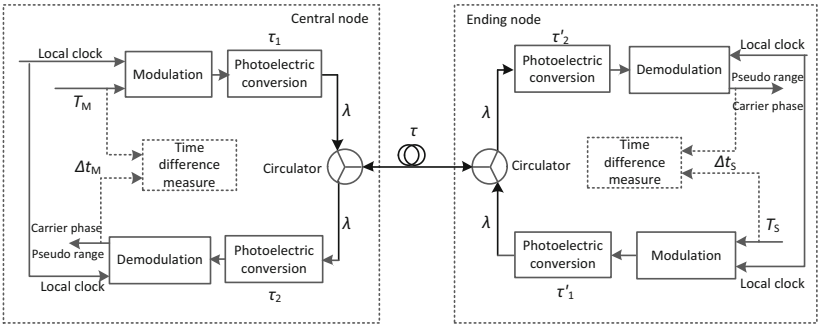


Fig. 2. Schematic: optical two-way time transfer system

Two-way optic-fiber time transfer synchronization system based on carrier phase measurement refers to the signal design method of satellite navigation system. Signal transmission time can be calculated by carrier phase measurement which is more accurate than by pseudo range measurement. The specific principle is explained as follows:

The center node and the end node use local references respectively. The baseband signal is generated with time information through pseudo code modulation, and carrier modulation, in which the selection of intermediate frequency is usually associated with AD sampling rate, such as several hundred MHZ. The generated signal is further modulated on the optical carrier by photoelectric module and transmitted through optic-fiber channel to the receiver. At the receiving end after photoelectric conversion, input signal is demodulated in the receiver with local pseudo code and carrier generated by local references. Then carrier phase measurements and the pseudo-range measurements

can be calculated at the same time to get the signal transmission time for synchronization using two-way transfer to eliminate time difference error [1].

As can be seen from Fig. 2, the same wavelength light and single optical fiber link with optical circulators between the two nodes ensure the consistency of two-way time transfer and facilitate the elimination of errors.

2.2 Characteristics of Phase Noise in Optic-Fiber Time Transfer

The mathematical model of phase noise in two-way optic-fiber time transfer is the basis of studying phase noise. Phase noise can be characterized from time domain and frequency domain [6]. In the time domain, phase noise presents as frequency stability, which can be described by Allan variance and modified Allan variance. In the frequency domain, phase noise can be described by various power spectral densities [6]. In addition, since phase noise is non-stationary, the study of it needs to be carried out with a certain model, and the power law spectrum model is widely recognized [7].

Based on this model, Allan variance of time signal transmitted through optic-fiber is measured to analyze the characteristics of phase noise in optic-fiber time transfer, so as to model the phase noise of signal transmitted through long-distance optical fiber. The 10 MHz crystal oscillator of BO2736LH1C509HC10 is selected as the frequency source in the test, the superstable constant-temperature crystal oscillator of OSA8607 BVA is used as the reference. Allan variance of the signal was measured by TSC 5125A phase noise analyzer, and 40 km optic-fiber time transfer was carried out under laboratory conditions. The analysis results are as follows.

Allan variance curve of the signal at the sending end is shown in Fig. 3. The short-term stability of phase noise is mainly affected by white frequency noise and random walk frequency noise, and the stability is about 10^{-12} (1 s).

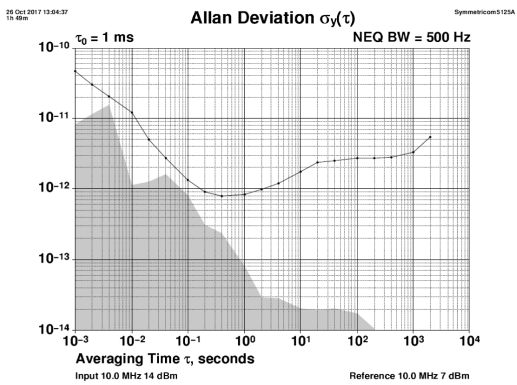


Fig. 3. Curve: Allan deviation at sending end

As shown in Fig. 4, the phase noise model of the signal at the receiving end does not change after 40 km optic-fiber transfer. The white frequency noise component further deteriorates compared with the transmitting end, and the random walk frequency noise component increases a little at the same time.

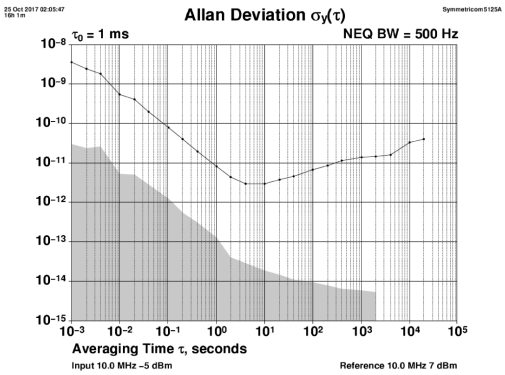


Fig. 4. Curve: Allan deviation after 40 km optical transfer

From the above analysis, it can be seen that: (1) the phase noise model does not change and still conforms to the distribution characteristics of power law spectrum model after optic-fiber transfer. (2) in terms of short-term stability, the white frequency noise component deteriorates, and increases its influence on frequency stability. Therefore, it can be considered that optical fiber transmission introduces a large white frequency noise component into the phase noise. When build the model of phase noise at the receiving end, white frequency noise and the significant phase noise component of the local frequency source are mainly considered.

According to Ref. [8], Allan variance, a characterization of frequency stability, can be expressed as:

$$\sigma_y^2(\tau) = \frac{q_1}{\tau} + \frac{q_2}{3}\tau \tag{2.1}$$

In Eq. (2.1), q_1 represents the driving noise term of white frequency noise, q_2 represents the driving noise term of random walk frequency noise, τ is the sampling interval of Allan variance. Allan variance curve can be used to fit the parameter values of q_1 and q_2.

Note that

$$Y = (A \quad B)_{n\times 2} \cdot \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}_{2\times 1} = H_{n\times 2} \cdot Q_{2\times 1} \tag{2.2}$$

$$A = \left(\frac{1}{\tau} \quad \frac{1}{2\tau} \quad \cdots \quad \frac{1}{n\tau}\right)_{n\times 1}^{T}, B = \left(\frac{\tau}{3} \quad \frac{2\tau}{3} \quad \cdots \quad \frac{n\tau}{3}\right)_{n\times 1}^{T}$$

In Eq. (2.2), Y is N × 1 measurement matrix of Allan variation, H is defined as N × 2 matrix made up of two coefficient vectors, A and B. Column vector A is N × 1, the same as B. Q is a 2 × 1 parameter-estimated matrix of q_1 and q_2, which can be obtained by Least Square Estimation (LSE).

$$Q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \left(H^T H\right)^{-1} H^T Y \tag{2.3}$$

Allan variance curve of B-type crystal oscillator was measured experimentally, and the parameter values of phase noise model obtained by fitting are shown in Table 1.

Table 1. Fitting values of phase noise model

Parameters	Sending end	After 40 km optic-fiber transfer
q_1	2.025×10^{-24}	1.163×10^{-20}
q_2	4.317×10^{-26}	2.553×10^{-25}

As can be seen from the data in the table, the value of q_1 that represents white frequency noise increases significantly after optic-fiber transfer, which increases about 4 orders of magnitude, indicating that the white frequency noise component will be amplified after the optical fiber transmission in terms of short-term stability. At the same time, the random walk frequency noise component also amplifies to a certain extent, which conforms to the phase noise distribution reflected by Allan variance curve of short-term stability.

It can be considered that in the actual scenarios, phase noise will be further degraded after transmitting over hundreds or thousands of kilometers in optical fiber, as a result of that it will affect the carrier phase measurement in the receiving process.

2.3 High-Precision Carrier Tracking Model Based on Kalman Filter

Since kalman filter has better performance for non-stationary noise than traditional PLL, a high-precision carrier phase tracking algorithm based on kalman filter is proposed to reduce phase noise in optical fiber signal on the basis of the known characteristics of phase noise model in optic-fiber transfer. The model on the basis of the original carrier tracking loop, adopts kalman filter instead of loop filter to estimate carrier phase difference of the next sample point and adjust the output phase and frequency of the numerical control oscillator (NCO). It is more quick to adjust local carrier phase close to the input signal, and better for phase noise filtering, which further improves the carrier phase tracking precision.

