Lecture Notes in Electrical Engineering 304

Jiadong Sun
Wenhai Jiao
Haitao Wu
Mingquan Lu
Editors

China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II

Springer

Lecture Notes in Electrical Engineering

Volume 304

For further volumes:
http://www.springer.com/series/7818

About this Series

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

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

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer's high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

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

Jiadong Sun · Wenhai Jiao
Haitao Wu · Mingquan Lu
Editors

China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II

Springer

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

Wenhai Jiao
China Satellite Navigation Office
Beijing
China

Haitao Wu
Navigation Headquarters
Chinese Academy of Sciences
Beijing
China

Mingquan Lu
Department of Electronic Engineering
Tsinghua University
Beijing
China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
ISBN 978-3-642-54742-3 ISBN 978-3-642-54743-0 (eBook)
DOI 10.1007/978-3-642-54743-0
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014937269

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

China's BDS Navigation Satellite System (BDS) has been independently developed, which is similar in principle to global positioning system (GPS) and compatible with other global satellite navigation systems (GNSS). The 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 5th China Satellite Navigation Conference (CSNC 2014) is held on May 21–23, 2014, Nanjing, China. The theme of CSNC 2014 is BDS Application—Innovation, Integration and Sharing, which covers a wide range of activities, including technical seminars, academic exchange, forum, exhibition, lectures, as well as ION panel. The main topics are as:

1. BDS/GNSS Navigation Applications
2. Satellite Navigation Signal System, Compatibility and Interoperability
3. Precise Orbit Determination and Positioning
4. Atomic Clock Technique and Time-Frequency System
5. Satellite Navigation Augmentation and Integrity Monitoring
6. BDS/GNSS Test and Assessment Technology
7. BDS/GNSS User Terminal Technology
8. Satellite Navigation Models and Methods
9. Integrated Navigation and New Methods

The proceedings have 171 papers in nine topics of the conference, which were selected through a strict peer-review process from 479 papers presented at CSNC 2014.

We thank the contribution of each author and extend our gratitude to 165 referees and 36 session chairmen who are listed as members of an editorial board. The assistance of CNSC 2014's organizing committees and the Springer editorial office is highly appreciated.

Jiadong Sun

Editorial Board

The 5th China Satellite Navigation Conference (CSNC 2014)

Scientific Committee

Chairman
Jiadong Sun, China Aerospace Science and Technology Corporation

Vice-Chairman
Rongjun Shen, China
Jisheng Li, China
Qisheng Sui, China
Zuhong Li, China Academy of Space Technology
Shusen Tan, Beijing Satellite Navigation Center, China

Executive Chairman
Jingnan Liu, Wuhan University
Yuanxi Yang, China National Administration of GNSS and Applications
Shiwei Fan, China

Committee Members (By Surnames Stroke Order)
Xiancheng Ding, China Electronics Technology Group Corporation
Qingjun Bu, China
Liheng Wang, China Aerospace Science and Technology Corporation
Yuzhu Wang, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences
Guoxiang Ai, National Astronomical Observatories, Chinese Academy of Sciences
Shuhua Ye, Shanghai Astronomical Observatories, Chinese Academy of Sciences
Zhaowen Zhuang, National University of Defense Technology
Qifeng Xu, PLA Information Engineering University
Houze Xu, Institute of Geodesy and Geophysics, Chinese Academy of Sciences
Guirong Min, China Academy of Space Technology
Xixiang Zhang, China Electronics Technology Group Corporation
Lvqian Zhang, China Aerospace Science and Technology Corporation

Junyong Chen, National Administration of Surveying, Mapping and Geoinformation
Benyao Fan, China Academy of Space Technology
Dongjin Luo, China
Guohong Xia, China Aerospace Science and Industry Corporation
Chong Cao, China Research Institute of Radio Wave Propagation (CETC 22)
Faren Qi, China Academy of Space Technology
Sili Liang, China Aerospace Science and Technology Corporation
Shancheng Tu, China Academy of Space Technology
Rongsheng Su, China
Zhipeng Tong, China Electronics Technology Group Corporation
Ziqing Wei, Xi'an Institute of Surveying and Mapping

Organizing Committee

Secretary General
Haitao Wu, Navigation Headquarters, Chinese Academy of Sciences

Vice-Secretary General
Wenhai Jiao, China Satellite Navigation Office
Tao Peng, Academy of Satellite Application, China Aerospace Science and Technology Corporation
Wenjun Zhao, Beijing Satellite Navigation Center
Tao Peng, Satellite Application Research Institute of China Aerospace Science and Technology Corporation
Chuang Shi, Wuhan University
Weina Hao, Navigation Headquarters, Chinese Academy of Sciences

Committee Members (By Surnames Stroke Order)
Qun Ding, Xi'an Institute of Navigation Technology
Miao Yu, Beijing Institute of Space Science and Technology Information
Yinhu Ma, Navigation Satellite Systems Engineering Center, China Academy of Aerospace Electronics Technology
Ying Liu, China Satellite Navigation System Administration Office
Hao Zhang, International Cooperation of China Satellite Navigation System Administration Office
Wangming Yang, Nan Jing High Technology Industrial Development Zone Management Committee
Mingquan Lu, Tsinghua University
Xiuwan Chen, Peking University

Shuanggen Jin, Shanghai Astronomical Observatories, Chinese Academy of Sciences
Xiang'an Zhao, China Defense Science and Technology Information Center
Jing Zhao, Ministry of Science and Technology National Remote Sensing Center
Yuehu Hu, China Aerospace Science and Technology Corporation 20th Institute
Zhong Dou, National Time Service Center of Chinese Academy of Sciences

Contents

Part III Satellite Navigation Models and Methods

Part I
Satellite Navigation Signal System, Compatibility and Interoperability

Chapter 1
The Research of the High-Altitude BDS Weak Signal Acquisition

Changyuan Wen, Fuzhan Yue, Yuehua Qiu and Ting Ke

Abstract Aiming at the application for High-Earth orbit spaceflight navigation, the paper is about the weak BDS signal acquisition on High-Earth orbit. First, this paper achieves the receiving power and Doppler-shift of the High-Earth orbit receiver by simulating the navigation signal characteristic which based on the STK (Satellite Tool Kit). In order to solve the problem of the secondary code witch modulated on BDS B1I and improve the acquisition sensibility, this paper analyses the autocorrelation characteristic of the data which modulated with the NH code, comes up with the resolution of using the alternate half-bit method by combining the coherent integration with the non-coherent integration. The result indicates that it can successfully acquire the −178 dBW BDS B1I signal under the mentioned method.

Keywords High-earth orbit · Weak signal acquisition · Coherent/Non-coherent · The alternate half-bit method · Neumann–Hoffman code searching

1.1 Introduction

High Earth Orbit (HEO) including Geostationary Orbit (GEO) and Highly Eccentric Earth Orbital (HEEO), are generally higher than the maximum height of its orbit 20,000 km. In recent years, in terms of land and sea communications, meteorological, educational applications, live television, disaster warning, HEO plays an important role. The high-altitude spacecraft navigation positioning and orbit determination has become one of the urgent needs to develop new technology. The satellite navigation system with the characteristics of global, all-weather,

C. Wen (✉) · F. Yue · Y. Qiu · T. Ke
Space Star Technology Co., Ltd, Beijing 10086, China
e-mail: achang11241124@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_1,

continuous and high precision has gradually become the principal means of navigation for positioning and orbit determination [1].

But due to high-altitude spacecraft orbital altitude higher than that of navigation constellation, needs to receive signals from the earth the other side of the navigation satellite receiver. Keeping out of the earth as well as propagation distance increasing lower the signal received power, which lead to reduce the presence of the available satellite, short duration of available navigation signals and weak receiving signal power by using the regular technology.

The NASA has a series of flight test for high-altitude GPS receiver application, developed such as the Pivot and the Navigator receiver for deep space exploration. High-altitude test shows that on high-altitude receiving power will be lower than the ordinary spaceborne receiver power more than 10 dB, doppler frequency shift in L1 band will be around 10 KHZ, the Navigator which uses 10 ms coherent integration when $C/N_0 = 25$ dB-Hz successful acquisition probability is only 50 %, larger deficiencies still exist for high-altitude applications [2]. The literature [3] studied the GNSS high-altitude satellite positioning related factors, but did not give a reasonable signal processing scheme .the literature [4] put forward navigation constellation with navigation signal constellation outside for locating scheme, but involved navigation constellation space redesign, implementation is difficult at present.

The above studies were all for the GPS orbit application, and domestic and foreign high-orbit applications are not seen on the BDS literature. This paper aiming high-altitude weak navigation signal with short duration carries out the high-orbit spacecraft BDS weak signal acquisition technology research.

1.2 BDS Navigation Signal Characteristics on High Earth Orbit

High-altitude navigation receiver needs to receive navigation signals from across the earth. Take GEO orbit for example, under the analysis of space geometric relationships between high-altitude with BDS MEO satellite constellation, and calculating the corresponding power link. Figure 1.1, the ε for BDS satellite antenna of unilateral coverage point of view, Angle φ for BDS satellite main beam unilateral, Angle θ for BDS satellite and earth tangent of unilateral. α for receiver and BDS satellite connection with the receiver to the Angle between the direction of the centre of the earth, β for receiver and BDS connect with BDS satellites pointing in the direction of the Angle between the direction of the center of the earth. In diagram of AC segment caused by earth sheltered, signal is not available.

Space navigation signal propagation link budget equation is:

$$P_R = P_T + G_T + G_R + 20\log\left(\frac{\lambda}{4\pi d}\right) - L_A \tag{1.1}$$

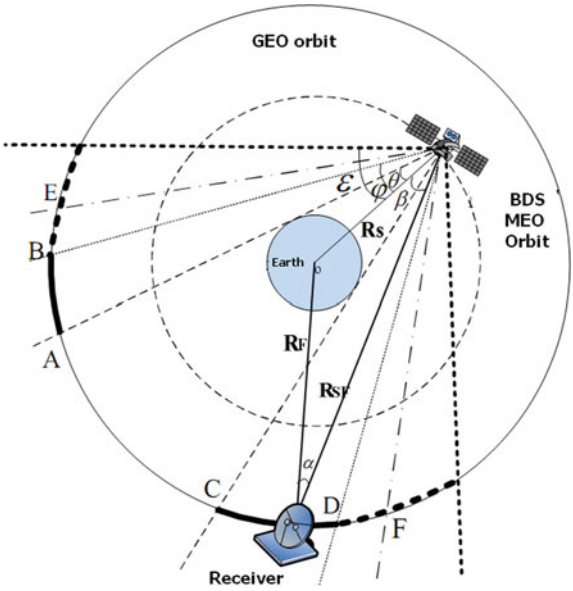


Fig. 1.1 The geometry model between GEO and BDS

P_R for the received power of the receiver; P_T for BDS satellite transmission power; G_R and G_T respectively BDS antenna gain and receive gain; $20\log_{10}(\lambda/4\pi d)$ free space for the signal propagation loss, linear distance d between receiver and navigation satellite, λ for signal transmission wavelength; L_A loss of the atmosphere when the navigation signal through the atmosphere. BDS B1 frequency carrier frequency of 1561.098 MHz; L_A is negligible in high-altitude computation.

As a result of the BDS EIRP undisclosed, this paper uses the GPS IIR satellite L1 frequency EIRP for reference, receiving antenna using 0 dB gain, as shown in Fig. 1.2. Analysis of the introduction of the concept of equivalent gain here, link equivalent gain is the launch of the navigation constellation gain characteristic, the decline of signal propagation characteristics, caused by the receiving antenna stance change characteristics and receives all projection to the receiving terminal antenna gain characteristics. Through equivalent gain to include all gain characteristic in the navigation signal link, at the same time meet the demand of visualization, dynamic simulation input equivalent gain, specific simulation Settings see reference [5].

Keeping out because of the earth would cause the equivalent gain under 8.3° shadow area, the receiving signals mainly concentrates in the angle between 8.3° and +30°. The gain distribution shows that to receive navigation satellite side lobe signal, the receiving power must be below −174 dBW, the power demands of the second side lobe reception is below −178 dBW. Using STK, with the global layout after completion of construction of BDS constellation, analysis GEO orbit corresponding to the distribution of visible satellite BDS. BDS complete constellation

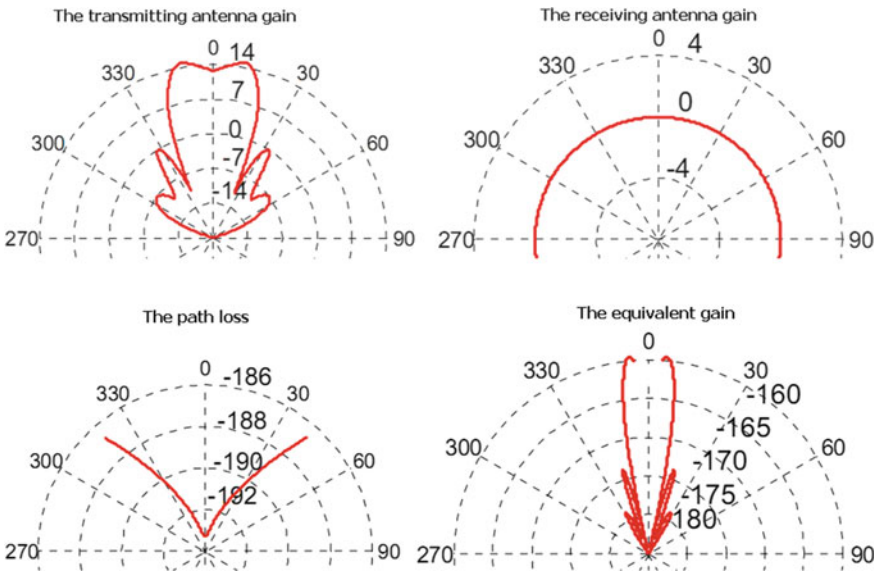


Fig. 1.2 The equivalent gain of BDS on GEO

Table 1.1 The number of the BDS on GEO in different power

Segment percentage (%)	>2	>3	>4	>5	>6	>7	>8
−173 dBW	18.53	0.97	0	0	0	0	0
−174 dBW	20.54	2.91	0	0	0	0	0
−175 dBW	21.24	4.3	0	0	0	0	0
−176 dBW	92.37	68.15	37.61	14.71	5.2	0.62	0
−177 dBW	100	100	99.53	91.7	75.78	64.68	51.77
−178 dBW	100	100	100	100	100	100	99.51
−179 dBW	100	100	100	100	100	100	100

configuration according to the ICD document is set to 5 GEO satellite, 27 MEO satellite and 3 IGSO satellites [6]. GEO receiver points east longitude 120° (Table 1.1).

Can be seen from the above statistics, GEO orbit in the receiving power −178 dBW, 4 satellites above the segment accounts for 100 % of all the way, the following simulates the value of the corresponding PDOP under −178 dBW (Figs. 1.3, 1.4, 1.5).

According to the result of simulation shows that applies to the GEO BDS reception power should guarantee at least −178 dBW, doppler frequency shift in ±12 kHz, which can satisfy the GEO orbit navigation application. In addition, the duration of certain signals will be less than 15 min or less, so the signal acquisition should be faster.

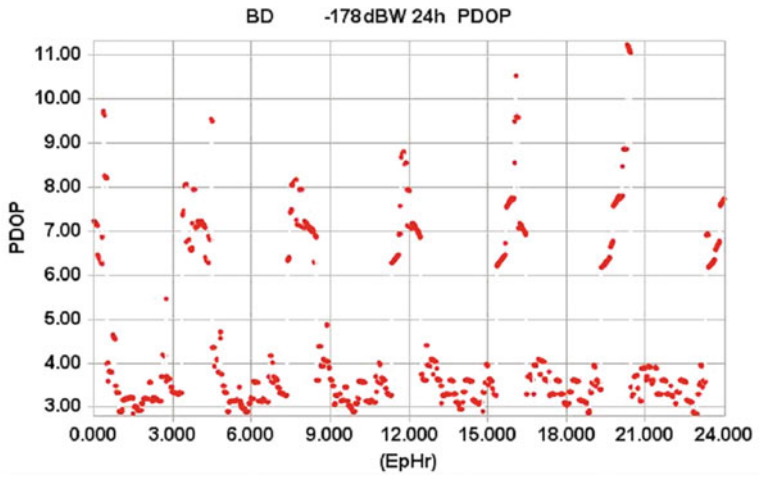


Fig. 1.3 The PDOP value on GEO

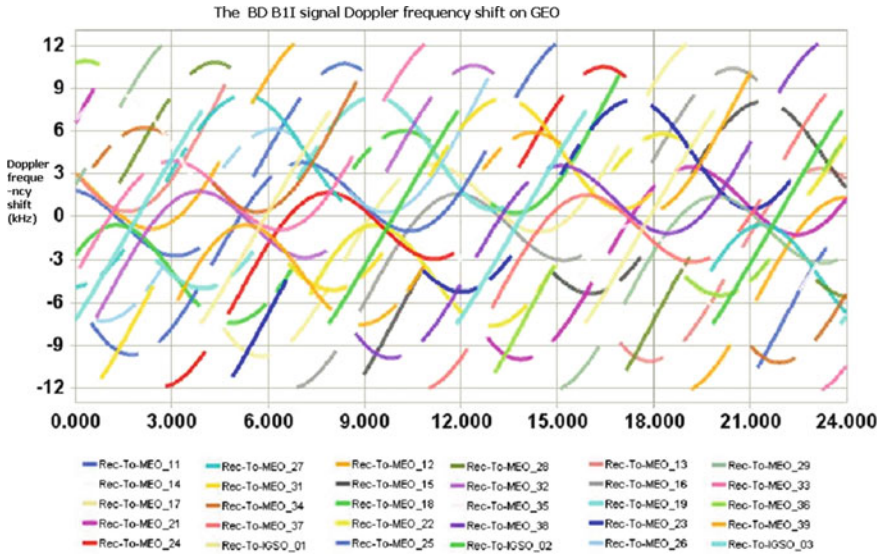


Fig. 1.4 Doppler of the BDS B1 on GEO in 24 h for −178 dBW

1.3 BDS Weak Signal Acquisition Technology Analysis

Navigation signal acquisition is essentially the navigation signal doppler ranging code phase and carrier frequency of the two-dimensional search process. From the previous section to receive high-altitude navigation signal feature analysis, in

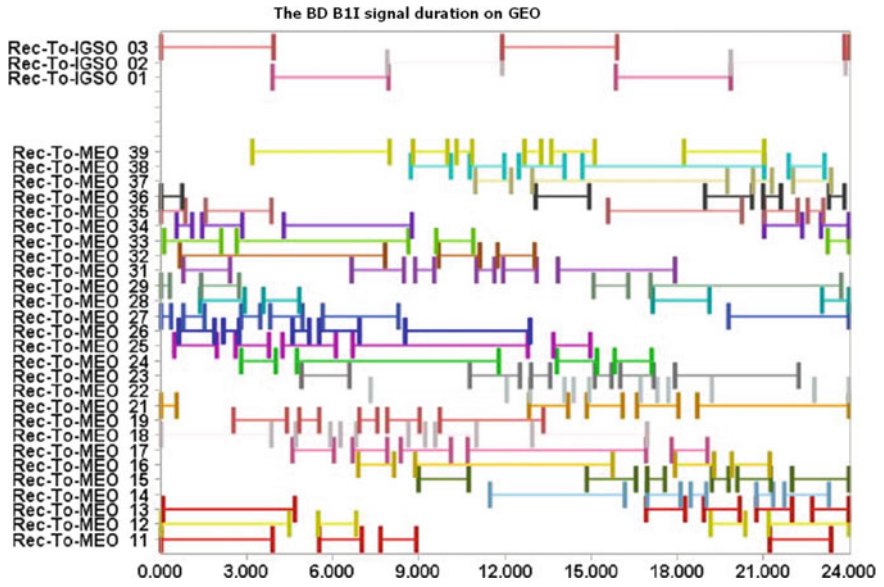


Fig. 1.5 The duration time of the visible BDS on GEO

order to achieve positioning throughout the GEO orbit, the power must be at least −178 dBW, in addition to consider frequency mismatch and the loss of code synchronization, so it need extra gain for baseband processing.

Now the methods of increasing the baseband signal processing gain include the coherent integration and non-coherent integration. The biggest difference between BDS and GPS in weak signal acquisition is that BDS B1I has secondary coding, which will lead to coherent integration length limited. The following introduces processing gain and secondary coding solutions.

1.3.1 The Integration Processing Gain

1.3.1.1 The Coherent Integration and Non-coherent Integration Gain

Coherent integration is the sum of the square results of N correlator, to enhance the signal-to-noise ratio. In the navigation signal acquisition, coherent integration is the first step, the coherent processing gain can be calculated by the bandwidth, and bandwidth is related to the data length. No-coherent integration with coherent integration results can output extra SNR gain. Non-coherent integration uses the square root of sum of squares of the I and Q coherent integration. Square arithmetic is the cause of signal energy increase, at the same time, the noise energy also will increase, and the resulting SNR loss called square loss.

Table 1.2 The gain and data length for different methods of the integration unit (dB/ms)

	1 ms	2 ms	5 ms	10 ms	20 ms
1 times	0/1	3/2	6.99/5	10/10	13/20
2 times	2.75/2	5.75/4	9.75/10	12.75/20	15.75/40
5 times	6.13/5	9.13/10	13.12/25	16.13/50	19.13/100
10 times	8.47/10	11.47/20	15.46/50	18.47/100	21.47/200
20 times	10.61/20	13.61/40	17.6/100	20.61/200	23.61/400
25 times	11.26/25	14.26/50	18.25/125	21.26/250	24.26/500

Normally, using coherent integration with non-coherent way to enhance the gain, the low SNR signal which after many Non-coherent integration, can be better to separate signal and no signal by the signal amplitude probability density curve of valuations. The Table 1.2 for different coherent integration time with different number of non-coherent integration corresponding gain value as well as the required data length, the horizontal axis represents the cumulative coherence length of the longitudinal axis of a non-coherent integration times, the $P_{fa} = 1e - 7$, $P_d = 0.99$.

From the graph, the same processing time, the longer the coherent integration time, bring the gain of the greater appreciation, therefore, it need to consider in the design of general extension of the coherent integration time length as much as possible.

The error estimate of the actual value from the Doppler shift caused by the carrier can be expressed as

$$L_{acq} = 20\log_{10}[\sin c(\Delta fT)] \tag{1.2}$$

Δf is carrier frequency error, T is coherent integration time.

The average code synchronization loss of $10 \times \log10\,(1/8)\,(1 - 2) = 1.16$ dB.

1.3.2 Secondary Code Features

BDS MEO/IGSO B1I signal has a secondary code yards wide 1 ms. In condition of the secondary code phase unknown, it will limit the coherent integration time length, and influence gain promotion effect. BDS secondary code refers to the rate of 50 bps navigation message on Neumann–Hoffman code modulation (NH code correlation value maximum, when the complete synchronization features as shown in Fig. 1.6a). NH code cycle for a bit the width of the navigation information, NH code 1 bit width is the same as the ranging code cycle. D1 navigation message a bits of information in the width of 20 ms, ranging code cycle is 1 ms. NH code is with the type (0,0,0,0,0,1,0,0,1,1,0,1,0,1,0,0,1,1,1,0) and 1 kbps code rate, the code width is 1 ms [6].

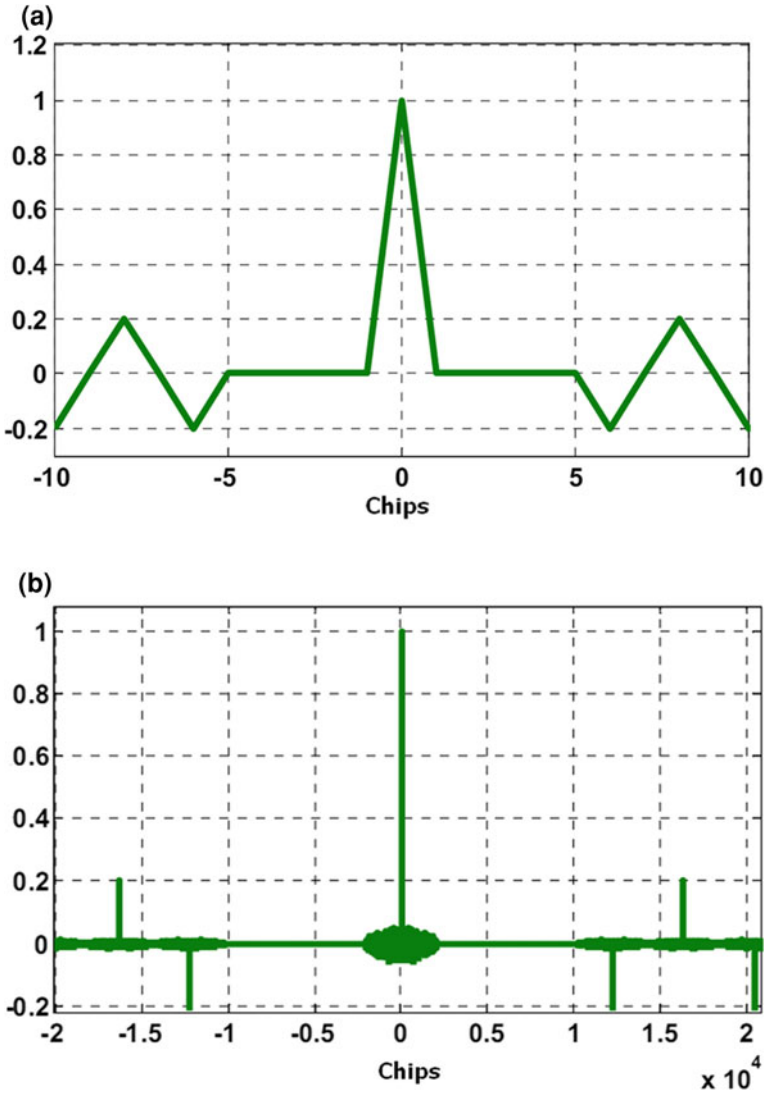


Fig. 1.6 The characteristic of NH code. **a** The autocorrelation of NH code **b** The autocorrelation of 20 ms data with NH code

NH Code extends the period ranging code BDS, and improves the ranging code autocorrelation and cross-correlation properties. But by the Fig. 1.6b the simulation results can be seen, although secondary NH code modulation brought autocorrelation side-lobe to fold the overweight reduced overall, but in the individual code phase delay autocorrelation side-lobe can reach 14 dB (0.2 times that of the main peak), will be a certain effects on the BDS signal capture performance [7].

1.3.3 Weak Signal Acquisition Process

BDS B1I weak signal processing, it is necessary to extend the coherent integration time length as long as possible, and the unknown NH code phase can limit the length of time of coherent integration, influence gain promotion effect. From Fig. 1.6b NH code modulation 20 ms data characteristics, the modulation NH code after the data stack of overweight has good autocorrelation properties, so under the condition of NH code phase unknown, 20 ms data can be correlated, by detecting the corresponding correlation peak to decide whether captured successfully.

This process needs to consider the navigation message data bit transition, so every 20 ms will be affected by the transition of the navigation message. Here use the 'alternate semi-bit method' [10], take any point, behind this point 20 ms received data is divided into two sections 10 ms, guaranteed of a data bit navigation message data is completely in the same time. Then calculated using the non-coherent integration, the maximum output of both is the corresponding estimates.

Some high-altitude signal duration is shorter, in order to guarantee the speed of acquisition considering the fast acquisition based on FFT/IFFT, determine the BDS B1I weak signal capture flow chart shown in Fig. 1.7, the input time length for 10 ms signals are stored, and zero padding to FFT transform after 20 ms. Local 20 ms modulation with NH code of conjugate FFT transform after the data, and the above after FFT transform is multiplied by the input data, then the results are IFFT transformation. Use semi-bit method, followed by the odd and even number sequences incoherent sequences corresponding cumulative times, the last successful capture judged by peak detection signal.

Usually the frequency interval of the frequency resolution to capture the maximum of 2/3, resulting in frequency mismatch loss of −1.65 dB, an average loss of synchronization code −1.16 dB, considering a certain margin, the frequency mismatch, the code will not result in the alignment of at least 3 dB loss. Baseband processing 21 dB gain must be provided to complete the capture of the weak signal, the coherent integration of 10 ms gain $10\log_{10}(10) = 10$ dB, $P_{fa} = 10^{-7}$ false-alarm probability and detection probability $P_d = 0.99$, 25 times non-coherent integration gain of 11.26 dB, can satisfy the requirements of the baseband signal processing.

1.4 Algorithm Performance Simulation Analysis

For simulating the navigation signal acquisition algorithm performance, the simulation is modeled by Matlab Simulink, with the simulation parameters in the Table 1.3 (Fig. 1.8).

According to the simulation results, the power −178 dBW, 10 ms coherent integration with the 25 non-coherent accumulation, can successfully capture the BDS B1I signal. According to the results in the code piece of delay capture peak,

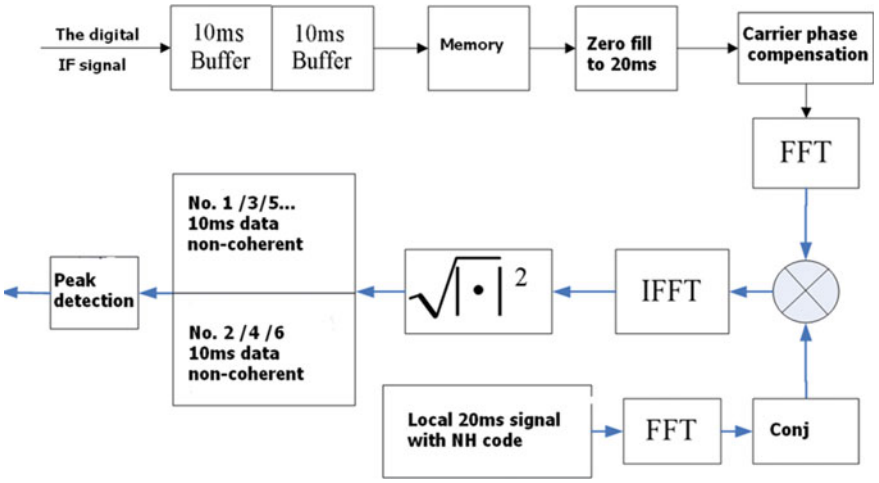


Fig. 1.7 The block diagram of the acquisition

Table 1.3 The parameter of the acquisition simulation

Parameters	Value
Signal	BDS B1I RPN = 7
Power	−178 dBW (26 dB/Hz)
Code delay	500 chips (NH code delay: 15 chips)
Doppler	10 kHz
Coherent integration data length	10 ms
Non-coherent integration times	25

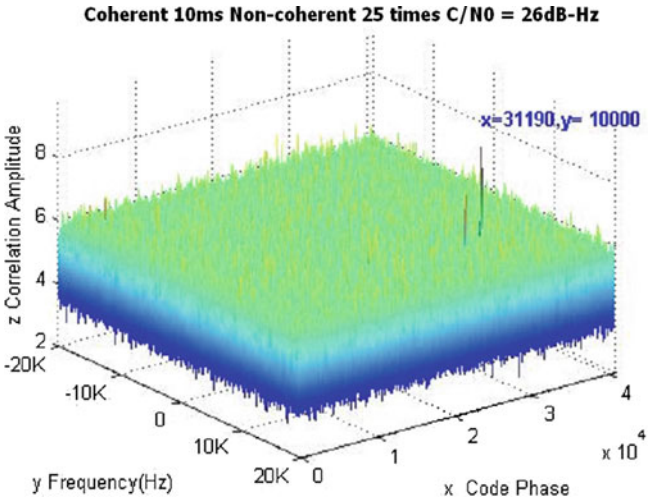


Fig. 1.8 The result of acquisition for 26 dB-Hz BDS signal

total 20 ms data code length of 40,920 chips, 31,190 and 2,046 modulus, get corresponding ranging code delay of 500 chips, integer part is 15, namely NH code of the code phase of 15 chips.

This scheme uses the autocorrelation characteristic of the NH modulation code data, avoid to search of NH code phase, reduce the dimension of the search and processing time. But 10 ms data directly to FFT, the amount of data processing is too large, it can use sectional FFT processing, reducing the consumption of hardware logic resources [8].

1.5 Summary

This main job: (1) through the equivalent gain simulation, proposed indicators to capture GEO orbit weak signal (carrier to noise ratio, Doppler range). (2) Verified the feasibility of removing the NH code algorithm for weak signal acquisition. Simulation results show −178 dBW sensitivity, able to meet the high-orbit spacecraft navigation applications, the use of alternative methods for semi-bit data is modulated 20 ms NH code to search FFT/IFFT-based approach, able to successfully capture the signal. Weak signal acquisition scheme is not considered BDS GEO satellite broadcast message information rate 500 bps of D2 for high-altitude applications, it can use the navigation message stripped ground prior information provided to improve sensitivity.

This paper mainly aims at the research of the High-Altitude BDS weak signal acquisition, After the successful capture of the actual code phase and get a rough estimate of the Doppler frequency, the signal needs to be sent after the capture signal tracking loop. High-Altitude under the environment of weak signal tracking is different from traditional loop design; the subsequent will follow the above work.

References

1. Davis G, Moreau M, Carpenter R, Bauer F GPS-based navigation and orbit determination for the AMSAT AO-40 satellite. In: AIAA, ID: 20030025378
2. Bamford WA, Heckler GW (2009) A GPS receiver for high-altitude satellite navigation. IEEE J Sel T Sign Proces 3(4):599–612
3. Honglei Q, Minmin L (2008) Study on high orbit satellite positioning technology based on GNSS. J Space Sci 28(4):316–325
4. He QJ, Sun Q (2010) New scheme to realize high orbit satellite autonomous navigation using GNSS. J Spacecr TT&C Technol 8
5. Yue F, Liang Z, Lv T, Qiu Y (2012) Analysis and simulation research of method of full link receiver for GPS satellite. The third annual meeting of China satellite navigation, 2012
6. BDS-SIS-ICD-2.0 http://www.beidou.gov.cn//2013/12/26
7. Xu D, Kou Y (2011) The dual component combined tracking of GPS L5 signal. The second Chinese satellite navigation conference special for 2011
8. Psiaki ML (2001) Block acquisition of weak GPS signals in a software receiver. In: Proceedings of Institute of Navigation GPS, Sep 2001, pp 2838–2850

Chapter 2
GNSS Message Structure Optimization

Yun Zhou, Yansong Meng, Xiaoxia Tao, Lei Wang and Zhe Su

Abstract The structure designing of satellite clock correction,ephemeris and system time parameters in navigation message directly affects navigating performance such as the time to first fix. Considering the message reading point as a random variable and creating ephemeris acquisition time as a function of the reading time, we integrated the acquisition time of the ephemeris in GPS L2C CNAV random data block and L1C CNAV2 stationary message structure by 95 % confidence level. The designing discrepancy of the GPS, Galileo data structure which is broadcasted on different frequency and its influence to the first time to access positioning information was emphatically discussed. Finally, we presented the key factor of navigation message structure designing to reduce the time to first fix that was ephemeris being consistent during effective time interval, almanac frame proportion of the total frame size should be appropriate and message parameters should be more compact.

Keywords GNSS message structure · Time to first fix

2.1 Introduction

Time to first fix (TTFF) is an important performance evaluating navigation message structure design. Satellite clock correction, ephemeris (CED) and system time (GST) directly affect the TTFF. Specifically, the acquisition time of CED, GST is determined by the total amount of data, broadcast rate, repetition and structure. Broadcasted on GPS L2C in the form of data blocks, CNAV message's CED and

Y. Zhou (✉) · Y. Meng · X. Tao · L. Wang · Z. Su
Institute of Space Radio Technology, Xi'an, China
e-mail: nikkochou@163.com

Y. Meng
e-mail: iamyantao@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_2,

Table 2.1 Max broadcast interval of GPS L2C CNAV

Message data	Message type	Max intervals (s)
Ephemeris	10, 11	48
Clock	Type 30's	48

Fig. 2.1 One arrangement of GPS L2C CNAV data block

GST parameters are designed in 10, 11 and 30 types of data blocks, according to requirement to broadcast, while CNAV2 on GPS L1C frequency planning to use three unequal sub-frame to broadcast messages, and positioning information such as CED and GST being designed for two parts which one varied and the other one does not in the active interval of time. For Galileo I /NAV and F /NAV adopt the way of page to broadcast messages.

Discussing the message structure differ from GPS and Galileo and its influence to TTFF, drawing lessons from the advantages of them, we proposed the message structure designing elements to reduce the time to first fix.

2.2 L2C CNAV Parameter Acquisition Time

Satellite clock correction, ephemeris and system time parameters of GPS L2C CNAV are arranged in paragraphs 10, 11 and 30 types of data blocks, each block lasting 12 s. Maximum broadcasting interval for each data block was shown in Table 2.1.

Assumed that type 10, 11 and 30 data block was broadcasted only once within 48 s, then had the following four conditions, which in addition to 10,11and 30, the another data block called *other type*. A kind of arrangement situation was shown in Fig. 2.1.

Each data block of GPS L2C CNAV was arranged by data bits following parity bits, and one data bit not being obtained would lead to cyclic redundancy check (CRC) checksum failure that also resulted in the whole data block bit stream being invalid, as shown in Fig. 2.2. Then, we should emphatically analyze the acquisition time of CED and GST parameters at $t = 0^+$, $t = 12^+$, $t = 24^+$, $t = 36^+$ reading point. Specifically the following four cases were analyzed for the first mode of Fig. 2.1.

- Reading *other type* data block from $t = 0^+$
  Assuming that the 10, 11 and 30 data blocks were broadcasted only once within 48 s, so the $T_{CED+GST}$ was equal to 48 s.

Fig. 2.2 CNAV data block composition and reading point

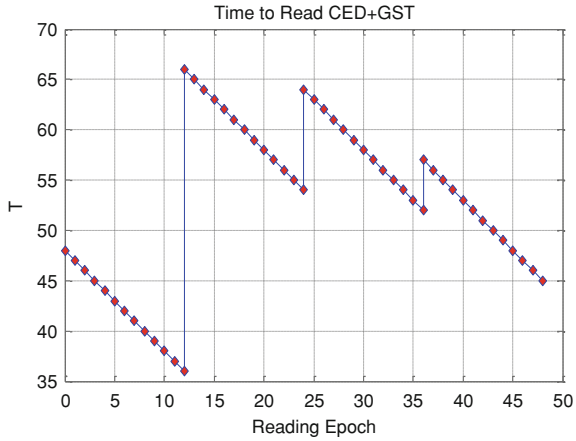


Fig. 2.3 CNAV reading point and CED, GST acquisition time

- Reading *type 10* data block from $t = 12^{+}$.
  of data blocks leading to the other bits being invalid, we should require the data from the next 48 s which was broadcasted by random way, so $T_{10} = (12 + 24 + 36 + 48) \times 25\ \% = 30$ s where T_{10} was indicated for the acquisition time of data *type 10* within the next 48 s. Thus, reading from $t = 12^{+}$, $T_{CED+GST}$was equal to 66 s (30 + 36).
- Reading *type 11* data block from $t = 24^{+}$.
  The data *type 10 and 11* bits being invalid, we should obtain the parameters in data *type 10* and *11* from the next 48 s. The acquisition time of them *was* 24, 3, 48, 48, 48, 36 s respectively by six random arrangements, so $T_{10+11} = 24 \times (1/6) + 36 \times (2/6) + 48 \times (3/6) = 40$ s. Thus, $T_{CED+GST}$ was equal to 64 s (24 + 40).
- Reading *type 11* data block from $t = 36^{+}$.
  The same as above, data *type 30* was ineffective, and we should obtain the parameters from the next 48 s, so $T_{10+11+30} = 48 \times (3/4) + 36 \times (1/4) = 45$ s. Thus, $T_{CED+GST}$was equal to 57 s (12 + 45).

Based on the above analysis of four cases, we draw the reading time t and$T_{CED+GST}$, shown in Fig. 2.3. The curve exhibited step at $t = 0^{+}$, $12^{+}24^{+}$, 36^{+} corresponds to the results of the above analysis and the rest were linear monotonically decreasing.

Assuming that the reading point t was uniformly distributed on a full frame period (here, 48 s), we obtained the probability density function (PDF) by calculating the occurrence frequency of $T_{CED+GST}$within each section described as a

function where T indicated the value of $x(t)$ that was the acquisition time of CED and GST parameters.

$$f(T) = \begin{cases} 1/48 & 36 \le T < 45 \\ 2/48 & 45 \le T < 48 \\ 1/48 & 48 \le T < 52 \\ 2/48 & 52 \le T < 54 \\ 3/48 & 54 \le T < 57 \\ 2/48 & 57 \le T < 64 \\ 1/48 & 64 \le T \le 66 \end{cases} \tag{2.1}$$

Based on 95 % confidence level, the probability density function substituted into the equation (2.2) [1] and we integrated the acquisition time of CED and GST parameters broadcasted on GPS L2C CNAV, and it was 63.8 s.

$$F(T_{CED+GST}) = \int_{-\infty}^{T_{CSE+GST}} f(t)dt = 0.95 \tag{2.2}$$

2.3 L1C CNAV2 Parameter Acquisition Time

The first sub-frame of GPS L1C CNAV2 message comprised the reference time, i.e., the numbers of 18 s interval within 2 hours epoch time, the rest GST parameters including second of week, week number and other system time information were arranged in the second sub-frame as well as the CED information [2]. So, the entire data of CED and GST should be decided by the whole bits in first and second sub-frame. For CNAV2, the data rate was 50 bps, encoded as 100sps. The first sub-frame of it lasted 0.52 s for a total 52 syb (corresponding 9 bits before encoding), the second sub-frame lasted 12 s for 1200 syb (600 bits before encoding), and the third sub-frame lasted 5.48 s for total 548 syb (274 bits before encoding). We considered the channel coding, cyclic redundancy check and the encoded symbol rate to analyze the acquisition time of the CED and GST parameters.

The CED parameters in second sub-frame of CNAV2 remained unchanged in a fairly long period of time (several minutes or even hours). Therefore, when some bits were missed, i.e., not reading the message from the beginning of the second sub-frame, only needed to re-read the missed message from the next corresponding position of the second sub-frame. For example, when reading the message from 1.52 s, the 52 symbols of the first sub-frame and the 100 symbols of the second sub-frame were not acquired, and then only needed to read the rest symbols of the

Table 2.2 GPS L1C CNAV2 reading point and time to access CED and GST

Reading point	Time to CED GST	Reading point	Time to CED GST
0	12.52	9.52	18
0+	18.52	10.52	18
0.52	18	11.52	18
0.52+	18	12.52	18
1.52	18	*12.52+*	18
2.52	18	13.52	17
3.52	18	14.52	16
4.52	18	15.52	15
5.52	18	16.52	14
6.52	18	17.52	13
7.52	18	18	12.52
8.52	18	*18+*	18.52

frame and then read the missed 152 symbols from the next frame, because the 100 symbols of the next second sub-frame were identical to the corresponding positions of the previous frame (the parameters of the first sub-frame might be changed). Therefore, when the reading point from 1.52 s to 12.52 s, the $T_{CED+GST}$ was equal to 18 s which were the interval time of CNAV2. All $T_{CED+GST}$ values were shown in Table 2.2.

Assuming that the reading point t was uniformly distributed on a full frame period (here, 18 s), we obtained the PDF described in Eq. (2.3). Based on 95 % confidence level, we integrated the acquisition time of CED and GST parameters broadcasted on GPS L2C CNAV2, and it was 18.02 s, approximately same as the CNAV2 interval time. It was associated with CED and GST of CNAV2 being designed for two parts which one varied and the other one does not in the active interval of time.

$$f(T) = \begin{cases} 2/18 & 12.52 \le T < 13 \\ 1/18 & 13 \le \mathrm{T} < 18 \\ 11/18 & \mathrm{T} = 18 \\ 2/18 & 18 < \mathrm{T} \le 18.52 \end{cases} \tag{2.3}$$

2.4 Galileo Parameter Acquisition Time

The relationship between message structure and TTFF of GPS NAV, Galileo F/NAV and I/NAV, *Marco Anghileri*a proposed a detailed exposition and we only used the conclusion. To illustrate the impact of the message structure to TTFF, we normalized data rate to 50 bps to compare CED and GST acquisition time at 95 % confidence level. For I/NAV, the data rate was 125 bps, 2.5 times to 50 bps, so

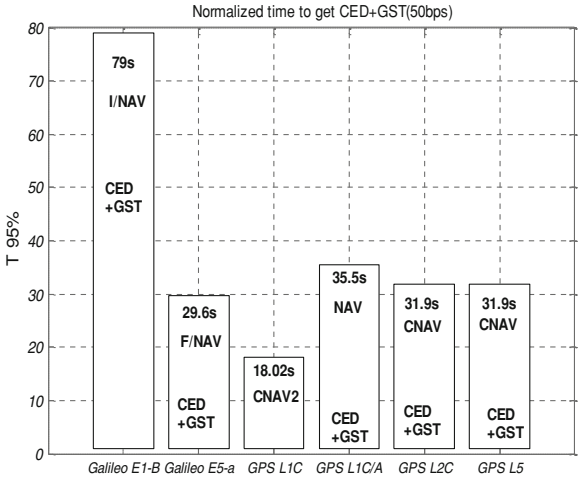


Fig. 2.4 CED and GST acquisition time for 50 bps of different data

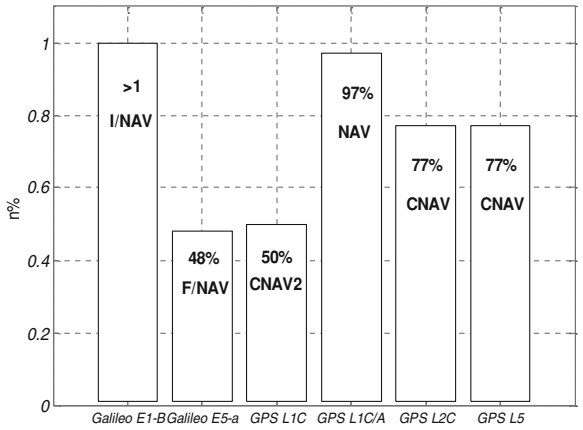


Fig. 2.5 The coefficient for different GPS and Galileo data

$T_{CED+GST}$ was 79 s, while for a data rate of 25 bps, $T_{CED+GST}$ was half of the original value [3]. The normalized acquisition time of GPS and Galileo message broadcasted on different frequency was shown in Fig. 2.4. To get ephemeris and clock parameters, CNAV2 was only 18.02 s, while the I/NAV was 79 s for data rate of 50 bps.

Next, we emphatically analyze the relationship from the message structure, the amount of data, and the repeat time to TTFF. Defined a coefficient η where T_{data} was indicated the minimum time to get CED and GST theoretically, that was the reading point beginning from $t = 0$, and $T_{CED+GST}$ was indicated the statistical time to get all CED and GST firstly by 95 % confidence level, considering the message reading point as a random variable. The coefficient $\eta$of different data was compared in Fig. 2.5.

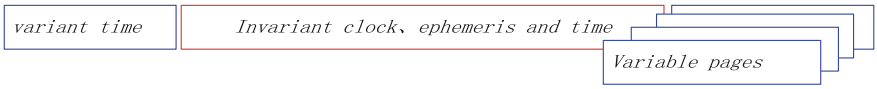


Fig. 2.6 Message structure of GPS L1C CNAV2

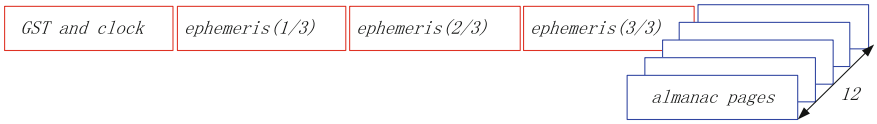


Fig. 2.7 Message structure of Galileo F/NAV

$$\eta = |T_{CED+GST} - T_{data}|/T_{data} \times 100\,\% \tag{2.4}$$

The coefficient of I/NAV was the largest (>1), that was to say its structure design was the worst just from TTFF, and CNAV2 was approximate to F/NAV, being the best of all message structure. In fact, the coefficient also reflected the impact of the structure of almanac to TTFF.

2.5 GPS and Galileo Message Structure Analysis

Comprehensive comparison of Figs. 2.4 and 2.5, the factor $T_{CED+GST}$ of GPS L1C CNAV2 was 18.02 s and the coefficient η was 50 %, and the structure of it had very important reference value, redrawn in Fig. 2.6 [4]. The CED and GST of CNAV2 were compressed into a separate sub-frame of 600 bits which was equivalent to the 900 bits information in 1, 2 and 3 sub-frame of NAV and to the data type 10 and 11 of CNAV. The almanac parameters were designed in a separate sub-frame of 250–300 bits, broadcasted by the way of paging, which was different from the NAV that fourth and fifth sub-frame broadcasted almanac as well as the CNAV that broadcasted almanac by data block manner.

In addition, another important design idea of CNAV2 was considering the consistency of CED and part of GST during the active interval, so when some bits of CED and GST were missed, we could wait for about 56 s (when 50 bps) to obtain the remaining bits from the next frame for the bits obtained from the previous frame still valid. The coefficient η was 48 % for Galileo F/NAV and the structure of it was redrawn in Fig. 2.7 [5].

The frame structure of F/NAV almanac parameters was similar to CNAV2 and the difference was that the former was 250 bits and the later was 274 bits. That was to say, when the data rate was 50 bps, we could obtain the CED and GST parameters from the next frame after 5 s, and for CNAV2 just needed 5.48 s. This also explained the reason why the coefficient of F/NAV (48 %) was slightly lower

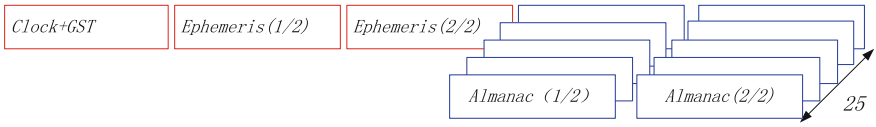


Fig. 2.8 Message structure of GPS L1C/A NAV

Fig. 2.9 Message structure of Galileo I/NAV

than CNAV2 (50 %). However, $T_{CED+GST}$ of F/NAV was more than CNAV2 about 11.6 s with 50 bps data rate. The reason of it was the total bits of CED and GST of F/NAV more than CNAV2 about 400 bits. Furthermore, these bits of F/NAV were divided into four sub-frames to broadcast and each sub-frame utilized the 24-CRC, which caused the consistency decreased and redundancy increased.

The frame structure analysis of GPS L1C CNAV2 and Galileo F/NAV showed that the design of almanac frame also had an important impact on CED and GST acquisition time. Then we analyzed the almanac frame structure of GPS L1C/A NAV and Galileo I/NAV which were redrawn in Figs. 2.8 and 2.9.

For L1C/A NAV, the almanac parameters were arranged in two sub-frames, which lasted 12 s, we needed to obtain the ephemeris parameters from the next frame if we did not require them, the time was about 2.4 times to CNAV2 and F/NAV, and about 2 times for the coefficient η. For Galileo I/NAV, shown in Fig. 2.9, the front five sub-frames had the similar structure with L1C/A NAV, the middle five were reserved pages, and the last five were for the remain bits of CED and GST as well as two blank pages. The structure of it was not designed compactly for CED and GST, which gone against reducing TTFF. This also explained the reason why it's coefficient $\eta > 1$ and it's $T_{CED+GST}$ was 79 s with data rate of 50 bps, which was the maximal of all message structure.

2.6 The Key Factor to Reduce TTFF

We presented some navigation message structure with low acquisition time of CED and GST as well as low TTFF. Firstly, we considered the factor of η, that was the almanac frame having appropriate proportion of the total frame size: too bigger was not conducive to reduce the CED and GST acquisition time and too small gone against the average speed to obtain the almanac and ionosphere parameters, which were all undesirable for navigating; Secondly, the consistency of the ephemeris parameters during the active interval time should be considered; Finally, the CED and GST parameters should be designed in a independent sub-frame to avoid being separated by the parity bits, that was to be designed more

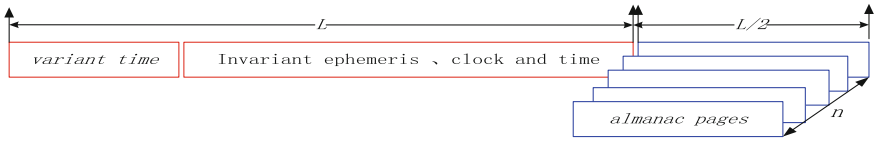


Fig. 2.10 Single structure of message

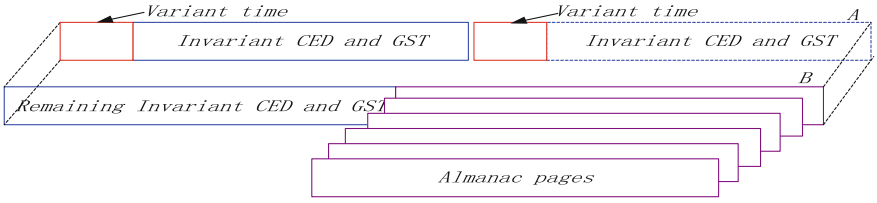


Fig. 2.11 Dual structure of message

compactly, however, it was associated with the channel coding and message checking ways.

The two message structure, *Single* and *Dual* (J. W. Betz) were analyzed as followed. For *Single*, the almanac frame was about half of the size of CED and GST frame shown in Fig. 2.10, where η was about 50 %. The all ephemeris parameters were designed in a independent sub-frame and some GST parameters which changed during the effective interval were arranged in a previous sub-frame, which was similar to CNAV2 structure. The benefits of this design would not only help to reduce the TTFF, but also for the message robustness.

The two pasts A and *B* of *Dual* structure were simultaneously broadcasted, shown in Fig. 2.11, where part *A* did not broadcast almanac pages and just repeated CED and GST information. For lower TTFF, the amount of essential data must be compendious, and if the total frame were 900, 450 bits needed for CED and GST in part *A*, and just 9 s with 50 bps data rate to first obtain positioning information, and then just 6 s with 75 bps. However, the reduction of the amount of data means that the accuracy of satellite orbit model would inevitably be reduced as well as the user positioning accuracy, which might lead to 5 m user range error (URE).

The remaining CED, GST and almanac parameters were broadcasted in part *B*. We should note that the almanac frame size in part *B*, with fewer remaining CED and GST, was longer than *Single* structure, that was to say the corresponding η may larger, however, the CED and GST were repeated twice in part *A* and the corresponding η of it almost equal to zero. So the impact of the longer almanac frame size in part *B* would not be considered.

Under the same conditions of 50 bps data rate and 900 bits frame size, the first time to get CED and GST of *Dual* structure was the shortest, only 9 s, compared to each message frame structure of GPS and Galileo.

Higher almanac rate was benefit for accessing to the constellation information rapidly, which would help reduce the capture time. For *Single* structure, the almanac average data rate was 15 bps where the frame length was set to 270 bits and the average time to get an almanac frame was18 s. For *Dual*, the almanac data rate was 27.8 bps where the frame length was 500 bits and the time was 18 s which was arranged in part *B*. Comparative analysis shown that the *Dual* message structure had great reference value for lower TTFF.

2.7 Conclusion

We emphatically discussed the relationship between message structure design and TTFF. The basic conclusions were following. Firstly, under the conditions that the maximum broadcast interval was 48 s and the GPS L2C CNAV was broadcasted by the way of data block, the acquisition time of CED and GST was about 64 s. Secondly, the CNAV2 only needed 18.02 s that was the minimum time compared to other GPS and Galileo navigation message structure to obtain the CED and GST parameters at the same data rate and total amount, which was closely related to the almanac frame structure and the consistency of ephemeris parameters. Then, the first time to access to the positioning information of *Dual* structure was smaller, only 9 s, as well as a higher almanac average data rate, up to 27.8 bps. Finally, we presented the key factor of navigation message structure designing to reduce the TTFF which was mainly about the following three points: (1) the almanac frame proportion of the total frame size should be appropriate; (2) the ephemeris parameters should be consistent during the effective interval; (3) the message parameters should be more compact.

References

1. Anghileri M (2009) Estimating the time to first fix for GNSS signals theory and simulation results. In: Proceedings of the fourth European workshop on GNSS signals and signal processing
2. Betz JW et al (2006) “L1C signal design options,” ION-NTM-2, Institute of Navigation, Washington
3. Garcia-Pena A, Macabiau C, Escher A-C, Boucheret M-L, Ries L (2009) Comparison between the future GPS L1C and GALILEO E1 OS signals data message performance, In: proceedings of the 2009 international technical meeting of the institute of Navigation, 26–28
4. IS-GPS-800, Navstar GPS Space Segment / Navigation user segment interfaces. USA: Navstar GPS Joint Program Office
5. GALILEO ICD, “GALILEO Open Service—Signal in Space Interface Control Document (OS SIS ICD), Draft 1”, European Space Agency / European GNSS Supervisory Authority, February 2008

Chapter 3
Interlacing Dual-Sideband Enhanced Alternative BOC (IDEAl-BOC) Modulation and Its Application for GNSS

Tao Yan, Jiaolong Wei, Zuping Tang, Bo Qu and Zhihui Zhou

Abstract Alternative binary offset carrier (AltBOC) modulated signal can combine two QPSK signals located at two close frequency bands into a constant envelope signal. It can achieve better code tracking accuracy and anti-multipath performance, thus Galileo E5 signal uses AltBOC (15, 10) modulation. However, the four signal components of AltBOC modulation have to be allocated the same power. To solve this problem, we propose a constant envelope dual-sideband enhanced alternative BOC (DEAl-BOC) modulation technique. In the case of not increasing signal generation complexity, DEAl-BOC technique can adjust the power ratio of four signal component according to different requirements, and only the phase look-up table needs to be updated. When the combined efficiency is decreased slightly, the signal generation complexity can be reduced by half. Based on DEAl-BOC, we present the interlacing dual-sideband enhanced alternative BOC (IDEAl-BOC) modulation. The main advantage is that IDEAl-BOC can improve further the combined efficiency while keeping power ratio constant, this means IDEAl-BOC can reach higher combined efficiency than AltBOC for the case of equal power ratio.

Keywords Constant envelope modulation · Power ratio · AltBOC · IDEAl-BOC · GNSS

3.1 Introduction

Constant envelope alternative binary offset carrier (AltBOC) modulation is the modulation type of Galileo E5 signal [1]. This modulation scheme can multiplex two QPSK signals whose carrier frequencies are close into a constant envelope

T. Yan · J. Wei (✉) · Z. Tang · B. Qu · Z. Zhou
The Department of Electronic and Information Engineering,
Huazhong University of Science and Technology, Wuhan 430074, China
e-mail: jlwei@mail.hust.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_3,

signal to maximize the amplifier efficiency. Thus it reduces implementation complexity obviously [2]. In addition, AltBOC (15, 10) modulation is also a candidate for BeiDou B2 signal [3]. The performance of AltBOC modulation is outstanding. However, it is not flexible enough due to its fixed equal power allocation.

For this problem, some related researches have been done. In literature [4], POCET method is proposed to combine GNSS signals at different carrier frequencies. This method is achieved by phase rotation. However, the operation to optimize the phase at each sample time is complex. Moreover, this analytical expression is difficult to obtain. Recently, time multiplex technology is integrated to AltBOC modulation. TD-AltBOC [5] and TMOC-QPSK [6] modulation are both this kind of technologies, which could simplify the generating and receiving process, and achieve similar power spectrum density and function with Galileo AltBOC modulation. For the two time multiplexed techniques, the power of four signal components has to be equal. A nice work is made by Yao and Lu [7, 8]. He proposed a new method based on subcarrier reconstruction to achieve dual-frequency constant envelope multiplex. This technique is referred to asymmetric constant envelope double-sideband (ACED) multiplexing [9]. In this technique, power allocation ratio can be adjusted. Nevertheless the frequency of control clock is different for different power allocation ratio, it is more difficult to achieve ACED than AltBOC. In addition, a generalized constant-envelope DualQPSK modulation is proposed in [10]. It can be seen as an extended AltBOC signal, which has unbalanced power on the upper and the lower sidebands. A limitation of this method is that the two signal components at the same sideband have equal power.

In this paper, we firstly propose a constant envelope dual-sideband enhanced alternative BOC (DEAl-BOC) modulation technique in Sect. 3.2. The modulation signal has all the advantages of AltBOC signal. At the same time, DEAl-BOC modulation enhances the flexibility of power allocation ratio. Different from ACED, DEAl-BOC technique can adjust the power ratio of four signal components; meanwhile signal generation complexity is not increased. In some cases, the signal generation complexity can be reduced by half. In Sect. 3.3, based on DEAl-BOC, we introduce the interlacing technique, and then the interlacing dual-sideband enhanced alternative BOC (IDEAl-BOC) modulation is presented. In Sect. 3.4, simulation results are shown and compared with ACED. Finally some conclusions are concluded.

3.2 DEAl-BOC Modulation

DEAl-BOC can be seen as an enhanced AltBOC modulation, whose power allocation ratio can be adjusted. To achieve constant envelope AltBOC modulation with flexible power assignment, we firstly assume that a general baseband expression of DEAl-BOC in a sub-period can be written as (3.1).

$$\begin{aligned} s = {} & (s_1 + jp_2 s_2)(a - jb) + (p_3 s_3 + jp_4 s_4)(a + jb) \\ & + (p_5 s_2 s_3 s_4 + jp_6 s_1 s_3 s_4)(c - jd) + (p_7 s_1 s_2 s_4 + jp_8 s_1 s_2 s_3)(c + jd), \end{aligned} \tag{3.1}$$

where s_1, s_2, s_3 and s_4 are desired binary signal components. p_5, p_6, p_7, p_8, a, b, c and d are unknown parameters to be determined. Assume pseudorandom codes of these four signals are ideal. s_1 and s_2 are in-phase and quadrature-phase components at lower sideband. s_3 and s_4 are in-phase and quadrature-phase components at upper sideband. The values of p_2, p_3, and p_4 are determined according to power allocation ratio.