In the high-precision carrier phase measurement of optic-fiber time transfer, the carrier phase measurement value is taken as the observation value. According to the phase noise model and kalman filter model, the state equation can be established by observing the change of phase and signal frequency.

Measurement equation shows as

$$y(k) = \boldsymbol{h}^T \cdot \boldsymbol{X}(k) + w(k) \tag{2.4}$$

In Eq. (2.4), $\boldsymbol{X}(k) = \begin{bmatrix} \varphi(k) \\ f(k) \end{bmatrix}$, $h = [1 \quad 0]^T$, $w(k)$ is the measurement noise, whose noise variance is σ_n^2.

State equation shows as

$$\begin{bmatrix} \varphi_{k+1} \\ f_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi_k \\ f_k \end{bmatrix} + B \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{2.5}$$

In Eq. (2.5), $A = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}$, v_1 is the driving noise term of phase noise, v_2 is the driving noise term of frequency drift, they are white noise. The matrix B reflects the magnitude and interrelation of the driving noise, and its covariance matrix is $Q = B \cdot B^T$.

$$Q = \begin{bmatrix} q_1 T + \frac{q_2 T^3}{3} & \frac{q_2 T^2}{2} \\ \frac{q_2 T^2}{2} & q_2 T \end{bmatrix} \tag{2.6}$$

Parameter q_1 and q_2 can be obtained from Sect. 2.2, represent the reference value of white frequency noise and random walk frequency noise in phase noise model respectively.

So far a carrier phase tracking model has been established based on kalman filter. According to the analysis in Sect. 2.2, it is considered that the phase noise of the received signal mainly includes white frequency noise and random walk frequency noise. For more generalized cases, such as the introduction of frequency scintillation noise, a similar modeling method can be used, which can be referred to Ref. [9].

3 Simulation and Analysis

3.1 Phase Noise Simulation

According to Sect. 2.2, the phase noise mainly includes white frequency noise and random walk frequency noise in two-way optic-fiber time transfer. After a long-distance optical fiber transmission, white frequency noise component increases a lot. So based on the principle and model of phase noise, the observation data of phase noise can be generated using the method of Ref. [9] and be added to the received signal for simulation, so as to achieve the performance of carrier phase tracking.

Simulation condition: sampling rate is 125 MHz, intermediate frequency is 31.25 MHz, signal carrier to noise ratio (CNR) is 60 dBHz, parameter value of phase noise model takes 10^{-12} orders of magnitude at the sending end, parameter value of phase noise model takes 10^{-10} orders of magnitude after long-distance optical transmission. Signal carrier phase at the sending end is generated with phase noise of local clock, as shown in Fig. 5. Since the phase noise of the local clock is its inherent characteristic and cannot be eliminated, Fig. 5 is considered as the real phase of the references.

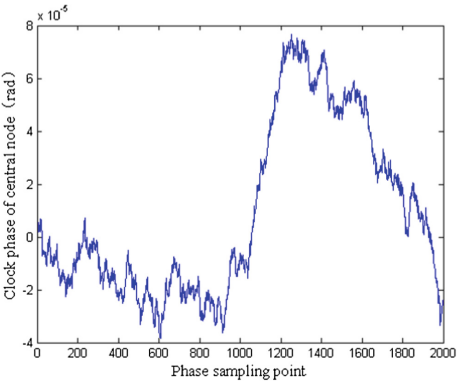


Fig. 5. Carrier phase at sending end

The signal phase at the receiving end containing phase noise introduced by optic-fiber time transfer is generated by simulation, as shown in Fig. 6, namely the observation phase of the model in Sect. 2.3.

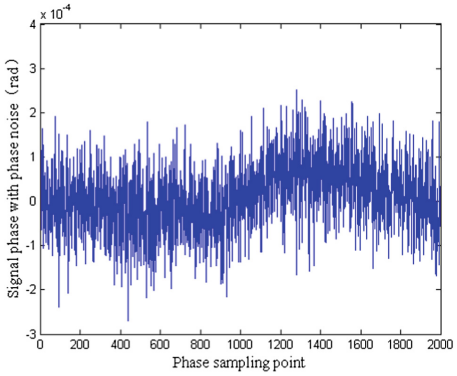


Fig. 6. Carrier phase at receiving send

As can be seen from the figure above, the observation signal phase change of optic-fiber time transfer at receiving end is at the order of 10^{-4}rad, which can have an impact on the time synchronization accuracy of about 20 ps under sub-nanosecond condition.

3.2 Carrier Tracking Performance of Kalman Filter

According to the carrier tracking model based on kalman filter proposed in Sect. 2.3, phase noise introduced by optic-fiber time transfer is all included in the system state model. The observed noise is the loop thermal noise that affects phase measurement, and the tracking analysis is based on this model. The kalman filtering process is shown as follows.

Prediction:

$$X(k|k-1) = AX(k-1|k-1) \tag{3.1}$$

Minimum predictive MSE:

$$M(k|k-1) = AM(k-1|k-1)A^T + BQB^T \tag{3.2}$$

The next step is to update the process, which mainly includes calculating kalman gain, updating state quantity and updating filter error variance matrix, as follows.

Kalman gain:

$$K(k) = \frac{M(k|k-1)h^T(k)}{\sigma_n^2 + h(k)M(k|k-1)h^T(k)} \tag{3.3}$$

Updating state quantity:

$$X(k|k) = X(k|k-1) + K(k)(Z(k) - h(k)X(k|k-1)) \tag{3.4}$$

Updating MSE:

$$M(k|k) = (I - K(k)h(k))M(k|k-1) \tag{35}$$

It should be noted that the measurement noise of phase observation quantity includes not only the system observation noise, but also the loop thermal noise. Under the condition of a carrier noise ratio of 60 dBHz, the phase jitter caused by thermal noise is about 10^{-4} [5].

The phase of the receiver signal generated by simulation in Fig. 6 is taken as input observation, and the filtering results of the model are shown in Fig. 7.

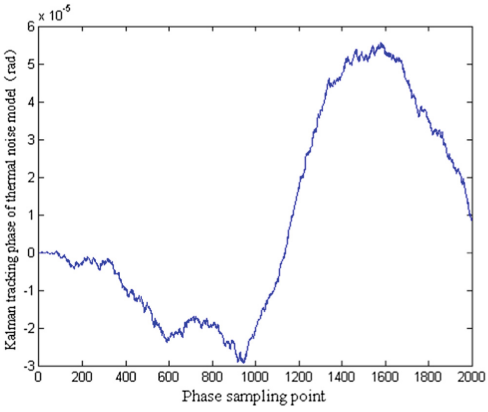


Fig. 7. Phase tracking based on kalman filter

As it shows, kalman filter whose measurement noise is thermal noise has a better filtering effect and can effectively reduce phase jitter caused. The measurement accuracy has been improved by an order of magnitude. Under this model, the tracking phase error of kalman filter is shown in Fig. 8.

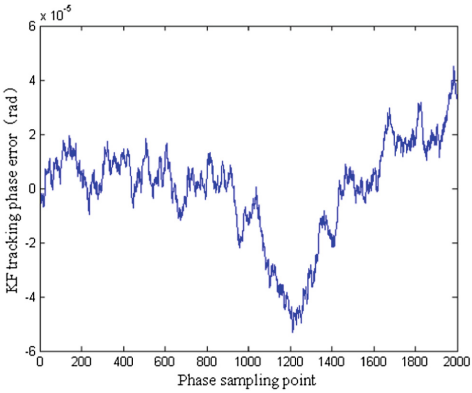


Fig. 8. Phase tracking error based on kalman filter

As can be seen from the figure above, the measurement accuracy of carrier phase reaches the order of 10^{-5}, which has a good tracking performance. The mean deviation of phase tracking error within this period is 1.82×10^{-5}.

3.3 Tracking Performance of Kalman Model Optimized by Phase Noise and Thermal Noise

In the actual two-way optic-fiber time synchronization, phase noise of the clock at the sending end cannot be eliminated due to the characteristics of the clock itself, so it can be modeled as its driving noise in the state model in Sect. 2.3. However, phase noise introduced in the optical fiber transmission process, including the phase noise introduced by local frequency reference at the receiving end, can be filtered together with the loop thermal noise as the measurement noise of the system, and the phase change result of the clock at the sending end can be retained to truly achieve time synchronization.

Therefore, in this section, under the condition that the measurement noise includes phase noise and thermal noise, the carrier tracking model based on kalman filter is optimized, mainly to change the expression of kalman gain, as shown below.

Kalman gain:

$$K(k) = \frac{M(k|k-1)h^T(k)}{\sigma_n^2 + h(k)M(k|k-1)h^T(k) + h(k)Q_1h^T(k)} \tag{3.6}$$

In Eq. (3.6), Q_1 is the driving noise term of phase noise model at the receiving end, in order to add it to the measurement noise, σ_n^2.

The phase of the receiver signal generated by simulation in Fig. 6 is taken as input observation, and the filtering results of the model are shown in Fig. 9.

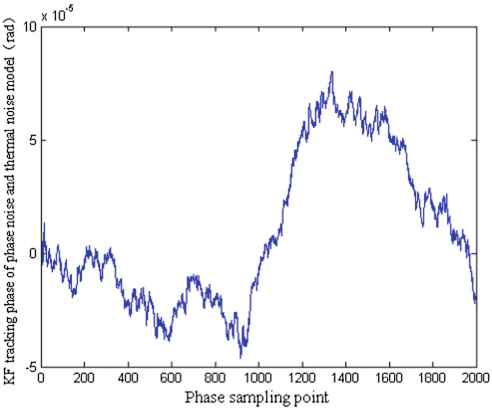


Fig. 9. Phase tracking based on kalman filter

As can be seen from the picture above, with the joint optimization of phase noise and thermal noise, carrier tracking model based on kalman filter has better noise filtering performance, can better reflect the true signal phase at the sending end, and can improve the accuracy of carrier phase measurement in this model. The tracking phase error of kalman filter is shown in Fig. 10.

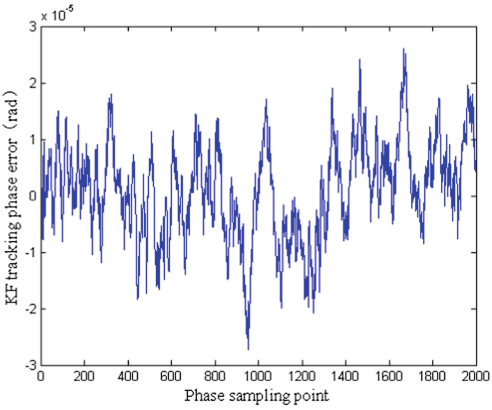


Fig. 10. Phase tracking error based on kalman filter

As it shows, the tracking phase error of the optimized model is further reduced, and the mean deviation of tracking error within this period is 8.49×10^{-6}. The measurement accuracy of carrier phase is further improved together with the time synchronization accuracy.

The comprehensive comparison of phase tracking results of the two models is shown in Fig. 11.

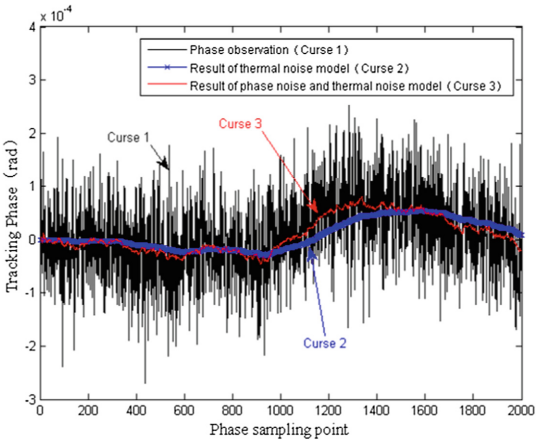


Fig. 11. Tracking performance of two model

As can be seen from the figure above, noise model of thermal noise and phase noise is more accurate so that it can adjust the tracking phase more quickly when the observation phase is fluctuating. In all, the kalman filter under the joint optimization of thermal noise and phase noise has better performance in phase noise rejection, higher phase tracking accuracy and stability, and can better reflect the phase change characteristics of the clock at the central node, as well as higher time synchronization accuracy.

4 Conclusion

This paper analyzes the characteristics of phase noise in optic-fiber time transfer, establishes the phase noise model in the two-way optic-fiber time transfer and proposes the high-precision carrier tracking model based on kalman filter. Furthermore, the observation noise model of kalman filter is optimized under the condition of phase noise and thermal noise, and carrier tracking performance is analyzed under the influence of phase noise.

Simulation results show that the high-precision carrier tracking model based on kalman filter has a good effect on phase noise rejection, can improve the precision of carrier phase measurement and time synchronization, enhance the stability of the system, and is conducive to the realization of long-distance high-precision two-way optic-fiber time transfer under sub-nanosecond condition.

References

1. Yang W (2014) Research on key technologies of high precision inter-station two-way time and frequency transfer. University of National Defense Technology
2. Piester D, Fujieda M, Rost M et al (2009) Time transfer through optical fibers (TTTOF): first results of calibrated clock comparisons. Santa Ana Pueblo, New Mexico, USA
3. Rost M, Piester D, Yang W et al (2012) Time transfer through optical fibers over a distance of 73 km with an uncertainty below 100 ps Metrology. Metrologia 49:772–778
4. Wang J, Zhuang Y, Zhou Q, Li D (2010) Phase-noise elimination of GPS receiver using improved Kalman filter. J Data Acquisition Process 25(5):611–614
5. Xue S, Huang X, Liu Z, et al (2017) Optimal carrier tracking PLL in the presence of phase noise in GNSS receiver. In: Proceedings china satellite navigation conference (CSNC) 2017, vol I
6. Liu Z (2013) Research on the characteristics of phase noise in high-precision receiver RF front-end. University of National Defense Technology
7. Yang W (2008) The study of phase noise analysis and identification. University of National Defense Technology
8. Galleani L, Tavella P (2010) Time and the Kalman filter: applications of optimal estimation to atomic timing. IEEE Control Syst Mag 30(2):44–65
9. Huang X, Gong H, Zhu X, Ou G (2011) A real-time anomaly monitoring algorithm for satellite clock based on Kalman filter. J Astronaut Metrol Meas 31(5):6–11

Analysis of Sagnac Correction for Time Transfer in Optical Fibers

Bin Wang[1,2(✉)] and Junping Chen[1]

[1] Shanghai Astronomical Observatory, CAS, Shanghai, China
{binw, junping.chen}@shao.ac.cn
[2] Shanghai Key Laboratory of Space Navigation and Positioning Techniques, Shanghai, China

Abstract. Over the past few years, atomic clocks have been improving and are now reaching stabilities and accuracies of a few parts in 10^{18} in fractional frequency. Fiber-based time and frequency transfer techniques have demonstrated excellent performance in the comparison of state-of-art optical atomic clocks over thousands of kilometers. For highly accurate time and frequency transfers, relativistic effects which affect the signal propagation in optical fibers, need to be taken into account. The most important is the Sagnac correction because of the computation complexity and also the non-reciprocity in the two-way time transfer. Sagnac correction is an important source of uncertainty in fiber-based time and frequency transfer. Besides, not all important parameters are known with sufficient precision when we compute the Sagnac correction such as the large position error of fiber nodes and also the sparse fiber nodes. It is necessary to evaluate the Sagnac correction due to imperfect knowledge of parameters. In our work, several simulation fiber links in China are analyzed as specific examples to evaluate the influences of imperfect knowledge on the accuracy of the time transfer. The results show that in order to ensure one picosecond precision of time transfer using optical fiber, position accuracy of nodes should be higher than 500 m when the information of enough number of nodes can be obtained.

Keywords: Time transfer · Sagnac correction · Optical fibers

1 Introduction

Precise time and frequency are the most demanding physical quantities in the development of science and technology. Optical clocks, due to their unprecedented precision [1] and accuracy [2], are already being used in the experiment of physical theories [3], and will play an important role in the redefinition of the Intentional System of Units and the development of global timescales [4]. Traditional satellite-based techniques used for long distance comparison of frequency standards, are no longer capable of supporting the required accuracy of optical clocks, currently at the 10^{-18} level [2]. However, fiber-based time and frequency transfer techniques have demonstrated excellent performance in the comparison of state-of-art optical atomic clocks over thousands of kilometers [5], and such fiber network would be one of the key technologies for the application of new generation of atomic clocks in many fields, such as geodesy, radio astronomy, spectroscopy, and so on.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 577–585, 2019.
https://doi.org/10.1007/978-981-13-7751-8_55

Requirements on accuracy and stability of time and frequency transfer push the evaluation of the phenomena affecting the signal propagation in fiber link, and also the uncertainty of correction models to a new level. Uncertainty of time transfer mainly arises from the employed time modems and the fiber link. High-precision calibration can be performed to reduce the influences of delays in time transfer modems [6]. While fiber link is subject to various factors, such as temperature variations, vibrations, acoustic noise, relativistic effects, and so on. Among them, relativistic effects can be precisely modeled [7], but difficult to correct in some cases due to poor knowledge of fiber link information [8].