We can derive signal's envelope in a sub-period from (3.1). Variable A is used to denote envelope. To ensure variable A is constant in a sub-period for 16 different possible signal states, the following equations are established.

$$\begin{cases} (p_4 p_7 - p_3 p_8)(ad - bc) = 0 \\ p_3\left(a^2 - b^2\right) + p_5 p_7\left(c^2 - d^2\right) + (p_2 p_8 + p_4 p_6)(ac - bd) = 0 \\ (bc + ad)(p_3 p_6 + p_2 p_7) - 2abp_4 - 2cdp_5 p_8 = 0 \\ (ac + bd)(p_5 + p_3 p_7 + p_2 p_6 + p_4 p_8) = 0 \\ (ad - bc)(p_6 - p_2 p_5) = 0 \\ p_2 p_4\left(a^2 - b^2\right) + p_6 p_8\left(c^2 - d^2\right) + (p_7 + p_3 p_5)(ac - bd) = 0 \\ 2abp_2 p_3 + 2cdp_6 p_7 - (bc + ad)(p_8 + p_4 p_5) = 0 \end{cases} \tag{3.2}$$

Then envelope A can be calculated by

$$A^2 = \left(a^2 + b^2\right)\left(1 + p_2^2 + p_3^2 + p_4^2\right) + \left(c^2 + d^2\right)\left(p_5^2 + p_6^2 + p_7^2 + p_8^2\right) \tag{3.3}$$

It is easy to verify that if $\{p_5, p_6, p_7, p_8, a, b, c, d\}$ is a solution of (3.2), (3.4) and (3.5) are also solutions of (3.2). These interesting properties are helpful to generate subcarrier waveforms.

$$\{p_5, p_6, p_7, p_8, ka, kb, kc, kd\}; \tag{3.4}$$

$$\begin{aligned} & \{p_5, p_6, p_7, p_8, b, a, d, c\}; \{p_5, p_6, p_7, p_8, -a, b, -c, d\}; \\ & \{p_5, p_6, p_7, p_8, a, -b, c, -d\}. \end{aligned} \tag{3.5}$$

Similar to AltBOC modulation, each subcarrier period is sub-divided in n equal sub-periods. For example, a period of subcarrier can be divided in 8 equal sub-periods, it can be expressed as:

$$\begin{aligned} sc_{\cos,S}(t) &= \sum_{i=1}^{8} SC^i_{\cos,S}\, p_{T_s/8}(t), sc_{\sin,S}(t) = \sum_{i=1}^{8} SC^i_{\sin,S}\, p_{T_s/8}(t), \\ sc_{\cos,P}(t) &= \sum_{i=1}^{8} SC^i_{\cos,P}\, p_{T_s/8}(t), sc_{\sin,P}(t) = \sum_{i=1}^{8} SC^i_{\sin,P}\, p_{T_s/8}(t), \end{aligned} \tag{3.6}$$

where $p_{Ts/8}(t)$ is the rectangular function which is unity from $0 < t < T_s/8$ and zero elsewhere. T_s is the period of periodic side-band sub-carrier functions. And

$$\begin{aligned} SC_{\cos,S} &= \{a, b, -b, -a, -a, -b, b, a\}; SC_{\sin,S} = \{b, a, a, b, -b, -a, -a, -b\}; \\ SC_{\cos,P} &= \{c, d, -d, -c, -c, -d, d, c\}; SC_{\sin,P} = \{d, c, c, d, -d, -c, -c, -d\}, \end{aligned} \tag{3.7}$$

Once generation scheme of subcarriers is determined, our DEAl-BOC modulation signal for one certain power allocation ratio is expressed as:

$$\begin{aligned} s(t) = &\frac{1}{A}\left(s_1(t) + jp_2 s_2(t)\right)\left[sc_{\cos,S}(t) - jsc_{\sin,S}(t)\right] \\ &+ \frac{1}{A}\left(p_3 s_3(t) + jp_4 s_4(t)\right)\left[sc_{\cos,S}(t) + jsc_{\sin,S}(t)\right] \\ &+ \frac{1}{A}\begin{pmatrix} p_5 s_2(t)s_3(t)s_4(t) + \\ jp_6 s_1(t)s_3(t)s_4(t) \end{pmatrix}\left[sc_{\cos,P}(t) - jsc_{\sin,P}(t)\right] \\ &+ \frac{1}{A}\begin{pmatrix} p_7 s_1(t)s_2(t)s_4(t) + \\ jp_8 s_1(t)s_2(t)s_3(t) \end{pmatrix}\left[sc_{\cos,P}(t) + jsc_{\sin,P}(t)\right] \end{aligned} \tag{3.8}$$

The combination efficiency of DEAl-BOC for a certain power assignment is:

$$\eta = \frac{(a^2 + b^2)\left(1 + p_2^2 + p_3^2 + p_4^2\right)}{A^2} \times 100\,\% \tag{1.9}$$

We expect that combination efficiency is as large as possible while keeping envelope constant. Next, we discuss two special cases.

Case One The power ratio between data component and pilot component is 1:1 in case one, but the total power ratio between lower sideband and upper sideband is adjustable. Namely $p_2 = p_1 = 1$, $p_3 = p_4 = \gamma$. Then the total power ratio is $1{:}1{:}\gamma^2{:}\gamma^2$. We get several optimal numerical solutions for different total power ratios. Then we can find that the optimal solution in this case has to satisfy the following conditions:

$$ac + bd = 0; \frac{a}{b} = \sqrt{2} + 1; p_5 = p_6; p_7 = p_8 \tag{3.10}$$

An analytical solution close to optimal solution is given out in (3.11). Figure 3.1a shows the combination efficiency of optimal solution and analytical solution versus power ratio γ^2. It illustrates that the difference of power between lower sideband and upper sideband is greater; the maximal combination efficiency is higher. In addition, we can find that the difference of combining efficiency between analytical solution and numerical optimal solution is less than 1 %. In this case, the complex subcarrier is the same with AltBOC modulation. Thus the generation complexity is also the same with AltBOC.

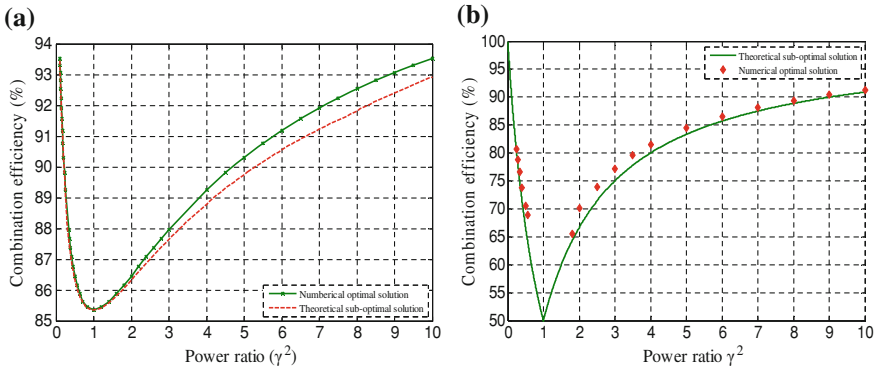


Fig. 3.1 Combination efficiency versus power ratio γ^2. **a** Case one. **b** Case two

$$\begin{cases} \left\{ \begin{array}{l} a=\sqrt{2}+1, b=d=1, c=1-\sqrt{2}, \\ p_5=p_6=\dfrac{1-\gamma^2}{\gamma^2 c}, p_7=p_8=\dfrac{\gamma}{(1-2\gamma^2)c} \end{array} \right\} \text{for} \quad \gamma^2 > 7.7582 \\ \left\{ \begin{array}{l} a=\sqrt{2}+1, b=d=1, c=1-\sqrt{2}, \\ p_5=p_6=p_7=p_8, p_5=\dfrac{\gamma+1-\sqrt{\gamma^2+6\gamma+1}}{2c} \end{array} \right\} \text{for} \quad 0.2955 \le \gamma^2 \le 7.7582 \\ \left\{ \begin{array}{l} a=\sqrt{2}+1, b=d=1, c=1-\sqrt{2}, \\ p_5=p_6=\dfrac{-\gamma^2}{c}, p_7=p_8=\dfrac{(\gamma^2-1)\gamma}{(\gamma^2+1)c} \end{array} \right\} \text{for} \quad 0<\gamma^2<0.2955 \end{cases} \tag{3.11}$$

Case Two The power at lower sideband and upper sideband is the same in case two, but the power ratio between data component and pilot component is adjustable. Namely $p_3 = p_1 = 1$, $p_2 = p_4 = \gamma$, and $\gamma \neq 1$. The total power ratio is $1{:}\gamma^2{:}1{:}\gamma^2$. Similar to the earlier, we first obtain several optimal numerical solutions for different γ. Let a be 1. We can find that the optimal solution has to satisfy the following conditions:

$$\begin{cases} b=a=1, c=d \\ c(p_6+\gamma p_7)-\gamma-c^2 p_5 p_8=0 \\ (p_5+p_7+\gamma p_6+\gamma p_8)=0 \\ \gamma+c^2 p_6 p_7-c(p_8+\gamma p_5)=0 \end{cases} \tag{3.12}$$

A solution of (3.2) in this case is:

$$\begin{cases} a=b=c=d=1, p_5=p_7=1, p_6=p_8=-\dfrac{1}{\gamma} & \text{for } \gamma>1 \\ a=b=c=d=1, p_5=p_7=-\gamma^2, p_6=p_8=\gamma & \text{for } 0<\gamma<1 \end{cases} \tag{3.13}$$

Figure 3.1b compares the combination efficiencies between theoretical sub-optimal solution and numerical optimal solution. The greater are power differences between data component and pilot component, the closer are combination efficiencies between theoretical sub-optimal solution and numerical optimal solution.

We can see from (3.7) and (3.13) that the subcarrier period is in fact subdivided in 4 equal sub-periods for case two. The subcarriers are complex binary offset carrier. Thus the signal generation complexity can be reduced by half.

3.3 IDEAl-BOC Modulation

Above two special case of DEAl-BOC have some different advantages. For case one, the main advantage is that the combining efficiency is higher than AltBOC when the total power ratio between lower sideband and upper sideband is not equal to 1. Unfortunately, we more likely allocate equal power to lower sideband and upper sideband. For case two, the required frequency of control clock is half frequency of AltBOC, due to its binary subcarrier. The signal generation complexity can be therefore reduced. However, the combining efficiency is lower than 80 % when the power ratio of pilot and data components at one sideband is less than 4. To eliminate these restrictions, we introduce the interlacing technique, and then propose the interlacing dual-sideband enhanced alternative BOC (IDEAl-BOC) modulation.

IDEAl-BOC modulation signal is generated by interlacing multiple DEAl-BOC signals with different power allocation ratios. This means that the power ratio of four signal components is changing with time. To simplify signal generation, we select two DEAl-BOC signals to interlace. During a code period, the ratio of two DEAl-BOC signals is designed to meet desired power allocation scheme. Namely the IDEAl-BOC can be denoted by

$$\text{IDEAl-BOC}\,(m,n,r) = \alpha\text{DEAl-BOC}\,(m,n,r_1) + (1-\alpha)\,\text{DEAl-BOC}\,(m,n,r_2), \tag{3.14}$$

where $m \times 1.023$ MHz is the subcarrier frequency, $n \times 1.023$ Mcps is the code rate. r_1 and r_2 are the power allocation ratio of two DEAl-BOC signals, respectively. α is the proportion of the first DEAl-BOC signal.

Two DEAl-BOC signals can be choose as DEAl-BOC (m, n, 1:1:γ^2:γ^2) and DEAl-BOC (m, n, 1:1:1/γ^2:1/γ^2), and $\gamma > 1$. If the two signals occur with the same probability in IDEAl-BOC signal, the four signal components have the same power. At the same time, the IDEAl-BOC signal can maintain the high combining

efficiency of DEAl-BOC (m, n, 1:1:γ^2:γ^2). This means IDEAl-BOC can achieve higher combined efficiency than AltBOC for the case of equal power ratio. Two DEAl-BOC signals can also be choose as DEAl-BOC (m, n, 1:γ^2:1:γ^2) and DEAl-BOC (m, n, 1:1/γ^2:1:1/γ^2), and $\gamma > 4$. If they have the same proportion in IDEAl-BOC, the four signal components of IDEAl-BOC have the same power. At the same time, the signal generation complexity can be reduced by half. We can also achieve an IDEAl-BOC signal with power ratio 1:3:1:3, namely

$$\begin{aligned}\text{IDEAl-BOC}(m,n,1:3:1:3) = &\frac{5}{6}\text{DEAl-BOC}(m,n,1:7:1:7)\\ &+\frac{1}{6}\text{DEAl-BOC}\left(m,n,1:\frac{1}{7}:1:\frac{1}{7}\right),\end{aligned} \tag{3.15}$$

This IDEAl-BOC signal can be generated by phase lookup table. The phase lookup table includes two DEAl-BOC signals' phase lookup table. We list the phase lookup table of IDEAl-BOC (15, 10, 1:3:1:3) in Table 3.1.

Figure 3.2 shows the normalized PSD for DEAl-BOC and IDEAl-BOC signals. The DEAl-BOC (15, 10, 1:1:1:1) is the Galileo AltBOC (15, 10) modulation. The green dashed shows that more power is allocated to main-lobes for IDEAl-BOC (15, 10, 1:1:1:1) than AltBOC. This is because that combining efficiency is more than 91 %, which is higher than AltBOC signal's 85.36 %.

3.4 Simulation Results

In this section, our proposed modulation signal is analyzed by simulation and compared with ACED. With no loss of generality, we make the following assumption: Code rate is taken as 10.23 Mcps, subcarrier frequency is 15 × 1.023 MHz. Four PRN codes exploit Galileo E5 signal's primary codes, no secondary codes are considered, so code length is 10,230, and code period is 1 ms.

Figure 3.3 is the signal's constellation diagram when four signals' power ratio is 1:1:3:3. Figure 3.3a is the case of DEAl-BOC, and Fig. 3.3b is ACED. From the view of time domain, in order to generate this signal, the required control-clock frequency for DEAl-BOC is 8 times of subcarrier frequency, but there are 32 possible phases in Fig. 3.3a, which means 32 phase values have to be stored in phase lookup table. Figure 3.3b is the corresponding constellation diagram of ACED. There are only 12 possible phases.

Figure 3.4 presents the constellation diagram when signals' power ratio is 1:3:1:3. The IDEAl-BOC signal is generated according to Table 3.1. For this signal, the required frequency of control-clock signal is only 4 times of subcarrier frequency. From Fig. 3.4a, we find that there are 12 possible phases, but the four phases located at I axis and Q axis have bigger occurrence probability than other 8 phases. Figure 3.4b is the corresponding constellation diagram of ACED. We can

Table 3.1 Phase lookup tables of DEAl-BOC signals whose power ratio are 1:7:1:7 and 1:1/7:1:1/7

s_1		−1	−1	−1	−1	−1	−1	−1	−1
s_3		−1	−1	−1	−1	1	1	1	1
s_2		−1	−1	1	1	−1	−1	1	1
s_4		−1	1	−1	1	−1	1	−1	1
i_{Ts}	$t' = t \bmod T_s$	Phase angle in degree (°)							
0	$[0,T_s/4)$	−131.4/−138.6	180.0/180.0	0.0/180.0	131.4/138.6	−90.0/90.0	138.6/131.4	41.4/48.6	90.0/90.0
1	$[T_s/4,2T_s/4)$	48.6/41.4	180.0/180.0	0.0/0.0	−48.6/−41.4	90.0/90.0	138.6/131.4	41.4/48.6	−90.0/90.0
2	$[2T_s/4,3T_s/4)$	48.6/41.4	0.0/0.0	180.0/0.0	−48.6/−41.4	90.0/−90.0	−41.4/−48.6	−138.6/−131.4	−90.0/−90.0
3	$[3T_s/4,T_s)$	−131.3/−138.6	0.0/0.0	180.0/180.0	131.4/138.6	−90.0/−90.0	−41.4/−48.6	−138.6/−131.4	90.0/−90.0
s_1		1	1	1	1	1	1	1	1
s_3		−1	−1	−1	−1	1	1	1	1
s_2		−1	−1	1	1	−1	−1	1	1
s_4		−1	1	−1	1	−1	1	−1	1
i_{Ts}	$t' = t \bmod T_s$	Phase angle in degree (°)							
0	$[0,T_s/4)$	−90.0/−90.0	−138.6/−131.4	−41.4/−48.6	90.0/-90.0	−48.6/−41.4	180.0/0.0	0.0/0.0	48.6/41/4
1	$[T_s/4,2T_s/4)$	90.0/−90.0	−138.6/−131.4	−41.4/−48.6	−90.0/−90.0	131.4/138.6	180.0/180.0	0.0/180.0	−131.4/−138.6
2	$[2T_s/4,3T_s/4)$	90.0/90.0	41.4/48.6	138.6/131.4	−90.0/90.0	131.4/138.6	0.0/180.0	180.0/180.0	−131.4/−138.6
3	$[3T_s/4,T_s)$	−90.0/90.0	41.4/48.6	138.6/131.4	90.0/90.0	−48.6/−41.4	0.0/0.0	180.0/0.0	48.6/41.4

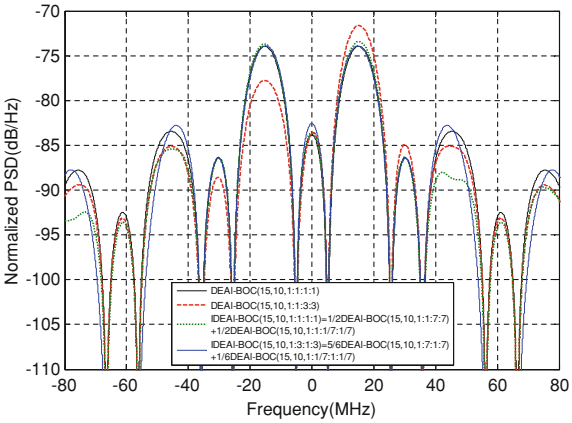


Fig. 3.2 Normalized PSD for different power assignment schemes

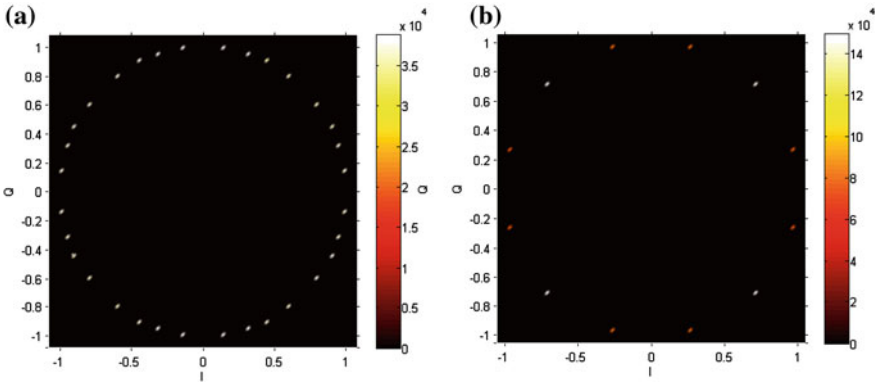


Fig. 3.3 The constellation diagram when four signals' power ratio is 1:1:3:3. **a** DEAI-BOC. **b** ACED

find that Fig. 3.4b is similar with Fig. 3.4a in this case, but ACED requires higher frequency of control-clock, namely 12 times of subcarrier frequency.

In order to analyze fact signal components' power ratio, we obtain the cross correlation function when receiving different signal components. The generated signal is baseband signal, received signal's bandwidth is 71.61 MHz, and filter is ideal. Data length of received signal is 1 ms, namely one primary code period. To obtain desired cross correlation function, we correlate received signal with corresponding signal component. For example, to obtain the cross correlation function between received signal with pilot component at lower sideband, first multiply received signal with exp ($j2\pi f_s t$), which moves the centre frequency of lower sideband signal to zero frequency, then correlate the signal with pilot component at lower sideband.

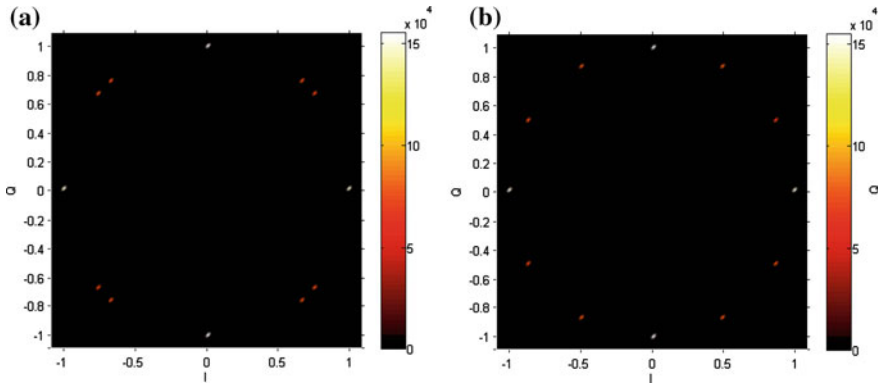


Fig. 3.4 The constellation diagram when four signals' power ratio is 1:3:1:3. **a** IDEAl-BOC. **b** ACED

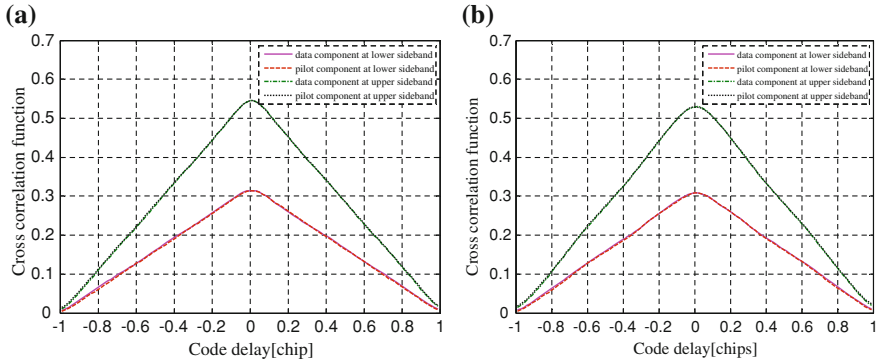


Fig. 3.5 Cross correlation function for four signal component when four signals' power ratio is 1:1:3:3. **a** DEAl-BOC. **b** ACED

Above two signals are analyzed. We correlate received signal with four signal components in simulation, namely data and pilot component at lower sideband, and data and pilot component at upper sideband. Thus four cross correlation curves are obtained for each signal. Results are presented in Figs. 3.5 and 3.6, respectively. In Fig. 3.5a, the maximum cross correlation values for four signal components are 0.315, 0.315, 0.546 and 0.546, so fact power ratio for the four signal components is 1.000:1.000:2.998:2.998. In Fig. 3.5b, the maximum cross correlation values for four signal components are 0.3086, 0.3079, 0.5294 and 0.5293, so fact power ratio for the four signal components is 1.000:0.995:2.9427:2.9415. For this signal, designed power ratio is 1:1:3:3. In this case, DEAl-BOC has obviously more received power than ACED.

In Fig. 3.6a, the maximum cross correlation values for four signal components are 0.292, 0.513, 0.293 and 0.513, so fact power ratio for the four signal components

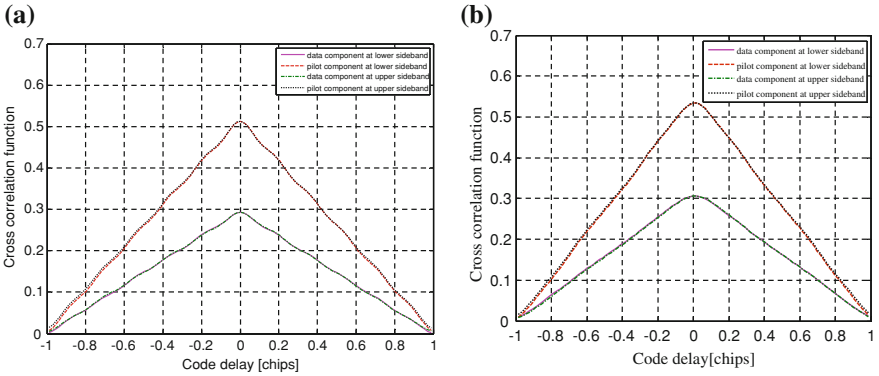


Fig. 3.6 Cross correlation function for four signal component when four signals' power ratio is 1:3:1:3. **a** IDEAl-BOC. **b** ACED

is 1.000:3.074:1.0015: 3.074. In Fig. 3.6b, the maximum cross correlation values for four signal components are 0.3065, 0.5339, 0.3066, 0.5340, so fact power ratio for the four signal components is 1.000:3.0351:1.0011:3.0364. For this signal, designed power ratio is 1:3:1:3. In this case, ACED has slightly more received power than proposed general AltBOC.

From above, we know that fact power ratio for each signal components is consistent with theoretical analysis, which demonstrates that the power ratio is adjustable in our method.

3.5 Conclusion

In this paper, we propose the DEAl-BOC and IDEAl-BOC modulated signals. DEAl-BOC technique is an enhanced AltBOC modulation, whose four signal components' power allocation ratio can be adjusted. At the same time, signal generation complexity is not increased. IDEAl-BOC modulation signal is generated by interlacing multiple DEAl-BOC signals with different power allocation ratios. The main advantage is that IDEAl-BOC can improve further the combined efficiency while keeping power allocation ratio constant, this means IDEAl-BOC can achieve higher combined efficiency than AltBOC for the case of equal power ratio. When four signals' power ratio is 1:3:1:3, the signal generation complexity can be reduced compared to ACED. Thus, for the GNSS signal design, IDEAl-BOC would be a better choice.

References

1. European Union, European GNSS (Galileo) (2010) Open service signal in space interface control document, OS SIS ICD, Issue 1, pp 4–7
2. Lestarquit L, Artaud G, Issler J-L (2008). AltBOC for dummies or everything you always wanted to know about AltBOC. In: Proceedings of the 21st international technical meeting of the satellite division of the institute of navigation, pp 961–970
3. ZuPing T, HongWei Z, XiuLin H YiHang R, YuQi L, YanLin Z (2010) Research on performance evaluation of compass signal. Scientia Sinica Phys Mech Astron 40(5):592–602
4. Dafesh PA, Cahn CR (2011) Application of POCET method to combine gnss signals at different carrier frequencies. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011), pp 1201–1206
5. ZuPing T, HongWei Z, JiaoLong Wei (2011) TD-AltBOC: a new compass B2 modulation. Sci China-Phys Mech Astron 54(6):P1014–P1021
6. Shivaramaiah NC, Dempster AG (2013) Time-multiplexed offset-carrier QPSK for GNSS. IEEE Trans Aerosp Electron Syst 49(2):1119–1138
7. Yao Z, Lu M (2012) Dual-frequency constant envelope multiplex with non-equal power allocation for GNSS. Electron Lett 48(25):1624–1625
8. Yao Z, Lu M (2013) Constant envelope combination for components on different carrier frequencies with unequal power allocation. In: Proceedings of the 2013 international technical meeting of the institute of navigation, San Diego, pp 629–637
9. Yao Z, Lu M (2013) ACED multiplexing and its application on BeiDou B2 band. In: Proceedings of China satellite navigation conference (CSNC 2013), Lecture notes in electrical engineering, vol 244, pp 25–34
10. Zhang Kai (2013) Generalized constant-envelope dual QPSK and altBOC modulations for modern GNSS signals. Electron Lett 49(21):1335–1337

Chapter 4
CBFH: Coherent Binary Frequency-Hopping Multiplexing for BeiDou B2 Signal

Zhihui Zhou, Jiaolong Wei, Zuping Tang, Tao Yan and Xuan Xia

Abstract AltBOC (15, 10) modulation can carry different information on its lower and upper lobes. The two lobes can be received and processed separately as BPSK (10) signal, or coherently processed to achieve better performance. However, AltBOC modulation has lower efficiency and great satellite implementation and receiver processing complexity. TD-AltBOC was proposed to reduce the complexity, and ACED (Asymmetric Constant Envelope Double-sideband) was proposed to improve flexibility of power allocation among the four signals as a cost of complexity increasing further. Nevertheless the correlation power efficiency of these methods is 81.06 %. A novel multiplexing scheme named as CBFH (Coherent Binary Frequency-Hopping) is presented in this paper. The signal generation principle and receiving method are given, and its power spectrum, power efficiency, processing flexibility and complexity are compared with AltBOC, TD-AltBOC and ACED. The Results show that the spectrum of CBFH is similar to AltBOC, TD-AltBOC and ACED. CBFH has the same performance as TD-AltBOC when two-level sub-carrier was adopted; CBFH has the best performance when its complexity as the same as AltBOC or ACED, or CBFH has the lowest complexity when its performance as the same as AltBOC or ACED. CBFH will be a good solution for COMPASS B2 navigation signal.

Keywords GNSS · BeiDou · CBFH · Multiplexing

Z. Zhou · J. Wei (✉) · Z. Tang · T. Yan · X. Xia
The Department of Electronics and Information Engineering,
Huazhong University of Science and Technology, Wuhan, China
e-mail: jlwei@mail.hust.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_4,

4.1 Introduction

In the Galileo system, an AltBOC (15, 10) modulation method is adopted in its E5a and E5b band [1], AltBOC is a new method proposed by Galileo to achieve constant-envelope modulation for dual-frequency signals, it can carry different information on its lower and upper lobes. The two lobes can be received and processed separately as BPSK (10) signal, or coherently processed to achieve better performance [2]. So it also becomes a modulation baseline for B2 band of BeiDou system. However, in order to achieve constant-envelope modulation for dual-frequency signals, AltBOC modulation must increase its subcarrier conversion rate to 8 times of its subcarrier frequency, and the number of corresponding level is increase to 4. Obviously, such a subcarrier will increase the complexity of signal generation and receiving. In [3], a TD-AltBOC was proposed to simplify achievement of signal generation and receiving, in TD-AltBOC only binary level subcarrier is needed, and its subcarrier conversion rate is 4 times of subcarrier frequency. And constant-envelope modulation methods for dual-frequency signals with adjustable power ratio are researched by Zheng Yao and Kai Zhang, in [4], a dual-frequency modulation method was proposed which can be used to the situation of different power ratio between up-band and lower-band, at the same time, it can increase the power efficiency; a more general dual-frequency modulation method ACED was proposed by Zheng Yao [5], it can achieve constant-envelope modulation for dual-frequency signals with arbitrary power ratio; However, take 1:3:1:3 as a example, its subcarrier conversion rate is 12 times of subcarrier frequency which is more complex than AltBOC, it is very difficult for ACED signal generation and receiving. Moreover a common shortcoming of AltBOC, TD-AltBOC and ACED is that their power efficiency are all 81.06 %.

In this paper, a CBFH (coherent binary frequency-hopping) multiplexing technology is proposed, a frequency-hopping technology is used to up-band and lower-band, it can be used to the situation of that up-band and lower-band has equal power, on each band, the power ratio between data-component and pilot-component can adjust arbitrarily. Moreover, it has flexible baseband waveform, when its subcarrier conversion rate is 4 times of subcarrier frequency, it has the same performance and complexity with TD-AltBOC; and the higher subcarrier conversion rate of CBFH is, the higher the power efficiency of CBFH is. The baseband waveform and power spectral density subcarrier with 4 or 8 times of subcarrier frequency are presented, and the power spectral, power efficiency and complexity of CBFH is compared with AltBOC, TD-AltBOC and ACED, the results show that CBFH has the same power spectral shape with AltBOC, TD-AltBOC and ACED but more higher power efficiency and lower complexity than AltBOC and ACED. CBFH is a feasible candidate modulation scheme of BeiDou B2 band.

This paper is organized as follows. In the Sect. 4.2, we present the principle of CBFH. Then in the Sect. 4.3, we introduced the power spectral expression of CBFH. The performance of our method is presented in the Sect. 4.4. Finally Sect. 4.5 presents the summary for this paper.

4.2 Principle of CBFH

4.2.1 Signal Model

There are two side-band B2a and B2b are contained on BeiDou B2 band, the center-frequency of B2a is 1176.45 MHz which is the same with GPS L5C, and the center-frequency of B2b is 1207.14 MHz which is the same with BeiDou local system B2 signal. Each band consists of data-component and pilot-component, we named it B2a_D, B2a_P, B2b_D and B2b_P, and the powers of these four signals are equal.

In order to transmit these four signals through the same HPA and antenna, we can set the center-frequency of B2 to 1191.795 MHz, and shift the lower-band signals and up-band signals to their own center-frequency by multiplied the signals by a linear complex subcarrier, the expression of linear complex subcarrier is:

$$SC_U(t) = e^{j2\pi f_{sc}t} = \cos(2\pi f_{sc}t) + j\sin(2\pi f_{sc}t)$$
$$SC_L(t) = e^{-j2\pi f_{sc}t} = \cos(2\pi f_{sc}t) - j\sin(2\pi f_{sc}t)$$

Here, $SC_L(t)$, $SC_U(t)$ are the linear complex subcarrier of lower-band and up-band respectively, f_{sc} is the frequency bias, so the baseband signals of lower-band and up-band are:

$$s_L(t) = (a_D d_L(t) c_{LD}(t) + j a_P c_{LP}(t)) \cdot SC_L(t)$$
$$s_U(t) = (b_D d_U(t) c_{UD}(t) + j b_P c_{UP}(t)) \cdot SC_U(t)$$

Here, $d_L(t)$, $d_U(t)$ are the data waveform of lower-band and up-band respectively, $c_{LD}(t)$, $c_{LP}(t)$ are the PRN code waveform of data-component and pilot-component of lower-band, $c_{UD}(t)$, $c_{UP}(t)$ are the PRN code waveform of data-component and pilot-component of up-band, a_D, a_P are the amplitude of data-component and pilot-component of lower-band, b_D, b_P are the amplitude of data-component and pilot-component of up-band.

In order to achieve constant-envelope of the composite signal, a Frequency-Hopping alike technology is presented, it is that select the lower-band or up-band to transmit according to the sign of a pseudorandom sequence, and we named it coherent binary frequency-hopping (CBFH) method. In this method, the only constraint is that the power of four signals should satisfy the equation $a_D^2 + a_P^2 = b_D^2 + b_P^2$, in other word, the power of lower-band and up-band must be equal; the power allocation between data-component and pilot-component in each band can be adjusted arbitrary. So, we can achieve constant-envelope transmitting of the composite signal without introducing inter-modulation products.

In fact, we can take a square wave sequence to replace the pseudorandom sequence to simplify the CBFH signal generation and receiving, which is equivalent to time division multiplexing, thus lower-band is transmitted at odd time slot and up-band is transmitted at even time slot, each of lower-band and up-band is a QPSK modulation signals, Fig. 4.1 depicts the timing relationship between lower-band and up-band.

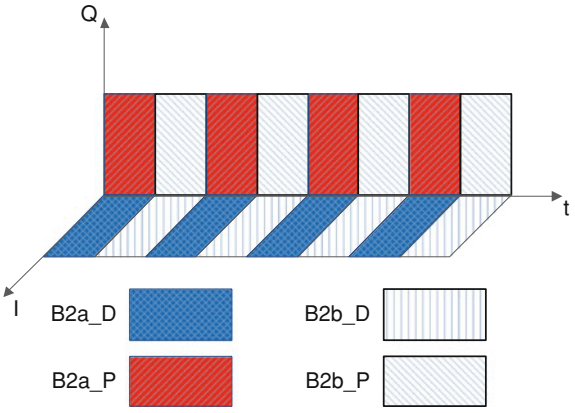


Fig. 4.1 Time slots allocation of CBFH signal components

4.2.2 Two Level and Four Level Subcarrier

The foregoing linear complex subcarriers continuous amplitude, in general the code shape of navigation is square wave or staircase waveform, so we need to sample the continuous complex subcarriers, and the sample rate must be not less than $4f_{sc}$, the sample rates of AltBOC, TD-AltBOC and ACED are $f_{sc\_AltBOC} = 4f_{sc} = 120 \times 1.023MHz$, $f_{sc\_TD-AltBOC} = 8f_{sc} = 60 \times 1.023MHz$ and $f_{sc\_ACED} = 12f_{sc} = 180 \times 1.023MHz$ respectively. And two kinds of sample rate will be considered in CBFH, they are $f_{sc\_CBFH\_4} = 60 \times 1.023MHz$ and $f_{sc\_CBFH\_8} = 120 \times 1.023MHz$, their subcarrier can present as follow:

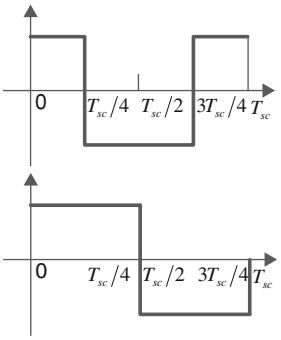
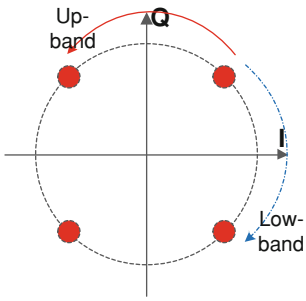
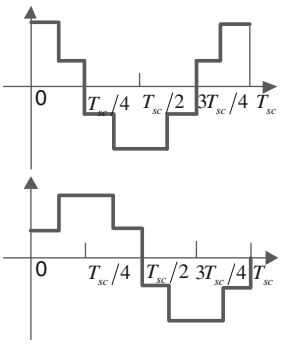
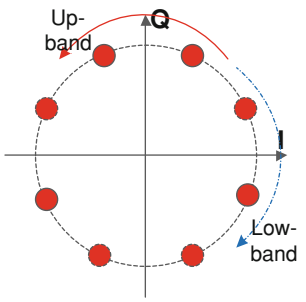
$$SC_{k,U}(t) = SC_{k,\cos} + jSC_{k,\sin}$$
$$SC_{k,L}(t) = SC_{4k,\cos} - jSC_{k,\sin}$$

Here $SC_{k,\cos}$ and $SC_{k,\sin}$ are subcarrier with sample rate of k times of subcarrier frequency, in this paper, k = 4 or 8. Table 4.1 presents the waveform and constellation of these two subcarriers, from Table 4.1 we know that composite signal is a 4PSK signals with sample rate of 4 times of subcarrier frequency or a 8PSK signals with sample rate of 8 times of subcarrier frequency.

4.2.3 Power Spectral Density of Complex Subcarrier

Figures 4.2 and 4.3 depict the power spectrums of complex subcarrier with 4 and 8 times of subcarrier frequency respectively. The red "△" symbol presents the PSD of up-band subcarrier and the blue "▲" presents the PSD of lower-band subcarrier, from these two figures we can obtain that when sample rate of subcarrier is $4f_{sc}$, the power ratio between the power at $\pm f_{sc}$ and the total power is 81.06 %; and when sample rate of subcarrier is $8f_{sc}$, this ratio is up to 94.96 % which is 13.9 %

Table 4.1 Waveform of subcarriers and constellation

Sample rate of complex subcarrier	Waveform	Constellation
$4f_{sc}$		
$8f_{sc}$		

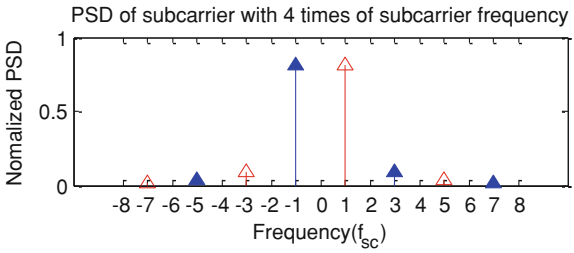


Fig. 4.2 PSD of complex subcarriers with $4f_{sc}$ sample rate

more than $4f_{sc}$ sample rate, and moreover the harmonic component which most close to center-frequency is at $\pm 7f_{sc}$. In addition, as the increasing of sample rate of subcarrier, the efficiency of subcarrier will higher and higher, but when sample rate of subcarrier exceed to $8f_{sc}$, the increase of efficiency is slower and slower, the $8f_{sc}$ of subcarrier is realizable, so from the point of power efficiency, $8f_{sc}$ is a suitable option of subcarrier sample rate.

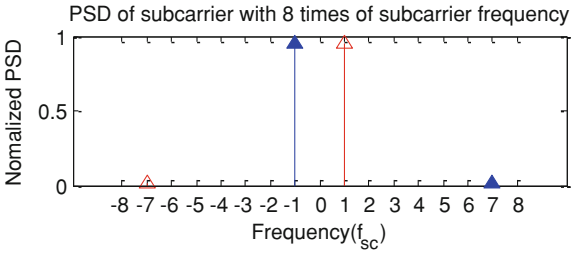


Fig. 4.3 PSD of complex subcarriers with $8f_{sc}$ sample rate

4.3 Power Spectral Density of CBFH Signal

There are two side-band B2a and B2b are contained on BeiDou B2 band and their center-frequency is 1176.45 and 1207.14 MHz respectively, Each band consist of data-component and pilot-component, we named it B2a_D, B2a_P, B2b_D and B2b_P, and the power ratio of four signals is 1:1:1:1, after tedious mathematical derivation one can obtain the power spectral density (PSD) expression of CBFH signals with $4f_{sc}$ sampled subcarrier and $8f_{sc}$ sampled subcarrier.

$$G_{CBFH\_4}(f) = \frac{1}{2\pi^2 f^2 T_c} \cdot \frac{\cos^2(\pi f T_c)[1-\cos(\varphi)]}{\cos^2(\varphi)}$$

$$G_{CBFH\_8}(f) = \frac{\cos^2(\pi f T_c)\left[\begin{array}{l}\sqrt{2}\cos^2(\varphi) - 2\sqrt{2}\cos(\varphi)\cos\left(\frac{\varphi}{2}\right) \\ -\sqrt{2}\cos(\varphi) + 4\left(\sqrt{2}-1\right)\cos\left(\frac{\varphi}{2}\right) + 2\left(2-\sqrt{2}\right)\end{array}\right]}{2\pi^2 f^2 T_c \cos^2(\varphi)}$$

Here $\varphi = \pi f / 2f_{sc}$, f_{sc} is subcarrier frequency, T_c is the width of PRN code. Figure 4.4 shows the PSD of CBFH signals with $4f_{sc}$ sampled subcarrier and $8f_{sc}$ sampled subcarrier and AltBOC signals, From Fig. 4.4 we can get that CBFH signals with $4f_{sc}$ sampled subcarrier has the power spectral density very close to AltBOC signals; and CBFH signals with $8f_{sc}$ sampled subcarrier has the same shape with the others, but it has the highest power at $\pm f_{sc}$, moreover, there is not harmonic at $\pm 3f_{sc}$ and $\pm 5f_{sc}$.

4.4 Receiver Complexity and Performance Analysis

The receiving of CBFH signal is the same as AltBOC, TD-AltBOC and ACED, two lobes can be received and processed separately as QPSK(10) signal, its acquisition tracking and demodulation algorithms are the same as GPS L5C; or two lobes coherently processed to achieve better performance.

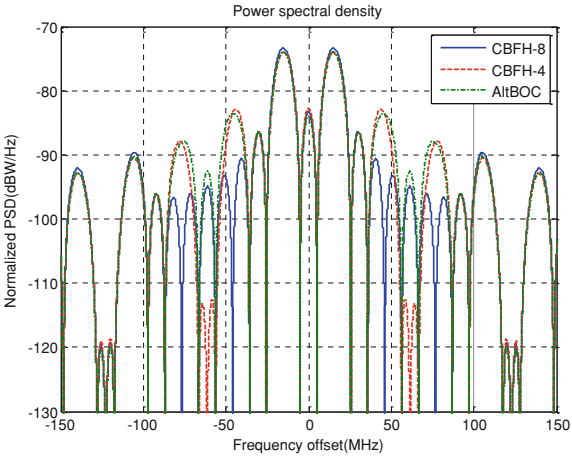


Fig. 4.4 PSD of CBFH and AltBOC signals

In respect of the complexity of CBFH signals generation, receiving and processing, CBFH with $4f_{sc}$ subcarrier sample rate has the same complexity as TD-AltBOC, which less than AltBOC and ACED, the sample rate of subcarrier of AltBOC is $8f_{sc}$ and ACED is $12f_{sc}$, and $12f_{sc}$ sampled subcarrier is too difficult to implement. What more important is that the level of CBFH on this situation is 2 levels, and AltBOC and ACED are 4 levels, so when receiving the signals with a matched waveform, less hardware resources and less power consumption will be cost by CBFH method. If CBFH with $8f_{sc}$ subcarrier sample rate, TD-AltBOC will not be considered here, it has the same subcarrier sample rate as AltBOC, but still less than ACED; in addition, because the absence of inter-modulation products, so the subcarrier of CBFH can generate by sample hold device, while the AltBOC signal and ACED signal generation must be adopt lookup table which is more complex than sample hold device.

In respect of power efficiency, CBFH with $4f_{sc}$ subcarrier sample rate has the same power efficiency as AltBOC, TD-AltBOC and ACED, it is 81.06 %, in other word, the power loss (consist of the loss of sampled subcarrier and the loss caused by the introducing of inter-modulation products) is 0.91 dB. When subcarrier sample rate of CBFH is up to $8f_{sc}$, we can obtain that the power efficiency is 94.96 %, the power loss is 0.22 dB which has 0.69 dB decreased than AltBOC, TD-AltBOC and ACED. So the later of this paper, CBFH with $8f_{sc}$ subcarrier sample rate are adopted in analyzing.

CBFH signal has the same power spectral density shape as AltBOC, TD-AltBOC and ACED, so the range performance and the performance of resisting multipath depend on the power, from the foregoing analysis, CBFH has higher power efficiency than AltBOC, TD-AltBOC and ACED, so when the total power is constant, CBFH has the highest power, so CBFH has the best performance.

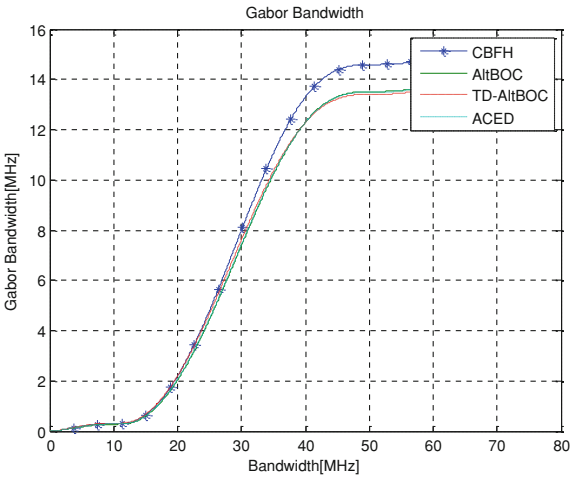


Fig. 4.5 Gabor bandwidth of CBFH, AltBOC, TD-AltBOC and ACED signals

Figure 4.5 depicts the Gabor bandwidth [6] of the four methods, we can see that CBFH has larger Gabor bandwidth than AltBOC, TD-AltBOC and ACED, especially when the receiving bandwidth is larger than 50 MHz, the Gabor bandwidth of CBFH is 1 dB more than AltBOC, TD-AltBOC and ACED.

4.5 Conclusion

In this paper, we proposed a CBFH method to achieve dual-frequency multiplexing, if the power of lower-band and up-band are equal, the power ratio between data-component and pilot-component at each band can adjust arbitrary with our method, and it has the same shape of power spectral density as AltBOC, TD-AltBOC and ACED, at the same time our method has the lowest complexity and highest power efficiency among CBFH, AltBOC, TD-AltBOC and ACED, moreover, CBFH will have better range performance and multipath resisting performance.

In a word, CBFH is a feasible candidate modulation scheme of BeiDou B2 band.

References

1. Rebeyrol E, Julien O, Macabiau C, Ries L, Delatour A, Lestarquit L (2007) Galileo civil signal modulations. GPS Solutions 11(3):159–171
2. Laurent L, Geraldine A, Jean-Luc I (2008) AltBOC for dummies or everything you always wanted to know about AltBOC. In: ION GNSS 21st international technical meeting of the satellite division, 2008
3. Zuping T, Hongwei Z, Jiaolong W et al (2011) TD-AltBOC: a new compass B2 modulation. Sci China: Phy Mech Astron 54(6):1014–1021

4. Zhang K (2013) Generalised constant-envelope DualQPSK and AltBOC modulations for modern GNSS signals. Electron Lett 49(21):1335–1337
5. Yao Z, Lu M (2013) ACED Multiplexing and Its Application on BeiDou B2 Band. China Satellite Navigation Conference (CSNC) 2013. Proceedings Lecture Notes inElectrical Engineering 244:25–34
6. Tang Z, Zhou HW, Hu XL, Ran YH, Liu YQ, Zhou YL (2010) Evaluate of compass navigation signals. Sci Sinica (Physica, Mechanica Astronomica) 5:228.4

Chapter 5
Navigation Signal Acquisition Method Based on Multiple DBZP Fusion

Tian Jin, Jianlei Yang, Zhigang Huang and Honglei Qin

Abstract With the upcoming of modernized global positioning system (GPS) and Galileo, the new generation global navigation satellite systems (GNSS) will use signals with equal spreading code and bit periods, resulting in a potential bit sign transition in each primary code period of the received signal segments. The double block zero padding (DBZP) method and its improved versions overcome the correlation peak loss caused by bit sign transitions, but 50 % of the correlation peak energy is lost because it only uses one correlation period accumulation during two correlation time processing. To solve the problem, an acquisition method based on multiple double block zero padding fusion (MDBZPF) has been proposed to reduce the correlation loss by fusing three different DBZP strategies. The theoretical and simulated resulted show the proposed method can provide 2–3 dB acquisition sensitivity improvement compared with the DBZP method, and reduce the average acquisition time about 5–70 %.

Keywords Navigation signal · Acquisition · Correlation peak loss · Multiple DBZP fusion · DBZP

5.1 Introduction

To meet the growing demand of navigation and positioning services, tiered codes and spreading code with the same period as data transition will be introduce in the GPS modernization and Galileo signal structure design process (such as L1C, L5C, E1OS, E6CS and E5). The bit sign transition could possibly occur in any primary code period within the received GNSS signals processed in the acquisition stage. If

T. Jin · J. Yang (✉) · Z. Huang · H. Qin
School of Electronic and Information Engineering, BeiHang University, Beijing, People's Republic of China
e-mail: yangjianlei@ee.buaa.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_5,

traditional methods (such as non-coherent (NCH) method [1], real differentially coherent (RDF) method [2, 3], complex differentially coherent (CDF) [4, 5] and double differentially (DDF) [6] method, etc.) are used to perform the correlation, the data bit transition occurring within an integration time may easily cause the correlation peak energy loss seriously.

To solve the data transition problems, Lin and Tsui [7, 8] proposed the double block zero padding (DBZP) method, which uses two consecutive data bit period to ensure that there is at least one accumulation interval that do not contain data bit transitions. However, this method requires estimating the most reliable data edge combination and has high computation complexity. To reduce the computation complexity, researchers proposed modified double block zero padding (MDBZP) [9], improved fast modified double block zero padding (FMDBZP) [10] and improved fast modified double block zero padding (IFMDBZP) [11] methods. DBZP technology has been successfully applied to the acquisition of GPS L5 and Galileo E5 [12, 13]. Although the above DBZP methods solve the correlation peak loss caused by the presence of bit sign transitions, 50 % of the signal power is lost because it only uses one correlation accumulation during two correlation periods of processing.

To use more correlation accumulation time, this paper proposes a multiple double block zero padding fusion (MDBZPF) acquisition method. Theoretical analyses of DBZP and MDBZPF methods are performed. In order to support the theoretical analyses, Monte Carlo simulations have been used to evaluate the performances of the proposed method. The analysis and simulation results demonstrate that MDBZPF method greatly reduce the correlation peak loss compared with DBZP method, improve the signal detection performance and reduce average acquisition time (MAT).

5.2 System Model

When code phase and carrier frequency of local replicated signal match with the incoming signal, the normalized in-phase (I) and quadrature-phase (Q) branches coherent integrator output can be modeled as [14]

$$\begin{cases} U_{I,i} = \sqrt{(C/N_0)N_{coh}T_c}D_iR(\Delta\tau_i)\sin c(\pi\Delta f_i N_{coh}T_c)\cos(\Delta\phi_i) + N_{I,i} \\ U_{Q,i} = \sqrt{(C/N_0)N_{coh}T_c}D_iR(\Delta\tau_i)\sin c(\pi\Delta f_i N_{coh}T_c)\sin(\Delta\phi_i) + N_{Q,i} \end{cases} \tag{5.1}$$

where C/N_0 is the carrier to noise ratio of signal, N_{coh} is the coherent accumulation times, T_c is the period of pseudorandom noise (PN) code, i is the epoch, D_i is the data bit, $\Delta\tau_i$ is PN code phase estimation error, $\Delta\phi_i$ is carrier phase estimation error, Δf_i is carrier frequency estimation error, $R(\cdot)$ is the auto-correlation function of pseudorandom noise (PN) code, $N_{I,i}$ and $N_{Q,i}$ are the noise variance of I and Q branches respectively, and both of them follow the standard normal distribution.

Assuming that data bit is not changed during coherent accumulation time, the detector can be expressed as

$$F(\Delta f_i, \Delta\tau_i) = U_{I,i}^2 + U_{Q,i}^2 = (C/N_0)N_{coh}T_c R^2(\Delta\tau_i)\sin c^2(\pi\Delta f_i N_{coh}T_c). \quad (5.2)$$

According to (5.2), the detector is a function of Δf_i and $\Delta\tau_i$. When the PN code phase and carrier frequency of incoming signal and local replicated signal are completely matched, $\Delta f_i = 0$, $\Delta\tau_i = 0$, and the detector $F(\Delta f_i, \Delta\tau_i)$ reaches the maximum $(C/N_0)N_{coh}T_c$.

If hypothesis H_0 and H_1 represent two circumstances that the signal does not exist and exist respectively, the probability density function (PDF) corresponding to H_0 and H_1 can be written as follows [15]

$$p(y|H_0) = \frac{1}{2}\exp\left(-\frac{y}{2}\right)(y \geq 0) \quad (5.3)$$

$$p(y|H_1) = \frac{1}{2}\exp\left(-\frac{y + (C/N_0)N_{coh}T_c}{2}\right) I_0\left(\sqrt{y(C/N_0)N_{coh}T_c}\right) \quad (y \geq 0). \quad (5.4)$$

The false alarm and detection probabilities are as follows

$$P_{fa} = \exp\left(-\frac{V_t}{2}\right) \quad (5.5)$$

$$P_d = Q\left(\sqrt{(C/N_0)T_{coh}}, \sqrt{V_t}\right) \quad (5.6)$$

where V_t is the detection threshold, and $Q(\alpha, \beta)$ is the first-order Marcum-Q function. The generalized Marcum-Q function of M order is defined as

$$Q(\alpha, \beta) = \frac{1}{\alpha^{M-1}}\int_{\beta}^{\infty} x^M \exp\left(-\frac{x^2 + \alpha^2}{2}\right) I_{M-1}(\alpha x)dx. \quad (5.7)$$

5.3 DBZP Method

The principle of the DBZP method is shown in Fig. 5.1, where D_a, D_b and D_c are data bits, which equals to +1 or –1, T_{coh} is the coherent accumulation time, $mT_{coh}(0 \leq m \leq 1)$ is the position of data bit transitions. If the incoming signal is represented as S_{D_a,D_b,D_c}, the signal have eight possible combinations: $S_{+1,+1,+1}$, $S_{-1,+1,-1}$, $S_{-1,+1,+1}$, $S_{-1,-1,+1}$, $S_{-1,-1,-1}$, $S_{+1,-1,+1}$, $S_{+1,-1,-1}$ and $S_{+1,+1,-1}$, F_{D_a}, $D_b, D_c(\Delta f_i, \Delta\tau_i)$ is the detector corresponding to S_{D_a,D_b,D_c}.

In the process of acquisition, since the impact on acquisition performance of the signal $S_{+1,+1,+1}$, $S_{-1,+1,-1}$, $S_{-1,+1,+1}$, $S_{-1,-1,+1}$ is the same as $S_{-1,-1,-1}$,

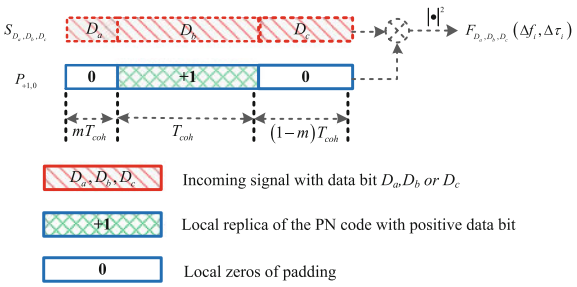


Fig. 5.1 Scheme of DBZP method

$S_{+1,-1,+1}$, $S_{+1,-1,-1}$ and $S_{+1,+1,-1}$. Therefore, this paper uses the first four signals as illustration.

DBZP method only uses one data bit interval for accumulation. Letting $N_{coh}T_c = T_{coh}$ in (5.6), the false alarm probability is the same as (5.5), and detection probability of DBZP method can be obtained as

$$P_d^{DBZP} = Q\left(\sqrt{(C/N_0)T_{coh}}, \sqrt{V_t}\right). \tag{5.8}$$

According to (5.8), the performance of ZP method is not impacted by data bit transition, but two data bit interval correlation is performed that only uses one data bit interval for accumulation, and therefore 50 % of the signal power is lost.

5.4 MDBZPF Method

5.4.1 MDBZPF Method Description

In order to overcome correlation peak energy loss, this paper proposed the MDBZPF method, as shown in Fig. 5.2. For the local replicated PN code, there are $P_{0,+1}$ (creating one PN code and padding with one PN code period of zeros ahead of PN code) and $P_{0,+1\,-1}$ (replicating $P_{0,+1}$ and the second half part of PN code with the negative bit) besides the $P_{+1,0}$ in DBZP method as shown in Fig. 5.3.

The local PN code $P_{+1,0}$, $P_{0,+1}$ and $P_{0,+1\,-1}$ are adopted to correlate with incoming signal respectively, the corresponding correlation sequence $corr_{0,+1}(k)$, $corr_{+1,0}(k)$ and $corr_{0,+1\,-1}(k)$ can be obtained, where k represent the k-th correlation value. If the incoming signals are $S_{+1,+1,+1}$ and $S_{-1,+1,-1}$, the coherent accumulated values $corr_{+1,0}(k) + corr_{0,+1}(k)$ and $corr_{+1,0}(k) - corr_{0,+1}(k)$ contain two data bit interval accumulations. If incoming signals are $S_{-1,+1,+1}$ and $S_{-1,-1,+1}$, select the optimum correlation peak from $corr_{+1,0}(k) + corr_{0,+1}(k)$, $corr_{+1,0}(k) - corr_{0,+1}(k)$, $corr_{+1,0}(k) + corr_{0,+1\,-1}(k)$, $corr_{+1,0}(k) - corr_{0,+1\,-1}(k)$ as the final detectors by the optimal function $f_{optimal}(\bullet)$. Then the detector $F_{D_a,D_b,D_c}(\Delta f_i, \Delta\tau_i)$ corresponding with S_{D_a,D_b,D_c} can be obtained.

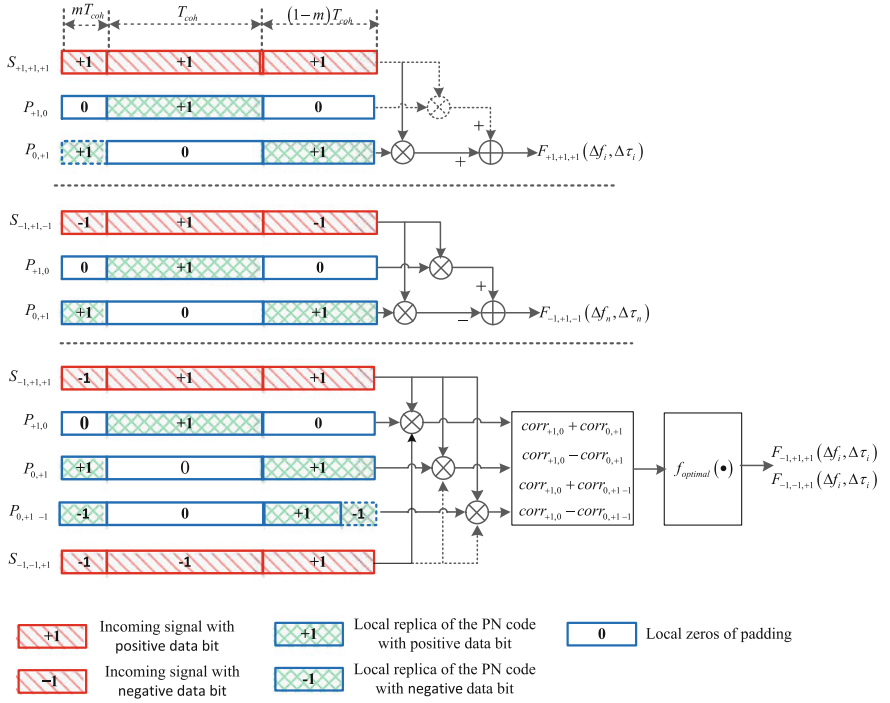


Fig. 5.2 Scheme of MDBZPF method

Fig. 5.3 Local PN code structure of MDBZPF method

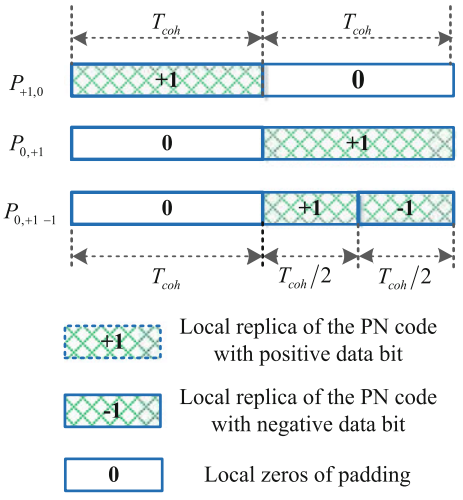


Table 5.1 Attenuation factor of MCSF

Signal type	Local PN code		
	$P_{+1,0} + P_{0,+1}$	$P_{+1,0} + P_{0,+1\,-1}$	
	$0 \le m \le 1$	$0 \le m \le 0.5$	$0.5 < m \le 1$
$S_{+1,+1,+1}$	1	0	0
$S_{-1,+1,-1}$	1	0	0
$S_{-1,+1,+1}$	$1-m$	$m+\frac{1}{2}$	$\frac{3}{2}-m$
$S_{-1,-1,+1}$	m	$m+\frac{1}{2}$	$\frac{3}{2}-m$

5.4.2 Performance Analysis of MDBZPF Method

According to Fig. 5.2, the coherent integration time is $2T_c$, and letting $N_c = 2$ in (5.6). If the data bit transition occurs at $mT_c (0 \le m \le 1)$, attenuation factors $g(m)$ will introduced, then the amplitude of detector is reduced to $g(m)^2 2T_c C/N_0$ from $2T_c C/N_0$. The attenuation factors are shown in Table 5.1.