Model of relativistic corrections with an uncertainty of 1 ps has been proposed by Geršl et al. [7]. The time corrections for one way time transfer using fiber link are composed of three terms, Newtonian term, Sagnac correction and also the Shapiro correction. Newtonian and Shapiro corrections can be compensated by the two-way time transfer, while not for the Sagnac correction. Therefore, Sagnac correction is an important source of uncertainty in the long haul optical time transfer.

Sagnac correction depends on the area of equatorial projection of the surface swept by the earth radial vector moving along the optical fiber. Difficulty in the computation of Sagnac correction mainly results from the limited knowledge of the route of optical fiber. The real coordinates of fiber route are often not known exactly or confidential for internal use of optical network operator only. Therefore, several approximated methods have been proposed [7–12]. These methods can be classified into four types of methods. The first two methods are sphere approximation [11] and ellipsoid approximation [9] of the earth when only the coordinates of the end points are known. The third method is used in the case that some nodes are exactly known an also the distance between two adjacent nodes [7]. The fourth method is used in the case of benchmark real-like optical paths [8]. As for the confidence interval, there are also four types of methods. First method uses the sector area of the parallels passing through the ends of the fiber link [9]. Second method uses the sector area of the parallels with the fixed length of fiber link [10]. Third method uses the maximum and minimum triangle area constrained by the fixed length [7], while fourth method uses the maximum semicircle area constrained by the fixed length [12].

In our work we do some analysis of Sagnac correction on china fiber link. Due to the very limited knowledge of the optical fiber that only the coordinates of some nodes are known, we have to find some real-like transmission lines for analysis. It's known that fiber cables on land are typically installed in ducts for utility distribution, such as gas and electricity, often running across metropolitan areas and along highways, routes planning function of map API (Application Programming Interface) is used to generate a real-like optical-fiber link. Based on the generated optical-fiber link routes, influences of imperfect information (especially the position error and also the sparse degree of fiber nodes) on Sagnac correction are analyzed.

2 Sagnac Correction in the Fiber-Based Time Transfer

Systematic relativistic description of signal propagation in optical fibers can be found in [7], and the time correction for one way time transfer over optical fiber is also derived with an uncertainty of 1 ps:

$$\Delta t_{\pm} = \frac{1}{c}\int_0^L n\,\mathrm{d}l \pm \frac{1}{c^2}\int_0^L \mathbf{v}\cdot\mathbf{s}_l\,\mathrm{d}l \pm \frac{1}{c^3}\int_0^L n\left(w + v^2/2\right)\mathrm{d}l \tag{1}$$

where c is the light speed, n is an effective refractive index, $\mathbf{v}$ is velocity vector of a fiber point, $\mathbf{s}_l$ is the tangent vector field with parameter l, w is earth potential, $v^2 = \mathbf{v}\cdot\mathbf{v}$. The first c^{-1} term is Newtonian term and the third c^{-3} term is Shapiro correction. These two terms are independent of direction, and can be compensated in two-way time transfer. The main part of the second term (c^{-2}) in above formula is the Sagnac correction and comes from the earth rotation. Sagnac correction can also be expressed as:

$$\Delta t_{\pm sagnac} = \frac{2\Omega A_{\pm}}{c^2} \tag{2}$$

where Ω is the nominal mean value of earth's angular velocity, numerically 7.2921115e-5 rad/s [13], and contribution of earth tides on the earth's angular velocity is omitted. $A_{\pm}$ corresponds to the equatorial projection area of the surface swept by the geocentric radial vector moving along the optical fiber. Besides, Sagnac correction depends on the direction of signal propagation, when the signal propagates along the direction of earth rotation, the sign is positive, otherwise, the sign is negative. Above equation assumes that $A_{\pm}$ is directed area dependent of the relative position of the two ends of optical fiber or segments of optical fiber. It will be shown that this assumption can be used to significantly simplify the numerical calculation process of Sagnac correction for a long-haul optical fiber.

3 Comparison of Sagnac Correction Methods

In order to obtain exact Sagnac correction, Sagnac area should be computed as precise as possible. However, the real coordinates of optical fiber are often not known exactly or confidential for internal use of optical network operator due to security reasons. Therefore, several approximation methods of Sagnac correction have been proposed, and also the confidence interval computation. In the following, comparison of these methods is performed.

Four methods can be found for Sagnac correction computation. First method (Method I) approximates the earth as an ideal sphere, when the ends of fiber link are known, the arc of great circle is used to approximate the real fiber link [11], as shown in Fig. 1. When computing the Sagnac correction, the following procedures can be used. Firstly, normal vector of plane OAB can be obtained through the cross product of radial vectors $\boldsymbol{OA}$ and $\boldsymbol{OB}$, then we have the equation of route L_{AB} on the great circle with combination of OAB plane equation and the sphere equation. After the cancellation of z variable, we can get the equation of $L_{A'B'}$ on the xoy plane, and projected area can be obtained through the integral.

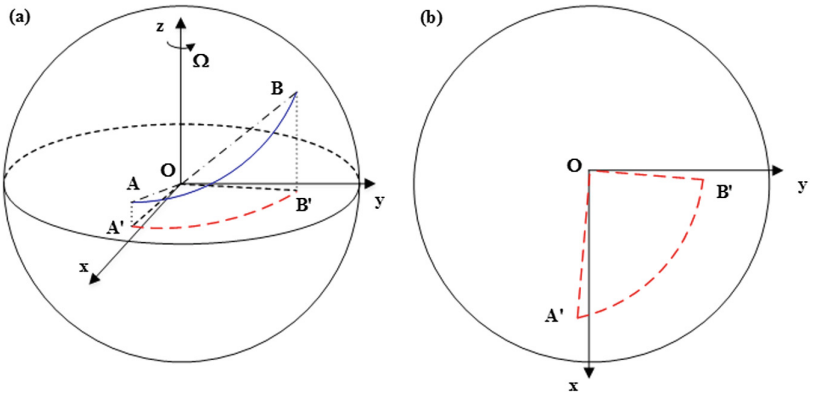


Fig. 1. Sagnac correction computation in the case of sphere approximation of the earth. The fiber link is shown in the left subplot (a), projected Sagnac area on the equatorial plane is shown in (b).

Second method (Method II) approximates the earth as an ellipsoid, when the ends of fiber link are known, the meridians and the parallels are used to approximate the real fiber link [9, 10], as shown in Fig. 2. Even though the real optical fiber may runs beyond the region between the meridians and parallels through its two ends, the projected area of the circular sectors (B′OC′ and A′OD′) closed by the projections of meridians and parallels on the equatorial and the earth radial vector would be a good approximation of the Sagnac area. With the known coordinates of two ends, parallels radius can be easily computed, then we can obtain the area of B′OC′ and A′OD′. Sagnac correction can be considered as the mean value of two corrections corresponding to the area of B′OC′ and A′OD′. Besides, uncertainty of the estimation can also be obtained using the difference between the area of B′OC′ and A′OD′.

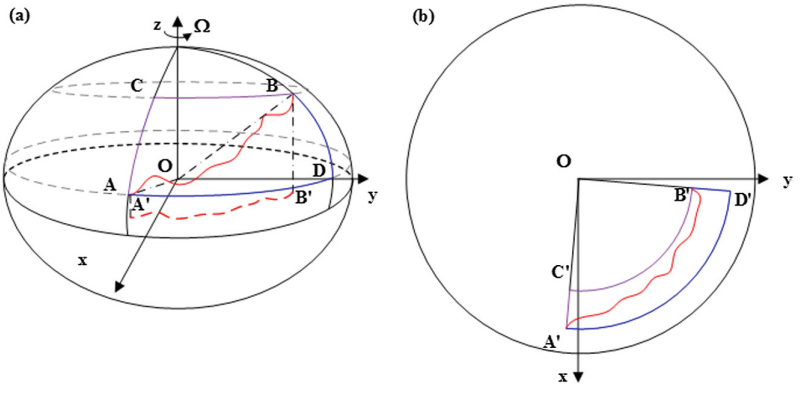


Fig. 2. Sagnac correction computation in the case of ellipsoid approximation of the earth. The fiber link is shown in the left subplot (a), projected Sagnac area on the equatorial plane is shown in (b). The violet and blue lines correspond to the considered projection of hypothetical fiber link.