According to (5.4), the PDF of detector $F_{D_a,D_b,D_c}(\Delta f_i, \Delta\tau_i)$ when bit transition occurs at $mT_c (0 \le m \le 1)$ is

$$\begin{aligned} &p(y|H_1, g(m), 0 < m < 1) \\ &= \frac{1}{2}\exp\left(-\frac{y + g(m)^2 2T_c C/N_0}{2}\right) I_0\left(\sqrt{y g(m)^2 2T_c C/N_0}\right). \end{aligned} \tag{5.9}$$

Consequently, the detection probability can be expressed by

$$P_d = \int_{V_t}^{\infty} p_0(y|H_1, g(m), 0 < m < 1)dy = Q\left(\sqrt{g(m)^2 2T_c C/N_0}, \sqrt{V_t}\right). \tag{5.10}$$

Based on Table 5.1 and (5.10), Fig. 5.4 shows the detection probabilities by performing the correlation between $S_{-1,+1,+1}$, $S_{-1,-1,+1}$ and $P_{+1,0} + P_{0,+1}$ (corresponding with $corr_{+1,0}(k) + corr_{0,+1}(k)$ and $corr_{+1,0}(k) - corr_{0,+1}(k)$), $P_{+1,0}$+$P_{0,+1\text{-}1}$ (corresponding with $corr_{+1,0}(k) + corr_{0,+1\,-1}(k)$ and $corr_{+1,0}(k) - corr_{0,+1\,-1}(k)$) respectively. Simulated MBOC (6,1,1/11) of B1C signal [16] generating by TMBOC mode is used, and false alarm $P_{fa} = 10^{-3}$, carrier-to-noise ratio C/N_0 = 27 dB-Hz.

According to Fig. 5.4, optimal function $f_{optimal}(\bullet)$ of $S_{-1,+1,+1}$ and $S_{-1,-1,+1}$ can be obtained as

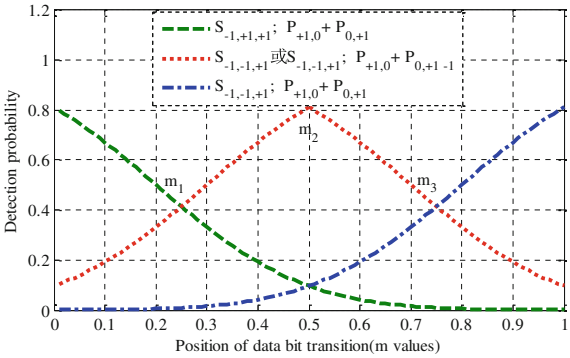


Fig. 5.4 Detection probabilities of $S_{-1,+1,+1}$, $S_{-1,-1,+1}$ correlation with $P_{+1,0} + P_{0,+1}$, $P_{+1,0} + P_{0,+1\,-1}$

$$f_{optimal}(m) = \begin{cases} Q\left(\sqrt{(1-m)^2 2T_c C/N_0}, \sqrt{V_t}\right) & 0 \le m \le m_1 \\ Q\left(\sqrt{\left(m+\frac{1}{2}\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) & m_1 \le m \le m_2 \\ Q\left(\sqrt{\left(\frac{3}{2}-m\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) & m_2 \le m \le m_3 \\ Q\left(\sqrt{m^2 2T_c C/N_0}, \sqrt{V_t}\right) & m_3 \le m \le 1. \end{cases} \tag{5.11}$$

According to Fig. 5.4 and (5.10) the following three equations is established,

$$Q\left(\sqrt{(1-m_1)^2 2T_c C/N_0}, \sqrt{V_t}\right) = Q\left(\sqrt{\left(m_1+\frac{1}{2}\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) \tag{5.12}$$

$$Q\left(\sqrt{\left(m_2+\frac{1}{2}\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) = Q\left(\sqrt{\left(\frac{3}{2}-m_2\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) \tag{5.13}$$

$$Q\left(\sqrt{\left(\frac{3}{2}-m_3\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) = Q\left(\sqrt{m_3^2 2T_c C/N_0}, \sqrt{V_t}\right). \tag{5.14}$$

Solving the Eqs. (5.12), (5.13) and (5.14), three intersection abscissa values m_1, m_2 and m_3 in Fig. 5.4 can be obtained as follows

$$\begin{cases} m_1 = \dfrac{1}{4} \\ m_2 = \dfrac{1}{2} \\ m_3 = \dfrac{3}{4}. \end{cases} \tag{5.15}$$

Because $m \sim U(0,1)$, according to (5.9), (5.11) and (5.15), the PDF of $S_{-1,+1,+1}$ and $S_{-1,-1,+1}$ can be can be derived as follows

$$\begin{aligned} p\left(y|H_1, S_{-1,+1,+1}\right) &= p\left(y|H_1, S_{-1,-1,+1}\right) \\ &= \int_0^{1/4} \frac{1}{2}\exp\left(-\frac{y+(1-m)^2 2T_cC/N_0}{2}\right) I_0\left(\sqrt{y(1-m)^2 2T_cC/N_0}\right)\, dm \\ &+ \int_{1/4}^{1/2} \frac{1}{2}\exp\left(-\frac{y+\left(m+\frac{1}{2}\right)^2 2T_cC/N_0}{2}\right) I_0\left(\sqrt{y\left(m+\frac{1}{2}\right)^2 2T_cC/N_0}\right)\, dm \\ &+ \int_{1/2}^{3/4} \frac{1}{2}\exp\left(-\frac{y+\left(\frac{3}{2}-m\right)^2 2T_cC/N_0}{2}\right) I_0\left(\sqrt{y\left(\frac{3}{2}-m\right)^2 2T_cC/N_0}\right)\, dm \\ &+ \int_{3/4}^{1} \frac{1}{2}\exp\left(-\frac{y+m^2 2T_cC/N_0}{2}\right) I_0\left(\sqrt{ym^2 2T_cC/N_0}\right)\, dm. \end{aligned} \tag{5.16}$$

For the incoming signal $S_{+1,+1,+1}$ and $S_{-1,+1,-1}$, local replicated PN code $P_{+1,0} + P_{0,+1}$ is adopted. Putting $g(m) = 1$ into (5.9), the PDF of $S_{+1,+1,+1}$ and $S_{-1,+1,-1}$ can be can be derived as follows

$$\begin{aligned} p\left(y|H_1, S_{+1,+1,+1}\right) &= p\left(y|H_1, S_{-1,+1,-1}\right) \\ &= \frac{1}{2}\exp\left(-\frac{y+2T_cC/N_0}{2}\right) I_0\left(\sqrt{y2T_cC/N_0}\right). \end{aligned} \tag{5.17}$$

Assuming that the probabilities of $S_{+1,+1,+1}$, $S_{-1,+1,-1}$, $S_{-1,+1,+1}$ and $S_{-1,-1,+1}$ are $1/4$, which can be expressed as

$$p\left(S_{+1,+1,+1}\right) = p\left(S_{-1,+1,-1}\right) = p\left(S_{-1,+1,+1}\right) = p\left(S_{-1,-1,+1}\right) = \frac{1}{4}. \tag{5.18}$$

According to (5.16), (5.17) and (5.18), the PDF of MDBZPF can be obtained as

$$\begin{aligned} p^{MDBZPF}(y|H_1) &= p\left(y|H_1, S_{+1,+1,+1}\right)p\left(S_{+1,+1,+1}\right) + p\left(y|H_1, S_{-1,+1,-1}\right)p\left(S_{-1,+1,-1}\right) \\ &+ p\left(y|H_1, S_{-1,+1,+1}\right)p\left(S_{-1,+1,+1}\right) + p\left(y|H_1, S_{-1,-1,+1}\right)p\left(S_{-1,-1,+1}\right). \end{aligned} \tag{5.19}$$

The detection probability of MDBZPF can be derived as follows

$$
\begin{aligned}
P_d^{MDBZPF} &= \int_{V_t}^{\infty} p(y|H_1)dy \\
&= \frac{1}{2}\int_0^{1/4} Q\left(\sqrt{(1-m)^2 2T_c C/N_0}, \sqrt{V_t}\right) dm + \frac{1}{2}\int_{1/4}^{1/2} Q\left(\sqrt{\left(m+\frac{1}{2}\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) dm \\
&\quad + \frac{1}{2}\int_{1/2}^{3/4} Q\left(\sqrt{\left(\frac{3}{2}-m\right)^2 2T_c C/N_0}, \sqrt{V_t}\right) dm + \frac{1}{2}\int_{3/4}^{1} Q\left(\sqrt{m^2 2T_c C/N_0}, \sqrt{V_t}\right) dm \\
&\quad + \frac{1}{2} Q\left(\sqrt{m^2 2T_c C/N_0}, \sqrt{V_t}\right).
\end{aligned}
\tag{5.20}
$$

Putting $2T_c C/N_0 = 0$ into (5.19) the PDF under hypothesis H_0 can be obtained as (5.3). Consequently, the false alarm probability P_{fa}^{MDBZPF} of MDBZPF can be obtained as (5.5).

5.5 Simulation Analysis and Performance Comparison

To assess the performance of the MDBZPF method, Monte Carlo simulation is performed to the detection probability, the receiver operating characteristic (ROC) and MAT; and comparisons between MDBZPF and DBZP methods are made. The simulations use MBOC(6,1,1/11) of B1C signal [16] generating by TMBOC mode with random data bit transitions and PN code phases. The length of PN code is 10230 chips, and over 10^5 simulations are conducted.

5.5.1 Detection Probability

In Fig. 5.5 the detection probabilities for DBZP and MDBZPF methods have been plotted under different C/N_0 and false alarm$P_{fa} = 10^{-3}$.

From Fig. 5.5, it can be found that under the same detection probability MDBZPF method has 2-3 dB improvement in sensitivity compared with DBZP method.

5.5.2 Roc

ROC represents the detection probability under different probabilities of false alarm. Figure 5.6 shows ROC curves of DBZP and MDBZPF.

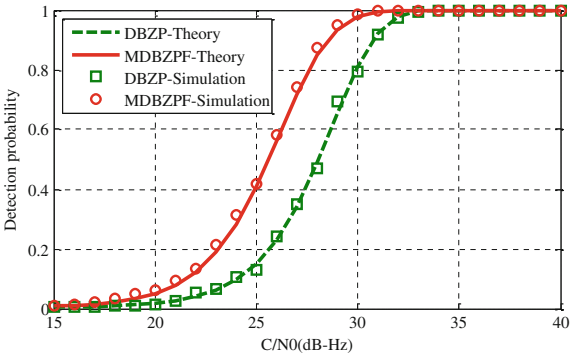


Fig. 5.5 Detection probabilities comparison among different acquisition methods

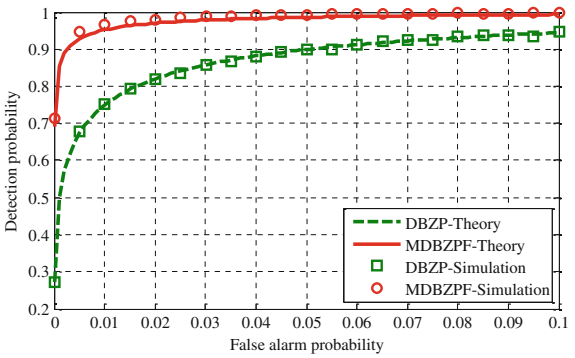


Fig. 5.6 ROC comparison among different acquisition methods

From Fig. 5.6, we can see that under certain probability of false alarm, the probability of detection for MDBZPF method is obviously higher than DBZP method. The advantage of MDBZPF in ROC curves also proves that it is a better acquisition method than DBZP method.

5.5.3 Mat

MAT [17] is an important indicator to assess acquisition performance, and the mean acquisition time of single dwell time search process is defined as

$$T_{MAT} = \frac{2 + (2 - P_d)\left(N_{code}N_{fd} - 1\right)\left(1 + k_{fa}P_{fa}\right)}{2P_d} T_c N_{nc} \tag{5.21}$$

where P_d is the detection probability, P_{fa} is the false alarm probability, T_c is the coherent integration time, N_{nc} is the non-coherent integration time, N_{code} is the bins of code phase search, N_{fd} is the cells of Doppler search, k_{fa} is the penalty factor associated with a false alarm state, so that $k_{fa}T_cN_{nc} = 1s$.

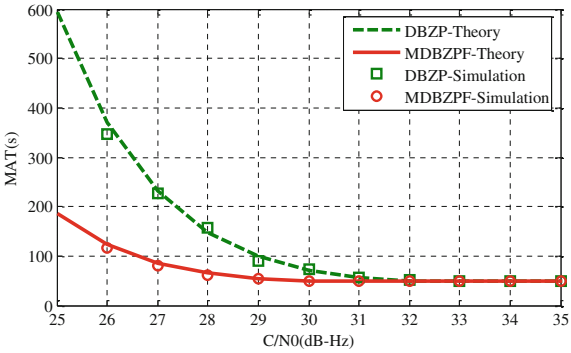


Fig. 5.7 MAT comparison among different acquisition methods

Assuming that Doppler range is −5 to +5 KHz, frequency search step is 25 Hz and code search step is 0.5 chips, frequency search cells $N_{fd} = 401$ and code search bins $N_{code} = 40920$. The MAT values of two methods are shown in Fig. 5.7 under false alarm $P_{fa} = 10^{-3}$.

The MDBZPF method show better performance with MAT reduced by 5–70 % compared with DBZP method when the two methods used same Doppler search step.

In the above analysis, the two methods used same Doppler search step. In practical environment, coherent accumulation time of DBZP method is half of the MDBZPF method, therefore the Doppler search step of DBZP can set to be two times of MDBZPF method.

5.6 Conclusions

In this paper, a MDBZPF acquisition method for reducing the correlation peak energy loss has been presented. The detection probability, ROC and MAT in MDBZPF have been discussed in MDBZPF method, and compared with DBZP method. The results show that MDBZPF method will introduce less correlation peak energy loss. Detection sensibility of MDBZPF is 2–3 dB higher than DBZP, and reduce MAT by 5–70 %. The proposed method can improve the performance of signal detection in the present of bit sign transition.

References

1. Viterbi AJ (1995) CDMA: principles of spread spectrum communication. Addison-Wesley, Reading
2. Zarrabizadeh RH, Sousa ES (1997) A differentially coherent PN code acquisition receiver for CDMA systems. IEEE Trans Commun 45(11):1456–1465

3. Ayaz AS (2005) Analysis of differential acquisition methods by using Monte-Carlo simulations. In: Proceedings of the 18th international technical meeting of the satellite division of The Institute of Navigation (ION GNSS 2005), Long Beach, CA
4. Elders-Boll H, Dettmar U (2004) Efficient differentially coherent code/Doppler acquisition of weak GPS signals. In: Proceedings of 2004 IEEE eighth international symposium on spread spectrum techniques and applications
5. Yu W, Zheng B, Watson R, Lachapelle G (2007) Differential combining for acquiring weak GPS signals. Signal Process 87(5):824–840
6. Jin T, Lu F, Liu Y, Qin H, Luo X (2013) Double differentially coherent pseudorandom noise code acquisition method for code-division multiple access system. IET Signal Proc 7(7):587–597
7. Lin DM, Tsui JBY, Howell D (1999) Direct P (Y)-code acquisition algorithm for software GPS receivers. In: Proceedings of the 12th international technical meeting of the satellite division of The Institute of Navigation (ION GPS 1999), Nashville, TN
8. Lin DM, Tsui JB (2000) Comparison of acquisition methods for software GPS receiver. In: Proceedings of the 13th international technical meeting of the satellite division of The Institute of Navigation (ION GPS 2000), Salt Lake City, UT
9. Ziedan NI, Garrison JL (2004) Unaided acquisition of weak GPS signals using circular correlation or double-block zero padding. In: Proceedings of position location and navigation symposium (PLANS 2004), IEEE
10. Heckler GW, Garrison JL (2009) Implementation and testing of an unaided method for the acquisition of weak GPS c/a code signals. Navigation 56(4):241–259
11. Zhang W, Ghogho M (2010) Improved fast modified double-block zero-padding (FMDBZP) algorithm for weak GPS signal acquisition. In: Proceedings of 18th European signal processing conference (EUSIPCO 2010), Aalborg, Denmark
12. Yang C, Hegarty C, Tran M (2004) Acquisition of the GPS L5 signal using coherent combining of I5 and Q5. In: Proceedings of the 17th international technical meeting of the satellite division of The Institute of Navigation (ION GNSS 2004), Long Beach, CA
13. Leclere J, Botteron C, Farine PA (2013) Modified parallel code-phase search for acquisition in presence of sign transition. In: Proceedings of 2013 international conference on localization and GNSS (ICL-GNSS 2013), Turin
14. Yao Z, Cui X, Lu M, Feng Z (2009) Dual update-rate carrier tracking technique for new generation global navigation satellite system signals in dynamic environments. IET Radar Sonar Navig 3(3):203–213
15. Borio D, Camoriano L, LoPresti L (2008) Impact of GPS acquisition strategy on decision probabilities. IEEE Trans Aerosp Electron Syst 44(3):996–1011
16. Yang Y (2009) China reveals updated Compass/Beidou-2 GNSS signal plan. In: Inside GNSS September/October 2009. Available via presentation. http://www.insidegnss.com/node/1624
17. Holmes JK (1982) Coherent spread spectrum systems. Wiley-Interscience, New York

Chapter 6
Weak Signal Acquisition Performance for Different Pilot Channel Design Options

Yingxiang Liu, Huaming Chen, Zhicheng Lv and Feixue Wang

Abstract Pilot Channel without navigation message modulation is added in next generation GNSS for enhancing the acquisition performance in challenge environment. The pre-detection integration time of pilot channel is potential to exceed the bit width, so the receiver could obtain better acquisition sensibility. However, the benefit from pilot channel has not been thoroughly analyzed. The equivalent ideal detector factor is used to compare the performance of different power allocation options of data and pilot channel. The results show that when the data and pilot channel has the same power, the acquisition sensibility is not better than that of traditional signal in normal conditions. The conclusion has significance on the BDS modernized signal design.

Keywords GNSS · Weak signal acquisition · Pilot channel

6.1 Introduction

With the rapid development of satellite navigation technology, the application scope of GPS has far exceeded the initial design goals. In urban canyons and jungle areas, the occlusion of buildings and trees make the received signal power much lower than that in the open area of about 20–30 dB, which is about −150 to −160 dBm [1]. In traditional global navigation satellite system (GNSS), the coherent integration time is hard to exceed the navigation data width because of navigation data modulation, which restricts the application of satellite navigation in weak signal environments.

Y. Liu · H. Chen · Z. Lv · F. Wang (✉)
Satellite Navigation and Positioning R&D Center, School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China
e-mail: wangfeixue_nnc@163.com

Y. Liu
e-mail: inshine1986@gmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_6,

In order to improve the availability in weak signal environments, pilot channel without navigation message modulation is widely used in modernized GNSS. The pilot channel was initially proposed in Transit navigation system in the 1960s for the first time, and some scholars also suggested that it could be used in GPS signals [2, 3]. However, due to the limit of integrated circuits (IC) technique at that time, navigation receiver is impossible to operate under weak signal environment. If the signal power is allocated to data and pilot channels, it is impossible to enhance the sensitivity, but increase the design complexity of the receiver. So the development of the IC techniques makes the pilot channel widely used in the next generation GNSS.

The design of pilot channel includes multiplexing techniques and power allocation options. The multiplexing techniques of data and pilot channels could be categorized into modulated on two orthogonal carriers [4], modulated on single carrier [5], and modulated on single carrier by time-division data multiplex [6, 7]. Although the implementation of signal acquisition for the three multiplexing techniques is different, the acquisition performance is the same when the parameters are identical. The research for pilot channel acquisition is currently focused on the combination of data and pilot channel [8], but the acquisition performance for pilot channel with different power allocation options have not been thoroughly compared.

The signal model of data and pilot channel with different power allocation option is given in the paper at first, and then the equivalent ideal detector factor is used to compare the detection performance. In the end suggestion has been made according to the conclusion.

6.2 Signal Model

Signal acquisition is essentially a signal detection problem of unknown power, code phase delay and Doppler frequency. Take the QPSK modulation as an example, the signal model of data and pilot channel can be expressed as the following binary hypothesis:

$$H_0: \quad s(t) = n(t) \tag{6.1}$$

$$\begin{aligned} H_1: \quad s(t) = &\sqrt{\eta 2C}d(t)c_d(t-\tau)\cos[2\pi(f_0+f_d)t+\theta_0] \\ &+\sqrt{(1-\eta)2C}c_p(t-\tau)\sin[2\pi(f_0+f_d)t+\theta_0]+n(t) \end{aligned} \tag{6.2}$$

where C represents the total signal power of data and pilot channel, η and $1-\eta$ is the power ratio of data and pilot channel respectively, $c_d(t)$ and $c_p(t)$ is the ranging code, τ is the time delay, f_0 is the nominal carrier frequency, f_d is the Doppler frequency, θ_0 is the initial carrier phase, $n(t)$ represents Gaussian white noise, and its unilateral power spectral density is N_0, so carrier to noise ratio is C/N_0.

In the signal model, the secondary code modulation on data and pilot channel is ignored. This is because under weak signal acquisition the ranging code and secondary code is always considered as a whole in order to increase the pre-detection integration time. On that condition, the bit synchronization is completed while signal is acquired.

Three power allocation options is adopted in modernize GNSS signals, which are:

Option 1, all the power is devoted on data channel, and $\eta = 1$;

Option 2, the power is allocated on the data and pilot channels with the ratio of 1:1, and $\eta = 1/2$;

Option 3, the power is allocated on the data and pilot channels with the ratio of 1:3, and $\eta = 1/4$.

Option 4, which all the power is devoted on pilot channel and $\eta = 0$, is also analyzed in the paper. Considering, dual frequency receivers could be the mainstream in the future, the option is theoretically feasible.

6.3 Statistical Detector

In order to reduce the Doppler frequency loss and eliminate the impact of sign reverse, the segmental coherent integration-video non-coherent integration acquisition is usually used in signal acquisition, and its implementation structure is shown in Fig. 6.1.

For the power allocation option 1 and 4, only correlation values of single channel needs to be calculated, so the correlators is only half that of option 2 and 3 if the total integration time is identical.

Assuming the pre-detection integration time of data and pilot channel is T_d and T_p, respectively, and the expression of the correlation values $v_d[k_1]$ and $v_p[k_2]$ is

$$v_d[k_1] = \int_{k_1 T_d}^{(k_1+1)T_d} s(t) c_d(t) \exp(-j2\pi f_0 t) dt \tag{6.3}$$

$$v_p[k_2] = \int_{k_2 T_p}^{(k_2+1)T_p} s(t) c_p(t) \exp(-j2\pi f_0 t) dt \tag{6.4}$$

where k_1 and k_2 is the index of the correlation values of data and pilot channels, which ranges from $[0, K_1 - 1]$ and $[0, K_2 - 1]$. Obviously, the total integration time of data and pilot channels is identical, so that $K_1 T_d = K_2 T_p$.

To increase the signal to noise ratio, the correlation values are non-coherently summed:

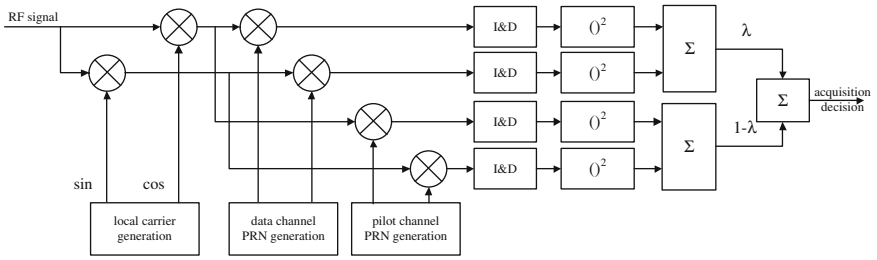


Fig. 6.1 Scheme of segmental coherent integration-video non-coherent integration acquisition

$$\begin{aligned} V &= \lambda \sum_{k_1=0}^{K_1-1} |v_d[k_1]|^2 + (1-\lambda) \sum_{k_2=0}^{K_2-1} |v_p[k_2]|^2 \\ &\triangleq \lambda V_d + (1-\lambda) V_p \end{aligned} \tag{6.5}$$

where λ is the combination coefficient, V_d and V_p denote the accumulated correlation values of data and pilot channels.

Obviously, for the option 1 and 4, λ is 1 and 0 respectively. The optimum coefficient for option 2 and 3 is analyzed herein. The correlation values of data and pilot channels under hypothesis H_0 and H_1 are Gaussian random variables:

$$H_0: \quad v_d[k_1] \sim \mathrm{N}\left(0, 2\sigma_d^2\right) \tag{6.6}$$

$$v_p[k_2] \sim \mathrm{N}\left(0, 2\sigma_p^2\right) \tag{6.7}$$

$$H_1: \quad v_d[k_1] \sim \mathrm{N}\left(d[k_1]\alpha_1 \mathrm{e}^{j(\pi f_d T_d + \theta_0)} \mathrm{e}^{j2\pi k_1 f_d T_d}, 2\sigma_d^2\right) \tag{6.8}$$

$$v_p[k_2] \sim \mathrm{N}\left(\alpha_2 \mathrm{e}^{j\left(\pi f_d T_p + \theta_0\right)} \mathrm{e}^{j2\pi k_2 f_d T_p}, 2\sigma_p^2\right) \tag{6.9}$$

where $d[k_1]$ represents the value of d(t) at the time of $k_1 T_d$, α_1 and α_2 is the magnitude of correlation values of the data and pilot channels, which is

$$\alpha_1 = \frac{\sqrt{\eta 2C}}{2} R(\tau) \mathrm{sinc}(\pi f_d T_d) \tag{6.10}$$

$$\alpha_2 = \frac{\sqrt{(1-\eta) 2C}}{2} R(\tau) \mathrm{sinc}\left(\pi f_d T_p\right) \tag{6.11}$$

where $R(\tau)$ is the autocorrelation function of the ranging code.

Under the two hypotheses, the real and imaginary parts of the data /pilot channel correlation values are independent and have equal variances, which are σ_d^2 and σ_{p}^2, and the expressions are

$$\sigma_d^2 = \frac{N_0}{4} T_d \tag{6.12}$$

$$\sigma_p^2 = \frac{N_0}{4} T_p. \tag{6.13}$$

Under the hypothesis H_0, $|v_d[k_1]|^2$ and $|v_p[k_2]|^2$ are the central χ^2 distributed random variables with degrees of freedom 2, and the mean and variance are

$$\mu_{d0} = 2\sigma_d^2, \mu_{p0} = 2\sigma_p^2 \tag{6.14}$$

$$\sigma_{d0}^2 = 4\sigma_d^4, \sigma_{p0}^2 = 4\sigma_p^4. \tag{6.15}$$

Under the hypothesis H_1, $|v_d[k_1]|^2$ and $|v_p[k_2]|^2$ are the non-central χ^2 distributed random variables with degrees of freedom 2, and the mean and variance are

$$\mu_{d1} = \alpha_1^2 + 2\sigma_d^2, \mu_{p1} = \alpha_2^2 + 2\sigma_p^2 \tag{6.16}$$

$$\sigma_{d1}^2 = 4\alpha_1^2\sigma_d^2 + 4\sigma_d^4, \sigma_{p1}^2 = 4\alpha_1^2\sigma_p^2 + 4\sigma_p^4. \tag{6.17}$$

If the post-detection integration times K_d and K_p is large, V_d and V_p can be approximated as Gaussian variables according to the law of large numbers, and the statistical detector V is also Gaussian variable because of the independence of V_d and V_p, and its mean value and variance are

$$H_0: \quad \mu_0 = \lambda K_d \mu_{d0} + (1-\lambda) K_p \mu_{p0} \tag{6.18}$$

$$\sigma_0^2 = \lambda^2 K_d \sigma_{d0}^2 + (1-\lambda)^2 K_p \sigma_{p0}^2 \tag{6.19}$$

$$H_1: \quad \mu_1 = \lambda K_d \mu_{d1} + (1-\lambda) K_p \mu_{p1} \tag{6.20}$$

$$\sigma_1^2 = \lambda^2 K_d \sigma_{d1}^2 + (1-\lambda)^2 K_p \sigma_{p1}^2. \tag{6.21}$$

Therefore, the detection probability of the test statistic is

$$P_d = Q\left(Q^{-1}(P_{fa}) \frac{\sigma_0}{\sigma_1} - \frac{\mu_1 - \mu_0}{\sigma_1} \right) \tag{6.22}$$

where P_{fa} is the preset false alarm probability, $Q(x)$ is the right tail function, which is

$$Q(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} t^2 \right) dt. \tag{6.23}$$

By substituting the mean value and variance under the two assumptions into Eq. (6.20), the detection performance under different coefficient λ can be obtained. Since the relationship between the coefficient λ and the detection performance is so complex to analytically expressed, the optimal coefficient can only be obtained by numerical calculation. In the following analysis, the optimal coefficient is used.

6.4 Detector Loss

In order to comparison the detection performance for different power allocation options of data and pilot channel, the equivalent ideal detector factor is used herein [9]. The equivalent ideal detector factor of data and pilot channel is

$$D_d = \frac{K_1 SNR_d^2}{SNR_d + 2.3}, \; D_p = \frac{K_2 SNR_p^2}{SNR_p + 2.3} \tag{6.24}$$

where in SNR_d and SNR_p is the signal to noise ratio of the detector of the data and the pilot channels, which is

$$SNR_d = \eta C / N_0 T_d R^2(\tau) \mathrm{sinc}^2(\pi f_d T_d) \tag{6.25}$$

$$SNR_p = (1-\eta) C / N_0 T_p R^2(\tau) \mathrm{sinc}^2(\pi f_d T_p). \tag{6.26}$$

Therefore, the equivalent ideal detector factor of the non-coherent combination of the data and pilot channel is

$$D = \frac{\lambda^2 D_d + 2\lambda(1-\lambda)\sqrt{D_d D_p} + (1-\lambda)^2 D_p}{\lambda^2 + (1-\lambda)^2}. \tag{6.27}$$

According to the above equation, the equivalent ideal detector factor for different power allocation options can be obtained, denoted as Di (i = 1–4). The relative detector loss is defined as

$$L_i = \frac{D_i}{D_1}. \tag{6.28}$$

Since the acquisition scheme for options 2 and 3 needs to calculate the correlation values of both data and pilot channels, the total integration time of options 2 and 3 is only half that of options 1 and 4 under the limitation of identical hardware complexity. Therefore, when calculating the relative detector loss, the impact of unequal integration time should be considered.

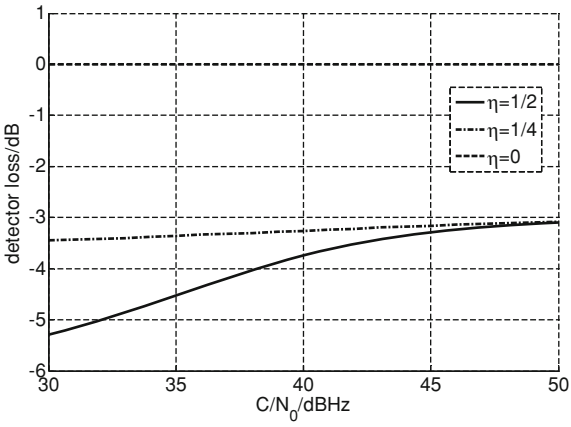


Fig. 6.2 Detector loss under different C/N_0, $T_d = T_p = 1$ ms

6.5 Performance Comparison

The detector loss is related with input carrier to noise ratio, pre-detection integration time, code phase bias, Doppler frequency error. Since the code phase and Doppler frequency error is a random variable, in order to simplify the analysis process, the optimal performance is used for comparison, thus ignoring the loss of the code phase and Doppler frequency error.

The pre-detection integration time of data and pilot channel can be divided into two situations:

1. Under the environment of high carrier to noise ratio, the pre-detection integration time of data and pilot channel is identical and smaller than the bit width;
2. Under the environment of low carrier to noise ratio, the pre-detection integration time of data channel equals the bit width and that of pilot channel is larger than the bit width.

In the next generation GNSS, the symbol rate of 100 sps is widely used [4, 5]. Here take the symbol width of 10 ms as an example, when the pre-detection integration time of data and pilot channel is less than the bit width, the detector loss for three options is shown in Figs. 6.2, 6.3, 6.4 and 6.5.

From the above analysis results, when the pre-detection integration time is less than the bit width, the detection performance is identical for options 1 and 4; while that for option 2 and 3 is far below.

When the carrier to noise ratio is relatively low and the Doppler frequency uncertainty is small, the pre-detection integration time of pilot channel can exceed the bit width. The detector loss for three options is illustrated in Figs. 6.6, 6.7, 6.8 and 6.9.

It can be seen from the figure, when the pre-detection integration time is larger than the bit width, the detection performance for option 4 is much better than that

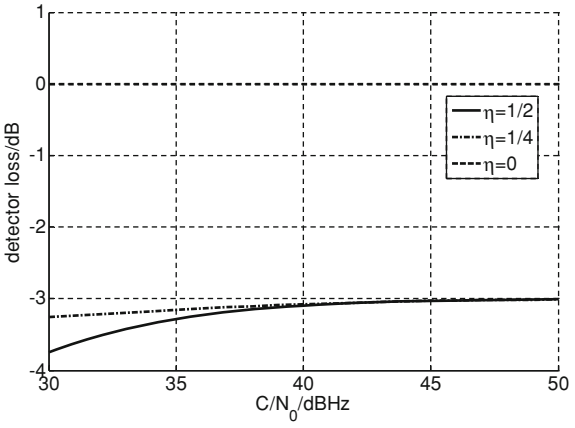


Fig. 6.3 Detector loss under different C/N_0, $T_d = T_p = 10$ ms

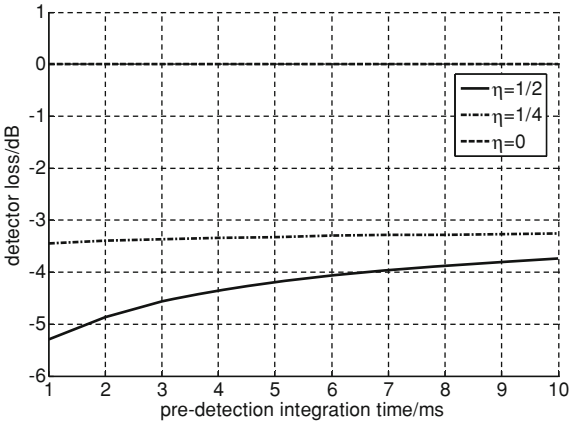


Fig. 6.4 Detector loss under different T_d and T_p, $C/N_0 = 30$ dBHz

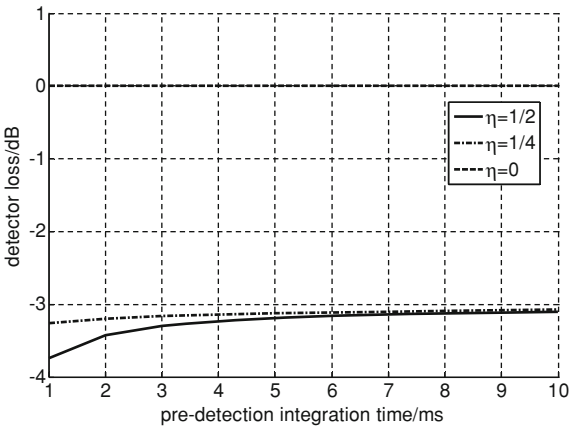


Fig. 6.5 Detector loss under different T_d and T_p, $C/N_0 = 40$ dBHz

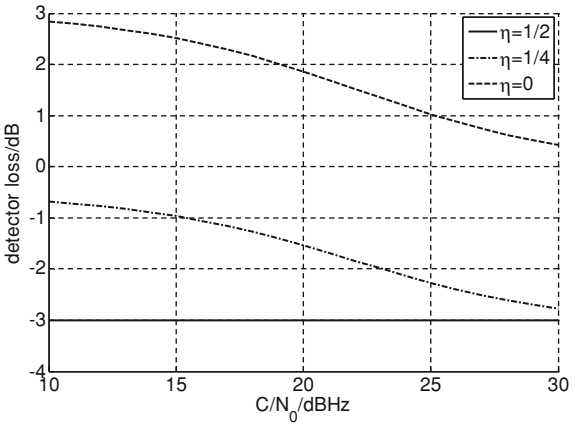


Fig. 6.6 Detector loss under different C/N_0, $T_p = 20$ ms

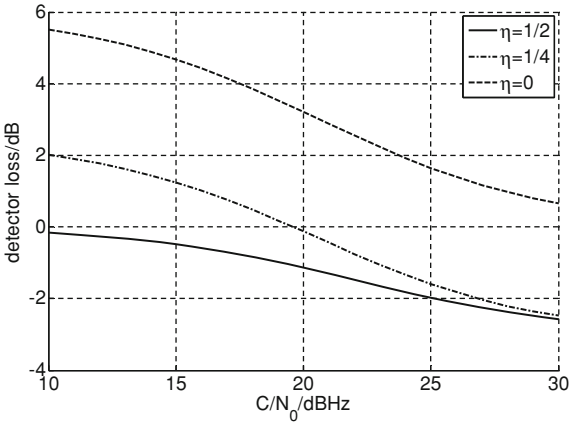


Fig. 6.7 Detector loss under different C/N_0, $T_p = 40$ ms

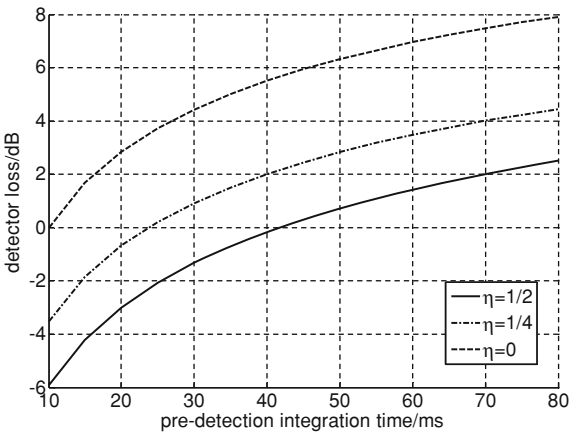


Fig. 6.8 Detector loss under different T_p, $C/N_0 = 10$ dBHz

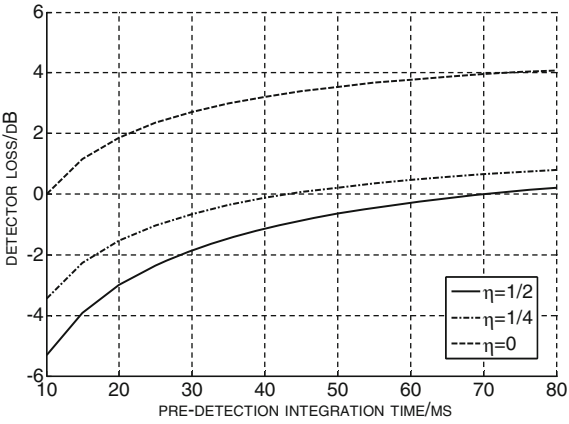


Fig. 6.9 Detector loss under different T_p, C/N_0 = 20 dBHz

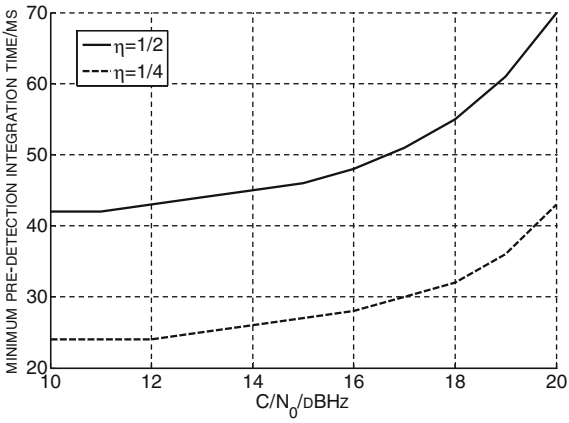


Fig. 6.10 Minimum pre-detection integration time of different C/N_0

of option 1. And only when the pre-detection time is much longer than the bit width, the detection performance for option 2 and 3 is better than that for option 1.

In order to achieve the same detection performance, the shortest pre-detection integration time of pilot channel is shown in Fig. 6.10.

It can be seen from Fig. 6.10, when the carrier-to-noise ratio is 10 dBHz, in order to achieve the same acquisition performance with option 1, the minimum pre-detection integration time is 42 ms for option 2, and that for option 3 is 24 ms. In fact, due to the Doppler frequency uncertainty, the pre-detection integration time of the pilot channel is difficult to reach 42 ms. Therefore when compared with traditional data channel, the widely used data and pilot channel with 1:1 power ratio is difficult to enhance the acquisition sensitivity under weak signal environments. When all of the signal power is allocated to the pilot channel, it can significantly improve the acquisition sensitivity under weak signal environment.

6.6 Conclusions

This paper analyzes the acquisition performance for different power allocation options of data and pilot channels. The results show that due to the Doppler frequency uncertainty, it is almost impossible to improve the acquisition sensitivity under weak signal environment when the power allocation ratio of 1:1 is adopted; while the acquisition sensitivity could be improved under the ratio of 1:3.

According to the conclusions, in order to improve availability of the BDS signals in weak signal environments, the power allocation ratio of 1:3 should be used, or all the energy is devoted to the pilot channel. Considering the dual-frequency receiver will become prevalent in the future, signal only with pilot channel is entirely feasible.

References

1. Gang X (2009) GPS receiver and the design principles. Electronic Industry Press, Beijing
2. Stansell A, Hudnut W, Keegan G (2010) GPSL1C: enhanced performance receiver design suggestion and key contributions. In: 23rd international technical meeting of the satellite division of the institute of navigation
3. Betz JW (2013) Signal structures for satellite-based navigation: past, present and future. Inside GNSS 8:34–42
4. Global Positioning System Wing (2010) Navstar GPS space segment/user segment L5 interface IS-GPS-705 revision A
5. Global Positioning System Wing (2010) Navstar GPS space segment/ user segment L1C interface IS-GPS-800 revision A
6. Global Positioning System Wing (2010) Navstar GPS space segment/ navigation user interface IS-GPS-200 revision E
7. Dafesh P (2002) Description and analysis of time-multiplexed M-code data. In: Proceedings of US institute of navigation AM conference, Albuquerque, NM, 24–26 Sept 2002, pp 598–610
8. Borio D, O'Driscoll C (2009) Coherent, non-coherent and differential coherent combining techniques for the acquisition of new composite GNSS signals. IEEE Trans Aerosp Electron Syst 45(3):1227–1240
9. Wang F (1998) Direct sequence spread spectrum signal all-digital fast acquisition. Changsha: National University of Defence Technology University, Graduate School PhD thesis

Chapter 7
Tracking and Multipath Performance Analysis of the MBOC Approximation by BOC Modulation in Band-Limited Receiver Channels

Xinhui Lin, Feng Liu, Yiqing Zhang and Lei Zhang

Abstract The MBOC modulation is the main interoperable signal which is adopted by the modernized GPS L1C and the Galileo E1 signal. It has better tracking and multipath performance with the respect to BOC modulation, but also brings the problem that should be required to deal with two different rates of sub-carrier BOC signals are multiplexed, and require a higher sampling rate and bandwidth of the front end which increased design complexity of the receiver. In this paper, the excellent performance of the MBOC signal and the relatively simple receiver tracking structure of BOC signal is considered, the tracking and multipath performance of the MBOC approximation by BOC modulation is analyzed. The impact of approximate tracking to the typical anti-multipath algorithms such as narrow correlation and Double Delta is researched based on multipath error envelope. The problem which too many zero crossings of code discriminator S-curve leads to the deviation of the loop control is analyzed. Focused on a more realistic situation of limited bandwidth, the multipath performance and tracking accuracy is analyzed. Simulation results show that the smaller front-end bandwidth, the approximation performance is closer to a complete tracking. And it can simplify the structural design of the receiver channel while the tracking performance is not a significant decrease. The result has a certain reference value for simplifying channel design of MBOC receiver and reducing the project implementation complexity.

X. Lin (✉) · F. Liu · L. Zhang
School of Information and Electronics, Beijing Institute of Technology, Beijing, People's Republic of China
e-mail: linxinhui@bit.edu.cn

F. Liu
e-mail: bit_liufeng@bit.edu.cn

L. Zhang
e-mail: aerolong@bit.edu.cn

Y. Zhang
Beijing Global Information Center of Application and Exploitation, Beijing, People's Republic of China
e-mail: yiqzhang@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_7,

Keywords MBOC · BOC · Approximation tracking · Multipath error envelope · Band-limited receiver channels

7.1 Introduction

The MBOC modulation is the interoperable signal which is adopted by GPS and Galileo, and is used for the modernization of GPS-L1C signal and Galileo-E1 signal. MBOC signal is mixed together from the BOC (1,1) and BOC (6,1) according to a certain proportion. There are two implementations which L1C uses time division multiple subcarrier TMBOC (6,1,4/33), while E1 uses composite subcarrier CBOC (6,1,1/11).

Currently, tracking methods for MBOC signal are the use of BOC approximation and using MBOC complete tracking. When using MBOC signal complete tracking, the problem of multiplexed use of two different rates BOC subcarrier signal should be deal with. The high-frequency portion of the MBOC signal require a higher sampling rate, the wider the bandwidth of the RF front end and a larger amount of system operation, which increases the complexity of the receiver design. But current most miniaturized receivers toward the civilian market often have limited resources, so the approximate tracking is more suitable to the small size receiver. Many scholars home or abroad have studied the multipath and tracking performance of MBOC and BOC signal [1]. And literature [2] also analyzed capture, the traditional narrow correlation tracking and interference performance of BOC approximation of MBOC signal. But the approximation tracking performance in a limited bandwidth receiver channel and multipath rejection performance when using Double Delta loop, such articles are relatively rare.

This paper analyzes the tracking and multipath performance of MBOC approximate by BOC signal. The performance differences based on the multipath error envelope is compared, and the approximate tracking performance is analyzed in the limited bandwidth channel which is closer to the actual situation. Simulation results show that the bandwidth of the receiver front end is smaller, the approximation performance is closer to a complete tracking, so it is more suitable for the miniaturized receivers with limited resources.

7.2 MBOC Signal and Features

7.2.1 MBOC Signal

MBOC signal is mixed together from the BOC (1,1) and BOC (6,1) according to a certain proportion. There are two kinds of implementations of MBOC modulation which are time division multiple subcarrier Time multiplexed Binary Offset

Carrier (TMBOC) and composite subcarrier Composite Binary Offset Carrier (CBOC). Galileo-E1OS uses CBOC(6,1,l/11) modulation, GPS-LIC uses TMBOC(6, l,4/33) modulation [3].

The new satellite navigation system exist two signal channels—pilot and data, MBOC (6,1,1/11) will join the pilot channel and data channels power spectral density as part of its power spectral density, the expression is as follows:

$$G_{MBOC}(f) = \frac{10}{11} G_{BOC}(1,1)(f) + \frac{1}{11} G_{BOC}(6,1)(f) \tag{7.1}$$

wherein, G_BOC (1,1) represents BOC(1,1) power spectral density, G_BOC (6,1) represents BOC(6,1) power spectral density.

GPS-L1C use TMBOC to achieve MBOC. Definition is that the data channel uses BOC(1,1), the pilot channel uses TMBOC(6,1,4/33). The sub-carriers of TMBOC(6,1,4/33) are synthesized in a manner of time division multiplexing by are BOC (1,1) and BOC (6,1), i.e., a predetermined length of the chip, the chip in this group, the position 1,5,7 and 30 is the BOC (6,1), other locations are BOC (1,1). BOC (6,1) is the proportion of 4/33, the power ratio of pilot and data channel is 1:3, the formula is expressed as:

$$G_{MBOC}\left(6,1,\frac{1}{11}\right)(f) = \frac{1}{4} G_{Data}(f) + \frac{3}{4} G_{pilot}(f) \tag{7.2}$$

wherein,

$$G_{Pilot}(f) = \frac{29}{33} G_{BOC(1,1)}(f) + \frac{4}{33} G_{BOC(6,1)}(f)$$
$$G_{Data}(f) = G_{BOC(1,1)}(f)$$

And Galileo-E1 achieves MBOC using CBOC. CBOC (6,1, l/11) subcarrier is based on the power ratio of 10:1 in the time domain based on a weighted superposition BOC (1,1) and BOC (6,1). The modulation is achieved by configuring symbols from the four different amplitude levels and is the weight of the composite amplitude. And then slip up data and pilot, the power of the pilot and data channel are the same. After the time-domain power spectrum weighted superposition, the formula is as follows:

$$\begin{aligned} G_{MBOC}\left(6,1,\frac{1}{11}\right)(f) &= \frac{1}{2} G_{Data}(f) + \frac{1}{2} G_{pilot}(f) \\ &= \frac{10}{11} G_{BOC}(1,1) + \frac{1}{11} G_{BOC}(6,1)(f) \end{aligned} \tag{7.3}$$

7.2.2 MBOC Characteristics

The power spectral density comparison curve of TMBOC, CBOC, BOC (1,1) signal is shown in Fig. 7.1. It can be seen that the MBOC spectrum is split into two symmetrical from the center portion, so that GPS-L1C and Galileo-E1

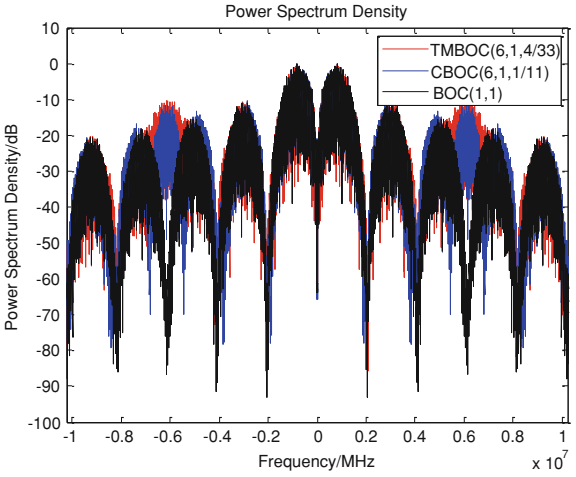


Fig. 7.1 Power spectrum density of TMBOC, CBOC, BOC. The *black*, *blue* and *red curves* are power spectrum density of TMBOC, CBOC, BOC

signals do not overlap in the frequency, which show good compatibility and interoperability. Compared with the BOC (1,1), MBOC has more high frequency components.

The comparison chart of the BOC (1,1), TMBOC (6,1,4 /33) and a CBOC (6,1, l/11) autocorrelation function is shown in Fig. 7.2. As shown, correlation peak of TMBOC and CBOC will have slight differences due to the generation of subcarriers in different ways. A narrower and sharper autocorrelation peak curve is got, which can improve the accuracy of the pseudo-code tracking, and multipath rejection performance of the algorithm will also be raised.

7.3 Multipath Performance Analysis of Approximation Tracking Based on Multipath Error Envelope

Compared with the traditional BPSK, autocorrelation function of BOC has a different so larger than the BPSK signal. Autocorrelation function of BOC class signal appears secondary peaks, which may lead to false lock, thus requiring additional consideration of the problem of fuzzy degree eliminate. The performance of multipath algorithm applied to BPSK modulation may also have changed in the BOC signal [4]. Thus this paper, analyzes effect on the classic narrow correlation and Double Delta of approximation tracking.

The multipath error envelope curve of BOC(1,1) approximation tracking, and full TMBOC(6,1,4/33), CBOC (6,1, l/11) tracking for narrow correlation which spacing width is 0.1 under an ideal unlimited bandwidth conditions is shown in Fig. 7.3a. It can be drawn from the figure: the trend of approximate tracking and full tracking is similar; overall tracking error of full tracking is less than

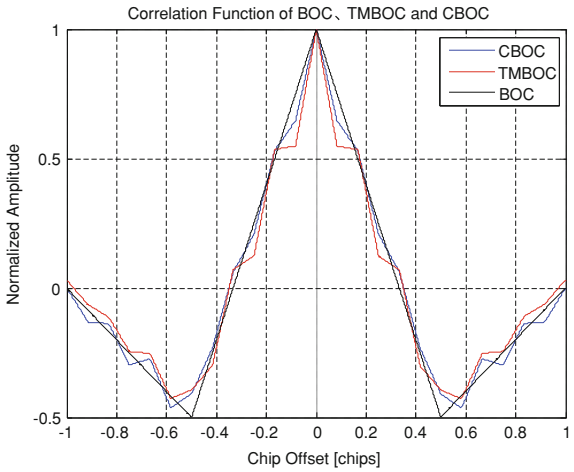


Fig. 7.2 Correlation function of TMBOC, CBOC and BOC. The *black*, *blue* and *red curves* are correlation function of TMBOC, CBOC, BOC

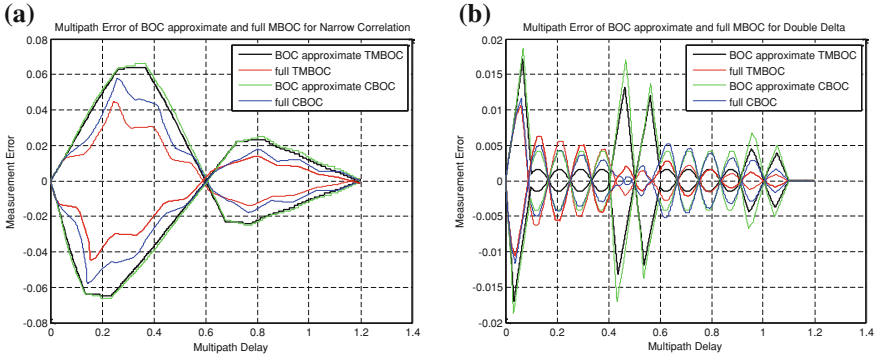


Fig. 7.3 Multipath error of BOC approximate and full MBOC **a** The *black*, *red, green* and *blue curves* are multipath error for narrow correlation. **b** The *black*, *red, green* and *blue curves* are multipath error for double delta

approximate tracking. Since the BOC signal has a sharper correlation peak, when multipath chip in the long delay, anti-multipath effects of approximate tracking is better than traditional BPSK signal. Multipath performance of TMBOC and CBOC are similar, while the correlation spacing width is smaller, multipath rejection performance of TMBOC is better than CBOC, because the high frequency component makes TMBOC spectrum sharper peaks.

The multipath error envelope curve of BOC(1,1) approximation tracking, and full TMBOC(6,1,4/33), CBOC (6,1, l/11) tracking for Double Delta is shown in Fig. 7.3b. Overall tracking error of full tracking is less than approximate tracking. Compared with Fig. 7.3a, Double Delta Technology multipath error magnitude is smaller, and therefore performance of multipath rejection is

superior than narrow correlation. Especially in the 0.2–0.4 and 0.6–0.8 chip multipath delay, multipath error envelope of Double Delta is significantly less than the narrow correlation.

The discrimination curve of BOC(1,1) approximation tracking, and full TMBOC(6,1,4/33), CBOC (6,1, l/11) tracking for narrow correlation which spacing width is 0.1 is shown in Fig. 7.4a. As can be seen from the figure, whether it is the use of BOC(1,1) approximation tracking or using full MBOC tracking, discrimination function has a number of zero-crossings, if multipath delay is large, may cause false lock, so a bump-jumping auxiliary tracking is necessary. From the graph can be seen, the pulling range of discrimination function is similar for both the full tracking and approximately tracking. The narrow correlation is about [−0.1, 0.1]. But the slope of traction the range is greater, which is affected by the high frequency component, the autocorrelation function of the MBOC is sharper than the BOC (1,1).

The discrimination curve of BOC(1,1) approximation tracking, and full TMBOC(6,1,4/33), CBOC (6,1, l/11) tracking for Double Delta is shown in Fig. 7.4b. The autocorrelation function of BOC and MBOC signal has secondary peaks, when using Double-Delta technology, should ensure that the interval correlation between the outermost is smaller than the width of the autocorrelation function peak. Compared to the narrow correlation technologies, BOC (1,1) signal using Double-Delta technology discrimination function has a greater number of zero crossing, so the use of assisted verification techniques to avoid false lock is necessary. Pulling range of the Double Delta technique is smaller, and therefore the tracking loop stability should be considered.

7.4 The Bandwidth Limited Effect on the Multipath Performance of Approximate Tracking

Bandwidth limited effect is the important factor of tracking accuracy and multipath suppression capability in satellite navigation system. Spread spectrum modulation of satellite navigation signal is infinitely wide based on spectrum theory, but the bandwidth of the satellite transmitter and receiver reception bandwidth is limited, with the result that the effect of limited bandwidth. Limited bandwidth will smooth the peak top of correlation function, thereby weakening the effect of suppressing multipath loop [5]. The correlation peak of BOC signal in different bandwidth of the front end is shown in Fig. 7.5

Figure 7.6a is a graph multipath error envelope of BOC (1,1) approximate tracking the TMBOC (6,1,4 /33) and the full tracking for narrow correlation under limited bandwidth conditions. Here are the front-end bandwidth using 4 and 14 M, and will be a comparison between the two. When the receiver front-end bandwidth is 4 M, the multipath error envelope of approximation tracking and full tracking is almost identical. The current side bandwidth expanded to 14 M later, the

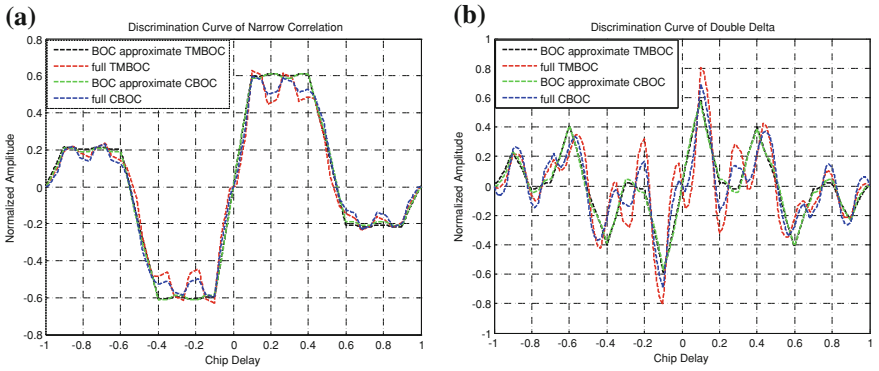


Fig. 7.4 Discrimination curve of BOC approximate and full MBOC **a** The *black*, *red*, *green* and *blue curves* are discrimination curve for narrow correlation. **b** The *black*, *red, green* and *blue curves* are discrimination curve for double delta

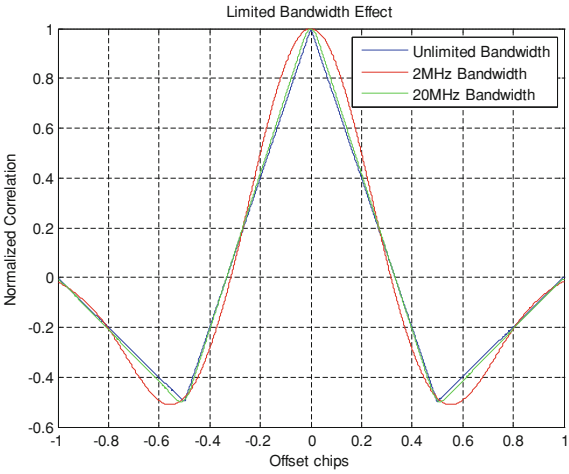


Fig. 7.5 Limit bandwidth effect. The *blue*, *red* and *green curves* are correlation function in different bandwidth of the front end

advantage of sharper correlation peak of full MBOC tracking is manifested, enhance the inhibitory effect of multipath.

Figure 7.6b is a graph multipath error envelope of BOC (1,1) approximate tracking the TMBOC (6,1,4/33) and the full tracking for Double Delta under limited bandwidth conditions. As can be seen, Double Delta technology has better anti-multipath performance than the narrow correlation technologies. And in nearly infinite bandwidth 14 M-bandwidth conditions, the full track will have a greater promotion to multipath performance. After a 4 M front-end filter, multi-path rejection performance of approximate tracking and full tracking is closer, because the sharp correlation peak by a smooth.

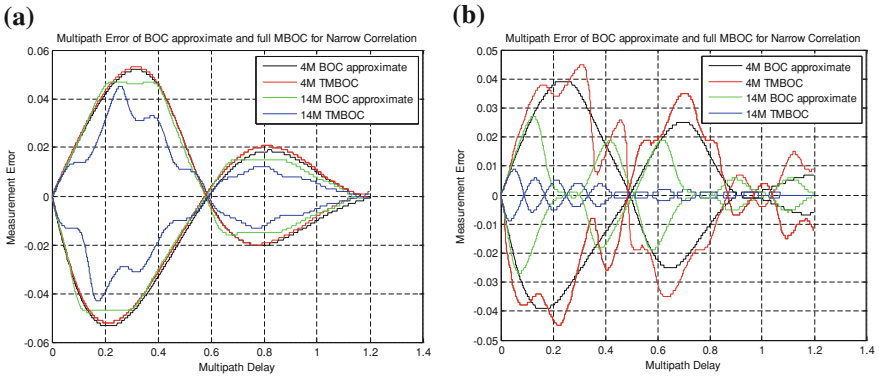


Fig. 7.6 Multipath error of BOC approximate and full MBOC in bandwidth channel **a** The *black*, *red*, *green* and *blue curves* are multipath error for narrow correlation **b** The *black*, *red, green* and *blue curves* are multipath error for double delta

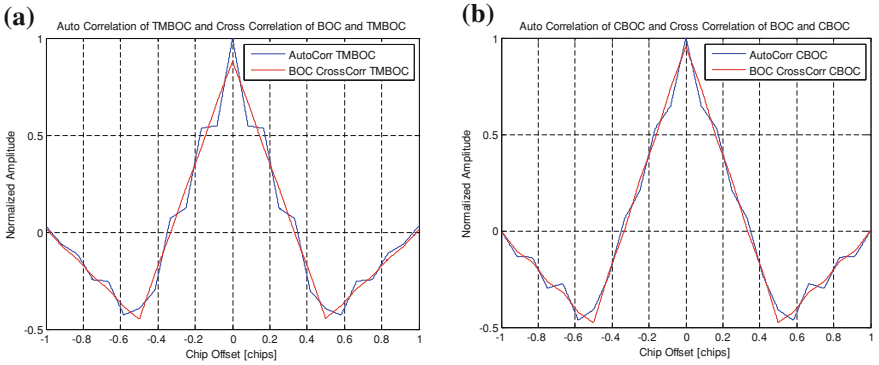


Fig. 7.7 Auto correlation of MBOC and cross correlation of BOC and MBOC **a** The *blue* and *red curves* are auto correlation of TMBOC and cross correlation of BOC and TMBOC. **b** The *blue* and *red curves* are auto correlation of CBOC and cross correlation of BOC and CBOC

7.5 Approximation Tracking Accuracy Simulation Under the Bandwidth Limited Condition

Autocorrelation function of TMBOC (6,1,4 /33) signal and BOC (1,1) approximate correlation function are shown in Fig. 7.7a. Red curve is the correlation using BOC (1,1) approximate, the problem with power down can be found, can be calculated by the loss of signal is 1.12 dB.

Similarly, promote it to the CBOC (6,1,1/11) signal, if used BOC (1,1) approximate CBOC signal, the correlation function is shown in Fig. 7.7b. By calculating, its SNR loss is only 0.4 dB.

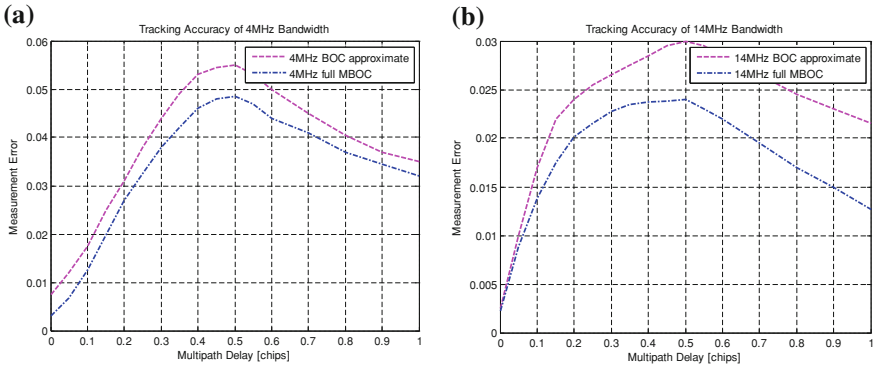


Fig. 7.8 Tracking accuracy of the BOC approximate and full MBOC **a** The *blue* and *red curves* are tracking accuracy of 4 MHz Bandwidth **b** The *blue* and *red curves* are tracking accuracy of 14 MHz bandwidth

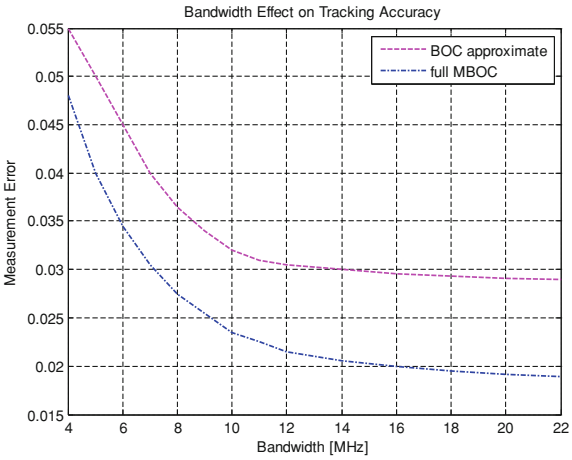


Fig. 7.9 Bandwidth effect of tracking accuracy. The *blue* and *red curves* are bandwidth effect of tracking accuracy of the BOC approximate and full MBOC

In order to analyze the limited bandwidth effect on the tracking precision of approximation MBOC, the effective range of the multipath is analyzed and full tracking and approximation tracking is compared. If only the way in the case of multipath, and the correlation distance $d = 0.1$, in different bandwidth full tracking and approximate tracking accuracy is shown in Fig. 7.8a, b. From the figure can be obtained: in the case of wide signal bandwidth, the multipath performance is similar to the unlimited bandwidth, MBOC full advantage can be tracked reflects about 12 % higher than the approximate tracking accuracy. If the signal bandwidth is narrow, the effects on approximate tracking and full tracking is very closer, the smaller the bandwidth, the proximity of the more obvious.

Through Monte Carlo simulation, tracking accuracy of the BOC approximate and full MBOC tracking in different bandwidth in the relevant spacing 0.1 is shown in Fig. 7.9. As can be seen in the context of the smaller receiver front-end bandwidth, the effect of approximate tracking and full tracking is closer and larger bandwidth of the receiver front end can reflect MBOC full tracking advantage.

7.6 Conclusion

This paper analyzes the MBOC signal characteristics and tracking and multipath performance of the BOC approximate tracking MBOC signal. The results show that the loss caused by the approximate trace within the acceptable range, and the multipath performance of the BOC approximate tracking is similar to full MBOC tracking, and in the context of the smaller receiver front-end bandwidth, the effect of full tracking and approximate tracking is closer, so this method is suitable for the most miniaturized receiver.

References

1. Insigler M, Eissfeller B (2003) Comparison of multipath mitigation techniques with consideration of future signal structures. Proc ION GPS/GNSS 2003:2584–2592
2. Macchi F, Petovello MG, Lachapelle G (2008) Consequences of MBOC approximation by BOC modulation. In: ION GNSS 21st international technical meeting of the satellite division, Savannah, GA, 16–19 Sept 2008
3. Hein GW, Wallner S (2007) The MBOC modulation: a final touch for the Galileo frequency and signal plan. In: ION GNSS, USA, pp 1515–1529
4. Irsigler M (2005) Criteria for GNSS multipath performance assessment. In: Proceedings of the 18th international technical meeting of the satellite division of the institute of navigation, pp 166–177
5. Kaplan ED, Hegarty CJ (2006) Understanding GPS principle and applications, 2nd edn. Artech House, Boston

Chapter 8
A Proposed Selection and Allocation of Spread-Spectrum Code for Compass B1C Signal

Jue Wang, Yao Wang and Zaixiu Yang

Abstract In the future, BD Satellites Navigation System (Compass) is going to use B1C signal for compatibility and interoperation with such GNSS as GPS and Galileo. As an important step in B1C signal design, the design of spread-spectrum (SS) code involves construction, selection and allocation of the codes. Based on an analysis of the L1-band pseudo-code characteristics of GPS and Galileo, the researcher optimally selected 100 SS codes from the optimal codes set of 10,243 bits-long Weil codes using greedy algorithm. With the autocorrelation up to −31.57 dB and the cross correlation up to −27.44 dB, all these SS codes are superior to GPS L1C and Galileo E1OS pseudo codes in performance. Besides, a pseudo-code allocation method developed under the minimum ranging error principle was adopted for rational allocation of the data and pilot branches for the proposed B1C pseudo codes. The simulation and comparative analysis showed that the findings were of great significance for the signal design of Compass in the future.

Keywords Satellite navigation system · B1C · Weil code · PRN · Autocorrelation · Cross correlation

J. Wang (✉) · Y. Wang · Z. Yang
The 54th Institute of China Electronic Technology Corporation,
Shijiazhuang Hebei 050081, China
e-mail: wangjue_cetc@139.com

Y. Wang
e-mail: wy_hrbeu@163.com

J. Wang
Hebei Satellite Navigation Technology and Equipment Technlogy Center,
Shijiazhuang Hebei 050081, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_8,

8.1 Introduction

BeiDou Satellites Navigation System (Compass) has been providing positioning, navigation and time services in the Asia-Pacific region. It is planned to possess global service capacity by 2020 in the future. As the core and the carrier to fulfill the fundamental functions of a satellite navigation system, navigation signal system plays a decisive role in the critical performances of navigation system such as positioning/velocity measurement/time service precision, compatibility, interoperability, confidentiality and anti-jamming capability. In July 2009, CNAGA announced the latest scheme of Compass signal [1], in which B1C signal is the key to global promotion of Compass in the future, especially the design of SS codes. The design considerations mainly include (1) construction method (i.e., what technique or principle to generate the code blocks) (2) optimal selection method (i.e. how to select the code blocks of good performance from the candidates) (3) allocation method (i.e., what method to allocate the SS code rationally).

L1 frequency is one of important public service frequencies of GNSS. Both GPS and Galileo adopt new SS codes at this frequency point and achieve good performance. B1C signal represents a new generation of civils signal in China, and the design of its SS code is faced with huge challenges. The researchers selected optimally 100 SS codes of 10,243 bit length with greedy algorithm under the selecting rules of balance and correlation. The data and pilot branches were rationally allocated with the pseudo-code allocation method developed under minimum ranging-error rule. The results of simulation and analysis show that the innovative SS code proposed in this paper is one of the optimal schemes for B1C signal of Compass.

8.2 Characteristic Analyses of Overseas SS Codes

8.2.1 SS Codes of GPS L1C Signal

L1C SS code uses a new design approach and code selection criteria. It adopts Weil code [2] based on number theory instead of the shift register-based design scheme, which has better correlated characteristics under the conditions of the same code length. L1C Weil code takes 10,223 as its prime number and the inadequate length is achieved by means of filling. The generation method is specific as follows:

(1) Use the prime number of 10,223 as generation basic number to generate 5,111 Weil codes of 10,223 bits length.
(2) Complement the 10,223 bit-long Weil code by means of insertion in the mode of 1001011 to generate 52,249,753 code blocks by different insertion positions.

(3) Select optimally 739 code blocks under the threshold rules of even cross-correlation −31 dB/odd autocorrelation −28 dB/even cross-correlation −27.2 dB and odd cross-correlation −26.2 dB.
(4) Further optimize the selection for 420 code blocks and finally selected optimally the even cross-correlation −28 dB/odd cross-correlation −26.5 dB.

8.2.2 SS Codes of Galileo E1 OS Signal

Another new approach was taken in the design of SS code for E1 OS signal, which is known as storage random code [4]. The code generation method is not closely associated with mathematical theory anymore. Its design concept is to prestore the pseudo random code at the transmit and receive ends, which does not require recalculation on site. The key issues to the design of storage random code were code selection criterion and computation complexity. Galileo signal designers proposed the concept of autocorrelation sidelobe and zero correlation interval on the basis of the original correlation constraint, and used genetic algorithm in computation.

8.3 Optimal Selection of SS Code for B1C Signal Based on Greedy Algorithm

8.3.1 Selection of Code Block Sets

The optimal selection of SS codes was purposed to pick up a smaller code blocks set with better performance from a larger one according to specific performance measures. The optimal selection method consists of candidate code block sets, definition of performance measures and optimal selection algorithm. The code block sets are preferably close to the lower limit of Welch. Many options were available including Gold, Kasami, Bent and other codes. However, under the constraint condition of 10,230 code length, Weil code was the one the easiest to find good correlation performance. Because some other codes require large numbers of patches or truncations to achieve the length of 10,230 bit, though their performances might seemingly be better than Weil code. In this process, the correlation performance usually was lost considerably; as a result, it was hard to pick up good codes. For this reason, Weil code was finally selected to serve as the optimal code block set of B1C signal.

There are two prime numbers close to 10,230 in Weil code: 10,223 and 10,243 respectively. The latter is our choice for the former has been used by L1C signal, as shown in Fig. 8.1. Its construction method and steps are as shown below:

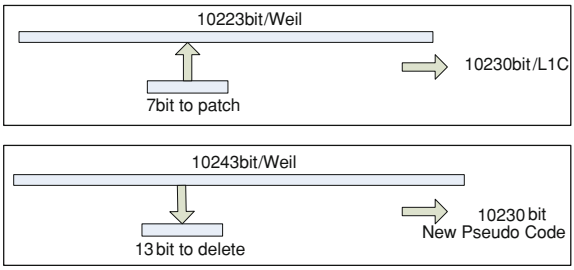


Fig. 8.1 Comparison of proposed SS code versus GPS L1C SS code

(1) Generate a Legendre sequence L of 10,243 bit length, of which the definition is the same as the reference document [5], and generate the prime number p = 10,243.
(2) Generate a Weil code blocks set W of 10,243 bit length, in which the code block $W(i,k) = L(k) \oplus L(i+k)$, where i represents the index number of Weil code $1 \le i \le 5{,}121$ and k represents the element suffix. Give p modular computation when $k = 1 \ldots 10{,}243$.
(3) Generate a periodic extension sequence set WW of $10{,}243 \times 2$ length, where the code block $WW(i) = [W(i)\ W(i)], 1 \le i \le 5,121$.
(4) Generate a SS code blocks set C of 10,230 bit length, in which $C(i, j)$ represents the extraction starts from the jth position of $WW(i)$ and the extracting length is 10,230 bits, $1 \le j \le 10{,}243$.

With the above method, traverse all the 5,121 shift differences and 10,243 extracting start points to achieve 52,454,403 combinations, i.e., the same number of candidate code blocks in Set C. The optimal selection of code blocks is to select 100 code blocks from the candidates to serve as the SS codes of B1C signal.

8.3.2 Performance Measure of SS Codes

Performance measure is the reference and criteria for selecting and evaluating SS codes. A Number of considerations must be given in the design of performance measure such as service type, signal environment, transmitting platform, receiver platform, etc. The designers of GNSS signal structure provide a number of performance measure requirements including randomness, correlation and anti-jamming performance.

For SS code of length $L = 10,230$ $\mathbf{a} = (a_0, \ldots, a_{L-1})$, where $a_i \in \{-1, +1\}$, the performance measures are defined as follows:

Balance measure:

$$bal(\mathbf{a}) = |\sum_{i=0}^{L-1} a_i|$$

Autocorrelation measure:

$$EvenACF(a;\tau) = \sum_{i=0}^{L-1} a_i a_{i+\tau}$$

$$OddACF(\mathbf{a};\tau) = \sum_{i=0}^{L-\tau-1} a_i a_{i+\tau} - \sum_{i=L-\tau}^{L-1} a_i a_{i+\tau}$$

Cross-correlation measure:

$$EvenCCF(\mathbf{a},\mathbf{b};\tau) = \sum_{i=0}^{L-1} a_i b_{i+\tau}$$

$$OddCCF(\mathbf{a},\mathbf{b};\tau) = \sum_{i=0}^{L-\tau-1} a_i b_{i+\tau} - \sum_{i=L-\tau}^{L-1} a_i b_{i+\tau}$$

Maximum line-spectrum measure:

$$MSL(\mathbf{a}) = max(FFT(EvenACF(\mathbf{a}))).$$

Based on a comprehensive analysis of the above-mentioned performance measures, it is acceptable to classify them into two categories: individual performance measures and combined performance measures. The individual performance measures are only related to one code block in the SS code set, e.g., balance, autocorrelation function (ACF), etc. The combined performance measures are related to the entire code blocks set, typically e.g., cross-correlation function (CCF). In the optimal selection process of SS codes, the code block selections made on the basis of individual performance measures are classified as linear complexity, i.e., only N orders of computation are required in selection of N code blocks. Whereas, the code block selections made on the basis of combined performance measures are far more complicated than linear complexity, e.g., it is required to traverse C_N^M combinations as per Stirling Formula to pick up M code blocks of optimal cross correlation from N candidates.

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

It is clear that the above computation is of exponential complexity.

In summary of the construction and performance measure of B1C Weil codes, the design process of B1C SS code is composed of the following steps:

(1) Generated the candidate code block set and generated a SS code block set of 10,230-bit length with the prime number of 10,243 as structural parameter.
(2) An optimal selection was made as per individual performance measures, and the threshold value of the individual performance measures were defined. Deleted the code blocks that failed the conditions from the candidate code blocks set to get N code blocks.

(3) The optimal selection was made as per combined performance measures. The code blocks were divided into several groups and each group contained M code blocks for evaluating as per the combined performance measures. The optimal M code blocks were picked up.

In the above three steps, the computation complexity at Step (3) was more complicated than polynomial complexity and up to NP, i.e., it was impossible to get the result within a finite duration.

8.3.3 Cross Correlation Optimization Based on Greedy Algorithm

In the optimal selection process of SS codes, it is required to traverse C_N^M combinations to pick up M code blocks of optimal cross correlation from N candidates. It is a problem more complicated than polynomial complexity, though the computation is finite. Only the problem of very small scale is solvable within a finite duration.

Greedy algorithm is an algorithm effective for solution of NP problem. Its fundamental concept is to divide the computation process into several steps and make the objective function (performance measure) to the optimum at each step when downsizing the question. After several rounds of partial optimum computations, it is possible to get a solution with relatively optimal performance [6]. The optimization procedure of SS codes with greedy algorithm is as follows:

(1) Initialize the computation. Take N candidate code blocks as input condition and take the minimization of the maximum sidelobe as the objective function.
(2) Delete unwanted code blocks. Delete 1 code block from N code blocks under such a rule that the objective function of the rest N−1 code blocks can get the extremum value. Thus, a size of N−1 candidate code blocks set is acquired.
(3) Repeat Step (2) until the number of the remainder code blocks is M. Output the remainder code blocks as the final selection result.

8.3.4 Performance Evaluation

Using the above optimal selection methods of SS codes for Compass B1C signal, 100 SS codes of 10,230 bit length (known as CETC100 Scheme) was picked up after going through construction of Weil codes, optimal selection as per individual performance measures and optimal selection as per combined performance measures. A comparison between CETC100 Scheme and GPS L1C Scheme in terms of the key performance measures was made, as shown in (Table 8.1).

Table 8.1 CETC100 codes versus GPS L1C codes

Metrics	GPS L1C	GPS L1C Ext	CETC 100
Number of codes	126 (in dB)	420 (in dB)	100 (in dB)
Even ACF	−31.07	−31.07	−31.57
Odd ACF	−28.02	−28.02	−31.57
Even CCF	−27.21	−27.21	−27.44
Odd CCF	−26.21	−26.21	−27.40
MSL	−31.83	−30.85	−32.07

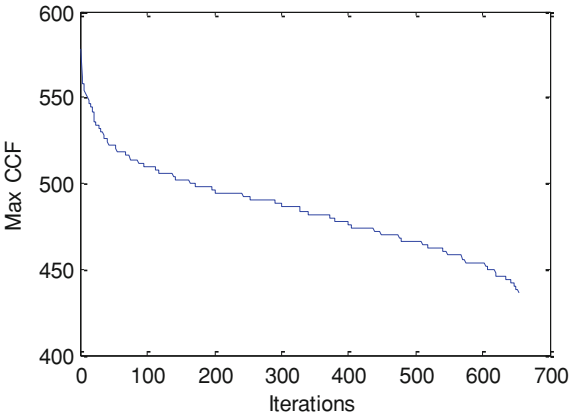


Fig. 8.2 Max CCF versus Iterations

In the design of SS codes for Compass B1C signal, the optimization of cross correlation was the most complex part, which however was resolved with greedy algorithm as linear complexity. Specifically, set N = 754 and M = 100 as input parameter and after repeating N−M = 654 times of iterative deletions, the sub-optimal solution of the cross correlation was worked out. The process of optimal selection took 38 s totally. The cross-correlation function performance changed along with the iteration times, as shown in Fig. 8.2.

8.4 Pseudo Code Allocation Under the Principle of Minimum Ranging Error

8.4.1 Allocation Method

When tracking the data or pilot channel of B1C signal separately, the cross-correlation interference inside the channel would cause code tracking error. It is possible to reduce the cross-correlation interference inside the channel by means of rational allocation of SS codes. Since the pilot signal is introduced in mainly for the purpose of improving the signal tracking performance, the evaluation criterion

proposed here for optimization is to assure minimum mean code tracking error when tracking pilot signal separately.

Assuming that it is required to complete SS code allocation for N data/pilot channels (i.e., there are a total of 2 N SS code sequences), the data and pilot SS code families are defined below respectively.

$$\mathbf{a}^{\mathrm{d}} = [\, a_1 \quad a_2 \quad \cdots \quad a_{2N} \,]$$
$$\mathbf{a}^{\mathrm{p}} = (\mathbf{a}^{\mathrm{d}})^{\mathrm{T}}$$

where, a_k (k=1, 2, …, 2 N) represents the known SS code sequences. It must be noted that $\mathbf{a}^{\mathrm{d}}$ and $\mathbf{a}^{\mathrm{p}}$ are defined only temporarily for the follow-up analysis. To know the pairing relationship between optimized data and pilot SS codes, it is required to figure out the S-curve offsets of all the possible SS code pairs. The error matrix is defined below:

$$\mathbf{B}_{2N\times 2N} = (b_{ij})_{2N\times 2N} = \mathrm{bias}(\mathbf{a}^{\mathrm{d}}, \mathbf{a}^{\mathrm{p}})$$
$$= \begin{bmatrix} \mathrm{bias}(a_1, a_1) & \mathrm{bias}(a_1, a_2) & \cdots & \mathrm{bias}(a_1, a_{2N}) \\ \mathrm{bias}(a_2, a_1) & \mathrm{bias}(a_2, a_2) & \cdots & \mathrm{bias}(a_2, a_{2N}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{bias}(a_{2N}, a_1) & \mathrm{bias}(a_{2N}, a_2) & \cdots & \mathrm{bias}(a_{2N}, a_{2N}) \end{bmatrix}$$

where (a_i, a_j) (i, j = 1, 2,…, 2 N) represents the possible SS code pair, i.e., a_i represents data channel SS code and a_j represents pilot channel SS code; bias (a_i, a_j) represents the absolute value of S-curve offset resulting from cross-correlation interference inside the channel when tracking pilot signal separately. One SS code sequence can only be allocated to data or pilot channel; therefore, the principal diagonal element of B is virtually unachievable.

It is purposed to select the optimal pairing mode of data/pilot so as to assure minimum code-tracking mean error, which can be converted into minimizing cost function 错误！未找到引用源

$$z = \sum_{i=1}^{2N} \sum_{\substack{j=1 \\ i\neq j}}^{2N} v_{ij} \cdot \mathrm{bias}(a_i^{\mathrm{d}}, a_j^{\mathrm{p}})$$

where,

$$v_{ij} = \begin{cases} 1, & \text{选择 } (a_i^{\mathrm{d}}, a_j^{\mathrm{p}}) \\ 0, & \text{未选择}(a_i^{\mathrm{d}}, a_j^{\mathrm{p}}) \end{cases}.$$

The condition restricting the selection of pairing relationship is as follows:

$$\sum_{i=1}^{2N} \left(v_{ij} + v_{ji}\right) = 1,\ j = 1,\ 2,\ \cdots,\ 2N.$$

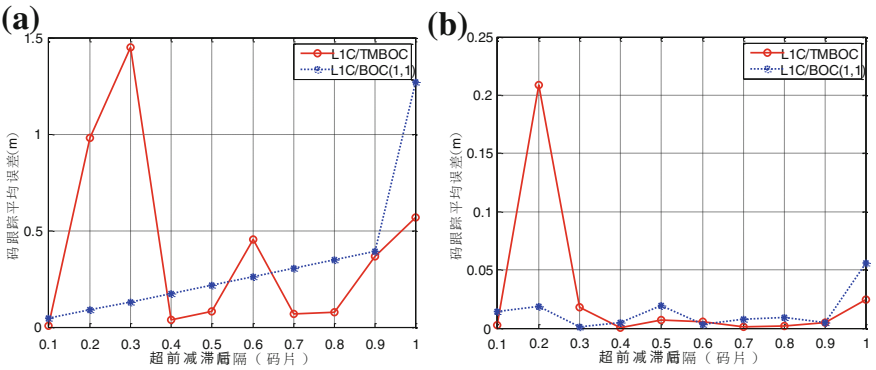


Fig. 8.3 Mean code tracking errors of GPS L1C signal, **a** before optimization, **b** after optimization

In other words, there is only one element in the *i*th row or the *i*th column that is possible to be selected. In simple words, this optimization is virtually to select N elements at B so as to assure the sum of their absolute values is the minimum, though the positions of these N elements are stringently restricted.

Since the value of N is large, it is inadvisable to make the calculation with enumeration method, for the computational complexity would be excessive. In fact, the pairing optimization may be regarded as a classical problem in operations research, i.e., assignment problem. Therefore, this pairing optimization problem may be solved with Hungary algorithm [8, 9] For the solution process of this problem, refer to the reference documents [8, 9].

8.4.2 Allocation Results and Analyses

The SS codes of GPS L1C signal were reallocated with the above-mentioned method to verify its validity. Before and after the optimization, the code-tracking mean errors changed at various lead-lag intervals, as shown in Fig. 8.3.

It can be seen that the ranging code errors resulting from the cross-correlation of GPS L1C SS codes reduced noticeably after optimization. It must be noted that the code-tracking mean error is considerably susceptible to correlation intervals, which is mainly caused by the different modulation modes of data and pilot. After this allocation method was applied to the proposed SS codes of B1C signal, the allocation scheme of SS sequences was acquired as shown in the table below. Where, the PRN line shows the SVN of Compass, the Data line shows the SS code number assigned to data channel and the Pilot line shows the SS code number assigned to pilot channel after optimization.

Figure 8.4 shows the code-tracking mean errors when tracking the pilot or data channels separately.

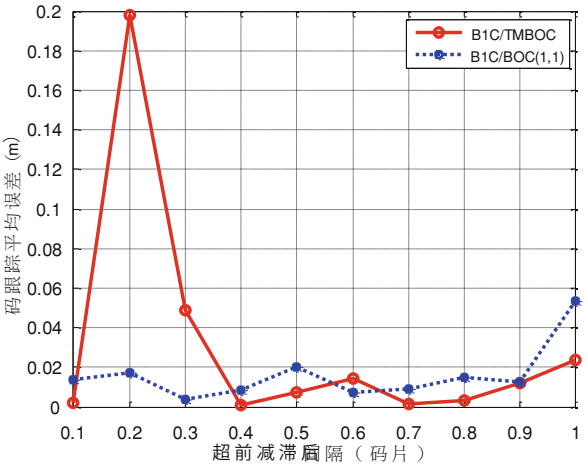


Fig. 8.4 Mean code tracking errors (B1C/TMBOC and B1C/BOC (1,1))

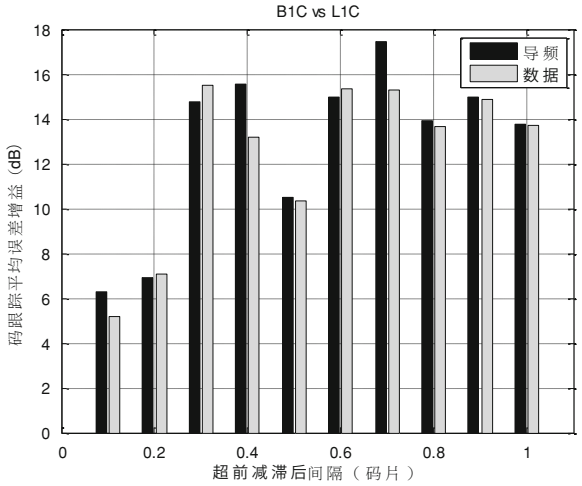


Fig. 8.5 Code tracking error gain: B1C versus L1C

Figure 8.5 shows the gains of mean code-tracking errors when tracking the pilot or data channels separately relative to L1C. It is clear that the effect of cross-correlation interference inside the signal channel to code tracking accuracy is noticeably less than that to L1C.

8.5 Conclusion

In a satellite navigation system, all the messages are modulated by means of SS codes. The well-behaved SS codes may improve the system positioning performance, antijamming capability as well as compatibility and interoperability with

other systems [10]. This paper proposed a SS-code design approach for the B1C signal of the future-oriented Compass. Detailed analyses were given to the construction method, optimal selection method and allocation method of the concerned code blocks in addition to a comparative analysis to the SS codes of GPS L1C signal. The analysis results showed that the SS codes proposed here were superior to those of GPS L1C signal in terms of performance, which may be considered as a candidate proposal for the B1C signal of the Compass.

References

1. CNAGA (2009) COMPASS view on compatibility and interoperability [EB/OL]. http://www.unoosa.org/pdf/icg/2009/workgroupinterop/04.pdf
2. Rushanan JJ (2006) Weil sequences: a family of binary sequences with good correlation properties [C]. ISIT 2006, Seattle, pp 1648–1652
3. GPS Joint Program Office (2007) GPS interference specification. IS-GPS-800 [S]. CES, EI Segundo
4. Wallner S, Rodriguez JAA, Hein GW. Galileo E1 OS and GPS L1C pseudo random noise codes [C]
5. Shusen T, Bing Z, Shengtao G (2011) Research on COMPASS navigation signals of China [J]. Chinese Space Sci Technol 4:7–14
6. Zou H (2012) Ex nihilo: algorithmica logos, 2nd edn [M]. China Machine Press, Beijing, pp 71–95
7. Hillier FS, Lieberman GJ (2007) Introduction to operations research, 8th edn [M]. Publishing House of Tsinghua University, Beijing, pp 338–345
8. Winston WL (2006) Operations research: applications and algorithms, 4th edn [M]. Publishing House of Tsinghua University, Beijing, pp 406–452
9. Guan Z (2008) Algorithm and implementation of assignment problem [D]. East China Normal University, Shanghai
10. Kaplan ED, Hegarty CJ (2007) Understanding GPS: principle and application, 2nd edn [M]. Publishing House of Electronics Industry, Beijing, pp 125–130

Chapter 9
The Performance of Interferometric Ranging Systems with Hopping Frequencies in Benign Environment and Under Partial-Band Jamming

Yue Zhang, Wangdong Qi, Weiheng Dai, Jing Lv and Guangxia Li

Abstract Interferometric Ranging with Hopping Frequencies (IRHF) is a novel ranging technique with application in anti-jamming navigation and localization systems. This paper explores the performance of IRHF systems in benign environments and under partial-band jamming (PBJ) conditions. The outlier probability is used to evaluate the ranging accuracy of IRHF systems. Employing an approximation tool called the Union Bound (UB), we derive a formula of the outlier probability of the maximum likelihood estimator (MLE) with respect to a few system parameters. Further analysis reveals that the worst-case partial-band jamming for IRHF is full-band jamming. Computer simulations show that the theoretical result is good enough to predict the performance of IRHF under different conditions.

Keywords Interferometric ranging with hopping frequencies (IRHF) · Partial-band jamming (PBJ) · Maximum likelihood estimator (MLE) · Union bound (UB)

9.1 Introduction

It is well-known that the Global Navigation Satellite Systems are vulnerable to jamming. In this paper, we put forward a novel ranging technique called Interferometric Ranging with Hopping Frequencies (IRHF) [1, 2] to provide a more

Y. Zhang · W. Qi (✉)
Center of Network Engineering, PLA University of Science and Technology, Nanjing 210007, China
e-mail: wangdongqi@gmail.com

Y. Zhang
e-mail: zhyemf@gmail.com

W. Dai · J. Lv · G. Li
College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_9,

jamming-resilient ranging scheme for navigation systems. Combining Frequency Hopping Spread Spectrum (FHSS) and a recently appeared technique called Radio Interferometric Ranging [3, 4], IRHF provides a new approach to creating anti-jamming navigation and localization systems.

To understand the performance of IRHF, we explore the outlier probability of the Maximum Likelihood Estimator (MLE) both in environments without intentional interference and under Partial-Band Jamming (PBJ), a major jamming threat faced by hopping-frequency systems [5].

It is straight forward to describe the ranging accuracy of a phase-based ranging system for a given set of frequencies in terms of Mean Square Error (MSE) against measurement errors of phases. IRHF, however, cannot be dealt with in this way because the set of frequencies employed in a measurement are randomly chosen from the total available frequencies. The probability of outlier is introduced in this paper to describe the ranging accuracy of an IRHF system. According to the method of Union Bound (UB) [6] based on the concept of 'Average Ambiguity Function (AAF)' [7], we obtain a closed-form expression of outlier probability for the MLE of IRHF in the whole signal-to-noise ratio (SNR) region, under both benign environments and partial-band jamming. Based on the results above, the worst-case jamming is proved to be full-band jamming.

The rest of this paper is organized as follows. The system model of IRHF is given in Sect. 9.2. The ranging performances of IRHF without jamming and with partial-band jamming are presented in Sects. 9.3 and 9.4 respectively. Section 9.5 gives the simulation results. We discuss the results and make a few conclusions in Sect. 9.6.

9.2 System Model and Estimator of IRHF

9.2.1 System Model

IRHF requires four nodes A, B, C and D to compose a ranging unit. A and B transmit sine waves at two close frequencies simultaneously, and two other nodes C and D measure the phases of the received signals. We denote the signal received by C or D is

$$s_{XY}(t) = a_{XY}\cos(2\pi f_X(t - t_X) + \varphi_{XY}) + w_{XY} \tag{9.1}$$

where X is either A or B, and Y is either C or D. The amplitude, carrier frequency and the transmitting time is denoted by a_{XY}, f_X and t_X respectively. The relationship between the phase shifts φ_{XY} and the distance d_{XY} from X to Y is $\varphi_{XY} = \frac{2\pi f_X}{c} d_{XY}$. w_{XY} is independent identically distributed (i.i.d) Gaussian White Noise and $w_{xy} \sim \mathbb{N}(0, N_0/2)$. We denote the SNR as a_{XY}^2/N_0.

In order to eliminate the effect caused by time synchronization error, a combination of the observed phases $\varphi = (\varphi_{AC} - \varphi_{BC}) - (\varphi_{AD} - \varphi_{BD})$ is used to resolve distance. The phase combination φ is a function of distances between the four nodes

A, B, C and D. After we get a set of the combined phases from a series of measurements with multiple carrier frequencies, we have the observation that [1, 8]

$$\varphi_i \approx 2\pi\frac{f_i}{c}d_0 + \theta + n_i \,(\mathrm{mod}\, 2\pi),\ i = 1, \ldots, M \tag{9.2}$$

where c is the speed of signal propagation, $\mathbf{f} = \{f_1, \ldots, f_M\}$ denotes the set of M measurement frequencies, and $\mathbf{\Phi} = \{\varphi_1, \ldots, \varphi_M\}$ is the corresponding estimated phases with error n_i. the distance $d_0 = d_{AD} - d_{AC} + d_{BC} - d_{BD}$ is a linear combination of the distances between the four nodes. $\theta = \frac{2\pi\delta}{c}(d_{AD} - d_{AC} - d_{BC} + d_{BD})$ is some constant related to d_0. measurement frequencies $\mathbf{f}$ are chosen randomly within bandwidth B. We assume that all frequencies are multiples of the system's minimum frequency interval $f_{\min}$, then $f_i = (k_0 + k_i)f_{\min}$, and $N = B/f_{\min} + 1$ is the number of maximum chosen frequencies with $k_0 f_{\min}$ be the smallest one. The positive integer k_i distributes uniformly in [0, $N-1$].

In order to describe the estimating error n_i, we assume that the samples number m, decided by hopping rate and sample bandwidth, is large enough to ensure that the variance of n_i equals to the Cramer-Rao Bound [9]. In the sense of statistic we suppose that all the received signals are with same amplitude $a_{XY} = A_0$ and same variance of noise $N_0/2$. At this time we have $n_i \sim \mathbb{N}(0, \sigma^2)$ and $\sigma^2 = 4N_0/(mA_0^2)$.

9.2.2 Maximum Likelihood Estimator of IRHF

According to the conclusion that (9.2) is equivalent to the single-tone frequency estimation model in mid and high SNR [1, 8], we have $\exp(j\varphi_i) = \exp(j(2\pi\frac{f_i}{c}d_0 + \theta)) + z_i$, where z_i is i.i.d Complex AWGN with variance $2\sigma^2$. Then the joint distribution function of output $\exp(j\mathbf{\Phi})$ at $\mathbf{f}$ with the unknown parameter vector $\mathbf{A} = [d_0, \theta]^T$ is

$$f(\mathbf{\Phi}; \mathbf{A}) = \left(\frac{1}{2\pi\sigma^2}\right)^M \exp\left[-\frac{1}{\sigma^2}\left(\sum_{i=1}^{M}(a_i - \mu_i)^2 + \sum_{i=1}^{M}(b_i - v_i)^2\right)\right] \tag{9.3}$$

where $a_i = \mathbf{Re}(\exp(j\varphi_i))$, $b_i = \mathbf{Im}(\exp(j\varphi_i))$, $\mu_i = \cos(2\pi\frac{f_i}{c}d_0 + \theta)$, and $v_i = \sin(2\pi\frac{f_i}{c}d_0 + \theta)$.

As a result, the MLE of $\hat{d}_0$ is obtained by maximizing the criterion function

$$V(d) = \left|\sum_{i=1}^{M}\left\{\exp(j\varphi_i)\exp(-j2\pi\frac{f_i}{c}d)\right\}\right| \tag{9.4}$$

where $d \in [d_0 - \frac{d_{\max}}{2}, d_0 + \frac{d_{\max}}{2}]$, and $d_{\max}$ is the range of interest within the unambiguous distance of IRHF.

9.3 Ranging Performance in Benign Environment

The MLE of IRHF is a nonlinear process with a threshold effect [6]. When the SNR decreases below the threshold point, the MSE of estimator increases significantly due to the global errors also known as 'outlier' [10]. For IRHF the ranging accuracy often locates in the region of threshold because of the hopped frequencies. Similar to the concept of 'bit error probability' in the arena of digital communication, the 'outlier probability' plays an important role to describe the ranging performance of IRHF.

The regions of the parameter space where outliers may occur are identified by ambiguity function, which is obtained by evaluating the criterion function over the range of parameter searched when the data is noise free [10]. Furthermore, an outlier denotes the event that the global maximum in criterion function is outside the mainlobe area of ambiguity function [6].

The variation of measurement frequencies in IRHF, along with which ambiguity functions fluctuate, brings difficulty in analysis. The similar problem occurs in Random Signal Radars, where 'Average Ambiguity Function (AAF)' plays an important role in describing radar performance [7]. It follows that we define the AAF of IRHF as the expectation of (9.4) replacing φ_i by (9.2) with noise free that

$$\bar{G}(d) = E\left|\sum_{i=1}^{M} \exp\left(j2\pi\frac{f_i}{c}(d-d_0)\right)\right| = \left|\frac{M\sin(\pi(d-d_0)\frac{Nf_{\min}}{c})}{N\sin(\pi(d-d_0)\frac{f_{\min}}{c})}\right| \tag{9.5}$$

We denote P_o as outlier probability. To simplify the evaluation of P_o, the AAF can be discretized at the sidelobe peaks and the criterion function can be evaluated only at d_n, $n = 0, 1, \ldots, N_p$, where d_ns are positions and N_p is the number of sidelobe peaks of the AF. Under this discretization, the P_o can be approximated in the probability of the union of events that one of the sidelobe peaks of AAF is higher than the mainlobe [6]. With the help of union bound, we have

$$P_o \approx \sum\nolimits_{n=1}^{N_p} p_n \tag{9.6}$$

where $p_n = \Pr[V(d_0) < V(d_n)]$, always called 'pairwise error probability'. d_ns are positions and N_p is numbers of the sidelobe peaks. It follows from (9.5) that

$$N_p = \left\lfloor \frac{c}{f_{\min}} \Big/ \frac{c}{Nf_{\min}} \right\rfloor - 2 = N - 2 \tag{9.7}$$

and

$$d_n = d_0 + (-1)^n \frac{c}{Nf_{\min}}(\lceil N/2 \rceil + 0.5) \tag{9.8}$$

We now turn to the derivation of the pairwise error probability p_n. Let

$$y_0 = \sum_{i=1}^{M} \exp(jn_i) \tag{9.9}$$

$$y_n = \sum_{i=1}^{M} \exp\left(j2\pi\frac{f_i}{c}(d_0 - d_n) + jn_i\right) \tag{9.10}$$

Then $V(d_0) = y_0$ and $V(d_n) = y_n$. Consequently we have $p_n = \Pr(|y_0| - |y_n| < 0)$ $= \Pr\left(|y_0|^2 - |y_n|^2 < 0\right)$.

Computing p_n depends on the distribution of y_0 and y_n, of which the calculation is quite involved. it is observed that y_0, as well as y_n, is the summation of M i.i.d random variables. In view of central-limit theorem, it is quite enough to guarantee both y_0 and y_n submit approximately to Gaussian distribution if $M \geq 30$, and the number M in IRHF always meets the requirement.

As a result, considering y_0 and y_n are correlated, we will get p_n resorting to appendix B in [11], only with effort to calculate the means and second order moments of y_0 and y_n.

Based on the definition of character function, we have $X = E[\exp(jx)] = e^{-\sigma^2/2}$ if x is submitted to normal distribution. Then some routine computation gives that the means of y_0 and y_n are

$$\begin{aligned} E[y_0] &= MX \\ E[y_n] &= Mr_nX \end{aligned} \tag{9.11}$$

While the second order moments are

$$\begin{aligned} \mathrm{var}[y_0] &= 0.5M(1 - X^2) \\ \mathrm{var}[y_n] &= 0.5M(1 - |r_n|^2X^2) \\ \mathrm{cov}[y_0, y_n] &= 0.5M(r_n)^*(1 - X^2) \end{aligned} \tag{9.12}$$

where $r_n = E[\exp(j2\pi\frac{f_i}{c}(d_0 - d_n))]$ is interpreted as the relative sidelobe level of the nth sidelobe of AAF. Substitute formulas from (9.11) and (9.12) into formula B-21 of [11], we have

$$p_n = Q_1(a, b) - \frac{v}{1+v} I_0(ab) \exp[-\frac{1}{2}(a^2 + b^2)] \tag{9.13}$$

Here, Q1(·,·) is Marcum's Q function, $I_0(\cdot)$ is modified Bessel function of the first kind and order 0. Let$A_n = 1 - |r_n|^2$, $S = 4(X^{-4} - X^{-2})$, then $v = \frac{\sqrt{1+S/A_n}-1}{\sqrt{1+S/A_n}+1}$, $a = \sqrt{\frac{M}{A_n+S}\left(2X^{-2} - A_n - \sqrt{A_n^2 + A_nS}\right)}$, $b = \sqrt{\frac{M}{A_n+S}\left(2X^{-2} - A_n + \sqrt{A_n^2 + A_nS}\right)}$.

9.4 Capability Against Partial-Band Jamming

Anti-jamming capability of FHSS communication system is always related to bit error probability [12, 13], which reflects the reliability of communication system. Similar to bit error probability, the increased outlier probability caused by interference is suitable to indicate the reliability of ranging system. So next we describe the anti-jamming capability of IRHF resorting to outlier probability of ranging under partial-band jamming. The worst-case jamming as well as the minimum Signal-to-jamming power ratio (SJR) required for certain probability with different fraction of jamming are analyzed.

9.4.1 Jamming Model

The PBJ considered here is modeled as a zero-mean Gaussian random process with a flat power spectral density $N_J/2\beta$ over a fraction β of total bandwidth B and zero elsewhere. Let SJR be A_0^2/N_J, and the phase estimating error be n_i'. Then we find that

$$n_i' \sim \begin{cases} \mathbb{N}(0,\sigma^2), & \text{if } f_i \text{ is not jammed} \\ \mathbb{N}\left(0,\sigma^2+\sigma_J^2\right), & \text{if } f_i \text{ is jammed} \end{cases} \tag{9.14}$$

where $\sigma_J^2 = 4N_J/\left(\beta m A_0^2\right)$ and σ^2 is denoted in Sect. 9.2. The measurement frequencies will be jammed with probability β and not jammed with probability $1-\beta$. So the distribution of n_i' is a mixture of two normal distributions. According to the feature of Mixture Distribution [14, 15], we have the moments of n_i' that

$$E\left[n_i'\right] = 0 \tag{9.15}$$

$$\mathrm{var}\left[n_i'\right] = \sigma'^2 = \beta\left(\sigma^2+\sigma_J^2\right) + (1-\beta)\sigma^2 = \sigma^2 + \beta\sigma_J^2 \tag{9.16}$$

It is reasonable to assume that the system don't realize whether the measurement frequencies are jammed or not, and substitute all of φ_i back to (9.4) to estimate distance.

9.4.2 Outlier Probability Under Partial-Band Jamming

We denote P_o' as outlier probability under partial-band jamming. With the help of union bound as before, we have

$$P'_o \approx \sum_{n=1}^{N_p} p'_n \tag{9.17}$$

where $p'_n = \Pr[V(d_0) < V(d_n)]$, d_ns are positions and N_p is numbers of the sidelobe peaks as before. The difference between the benign and the jamming environment is that the n_i of (9.2) turns into n'_i. Let

$$y'_0 = \sum_{i=1}^{M} \exp(jn'_i) \tag{9.18}$$

$$y'_n = \sum_{i=1}^{M} \exp\left(j2\pi\frac{f_i}{c}(d_0 - d_n) + jn'_i\right) \tag{9.19}$$

Then the key step is to calculate the probability $p'_n = \Pr(|y'_0| < |y'_n|)$. By denoting $X' = E[\exp(jn'_i)]$, a simple manipulation based on the moments of Mixture Distribution [14, 15] gives that

$$X' = \beta e^{-(\sigma^2+\sigma_j^2)/2} + (1-\beta)e^{-\sigma^2/2} \tag{9.20}$$

Clearly the expression of p'_n is with the same form as p_n that

$$p'_n = Q_1(a', b') - \frac{v}{1+v} I_0(a'b') \exp[-\frac{1}{2}(a'^2 + b'^2)] \tag{9.21}$$

where $A_n = 1 - |r_n|^2$, $S' = 4(X'^{-4} - X'^{-2})$, and $v = \frac{\sqrt{1+S'/A_n}-1}{\sqrt{1+S'/A_n}+1}$, $a = \sqrt{\frac{M}{A_n+S'}\left(2X'^{-2} - A_n - \sqrt{A_n^2 + A_n S'}\right)}$, $b = \sqrt{\frac{M}{A_n+S'}\left(2X'^{-2} - A_n + \sqrt{A_n^2 + A_n S'}\right)}$.

9.4.3 The Worst-Case Partial-Band Jamming

The jammer's optimum strategy is to select the value of β that maximizes the outlier probability. Differentiating (9.21) and solving for the extremum with the restriction $0 \le \beta \le 1$ is a straightforward way, while the calculation is quite involved.

It is noticed that the outlier probability of IRHF in either benign environments or jamming circumstances is with the same form, and the role of X' in (9.21) is equivalent as X in (9.13). Due to the fact that X is inversely proportional to σ^2 and (9.13) is increasing directly with σ^2 as a matter of common sense (9.13) is a monotone decreasing function of X. Similarly (9.21) is also a monotone decreasing function of X'. As a result, the minimum value of X' leads to the worst-case partial-band jamming. Differentiating X' with respect of β gives rise to

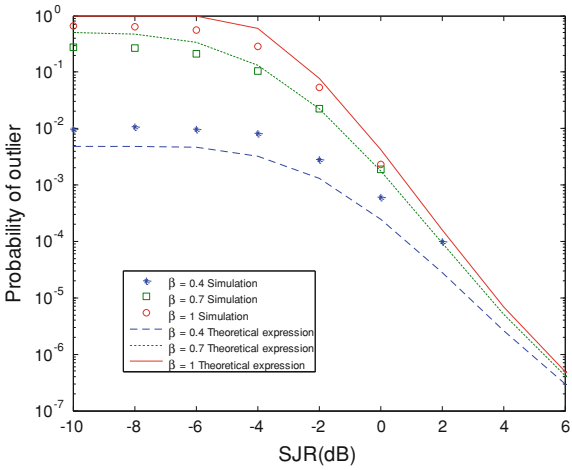


Fig. 9.1 Outlier probability of IRHF under partial-band jamming versus SJR

$$\frac{dX'}{d\beta} = e^{-\sigma^2/2}\left[\left(1+\frac{N_J}{2mA_0^2\beta}\right)\Big/ e^{N_J/\left(2mA_0^2\beta\right)} - 1\right] \tag{9.22}$$

The fact $1+\frac{N_J}{2mA_0^2\beta}$ is constantly smaller than $e^{N_J/\left(2mA_0^2\beta\right)}$ shows that the value of (9.22) is negative, so X' is inversely proportional to β. It follows that IRHF suffers from worst-case jamming when $\beta = 1$, implying full-band jamming. The jammer has to allocate the power through as large portion of the whole bandwidth as possible to obtain optimum jamming.

9.5 Simulation Results

To check the accuracy of theoretical approximations in sections above, Mont Carlo simulations of MLE for IRHF under PBJ is performed in Fig. 9.1. The specification of the simulations is as follow. The bandwidth B of transceiver is 15 MHz, and the minimum frequency interval $f_{\min}$ is 1 kHz. The SJR ranges from -10 to 6 dB, the number of Mont Carlo trials is 10^4 for each SJR, and in each trial 31 frequencies of measurement are randomly chosen from the available frequencies. We examine three different scenarios where β is 0.4, 0.7 and 1. During all the simulations the SNR is 20 dB.

It can be seen that the approximation of outlier probability in (9.17) is accurate in modeling the performance of IRHF under PBJ. Particularly, the expression (9.13) has also been verified since (9.17) equals to (9.13) when β is 1. It's notable that the larger β is, the more outlier IRHF suffers, which is coincident with the analysis in Sect. 9.4.2.

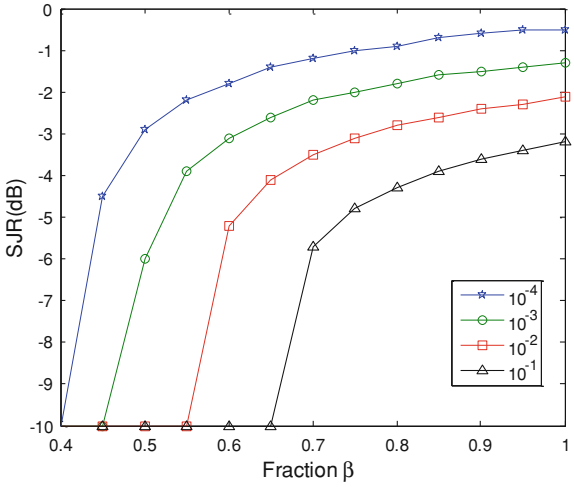


Fig. 9.2 Capability against partial-band jamming of IRHF

On the other hand, the ranging system is able to fight against jamming by adjusting signal emitting power. With the same specification as in Figs. 9.1 and 9.2 illustrates the minimum SJR for IRHF to reach certain outlier probability of 10^{-4}, 10^{-3}, 10^{-2} and 10^{-1} with different values of β. The evaluation is based on (9.17) and (9.21). It is apparent from Fig. 9.2 that in order to maintain the same precision, larger SJR is demanded if β is increasing. In addition, larger SJR leads to higher precision with β fixed.

9.6 Conclusion

This paper investigates the performance of IRHF, a novel ranging technique in constructing anti-jamming navigation systems, in Benign Environment and under Partial-Band Jamming. We have derived the approximated formulas of outlier probability with respect to some system parameters such as SNR, SJR, bandwidth, and the number of measurement frequencies. Then with the evaluation method of anti-jamming capability for ranging system proposed here, we have got a crucial result that the worst-case PBJ for IRHF is full-band jamming.

Theoretical expressions, be in very good agreement with the simulation results, provide basic of system analysis and parameters design for constructing anti-jamming navigation systems in practice.

References

1. Su Z, Wangdong Q, Jiang C et al (2014) Interferometric ranging with hopping frequencies: a new way to anti-jamming PNT systems. In: The 5th China satellite navigation conference (CSNC), Nanjing
2. Wei L, Wangdong Q, Peng L et al (2013) Method for selecting measurement frequencies based on dualp seudo-random code in radio interferometric positioning system. Chinese Patent CN102221695
3. Maroti M, Kusy B, Balogh G et al (2005) Radio interferometric geolocation. In: Proceedings of 3rd ACM international conference on embedded networked sensor systems (SenSys'05). San Diego, pp 1–12
4. Peng L, Wangdong Q, En Y et al (2011) Ground displacement measurement by radio interferometric ranging for landslide early warning. In: IEEE instrumentation and measurement technology conference (I2MTC), Binjiang, pp 1–6
5. Marvin KS, Jim KO, Robert AS et al (2002) Spread spectrum communications handbook. McGraw-Hill, New York, pp 470–471
6. Athley F (2005) Threshold region performance of maximum likelihood direction of arrival estimators. IEEE Trans Signal Process 53(5):1359–1373
7. Kaveh M, Cooper GR (2006) Average ambiguity function for a randomly staggered pulse sequence. IEEE Trans Aerosp Electron Syst 12(3):410–413
8. S. Zhang (2012) Extended beat radio interferometry. PLA University of Science and Technology
9. Rife DC, Boorstyn R (1974) Single tone parameter estimation from discrete-time observations. IEEE Trans Inf Theory 20(5):591–598
10. Richmond C (2006) Mean-squared error and threshold SNR prediction of maximum-likelihood signal parameter estimation with estimated colored noise covariances. IEEE Trans Inf Theory 52(5):2146–2164
11. Proakis JG (2006) Digital communications, 4th edn. Publishing House of Electronics Industry, Beijing, pp 943–948
12. Kang H, Stark WE (1998) Turbo codes for non coherent FH-SS with partial band interference. IEEE Trans Commun 46(11):1451–1458
13. Kegang P, Bangning Z (2009) Performance analysis of differential hopping system against partial-band jamming. Signal Process 25(8A):370–375
14. Frühwirth-Schnatter S (2006) Finite mixture and markov switching models. Springer, New York, pp 10–11
15. Behboodian J (1970) On the modes of a mixture of two normal distributions. Technometrics 12(1):131–139

Chapter 10
Optimized Asymmetric Dual-Frequency Constant Envelope Multiplexing Technique

Jiayi Zhang, Zheng Yao and Mingquan Lu

Abstract In next generation GNSS signal design, the increasingly various services draw a strong demand to the technique of multiplexing, among which the dual-frequency constant envelope multiplexing is one of the hottest issues and has been attracted great attention from researchers and system builders. ACE-BOC is a recently proposed technique, which allows the multiplexing of four or less arbitrary signals into a composite signal with constant envelope. In this paper, an optimization based low-complexity simplification is made for ACE-BOC. To meet the design requirements and constraints of BeiDou satellite navigation system (BDS), a typical implementation for B2 band is provided by using the optimized ACE-BOC. Analyses show that the optimized ACE-BOC achieves superior tracking performance to other representative dual-frequency CEM such as AltBOC and TD-AltBOC, and has equal or lower complexity in the implementation of signal generating. The optimized ACE-BOC provides a competitive alternate for BDS B2 signal with good performance on both transmitter and receiver.

Keywords Constant envelope multiplex · Optimization · Low complexity · Dual-frequency · Navigation signal design

10.1 Introduction

In signal design of the next generation GNSS, in order to provide users with diversified services with different receiving strategies and accuracy requirements, the technique of multiplexing multiple binary spreading signals on two adjacent frequencies into a constant envelope integral signal has been attached increasing

J. Zhang (✉) · Z. Yao · M. Lu
Department of Electronic Engineering, Tsinghua University, Beijing, China
e-mail: zhangjiayi12@mails.tsinghua.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_10,

attention. The dual-frequency constant envelope multiplexing (CEM) provides a viable design frame for the BeiDou satellite navigation system B2 band, where BDS plans to broadcast two QPSK signals located at B2a (1176.45 MHz) and B2b (1207.14 MHz) respectively, with different services carried on each of the two frequencies [1]. Distinct advantages can be obtained when utilizing the dual-frequency CEM in B2 band. In term of the transmitters, the constant envelope of composite signal ensures the share of amplifier chain on generator of B2a and B2b. In term of the receivers, the dual-frequency CEM signal can be treated jointly as a wideband signal, bringing potential for significant performance boost in the future.

In terms of the dual-frequency CEM, alternate binary offset carrier (AltBOC) is a representative technique, allowing the multiplexing of four signals into an integral 8-PSK signal [2], which has been adopted by the Galileo E5 and already in operation. TD-AltBOC [3] is a variant of AltBOC, in which data and pilot components are modulated with different binary offset carrier and combined in time-division (TD) multiplexing, thus reducing the complexity of signal generator by TD switching circuits. However, both AltBOC and TD-AltBOC require the input to be four signals with equal power, while in the modern GNSS signal design, higher power in pilot components are preferred for the promotion in tracking performance [4]. To counter these limitations, asymmetric constant envelope binary offset carrier (ACE-BOC) multiplexing [5, 6, 8] is proposed recently, which is prominent in the flexibility, allowing the combination of four or less signals with different modulation and arbitrary power allocation. However, the required driving clock rate of original ACE-BOC is relatively high for the digital circuit implementation. Due to the particularly strict demand in stability and reliability for hardware equipped on the payload of satellite, system builders tend to adopt the multiplexing techniques with lower complexity on the premise of operating performance.

In this paper, simplification method is proposed for ACE-BOC. The optimization based mathematical model of the signal as well as the generating approach is provided. Then, according to practical request of BDS B2 signal design, a specific implementation is presented by the optimized ACE-BOC. Performance are analyzed including the multiplexing efficiency, complexity and the inherent code tracking performance and compared with other dual-frequency CEM techniques. Results show that, the optimized ACE-BOC is applicable with a low driving clock rate in the multiplexer, while maintain the adjustability of the power distribution of services. The work in this paper provides an engineering facilitated improvement for the original ACE-BOC, which can be referred as a distinctive and competitive multiplexing solution for BDS B2 signal.

10.2 Decomposed form of ACE-BOC Multiplexing

ACE-BOC multiplexing is utilized to combine four or less composing signals into an integral signal with constant envelope. Considering multiplexing four binary baseband spreading signals, denoted by s_{UI}, s_{LI}, s_{UQ} and s_{LQ}, into a dual-frequency

split-spectrum signal, in which s_{UI} and s_{UQ} are modulated on the upper sideband with in-phase and quadrature-phase respectively, while s_{LI} and s_{LQ} are modulated on the lower sideband with in-phase and quadrature-phase respectively. Signals can be allocated with arbitrary power, denoted by P_{UI}, P_{LI}, P_{UQ} and P_{LQ} respectively. In addition to the four signals, inter-modulation (IM) components are modulated in the ACE-BOC composite signal, which are derived as $\bar{s}_{UI} = s_{UQ}s_{LI}s_{LQ}$, $\bar{s}_{UQ} = s_{UI}s_{LI}s_{LQ}$, $\bar{s}_{LI} = s_{UI}s_{UQ}s_{LQ}$ and $\bar{s}_{LQ} = s_{UI}s_{UQ}s_{LI}$.

In order to ensure the four signal s_{UI}, s_{LI}, s_{UQ} and s_{LQ} are modulated to its corresponding frequency and phase, every baseband signal as well as the IM should be modulated with a complex subcarrier. The decomposed expression of ACE-BOC signal [6] can be written as

$$\begin{aligned} s_{ACE-BOC}(t) = {} & s_{UI}(t)[sc_{S-UI}(t) + \mathrm{j}sc_{S-UI}(t - T_{sc}/4)] + \mathrm{j}s_{UQ}(t)[sc_{S-UQ}(t) + \mathrm{j}sc_{S-UQ}(t - T_{sc}/4)] \\ & + s_{LI}(t)[sc_{S-LI}(t) - \mathrm{j}sc_{S-LI}(t - T_{sc}/4)] + \mathrm{j}s_{LQ}(t)[sc_{S-LQ}(t) - \mathrm{j}sc_{S-LQ}(t - T_{sc}/4)] \\ & + \bar{s}_{UI}(t)[sc_{P-UI}(t) + \mathrm{j}sc_{P-UI}(t - T_{sc}/4)] + \mathrm{j}\bar{s}_{UQ}(t)[sc_{P-UQ}(t) + \mathrm{j}sc_{P-UQ}(t - T_{sc}/4)] \\ & + \bar{s}_{LI}(t)[sc_{P-LI}(t) - \mathrm{j}sc_{P-LI}(t - T_{sc}/4)] + \mathrm{j}\bar{s}_{LQ}(t)[sc_{P-LQ}(t) - \mathrm{j}sc_{P-LQ}(t - T_{sc}/4)] \end{aligned} \tag{10.1}$$

where sc_{S-UI}, sc_{S-LI}, sc_{S-UQ} and sc_{S-LQ} are real subcarriers for useful signal components. sc_{P-UI}, sc_{P-LI}, sc_{P-UQ} and sc_{P-LQ} are real subcarriers for IM terms.

10.3 Optimized Asymmetric Dual-Frequency CEM

10.3.1 General Optimization of Dual-Frequency Multiplexing

Not all subcarriers set can ensure the multiplexed signal to have a constant envelope. Except for the constraint of constant envelope, there also involves that the received power and relative phase of different components should be consistent with the design, and minimum the power loss induced by multiplexing. In this aspect, it is reasonable to simplify the original ACE-BOC by optimizing of the subcarrier set under the mentioned constraints.

For illustrative convenience, the derivation below is based on a typical and application promising power allocation of the dual-frequency multiplexing, as shown in Fig. 10.1a, that the pilot power is r times of the data power, and the upper and lower sideband are asymmetrical, denoted service Type I. The optimized ACE-BOC can be easily generalized and applied in other situations, as shown in Fig. 10.1b, c.

Under the power constrains of Type I, the number of different subcarriers in the signal generation of (10.1) is reduced from 8 to 4, with two subcarriers sc_{SI}, sc_{SQ} are modulated to signal components and sc_{PI}, sc_{PQ} for IMs.

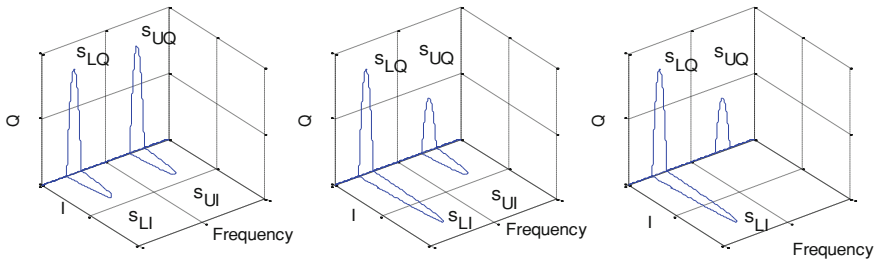


Fig. 10.1 Spectral sketches of three typical service types for BDS B2 signal. **a** Type I. **b** Type II. **c** Type III

The precondition of the simplification of ACE-BOC is that the subcarriers in a cycle should have less number of time slots than original ACE-BOC, and every slot is *equal in length*. In theory, the number of segments in a subcarrier cycle should be any multiples of 4, it is reasonable to assume it 8 for instance. Then 8 variables are used to describe a subcarrier, each lasting for $T_{sc}/8$. Following analysis will prove that the 8 variables can be further simplified in some special cases of constraints.