If the position of some nodes and also the lengths between two adjacent nodes of the optical fiber can be exactly known, then another method (Method III) can be used. In this case, actually two methods have been proposed for the computation of projected Sagnac area. One is the computation of circular sectors area with the fixed length of fiber route segment [10], as shown in the subplot (a) of Fig. 3. The other is the computation of triangular area with exact positions, while the fixed length of fiber route segment is used as a constraint condition for the computation of estimation uncertainty [7], as shown in the subplot (b) of Fig. 3.

The fourth method (Method IV) is used in the case of benchmark real-like optical paths [8], where position of more points are exactly known than that of aforementioned case. Triangle area in the equatorial plane with exact positions of fiber route segments are computed, and the length of fiber is used as a constraint condition to find the maximum semicircle for the computation of estimator uncertainty [12].

Method II, Method III and Method IV are compared by Czubla, Krehlik [10], it is shown that highest precision and accuracy can be obtained with Method IV, and accuracy of Method II and Method III is at the same level, while Method II has better precision than Method III. Slapak and Vojtech [8] compare all these four methods with the published information of several fiber links, and find that more detailed information of nodes is known, more precise Sagnac correction will be. In the case that when we have enough knowledge of the optical fibre path, the Method IV can achieve at least two times lower uncertainty.

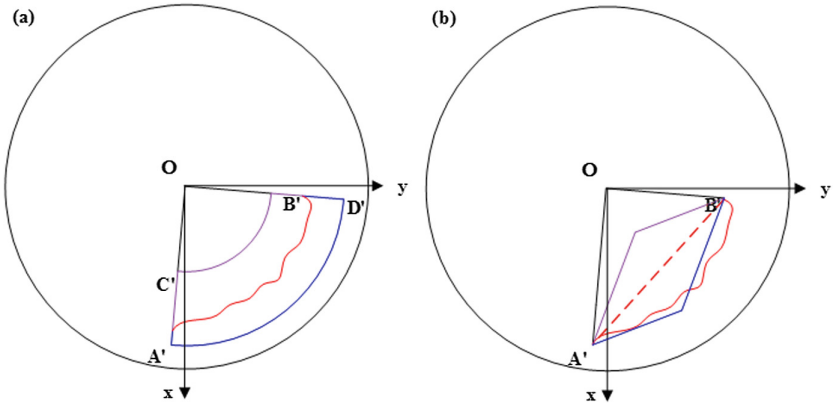


Fig. 3. Sagnac correction computation in the case of exact position and fixed length of fiber route segment. Circular sectors area with the fixed length of fiber route segment is shown in the left subplot (a), while triangle area with the fixed length of fiber route segment is shown in subplot (b).

4 Results and Discussion

Due to the confidential data of the fiber routes, we have very limited knowledge that only the coordinates of some nodes are known. In order to analyze the Sagnac correction of china fiber link, we have to find real-like transmission lines for analysis. It's

known that fiber cables on land are typically installed in ducts for utility distribution often running across metropolitan areas and along highways. Routes planning function of baidu map API (http://lbsyun.baidu.com/) is used to generate a real-like optical-fiber link. Based on the generated optical-fiber link routes, Sagnac correction and also their uncertainty from imperfect information (especially the fiber and station coordinates) are analyzed.

4.1 Software Validation

In order to analyze the Sagnac correction of China optical fiber used for precise time transfer, a software package has been developed. The automated calculation algorithm of Sagnac correction (Method IV) published by Šlapák et al. [12] is used, and some improvements have been made. Concept of directed triangle area is used, and procedure of loop detection is removed. As for the confidence interval of Sagnac correction estimated value, length of fiber is constrained to find the maximum semicircle area enclosed by a given ratio of chord length and arc length.

Published results of several benchmarking paths [8] are used as the reference to evaluate the software package. Six benchmarking paths are used for the validation, and the information about them can be found in the website (https://photonics.cesnet.cz/en/sagnac-benchmark). Comparison results are shown in Table 1. It can be seen that difference of Sagnac correction value is at tens of fs level. The large difference of uncertainty for some path (Bethel–Pakhachi) may result from the combination effect of long length and small points.

Table 1. Comparison of Sagnac corrections on benchmarking paths between the published results and our software package results.

Endpoints	Length [km]	Points	Published Sagnac correction [ps]	Our Sagnac correction [ps]
Berlin–Poznan	260	51	756.561 ± 0.415	756.586 ± 0.406
Berlin–Ustı nad Labem	265	89	145.421 ± 0.317	145.428 ± 0.104
Bethel–Pakhachi	1943	31	−4063.525 ± 63.392	−4063.589 ± 61.752
Dusseldorf	30	46	−11.910 ± 0.010	−11.910 ± 0.003
Karlskoga–Gustavberg	260	99	582.180 ± 0.243	582.197 ± 0.235
Sydney–Adelaide	1397	137	−5006.103 ± 10.037	−5006.545 ± 9.933

4.2 Sagnac Correction Analysis of Optical-Fiber Time Transfer

Eight cities, (Beijing, Wuhan, Guangzhou, Sanya, Nanjing, Shanghai, Xian, Chengdu), are used as the input parameter for routes planning of baidu map API, and seven optical fiber paths are output as the simulation results (Fig. 4).

Fig. 4. Simulation results of China optical fiber paths using baidu map API.

Sagnac corrections are computed based on the simulated seven paths, and the results are shown in Table 2. It can be seen that uncertainty of Sagnac corrections is at fs level. In the following, these results would be considered as the reference. The influences of optical-fiber nodes position error on Sagnac correction computation are analyzed from two aspects, three-dimensional position error and degree of sparsity of fiber nodes.

Table 2. The results overview of Sagnac corrections on seven simulated optical fiber paths

Endpoints	Length [km]	Points	Sagnac correction [ps]
Beijing–Wuhan	1184.5	4489	−673.807 ± 0.002
Wuhan–Guangzhou	996.6	4199	−433.567 ± 0.000
Guangzhou–Sanya	878.0	3510	−1917.522 ± 0.004
Wuhan–Nanjing	543.7	2171	1845.816 ± 0.000
Nanjing–Shanghai	296.8	1571	1108.437 ± 0.000
Wuhan–Xian	747.2	3939	−2203.347 ± 0.000
Xian–Chengdu	714.7	4515	−2023.673 ± 0.000

In order to analyze the influences of optical-fiber nodes position error on Sagnac correction, random position error of 10 m, 20 m, 50 m, 100 m, 200 m, 500 m, and 1 km, respectively on north, east, and up direction is added into the known position of fiber nodes. Then the Sagnac correction is computed, and compared with reference value. We repeat 5000 times of above processing, and obtain the statistics. Influences of position error on Sagnac correction for all these seven paths are almost the same. Analysis results of Beijing–Wuhan path are shown in Table 3. It can be seen that in order to ensure 1 ps precision of time transfer, position error of nodes should be smaller than 500 m. This result is consistent with the analysis results of Gersl et al. [7].

Table 3. Influences of optical-fiber nodes position error on Sagnac correction for Beijing–Wuhan path

Endpoints	Length [km]	Points	Sagnac correction [ps]
Beijing–Wuhan	1184.5	4489	−673.807 ± 0.002
Position error [m]	Mean [ps]	STD[ps]	RMS [ps]
5	0.000	0.017	0.017
10	0.000	0.035	0.035
20	0.000	0.069	0.069
50	0.000	0.173	0.173
100	0.000	0.345	0.346
200	0.000	0.691	0.692
500	−0.001	1.726	1.729
1000	−0.001	3.453	3.458

In order to analyze the influences of fiber nodes sparsity on Sagnac correction, several degrees, such as 5%, 10%, 20%, 30%, until 70%, are used. We also repeat 5000 times. The influences of nodes sparsity are little, because even 70% nodes are deleted, one node per 1 km is still guaranteed for all these seven simulation paths.

5 Conclusion

In order to analyze the Sagnac correction of china optical fiber links routes, planning function of baidu map API (http://lbsyun.baidu.com/) is used to generate a real-like optical-fiber link. Based on the generated optical-fiber link routes, magnitude of Sagnac correction and their uncertainty from imperfect information (especially the fiber and station coordinates) are analyzed. It is found that in order to ensure one picosecond precision of time transfer using optical fiber, position error of nodes should be smaller than 500 m.