10.3.2 Constructing the Optimization

In this work, the design of the optimized ACE-BOC is equivalent to an optimization problem. This subsections will develop the constraints and objective function to a solvable numerical optimization problem, which can be used in implementation design for the optimized ACE-BOC.

Firstly, the constraints of phase relationship of the input components can be derived from the Fourier series of subcarrier. In order to satisfy the phase orthogonality between received signals on channel I and Q, the subcarrier in a cycle should be symmetric to $T_{sc}/2$. And in order to modulate most power to the main lobe of the spectrum and minimum the useless high-ordered harmonics, the subcarrier should satisfy $sc(t) = -sc(t - T_{sc}/4)$ in every half cycle. So far, the number of subcarrier values is reduced to 2 for each components.

Secondly, the constraint that the multiplexed signal should have constant envelope is developed by using the tool of inter-modulation construction (IMC) [7], which describes the relation between the modulation phase and the subcarrier value for CEM of N signals as below.

$$\exp(\mathrm{j}\Theta(t)) = \mathbf{SC}(t) \tag{10.2}$$

In (10.2), $\mathbf{S}$ is a 2^N order Hadamard matrix, enumerating all the 2^N combination for N signals, $\mathbf{C}(t)$ is complex subcarriers vector for the corresponding signal term, and $\Theta(t)$ provides the modulation phase. By expanding the matrix expression in

the two dimensions of input state and time, one can get 16 × 8 quadratic equality constraints with variables of subcarrier values. Constant envelope is achievably ensured by these equations after duplicates removal and simplification.

The third constraint is that the received power of components should be as designed, for instance, the power ratio of pilot to data should be r for Type I. It can be easily implemented by constraining the average power of the corresponding subcarrier.

Finally, the objection of CEM design is to maximum the proportion of useful power in the total transmitting power, characterized by multiplexing efficiency.

$$\eta = \frac{\sum_{n=1}^{4} \left| \mathrm{E}\left\{ \frac{1}{T} \int_T s_{CEM}(t) \tilde{s}_n^*(t)\, \mathrm{d}t \right\} \right|^2}{A^2} \tag{10.3}$$

where A is the amplitude of multiplexed signal, the numerator is the sum of the average output of correlator for the four composing signals. In this work, discussion of the optimization is under the generally adopted non-matched receiving [8], that separately processes the 4 components as BPSK signals. With replica of local subcarrier $\tilde{s}(t) = \exp(\pm \mathrm{j}2\pi f_{sc} t)$, one can derive the object by using (10.3).

So far, the optimization problem is totally expanded, where $x_m^{(\cdot)}$ stands for the value of the mth time slot of the subcarrier $sc_{(\cdot)}$.

$$\begin{aligned}
\textbf{max} \quad & (x_1^{(SI)})^2 + (x_1^{(SQ)})^2 + \sqrt{2}(2-\sqrt{2})(x_1^{(SI)} x_2^{(SI)} + x_1^{(SQ)} x_2^{(SQ)}) \\
\textbf{s.t.} \quad & x_2^{(SI)} = x_2^{(PI)} \\
& x_2^{(SQ)} = x_2^{(PQ)} \\
& (x_2^{(SI)})^2 + (x_2^{(SQ)})^2 = 1/16 \\
& (x_1^{(SI)} - x_1^{(PI)})^2 = 1/4 \\
& (x_1^{(SQ)} - x_1^{(PQ)})^2 = 1/4 \\
& (x_1^{(SI)} + x_1^{(PI)})^2 + (x_1^{(SQ)} + x_1^{(PQ)})^2 = 1/4 \\
& r((x_1^{(SI)})^2 + (x_2^{(SI)})^2) - (x_1^{(SQ)})^2 - (x_2^{(SQ)})^2 = 0 \\
& x_1^{(PI)} < 0,\ x_1^{(PQ)} < 0 \\
& x_1^{(SI)} > 0,\ x_2^{(SI)} > 0,\ x_1^{(SQ)} > 0,\ x_2^{(SQ)} > 0,\ x_2^{(PI)} > 0,\ x_2^{(PQ)} > 0
\end{aligned} \tag{10.4}$$

The quadratic optimization problem can be solved by either analytical derivation or numerical method, such as sequential quadratic programming (SQP). Given the input power allocation of four or less signals, implementation for the optimized ACE-BOC can be obtained by solving the optimization problem, showing the flexibility inheriting from of the original ACE-BOC.

10.4 Typical Case of Implementation

By using the optimized ACE-BOC technique, specific implementation schemes under typical power allocation can be provided. Also, promising application are analyzed based on the characteristic of the service type.

A representative service type for dual-frequency CEM is the previous mentioned Type I, as shown in Fig. 10.1a, where two QPSK are set symmetrical to the central frequency, while the pilot power is β^2 times of the data power. As accepted, increasing the power of pilot components helps improving the tracking accuracy and robustness, thus enhancing the performance in extreme environments, such as low SNR and complex multipath. Type I reflects the current trend of new and modern GNSS signal design and is of good application prospects among the dual-frequency multiplexing schemes. The implementation presented below provides a multiplexing solution for BDS B2 civil signal.

By using the optimized ACE-BOC, we can derive the complete design for the Type I and BDS B2, with $P_{B2b-I} : P_{B2a-I} : P_{B2b-Q} : P_{B2a-Q} = 1 : 1 : 3 : 3$, including subcarrier value in Table 10.1, LUT for phase state in Table 10.2 and the corresponding phase in (10.5). The subcarrier waveform is shown in Fig. 10.2 and the Fresnel Plot of modulation phase is shown in Fig. 10.3.

$$\theta(k) = \left[\theta_1,\ \theta_2,\ \frac{\pi}{2} - \theta_2,\ \frac{\pi}{2} - \theta_1\right]^{\mathrm{T}} \quad \text{where } \theta_1 = 0.2371\,(\text{rad}), \quad \theta_2 = 0.3208\,(\text{rad})$$

$$\Theta_I(k) = \begin{cases} \theta\,(k), & k = 1, \ldots, 5 \\ \pi - \theta(k-5), & k = 6, \ldots, 10 \\ \pi + \theta(k-10), & k = 11, \ldots, 15 \\ 2\pi - \theta(k-15), & k = 16, \ldots, 20 \end{cases} \tag{10.5}$$

From both Fig. 10.2 and Table 10.2 we can see that the implementation derived from the optimized ACE-BOC requires a driving clock rate of $8f_{sc}$ in generator while the implementation of Type I give by the original ACE-BOC in [6, 8] requires a driving clock rate of at least $12f_{sc}$. In this aspect, the optimized ACE-BOC decreased the clock rate by 33.3 %, which facilitates the implementation of signal generator in stability and reliability.

The implementation of the optimized ACE-BOC can also be generated under arbitrary power allocation of the input components, which needs slight adjustment to the optimization proposed in this paper. The typical service Type II and Type III, as shown in Fig. 10.1b and c, are also implemented by using the optimized ACE-BOC, which cannot be listed here due to the space limitation. Evaluation of the optimized ACE-BOC under the three types in Fig. 10.1 are provided in the following section.

Table 10.1 Look up table for phase state for Type I with $P_{B2b-I} : P_{B2a-I} : P_{B2b-Q} : P_{B2a-Q} = 1 : 1 : 3 : 3$

s_{UI}	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1
s_{LI}	1	1	1	1	−1	−1	−1	−1	1	1	1	1	−1	−1	−1	−1
s_{UQ}	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1
s_{LQ}	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1
$i_{T_{sc}}$	Index k of phase state $\Theta_k(t)$															
0	4	20	20	16	5	8	2	15	5	12	18	15	6	10	10	14
1	3	10	20	17	5	9	1	5	15	11	19	15	7	10	20	13
2	13	10	20	7	5	9	1	5	15	11	19	15	17	10	20	3
3	14	10	10	6	15	8	2	5	15	12	18	5	16	20	20	4
4	14	10	10	6	15	18	12	5	15	2	8	5	16	20	20	4
5	13	20	10	7	15	19	11	15	5	1	9	5	17	20	10	3
6	3	20	10	17	15	19	11	15	5	1	9	5	7	20	10	13
7	4	20	20	16	5	18	12	15	5	2	8	15	6	10	10	14

The mapping from states k to the transmitting phase $\Theta(k)$ is given in (10.5). The time index $i_{T_{sc}}$ is given by $\mathrm{i_{Tsc}}$ = integer part$[8f_{sc}(t \bmod T_{sc})]$, with $i_{T_{sc}} \in \{0, 1, 2, 3, 4, 5, 6, 7\}$

Table 10.2 The optimized ACE-BOC subcarrier values for Type I in the first half cycle, while the other half can be obtained by even-symmetry

Time slot	1	2	3	4
$4SI_i$	$\sin\theta_1 + 1$	$\sin\theta_2$	$-\sin\theta_2$	$-\sin\theta_1 - 1$
$4SQ_i$	$\cos\theta_1 + 1$	$\cos\theta_2$	$-\cos\theta_2$	$-\cos\theta_1 - 1$
$4PI_i$	$\sin\theta_1 - 1$	$\sin\theta_2$	$-\sin\theta_2$	$-\sin\theta_1 + 1$
$4PQ_i$	$\cos\theta_1 - 1$	$\cos\theta_2$	$-\cos\theta_2$	$-\cos\theta_1 + 1$

$\theta_1 = 0.2371$ (rad) $\approx$ 13.5845 (deg), $\theta_2 = 0.3208$ (rad) $\approx$ 18.3805 (deg)

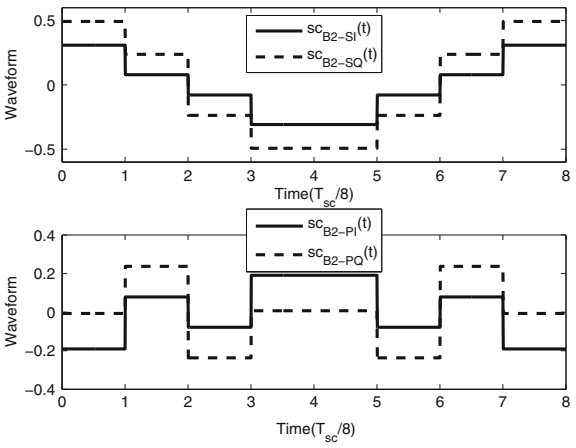


Fig. 10.2 Waveform of subcarriers for Type I

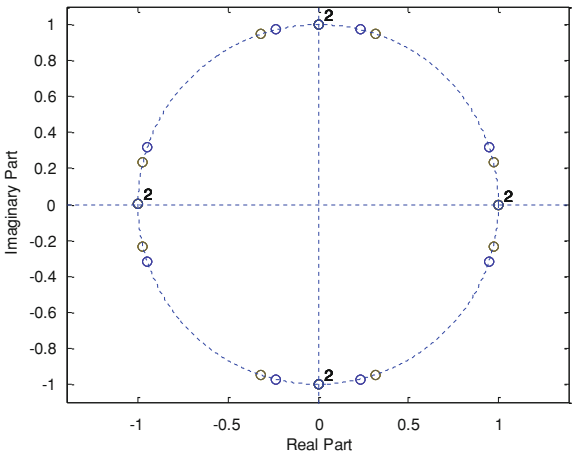


Fig. 10.3 Transmitting phase of Type I

10.5 Performance Analyses

10.5.1 Multiplexing Efficiency

Dual-frequency CEM enables various receiving methods for different users. In terms of consumer electronics products, such as handheld receivers, they prefer to receive one of the two sideband of the wideband signal and use QPSK-like processing. In terms of high accuracy receivers, the wide front-end bandwidth allows dual-band signals received simultaneously, bringing better performance in tracking accuracy and anti-multipath along with wideband signal.

In the non-matched receiving, local replica $\tilde{s}(t) = \exp(\pm j2\pi f_{sc}t)$ is selected. The non-matched multiplexing efficiency shows the receiving performance of most of the wideband B2 receivers. While in the matched receiving, local replica is the same multi-level subcarriers as used in signal generation. The multiplexing efficiency obtained in this way shows the upper bound of the potential performance of the signal. As the technique of receiving and processing improves, it will be achieved in the future by taking advantage of the higher harmonics in multi-level subcarrier.

Table 10.3 shows the multiplexing efficiency of the optimized ACE-BOC under matched receiving and non-matched receiving. From Table 10.3, we can see that the multiplexing efficiency of Type I and Type III is slightly lower than the AltBOC and the original ACE-BOC, while further discussion on the tracking performance will show that it is still competitive among the existing multiplexing techniques. The Type II is higher than the two existing CEM techniques in both matched receiving and unmatched receiving.

Table 10.3 Multiplexing efficiency of the optimized ACE-BOC in Type I, Type II and Type III, along with AltBOC and the original ACE-BOC Type I for comparison, where $\beta^2 = 3$

	AltBOC	Original ACE-BOC	Optimized ACE-BOC		
			Type I	Type II	Type III
Matched receiving	0.8536	0.8720	0.8017	0.8796	0.8146
Non-matched receiving	0.8106	0.8106	0.7557	0.8353	0.5515

10.5.2 Tracking Performance

Tracking accuracy is an essential performance indicator for satnav system. Leaving enough potential for tracking accuracy is one of the guideline of new and modern GNSS signal design. The tracking performance can be evaluated by the Cramer-Rao Lower Bound (CRLB) of code-tracking error [9].

$$\sigma^2_{CELP} = \frac{B_L(1 - 0.5B_LT)}{(2\pi)^2 a_P \eta \frac{C_s}{N_0} \beta^2_{rms}} \tag{10.6}$$

where B_L is the one-sided equivalent rectangular bandwidth of code-tracking loop, T is the integration time, β_{rms} is the root mean square (RMS) bandwidth, a_p is the proportion of the pilot power and η is the multiplexing efficiency. It should be noted that (10.6) is based on the tracking of pilot components.

Figure 10.4 shows the LB of code-tracking error in matched receiving and non-matched receiving with C_s/N_0 from 25 to 45 (dB-Hz). The power allocation is as Type I for the optimized and the original ACE-BOC with $\beta^2 = 3$. It should be noted that in the simulation results, pure pilot components are used in the simulation of tracking.

From Fig. 10.4 one can see that the tracking error of the optimized ACE-BOC is slightly worse than original ACE-BOC, but have notable advantage over AltBOC and TD-AltBOC in both matched and non-matched receiving. Observation of Fig. 10.4 shows that the tracking error variance of the optimized ACE-BOC and the original ACE-BOC maintain a constant ratio of 1.07, while the ratio of the optimized ACE-BOC to the AltBOC and TD-AltBOC is 0.72 in the non-matched receiving. That means, utilizing the optimized ACE-BOC, BDS B2 will obtain effectively reduction in the tracking error. With the same transmitting power, compared with AltBOC, the optimized ACE-BOC will bring BDS B2 users a 1.46 dB improvement in tracking threshold.

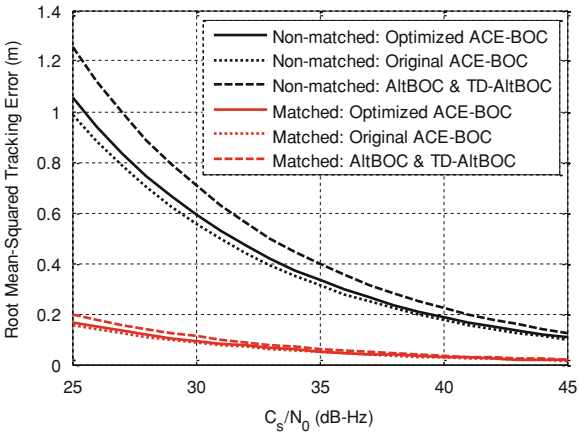


Fig. 10.4 The lower bound of code-tracking error (m), with $B_L = 1$ Hz, $T = 20$ mS

10.5.3 Complexity of Implementation

Considering the BDS B2 civil use, the dual-frequency CEM signal has a code rate of $f_c = 10.23$ MHz and subcarrier rate of $f_{sc} = 15.345$ MHz. The optimized ACE-BOC has 8 segments of equal length in a subcarrier cycle, requiring a clock rate of $8f_{sc} = 120 \times 1.023$ MHz, which is same as AltBOC. However, the original ACE-BOC requires a clock rate of $12f_{sc} = 180 \times 1.023$ MHz or more. As a result, optimized ACE-BOC requires only 66.7 % or lower driving clock rate than the original ACE-BOC, and has the same implementation complexity with AltBOC.

In terms of the interoperability, the optimized ACE-BOC can be received like QPSK, without hardware change of existing receivers for GPS L5 and Galileo E5. While TD-AltBOC needs a switching unit in receiver, leading to additional complexity for multi-system user device.

10.6 Conclusion

Based on the constraints and requirements of BeiDou navigation satellite system on B2 band, this paper proposes an optimization based general dual-frequency constant envelope multiplexing technique for arbitrary no more than four signals, which can be seen as a low-complexity method of ACE-BOC. Design philosophy, optimizing process and generation approach are presented. Aiming at the BDS B2 civil use, based on the proposed technique, a distinctive multiplexing service type is proposed with both the implementation schemes and performance analyses. From the analysis results one can see that while maintaining the advantage of high flexibility in power allocation design, the optimized ACE-BOC counters the disadvantage of high clock rate in the generation of the original ACE-BOC and

achieves significant performance advantage than AltBOC and TD-AltBOC in pilot receiving. This paper provides a practical and competitive solution for BDS B2 signal.

Acknowledgments This work is supported by the National Science Foundation of China (NSFC), Grant 61201190.

References

1. ICG (2010) Current and planned global and regional navigation satellite systems and satellite based augmentations systems. In: International committee on global navigation satellite systems providers forum, New York
2. EU (2010) European GNSS (Galileo) Open service signal in space interface control document, Issue 1
3. Tang Z, Zhou H, Wei J et al (2010) TD-AltBOC: a new COMPASS B2 modulation. In: China satellite navigation conference 2010, Shanghai
4. Betz J (2013) Something old, something new: signal structures for satellite-based navigation: past, present, and future, Inside GNSS
5. Yao Z, Lu M (2012) Dual-frequency constant envelope multiplex with non-equal power allocation for GNSS. IET Electron Lett 48(25):1624–1625
6. Yao Z, Lu M (2013) Design, implementation, and performance analysis of ACE-BOC modulation. In: 26th international technical meeting of the satellite division of the institute of navigation, ION GNSS, Nashville
7. Zhang XM, Zhang X, Yao Z et al (2012) Implementations of constant envelope multiplexing based on extended interplex and inter-modulation construction method In: 25th international technical meeting of the satellite division of the institute of navigation, ION GNSS, Nashville
8. Yao Z, Lu M (2013) Constant envelope combination for components on different carrier frequencies with unequal power allocation. In: International technical meeting of the institute of navigation, San Diego, pp 629–637
9. Betz JW, Kolodziejski KR (2009) Generalized theory of code tracking with an early-late discriminator part I: lower bound and coherent processing. IEEE Trans Aerosp Electron Syst 45(4):1538–1556

Chapter 11
Performance Evaluation of Interoperable Signals on B2 Band

Su Chengeng, Tang Zuping and Meng Yinan

Abstract Interoperability has become an important trend among future global satellite navigation systems. As the interoperable signal options proposed on B2 band, this paper evaluated AltBOC, TD-AltBOC and ACED in aspects of ranging accuracy, sensibility, Time to First Fix (TTFF), anti-interference performance, anti-multipath performance, multiplexing efficiency, cross correlation property and complexity. It can be a reference for COMPASS signal design.

Keywords Interoperability · GNSS · Signal structure · Performance evaluation

11.1 Introduction

Interoperability means without significant increment of using cost, better service can be enjoyed for users who use signals from multiple satellite navigation systems than for those who use signals from single system. Nowadays, interoperability has become a key factor in frequency selection, signal design, negotiation and coordination issues for each satellite navigation system.

According to current situation of signal design, the interoperability work on L1/E1 and L5/E5 a bands has been realized for GPS and Galileo which share the same center frequency and spectrum structure, their difference exists in modulation method, PRN code and message structure; the interoperability of GPS and QZSS has reached the maximum degree on L1, L2 and L5 bands: the center frequency, spectrum structure, modulation method, PRN code, message structure and satellite

S. Chengeng (✉) · T. Zuping
Beijing Institute of Tracking and Telecommunication Technology, Beijing 100094, China
e-mail: hancecilia@126.com

M. Yinan
Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_11,

numbers are basically the same while several message parameters are different; GLONASS is developing the work of future modernized signal design on L1 and L5 bands actively for the purpose of interoperability considering the center frequency of its civil signal is different from other systems. It's seen that the emphasis of interoperable signal design will be focused on L1/B1 and L5/E5/B2 bands for each system. The level of interoperability among those systems is also different, considering the cost of interoperable receiver is most sensitive to carrier frequency and front-end bandwidth, same carrier frequency and front-end bandwidth is taken as the basis of interoperability. Interoperability has been fully considered for COMPASS system on B1/L1 and B2/L5 bands by sharing the same frequency point and similar spectrum structure and front-end bandwidth with the interoperable signal of GPS and Galileo, different levels of interoperability can be met by designing modulation method, PRN code and navigation message flexibly.

AltBOC modulation method proposed by CNES is used by Galileo system on E5 band which ingeniously realizes the constant envelope multiplexing of four signals [1]. AltBOC is also a derivation of BOC modulation in which different PRN code can be modulated for its upper and lower lobes respectively, this method makes two signals able to be transmitted with single power amplifier; for the receiver, AltBOC signal can be either received independently to achieve the performance of BPSK signal or received jointly as large bandwidth signal with both side lobes to achieve its maximum tracking accuracy and anti-multipath performance.

Self-innovative modulation methods TD-AltBOC [2] and ACED [3] are also proposed by the native scholars. In a chip-by-chip time division way, at any time TD-AltBOC has to transmit only two signal components and thus constant envelope multiplexing can be realized without adding product components. ACED makes the power ratio of pilot and data components adjustable, it's a breakthrough in signal power ratio assignment which is unadjustable for AltBOC.

Using modulation method similar to AltBOC on B2 band is beneficial for the interoperability with GPS and Galileo systems and the flexibility of receiver configuration, it takes the development of future ultra wide band high accuracy receiver into consideration and meets the demand of high positioning accuracy scenarios such as urban vehicle accurate navigation, geodesy in complicated environment, civil aviation, etc. In this paper the performance of the three interoperable signal options on B2 band are preliminarily analyzed, more discussion on this subject is expected to be seen.

11.2 Interoperable Signal Schemes on B2 Band

The core task of the interoperable signal design on B2 band is the design of constant envelope multiplexing method in which way B2a and B2b signals can be broadcasted in the same center frequency, existing multiplexing methods include AltBOC, TD-AltBOC and ACED.

11.2.1 AltBOC

Baseband signal of AltBOC(15, 10) modulation method of which the subcarrier frequency is 15.345 MHz can be mathematically expressed as following expression [4]. B2a and B2b signals can be received independently according to the fact that they can be regarded as two separate QPSK(10) signals centering their carrier frequency at 1176.45 and 1207.14 MHz respectively.

$$\begin{aligned} s_{B2}(t) = & \frac{1}{2\sqrt{2}}\left(S_{B2a\_data} + jS_{B2a\_pilot}\right)\left[sc_{B2-S}(t) - jsc_{B2-S}\left(t - T_{SC,B2}/4\right)\right] + \\ & \frac{1}{2\sqrt{2}}\left(S_{B2b\_data} + jS_{B2b\_pilot}\right)\left[sc_{B2-S}(t) + jsc_{B2-S}\left(t - T_{SC,B2}/4\right)\right] + \\ & \frac{1}{2\sqrt{2}}\left(\bar{S}_{B2a\_data} + j\bar{S}_{B2a\_pilot}\right)\left[sc_{B2-P}(t) - jsc_{B2-P}\left(t - T_{SC,B2}/4\right)\right] + \\ & \frac{1}{2\sqrt{2}}\left(\bar{S}_{B2b\_data} + j\bar{S}_{B2b\_pilot}\right)\left[sc_{B2-P}(t) + jsc_{B2-P}\left(t - T_{SC,B2}/4\right)\right] \end{aligned}$$

where the dashed signal components $\bar{S}_{B2a\_data}$, $\bar{S}_{B2a\_pilot}$, $\bar{S}_{B2b\_data}$ and $\bar{S}_{B2b\_pilot}$ are the product signals according to the following equations

$$\begin{aligned} \bar{S}_{B2a\_data} &= S_{B2a\_pilot}S_{B2b\_data}S_{B2b\_pilot} \\ \bar{S}_{B2b\_data} &= S_{B2b\_pilot}S_{B2a\_data}S_{B2a\_pilot} \\ \bar{S}_{B2a\_pilot} &= S_{B2a\_data}S_{B2b\_data}S_{B2b\_pilot} \\ \bar{S}_{B2b\_pilot} &= S_{B2b\_data}S_{B2a\_data}S_{B2a\_pilot} \end{aligned}$$

The parameters sc_{B2-S} and sc_{B2-P} are the signal and product components of the four level subcarrier:

$$\begin{aligned} sc_{B2-S}(t) &= \sum_{l=-\infty}^{\infty}\sum_{i=0}^{7} AS(i)p_{T_{SC,B2}/8}\left(t - iT_{SC,B2}/8 - lT_{SC,B2}\right) \\ sc_{B2-P}(t) &= \sum_{l=-\infty}^{\infty}\sum_{i=0}^{7} AP(i)p_{T_{SC,B2}/8}\left(t - iT_{SC,B2}/8 - lT_{SC,B2}\right) \end{aligned}$$

where $p_{T_{SC,B2}/8}(t)$ is given by

$$p_{T_{SC,B2}/8}(t) = \begin{cases} 1 & 0 \le t < T_{SC,B2}/8 \\ 0 & \text{others} \end{cases}$$

where the coefficients AS_i and AP_i are according to Table 11.1.

The subcarrier waveform for a duration of one cycle is shown in Fig. 11.1.

The baseband complex signal $S_{B2}(t)$ of AltBOC can be described as an 8-PSK signal equivalently. The corresponding constellation diagram is shown in Fig. 11.2.

$$S_{B2}(t) = \exp\left(j\frac{\pi}{4}k(t)\right) \quad \text{with} \quad k(t) \in \{1, 2, 3, 4, 5, 6, 7, 8\}.$$

Table 11.1 AltBOC subcarrier coefficients

i	0	1	2	3	4	5	6	7
$2AS(i)$	$\sqrt{2}+1$	1	-1	$-\sqrt{2}-1$	$-\sqrt{2}-1$	-1	1	$\sqrt{2}+1$
$2AP(i)$	$-\sqrt{2}+1$	1	-1	$\sqrt{2}-1$	$\sqrt{2}-1$	-1	1	$-\sqrt{2}+1$

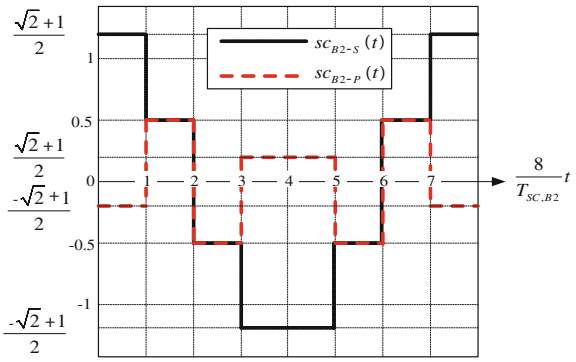


Fig. 11.1 AltBOC subcarrier waveforms

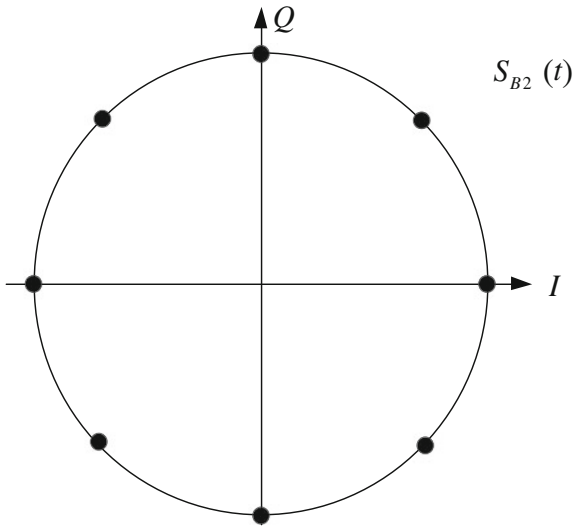


Fig. 11.2 Diagram of the 8-PSK phase state of AltBOC signal

11.2.2 TD-AltBOC

The data components and pilot components of upper and lower band are transmitted with time divided mode for TD-AltBOC(15, 10) modulation, its timing relationship is shown in Fig. 11.3.

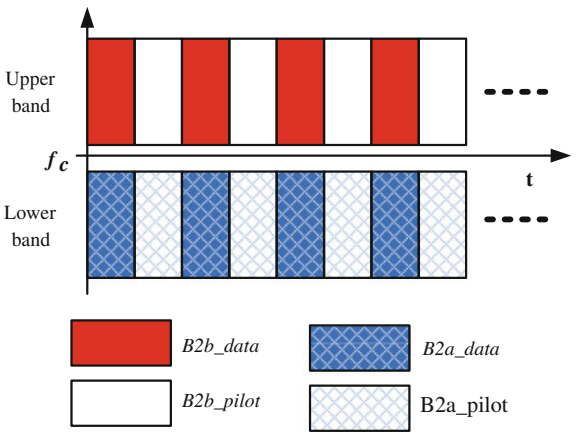


Fig. 11.3 Transmission timing relationship of TD-AltBOC signal components

The baseband signal of TD-AltBOC(15, 10) can be mathematically expressed as

$$\begin{aligned} s(t) = &\left[S_{B2a\_data}(t) + S_{B2a\_pilot}(t)\right]\left[SC_{B2,\cos}(t) - jSC_{B2,\sin}(t)\right] \\ &+ \left[S_{B2b\_data}(t) + S_{B2b\_pilot}(t)\right]\left[SC_{B2,\cos}(t) + jSC_{B2,\sin}(t)\right] \end{aligned}$$

where $SC_{B2,\cos}(t)$ and $SC_{B2,\sin}(t)$ are sine subcarrier and cosine subcarrier respectively and they are defined as following equations, frequency of these subcarriers f_{SC_B2} are both 15.345 MHz

$$\begin{aligned} SC_{B2,\cos}(t) &= \text{sign}(\mathbf{cos}(2\pi f_{SC\_B2}t)) \\ SC_{B2,\sin}(t) &= \text{sign}(\mathbf{sin}(2\pi f_{SC\_B2}t)) \end{aligned}$$

It's worth noting that B2a and B2b signals can be regarded as two separate BPSK(10) signals with the carrier frequency at 1176.45 and 1207.14 MHz respectively when they are solely received, for each BPSK(10) signal the corresponding data components and pilot components are time division multiplexed.

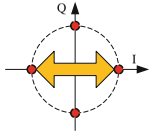
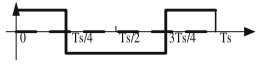
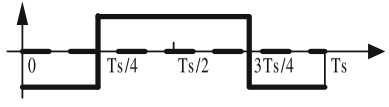
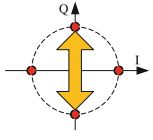
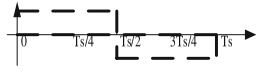
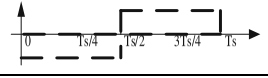
The baseband complex signal $s_{B2}(t)$ of TD-AltBOC(15, 10) can be described as a QPSK signal equivalently, its phase state is shown in Table 11.2.

In Table 11.2 T_s is the subcarrier cycle which equals to $1/f_{SC\_B2}$, when current time slot corresponds to odd numbered chip, we have $S_a = S_{B2a\_data}$ and $S_b = S_{B2b\_data}$, otherwise we have $S_a = S_{B2a\_pilot}$ and $S_b = S_{B2b\_pilot}$. The waveform described in solid line corresponds to in-phase component and waveform in dashed line corresponds to quadrature component.

11.2.3 ACED

The power ratio of pilot to data components of upper band or lower band is adjustable for ACED, the output of ACED multiplexer can be expressed as a 12-PSK signal when pilot to data power ratio is set to 3:1

Table 11.2 Constellation diagram and waveforms of TD-AltBOC modulation

S_a	S_b	Constellation	Waveforms
1	1		
−1	−1		
−1	1		
1	−1		

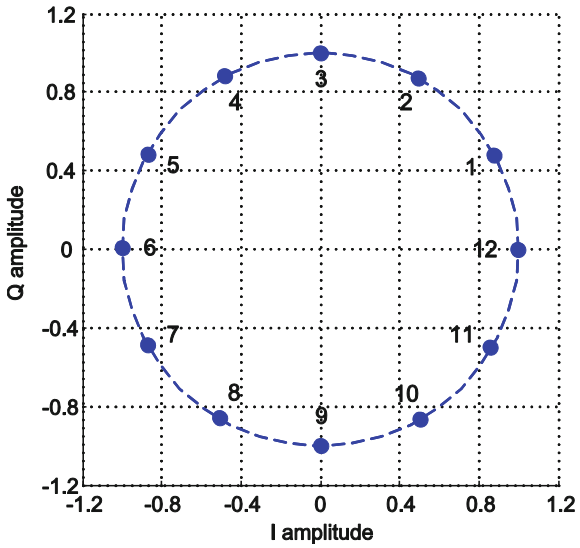


Fig. 11.4 Constellation diagram of ACED signal

$$\tilde{s}_l(t) = \exp\left(j\frac{\pi}{6}k(t)\right)$$

where $k(t) \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, the constellation diagram is shown in Fig. 11.4.

The normalized power spectrum density curves of these three modulation methods are shown in Fig. 11.5, it can be observed that the spectral shape of these three modulation methods are basically the same within the transmission bandwidth, therefore their performances in ranging accuracy, anti-interference and anti-multipath are also basically the same. Considering the power assignment of ACED(15, 10) is different from AltBOC(15, 10) and TD-AltBOC(15, 10), their performances in carrier tracking, demodulation (which further affects the TTFF) and application complexity will show diversity, the following part will focus on the analysis of such performances.

11.3 Performance Evaluation of Interoperable Signals on B2 Band

11.3.1 Ranging Accuracy

The Gabor bandwidth of the three modulation methods and single side band signal is shown in Fig. 11.6 where we can find out that the Gabor bandwidth of AltBOC, TD-AltBOC and ACED are equal in the transmission bandwidth of 71.61 MHz,

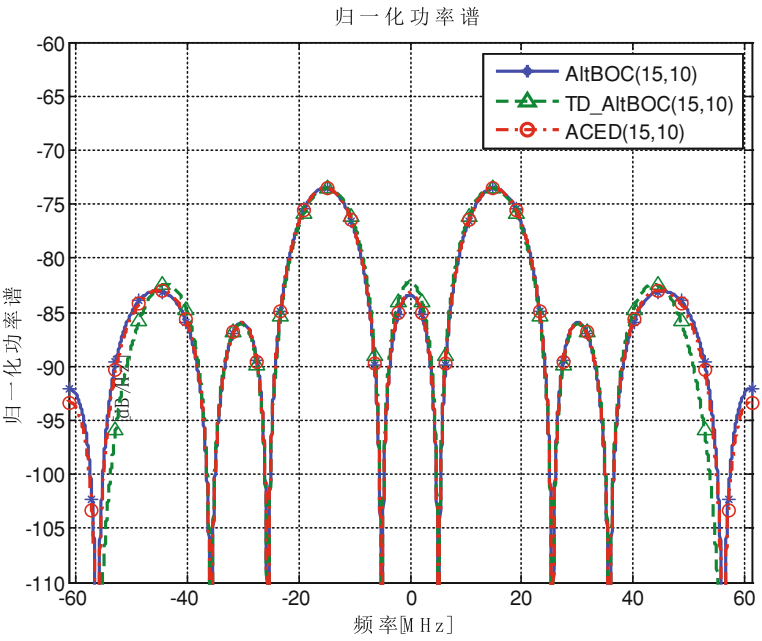


Fig. 11.5 PSD curves of three modulation methods

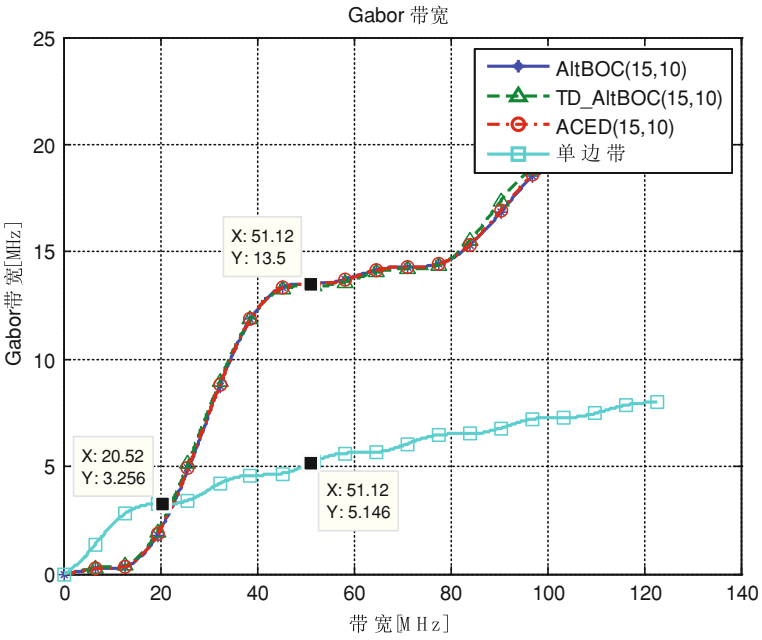


Fig. 11.6 Gabor bandwidth of three modulation methods

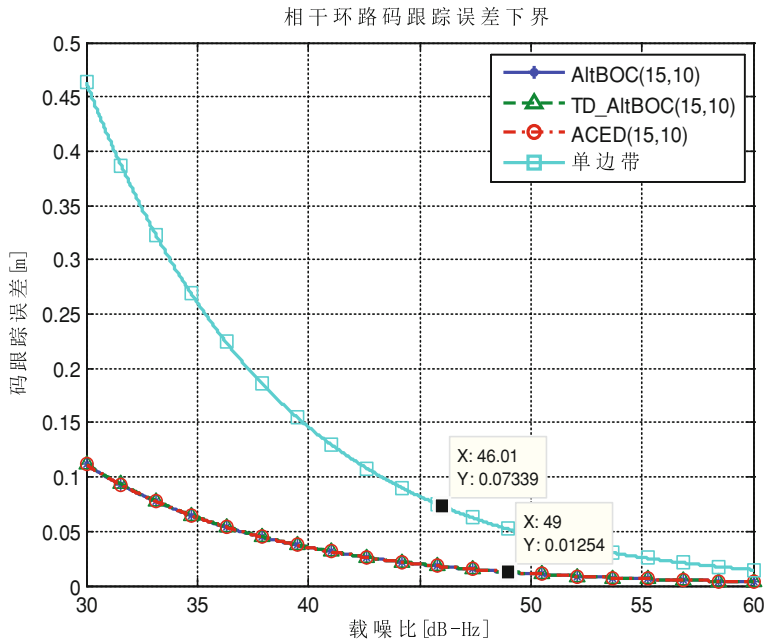


Fig. 11.7 Cramer-Rao lower bound

which means their performance in code tracking accuracy is the same. For receiving bandwidth less than 51.15 MHz a rapid raise of Gabor bandwidth value can be observed with the increment of receiving bandwidth, while for receiving bandwidth between 51.15 and 71.61 MHz the raise of Gabor bandwidth is relatively slow. Therefore, for the receiving of B2 signal with double side band, a relatively high ranging accuracy can be achieved when only main lobes are retained; Gabor bandwidth of single side band signal is far less than double side band, but it's equivalent to BPSK signal for the receiving process, this makes single side band receiving suitable for users sensitive to cost because of the low demanding of cost and power.

The Cramer-Rao lower bounds and code tracking mean square error of EMLP loop for the double side band receiving of the three modulation methods and for the single sideband receiving are shown in Figs. 11.7 and 11.8 respectively, where the receiving bandwidth for double side band receiving is 51.15 MHz and for single side band receiving is 20.46 MHz. As analyzed above, the Cramer-Rao lower bounds and code tracking mean square error of EMLP loop for AltBOC, TD-AltBOC and ACED are equal, while the code tracking mean square error for single side band receiving is always greater than double side band signal. When CNR is 46 dB-Hz, code tracking mean square error for double and single side band receiving is 0.01 and 0.07 m respectively.

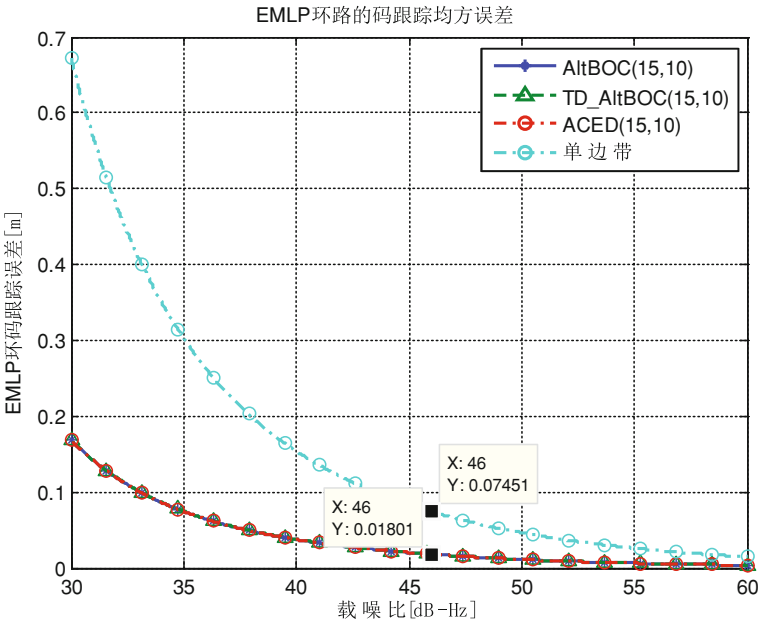


Fig. 11.8 Code tracking MSE of EMLP loop

11.3.2 Sensitivity

11.3.2.1 Acquisition Threshold

When receiving bandwidth is set to 50 MHz, the in-band power of the three methods takes up 77 % of total power. When only pilot channel is acquired, for AltBOC and TD-AltBOC, the in-band power of pilot channel takes up 39 % of total power, for ACED, that value is 58 %. In that assumption, ACED is better than AltBOC and TD-AltBOC in acquisition performance for both single side band and double side band receiving.

11.3.2.2 Tracking Threshold

Figure 11.9 shows the thermal noise jitter of the three modulation methods in different tracking modes where coherent integration time is 1 ms, bandwidth of carrier tracking loop is 15 Hz. Due to the fact that thermal noise jitter of carrier tracking is only determined by power assignment if for the same coherent integration time and tracking loop bandwidth, thermal noise jitter of carrier tracking of AltBOC and TD-AltBOC in Fig. 11.9 is the same, joint tracking can lead to better tracking performance considering only half power is assigned to pilot channel; the pilot to data

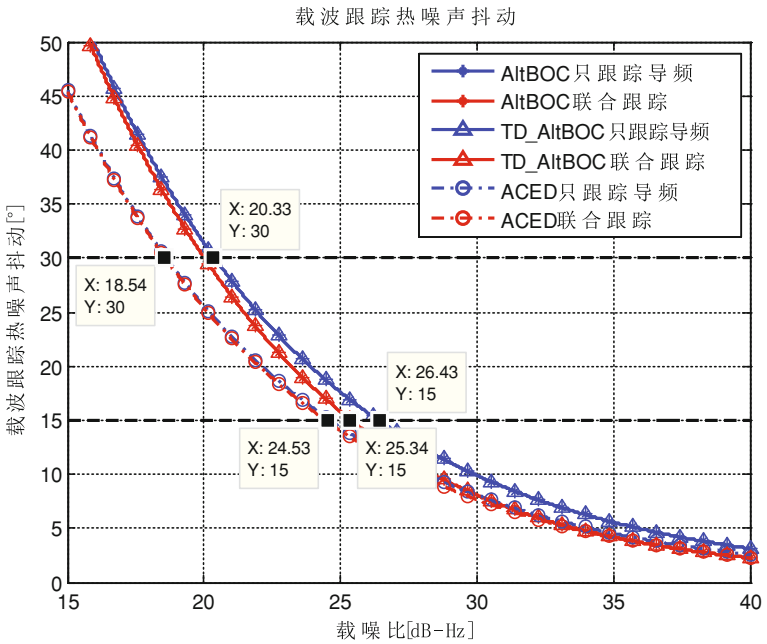


Fig. 11.9 Carrier thermal noise jitter and carrier tracking threshold of three modulation methods

Table 11.3 Carrier tracking threshold for different power ratio

Power ratio	Threshold	
	Carrier tracking threshold (dB-Hz)	
	Pilot only	Jointly
1:1	26.4	25.3
1:3	24.5	24.5

power ratio for ACED method is 3:1, so the thermal noise jitter of joint tracking and pilot-only tracking is similar. Table 11.3 gives out the tracking threshold values corresponding to the situation when thermal noise jitter of carrier tracking is 15°, for pilot-only and joint tracking, the tracking threshold corresponding to 1:3 power ratio is 1.9 dB and 0.8 dB lower than 1:1 power ratio respectively.

11.3.2.3 Demodulation Threshold

Demodulation threshold is irrelevant to modulation method and single or double side band receiving, it's only related to the power ratio of data component and information rate, more power ratio of data component means lower bit error ratio for the same CNR; higher information rate means higher bit error ratio for the

Table 11.4 Demodulation thresholds for different power ratio and different information rate

Power ratio	Information rate		
	50 bps	100 bps	500 bps
1:1	25.6	28.6	35.6
1:3	28.6	31.6	38.6

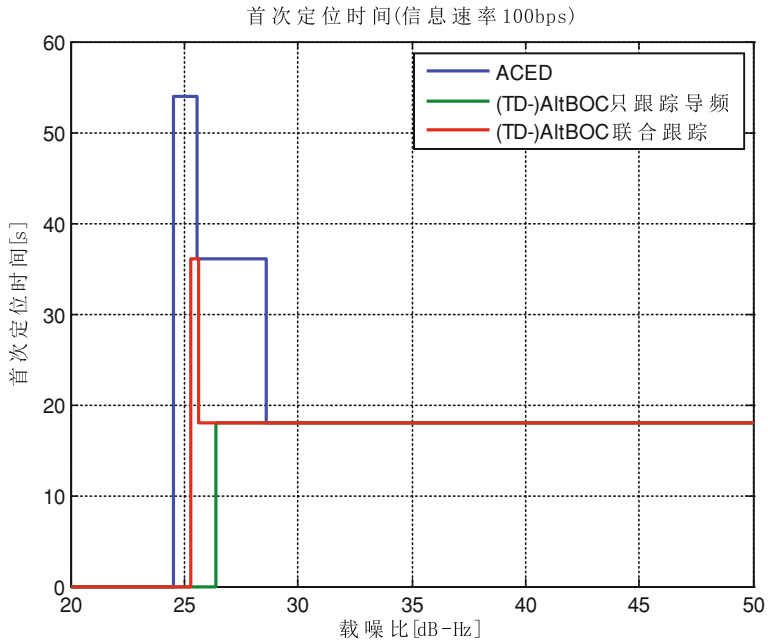


Fig. 11.10 TTFF performance when information rate is 50 bps

same CNR. The demodulation thresholds corresponding to the situation when bit error ratio is 10^{-5} for the power ratio of 1:1 and 1:3 are given in Table 11.4, where information rate is 50, 100 and 500 bps respectively, coding length is 2 and coding gain is 7 dB. It's seen that demodulation threshold for the power ratio of 1:3 is 3 dB higher than 1:1.

11.3.3 Time to First Fix

In the same way as the study of carrier tracking threshold and demodulation threshold in last section, the Time to First Fix (TTFF) performances of different methods in various CNR are shown in Figs. 11.10, 11.11 and 11.12 where TD-AltBOC is not included because its TTFF performance is the same as AltBOC. Assume that the navigation message structure is similar to that of GPS L1C signal [5], the TTFF value for ACED is 18 s when information rate is 50 bps, for the

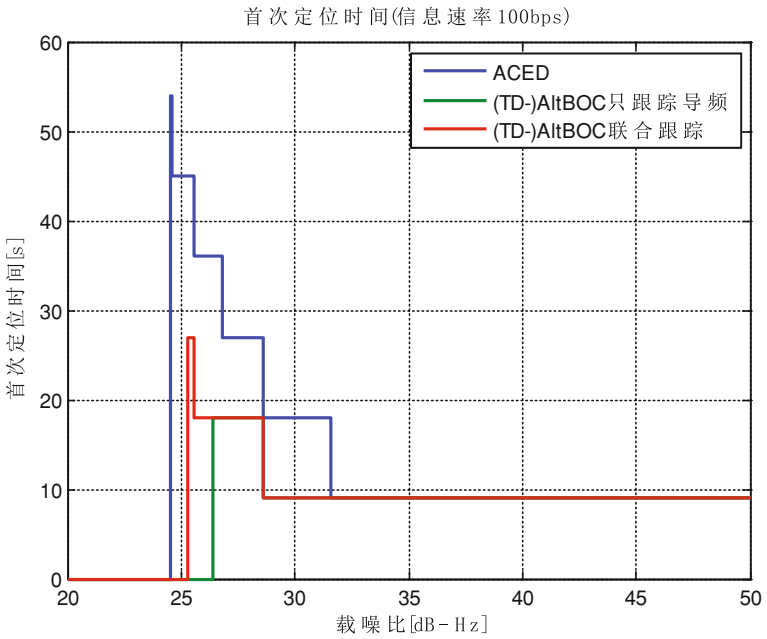


Fig. 11.11 TTFF performance when information rate is 100 bps

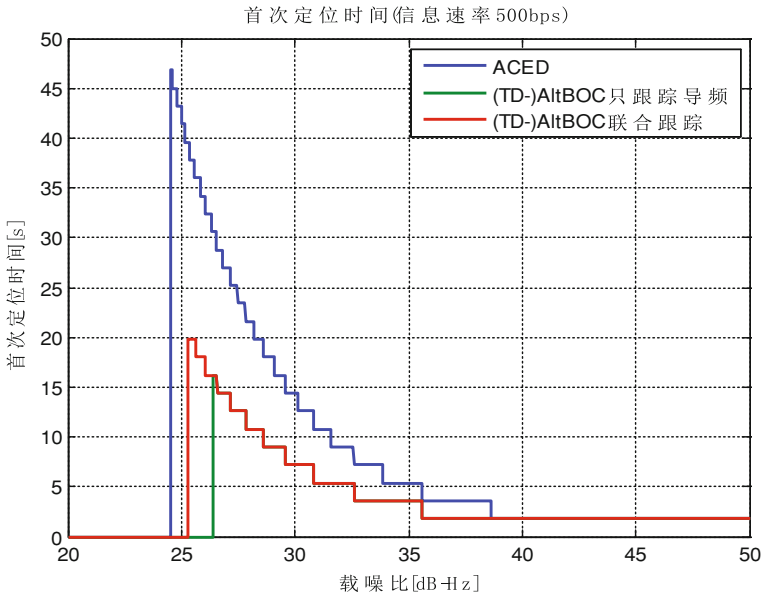


Fig. 11.12 TTFF performance when information rate is 500 bps

information rate of 100 and 500 bps, the TTFF will be 9 s and 1.8 s respectively; CNR can be increased by means of coherent integration if the CNR of received signal is lower than demodulation threshold, the increment is 10 log(n) dB for n times of coherent integration. A significant improvement of TTFF can be made by setting the power ratio 1:1 rather than 1:3 especially for the case of high information rate where large numbers of coherent integration times are needed for the power ratio of 1:3. For server cases with remarkable rise and fall of CNR, the improvement of CNR by coherent integration will be much worse and no longer obey the rule of 10 log(n) dB.

11.3.4 Anti-interference Performance

The anti-interference capacities of the three modulation methods are shown in Fig. 11.13 where we can conclude that they share the same anti-interference capacity and have the same performance in anti-interference.

11.3.5 Anti-multipath Performance

The multipath error envelopes and average multipath errors of the three modulation methods on B2 band are shown in Figs. 11.14 and 11.15 respectively. It's seen that their multipath error envelopes and average multipath errors are the same, therefore they share equivalent anti-multipath performance.

11.3.6 Multiplexing Efficiency

Multiplexing efficiency for AltBOC and ACED signal is 81.06 %, while for TD-AltBOC the multiplexing efficiency is 100 %. Therefore TD-AltBOC owns higher multiplexing efficiency and lower multiplexing loss.

11.3.7 Cross Correlation Property

According to statistical results, the mean square values of cross correlation functions of the three modulation methods are the same, but the maximum cross correlation values of AltBOC and ACED are lower because the phase division method they use makes the distribution of cross correlation values more concentrated. The maximum cross correlation value of TD-AltBOC signal is 2.3 dB higher than AltBOC and ACED signal because time division multiplexing makes

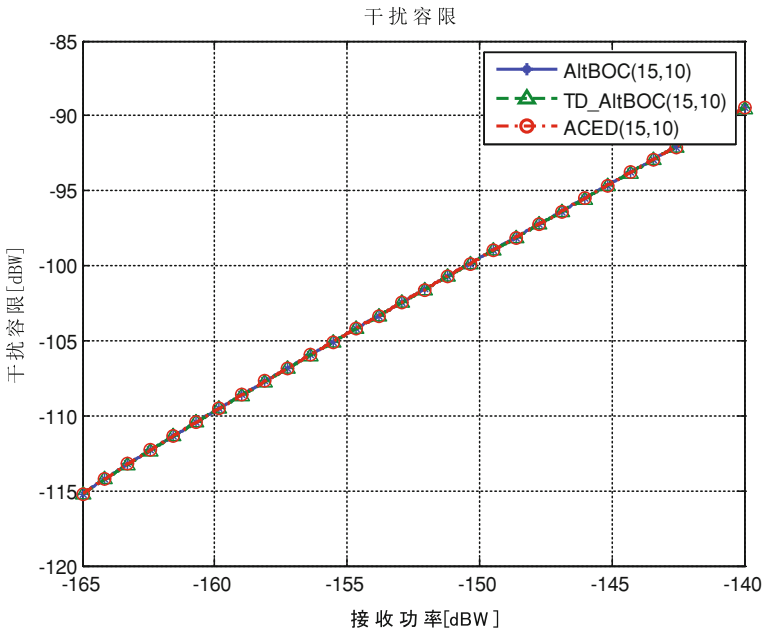


Fig. 11.13 Anti-interference capacities of three modulation methods

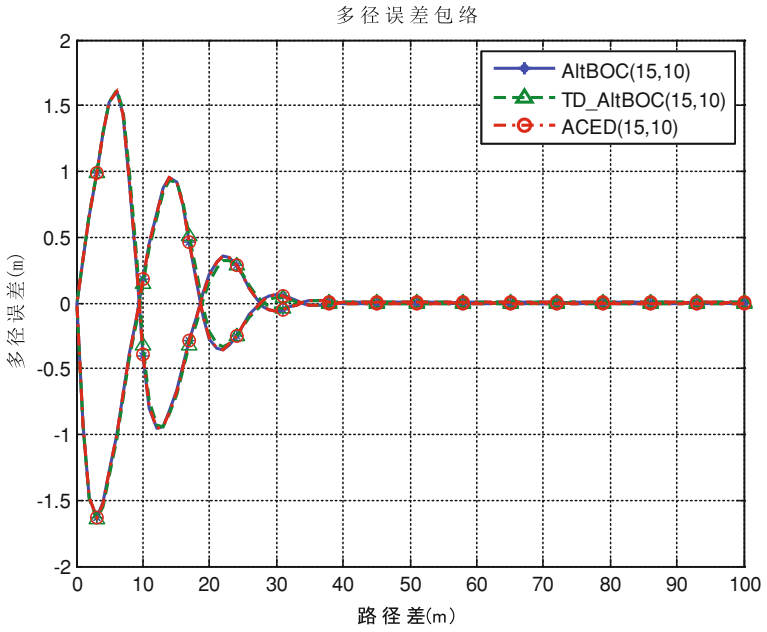


Fig. 11.14 Multipath error envelopes of three modulation methods

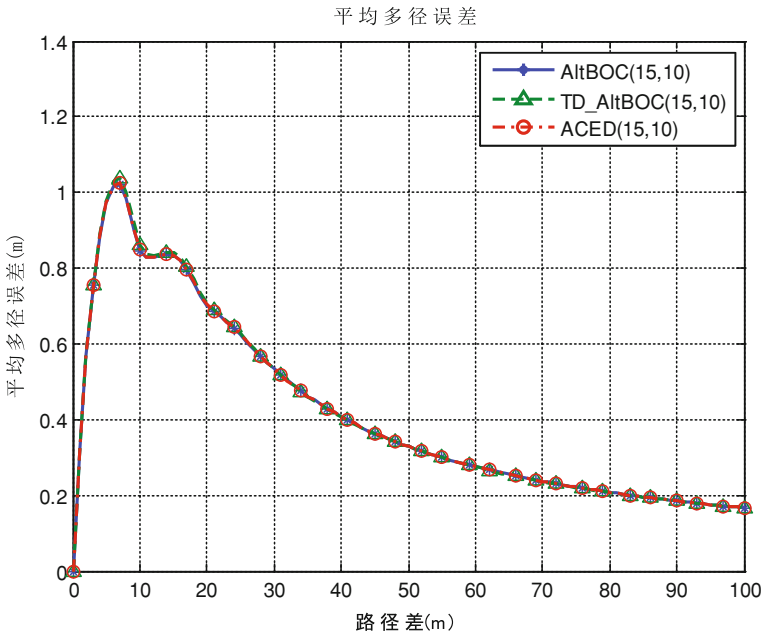


Fig. 11.15 Average multipath errors of three modulation methods

its code length half of the latters under the premise of the same code cycle. Considering that the code rate of B2 signal is 10.23 MHz and the spectral separation coefficient is quite low (lower than −70 dB/Hz), the effect of multi-address interference is much lower than noise, the influence of the difference of cross correlation property on multi-address performance will appear only when the CNR of multi-address signal is higher than 60 dB-Hz.

11.3.8 Complexity and Power Dissipation

11.3.8.1 Complexity

Because of the time division method in TD-AltBOC modulation, a hardware time division module is necessary in both satellite signal generation and signal receiving process; the sampling rate of signal generation in baseband is 60*1.023 MHz. For AltBOC method, the receiving of its single side band signal is simple because it's equivalent to QPSK signal; the sampling rate of signal generation in baseband should reach 120*1.023 MHz. The receiving complexity of ACED signal is same to AltBOC signal; for the case when power ratio is 1:3, the sampling rate of signal generation in baseband is 180*1.023 MHz which is more demanding for satellite implementation.

11.3.8.2 Power Dissipation

There're many factors can affect the power dissipation, we'll analyze from the perspective of digital processing part in receiver because the RF front-end is universal for them. This paper takes two parameters into consideration: necessary logical resource numbers and operation times in process of receiver's reference signal generation for matched receiving and single side band receiving respectively.

1. Matched receiving
TD-AltBOC, AltBOC and ACED are equivalent to 4PSK, 8PSK and 12PSK modulation respectively, the reference waveforms of TD-AltBOC signal can be generated by I, Q logical operation circuits where 4 logical operations are needed in each subcarrier cycle, the necessary number of logical resources and operation times is comparatively low; the generation of AltBOC and ACED reference signal, which is usually realized by phase look up table, is implemented by computation method where large number of logical resources and operation times are needed. For the method of phase look up table, the table size of TD-AltBOC is 32 bits which is accessed 4 times per subcarrier cycle, an extra time division switch is needed where the switchover operates once per chip duration; the table size of AltBOC is 384 bits which is accessed 8 times per subcarrier cycle; the table size of ACED is 768 bits which is accessed 12 times per subcarrier cycle; the switchover of state and accessing of look up table is the main source of power dissipation in above process, we also have that the power dissipation of time division switchover is less than look up table accessing, so the demanding of logical resource numbers and power for TD-AltBOC, AltBOC and ACED is in ascending order.
2. Single side band receiving
AltBOC and ACED signal in single side band receiving mode is equivalent to a QPSK signal while for TD-AltBOC it's equivalent to a time division QPSK signal, so the reference signal generation of these three methods can be realized by the logical operation of in-phase and quadrature channels. An extra time division switch is still needed for TD-AltBOC signal, but it requires quite few logical resources and power dissipation; on the other hand, there is only half time that the signal state of correlator's in-phase or quadrature channel may change in the process of time division signal, the power dissipation of logical unit will be much less if its state keep the same instead of changing forward, therefore both channels of TD- AltBOC receiver are in low power dissipation state for half time, which, ensures the power dissipation of TD-AltBOC receiver lower than AltBOC and ACED.

11.4 Conclusion Remarks

According to the evaluation results, AltBOC, TD-AltBOC and ACED modulation methods can meet the requirement of interoperability with Galileo and GPS on B2 band, they present similar performance in ranging accuracy, anti-interference and

anti-multipath. Taking AltBOC signal as a reference, the advantage of TD-AltBOC signal is the high multiplexing efficiency, low complexity, low power dissipation and its self-innovative property, but it requires time division processing module and shows worse cross correlation property under server conditions; the advantage of ACED is the flexible pilot to data channel power ratio assignment, self-innovative property, better acquisition and tracking performance comparatively, but it shows the weakness in demodulation performance and time to first fix, and the complexity is high for satellite signal generation.

References

1. Ries L, Lestarquit L, Armengou-Miret E et al (2002) A software simulation tool for GNSS2 BOC signal analysis [C]. In: Proceeding of ION GPS 2002,24–27 Sept 2002, Portland, pp 2225–2239
2. Tang Z, Zhou H, Wei J et al (2011) TD-AltBOC: a new COMPASS B2 modulation. In: CSNC2011 electronic proceedings
3. Yao Z, Lu M (2013) ACED multiplexing and its application on BeiDou B2 band. In: CSNC2013 electronic proceedings
4. European Space Agency Galileo Joint Undertaking (2006) Galileo open service signal in space interface control document (OS SIS ICD) [S]
5. Betz JW, Cahn CR, Dafesh PA et al (2006) L1C signal design options. In: ION NTM 2006

Chapter 12
PRIS: A Novel Complex Sequence Construction Method

Zuping Tang, Jiaolong Wei, Hongwei Zhou, Tao Yan, Zhihui Zhou and Bo Qu

Abstract Unexceptionally, modern global Navigation Satellite System (GNSS) adopt Direct Sequence Spread Spectrum (DSSS) signal structure, in which, correlation property of Pseudo Random Noise (PRN) code play an important role of affecting its rejection performance to Narrow Band Interference (NBI) and Multiple Address Interference (MAI). In order to mitigate the cross-correlation performance degradation due to tiered code construction, this paper proposes a new method for constructing complex sequence of GNSS civil signal. Based on any primary sequence of length L, Pseudo Random Initiate State (PRIS) transformation is used to construct a complex sequence of length NL, and a feasible generator scheme is presented. Simulation and analysis results shows, the acquisition complexity of PRIS sequence is similar to tiered code, and the correlation property of PRIS sequence is 4–5 dB better than tiered code, which approaches the performance of random sequence of length NL. The proposed method is applicable to civil GNSS signal designing, and is helpful to improve the receiving performance in challenge environment without distinct additional complexity, which will be attractive to GNSS users around the world.

Keywords PRN code · Cross-correlation · PRIS sequence · GNSS

Z. Tang (✉) · J. Wei · T. Yan · Z. Zhou · B. Qu
Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan, China
e-mail: tang_zuping@mail.hust.edu.cn

H. Zhou
Beijing Institute of Tracking and Telecommunication Technology, Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_12,

12.1 Introduction

Unexceptionally, modern global Navigation Satellite System (GNSS) adopt Direct Sequence Spread Spectrum (DSSS) signal structure, in which, correlation property of Pseudo Random Noise (PRN) code play an important role of affecting its rejection performance to Narrow Band Interference (NBI) and Multiple Address Interference (MAI). Code Length, the first parameter should be decision in GNSS PRN code design, usually is hard to decision. With longer code length, property of both Auto Correlation and Cross Correlation will be better, but complexity of code acquisition will inevitably increase. With shorter code length, faster acquisition is met, but good correlation property is missing, which is harmful to improve acquisition sensitivity in challenge environment.

PRN code with short code length is applied in traditional GPS C/A signal and repeats 20 times within single data bit duration [1], which causes three main drawbacks. Firstly, short code gives bad correlation protection, a strong signal from one satellite can crosscorrelate with the codes that a receiver uses to track other satellites,which may block reception of weak signals in challenge environments. Secondly, additional bit synchronization has to be established before data recovery. Thirdly, navigation signal's PSD with line spectral characteristic makes it vulnerable to narrow band interference. Therefore, two improvements are made in the process of GPS modernization [2] and Galileo signal [3] design. Firstly, longer primary code, code lengths chosen by GPS are generally 10,230, and for Galileo, its shortest code length [4] is 4,092. Secondly, introduction of secondary code, aiming at removing ambiguity of bit synchronization and smoothing signal spectra. Secondary codes as long as 1,800 chips are adopted in GPS L1C pilot channel [2], which makes the length of composite code reach 1.8414×10^7. To a certain degree, composite code with tiered structure adopted by modernized GNSS signal has taken both code correlation property and acquisition complexity into consideration [5].

However, there is still degradation of performance for tiered composite code compared with PRN code which has the same code length. For a long period of time, this has been considered as the unavoidable expense for the trade-off between performance and cost. Now comes the question: can this expense dramatically been lowered? As the answer to such question, this paper presents a new PRN code family named PRIS sequence, which attempt to eliminate the hardness to tradeoff between acquisition complexity and correlation performance. Based on the original sequence of length L, using a special transformation name PRIS, can generate PRIS sequence of length NL. It has attractive characters, in acquisition complexity, PRIS sequence is similar to tiered code, and in correlation property, PRIS sequence is much better than tiered code and approach the performance of random sequence of length NL. If PRIS sequence is applied in civil GNSS signals, receiving performance in challenge environment could be dramatically improved without distinct additional complexity, which will be attractive to GNSS users around the world.

12.2 Problem Description

Primary code and secondary code, the chip width of which equals the repeated period of primary code, are composited to tiered code in modernized GNSS signal. Secondary code can be viewed by receiver as a certain data sequence which performs similar to navigation message modulation. Due to the modulation of secondary code, repetition period of PRN code becomes longer, spaces between spectra lines become closer, better cross-correlation property is presented and thus improve the GNSS signal's rejection performance to NBI as well as multi address performance.

However, cross-correlation property of tiered code composed by primary code with the length of L and secondary code with the length of N is significantly degraded compared with PRN code with the length of NL. Figure 12.1 shows the correlation property of tiered code and random code respectively, where the primary code of tiered code is a Gold sequence with the length of 1,023 bit, and the length of secondary code is 1,000 bit; code length of random code is 1,023,000 bit. According to correlation statistical characteristics, correlation properties of tiered has degraded 4–5 dB compared with random code.

In addition, tiered code plays an unsatisfactory role in code tracking performance. Because of the strict periodicity of primary code and that different satellites using the same code rate and primary code length, primary code phase differences of different satellite signals change slowly. In realistic receiver system, coherent integration period is usually set to be a primary code period, which means the influence of multi address interference to correlator output is determined by cross-correlation property of primary code. So, strong coherence exists between neighboring coherent integration output amplitude when receiving navigation signals with tiered code, the code tracking bias caused by multi address is difficult to filter out.

12.3 PRIS Sequence

12.3.1 Definition

PRIS is a composite sequence which can be composed based on arbitrary PRN code sequence. Assume there is a PRN code sequence c_0 with the length of L, a PRIS with the length of NL can be constructed as follow

$$PRIS(c_0) = [c_{k_1}, c_{k_2}, \ldots c_{k_N}] \tag{12.1}$$

where $\{c_{k_i}, i = 1 \cdots N\}$ represents sequences obtained by cyclic shifting original sequence c_0 by k_i bits, i.e.

where $k_i \in \{0, 1, \ldots, L-1\}$. $c_{k_i}(l) = \begin{cases} c_0(l + k_i) & 0 \le l < L - k_i \\ c_0(l + k_i - L) & L - k_i \le l < L \end{cases}$ (12.2)

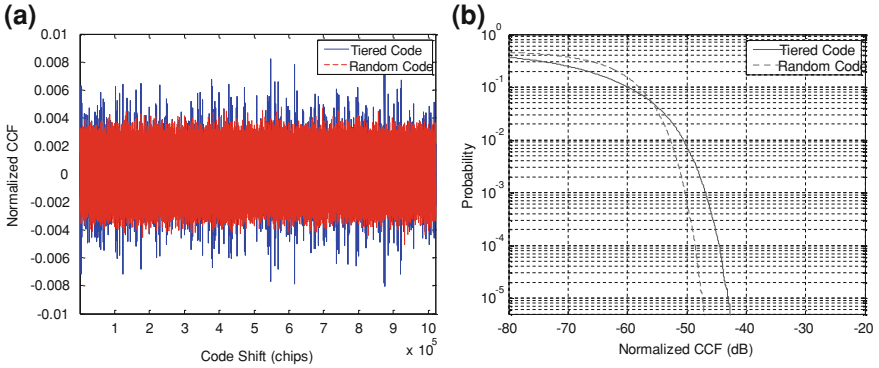


Fig. 12.1 Contrast of correlation property between tiered code and random code, **a** cross-correlation function, **b** cross-correlation statistical characteristics

It's clearly seen that PRIS sequence is composed of N sub-sequences which are generated by the same original sequence, the only difference among different sub-sequences is their initial phase. So, this is how the sequence gets its name: Pseudo Random Initial State (PRIS) sequence.

12.3.2 Generation Method

General generation method of PRIS sequence is shown in Fig. 12.2, where fc represents basic sequence generation clock, the frequency of which equals to the navigation signal code rate.

For the convenience of PRIS sequence generation, original sequence and pseudo initial phase sequence can be based on shift register. For the explanation of PRIS sequence generation, we choose Gold sequence as the original sequence and m sequence as the pseudo random initial phase to compose the PRIS sequence which is named Gold-m-PRIS sequence. The generation schematic diagram of Gold-m-PRIS sequence is shown in Fig. 12.3.

PRIS sequence in Fig. 12.3 can be described by three generator polynomials, an initial state and phase selection relation logic, their definitions are as follows

$$G_1 = 1 + a_{1,1}X + a_{1,2}X^2 + \cdots + a_{1,r}X^r \tag{12.3}$$

$$G_2 = 1 + a_{2,1}X + a_{2,2}X^2 + \cdots + a_{2,r}X^r \tag{12.4}$$

$$G_3 = 1 + a_{3,1}X + a_{3,2}X^2 + \cdots + a_{3,R}X^R \tag{12.5}$$

$$S_3 = s_{3,1}s_{3,2}s_{3,3}\cdots s_{3,R-1}s_{3,R}. \tag{12.6}$$

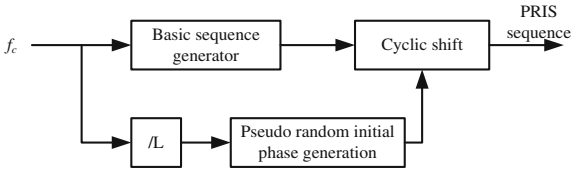


Fig. 12.2 PRIS sequence generation schematic diagram

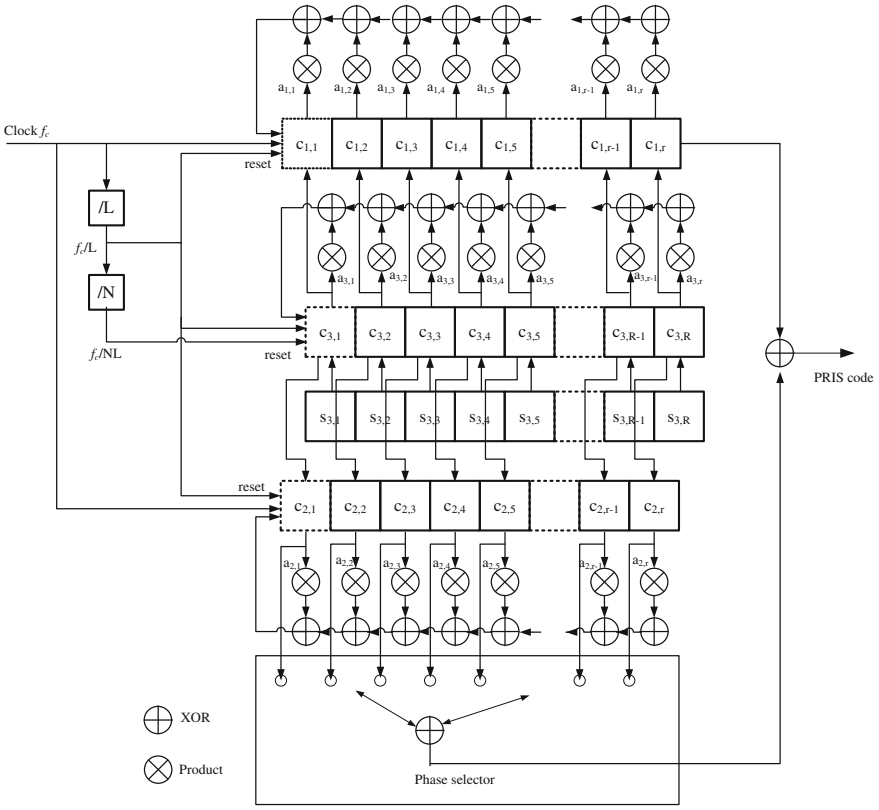


Fig. 12.3 Gold-m-PRIS sequence generation schematic diagram

Where G1 and G2 are the generator polynomials corresponding to original Gold sequence with the order r, which makes the maximum sequence length $2^r - 1$; G3 is the generator polynomial corresponding to m sequence with the order R, which makes the maximum sequence length $2^R - 1$; S3 is the initial phase of m sequence. Gold sequence is controlled and generated by clock with frequency fc equals to code rate, every L clock periods a reset is operated, by which the states of two

register groups C1 and C2 are set as the current state of m sequence resister. Which means, when the L + 1*th* rising edge arrives,

$$c_{1,1}c_{1,2}c_{1,3}\cdots c_{1,r} = c_{3,1}c_{3,2}c_{3,3}\cdots c_{3,r} \tag{12.7}$$

M sequence is controlled and generated by clock with frequency fc/L, i.e. it will update once for each Gold sequence period, and it will be reset for each N times of update, the initial state of register group C3 is defined by S3. Which means, when the N + 1*th* rising edge arrives

$$c_{3,1}c_{3,2}c_{3,3}\cdots c_{3,R} = s_{3,1}s_{3,2}s_{3,3}\cdots s_{3,R}. \tag{12.8}$$

For the convenience of determining the initial state of original sequence, the order R of m sequence should be selected obeying

$$R \geq r \quad \text{and} \quad 2^R - 1 \geq N\ . \tag{12.9}$$

When $R \neq r$, the specific mapping of the initial state of the reset original sequence and the m sequence register state can be user-defined.

12.4 Performance Analysis

For the purpose of comparing the performance of tiered code and PRIS code family, we choose GPS C/A code as the common original sequence with the length of 1,023 bit, the length of composite code is 1,023,000 bit.

12.4.1 Auto-Correlation Property

Figure 12.4 compares the auto-correlation property of tiered code, PRIS code and random code. From (a) it's shown that for almost every case, the side lobes of PRIS sequence are far less than tiered code. An insight of auto-correlation function property is shown in (b) where the phase difference is below 11,000 code chips. Recalling the fact that PRIS code and tiered are generated based on the same original sequence, this makes their auto-correlation property mainly determined by auto-correlation function of the original sequence and the amplitude tends to decrease linearly when code phase difference is less than original sequence code length. When code phase difference is equal to or greater than original sequence code length, the auto-correlation function of PRIS sequence presents a noise-like property similar to random code, whereas for auto-correlation function of tiered code significant peaks appear in integral multiple positions of original sequence code length. The main reason can be explained as: primary code sequence of tiered code is repeated periodically, when the code phase difference is an integral multiple of primary code length, auto-correlation property is totally determined by secondary code since the primary code is perfectly related.

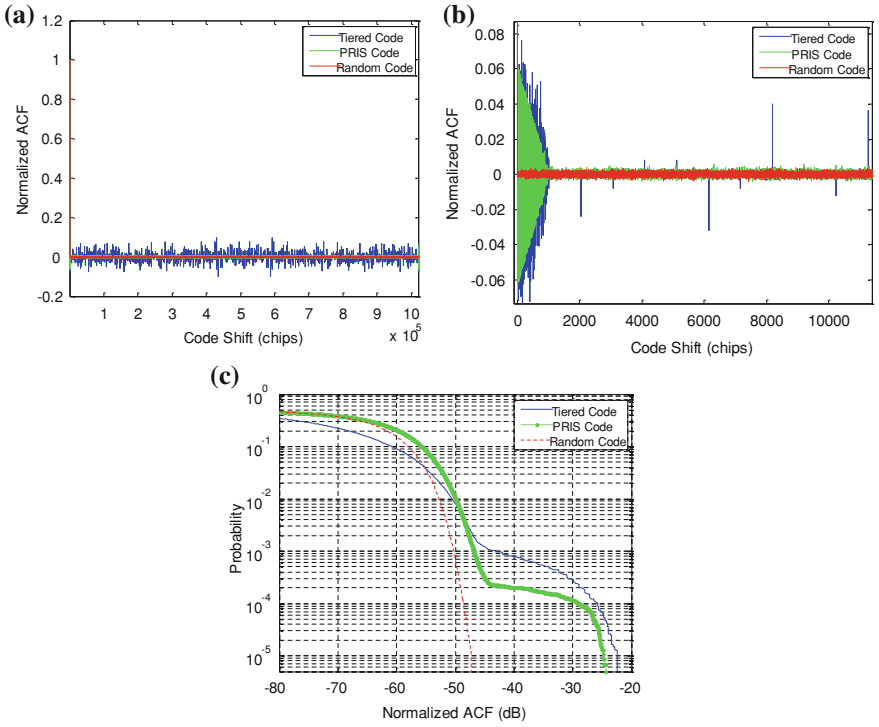


Fig. 12.4 Comparison of auto-correlation property, **a** normalized ACF (complete), **b** Normalized ACF (detail), **c** Statistical probability

For the three code sequences, statistical probability of auto-correlation side lobe amplitude is shown in Fig. 12.4c, where the ordinate means the probability of auto-correlation side lobe exceeding the corresponding amplitude given by abscissa. It can be seen that for PRIS sequence, the probability of appearing relatively large side lobes is far less than tiered code which is beneficial for weak signal acquisition. The auto-correlation property of PRIS sequence still shows a great distance compared with random code since its auto-correlation property is mainly determined by original sequence when code phase is less than L.

12.4.2 Cross-Correlation Property

Figure 12.5 compares the cross-correlation property of tiered code, PRIS code and random code. The complete cross-correlation function curves of three sequences are shown in (a). It can be seen that the maximum cross-correlation amplitude is much greater than PRIS sequence and random code. Partial property of cross-correlation function given in (b) show that cross-correlation amplitude of tiered

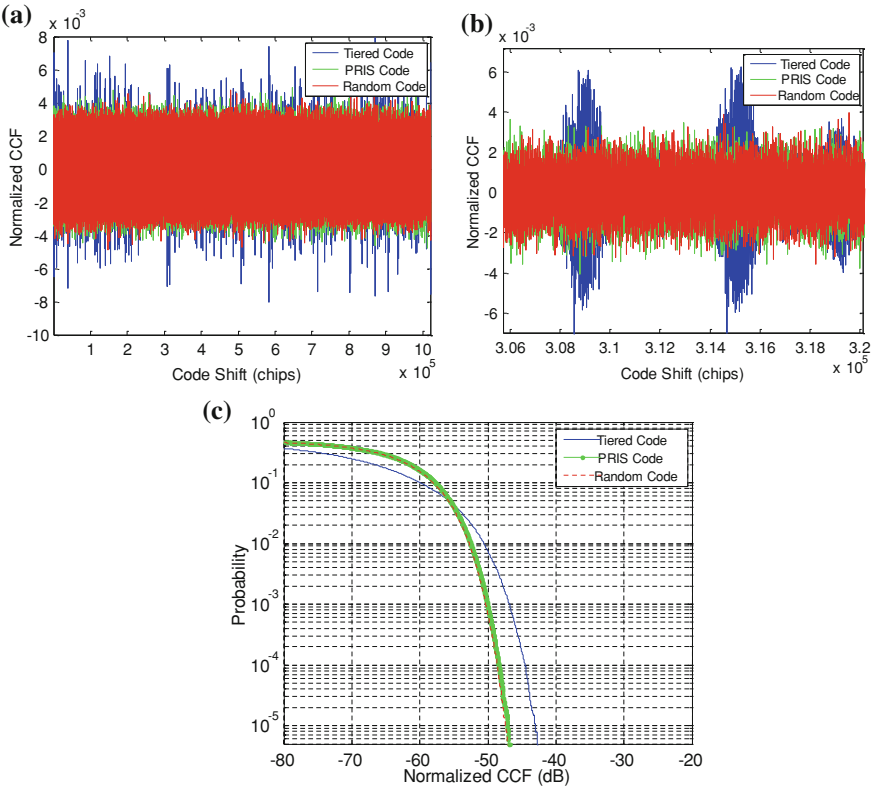


Fig. 12.5 Comparison of cross-correlation property, **a** cross-correlation function (complete), **b** cross-correlation function (partial), **c** cross-correlation statistical characteristic

code fluctuates significantly, while for PRIS sequence and random code it represents a noise-like property with flat amplitude. The probability of code sequences' normalized cross-correlation value exceeding certain amplitude is given in (c). From the probability statistical curve we can tell that cross-correlation property of PRIS sequence is almost the same as random code and outperforms tiered code, the greatest cross-correlation value of PRIS sequence is 4–5 dB lower than tired code. To sum up, the code isolation degree of PRIS sequence is better than tiered and that guarantees PRIS sequence better multi address performance.

12.4.3 Acquisition Complexity

Complexity of PRN code acquisition varies with different realization mode significantly. For the fair comparison of code sequence's influence on acquisition complexity, we take total the number of possible PRN code phases as the

evaluating index. Assume that no prior information about PRN code phase is offered to user, receiver has to search every possible code phase, and larger search space means higher acquisition complexity.

For tiered code, the number of possible primary code phase is L, and for secondary code, this number is N, the possible number of composite sequence code phase is L + N because of the strict timing relationship between primary and secondary code. For PRIS sequence, the number of possible original sequence code phase is L, initial phases of original sequences will change for N original sequence periods, but the discipline of change is known to receiver as prior information, so the possible number of code phase is also L + N. For random code, the number of possible code is equal to the code length N*L.

We can conclude that acquisition complexity for PRIS sequence and tired code is similar and that means PRIS sequence is easier to acquire than long random sequence. It's worth mention that existing acquisition methods are no longer suitable for PRIS sequence considering its original code sequences don't show up by simply repeating itself.

12.4.4 Code Tracking Performance Simulation

As mentioned in Sect. 12.2, for tiered code, slowly changing code tracking bias occurs in situation of multiple access interference which is difficult to filter out because of the simple repetition of primary code, thus code tracking performance becomes degraded. By means of simulation, this section compares the code tracking performance of PRIS sequence and tiered in presence of strong multiple access interference.

The simulation results of code tracking performance in situation of weak signal and strong MAI is shown in Figure 12.6. Following assumptions are made for the simulation: signal processed is from pilot channel (no data bit modulation), carrier frequency is 1575.42 MHz, modulation method is BPSK(1), front-end bandwidth is 2 MHz, CNR is 30 dB-Hz, multiple access signal power to desired signal power ratio is 30 dB, Doppler shift between multiple access signal and desired signal is 1,000 Hz, coherent integration time is 1 ms, original sequences for PRIS sequence and tiered sequence are both GPS C/A code, length of composite code is 1,023,000 bit.

According to the simulation results, a lock-lose is observed for navigation signal with tiered code at the 5th second of our simulation, while navigation signal with PRIS sequence is able to keep tracking steadily. By furthermore analysis we can explain the reason for the lock-lose of tiered code tracking: the code phase difference of multiple access signal and desired signal changes with the rate of 0.65 chip/s, the cross-correlation amplitude of multiple access signal and desired signal may exceed auto-correlation amplitude of desired signal when the code phase difference appears to be the position causing large cross-correlation amplitude, in this situation, code tracking loop will lock on the undesired cross-correlation peak and gradually get

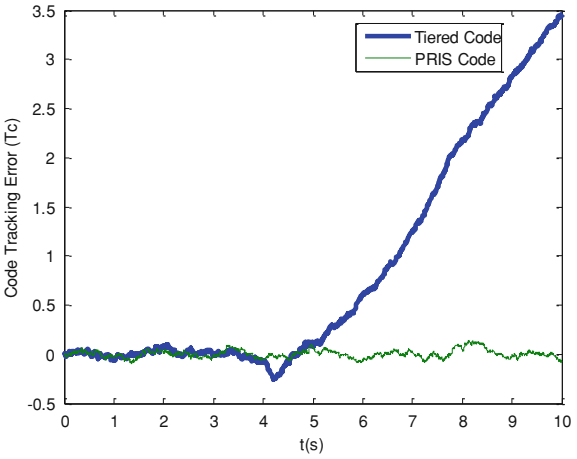


Fig. 12.6 Simulation results of code tracking simulation

fooled away due to the slip of code phase difference. Let's reconsider signal with PRIS sequence, cross-correlation amplitude won't keep a high value because the original sequence code phase of desired signal and MAI is always changing pseudo randomly, the behavior of MAI is closer to noise and won't lead to such fooling away problem.

12.5 Conclusion

In this paper we proposed a new PRN code named PRIS sequence, and provided a realization method for a specific instance. PRIS sequence archives the desired properties of civil signal, i.e. favorable cross-correlation property and low acquisition complexity. In auto-correlation property, the probability of appearing large auto-correlation side lobe for PRIS sequence is far less than tiered code; in cross-correlation property, the maximum cross-correlation of PRIS sequence is 4–5 dB lower than tiered code and reaches the level of random code; in acquisition complexity, PRIS sequence is similar to tiered code and far less than random code if for the same code length. Code tracking performance simulation shows that PRIS sequence offers more robust tracking performance than tiered code. PRIS sequence can be an optimized scheme for civil navigation signal.

References

1. Kaplan ED, Hegarty CJ (2005) Understanding GPS: principles and applications. Artech house, Boston
2. Rushanan JJ (2007) The spreading and overlay codes for the L1C signal. In: Proceedings of the 2007 national technical meeting of the institute of navigation. San Diego, pp 539–547

3. Wallner S, Avila-Rodriguez J-A, Hein GW (2007) Galileo E1 OS and GPS L1C pseudo random noise codes- requirements, generation, optimization and comparison. In: ION GNSS 20th international technical meeting of the satellite division, 25–28 Sept 2007, Fort Worth
4. European GNSS (Galileo) (2010) Open service signal in space interface control document. Publications Office of European Union, Luxembourg
5. Soualle F, Soellner M, Wallner S, Avila-Rodriguez J-A (2005) Spreading code selection criteria for the future GNSS Galileo. In: Proceedings of the European navigation conference GNSS 2005, Munich

Part II
Satellite Navigation Augmentation and Integrity Monitoring

Chapter 13
Application of Neural Network Aided Particle Filter in GPS Receiver Autonomous Integrity Monitoring

Ershen Wang, Tao Pang, Ming Cai and Zhixian Zhang

Abstract According to the measurement noise feature of GPS receiver and the sample impoverishment problem with the basic particle filter, an improved particle filter based on neural network algorithm is proposed. Using back-propagation (BP) neural network to adjust the particles with too high and too low weight, firstly, the larger weight particles are respectively splitted into two smaller weight particles. Then, abandoning the particles with very small weight, and adjust the particles with smaller weight by using the neural network. Therefore, the diversity of the sample particles is improved. The improved particle filter algorithm is combined with the likelihood ratio method for GPS receiver autonomous integrity monitoring (RAIM). By using the likelihood ratio as a consistency test statistic to achieve the fault detection, satellite fault detection is undertaken by checking the cumulative likelihood ratio of system state with detection threshold. By taking advantage of the relationship in statistical values between the total cumulative likelihood ratio and partial cumulative likelihood ratio, the number of fault satellite can be determined. Based on the real GPS raw data, the simulation results demonstrate that the improved particle filter under the conditions of non-Gaussian measurement noise can effectively detect and isolate fault satellite, and improve the performance of fault detection.

Keywords Global positioning system (GPS) · Receiver autonomous integrity monitoring(RAIM) · Particle filter · Neural network · Fault detection

E. Wang (✉) · T. Pang · M. Cai · Z. Zhang
School of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang, China
e-mail: wes2016@sau.edu.cn; wes2016@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_13,

13.1 Introduction

With the development of the global navigation satellite system (GNSS) and the growth of user performance requirements for GNSS service, for safety-critical applications of GNSS, such as aircraft and missile navigation systems, it is to be able to detect and exclude faults that could cause risks to the accuracy and integrity, so that the navigation system can operate continuously without any degradation in performance. Because it needs a long time for satellite fault monitoring to alarm through controlling the satellite navigation system itself, usually within 15 min to a few hours, that can't meet the demand of air navigation. As a result, to monitor the satellite fault rapidly in the client part, namely the Receiver Autonomous Integrity Monitoring (RAIM) has been researched a lot [1, 2].

Currently, RAIM algorithm includes two categories: one is to the snapshots algorithm use the current pseudorange observation, the other is the RAIM algorithm based on kalman filter. The snapshots algorithm mainly has Parity space method, the sum of least Squares of the Error (SSE) method, etc. [3]. In recent years, fault diagnostic method based on kalman filter has been widely used in nonlinear system fault diagnosis. And this algorithm requires measurement noise obey Gaussian distribution, and the actual measurement noise is difficult to strictly obey Gaussian distribution, the performance of the algorithm will downgrade much [4]. Because GNSS measurement error does not follow a Gaussian distribution perfectly [5], the kalman filter approach has to use an inaccurate error model that may cause performance degradation. Particle filter algorithm can better be suitable to any non-linear, non-Gaussian systems. And the algorithm has no any restrictions to the system process noise and measurement noise. The optimal state estimation can be easily gotten [6]. Therefore, compared to the kalman filter, the particle filter is more suitable for non-linear, non-Gaussian system in fault detection, navigation and positioning applications [7, 8]. But basic particle filter exists the degeneracy phenomenon and the sample impoverishment problem.

In order to solve these problem,the BP neural network algorithm is melted into basic particle filter. This paper describes the basic principle of the particle filtering algorithm and the particle filter algorithm based on neural network weights to adjust the for fault detection, establishes the GPS RAIM model and verify the validity of the algorithm through the measured GPS data.

13.2 BP Neural Network Aided Particle Filtering Algorithm

13.2.1 Basic Particle Filter Algorithm

Particle filters can deal with nonlinear/non-Gaussian dynamic systems as well as linear/Gaussian systems.

Let's consider the dynamic state space model below:

$$\boldsymbol{X}_k = \boldsymbol{f}_k(\boldsymbol{X}_{k-1}, \boldsymbol{v}_{k-1})$$
$$\boldsymbol{Z}_k = \boldsymbol{h}_k(\boldsymbol{X}_k, \boldsymbol{n}_k)$$

where x_k is a state vector, z_k is an output measurement vector, $f(.,.)$ and $h(.,.)$ are state transition function and measurement function respectively. v_{k-1} is the process noise vector independent of current state, and n_k the measurement noise vector independent of states and the system noise.

The basic principle of particle filter algorithm can be described as follows. First, based on the priori conditional distribution of system state vector the state space generate a group of random samples, these samples called particles, then based on the measurement value adjust constantly the particle weight and position of the particle distribution, modified initial priori conditional distribution. The algorithm is a recursive filtering algorithm, commonly used to handle non-gaussian and nonlinear systems state and parameter estimation [9].

A common problem of particle filter is the degeneracy phenomenon. A suitable measure of degeneracy of the algorithm is the effective sample size $\widehat{N}_{eff}$.

$$\widehat{N}_{eff} = \frac{1}{\sum_{i=1}^{N_s} \left(w_k^i\right)^2}$$

where w_k^i is the normalized weight.

Resampling particle filter can suppress weight degradation, but it introduces other problems, for example, the particles will be no longer independent and so on. And the particles that have high weights are statistically selected many times. This leads to a loss of diversity among the particles as the resultant sample will contain many repeated points. This problem, which is known as sample impoverishment, is severe in the case of small process noise.

13.2.2 BP Neural Network Combined with Particle Filter

ANN (Artificial Neural Networks) is a neural network simulation of animal behavior algorithm for parallel and distributed information processing, and the essence of this algorithm is to process information by adjusting internal connections among a large number nodes weight. BP neural network includes two aspects that are forward signal transmission and error back propagation, the topology model consists of input layer, hidden layer and output layer [10].

BP neural network algorithm is combined with basic particle filter, too high and too small weight particles can be adjusted to avoid the high-weight particles repeatedly sampled, and make low weight particles move to higher weight area, increasing the probability of being sampled, replacing the simple choice of

re-sampling. The detailed steps of weight adjusting particle filter based on neural network (NNWA-PF) can be described as the following.

Step 1 Initialization. According to the priori probability $p(x_0)$, the initial particles $\{x_0^i\}_{i=1}^{N_s}$ are generated,and each particle weight is $1/N_S$

Step 2 Weight updatation. At k time, according to the weight calculation formula to update the weights of particles, and normalized weights

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{0:k-1}^i, z_{0:k})},\ i = 1, 2, \cdots, N_s$$

$$\bar{\omega}_k^i = \omega_k^i \Big/ \sum_{i=1}^{N_s} \omega_k^i$$

Step 3 According to the particle weight, the particles are classified. Setting the k moment, particle weight classification threshold can be calculated by the following formula

$$A = sort(\bar{\omega}_k^1, \bar{\omega}_k^2, \cdots \bar{\omega}_k^{N_s})$$

$$\omega_{thr}(k) = A\left(round\left(N_s/3\right)\right)$$

If the particle x_k^i weight is less than the weight of classification threshold, that is $\bar{\omega}_k^i < \omega_{thr}(k)$, then the particles are classified as the low weight particle classification set, namely, $\{x_k^i\}_{l=1}^{N_L}$. Other else, the particles are classified as the high weight of particle set, namely $\{x_k^h\}_{h=1}^{N_H}$, when there are $N_L + N_H = N_s$.

Step 4 Weight splitting. Considering the high weight classification set of particles $\{x_k^h\}_{h=1}^{N_H}$. In the high weight particle set, the number of $q(q < N_H)$ too large weights particles splits into two smaller, weights halved particles. The total number of particles become large from N to $N + q = N_L + N_H + q$, in order to keep the total number of particles unchanged, abandoning low weight classification particle set $\{x_k^i\}_{l=1}^{N_L}$, from small to large values of the $q(q < N_L)$ weights and the corresponding particle

Step 5 Weight adjustment. After weights splitting step, the original weight matrix W becomes $\tilde{W}$. According to weight, by BP neural network adjustment, the number of q particles is taken from the matrix $\tilde{W}$

Step 6 After splitting weights and weights adjusted particle set, a new set of particles $\{\tilde{x}_k^{i'}, \bar{\omega}_k^{i'}\}_{i=1}^{N_s}$ is attained. Calculating the normalized importance weights and effective number of particles, if the effective number of particles is less than the threshold value, that is $\hat{N}_{eff} < N_{thr}$, then the new set of particles $\{\tilde{x}_k^{i'}, \bar{\omega}_k^{i'}\}_{i=1}^{N_s}$ re-sampling operation to obtain a set of particles after re-sampling $\{\tilde{x}_k^i, \bar{\omega}_k^i\}_{i=1}^{N_s}$. Otherwise, go to step 7

Step 7 State estimation. $\hat{x}_k = \sum_{i=1}^{N_s} \bar{\omega}_k^i \tilde{x}_k^i$

Step 8 State prediction. Using the state equation f predict unknown state parameters x_{k+1}^i, namely $x_{k+1}^i = f\left(\tilde{x}_k^i, v_k\right), i = 1, 2, \cdots, N_s$

Step 9 Time $k = k + 1$, go to Step 2.

13.3 Neural Network Aided Particle Filter for GPS RAIM

Fault detection means the failure of the system being monitored, and the fault isolation refers to classifying the type of system failure. GNSS RAIM must have an integrity monitoring system that contains two functions: 1) detection and exclusion of satellite faults and 2) estimation of the uncertainty of the position solutions. The system calculates decision variables and compares these to thresholds that have been set to satisfy the integrity requirements for the desired operation. If any decision variable exceeds the threshold, then the system concludes that the corresponding measurement has a fault, and excludes the detected fault.

The GPS dynamic system model is as follows:

$$X_k = F_{k-1} X_{k-1} + w_{k-1}$$

where, $X_k = [r_x, r_y, r_z, \Delta\delta]^T$ is three dimensional position and GPS-receiver clock offset, and F is a transition matrix which in the static case is an identity matrix [11]. The data of satellite coordinates (s_x^i, s_y^i, s_z^i), pseudo-range ρ^i and time error $\Delta\delta$ is got from the GPS receiver.

The problem of fault detection consists of making the decision on the presence or absence of faults in the monitored system. The GPS RAIM approach is designed. The failures to be detected affect only a subset of the system measurements. A reference system unaffected by failures is required for the GPS RAIM approach, which monitors the state estimate of GPS receiver [12, 13].

According to the coordinates (r_x, r_y, r_z) of the receiver, generate the initial set of N particles $\{x_0^A(\mathrm{i}):\mathrm{i} = 1, 2, \ldots N\}$ for main PF particle from the prior probability density function(PDF) $p(x_0)$,and the auxiliary PFs particles $\{x_0^A(\mathrm{i}):\mathrm{i} = 1, 2, \ldots, N\}$, $x_0^q(\mathrm{i}) = x_0^A(\mathrm{i})$ for i = 1, 2, …, N. The main PF processes all m measurements (m is total number of measurements), while the auxiliaries, process subset of measurements (m-1 measurements).

Repeating the following steps for each time k:

(1) State prediction. The particles of $\{x_0^A(\mathrm{i}):\mathrm{i} = 1, 2, \ldots N\}$ and $\{x_0^q(\mathrm{i}):\mathrm{i} = 1, 2, \ldots, N\}$ are introduced into the system state equation, using the formula (1) to obtain particles predicted values $x_{k|k-1}^A(i)$ and $x_{k|k-1}^q(i)$.
(2) Calculate the particles weight. Take the predicted values of the particles

$x^A_{k|k-1.}(i)$, $x^q_{k|k-1.}(i)$ and the i-th satellite position coordinates s^i_x, s^i_y, s^i_z and the time error $\Delta\delta$, and so on into the system measurement equation to obtain the predicted i-th satellite pseudo-range value ρ^{*i}. Take the ρ^{*i} and pseudo-range measurement value ρ^i into the weight calculation formula and normalize them to obtain the normalized particle weights $\tilde{\omega}^A_k(i)$ and $\tilde{\omega}^q_k(i)$.

(3) Likelihood evaluation. On receipt of the measurement y*k*, the likelihood of the predictive state samples from the main PF are evaluated as.

$$\tilde{\omega}^A_k(i) = p(z_k|x^A_{k|k-1}(i))$$

While the likelihood of the predictive state samples from the auxiliary PFs can be expressed as the following.

$$\tilde{\omega}^q_k(i) = p(z_k|x^q_{k|k-1}(i))$$

(4) LLR calculation. LLR $S^k_j(q)$ is computed by equation as follows.

$$S^k_j(q) = \sum_{r=j}^{k} ln \frac{\frac{1}{N}\sum_{\mathrm{i}=1}^{N} \tilde{w}^q_r(\mathrm{i})}{\frac{1}{N}\sum_{\mathrm{i}=1}^{N} \tilde{w}^A_r(\mathrm{i})}$$

(5) Decision function. Decision function for FD defined as the following equation:

$$\beta_k = \max_{k-U+1 \le j \le k} \max_{1 \le q \le Q} S^k_j(q)$$

(6) Fault detection.
If $\beta_k > \tau$ (the decision threshold value is τ), the fault alarm time is set to $t_a = t$ and jump to step (6), if $\beta_k > \tau$, then no fault, go to step (7).

(7) Fault isolation. In $k > t_a$, remove the accumulated LLR Q satellites largest subset of the satellite, namely $g = \arg \max_{1 \le d \le D} S^k_{t_a}(k > t_a)$ formula.g as failed satellite number, determining a failure satellite number of the satellite, the satellite in turn, can be isolated from the measured value.

(8) Status update. The importance weight $\omega^A_k(i)$ of the predictive state samples $\{x^A_{k|k-1.}(i): i = 1, 2, \ldots, N\}$ from the main PF and $\omega^q_k(i)$ of the predictive state samples $\{x^q_{k|k\text{-}1.}(i): i = 1, 2, \ldots, N\}$ from the auxiliary PFs are calculated as follows:

$$\omega^A_k(i) = \frac{\tilde{\omega}^A_k(i)}{\sum_{j=1}^{N} \tilde{\omega}^A_k(j)}, \ \omega^q_k(i) = \frac{\tilde{\omega}^q_k(i)}{\sum_{j=1}^{N} \tilde{\omega}^q_k(j)}$$

The samples $\{x_k^A(i):i = 1, 2, \ldots, N\}$ for the main PF and $\{x_k^q(i):i = 1, 2, \ldots, N\}$ for the auxiliaries are obtained by resampling $\{x_{klk\text{-}1}^A(i):i = 1, 2, \ldots, N\}$ and $\{x_{klk-1}^q(i):i = 1, 2, \ldots, N\}$ respectively. The particles of particle filter are updated.

13.4 Simulation Test and Results Analysis

Numerical simulations on GPS positioning illustrate the FDI performance of the proposed approach for GPS integrity monitoring.

13.4.1 Experiments Conditions

The experimental raw observation data are obtained by GPS receiver N220, the observation data including the satellite receiver position location information and pseudorange values. During the period of this experiment, there are 6 satellites used for PVT solution, the number of the satellite is 3, 15, 18, 19, 21, 26 respectively, and the corresponding pseudorange value can be expressed as $Y = (y_1, y_2, y_3, y_4, y_5, y_6)$. At the same time, the RCB-4H receiver monitors that the satellite is working normally. In order to simulate the fault when a satellite is failure, whether the algorithm could detection effectively, some error is intentionally added.

According to Banach fixed point theorem, an arbitrary n-dimensional to m-dimensional mapping can be achieved by a three-layer BP network, so a hidden layer BP neural network structure is selected. The initial weight of BP neural network is the state value of small weight particles, the initial input is the weight of small particles, neural network learning step length is 0.05, the number of neurons is 6, the incentive function is $1/(1 + e^{-x})$, the system observation equation is chosen as the learning sample function, the study process will be finished when the total error is less than 0.005.

13.4.2 Simulation Results and Analysis

To verify the feasibility and effectiveness of NNWA-PF algorithm used in GPS RAIM, firstly, without fault condition, the measurement data which contains no error has been applied in RAIM approach. The experimental results are shown in Figs. 13.1 and 13.2.

Figures 13.1 and 13.2 show respectively the decision function curve and accumulated LLR curve of NNWA-PF algorithm and PF algorithm FDI method

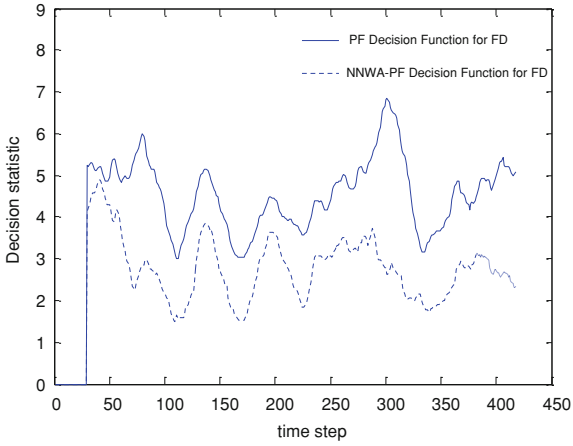


Fig. 13.1 Decision function for fault detection under nominal condition

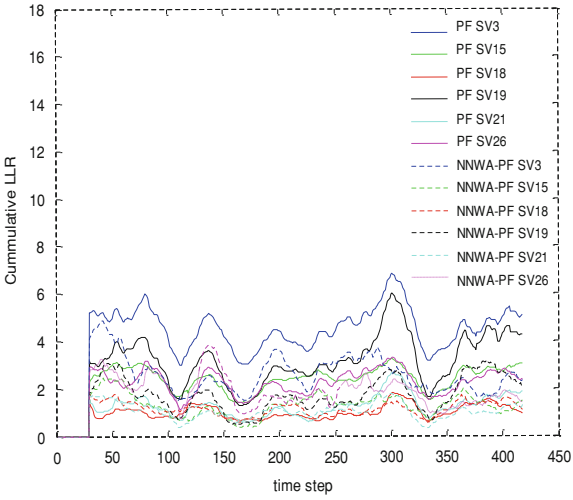


Fig. 13.2 Cumulative LLR for fault isolation under nominal condition

under nominal condition. As can be seen from Fig. 13.1, the decision function value of FDI method based on NNWA-PF exist smaller fluctuations. The decision function value of FDI method based on PF exist bigger fluctuations, which is obviously larger than NNWA-PF. If the fluctuation range of decision function β_k is too large, it may trigger an alarm when the system has not reached the alarm actually, thereby increase the probability of false alarm. For the selected alarm threshold, the system false alarm probability of FDI system using NNWA-PF algorithm is less than the FDI system using PF.

The simulation results based on NNWA-PF algorithm and basic particle filter algorithm used in RAIM are shown as Figs. 13.3 and 13.4 when a failure occurs intentionally.

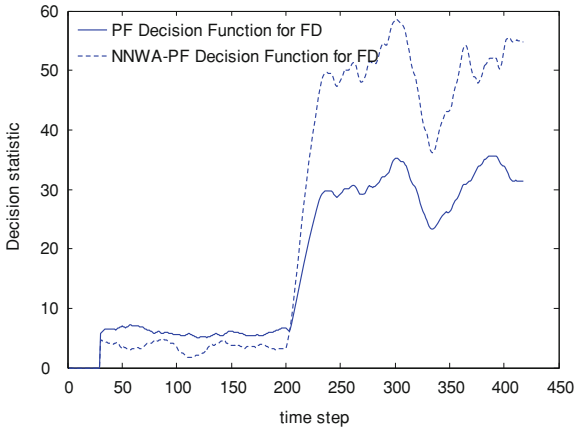


Fig. 13.3 Decision function for fault detection under fault condition

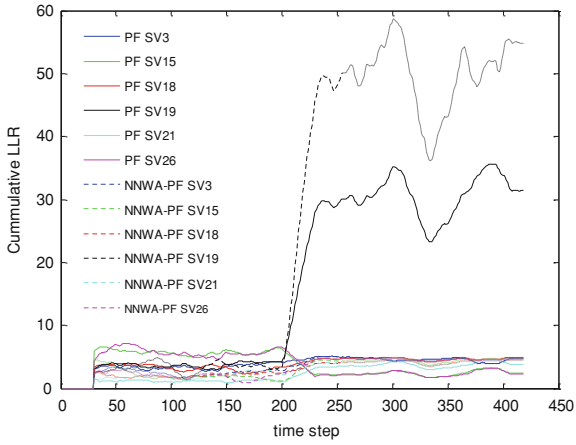


Fig. 13.4 Cumulative LLR for fault isolation under fault condition

As can be seen from Fig. 13.3, the decision function based on NNWA-PF appears a significant jump at the time of k = 205 and exceeds the detection threshold. However, the decision function based on PF appeared a significant jump at the time of $k = 210$. As can be seen in Fig. 13.4, before the fault occurs, the decision function value of FDI system using NNWA-PF is less than FDI system using the PF, indicating that FDI system using NNWA-PF has a less probability of false alarm than the FDI system using PF.

According to the error detection principle by using the cumulative LLR function described before, it can successfully identify that the 19th satellite has a fault in Fig. 13.4. Then, during the satellite failure the 19th satellite data for positioning solution should be discarded. It can be seen from Figs. 13.3 and 13.4, its detection performance of RAIM based on NNWA-PF's is superior to that of the basic particle filter and the logarithmic likelihood ratio (LLR) combining the RAIM algorithm.

13.5 Conclusion

An approach of fault detection and isolation (FDI) for GPS integrity monitoring by combining particle filters with neural network algorithm is proposed. Based on neural network weights adjustment particle filter (NNWA-PF) algorithm, the algorithm can effectively suppress weight degradation and sample impoverishment. The failure detection is undertaken by checking the system state logarithmic likelihood ratio (LLR). Through GPS receiver experiment platform NS220 samples the observation data, numerical simulations verified that in the non-Gaussian measurement noise environment the algorithm combined the neural network particle filter algorithm with logarithmic likelihood ratio is feasible and effective for GPS receiver autonomous integrity monitoring (RAIM). The NNWA-PF algorithm can improve the accuracy of state estimation and reduces the failure detection time. The results demonstrated that the detection performance is superior to that of the basic particle filter.

Acknowledgment This study is funded by National Natural Science Foundation of China(61101161), The Aeronautical Science Foundation of China(2011ZC54010) and the Joint Funds of the Natural Science Foundation of Liaoning Province(2013024003).

References

1. Xu XH, Yang CS, Liu RH (2013) Review and prospect of GNSS receiver autonomous integrity monitoring[J]. Acta Aeronautica et Astronautica Sinica 34(3):451–463 (in Chinese)
2. Yun Y, Kim D (2007) Integrity monitoring algorithms using filtering approaches for higher navigation performance:consideration of the non-gaussian gnss measurements[C]. In: Proceedings of ION GNSS 20th international technical meeting of the satellite division, Fort Worth, pp 3070–3071
3. Sun GL, Sun MH, Chen JP (2006) A study on time and set combined method for receiver integrity autonomous monitoring. Acta Aeronautica et Astronaut Sin 27(6):1171–1175. (in Chinese)
4. Mathieu J, Boris P (2011) Integrity risk of kalman filter-based RAIM[C]. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation. USA:ION, 3856–3867
5. Sayim I, Pervan B, Pullen S, Enge P (2002) Experimental and theoretical results on the LAAS sigma overbound. In: Proceedings of the ION GPS, Portland, pp 29–38
6. Gordonn J, Almond S, Looney DJ et al (1993) Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEEE Proc F Radar Sig Process 140(2):107–113
7. Vaswani N (2004) Bound on errors in particle filtering with incorrect model assumptions and its implication for change detection. Proceedings of IEEE international conference on acoustics, speech and signal processing, Montreal, II-729–32
8. F Gustafsson, F Gunnarsson et al (2002) Particle filters for positioning, navigation, and tracking. IEEE Trans Sig Process 50(2):425–437
9. Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process 50(2):174–188
10. Chen Y,Wang L, Huang S (2006) Neural network learning algorithm based on particle filter. Eng J Wuhan Univ 39(6):86–88 (in Chinese)

11. Kaplan E, Hegarty C (2006) Understanding GPS: principles and application, 2nd edn. Artech House, Norwood
12. Li P, Kadirkamanathan V (2001) Particle filtering based likelihood ratio approach to fault diagnosis in nonlinear stochastic systems. IEEE Trans Syst Man Cybern C 31(3):337–343
13. Rosihan R, Indriyatmoko A, Chun S et al (2007) Particle filtering approach to fault detection and isolation for GPS integrity monitoring. In: Proceedings of ION GNSS 19th international technical meeting of the satellite division, Fort Worth, pp 873–881

Chapter 14
Receiver Autonomous Integrity Monitoring for Fixed Ambiguity Precise Point Positioning

Shaojun Feng, Altti Jokinen, Washington Ochieng, Jianye Liu and Qinghua Zeng

Abstract There are still many challenges in Precise Point Positioning (PPP) including formulation of mathematical models, fast resolution of integer ambiguities, ambiguity validation and integrity monitoring. Research to date has focused on error modelling and ambiguity resolution. The ambiguity validation and integrity monitoring is still to be investigated in detail. Early research on PPP integrity has addressed the transferability of the Carrier phase based Receiver Autonomous Integrity Monitoring (CRAIM) algorithms developed for conventional Real Time Kinematic positioning (cRTK). However, there are significant differences between cRTK and PPP in the characteristics of the corresponding residual errors. For example, the satellite clock errors are removed in cRTK; while there are still satellites clock errors remaining in PPP after the application of correction products. The magnitude of these residual satellite clock errors depends on the quality of the products used. The residual errors in PPP are expected to be bigger than those in cRTK. These errors have significant negative impacts on ambiguity validation and integrity monitoring. This paper addresses these challenges.A Doubly Non-Central F distribution (DNCF) is justified for the use with popular ratio test for ambiguity validation. The residual errors in the PPP are characterised for the two key processes in RAIM, failure detection and derivation of protection levels. The correction products used for tests were from Centre National d'Etudes Spatiales (CNES). The GNSS measurement data used were from the American National Oceanic and Atmospheric Administration (NOAA). This selection is to ensure that data from same stations used to test the method are not part of the data sets for the generation of correction products. A dataset from 2 NOAA stations was used for testing. Test results show that the PPP algorithm with

S. Feng (✉) · A. Jokinen · W. Ochieng
Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK
e-mail: s.feng@imperial.ac.uk; dr.shaojunfeng@gmail.com

J. Liu · Q. Zeng
Nanjing University of Aeronautics and Astronautics, Nanjing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_14,

the DNCF based ambiguity validation can reach sub-decimetre accuracy. The protection levels calculated shown to over-bound the position errors all the time. The relative lower protection levels give the potential for the proposed method to be used in critical high accuracy applications.

Keywords Precise point positioning · RAIM · Doubly non-central F distribution

14.1 Introduction

The Precise Point Positioning (PPP) technique uses products from a global network of receivers and carrier phase measurements to enable single receiver users to access centimetre-level accuracy. Unlike conventional Real Time Kinematic positioning (cRTK), PPP does not require local reference stations. In both techniques, the ambiguity resolution is a crucial step. The level of difficulty in resolving correct ambiguities depends on the residual errors in the formulated observations. Correct ambiguity resolution is relatively easy in cRTK particularly in short baseline cases, where the common errors (e.g. satellite and receiver clock) are cancelled; while the correlated errors (e.g. ionosphere effect and troposphere delay) are mitigated. However, it is relatively difficult in PPP due to relatively higher residual errors in observations resulting difficult to obtain sufficiently accurate float ambiguity position solutions with narrow-lane or carrier measurements (e.g. L1). At the early stage, the carrier-phase ambiguities in PPP are estimated only as float values [23]. It was not possible to fix ambiguities to integers, because there are Fractional Cycle Bias (FCB) in carrier-phase measurements. The recent development in PPP products generation and modelling takes FCBs into account enabling PPP ambiguity resolved as integers [3, 9, 10, 16]. There are three different modelling methods for the integer ambiguity fixing: the decoupled clock, the single-difference between-satellites and the integer phase clock model. Even though the implementations of these methods are different, a comparison showed that the three methods are equivalent in terms of the integer property recovery, the system redundancy and the necessary corrections [19].

In order to provide the trust of the PPP solutions, an integrity monitoring process should be incorporated to the PPP algorithm in both the ambiguity resolution and positioning stage. The first stage is to validate the ambiguity and the second stage is to derive the protection levels. In order to ensure the correctness of resolved ambiguities, a validation process is followed. The validation is based on a test involving the construction of a test statistic and the determination of the corresponding threshold. The test statistic can be constructed using either the residuals of observations or the residuals of ambiguities. Multiple test statistics may be constructed using both. Depending on the nature of the test constructed, there are three types of scheme in the determination of if the test is passed: the use of constant threshold; the using of threshold determined from assumed distribution and the

simulation. An example of using constant threshold is the popular ratio test. Examples of using statistic distributions include F-distribution, t-distribution and Chi-squared distribution test etc. An example of simulation is the Ratio Test Integer Aperture (RTIA) method. Each method has its weaknesses. The ratio test with constant threshed has obvious disadvantage of ignoring the degrees of freedom. The distribution based method is good only if the statistical distribution used can accurately reflect the nature of the test statistic. Unfortunately, this is not the case. The RTIA method needs a large number samples simulation excluding it from time stringent applications.

The Carrier phase based RAIM (CRAIM) [7] was originally developed for cRTK to detect and exclude potential failures and to estimate protection levels. However, there are significant differences between cRTK and PPP in the characteristics of the corresponding residual errors. For example, the satellite clock errors are removed in cRTK; while there are still satellites clock errors remaining in PPP after the application of correction products. The magnitude of these residual satellite clock errors depends on the quality of the products used. The residual errors in PPP are expected to be larger than those in cRTK. These errors could have significant negative impacts on ambiguity validation and integrity monitoring. This paper addresses these challenges in both the ambiguity resolution and positioning stage. A recently defined Doubly Non-Central F distribution (DNCF) for popular ratio test has significantly advantages over other methods [8]. It is adopted in this paper. In addition, this paper proposes a new method to calculate the protection levels taking into account residual errors in correction products and the correlation of elements in design matrix due to the combination of measurements.

The GNSS measurement data used were from the American National Oceanic and Atmospheric Administration (NOAA). While the correction products were from Centre National d'Etudes Spatiales (CNES). This selection is to ensure that any user data are not part of the generation of products. The single-difference between-satellites PPP model is employed for test. Test results show that the DNCF based ambiguity validation has better performance than the other methods in terms of the percentage of correct ambiguities. The failure detection scheme is demonstrated to effectively detect and exclude potential failures, and the protection levels calculated shown to over-bound the position errors.

14.2 PPP Models

The state-of-the-art PPP error correction and models are employed in this paper. Satellite orbit, clock errors and Fraction Cycle Bias (FCB) are corrected using products from CNES [14, 15]. Tropospheric delay is estimated using the UNB3m model [17] and it is mapped to the range level using the Global Mapping Function (GMF) [1]. Tropospheric gradients are mapped to the range level using the Chen mapping function [2]. The state-of-the-art methods are used to corrected site-displacement effects such as solid earth tides and ocean loading [12].

The Between-Satellite-Difference (BSD) operation is applied to both code-phase and carrier-phase measurements to eliminate receiver clock and FCB errors. The BSD is calculated by selecting a satellite (typically the highest elevation satellite) as a base-satellite and differencing its measurements from other satellites' measurements. The ionosphere-free combination is used [4]. It is calculated based on L1 (ϕ_1) and L2 (ϕ_2) carrier-phase measurements in cycles.

Position is estimated by using an Extended Kalman Filter (EKF) [13, 22]. The estimated EKF states are position, tropospheric wet delay, tropospheric gradients and wide-lane and ionosphere-free carrier-phase ambiguities.

Wide-lane ambiguities are estimated in a geometry-free way using the Melbourne-Wubbena combination [18]. Wide-lane ambiguities are fixed to integers using a probabilistic based test [5]. Float narrow-lane ambiguities are calculated based on the float ionosphere-free $\left(\mathrm{b}_{\mathrm{c}}^{\mathrm{i}}\right)$ and fixed wide-lane ambiguities $\left(\mathrm{N}_{\mathrm{wl}}^{\mathrm{i}}\right)$ as in Eq. (14.1), where i is the satellite; f_1 and f_2 are the frequencies of the L1 and L2 measurements. A major challenge in fixed ambiguity PPP is the resolution of narrow-lane ambiguities. The current methods to resolve and validate narrow-lane ambiguities are discussed in next section. The wavelength of the narrow-lane combination is 10.7 cm, which makes narrow-lane ambiguity resolution vulnerable for different error sources.

$$b_{nl}^{i} = \frac{f_1 + f_2}{f_1} b_c^i - \frac{f_2}{f_1 - f_2} N_{wl}^i \tag{14.1}$$

Fixed ionosphere-free ambiguities can be calculated based on the fixed wide-lane $\left(\mathrm{N}_{\mathrm{wl}}^{\mathrm{i}}\right)$ and narrow-lane $\left(\mathrm{N}_{\mathrm{nl}}^{\mathrm{i}}\right)$ ambiguities as

$$b_c^i = \frac{f_1}{f_1 + f_2} N_{nl}^i + \frac{f_1 f_2}{f_1^2 - f_2^2} N_{wl}^i \tag{14.2}$$

Based on the observation models, user's position is obtained as EKF output. There is an interactive interface between EKF and the algorithms of ambiguity validation and calculation of protection levels.

14.3 Ambiguity Validation

Ambiguity must be solved for high accuracy positioning. In order to take proper account of the correlation and geometry-dependency of the ambiguities the Integer Least-Squares (ILS) method [20] is developed. The principle of the ILS method is to minimise the least-squares distance between the float ($\hat{\mathrm{a}}$) and fixed (a) ambiguity vectors as presented in Eq. (14.3), where ($\mathrm{Q}_{\hat{\mathrm{a}}}$) is the variance/co-variance matric of the float ambiguities [20]. The ILS search can be computationally demanding but it can be done in a computationally efficient way by employing the Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) method [20].

The principle of the LAMBDA method is making the integer least-squares search computationally efficient by de-correlating ambiguities.

$$\min_a \|\hat{a} - a\|^2_{Q_{\hat{a}}^{-1}}, a \in Z^n \tag{14.3}$$

Ambiguity validation when employing ILS can be done using the ratio test [6]. The ratio test statistic is calculated as in Eq. (14.4) where R_2 is the residuals of the second best integer ambiguity candidate vector; R_1 is the residuals of the best integer ambiguity candidate vector and k is the test acceptation threshold. The residuals R_i between the fixed ambiguity candidate vectors and float ambiguity vector are calculated as in Eq. (14.5).

$$\frac{R_2}{R_1} > k \tag{14.4}$$

$$R_i = (\hat{a} - \check{a}_i)^T Q_{\hat{a}}^{-1} (\hat{a} - \check{a}_i) \quad (i = 1, 2) \tag{14.5}$$

There are two barriers in finding the distribution of the Ratio: the requirement of the independency of R2 and R1; and the R1 is bias free. Obviously, none of them can be guaranteed. This makes the derivation of threshold (k) very difficult. Different methods have been developed to determine an estimate of the ratio test threshold (k). The simplest method is based on assuming a constant ratio threshold value. For example, the threshold value 3.0 is used in Geng et al. [8]. The F test is based on the assumption of both. The RTIA based on the assumption that bias can be detected and mitigated.

The DNCF method is based on the same test statistic used in ratio test. The DNCF distribution describes the statistic characteristics of the quotient of two independent non-central Chi distributions. The threshold determined from DNCF distribution can effectively over-bound the ratio test statistic in terms of confidence level where the numerator and denominator are correlated. Therefore, it acknowledges both correlation and the existence of bias in R1. For comparison purposes, Fig. 14.1 shows the relationship between the Confidence Levels (CL) and Degrees Of Freedom (DOF) for various thresholds.

It is obvious that the traditional fixed threshold method cannot reflect the confidence level needed. Using a fixed threshold results either in lower confidence level, or misses the opportunity to fix integer ambiguities. For example, a selection of threshold of 2.5 results in low confidence (<0.95) candidates being chosen when the DOF is less than 7; while at high confidence (>0.99) candidates are rejected when the DOF is larger than 18. The performance of the DNCF based validation with reasonable time constraint has been shown better than all existing methods [11].