Acknowledgements. This work is support by the NSFC (No. 11673050); the 863 (No. 2014AA123102); the open project program of the key lab of space navigation and positioning technology (No. KFKT-201705), Shanghai; Open research fund of state key laboratory of information engineering in surveying, mapping and remote sensing, Wuhan University (No. 18P01); Open Foundation of State Key Laboratory of Geodesy and Earth's Dynamics (SKLGED2019-3-1-E)

References

1. Hinkley N, Sherman JA, Phillips NB et al (2013) An atomic clock with 10(-18) instability. Science 341(6151):1215–1218
2. Nicholson TL, Campbell SL, Hutson RB et al (2015) Systematic evaluation of an atomic clock at 2 × 10 (-18) total uncertainty. Nat Commun 6:8
3. Godun RM, Nisbet-Jones PBR, Jones JM et al (2014) Frequency ratio of two optical clock transitions in Yb-171(+) and constraints on the time variation of fundamental constants. Phys Rev Lett 113(21):5

4. Riehle F (2015) Towards a redefinition of the second based on optical atomic clocks. C R Phys 16(5):506–515
5. Turza K, Krehlik P, Sliwczynski L (2018) Long haul time and frequency distribution in different DWDM systems. IEEE Trans Ultrason Ferroelectr Freq Control 65(7):1287–1293
6. Zhang H, Wu GL, Li XW et al (2017) Uncertainty analysis of BTDM-SFSW based fiber-optic time transfer. Metrologia 54(1):94–101
7. Geršl J, Delva P, Wolf P (2015) Relativistic corrections for time and frequency transfer in optical fibres. Metrologia 52(4):552–564
8. Slapak M, Vojtech J (2018) Real like transmission lines for Sagnac correction study. In: Proceedings of EFTF 2018, Torino, Italy (2018)
9. Łukasz Ś, Przemysław K, Albin C et al (2013) Dissemination of time and RF frequency via a stabilized fibre optic link over a distance of 420 km. Metrologia 50(2):133
10. Czubla A, Krehlik P, Sliwczyliski L et al (2017) Some approximated methods of calculation Sagnac correction for optical fiber time transfer. In: 2017 joint conference of the European frequency and time forum and IEEE international frequency control symposium, New York, pp 399–401. IEEE (2017)
11. Yu LQ, Lu L, Wang R et al (2013) Analysis of the Sagnac effect on the accuracy of the long haul optical fiber time transfer system. In: 2013 joint European frequency and time forum & international frequency control symposium, New York, pp 303–305. IEEE (2013)
12. Šlapák M, Vojtěch J, Velc R (2017) Automated numerical calculation of Sagnac correction for photonic paths. Opt Commun 389:230–233
13. Petit G, Luzum B (2010) IERS Conventions 2010. IERS Technical Note. No. 36. International Earth Rotation and Reference Systems Service (IERS), Frankfurt am Main

Dynamic Simulations of Hydrogen Atoms in Magnetic State Selector of MASER

Dongliang Cong(✉), Shuo Liu, Jing Li, and Lianshan Gao

Beijing Institute of Radio Metrology and Measurement, Beijing, China
congdongliangcg@163.com

Abstract. In this paper, a random method to simulate atomic behavior in magnetic state selector of atomic hydrogen maser (H-maser) is reported. Using this method, the trajectory of each atom and atomic distribution in the entrance plane to the storage bulb is provided. The effect of the geometry and properties of magnetic state selector on the results are also discussed. We take H_0 and L_3 as examples and discuss the number of effective hydrogen atoms which are focused into storage bulb. When H_0 = 0.9T, the number of effective atoms is at most and the optimum value of L_3 is between 60 mm and 100 mm in the simulation condition. This approach can be possessed to improve the performance of H-maser.

Keywords: Magnetic state selector · Atomic trajectory · Atomic distribution · Dynamic simulations · Random method

1 Introduction

The atomic hydrogen maser (H-maser) currently provides the highest available level of frequency stability [1–3]. In order to improve the performance of H-maser, one key method is increasing the number of effective hydrogen atoms which are focused into storage bulb [4]. Magnetic state selector is used to obtain population inversion in the hyperfine transition levels. A beam of hydrogen atoms is passed through a multipole magnet (magnetic state selector), and atoms in the F = 1, m_F = 1, 0 states are focused into storage bulb, through an aperture and atoms in the F = 1, $m_F = -1$ and F = 0, m_F = 0 states are defocused.

Using a random method to simulate atomic behavior is presented in this paper. Initial angle, velocity and states of atoms effusing out of discharge dissociator are taken into consideration in this simulation model. The possible trajectory of each atom is simulated and the atomic position distributions in hydrogen beam out of magnetic state selector are also reported. The effect of the geometry and properties of magnetic state selector on atomic trajectories, atomic position distributions and the atomic number in storage bulb are also discussed. Using those results, we can optimize the design of magnetic state selector to improve H-maser performance. Here we show the simulation method and those results.

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 586–593, 2019.
https://doi.org/10.1007/978-981-13-7751-8_56

2 Methods

2.1 Introduction of Magnetic State Selector

For beam devices designed to use velocity focusing trajectories to achieve maximum efficiency, the quadrupole or the hexapole state selector may be used, depending upon the physical properties of the particular atom. Quarupole state selectors have a magnetic field which increases linearly with radius r about the beam axis, while hexapole state selector fields increasing with the square of the radius r [4]. For H-maser, quarupole state selectors are used in this paper and its magnetic field H is given by (1).

$$H = H_0 r / r_0 \tag{1}$$

Here H_0 is poletip field, r_0 is bore radius. The force acting on an atom immersed in an inhomogeneous magnetic field H is given by (2).

$$F = \mu_{eff} grad\,(H) = \pm \mu_B H_0 / r_0 \tag{2}$$

Here μ_{eff} is the effective magnetic moment of an atom. μ_B is the Bohr magneton; the (+) sign corresponds to F = 1, $m_F = -1$ and F = 0, $m_F = 0$; the (−) sign, to F = 1, $m_F = 0$, and F = 1, $m_F = 1$.

Figure 1 shows the magnetic state selector geometry: a quadrupole magnet of length L_2 has its entrance plane at a distance L_1 from the source exit plane; the entrance plane to the storage bulb is at a distance L_3 from the exit plane of the magnet. The axial velocity of the atom is constant. The radial velocity is constant in regions a and c, and atoms in regions b accelerate for its are acted the constant force F. Atomic trajectories

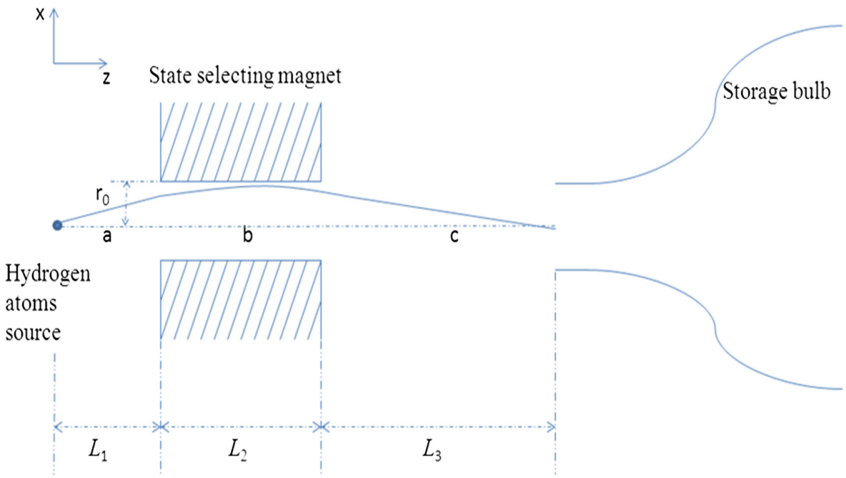


Fig. 1. Schematic view of the state selector geometry

leading from the source into the storage bulb are calculated by solving the equations of motion in regions a, b, and c.