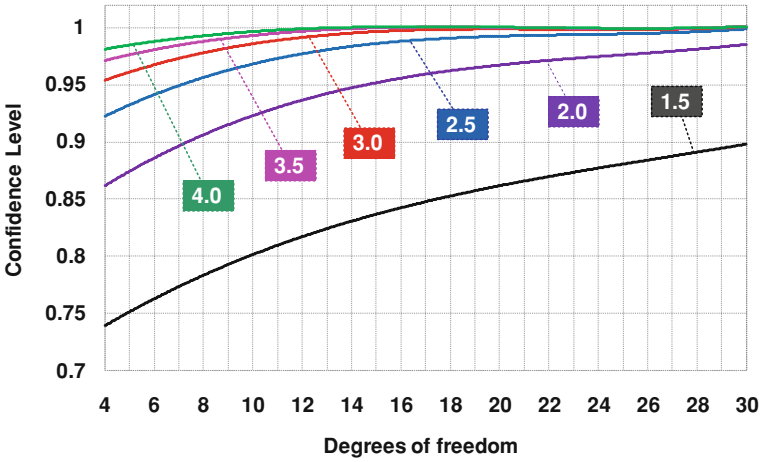


Fig. 14.1 The relationship between confidence level and DOF for various thresholds

14.4 Failure Detection and the Derivation of Protection Levels

There are extensive research efforts on conventional RAIM using pseudorange measurements. The RAIM has achieved a level of success for aviation involving failure detection and exclusion, and the derivation of protection levels. The protection level is the outer bound of positioning error. Both horizontal and vertical protection levels are used in order to decide if intended operation can be supported by satellite navigation system. However, there are not much works being done on integrity monitoring of PPP solutions, particularly on the derivation of protection levels.

In the failure detection process, three different test statistics are used when employing CRAIM. The test statistics are: total, code-phase only and carrier-phase only. Failure exclusion is attempted, if any of the test statistics is exceeding its threshold. The BSD approach used for GPS measurements requires failure exclusion to be applied separately for failures associated with the base-satellite and other satellites. In the case of the base-satellite failure, the magnitude of measurement residuals increases typically for each satellite pair. In the case of failures associated with other satellites, the magnitude of the residuals increases typically only for the measurements associated with the failed satellite. Thus, the base-satellite failure can be detected based on the contribution of residuals from different satellite pairs. The test used to detect the base-satellite failure is:

$$\frac{re_{largest} + re_{second_{largest}}}{\sum_0^n re(j)} < \frac{1.0}{n} * 3.5 \tag{14.6}$$

In this equation, *re* is the measurement residual vector and n is the number of measurements used to calculate the test statistic. This test is carried out separately for each test statistic. The measurement residuals are calculated using the same approach as for the test statistics. Both the largest and second largest residuals are used in the test, because of the potential of multiple failures. The test decision threshold as in (14.6) is chosen empirically.

To estimate realistic protection levels, a nominal bias calculated based on the estimated magnitudes of non-Gaussian errors such as satellite orbit, site-displacement and tropospheric delays are taken into account in the protection level estimation. In this paper, an empirically chosen nominal bias value of 5 cm is used. The value is selected based on the assumed error in satellite orbit corrections, tropospheric estimation and site-displacement estimation. When taking the nominal bias into account, horizontal protection level can be calculated as in Eq. (14.7) and vertical protection level as in Eq. (14.8).

$$HPL = k_H \sqrt{P_{local,11} + P_{local,22}} + \text{nominalBias} * ah \tag{14.7}$$

$$VPL = k_V \sqrt{P_{local,33}} + \text{nominalBias} * av \tag{14.8}$$

$$ah = \sum_{i=1}^{n} \left(\sqrt{G_{1,i}^2 + G_{2,i}^2}\right), \text{ where } G = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} \tag{14.9}$$

$$av = \sum_{i=1}^{n} \left|G_{3,i}\right|, \text{ where } G = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} \tag{14.10}$$

Where, P_{local} is the EKF P matrix converted to the local (East, North, Up) coordinates; k_H and k_V are factors calculated based on the allowed rate of missed detection and nominal Bias is the chosen nominal bias value. ah is calculated as in Eq. (14.9) and av as in Eq. (14.10). H and R are the EKF design and measurement variance matrices, respectively.

14.5 Results

GNSS measurement data from NOAA were used. For demonstration purpose, a random one hour data starting at 13:00, 12 June, 2013 from two different stations (MIQE and MTDT) were used. The stations were selected so that none of the stations were used to generate CNES products. This is to avoid any correlation between products generation and user positioning. The BSD PPP model was used to formulate observables. The ambiguities resolved were validated by DNCF method with a reasonable time constraint. The time required obtaining an initial ambiguity resolution and positioning errors are shown in Table 14.1. The horizontal position errors and horizontal protection levels (HPL) are shown in Figs. 14.2 and 14.4 for the stations MIQE and MTDT respectively. The vertical

Table 14.1 Positioning performance

Station code	Location	Epochs needed for fixing (s)	3D error (cm)	Horizontal error (cm)	Vertical error (cm)
MIQE	Michigan	1250	1.7 cm	0.8	1.5
MTDT	Montana	2145	3.2 cm	1.9	2.5

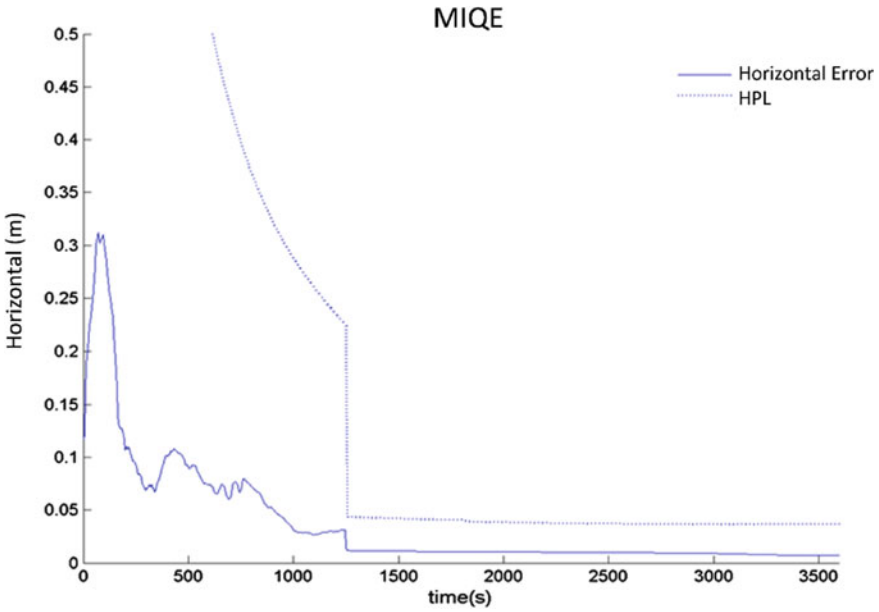
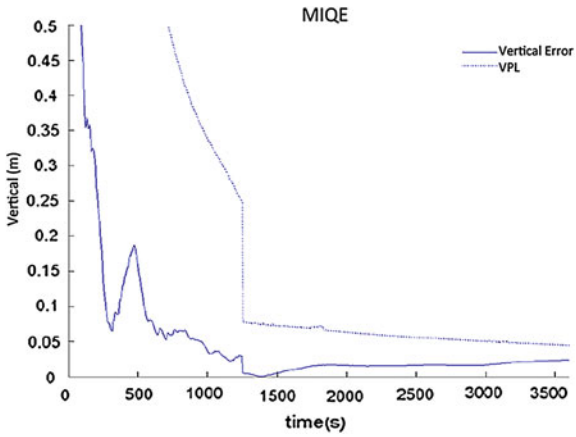


Fig. 14.2 The horizontal error and HPL of station MIQE

Fig. 14.3 The vertical error and VPL of station MIQE



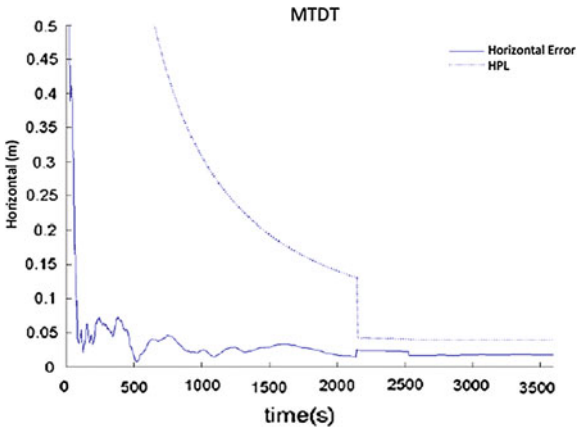


Fig. 14.4 The horizontal error and HPL of station MTDT

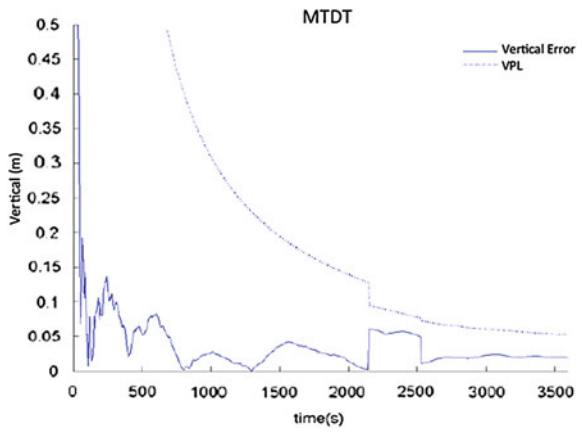


Fig. 14.5 The vertical error and VPL of station MTDT

position errors and vertical protection levels (VPL) are shown in Figs. 14.3 and 14.5 for the stations MIQE and MTDT respectively.

The test results in Table 14.1 show that the PPP algorithm can reach sub-decimetre level accuracy after the ambiguities are fixed. The high accuracy solutions also give the indication that the ambiguities are validated correctly. Figures 14.2, 14.3, 14.4, 14.5 show that the protection levels derived always over-bound position errors both in horizontal and vertical. Therefore, the users with the methods proposed in this paper are protected.

14.6 Conclusion

New methods for PPP ambiguity validation and integrity monitoring are proposed. Fixed ambiguity positioning can be carried out using the wide-lane and narrow-lane measurement combinations. The ambiguity validation is based on a recently defined Doubly Non-Central F distribution (DNCF) for popular ratio test with a reasonable time constraint. A new method is proposed to calculate the protection levels taking into account residual errors in correction products and the correlation of elements in design matrix. The two-stage integrity monitoring method can ensure the resolved ambiguity validated with confidence. The positioning accuracy can achieve sub-decimetre level and the protection levels always over-bound corresponding position errors. The relative low protection levels indicate the potential of using PPP to support critical high accuracy applications.

Acknowledgments This work is partially supported by the National Natural Science Foundation of China (NSFC: 61328301).

References

1. Boehmm J, Niell A et al (2006) Global mapping function (GMF): a new empirical mapping function based on numerical weather model data. Geophys Res Lett 33(7):1–4
2. Chen G, Herring TA (1997) Effects of atmospheric azimuthal asymmetry on the analysis of space geodetic data. J Geophys Res 102:20,489–20,502
3. Collins P (2008) Isolating and estimating undifferenced GPS integer ambiguities. In: Proceedings of the ION NTM, San Diego, CA
4. Dach R, Hugentobler U et al (2007) Bernese GPS Software Version 5.0 user manual, University of Bern, Switzerland
5. Dong D, Bock Y (1989) Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in California. J Geophys Res 94:3949–3966
6. Euler HJ, Schaffrin B (1991) On a measure for the discernability between different ambiguity solutions in the static-kinematic GPS-mode. IAG Symposia no 107, Kinematic Systems in Geodesy, Surveying, and Remote Sensing. Springer, pp 285–295
7. Feng S, Ochieng W, Moore T, Hill C, Hide C (2009) Carrier phase-based integrity monitoring for high-accuracy positioning. GPS Solutions 13:13–22
8. Feng S, Ochieng W et al (2012) Integrity monitoring for carrier phase ambiguities. J Navig 65:41–58
9. Ge M, Gendt G et al (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. J Geodesy 82:389–399
10. Geng J, Teferle FN, Shi C, Meng X, Dodson AH, Liu J (2009) Ambiguity resolution in precise point positioning with hourly data. GPS Solutions 13:263–270
11. Jokinen A, Feng S, Schuster W, Ochieng W, Yang L, Moore T, Hill C (2013) Improving ambiguity validation and integrity monitoring of precise point positioning (PPP). In: Proceeding of the ION GNSS, Nashville, TN, USA, 16–20 Sept 2013
12. Kouba J (2009) Guide to using International GNSS service (IGS) products. Geodetic Survey Division, Natural Resources Canada, Ottawa
13. Larson RE, Dressler RM, Ratner RS (1967) Application of the Extended Kalman filter to ballistic trajectory estimation. Stanford Research Institute, Menlo Park

14. Laurichesse D (2011) The CNES real-time PPP with undifferenced integer ambiguity resolution demonstrator. In: Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011), Portland, Oregon
15. Laurichesse D (2012) Phase biases estimation for undifferenced ambiguity resolution. In: PPP-RTK & open standards symposium. Frankfurt am Main, Germany
16. Laurichesse D, Mercier F (2007) Integer ambiguity resolution on undifferenced GPS phase measurements and its applications to PPP. In: Proceedings of the 20th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2007), Fort Worth, TX
17. Leandro R, Santos M, Langley RB (2006) UNB neutral atmosphere models: development and performance. In: Proceedings of ION NTM 2006, Monterey, California
18. Melbourne W (1985) The case for ranging in GPS based geodetic systems. In: Proceedings of the 1st international symposium on precise positioning with the global positioning system. Rockville, Maryland
19. Shi J, Gao Y (2013) A comparison of three PPP integer ambiguity resolution methods. GPS Solutions. doi:10.1007/s10291-013-0348-2
20. Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. In: The general meeting of the international association of Geodesy, Beijing, China
21. Teunissen P (1998) A class of unbiased integer GPS ambiguity estimators. Artif Satell 33:4–10
22. Wishner RP, Tabaczynski JA, Athans M (1969) A comparison of three non-linear filters. Automatica 5:487–496
23. Zumberge JF, Heftin MB et al (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102:5005–5017

Chapter 15
Deformation Characteristics of COMPASS Satellite Signals

Jiancheng Liu, Huafeng Yang, Jianjun Fan and Dawei Zhu

Abstract Stanford University's study declare that navigation signals of healthy satellite may be deformed which results in ranging bias, so micro-deformation characteristics of COMPASS satellite signal is studied based on measured data. Using of periodicity of PRN code for civil signals, data is collected at sampling rate which the interval of first point of each PRN code from the starting point is fixed. Rearranging the data, which makes that several PRN code waveforms in one data code incorporate into one PRN code waveform, so each chip has enough data points and data rate increases. A large-diameter dish antenna is used, and initial phase, frequency, secondary code and information code of sampling data are removed, then data is rearranged into one PRN code period, then average for noise reduction. Finally the clean COMPASS satellite signal waveform is compared with the real GPS satellite signal waveform and the TMB fault model.

Keyword COMPASS satellite signal · Sampling rate · Data processing · TMB fault model

15.1 Introduction

Satellite navigation system requires accurate information on the signal structure which satellite transmits. Signal transmission channel failure of navigation satellite in orbit can cause distortion of the navigation signal waveform, which affects the signal acquisition and tracking of the navigation receiver, and make the receiver is loss of lock or deviation from the actual measurement values [1, 2].

Current status of satellite navigation signal quality monitoring studies and reality applications as follows [1, 3–5]: (1) the antenna gain of a ordinary receiver

J. Liu (✉) · H. Yang · J. Fan · D. Zhu
Beijing Satellite Navigation Center, Beijing 100094, China
e-mail: ljc-mymail@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_15,

is very low, which lead to the received signal low SNR, even buried in noise, and therefore cannot distinguish details of the signal waveform; (2) the data sampling rate cannot meet the need to describe the time-domain waveform in detail for high rate PRN code. For Agilent VSA 89600, the sampling rates can be set to 46.08 MHz, and for COMPASS B1 in-phase channel signal whose code rate is 2.046 Mcps there are not more than 23 sampling points in one chip, so clear information of signal waveform cannot be got.

A large-diameter antenna is adopted to receive the GPS satellite L1 C/A signal in [1, 6], the data collected of 20 pseudo-code waveforms in one information code may be combined, the data rate is increased by 20 times. Conditions of the process are that each sample point position of the PRN code period is shifted at a fixed interval. The sampling rate can be determined according to the condition. PRN code period may be lengthened or shortened due to the relative motion of satellites to ground equipment. Therefore, the sampling rate relates with the Doppler velocity and increased multiples of data rate. After elimination of initial phase and frequency, then data combination, accumulating, finally the clean signal waveform of the healthy GPS satellites in the time domain was obtained.

The data preprocessing in [1, 6] does not include the elimination of secondary code and information code, and therefore multiples of the data rate increase is fixed at 20 times. For the COMPASS satellite signals, data preprocessing must including eliminating secondary code firstly. The method is presented in this paper, which is for data preprocessing of COMPASS satellite signals, including the elimination of the initial phase, the residual frequency, secondary code, navigation code, etc., so multiples of the data rate increase is not limited by the number of PRN code in one information code. Doppler velocity in the moment of collection is estimated using the periodicity of Doppler velocity of IGSO satellites according to multiple cycles of Doppler velocity data from a GNSS receiver near the large-diameter antenna.

15.2 Preprocessing Method of Collected COMPASS Satellite Signals

The data preprocessing step of this paper is:

(1) Eliminates the initial phase and frequency

In order to eliminate the initial phase and frequency of the data $s(\mathrm{k})$, adjusting f_{res} and φ_0, the collected data is processed as:

$$s^{'}(\mathrm{k}) = s(\mathrm{k})\ e^{-j2\pi f_{\mathrm{res}} k T_s - \varphi_0}$$

where T_s is the sampling period of the data.

Check the real part and imaginary part of the resulting data whether show pulse shape, and pulse width is multiple pulse width which is determined by PRN code rate. Adjusting f_{res} and φ_0 until satisfy that condition, and the real part is in-phase data. The resulting sequence includes PRN code, secondary code, and navigation code.

(2) Eliminates secondary code and navigation code

To eliminate secondary code and navigation code, it is required to find the beginning position of PRN code period and navigation code, and discard the front part of data which makes the first data point is in the starting position of PRN code period and navigation code. The method is described in the following.

Firstly according to the generator polynomial and initial state, the PRN code sequence is generated. The top sections of the data sequence modulated by secondary code and navigation information convert to -1 or 1 according to PRN period, and check the fragment whether appears in the PRN code to find the starting point. After finding the position, when discard the data of non-complete PRN period, so the new data starts from the starting point of the PRN period. For the new data, each sequence of 1 ms within 20 ms is same except for secondary code and information code, so that the polarity transformation of each sequence of 1 ms is caused entirely by secondary code and information code. According to the rule of the polarity transformation for multiple 1 ms sequences and Neumann-Hoffman coding, the starting point of information code is determined. After determining the position, discard the data of the non-complete navigation code, the new data starting point is from the start position of a navigation code.

After a series of data preprocessing eliminating the initial phase, the residual frequency, secondary coding, and navigation code, the data only contains PRN code. If the sample rate is set appropriately, then the data sequence can be combined to increase the data rate. The data rate is not constrained by the number of PRN code in one navigation code.

15.3 Data Processing of Collected COMPASS Satellite Signals

15.3.1 Data Collection Experiment

A large-diameter dish antenna is adopted to receive COMPASS satellite signals, including B1 band and B2 band. The received signal is processed by low noise amplifier and down converter into the IF signal at 71.61 MHz, then is input Agilent's vector signal analyzer 89600 after intermediate amplifier. One of the two band signals is manually selected to connection when collection. The connection is shown in Fig. 15.1.

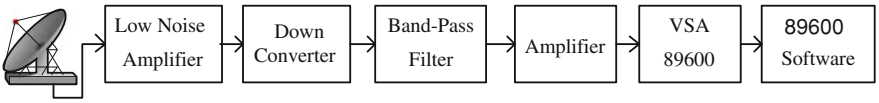


Fig. 15.1 COMPASS satellites signals collection experiment

Table 15.1 Sampling rates of different COMPASS satellite signals

COMPASS satellites	Bands	Doppler velocity measures (m/s)	Doppler velocity estimation (m/s)	Sampling rates (MHz)
IGSO-1	B1	−22.831	−2.840	46.079949916
	B2	−5.544	−5.549	46.079949788
IGSO-2	B1	−13.962	−13.987	46.079949587
	B2	−14.795	−14.889	46.079949431
IGSO-4	B1	−55.891	−55.837	46.079948352
	B2	−58.445	−58.474	46.079947769

Before the start of data collection, the sampling rate is needed, which is related to Doppler velocity, so a method to estimate Doppler velocity is described below. Although you can get real-time Doppler velocity from other receiving devices, but Doppler velocity will change during this time from calculating the sampling rate according to the Doppler velocity and setting sampling rate in collection equipment, to data storage, which resulting in sampling rate error, so Doppler velocity should be determined in advance. Read multiple days of Doppler velocity from a GNSS receiver near the large-diameter antenna, and use the periodicity of the Doppler velocity of IGSO satellites to find the change rules of the Doppler velocity. According to the change rules Doppler velocity in the moment of collection can be calculated before collection. The estimations and measures of Doppler velocity of multiple COMPASS satellites are listed in Table 15.1. As can be seen, the error is smaller than 0.1 m/s.

A large-diameter dish antenna tracks IGSO-1, IGSO-2 and IGSO-4 satellite respectively, and B1 frequency signal sampling rate is set, navigation signals of B1 band is collected, after the collection the B1 signal, disconnect the B1 signal and connect B2 signal to collect B2 band navigation signals. Each frequency sampling rate is shown in Table 15.1. The collecting time is 2 s. The storage data type which the VSA 89600 selects is mat type, and the file size is about 720 MB.

15.3.2 Data Processing Procedure

The data processing adopted by COMPASS satellite signal is including the elimination of the initial phase, the frequency, secondary coding, and navigation codes, then according to certain rules the data sequences are combined to get the PRN code waveform, finally accumulate and moving average to reduce the noise.

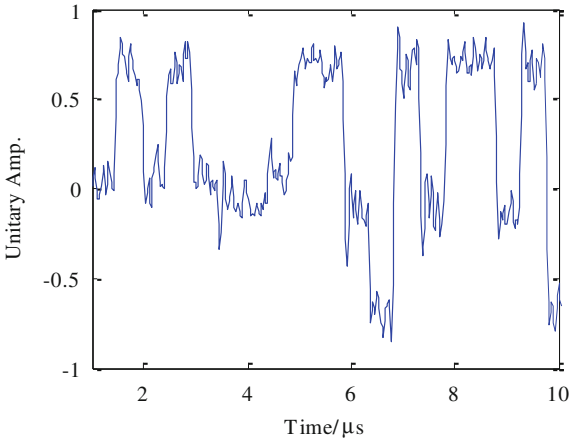


Fig. 15.2 Collected raw data of in-phase signals

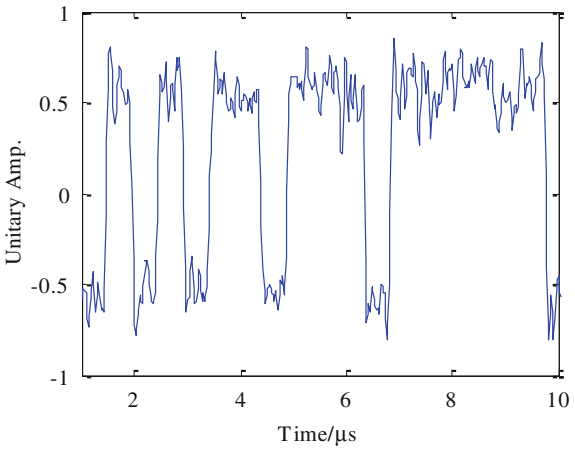


Fig. 15.3 Data with initial phase and frequency removed

Below an example of B1 is given to show the processing result of the COMPASS satellite collected signals.

(1) The results of elimination initial phase and frequency

Read the data collection of IGSO-1 satellite B1 band, and the initial phase and frequency of the data are eliminated. The data fragment of in-phase channel is shown in Fig. 15.2, the resulting data of elimination of the initial phase and frequency is shown in Fig. 15.3.

Comparison of Figs. 15.2 and 15.3, it can be seen that the resulting data after the elimination of the initial phase and the residual frequency can be expressed as a pulse shape, but the sample rate is not enough, each chip sample point is too small, and due to noise, amplitude undulating thus the PRN code waveform is not clear after eliminating the initial phase and frequency, and cannot accurately describe the true waveform.

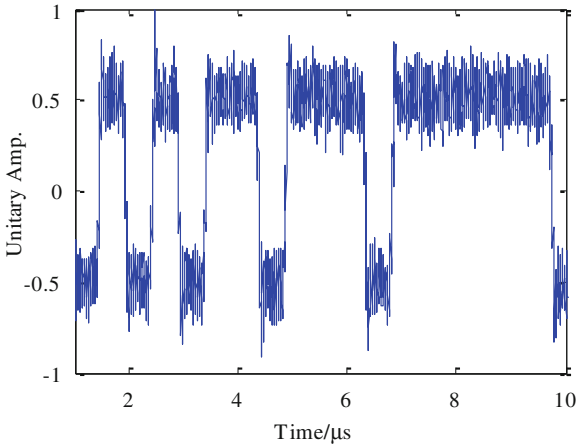


Fig. 15.4 Data after combination

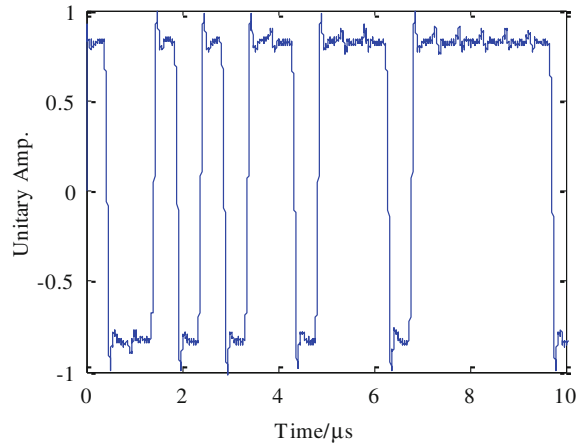


Fig. 15.5 Data after smoothness

(2) The results of elimination of secondary code and navigation code and data combined

The data which has eliminated the initial phase and frequency eliminates the secondary code and navigation code, and the data sequence according to certain rules are combined to obtain a high data sampling rate, as shown in Fig. 15.4.

As can be seen from Fig. 15.4, the combined data rate has been greatly improved, and the number of sampling points in one chip increases from 23 to 450, then the clarity of the pulse waveform is improved, but because of the influence of noise, the amplitude fluctuates excessively.

(3) The results of accumulating and moving average

The combination data accumulates the multiple periods of PRN code waveform, and further using 9 point moving average to decrease the effect of noise. The

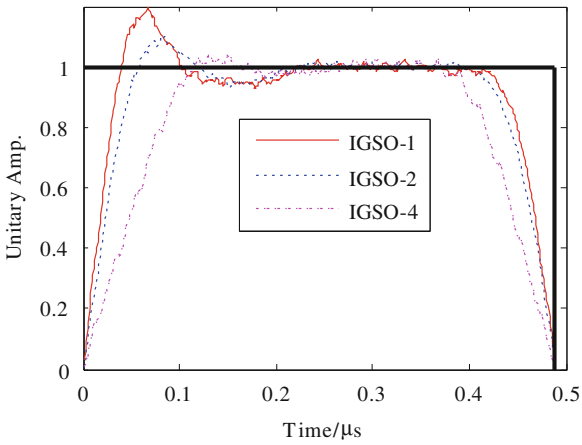


Fig. 15.6 Waveforms of COMPASS IGSO-1/2/4 satellites

result is shown in Fig. 15.5. As you can see from Fig. 15.5, the pulse waveform is clear after accumulating and moving average.

At present the IGSO-1, IGSO-2, IGSO-4 satellite I1 frequency point data processing have completed, the clear pulse waveforms of three satellites are obtained, as shown in Fig. 15.6.

As can be seen in Fig. 15.6, the COMPASS satellite signal waveform is not a ideal rectangular pulse, and there is a slight distortion. Comparison with a pulse waveform of GPS satellites in [6], it is shows that: (1) the ringing frequency of COMPASS satellite is lower than the GPS satellite. (2) The consistency of pulse waveforms between COMPASS satellites is inferior to GPS satellites.

15.4 Comparison of the Real Signal Waveform to TMB Distortion Model

The 2OS model is a satellite navigation signal distortion model approved by ICAO, is widely used in navigation signal quality monitoring [5–7]. The waveform distortions caused by the digital circuit fault, analog circuit fault, and a hybrid circuit faults are called Threat Model A (TMA), Threat Model B (TMB), and Threat Model C (TMC), wherein the TMB model is used to describe the oscillation characteristics of resulting pulse waveform. The measured pulse waveforms of IGSO-1/2/4 satellites are compared with the TMB model. Figure 15.7 shows the difference between the measured pulse waveform with the TMB model for different parameters. As can be seen, when changing the model parameters of the TMB model, it is still unable to match the measured IGSO-1/2/4 satellite pulse waveform, so the TMB model cannot fully describe the oscillation characteristics of the COMPASS satellite pulse waveforms.

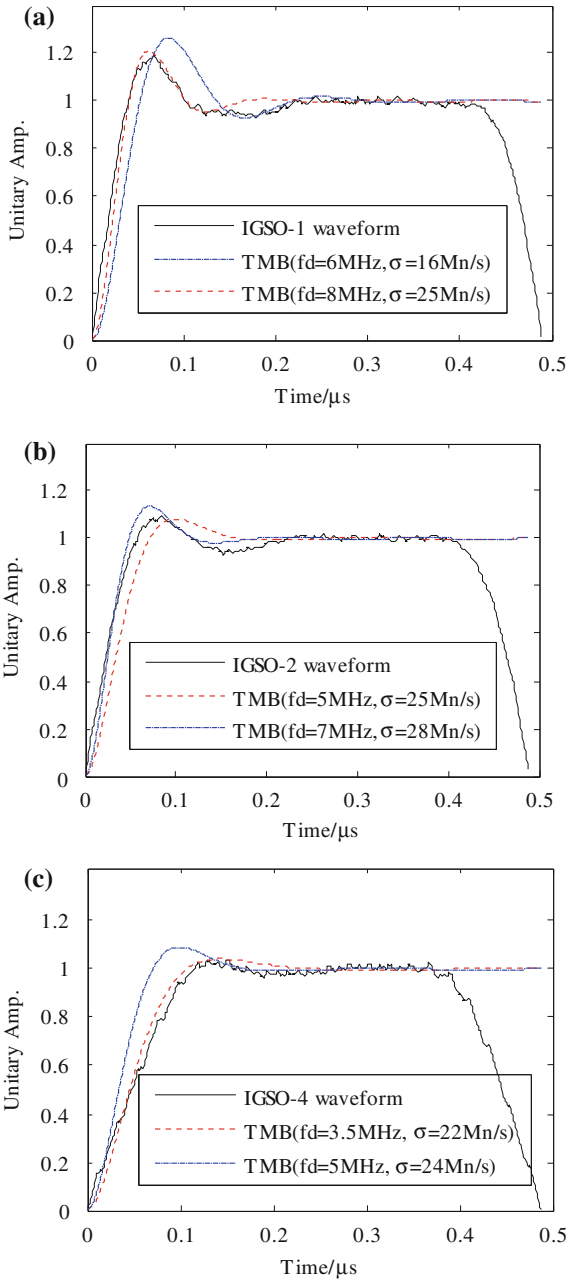


Fig. 15.7 Comparison of real waveforms with TMB model

15.5 Conclusions

In this paper, slight deformation characteristics of COMPASS satellite civil navigation signal waveform are studied based on the measured data. Data preprocessing methods for the COMPASS satellite signal and the method of pre-calculated Doppler velocity at the moment of collection are proposed. In this paper, a large-diameter dish antenna is used to receive navigation signals; a high SNR data is obtained. After the data pre-processing, data combination, accumulating and moving average, a clear pulse waveform is got, and there is a slight deformation. The comparison of the measured signals to GPS satellite signals is shown that: (1) the ringing frequency of COMPASS satellite is lower than the GPS satellites. (2) The consistency of pulse waveforms between COMPASS satellites is inferior to the GPS satellites. The TMB model cannot fully describe the oscillation characteristics of the COMPASS satellite pulse waveforms. Further research work is to establish distortion evaluation model based on measured pulse waveforms for assessing signal quality of the satellites in orbit.

References

1. Mitelman AM (2004) Signal quality monitoring for GPS augmentation systems. Ph.D. thesis, Standford University, California
2. Liu J, Sang H, Xu Y et al (2013) Effect of GNSS waveform distortions on SNR. Telecommun Eng 53(1):28–32
3. Lu X, Zhou H (2010) Study on analysis method of GNSS space signal quality. Sci SINICA Phys Mech Astron 40(5):528–533
4. Wang H, Liu C (2009) Measuring system for satellite navigation signal quality. GNSS World China 34(3):60–63
5. Wong G, Phelts RE, Walter T, Enge P (2011) Bounding errors caused by nominal GNSS signal deformations. In: Proceedings of the ION GNSS, ION GNSS, Manassas, VA, pp 2657–2664
6. Wong G, Phelts RE, Walter T, Enge P (2011) Alternative characterization of analog signal deformation for GNSS-GPS satellites. In: Proceedings of the 2011 international technical meeting of the institute of navigation, San Diego
7. Phelts RE (2001) Multicorrelator techniques for robust mitigation of threats to GPS signal quality. Ph.D. thesis, Standford University, California

Chapter 16
Modeling and Application of COMPASS Satellite Orbits and Clocks Predicted Correction

Yize Zhang, Junping Chen, Xiuqiang Gong, Bin Wu, Jiexian Wang, Sainan Yang and Mao Li

Abstract To evaluate the satellite orbit and clock errors of COMPASS, we compare the difference between broadcast ephemeris and precise orbits and clocks. Results show that the orbit errors present certain periodicity for GEO and IGSO satellites. The SISRE of COMPASS is larger than GPS and GLONASS overall. However, the broadcast ephemeris of COMPASS appears regular changes in short time. Based on this, we establish different rapid correction models for COMPASS satellite orbits and clocks error. Analysis proves that the result performs well when the length of the arc for the model is shorter than 30 min, and the linear model is best for satellite clock corrections. By using this model, we predict the corrections of the satellite orbits and clocks, which is used for pseudo-range and carrier phase positioning. Results show that the positioning accuracy is on the same level with precise products for pseudo-range and carrier phase, while the convergence time is longer. This research provides an important basis for the realization of sub-meter wide area augmentation of COMPASS.

Keywords Broadcast ephemeris error · Space-in-space range error (SISRE) · Predicted model · Single point positioning

16.1 Introduction

Broadcast ephemeris error including satellite orbits and clock errors is a very important factor for user positioning. A lot of researches have been focusing on the precision of broadcast ephemeris for GPS and GLONASS [1–5]. Many years'

Y. Zhang (✉) · J. Wang · M. Li
College of Surveying and Geo-Informatics, Tongji University,
Shanghai, China
e-mail: zhyize@163.com

Y. Zhang · J. Chen · X. Gong · B. Wu · S. Yang · M. Li
Shanghai Astronomical Observatory, Chinese Academy of Science,
Shanghai, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_16,

results show that the precision of GPS broadcast ephemeris has been improving over the years. For GPS broadcast ephemeris, the error on radius is smaller than that on track and normal direction, while the errors in three directions are below 2 m. For GLONASS, the precision is about 3–4 m. Yang [6] studies the GPS broadcast ephemeris error and its effect on positioning, and results show that the GPS broadcast ephemeris error has a great effect on GPS single point positioning, and little effect on relative positioning. Huang [7] analyzes the spectral characteristic of GPS broadcast ephemeris by using spectral analysis model, the results demonstrate that the GPS broadcast ephemeris has a periodical error wave form. After compensation, the precision of ephemeris was at the level of 0.5 m.

The COMPASS Navigation Satellite System, also named Beidou-2, was taken into formal operation on Dec. 27th, 2012. The interface control document is also released at that time [8].

To compensate the lack of basic GNSS navigation service, many countries are developing their own GPS or GLONASS augmentation system. For COMPASS, to achieve a better augmentation result, how to correct the errors on broadcast ephemeris and broadcast in an appropriate way is an important issue.

Based on the above background, we analyze the error of COMPASS broadcast ephemeris. On the basis of the regularity of COMPASS broadcast ephemeris error, we establish fast correction models to predict the errors. By broadcasting the parameters of the correction model, this model can be used to correct the error of broadcast ephemeris to improve positioning accuracy.

16.2 Strategic Analysis

To analyze the error of COMPASS broadcast ephemeris, we take the precise orbits and clocks as 'true' value, and compare it with broadcast ephemeris so that to get the satellite orbits and clocks error $(\Delta X\ \Delta \mathrm{Y}\ \Delta Z\ \Delta Clk)$ in CGCS2000 coordinate system.

To describe the orbit error correctly, we usually transfer the orbit error into radius, along-track and across-track direction, namely satellite orbit coordinate system. According to the state vector of satellites at the epoch, we can calculate the rotation matrix $G_{3\times3}$ to convert orbit error from ECEF to satellite orbit coordinate system:

$$[\Delta R\ \Delta T\ \Delta N\ \Delta Clk] = \begin{bmatrix} G_{3\times3} & 0_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix} (\Delta X\ \Delta \mathrm{Y}\ \Delta Z\ \Delta Clk) \tag{16.1}$$

where ΔR, ΔT, ΔN denote components in radial, along-track, across-track direction respectively, and ΔClk represents the satellite clock error.

As the broadcast error consist of satellite orbits and clocks, error we use the SISRE (Signal-In-Space Range Error) to evaluate the precision of broadcast ephemeris, and the RMS reflects the total precision. The expression of SISRE can be written as follow:

$$SISRE = \sqrt{(\Delta R - \Delta Clk)^2 + \frac{1}{49}(\Delta T^2 + \Delta N^2)}. \quad (16.2)$$

SHA analysis centre is now focusing on GPS/GLONASS/COMPASS multi-system orbit determination and positioning [10]. User can download GPS and GLONASS precise products through the internet freely [11].

16.3 COMPASS Broadcast Ephemeris Correction

When comparing broadcast ephemeris and precise products for GPS and GLONASS, the consistency of the coordinate time system and the correction of antenna phase centre should be taken into account [12]. The COMAPSS broadcast ephemeris is based on CGCS2000 reference frame, and the time is aligned to BTD, which has a difference of 14 s with GPST. The satellite coordinates are referred to the antenna phase centre [8]. The precise orbits and clocks provided by SHA is consistence with broadcast ephemeris, so there is no need to consider the time and orbits inconsistency.

In this paper, we use the COMPASS broadcast ephemeris from June 27th to June 30th, 2012. We compared the healthy satellite orbits and clocks with that provided by SHA. The orbit errors are then transferred to satellite RTN direction according to Eq. (16.1). Figure 16.1 show the orbit and clock errors for one GEO (C01), one IGSO (C08) and one MEO (C11), respectively.

From Fig. 16.1 we can see that the orbit errors present certain periodicity for GEO and IGSO satellites, especially in the normal direction. While for MEO satellites, the periodicity is not obvious. This may be due to that the tracking stations of COMPASS are not world-wide distributed, and the MEO satellite has the phenomenon of satellite arising and descending. For GEO, IGSO or MEO satellites, the error in radius is much smaller, this is because that the observations of the ground tracking stations is more sensitive than the other two directions, and the force model in normal and track direction is still inadequate. These make the precision in radius better than normal and track direction, similar to GPS and GLONASS [4, 5]. However, the broadcast ephemeris precision is worse than GPS or GLONASS.

From Fig. 16.1 we can also notice an phenomenon that error jumps exists in all directions, for that the COMPASS broadcast ephemeris provide a series of orbit parameters every one hour, and the adjacent two ephemeris is not consistent. However, in the same ephemeris, the orbit and clock error appears a regular change.

We also make a four-day SISRE statistics for all healthy satellites, the mean value and RMS distribution are shown in Fig. 16.2. From Fig. 16.2 we can see that the SISRE of COMPASS is between 3 and 10 m. The result is worse than GPS or GLONASS, resulting from the error of COMPASS broadcast ephemeris, which will affect the precision of user positioning.

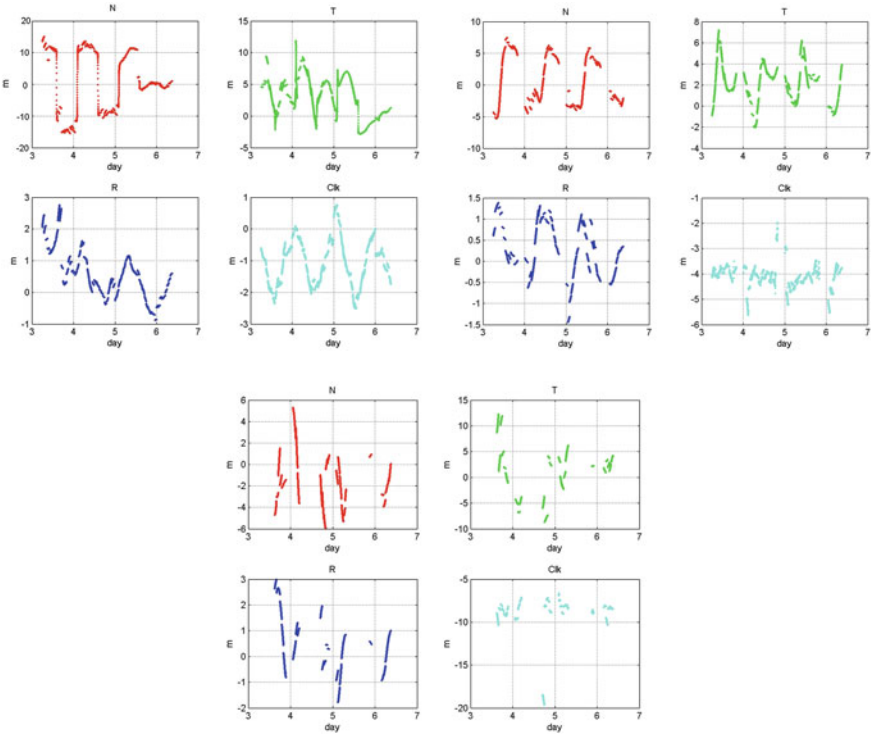


Fig. 16.1 Broadcast ephemeris errors for C01, C08 and C11

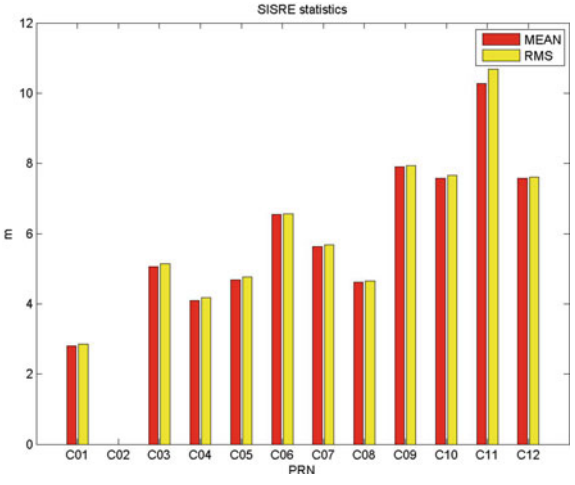


Fig. 16.2 SISRE statistics for each satellite

16.4 COMPASS Broadcast Ephemeris Corrections Model

According to Sect. 16.3, the error of COMPASS broadcast ephemeris varies much with a certain periodicity. However, the error appears a regular change in short time (such as in 1 h). So we can fast model the error of COMPASS broadcast ephemeris, and then predict the error of satellite orbits and clocks according to the model.

16.4.1 The Establishment of the Correction Model

Owning to the byte limit of broadcast ephemeris, the predict model should be simple but not reduce the predict accuracy. Huang used the combined model of spectral analysis model and AR model, and the precision was at 0.5 m level [7]. However, the model is too complex and not conductive for broadcasting.

Figure 16.1 shows that the error of broadcast ephemeris has a regular change in the same ephemeris. So we can use a linear or quadratic model to fit the errors. In order to find a best fitting model and fitting length, we use the linear and quadratic model to fit the error of broadcast ephemeris with a data length of 30 and 60-min. The fitting residuals for C04 on RTN direction and clocks are presented in Fig. 16.3, and Table 16.1 gives the statistics of fitting residuals for different fitting models and fitting length.

From Fig. 16.3 and Table 16.1 we can see that for each fitting model and different fitting length, the precision of fitting residuals is at centimeter-level or higher, which is due to the high precise orbits and clocks. While the fitting precision of clocks varies not much, this is because that the satellite orbit applies to the law of movement of dynamics. As for different fitting model and length, the quadratic model is better than linear model; the shorter the fitting length, the better fitting result we get. However, with the byte limitation of broadcast ephemeris, the bytes used in quadratic model are more than linear model; the shorter the fitting length, the more bytes should used. So how to choose a best fitting model and fitting length is a weigh issue.

16.4.2 Precision of Prediction Model

Having got the fitting model, the next issue is using it for prediction to correct the error of COMPASS broadcast ephemeris. Take C04 as example, Fig. 16.4 presents the satellite orbits and clocks predicting residuals by using the fitting model for prediction. The model here we choose is quadratic model, the fitting length and predicting length are both 60-min. Figure 16.4 demonstrates that the predicting error increases when predicting length extends, but all are less than 1 m in 1 h.

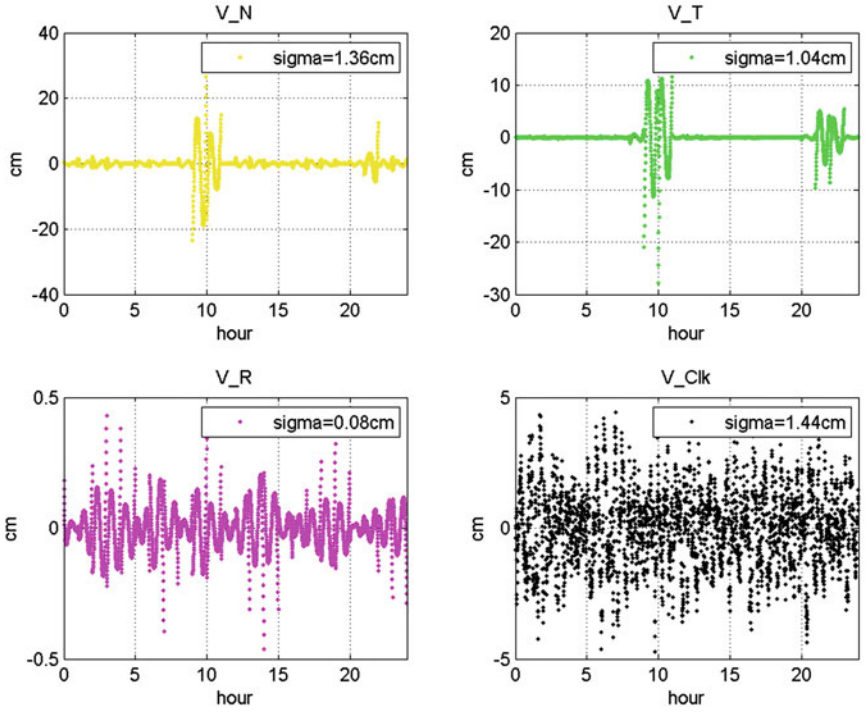


Fig. 16.3 Fitting residuals for C04 (quadratic + 60 min)

Table 16.1 Fitting residuals of different models and arc length for C04

Fitting model and fitting length	N/cm	T/cm	R/cm	Clk/cm
Linear + 60 min	3.98	4.69	0.75	1.81
Linear + 30 min	1.19	1.14	0.17	1.29
Quadratic + 60 min	1.36	1.04	0.08	1.44
Quadratic + 30 min	0.16	0.22	0.02	1.01

Table 16.2 lists the statistics of predicting residuals for orbits and clocks using different fitting models and predicting length. From the table we can see that the prediction precision of 30-min is much better than that of 60-min. As for clock prediction, a linear model is more suitable. This conclusion also applies to other satellites.

Figure 16.5 gives the statistics of SISRE after model prediction correction. Compared with Fig. 16.2, we can see that after broadcast ephemeris correction, the SISRE is at the level of 0.1–0.3 m, about 1/30 of that before correction.

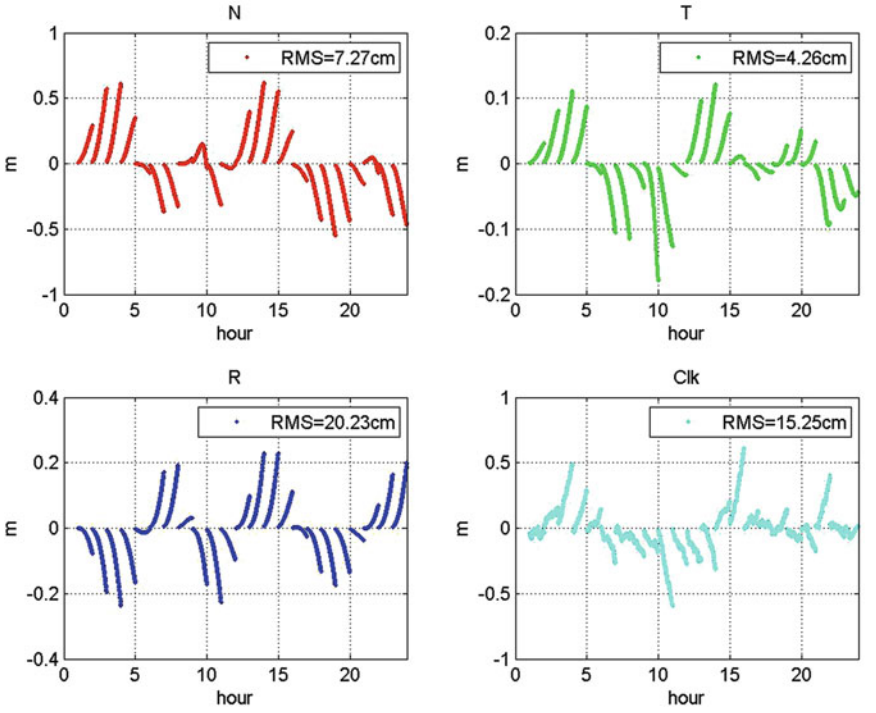


Fig. 16.4 Prediction residuals for C04

Table 16.2 Predict precision of different model and length for C04

Prediction model and length	N/cm	T/cm	R/cm	Clk/cm
Linear + 60 min	33.61	9.01	3.27	7.66
Linear + 30 min	8.25	2.54	2.48	5.53
Quadratic + 60 min	20.23	4.26	7.27	15.25
Quadratic + 30 min	4.76	1.17	0.43	12.85

16.5 The Application of Prediction Model

For users, the most important issue is positioning precision. After having the predicted model on broadcast ephemeris correction, the system can broadcast the parameter of the model to users via GEO. Thus user can correct the broadcast ephemeris and improve the positioning precision.

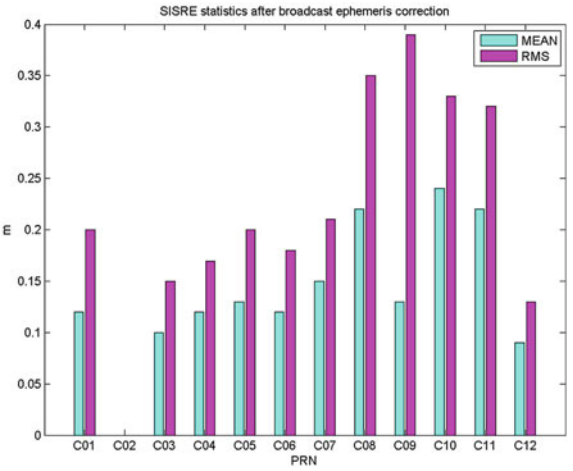


Fig. 16.5 SISRE statistics for each satellite after broadcast ephemeris correction

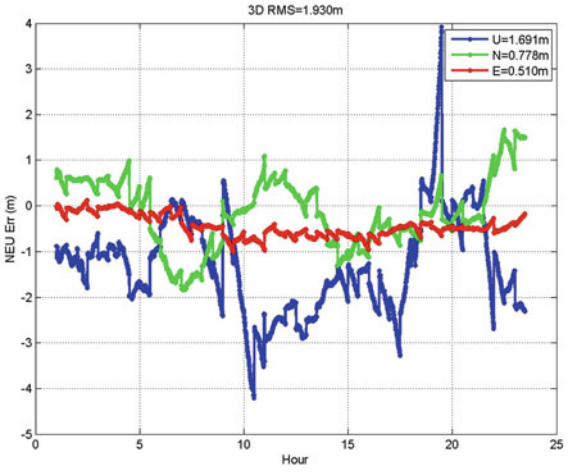


Fig. 16.6 Single point positioning using broadcast ephemeris after correction

16.5.1 Single Point Positioning

After receiving the correction model, one can correct the broadcast orbits and clocks for real-time pseudo-range positioning.

Figure 16.6 gives the COMPASS single point positioning result using the predicted model, the correction model is quadratic for orbits and linear model for satellite clocks, the length of fitting and prediction are both 30-min. Table 16.3 gives the result compared with precise orbits and clocks, and raw broadcast ephemeris.

From Fig. 16.6 and Table 16.3 we can conclude that the single point positioning result improves much after the correction of broadcast ephemeris. The precision is similar to precise orbits and clocks, which is due to the precision of prediction model.

Table 16.3 Positioning error using different orbits and satellites

	N/m	E/m	U/m	3D/m
BRD (raw)	3.016	1.893	3.585	5.039
BRD (prediction model)	0.778	0.510	1.691	1.930
Precise orbits and clocks	0.486	0.778	1.385	1.662

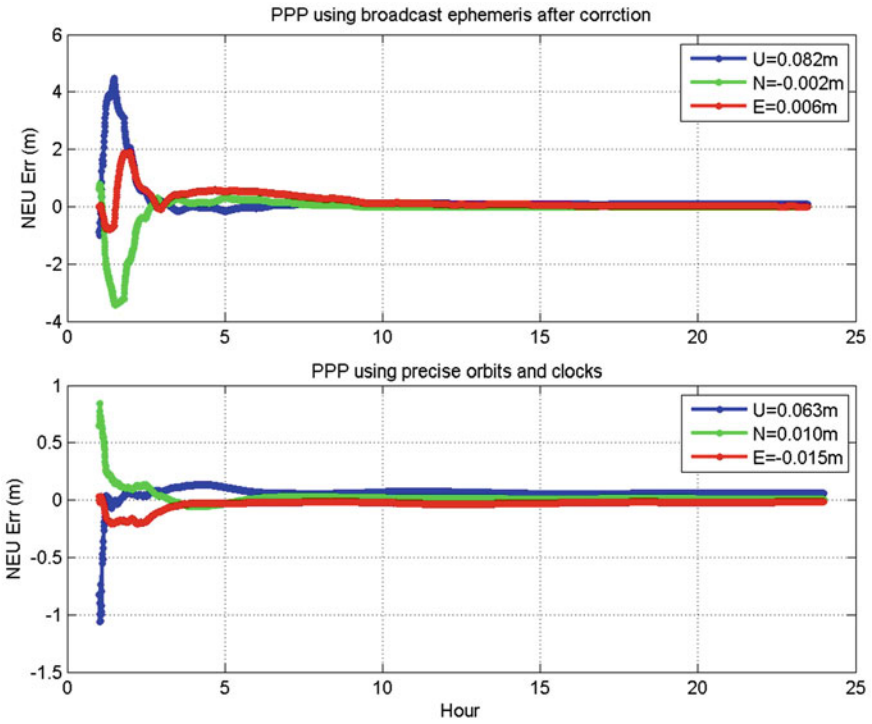


Fig. 16.7 Precise point positioning using broadcast ephemeris after correction

16.5.2 Precision Point Positioning

Due to the precision of broadcast ephemeris, the ambiguity can't convergence correctly for carrier phase positioning. According to the previous sub-section, the broadcast ephemeris can be corrected well by using the prediction model, and it is almost at the same level as precise products. So we can use the corrected broadcast ephemeris for carrier phase positioning.

Figure 16.7 shows the precise point positioning result for COMPASS users, the upper plot is the positioning result using the broadcast ephemeris after prediction model correction. The correction model is quadratic for orbits and linear model for

satellite clocks, the fitting and prediction length are both 30-min. The lower plot is the PPP result using the COMPASS precise orbits and clocks. By comparing the two plots, the convergence time is longer for modeled correction on broadcast ephemeris. However, the positioning precision is at the same level with precise orbits and clocks after convergence, and the final positioning error is (0.2, 0.6, 8.2) cm and (1.0, 1.5, 6.3) cm. This proves that this method can be used in precise point positioning.

16.6 Conclusion

In this paper, the COMPASS broadcast ephemeris error is analyzed by the comparison with the COMPASS precise ephemeris from SHA. On this basis, the broadcast ephemeris error model is established to predict the error correction and build the prediction model. The broadcast ephemeris errors can be corrected with this prediction model, thereby improving the users' real-time positioning accuracy. Some useful conclusions are as follows.

(1) There are certain periodicities in the GEO and IGSO orbit errors from COMPASS broadcast ephemeris, especially in their N direction, whereas the periodicity of MEO orbit errors is less obvious. Overall, the SISRE of COMPASS broadcast ephemeris is higher than the GPS and GLONASS.
(2) In a short time, regularities exist in COMPASS broadcast ephemeris error. Linear model or quadratic model can be used to fit the short time broadcast ephemeris errors. When the fitting arc length is 30–60 min, the fitting residuals are centimeter level.
(3) COMPASS broadcast ephemeris error can be predicted by the fitting model. The comparison results demonstrate that the longer time predicted, the larger prediction residuals produced. Better forecast results will be obtained in 30 min or less time. Linear model are recommended in the satellite clock error fitting and forecasting.
(4) User pseudo-range and carrier phase positioning accuracy can be greatly improved due to the broadcast ephemeris prediction model.

Acknowledgments This paper is supported by the 100 Talents Programme of The Chinese Academy of Sciences, the National High Technology Research and Development Program of China (Grant No. 2013AA122402), the National Natural Science Foundation of China (NSFC) (Grant No. 11273046 and 40974018), and the Shanghai Committee of Science and Technology (Grant No. 12DZ2273300, 13PJ1409900).

References

1. Warren D, Raquet J (2003) Broadcast versus precise GPS ephemesides: a historical perspective. GPS Solut 7(3):151–156
2. Wang D, Huang Z, Li R (2013) Integrity analysis of GPS signal-in-space error. In: China satellite navigation conference (CSNC) 2013 proceedings. Lecture notes in electrical engineering, vol 244. doi:10.1007/978-3-642-37404-3_17
3. Li Z, Li Z, Ding W, Li Z (2008) Error analysis of orbit determined by GPS broadcast ephemeris. J Geod Geodyn 28(1):50–54
4. Guo F, Zhang X, Li X, Hu Q (2009) Precision analysis on orbit and clock of GPS satellites broadcast ephemeris. Geomatics Inf Sci Wuhan Univ 34(5):589–592
5. Guo J, Meng X, Li Z, Nie Z (2011) Accuracy analysis of GLONASS satellite broadcast ephemeris. J Geod Geodyn 31(1):68–71
6. Yang R, Ou J, Wen D (2006) GPS broadcast ephemeris error and its effect on positioning. J Geomatics 31(1):1–3
7. Huang W, Wang W, Xi X (2010) Spec spectral characteristic analysis and prediction model research of navigation satellite broadcast ephemeris errors. Chin Space Sci Technol 3:12–18
8. BeiDou navigation satellite system signal in space interface control document. http://www.beidou.gov.cn/attach/2012/12/27/201212275f2be9ad57af4cd09c634b08d7bc599e.pdf
9. Creel T, Dorsey A, Mendicki P, et al (2006) Accuracy and monitoring improvements from the GPS legacy accuracy improvement initiative. In: Proceeding of the 2006 national technical meeting of the institute of navigation (NTM 2006), PP 664–684
10. Chen J, Wu B, Hu X, Li H (2012) SHA: the GNSS analysis center at SHAO. In: China satellite navigation conference (CSNC) 2012 proceedings. Lecture notes in electrical engineering, vol 160. doi:10.1007/978-3-642-29175-3_19
11. SHA GNSS AC. http://www.shao.ac.cn/shao_gnss_ac/
12. Heng L, Gao X, Walter T, Enge P (2010) GPS signal-in-space anomalies in the last decade data mining of 400,000,000 GPS navigation messages. In: Proceedings of the 23rd international technical meeting of the satellite division of the institute of navigation (ION 2010), pp 3115–3122

Chapter 17
Design and Evaluation of Wide Area Augmentation Navigation Message

Mengli Wang, Liuhai Ma, Tingfeng Xie and Wei Zhao

Abstract Navigation message is one of the important components of satellite navigation signals, its design will not only influence the system's performances including timeliness, integrality, flexibility, reliability and extendibility, but also user's use-cost. Different from other GNSS systems, the BDS system broadcasts not only the basic navigation messages but also augmentation messages including difference and integrity data. In this paper, the design project of BDS wide area augmentation navigation message is analyzed. And the design contents include message parameters, frame structure, data rate, message arranging format and broadcasting orders. On the basic of the designing, the augmentation navigation message's performances are evaluated. And the evaluation parameters include communication efficiency, parameters' expressing performances, broadcasting frequency, demodulation performance and signal detection performance. The conclusions can offer theory and engineering guidance for satellite navigation message design, and can apply in the design of BDS system's wide area augmentation navigation message directly.

Keywords Navigation signal · Augmentation navigation message · Data frame · Timeliness · Data rate

17.1 Introduction

Signal of navigation satellite includes three elements: carrier, ranging code and data code. The data codes modulated navigation and positioning information are broadcasting to the users by binary code stream, which is navigation message. At

M. Wang (✉) · L. Ma · T. Xie · W. Zhao
Beijing Satllite Nabigation Center, Beijing 100094, China
e-mail: compass2020@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_17,

present, there are build or building systems including GPS, GLONASS, Galileo, and Chinese BDS in the world. In addition, there are wide area augmentation systems such as WAAS and EGNOS. For different system scheme and service performances, the designs of navigation messages are different [1–5].

BDS is RNSS system which collects the basic navigation service with the differential /integrity service together. So it has the same functions with not only GPS, GLONASS and Galileo systems, but also WAAS and EGNOS systems. The constellation of BDS system constitutes with admixture satellites including GEO, IGSO and MEO. All the satellites broadcast the basic navigation messages, at the same time, the wide area augmentation messages is broadcasted by GEO satellites.

In this paper, the design scheme of BDS system's wide area augmentation message is analysed. And the analysis includes several aspects such as message parameter contents, frame structure, data rate, arrange style and broadcasting orders. On the basic of the design, the message evaluation methods are analysed by defining five evaluation targets. The five targets are broadcasting efficiency, expressing performances, broadcasting frequency, demodulation performance and signal capturing performance. From the five targets, the BDS message design scheme is evaluated.