2.2 Dynamic Simulations of Hydrogen Atoms in Magnetic State Selector

Firstly, choose some random numbers to describe characters of each atom. Those three parameters (α, v, e) are used to describe an atom which have initial angle α, velocity v and symbol of energy level e. The angular distribution of atoms effusing out of source collimator is given by (3) [5].

$$f(\alpha) = \frac{2}{\pi}\left\{\cos^{-1}\frac{\alpha}{\alpha_0} - \frac{\alpha}{\alpha_0}\left(1-\frac{\alpha^2}{\alpha_0^2}\right)^{1/2} + \frac{2\alpha}{3\alpha_0}\left[1-\left(1-\frac{\alpha^2}{\alpha_0^2}\right)^{3/2}\right]\right\} \tag{3}$$

Here α_0 is maximum initial angle of atoms out of the collimator. α_0 is defined by the radius to length ratio of the collimator. Considering the state selector symmetry, the value of α should be picked up between $-\alpha_0$ and α_0. Atoms whose angles are out of $-\alpha_0$–α_0 are not taken into considered, for those atoms do not reach to the entrance plane to the storage bulb.

Atomic velocity v relates to temperature of the discharge dissociator bulb. The thermal kinetic energy of the atom in the hydrogen atomic beam, 1/2 mv^2, is equal to 3/2 KT, and the velocity can be obtained from this equation. Here m is the hydrogen atomic mass, K is the Boltzmann constant, and T is temperature of the discharge dissociator bulb.

Symbol of energy level e is used to distinguish atomic energy level. It obeys even distribution between 0 and 1. If $e \in [0.5, 1]$, the atom is in the $F = 1$, $m_F = 1, 0$ energy level and is focused into storage bulb, or in the $F = 1$, $m_F = -1$ and $F = 0$, $m_F = 0$ is defocused.

In order to miniaturize physical package of hydrogen maser, the distance between the entrance plane of the state selector and the source exit plane is zero, $L_1 = 0$. There is not force acting on atoms in regions c, so atoms move uniformly and in a straight line. Atoms in regions b accelerate for the constant force F. The radial equation of motion for an atom in regions b is given by (4).

$$\ddot{r} = \pm\frac{\mu_B}{m}\frac{H_0}{r_0} \tag{4}$$

Here the (+) sign corresponds to $F = 1$, $m_F = -1$ and $F = 0$, $m_F = 0$; the (−) sign, to $F = 1$, $m_F = 0$, and $F = 1$, $m_F = 1$.

Then, according to above analysis, we use those random number and dynamic process to draw each atom's trajectory. Figure 2 shows atomic trajectories from atomic source to the entrance plane to the storage bulb. N of simulation condition presents initial number of atoms out of the collimator. The red lines are related to atoms in $F = 1$, $m_F = -1$ and $F = 0$, $m_F = 0$; the blue lines are trajectories of atoms in $F = 1$, $m_F = 0$, and $F = 1$, $m_F = 1$. For the sake of easy introduction, we use the phrase "effective atoms" to describe atoms in $F = 1$, $m_F = 0$, and $F = 1$, $m_F = 1$ and the

phrase "ineffective atoms" to indicate atoms in F = 1, $m_F = -1$ and F = 0, $m_F = 0$. So the blue lines are trajectories of effective atoms and the red trajectories are ineffective atoms. In Fig. 2, X-axis represents atomic radial position. Z-axis represents the distance from the source exit plane. The position of $Z = L_2$ is the exit plane of the magnet and the position $Z = L_2 + L_3$ is the entrance plane to the storage bulb. Obviously, the state selector captures the atoms in F = 1, $m_F = 0$, and F = 1, $m_F = 1$ and deflects strongly the atoms in F = 1, $m_F = -1$ and F = 0, $m_F = 0$. In the follow discussion of this paper, we only consider the atoms in F = 1, $m_F = 0$, and F = 1, $m_F = 1$, for the atoms in F = 1, $m_F = -1$ and F = 0, $m_F = 0$ do not enter into the storage bulb.

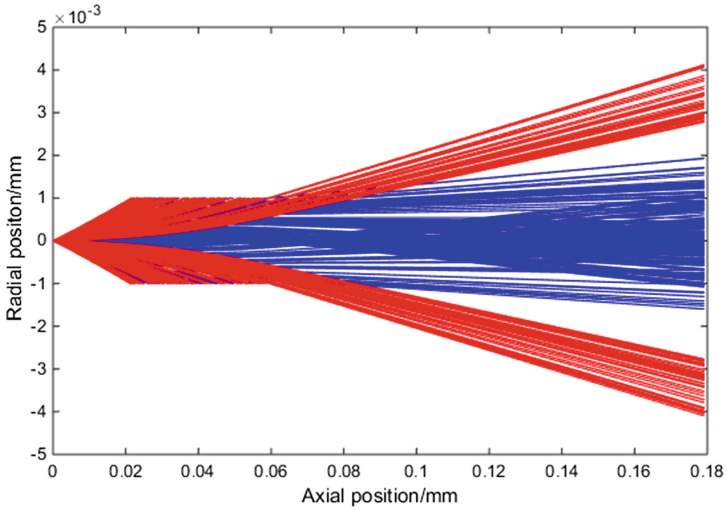


Fig. 2. Atomic trajectories diagram from atomic source to the entrance plane to the storage bulb. Simulation condition: N = 1000, v = 3000, $H_0 = 0.5T$, $r_0 = 1$ mm, $L_2 = 59$ mm, $L_3 = 120$ mm.

Figure 3 shows atomic distribution of radial position in the entrance plane to the storage bulb. X-axis represents radial position and Y-axis is about atomic number. Obviously effective atoms are focused in −1 mm–1 mm and ineffective atoms do not reach to the entrance plane to the storage bulb for those atoms run into the wall of magnetic state selector and are not considered.

Atomic distribution of radial position in the entrance plane to the storage bulb is related to atomic initial angle of atoms out of the collimator, shown in Fig. 4. X-axis represents initial angle of hydrogen atoms and Y-axis is about radial position in the entrance plane to the storage bulb. This curve indicates 85% of effective atoms reach the entrance plane and its initial angles are between −35 mrad and 35 mrad but only 82% of effective atoms enter into the storage bulb, if we assume the entrance diameter of the storage bulb is 4 mm. The impact of the geometry and properties of magnetic state selector on the number of effective atoms are discussed in the follow.

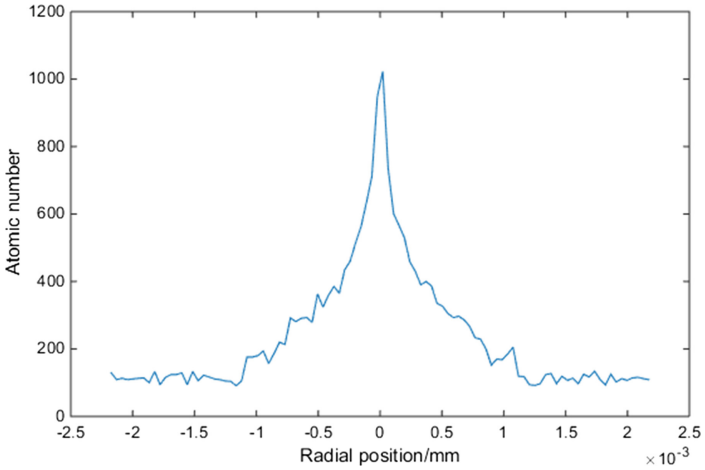


Fig. 3. Atomic distribution of radial position in the entrance plane to the storage bulb. Simulation condition: N = 100000, v = 3000, H_0 = 1T, r_0 = 1 mm, L_2 = 59 mm, L_3 = 120 mm.

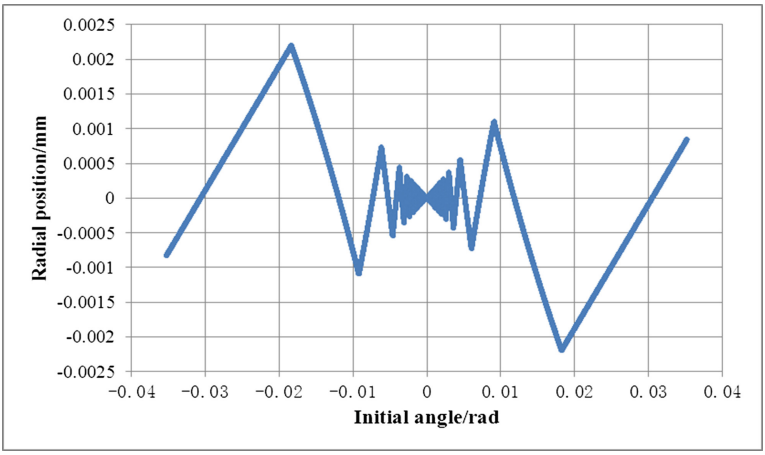


Fig. 4. Atomic initial angle versus atomic distribution of radial position in the entrance plane to the storage bulb. Simulation condition: N = 100000, v = 3000, H_0 = 1T, r_0 = 1 mm, L_2 = 59 mm, L_3 = 120 mm

3 Discussion

Based on the above method, first simulate the results with different poletip field H_0. The results are shown in Fig. 5. We can learn that the more atoms reach the entrance plane with H_0 increasing. But the number of atoms entering into the storage bulb is not linear with increased H_0, for the entrance diameter of the storage bulb limits a part of atoms enter into the bulb. Take poletip field H_0 = 1.5T as an example. Effective atoms of

initial angels between −43mrad and 43mrad can reach the entrance plane. The maximum value of atomic radial position is 3.3 mm and obviously this is larger than 2 mm which is entrance radius of the storage bulb, so atoms whose radial position are between 2 mm and 3.3 mm do not enter into the storage bulb as well as atoms between −2 mm and −3.3 mm. For H_0 = 1.5T, only 80.26% of effective atoms enter into the storage bulb. Figure 6 shows the percent of effective atoms entering into the storage bulb with increased poletip field H_0. The maximum value is 83.02% at H_0 = 0.9T in Fig. 6. The percent values increase sharply with H_0 from 0.5T to 0.9T, while the values decrease slowly with H_0 between 0.9T to 1.5T. 0.9T is optimum for magnetic state selector in this simulation condition.

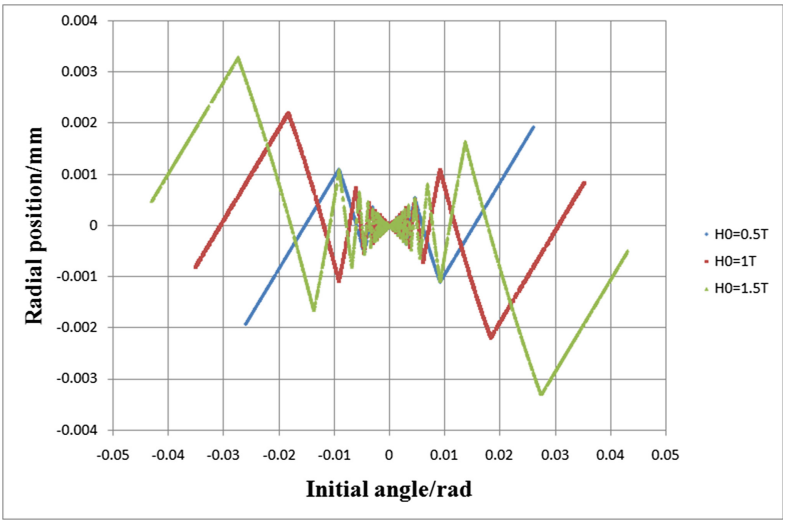


Fig. 5. Different Poletip fields impact on atomic distribution in the entrance plane to the storage bulb. Simulation condition: N = 10000, v = 3000, r_0 = 1 mm, L_2 = 59 mm, L_3 = 120 mm.

L_3, a distance from the exit plane of the magnet to the entrance plane to the storage bulb, plays an important role in miniaturizing H-maser. Figure 7 shows different L_3 impact on atomic distribution in the entrance plane to the storage bulb. From those curves, obviously the initial angles of atoms reaching the entrance plane are the same for different L_3. But the maximum value of atomic radial position increase with increased L_3. The number of atoms entering into the storage bulb is the same when L_3 is between 60 mm and 100 mm, and the number decreases sharply when L_3 is more than 100 mm, if we still assume the entrance diameter of the storage bulb is 4 mm. The reason is the maximum value of atomic radial position larger than 2 mm which is entrance radius of the storage bulb, so some atoms do not enter into the storage bulb. L_3 According to this consequence, we can optimize L_3 and reduce the volume physical package of H-maser.

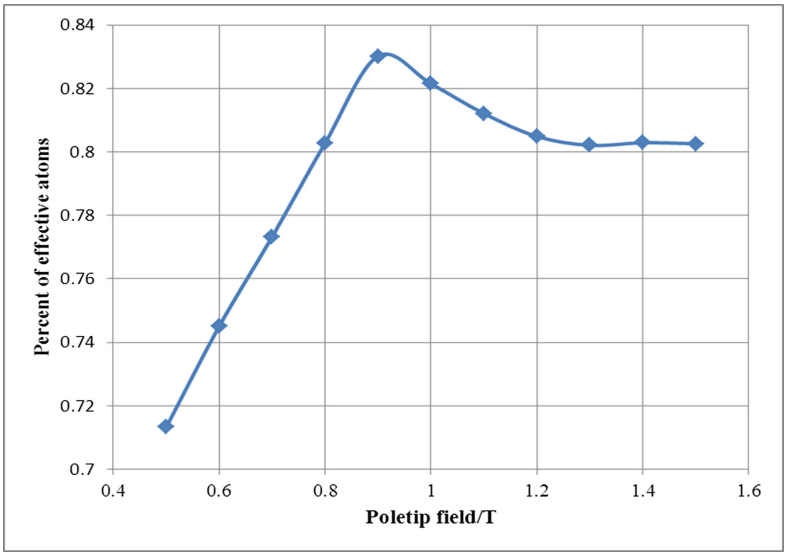


Fig. 6. Poletip field H_0 versus the percent of effective atoms entering into the storage bulb.

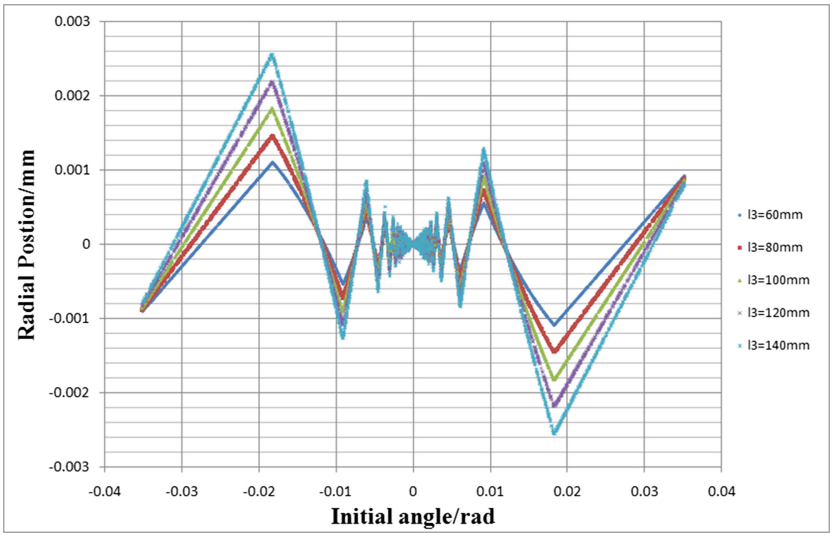


Fig. 7. Different L_3 impact on atomic distribution in the entrance plane to the storage bulb. Simulation condition: N = 10000, v = 3000, H_0 = 1T, r_0 = 1 mm, L_2 = 59 mm, L_3 = 60 mm.

4 Conclusions

In this paper, a random method to simulate atomic behavior in magnetic state selector of H-maser is presented. Initial angle, velocity and states of atoms effusing out of discharge dissociator are taken into consideration in this simulation model. Using this method, the trajectory of each atom and atomic distribution in the entrance plane to the storage bulb is provided. The impacts of the geometry and properties of magnetic state selector on the results are also discussed. We take H_0 and L_3 as examples and discuss the number of effective hydrogen atoms which are focused into storage bulb. When $H_0 = 0.9T$, the number of effective atoms is at most and the optimum value of L_3 is between 60 mm and 100 mm in the simulation condition. This approach can be possessed to discuss other parameters impacts on the results to optimize the performance of H-maser and to solve the similar random problem in other system.

References

1. Goldenberg HM, Ramsey NF (1962) Phys Rev 126(2):603–615
2. Peters H, Owings B, Oakley T (1987) 41st annual frequency control symposium, pp 75–81
3. Vessot RFC (2005) Metrologia 42:S80–S89
4. Peters HE (1981) Proceedings of the 13th annual precise and the time interval application and planning meeting, pp 645–665
5. Wang Y, Wang Q, Fu J, Dong T (1986) Principle of quantum frequency standard, pp 282–287. Science Press, Beijing (Ch)

Author Index

J. Sun et al. (Eds.): CSNC 2019, LNEE 562, pp. 595–597, 2019.
https://doi.org/10.1007/978-981-13-7751-8