17.2 Design Scheme of BDS Augmentation Message

For the global systems such as GPS and GLONASS, the navigation messages structure is arranged by sub-frame and page orders. The parameters include satellite ephemeris, satellite clock bias, ionosphere delay correction parameters, almanac data and time synchronization parameters. WAAS, EGNOS and GAGAN are all augmentation systems basing GPS system. And their navigation messages are designed according the unify standard of RTCA DO-229D [6]. The frame structure is data block structure basing the message type; the parameters include fast changing corrections, integrity information, GEO positioning information and ionosphere grid corrections. The RHCP signal is broadcasted by GEO satellite at L1 frequency. The modulation style is BPSK (1), the length of PRN code is 1023bits, and the data rate is 250 bps.

Compared with WAAS or EGNOS systems, the GEO satellite of BDS will broadcast two systems' augmentation messages which are BDS and GPS. The regional visible satellites are more than 24. In the following, the design scheme of BDS augmentation message is analyzed including five aspects which are parameters contents, frame structure, data rate, arranging format and broadcasting time orders.

(1) Parameters contents

Parameters of BDS augmentation message are basing RTCA DO-229D [6], and the design is optimized. Table 17.1 presents not only the information types of the augmentation message, but also the parameters contents and data quantity,

Table 17.1 Augmentation message contents

Information type	Information contents	Least broadcasting frequency (s)	Data quantity (bits)	
			Signal satellite/grid point	All data[a]
Type1	PRN mask, denoting if the satellite's data will be presented	300	One satellite 1 bit	80[b]
Type2	Fast correction and integrity information	6	One satellite 16 bit	384
Type7	UDRE fast correction attenuation factor	120	One satellite 4 bit	96
Type9	GEO satellite navigation message	120	Self-satellite 204 bit	204
Type10	Attenuation parameters	120	131 bit	131
Type12	BDS time/UTC shift parameters	3600	137 bit	137
Type17	SBAS satellite almanac	3600	One satellite 67 bit	1608
Type18	Ionosphere grid mask	600	IGP 210 bit	210
Type25	Long time error corrections	120	One satellite 103 bit	2472
Type26	Ionosphere delay correction	300	Every grid 13 bit	2730
Type28	Clock-ephemeris covariance matrix	120	One satellite 105 bit	2520

[a] Total data quantity of one group integrity parameters (all satellites and ionosphere grid)
[b] Supposing BDS and GPS constellation both have 40 satellites

parameters least broadcasting frequency and the quantity of one group integrity parameters. There into, the biggest number of two systems visible satellites is supposed as 24, and the grid number of ionosphere is supposed as 210.

(2) Frame structure

Considering the compatibility designing with the message of SBAS systems and BDS's other signals, the frame structure of BDS wide area augmentation message is designed as the Fig. 17.1. The length of every data frame is N bit. Every frame includes synchronization head of 8 bit, information type mask of 6 bit, CRC parity bit of 24 bit, and the rest bits is the field of message parameters which can write all kinds of the message parameters. The correction coding scheme is convolution-code with 1/2 coding rate. The value of N should be confirmed general by data rate, parameters broadcasting frequency and the correction coding scheme. And the value of N can be 300 bit, 250 bit, 200 bit or 125 bit.

(3) Data rate

Data rate can be confirmed general by the quantity of all the parameters and the least broadcasting frequencies of different information types. On the one hand, for

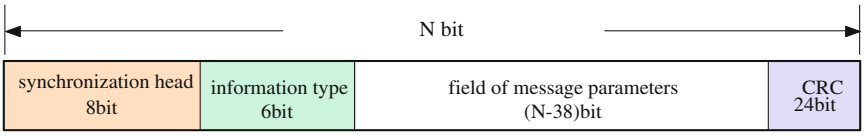


Fig. 17.1 Frame structure of the augmentation message

Table 17.2 Data rate of different frame length

Frame length N (bit)	300	250	200	125
Data rate Rb (bps)	300	250	200	250

every kind of the information, the data rate should be at least 100 bps to satisfy its all parameters broadcasting frequency. On the other hand, the data rate should be at least 130 bps for the parameters of all kinds of information. Super adding the communication resources impropriated by synchronization head, information type mask and CRC parity bit, the data rate should be higher than 160 bps when the value of N is 300 bit, 150 bps when the value of N is 250 bit, 140 bps when the value of N is 200 bit and 190 bps when the value of N is 125 bit. Considering stated spare designing for future expand and the duration of one frame designing with complete seconds, the data rate with different frame length design results is show in Table 17.2.

(4) Arranging format

According to the frame structure designed in Fig. 17.1, the message's arranging format can be designed easily by filling the parameters of every information kind in the field of message parameters orderly. Every message frame includes the parameters with only one kind of information type. According to the requirements of message format arranging and practice broadcasting, the parameters of different information types or parts parameters of the same information type can be newly grouped when the obligate bits are excessive. Parameters of the newly grouped information type are broadcasted in one frame. By this method, the message broadcasting efficiency will be optimized. For the information types with big data quantity such as type 25, type 26 and type 28, one frame can only broadcast several satellites' augmentation messages, an all the 24 satellites' information should be broadcasted by multi-frame. For the detailed message arranging format, there will not states superfluously.

(5) Broadcasting time orders

Broadcasting time orders refers to the satellite broadcasting orders of the different information types. And its design accordance is the broadcasting time delay and the least broadcasting frequency of different information types. There have many schemes of the broadcasting time orders for the different frame length. Basing on the information types' definition showing in Table 17.1, one design example of

Table 17.3 Broadcasting orders of different frame length

SOW (s)	Broadcasting time orders of different information types		
	N = 300, 250	N = 200	N = 125
6	2, 2, 25, 26, 17	2, 2, 2, 28, 7, 9	2, 2, 2, 2, 2, 25, 26, 28, 1, 7, 9, 28
12	2, 2, 25, 28, 26, 18	2, 2, 2, 10, 12, 9	2, 2, 2, 2, 2, 25, 26, 28, 18, 25, 9, 28
18	2, 2, 25, 28, 26, 10	2, 2, 2, 18, 25, 17	2, 2, 2, 2, 2, 25, 26, 28, 12, 25, 9, 28
24	2, 2, 25, 28, 26, 12	2, 2, 2, 28, 25, 18	2, 2, 2, 2, 2, 25, 26, 28, 18, 25, 25, 28
30	2, 2, 25, 28, 26, 17	2, 2, 2, 28, 25, 26	2, 2, 2, 2, 2, 25, 26, 28, 17, 25, 26, 28
36	2, 2, 25, 28, 26, 28	2, 2, 2, 28, 25, 26	2, 2, 2, 2, 2, 25, 26, 28, 18, 25, 26, 28
42	2, 2, 25, 28, 26, 9	2, 2, 2, 28, 25, 26	2, 2, 2, 2, 2, 25, 26, 28, 25, 25, 26, 28
48	2, 2, 25, 28, 26, 1	2, 2, 2, 28, 25, 26	2, 2, 2, 2, 2, 10,28, 7, 25, 25, 28
54	2, 2, 25, 28, 26, 26	2, 2, 2, 28, 25, 26	2, 2, 2, 2, 2, 10, 28, 26, 25, 25, 28
60	2, 2, 25, 28, 26, 17	2, 2, 2, 1, 25, 26	2, 2, 2, 2, 2, 26, 28, 17, 25, 25, 28
……	The orders will be repeatedly implemented in every 60 s		

broadcasting time orders is presented in Table 17.3. For the sake of convenient to describing, 6 s is taken as a unit, and the information types' broadcasting alignment of the first 60 s is designed particularly. Then the second, the third, and all the following 60 s, the alignment can be implemented repeatedly.

17.3 Performance Evaluation of BDS Augmentation Message

Performance evaluation of BDS augmentation message is carried by the following aspects: message broadcasting efficiency, parameters expressing performances, parameters broadcasting frequency, message demodulation performance and signal capturing performance.

(1) Message broadcasting efficiency

On the assumption that the message length of each frame is N bit, and the effective message parameters practically filling in is n bit. The message broadcasting efficiency is masked as Enav, which is defined as the percentage of the effective message parameters practically filling into the message length of each frame. Namely:

$$E_{nav} = (n/N) \times 100\ \%. \tag{17.1}$$

It is appearance to see that the bigger of the frame length N, the higher of the broadcasting efficiency. According to Fig. 17.1, if leave remain bits out of account, the message broadcasting efficiency is the percentage of the field of message parameters to one whole frame. Consequently, basing the frame structure of Fig. 17.1, the augmentation message broadcasting efficiency is evaluated and the

Table 17.4 Broadcasting efficiency of different frame length

Frame length N (bit)	300	250	200	125
Broadcasting efficiency Enav (%)	87	85	81	70

result is shown in Table 17.4. Moreover, owning to the remain bits in the field of the message parameters, if it is deducted from the efficiency parameters, the communication efficiency will be lower than shown in Table 17.4.

(2) Parameters expressing performances

Parameters expressing performances includes two targets of expressing range and expressing precision. The target of expressing range is used to evaluate the parameters range, and if it can satisfy the parameters' practice physics significance. The target of expressing precision is used to evaluate the parameters' resolving rate, and if it can satisfy the precision requirements in the certain extrapolating time interval. Supposing the bit number of the parameter is m, the scale factor is a, and the scale factor is the resolving rate. Then the expressing range of the parameter is L, and the L value is:

$$L = \begin{cases} [0, (2^m - 1) \cdot a] \text{ no sign bit} \\ [-(2^{m-1} - 1) \cdot a, (2^{m-1} - 1) \cdot a] \text{ the high bit is sign bit.} \end{cases} \quad (17.2)$$

It can be seen that if the bit number is more and the scale factor is bigger, the parameter's expressing range is bigger. If the scale factor is smaller, the parameter's expressing precision is higher. The practical expressing precision of the parameters is correlative with the extrapolate time interval. On the assumption that parameter A_1 is time's one rank constant, and A_2 is time's two rank constant. And their scale factors are $a_1$and a_2 respectively. If the extrapolate time interval is t_L, the practical expressing precision of parameters A_1 and A_2 are respectively equal $a_1 \cdot t_L$ and $a_2 \cdot t_L$ [2].

For there are excessive parameters of the wide area augmentation message, this paper only presents the evaluation method of the parameters' expressing performances, and the particular evaluation results will not states superfluously.

(3) Parameters broadcasting frequency

Parameters broadcasting frequency determines the delay time for the users to obtaining the parameters, and it will affect users' performances. For example, timeliness of the basic navigation messages will affect the first time of positioning, and timeliness of the wide area augmentation messages will affect the alert time of the integrity information. The parameters broadcasting frequency is designed according to the length of one frame, the data rate, broadcasting delay of the parameters and the quantity of the parameters. Corresponding to the broadcasting time orders designed in Table 17.3, the evaluation result of every kind of information types is presented in Table 17.5.

Table 17.5 Parameters broadcasting frequency evaluation

Information type	Requirements of broadcasting frequency (s)	The practical broadcasting frequency (/group)			
		N = 300 (s)	N = 250 (s)	N = 200 (s)	N = 125 (s)
Type1	300	60	60	60	60
Type2	6	6	6	6	6
Type7	120	60	60	60	60
Type9	120	60	60	60	60
Type10	120	60	60	60	60
Type12	3600	60	60	60	12
Type17	3600	120	150	480	390
Type18	600	60	60	60	60
Type25	120	60	72	120	90
Type26	300	60	78	180	180
Type28	120	60	72	120	90

Concluded from Table 17.5, the broadcasting frequency of all types of parameters can satisfy the requirements. The higher of the data rate, the faster of the broadcasting frequency and the better of the timeliness. If the data rate $R_b = 250$ bps, when $N = 250$, the broadcasting frequency is better than when $N = 125$. The biggest broadcasting frequency is decided by the frame length N and the data rate R_b, and the value is $f_{max} = N/R_b$. If the timeliness requirements is higher, such as the fast correction parameters, the broadcasting frequency can be the highest, which can be one time every frame.

(4) Message demodulation performance

The demodulation performance relating to the message design is data rate. Supposing the data rate is R_b (bps), the receiving ratio of carrier to noise is C/N_0 (dBHz), bit ratio of signal to noise is E_b/N_0 (dB), then the error bit rate is [7]:

$$P_e = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right) = \frac{1}{2} erfc\left(\sqrt{\frac{C}{N_0 \cdot R_b}}\right). \tag{17.3}$$

In Eq. (17.3), the error function $erfc(x)$ is a decreasing function with the variable of x. That is, when the carrier to noise ratio C/N_0 is equal, the lower of the data rate R_b, the smaller of the error bit rate P_e. When the limitation of the error bit rate is equal, if the data rate is lower, the limitation of carrier to noise rate will be lower. Taking $P_e = 10^{-6}$ and $P_e = 10^{-5}$ as examples. The calculation results of the carrier to noise ratio limitation with different data rate are presented in Table 17.6. And the coding plus is not taking account. It can be concluded from Table 17.6 that the limitation of carrier to noise ratio varies 0.8–1 dB in that three data rates.

(5) Signal capturing performance

The signal's capturing performance relating with the message designing is the synchronization head. In the navigation message, the Bark-code is often taken as

Table 17.6 The carrier to noise ratio limitation at different data rate

Data rate (bps)	Limitation of carrier to noise ratio (dBHz)	
	$P_e = 10^{-6}$	$P_e = 10^{-5}$
300	36.1	27.8
250	35.3	27.0
200	34.3	26.0

Table 17.7 Detection rate and shielded false alarm rate

J	P_A (%)		P_B (%)
	P_e = 10–6	P_e = 10–5	
4	100	100	36.33
5	100	100	63.67
6	100	100	85.55
7	100	100	96.48
8	100	99.99	99.61

the synchronization head, which is used to ulteriorly affirm the captured signal and eliminate the fault alarm [8]. As show in Fig. 17.1, the BDS augmentation message introduces the Bark-code of 8 bit as the synchronization head. Supposing the bit number of the synchronization head is N_p, the determination limitation of the same code number judging as the receiving signal is J, the detection rate of the signal is P_A, the false alarm rate is P_f, and the shielded false alarm rate on condition of false alarm is P_B. Then:

$$P_A = \sum_{i=J}^{N_p} C_{N_p}^{i} (1 - P_e)^i P_e^{N_p - i} \tag{17.4}$$

$$P_f = \sum_{i=J}^{N_p} C_{N_p}^{i} \left(\frac{1}{2}\right)^{N_p} \tag{17.5}$$

$$P_B = 1 - P_f. \tag{17.6}$$

Taking $P_e = 10^{-6}$ and $P_e = 10^{-5}$ as examples, the signal's detection rate and the shielded false alarm rate with different judging limitations is presented in Table 17.7.

The requirement of signal capture rate generally is 99 %. And it is confirmed from Table 17.7 that the signal detection rate can be satisfied with the requirements. When the same code judging limitation is set as 6, the shielded false alarm rate will be 85.55 %. Considering that the amount of false alarm will be shielded in the phase of capturing, the shielded false alarm rate of 85.55 % is acceptable.

17.4 Conclusions and Applications

The design of navigation messages is a process of multi-elements considered eclectically and at the same time the key element optimized considered. The design scheme will be restricted by excessive aspects which are system scheme, navigation signals' scheme, system performances, etc. In this paper, several doable design schemes of BDS augmentation messages are performed, and the design is basing on BDS global system schemes. Furthermore, the evaluation and comparison is performed aiming at the design schemes. The investigation conclusions afford theory and engineer guidance for satellite navigation system's message design and its performances evaluation. And the study fruits can be used in the designing of BDS system's augmentation message straightly.

It is necessary to be explained that the design scheme is primary, for the study is aiming at describing the design method and the evaluation factors. And the design has great optimization spaces in such aspects as parameters contents, arranging formation and broadcasting time orders.

References

1. Interface Specification (2006) IS-GPS-200, Navstar GPS space segment / navigation user interfaces. Revision D, 2006
2. Glonass Icd (2002) Glonass interface control document. Ver 5.0, 2002
3. Interface Control Document (2005) ICD-GPS-705, Navstar GPS space segment / user segment L5 interfaces. Revision 2
4. Interface Specification (2008) IS-GPS-800, Navstar GPS spaces segment / user segment L1C interfaces. Initial release
5. GAL OS SIS ICD (2010) European GNSS (Galileo) open service signal in space interface control document, Issue 1
6. SC-159 (2006) Minimum operational performance standards for global positioning system / wide area augmentation system airborne equipment. RTCA DO-229D, Dec 13, 2006, pp 27–31
7. Soloviev A, Graas FV, Gunawardena S (2009) Decoding navigation data messages from weak GPS signals. IEEE Trans Aerosp Electron Syst 45(2):660–666
8. Ziemer RE, Peterson RL (2005) Introduction to digital communication, 2nd edn. Prentice Hall, America, pp 306–321

Chapter 18
Validation of GNSS ARAIM Algorithm Using Real Data

Bin Li, Jizhang Sang and Yun Wu

Abstract This paper presents the research results on validation of the GNSS ARAIM (Advanced RAIM) algorithm, which is required in extending GNSS applications to areas demanding high integrity and safety, such as precision approach of LPV-200 phase before landing. In the ARAIM algorithm, the integrity risk of the GNSS systems is quantified by applying theory of multiple-hypothesis (MH). The multiple hypotheses are tested individually by using the method of solution separation (SS). Depending on different scenarios of the potential risks in the GNSS system, the given total tolerated integrity risk is allocated among the satellite-failure cases. For each case, the user's XPL (including HPL, and VPL which is more concerned in this paper), accuracy and EMT, are predicted. In this paper, the ARAIM algorithm is studied using real GNSS observation data from a number of IGS stations at different locations, and then the availability levels of the GPS and GPS/GLONASS systems are evaluated based on the navigation performance requirements of LPV-200. Results show that: (1) VPE calculated using GNSS observation data are consistent with the results of the predicted VPL, accuracy and EMT, which validates the ARAIM algorithm. (2) Under the ARAIM, the highest availability achieved in the GPS system is only 63.04 %, while availability in the GPS/GLONASS system achieves more than 99.0 %, which fully meets the navigation performance requirements of LPV-200.

Keywords Integrity · LPV-200 · Multiple hypothesis solution separation · ARAIM

B. Li (✉) · J. Sang · Y. Wu
School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
e-mail: libin@whu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_18,

18.1 LPV-200 Requirements and GNSS Ranging Error Characteristics

18.1.1 LPV-200 Requirements

Federal Aviation Administration (FAA) has claimed to provide LPV-200 capability worldwide without support of SBAS and Instrument Landing System (ILS) depending on GNSS during 2020–2025. Here, LPV-200 stands for Localizer Performance with Vertical guidance and indicates that decision height is 200 feet above ground level. According to International Civil Aviation Organization (ICAO), the performance requirements of LPV-200 are as follows [3]:

- Probability of Hazardously Misleading Information (HMI) must not exceed 1×10^{-7}/approach.
- False Alert Probability must not exceed 4×10^{-6}/15 s.
- Vertical Alert Limit (VAL) equals 35 m.
- Accuracy must not exceed 4 m, and Effective Monitoring Threshold (EMT) must not exceed 15 m.

18.1.2 GNSS Ranging Error Characteristics

Two sets of parameters below are defined in ARAIM algorithm. One is for accuracy, the other for integrity.

URE is the non-integrity-assured standard deviation of the range component of clock/ephemeris error and is used to evaluate accuracy. It is taken to be the half of each URA value [3].

URA is the least standard deviation of a distribution that bounds the distribution of the range component of clock/ephemeris error in the absence of a fault condition, and is used to evaluate integrity. Its values range from 0.5 to 2.4 m [3].

Two levels of bias magnitudes are defined in ARAIM algorithm. One is a typical magnitude of a bias in a nominal condition (nominal error). This magnitude is used for the evaluation of accuracy. It is fixed at 0.1 m. The other is the maximum bias magnitude used for the evaluation of integrity. The maximum bias magnitude is the maximum only under fault-free conditions. It is 0.75 m with a URA of 0.5 m, and will assume to be 0.5 m with a larger URA value.

Besides URA and URE, errors in the measurements also include tropospheric error, airborne multipath error, and user receiver noise. As integrity affects safety-of-life directly, the corresponding calculations overbound the extreme conditions, while accuracy calculations are based on realistic performance estimates. For these errors, two different airborne error models are used, one set for accuracy, the other for integrity [2, 3, 6]:

- Error model for accuracy

$$\begin{aligned}\sigma_{k,trop} &= \frac{(0.08)(1.001)}{\sqrt{0.002001 + \sin^2(EleAngle)}} \\ \sigma_{k,noise} &= 0.04 - 0.02(EleAngle^\circ - 5)/85 \\ \sigma_{k,mp} &= 0.18 - 0.1(EleAngle^\circ - 5)/85 \\ \sigma_{k,DF} &= a\sqrt{\sigma_{k,noise}^2 + \sigma_{k,mp}^2}\end{aligned}$$

then

$$\sigma_k^2 = URE_k^2 + \sigma_{k,DF}^2 + \sigma_{k,trop}^2 \tag{18.1}$$

- Error model for integrity

$$\begin{aligned}\sigma_{k,trop} &= \frac{(0.12)(1.001)}{\sqrt{0.002001 + \sin^2(EleAngle)}} \\ \sigma_{k,noise} &= 0.04 - 0.02(EleAngle^\circ - 5)/85 \\ \sigma_{k,mp} &= 0.13 + 0.53\exp(-EleAngle^\circ/10) \\ \sigma_{k,DF} &= a\sqrt{\sigma_{k,noise}^2 + \sigma_{k,mp}^2}\end{aligned}$$

then

$$\sigma_k^2 = URA_k^2 + \sigma_{k,DF}^2 + \sigma_{k,trop}^2 \tag{18.2}$$

EleAngle is the satellite elevation angle. $a = \sqrt{\left(\frac{f_1^2}{f_1^2 - f_5^2}\right)^2 + \left(\frac{f_5^2}{f_1^2 - f_5^2}\right)^2}$. The total variance σ_k^2 stands for the nth diagonal element of the weight matrix.

18.2 Realization of ARAIM Algorithm

As vertical geometry strength is weaker than the horizontal, stricter integrity is required in the vertical dimension [3]. So, the study of GNSS ARAIM algorithm in the paper focuses on the vertical dimension.

18.2.1 Multiple Hypothesis Solution Separation

Conventional RAIM assumes that only one fault-satellite happens once a time. It is calculated based on the single-frequency measurement. Its results tend to be conservative. Owing to the limitations above, conventional RAIM is not suitable for integrity monitoring in the LPV-200 anymore. Compared to RAIM algorithm, ARAIM algorithm uses multiple hypothesis solution separation (MHSS), which considers multiple underlying GNSS faults. The GNSS faults can be characterized both in terms of their causes and effects. So it shows a stricter and more flexible superiority to RAIM [4]. Besides, ARAIM focuses on the calculation of Vertical Position Level (VPL) and bases on the smooth pseudorange measurements with carrier phase measurements. It could remove the ionosphere delay. Meanwhile, the threats caused by multiple faults could be allocated through dynamic optimization to reduce the VPL value [1].

The total integrity risk in the ARAIM algorithm is calculated as follows:

$$\begin{aligned} P\{HMI\} = & P\{HMI/H_0\} \cdot P_0 + P\{HMI/H_1\} \cdot P_{sat1} \\ & + \cdots P\{HMI/H_n\} \cdot P_{satn} + P\{HMI/H_{mult}\} \cdot P_{mult} \\ & + P\{HMI/H_{const}\} \cdot P_{const}. \end{aligned} \tag{18.3}$$

$H_i (i = 1, 2, \ldots N)$ is the failure hypothesis, N is the number of satellites in view, $P\{HMI/H_i\}$ is the integrity risk under the hypothesis of H_i, and P_{sat} is the integrity failure rate on individual satellite.

Assuming that all the hypotheses are independent, integrity risk the user bearing could be calculated through the formula above. According to the requirements of ICAO, during the civil aviation navigation using GNSS, integrity risk in the vertical guidance is

$$P_{HMI\_Req} = 1 \times 10^{-7}/approach \tag{18.4}$$

If

$$P\{HMI\} < P_{HMI\_Req}. \tag{18.5}$$

The required GNSS integrity could be reached, and GNSS is available. When P_{sat} is small enough, such as 10^{-5}/approach, the probability of two-satellites-failure is less than P_{HMI_Req} dramatically. Thus, hypotheses of failures of two or more satellites could be removed from the formula (18.3) [7]. Assuming that the removed integrity risk is P_{HMI_Un}, then

$$\begin{aligned} P\{HMI\} = & P\{HMI/H_0\} \cdot P_0 + P\{HMI/H_1\} \cdot P_{sat1} \\ & + \cdots P\{HMI/H_n\} \cdot P_{satn} \\ & < P_{HMI\_Req} - P_{HMI\_Un}. \end{aligned} \tag{18.6}$$

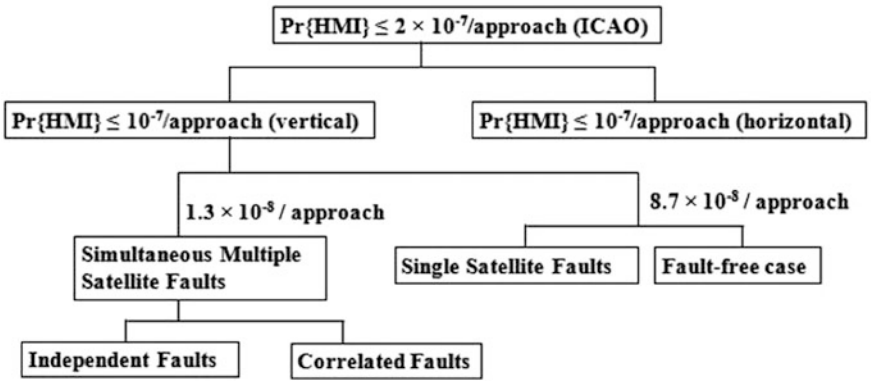


Fig. 18.1 Allocation of P{HMI} requirement for ARAIM [3]

In turn, if P_{sat} is known, given the integrity risk $P\{HMI\}$ and allocated it among all the hypotheses, VPL of corresponding hypothesis could be calculated (Fig. 18.1).

18.2.2 Calculation of VPL

The user' VPL is derived from the fault-free full-set VPL_0 and the faulted subsets VPL_n, given as follows:

$$VPL = \max\{VPL_0, VPL_n\},\ n = 1 \cdots N \tag{18.7}$$

where,

$$VPL_0 = K_{md,0} \times \sigma_{V,0} + \sum_{i=1}^{N} |S_0(3,i)| \times \text{Maximum\_bias}$$

$$VPL_n = D_n + K_{md,n} \times \sigma_{V,n} + \sum_{i=1}^{N} |S_n(3,i)| \times \text{Maximum\_bias}.$$

N is the number of satellites in view. S_0 and S_n are the projection matrixes of each corresponding hypothesis. D_n is the detection threshold for the nth test statistic.

$$D_n = K_{ffd,n} \times \sigma_{dV,n} + \sum_{i=1}^{N} |\Delta S_n(3,i)| \times \text{Nominnal\_error}. \tag{18.8}$$

Tab. 18.1 Necessary parameters settings [3, 5, 6]

P{HMI}	P_{fa}	P_{sat}	Normial error	Maxmium bias	GPS		GLONASS	
					URA	URE	URA	URE
8.7×10^{-8}/approach	4×10^{-6}/15 s	10^{-5}/approach	0.1 m	0.5 m	0.5 m	0.25 m	1.0 m	0.5 m

Three sets of coefficients are introduced here, and they are: $K_{ffd,n}$, $K_{md,0}$, $K_{md,n}$. $K_{ffd,n}$ is determined to meet accuracy and continuity requirement, $K_{md,0}$, $K_{md,n}$ are determined to meet integrity requirement. The continuity and integrity requirements can be met by selecting these coefficients as follows:

$$K_{ffd,n} = -Q^{-1}\left(\frac{Total\_P_{fa}}{2 \times N}\right) \tag{18.9}$$

$$K_{md,0} = -Q^{-1}\left(\frac{P\{HMI\}}{2 \times (N+1)}\right) \tag{18.10}$$

$$K_{md,n} = -Q^{-1}\left(\frac{P\{HMI\}}{2 \times P_{sat} \times (N+1)}\right). \tag{18.11}$$

The total false alert probability requirement is equally divided among the faults of all N satellites in view, and the total integrity risk is equally divided and allocated among the fault-free case and the cases which any single satellite in view has a fault. In the above, Q^{-1} is the inverse of the complement of the one-side standard normal cumulative distribution function (CDF).

All the calculations above are done in the ENU coordinate system.

18.3 Validation of GNSS ARAIM Algorithm Using Real Data

The paper firstly calculates the VPE of hofn station and other stations, including bake, gold and mdvj, using real GNSS (GPS only mode and GPS/GLONASS mode) data of ten continuous days. Then, the validation of ARAIM Algorithm is validated by comparing calculated VPE to the VPL, Accuracy, EMT predicting in the ARAIM. Finally, the availabilities in the GPS only mode and GPS/GLONASS mode of each station are evaluated, according to the LPV-200 performance.

18.3.1 Necessary Parameters Settings

(See Table 18.1)

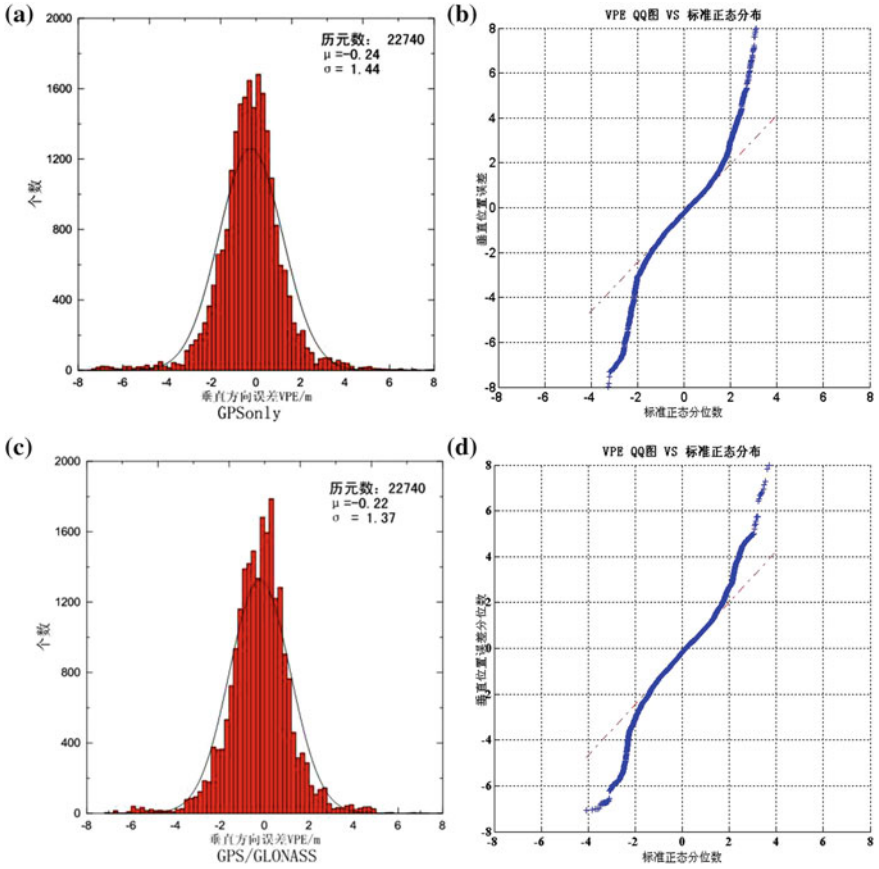


Fig. 18.2 Histogram and Q–Q Plot of VPE in GPS only and GPS/GLONASS modes

18.3.2 Vertical Position Error (VPE)

GNSS VPE of hofn station is calculated by ten continuous days' data. So its performance of range error characterization could be analyzed through two means: histogram and Quantile Quantile Plot.

Figure 18.2a shows the VPE histogram of GPS only mode. The black line stands for normal distribution curve by fitting the VPE histogram. As the mean value and standard deviation of the histogram show, the normalized VPE to the predicted accuracy is more likely to be a standard normal distribution. Figure 18.2b shows the VPE Q–Q Plot of GPS only mode, which displays the quantitative relationship between sample result quantiles of the normalized VPE and theoretical quantiles from a normal distribution. If the distribution of the former is normal, the plot will be close to linear. According to the histogram of the GPS only mode, the corresponding Q–Q plot, Fig. 18.2b, is almost a linear line

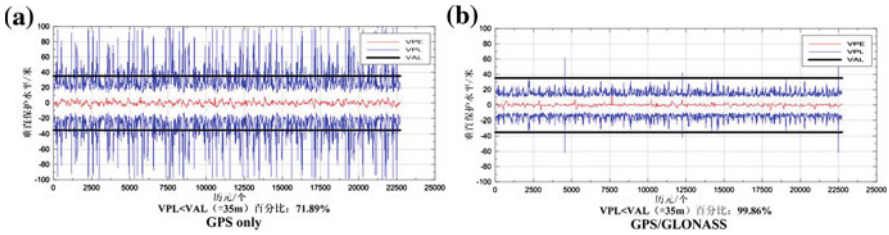


Fig. 18.3 VPL in GPS only and GPS/GLONASS modes

except outside the interval ranging from −2 to 2 m. That means even the histogram approximately looks like a standard normal distribution, though specifically there are some measurements whose range error statistics are not predicted well. Figure 18.2c, d show the VPE histogram and Q–Q Plot of GPS/GLONASS mode. Their explanations are the same as GPS only mode shows.

From the figures above, we can see that the VPE, calculated in two modes, reaches about 2 m. Comparing it with the accuracy estimated by ARAIM, we can analyze whether the accuracy could predict VPE well or not, thus to test the validation of ARAIM in predicting accuracy. Details can be seen in Sect. 18.3.4.

18.3.3 Vertical Protecting Level (VPL)

VPL, the bound of estimated VPE, is related to the requirements in aviation navigation safety. It is one significant aspect of ARAIM evaluation whether or not the VPL bounds the VPE well during navigation operation. VPL must be greater than VPE, but it cannot be too great enough to reduce the continuity of the navigation. However, HMI exists when the VPE is greater than the VPL for longer than the time-to-alert (TTA). So, it should be validated that whether the VPL bounds the VPE well. In the precise approach, we usually calculate the percentage of time that the VPL is less than VAL (35 m) out of total epochs to evaluate the availability GNSS works.

As Fig. 18.3 shows, VPL, predicted in ARAIM of two modes, are both greater than the corresponding VPE which is calculated by real GNSS measurement. We can conclude that HMI does not occur in the situation given the integrity risk and continuity risk. Comparing the predicted VPL and the VAL which is a specified value of LPV-200, there are many cases in which the value of the VPL is less than the VAL. The percentage of time that VPL is less than VAL is 71.89 % in GPS only mode, which is the availability. However, the value of GPS/GLONASS mode reaches 99.86 %, as Fig. 18.3b shows, which meets the requirement of LPV-200. Results above show that more satellites in view, during the GPS/GLONASS navigation, can increase the geometry strength greatly, reducing the VPL value and making it smoother and more steady than that during GPS only navigation.

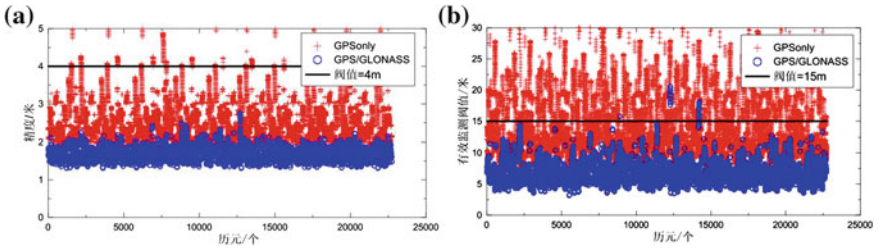


Fig. 18.4 Accuracy and EMT in GPS only and GPS/GLONASS modes

Table 18.2 Percentage of accuracy < 4 m and EMT < 15 m

	Accuracy (%)	EMT (%)
GPS only	97.07	62.74
GPS/GLONASS	100	99.95

18.3.4 Accuracy and EMT

95th percentile vertical accuracy requirement for LPV-200, which is accuracy, is less than 4 m. The EMT is an additional LPV-200 requirement on the vertical position error. GEAS has agreed to require that the 50th percentile absolute value of VPE, which is EMT, should be no greater than 15 m when an integrity alert occurs for the worst-case geometry and for the satellite whose integrity failure is the most difficult to detect. Figure 18.4 displays the accuracy and EMT of each mode.

Statistics of Table 18.2 shows the percentages of time that the accuracy is less than 4 m and EMT is less than 15 m.

From Fig. 18.4a, accuracy of GPS only mode concentrates on about 2 m or more, and 1.5–2.0 m of GPS/GOLNASS mode. The accuracy outside the interval shows a distinct randomness, indicating the accuracy estimated cannot predict the VPE well. Results achieved above are consistent with the statistics of VPE by Q–Q Plot in Sect. 18.3.2, indicating that ARAIM is available and valid in predicting accuracy. From Fig. 18.4b, EMT in the GPS only mode fluctuates dramatically, with its least value 5.7 m and biggest value reaching about 100 m or more. EMT in the GPS/GLONASS mode ranges from 4 to 10 m, shows a little centralized. As more satellites are available when two systems combine, the geometry strength will be enhanced and the reliability and stability of the systems will be improved. From the results above, we can get the conclusion that accuracy and EMT in GPS/GLONASS mode are both superior to the ones in GPS only mode, fully meeting the requirements of LPV-200.

Table 18.3 Availability under GPS only and GPS/GLONASS modes (the last column)

	VPL < 35 m (%)	Accuracy < 4 m (%)	EMT < 15 m (%)	Availability (%)
GPS only	71.89	97.07	62.74	59.18
GPS/ GLONASS	99.86	100	99.95	99.82

Table 18.4 VPL, accuracy, EMT, availability of other three IGS stations (*G* GPS only mode *G/G* GPS/GLONASS mode)

		VPL < 35 m (%)	Accuracy < 4 m (%)	EMT < 15 m (%)	Availability (%)
Bake	G	68.74	98.22	46.38	44.78
	G/G	99.77	100	99.83	99.71
Gold	G	74.08	97.59	69.47	63.04
	G/G	99.25	99.97	99.31	99.11
mdvj	G	57.77	99.04	59.01	49.25
	G/G	99.55	100	99.48	99.25

18.3.5 Availability

Availability is the percentage of time that navigation system works normally in the region it covers, when the accuracy, integrity and continuity are fully met. According to requirements of availability in ARAIM, the ideally ARAIM is declared available if the following three conditions are satisfied: VPL < 35 m, accuracy < 4 m, EMT < 15 m. The availabilities of both the modes are evaluated as Table 18.3 shows:

Comparing to the availability of GPS only mode, which is 59.18 %, the availability in GPS/GLONASS mode reaches up to 99.82 %, which meets the requirement of LPV-200 excellently. Results above show that t ARAIM is available and valid in estimating VPL, accuracy and EMT.

Results of other three IGS stations achieved are as Table 18.4 shows.

18.4 Conclusions

The paper firstly gives a deep study of ARAIM algorithm bases on multiple hypotheses. Then VPL, accuracy, EMT in the GPS only and GPS/GLONASS modes are estimated and analyzed by the ten-day's real GNSS measurement of four IGS stations. Finally, the availabilities in the GPS only and GPS/GLONASS modes of each station are evaluated. Results show that:

(1) The predicted VPL in ARAIM can bound the calculated VPE based on real GNSS measurement well. Additionally, in the GPS only mode, the biggest

percentage of time that VPL is less than VAL (35 m) is 74.08 %, and the least is 57.77 %. However, the biggest percentage in GPS/GLONASS mode reaches up to 99.86 % and the least is 99.25 %, more superior to the ones in the GPS only mode.

(2) Accuracy predicted in the ARAIM are both less than 4 m in the two modes. The percentage that accuracy less than 4 m in GPS/GLONASS mode is a little better than that in GPS only mode. This is consistent with the results of the real navigation. Meanwhile, when estimating EMT by ARAIM algorithm, all the EMT in GPS/GLONASS mode are less than 15 m, which are better than ones in GPS only mode.

(3) The biggest availability in GPS only mode of LPV-200 is 63.04 %, while the one in the GPS/GLONASS mode are all greater than 99 %, fully meeting the performance requirements of LPV-200.

Results above are achieved based on the static measurement of IGS stations. However, during the precise approach of LPV-200, the receiver is in the dynamic condition of high-speed. The quality of data the receiver gets is worse than that of the static measurement. So, results in ARAIM are relatively optimistic. With the combination of three or more navigation systems among GPS, Beidou, GLONASS and Galileo in the future, the geometry strength of the satellites will improve obviously. Thus, we can foresee that it will be sure to achieve LPV-200 capability worldwide at airports without local GNSS instrumentation relying on ARAIM.

References

1. Boris S, Pervan SPP, Christie JR (1998) A multiple hypothesis approach to satellite navigation integrity. Navig J Inst Navig 45(1):61–71 (Spring Printed in U.S.A)
2. Choi M, Blanch J, Akos D et al. (2011) Demonstrations of multi-constellation advanced RAIM for vertical guidance using GPS and GLONASS signals. Paper presented at proceedings of 24th international technical meeting of the satellite division of the institute of navigation Portland OR, 19–23 September
3. GEAS Panel (2010) Phase II of the GNSS evolutionary architecture study
4. Juan B, Todd W, Per E (2012) Satellite navigation for aviation in 2025. Proc IEEE 100:1821–1830
5. Todd W, Per E, Juan B, Boris P (2008) Worldwide vertical guidance of aircraft based on modernized GPS and new integrity augmentations. Paper presented at proceeding of the IEEE
6. Young CL, Michael PMc (2007) Feasibility analysis of RAIM to provide LPV-200 approaches with future GPS, ION GNSS. Paper presented at 20th international technical meeting of the satellite division, pp 25–28
7. Yun W, Jinling W, Yiping J (2013) Advanced receiver autonomous integrity monitoring (ARAIM) schemes with GNSS time offsets. Adv Space Res 52(1):52–61

Part III
Satellite Navigation Models and Methods

Chapter 19
Real-Time and Fast Retrieve the Coseismic Wave by GPS, Strong-Motion Combined Measurements and Broadcast Ephemeris

Rui Tu

Abstract In this study, we propose an approach of real-time and fast retrieving the coseismic waves by GPS, strong-motion combined measurements and broadcast ephemeris. Firstly, the velocity observation equations and state equations of the strong-motion measurements are introduced into Precise Point Positioning (PPP) solution model, the baseline shifts of the strong-motion are real-time estimated as unknown parameters like other positioning parameters by a Kalman filter. Then, the true velocity wave was retrieved by correcting the estimated baseline shifts, the true displacement wave was recovered by the velocity integrated displacement after de-trending a linear correction. The series of validation results have shown that, this method can fast and accurate retrieve the coseismic wave, the convergence time is smaller than one minute, and the precision are better than 2 mm/s and 2.5 cm for velocity and displacement respectively.

Keywords GPS and strong-motion combined measurements · Broadcast ephemeris · Coseismic wave retrieving · Real-time

19.1 Introduction

Strong-motion sensor is a kind of acceleration survey equipment, it with advantages of low-cost and high sensitivity, and is commonly used for earthquake monitoring and early warning [1–4]. However, it can only provide high-precision acceleration waves, the velocity and displacement waves are obtained by the acceleration integration [5–7]. As the tilting and/or rotation and some other reasons, the strong-motion records probably contain baseline shifts, they will be

R. Tu (✉)
The Germany Research Center for Geosciences, Telegrafenberg 14473 Potsdam, Germany
e-mail: turui-2004@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_19,

further amplified if integrate into velocity or displacement. Currently, many empirical methods are used for the baseline shift correction [8–12]. But most of them involve large and unquantifiable uncertainties compare with the geodetic survey results, and can not be operated in real-time.

With the development of space geodesy technology, especially for the GPS, it has been an essential measurements for real-time earthquake monitoring and early warning. The Precise Point Positioning (PPP) and network solution are two basic solution models to retrieve high precision displacements. However, as the problems of the GPS system designs and receiver sampling rate, the GPS can only obtain high precision displacement, the velocity and acceleration are easily polluted by the high-frequency noises [5–7].

How to use the GPS, strong-motion combined measurements and complement the advantages of each other, make it neither affects by the strong-motion's baseline shifts nor affects by the GPS noise pollution and can provide high precision coseismic information is a hot topic in GPS seismology. Currently, many researchers have studied an integrated processing of high-rate GPS and strong-motion records, their researches can be classified into two categories. The first one is loose integration, it mainly use the high-precision GPS displacement to correct the baseline shift and recover the coseismic waves, such as [13, 14, 15]. The second one is tight integration, the coseismic waves are directly estimated by the GPS and acceleration raw observations, such as [16, 17, 15].

For these integration methods, one important condition is that the GPS observation must be with high precision. The network solution can provide the relatively displacement with mm precision, but it needs many reference stations for simultaneous observation and is difficult for real-time operation [13]. The dynamic PPP technology can provide the absolute displacement with cm precision, but needs a long time initialization [18, 19]. In additional, the data solution needs high-precision ephemeris and with poor timeliness.

For the application of earthquake monitoring and early warning, the goals are not only accurate and reliable, but also are easy for real-time operation and with high efficient. Based on these backgrounds, this paper propose an approach of real-time and fast retrieving the coseismic wave by GPS, strong-motion combined measurements and broadcast ephemeris. Firstly, the velocity observation equations and state equations of the strong-motion measurements are introduced into the PPP solution model, the baseline shifts of the strong-motion are real time estimated as unknown parameters like other positioning parameters by a Kalman filter. Then, the true velocity wave was retrieved by correcting the estimated baseline shifts, the true displacement information was retrieved by the velocity integrated displacement after de-trending a linear correction. The experimental and real earthquake data validation results have shown that, it can fast and accurate retrieve the coseismic waves.

19.2 Methodology

The approach is based on GPS, strong-motion combined measurements and broadcast ephemeris product, the dynamic PPP was used to estimate the baseline shift while adding the strong-motion observation firstly, then the coseismic wave was recovered by correcting the baseline shift and de-trending the linear correction. In this paper, the acceleration records of the strong-motion were firstly removed the initial baseline shift and integrated into velocity for the combined solution to recover the coseismic waves.

19.2.1 Mathematical Model

While the GPS using the ionosphere-free observation and consider the strong-motion's baseline shifts, the observation equations of the tight integration model can be written as following.

$$\begin{bmatrix} L_c \\ P_c \end{bmatrix}_k = \begin{bmatrix} e_k & m_{z,k} & 1 & 1 \\ e_k & m_{z,k} & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ z_k \\ \delta t_k \\ b \end{bmatrix} + \begin{bmatrix} \varepsilon_{Lc} \\ \varepsilon_{Pc} \end{bmatrix}, \begin{matrix} \varepsilon_{Lc} \sim N(0, \delta_{Lc}^2) \\ \varepsilon_{Pc} \sim N(0, \delta_{Pc}^2) \end{matrix} \tag{19.1}$$

$$v_k = \dot{x}_k - u_k + w_{vk} \tag{19.2}$$

where, L_c and P_c are the observed minus computed phase and pseudorange observations from satellite to receiver respectively. v and u represent the strong-motion velocity observation and baseline shift. e is the unit direction vector from satellite to receiver. m is the wet tropospheric mapping function x, $\dot{x}$ denotes the vector of the receiver position and velocity. $z, \delta t, b$ are the tropospheric zenith delay, receiver clock and phase ambiguities. ϵ is the measurement noise, its variances is δ^2. k is the epoch number. The error components such as the antenna center offsets and variations, relativistic effect, tide loading and phase wind-up can be corrected with the existing methods [20, 21]. The baseline shifts are estimated as a random walk process, some other parameters like troposphere, receiver clock and phase ambiguities are treated the same as the traditional PPP.

The state equations for the station movement and baseline shifts are expressed as follows.

$$\begin{bmatrix} x_k \\ \dot{x}_k \\ u_k \end{bmatrix} = \begin{bmatrix} I & 0.5I\tau & 0.5I\tau \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} x_{k-1} \\ \dot{x}_{k-1} \\ u_{k-1} \end{bmatrix} + \begin{bmatrix} 0.5I \\ I \\ 0 \end{bmatrix} [v_k] + w_k \tag{19.3}$$

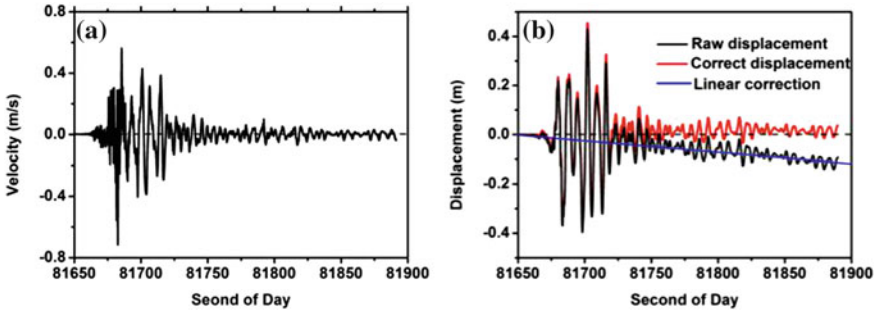


Fig. 19.1 De-trending the linear correction (**a** represents the recovered velocity, **b** represents de-trending the shift by a linear correction)

where, τ is the smallest sampling rate of the integrated sensors. w_k is the dynamic noise; with the expectation $E(w_k) = 0$ and variance $D(w_k) = \begin{bmatrix} \frac{\tau^3}{3}q_u^2 I & \frac{\tau^2}{2}q_u^2 I & 0 \\ \frac{\tau^2}{2}q_u^2 I & \tau q_u^2 I & 0 \\ 0 & 0 & \tau q_u^2 I \end{bmatrix}$ respectively, q_u is the power density of the baseline shift.

With these equations, a Kalman filter can be employed to solve the baseline shifts and some other parameters [22]. It is emphasized that as the GPS sample rate usually not matches the strong-motion's sample rate, so the filter can only be employed on the GPS sample points, otherwise only for the prediction [23].

19.2.2 De-Trending the Linear Correction

As the broadcast ephemeris with a low precision and the ambiguities and atmosphere parameters can not estimate accurately in a short period, the solved baseline shift may have a constant bias. It will results in the corrected velocity with a bias, it is smaller than 2 mm/s and can be negligible for the inversion analysis (shown in Fig. 19.1a). But it will lead to a large linear trend while integrated into displacement, as shown in Fig. 19.1b (in black). The trend is very stable in a short period and can be removed easily by a linear correction (in blue), then the coseismic displacements are recovered (in red).

19.2.3 The Implementation Process of the Combined System

Figure 19.2 shows the implementation flow of the approach. The first step, retrieve the high-frequency acceleration records from the strong-motion measurements, and integrated into velocity after remove the initial baseline shifts. The second

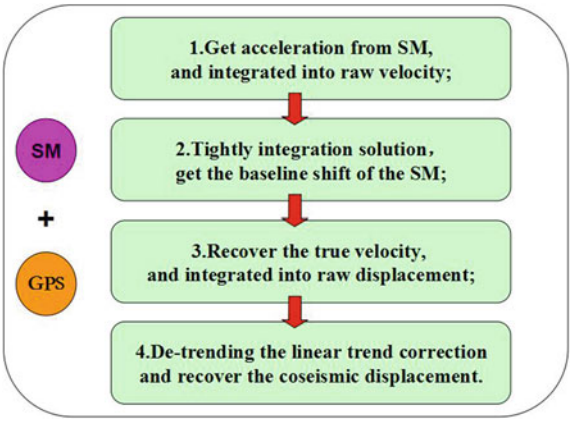


Fig. 19.2 The flow of the combined system

step, estimate the strong-motion's baseline shift by the tight integration of GPS raw observation and strong-motion velocity. The third step, recover the true velocity by correcting the estimated baseline shift and integrated it into raw displacement. The fourth step, de-trend the linear trend correction and recover the coseismic displacement.

19.3 Validation and Analysis

The following data was used for validation and analysis. The data is processed in a simulated real-time mode, and the broadcast ephemeris was used for data solution. The acceleration records of the strong-motion were removed the initial shifts by the empirical method of [12] firstly, then integrated into velocity time series for the combined solution.

19.3.1 Analysis of the Experimental Data

Firstly, the experimental dataset which carried out on December 2012 at GFZ was used for the validation. It contains a platform, which can be slide along a table, includes a dynamic GPS antenna, and a strong-motion sensor, the details can be found in [15]. The sampling rates are 50 and 100 Hz for the GPS and the strong-motion sensor respectively, we simulated eight simulations, the GPS data was utilized to 1 Hz for the data analysis. In Fig. 19.3, the left side shows the raw integrated velocity (black) and the estimated baseline shift (red). We can clearly see that for the eight simulations, they all have large baseline shift and the integrated velocity are seriously polluted. The middle side shows the time series of the corrected velocity, the velocity was shaken between the zero line when there is no

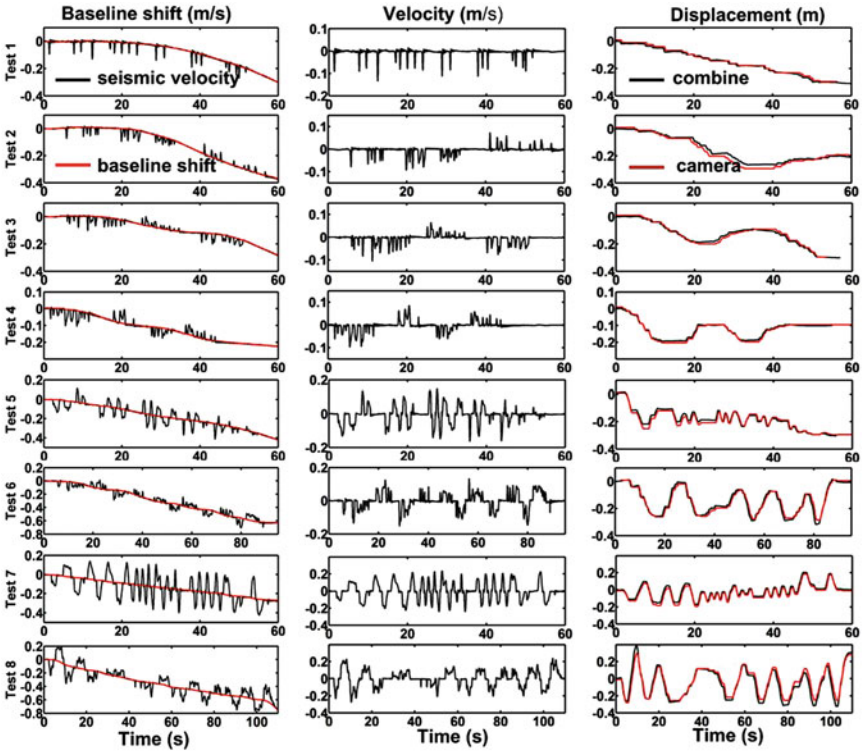


Fig. 19.3 The combined results of the eight simulations

movement, comparison to the raw velocity, it is corrected very well. The right side shows the recovered displacement (in black) and the reference displacement (in red) which provided by the high-frequency camera with a precision of 3 mm, they are consistent with each other very well.

19.3.2 Analysis of the Real Earthquake Data

Furthermore, we collected the co-located GPS and seismic data during the E1 Mayor-Cucapah earthquake (Mw 7.2, 4 April 2010, 22:40:42 UTC) to estimate the three-dimensional (3D) baseline shifts and recover the coseismic velocity and displacement. Four pairs of co-located GPS and strong-motion seismometer are selected and processed. The GPS station P496, which is located 73 km from the earthquake epicenter, is co-located with the seismic station 0931 of the Southern California Seismic Network (SCSN) with a distance of about 10 km. GPS station P496 located 74 km from the epicenter is about 140 m away from its co-located

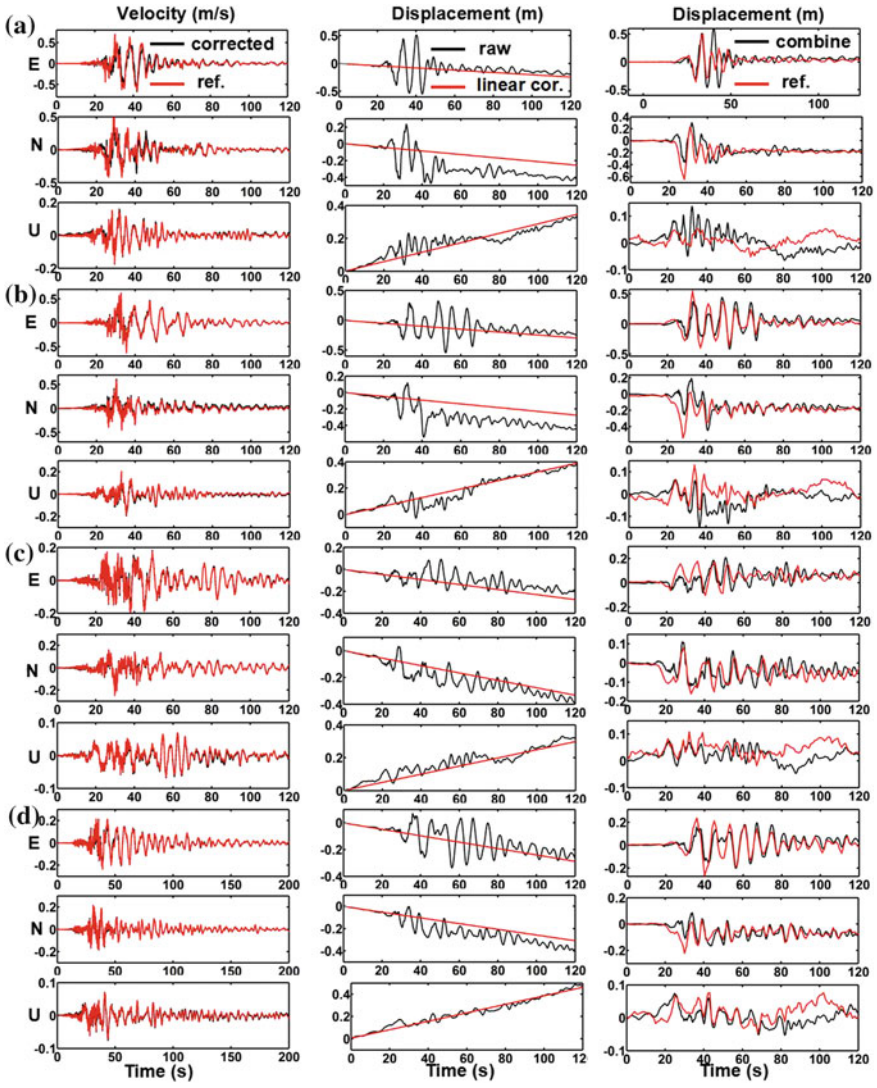


Fig. 19.4 The results of the four combined paries (**a**, **b**, **c**, **d** represents the combined paries of P494/0931, P496/5058, P500/5054, P744/5028 respectively)

seismic station 5058. GPS station P500 located 63 km from the epicenter is about 3.6 km away from its co-located seismic station 5054. GPS station P744 located 80 km from the epicenter is about 275 m away from its co-located seismic station 5028. The accelerometer data is 200 Hz, and the GPS data is recorded at 5 Hz. Figure 19.4a–d are the results of the four paris (P494/0931, P496/5058, P500/5054, P744/5028) in ENU components. The left side are the velocity comparison, where the reference velocity (red) was provided by the USGS post solution

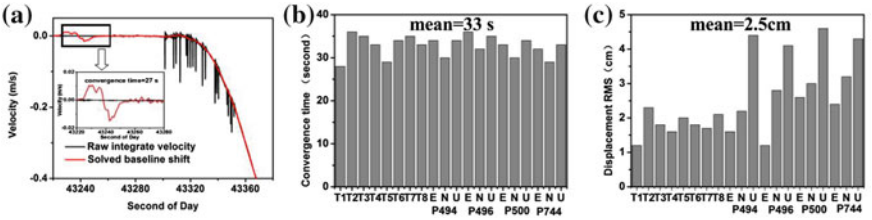


Fig. 19.5 **a** presents the convergence process of the estimated baseline shift, **b**,**c** presents the statics of the convergence time and precision respectively

(http://nsmp.wr.usgs.gov/datasets/20100404_2240.html), we can concluded that the recovered velocity waves are consistent with the USGS references. The middle side shows the raw integrated displacement (in black) and the linear correction (in red). The right side shows the comparison of the de-trend displacement (in black) and the GPS post solution displacement (in red). The real-time and fast retrieved displacement are well consistent with the post solutions in horizontal component, while in vertical component, as the atmosphere error is difficult to process, the recovered displacement with large uncertainties.

19.3.3 The Analysis of the Convergence and Precision

Figure 19.5 shows the convergence time and the RMS values of the displacement bias, where the convergence time was statical by the bias of the solved baseline shift is smaller than 2 mm/s (in Fig. 19.5a, the red color represents the solved baseline shift, the black color represents the integrated velocity). The camera recorded movement tracks (the precision is better than 3 mm) were used for the simulated experimental precision analysis, and for the real earthquake, the GPS post solution (the precision is better than 1 cm) was used as a reference. From the Fig. 19.5b, c the average convergence time is 33 s for the combined system, and the precision of recovered displacement is 2.5 cm, it is much better than the traditional PPP solutions.

19.4 Conclusions

This paper propose an approach of real-time and fast retrieving the coseismic wave by GPS, strong-motion combined measurements and broadcast ephemeris. In this method, the acceleration of the strong-motion was integrated into velocity firstly, then tightly integrated with the GPS raw observations to estimated the strong-motion's baseline shifts, finally the true coseismic wave was recovered by correcting the baseline shifts and removing a linear correction. After introducing the

algorithms, we make a validation analysis by two different datasets. The series results have shown that, the baseline shifts of the strong-motion are recognized and corrected effectively, and high-precision coseismic information is recovered, the precision are better than 2 mm/s and 2.5 cm for velocity and displacement respectively. Compare with some other combination methods, this approach not only has a short convergence time but also can provided high-precision and broadband displacement, velocity and acceleration waves. Compare with empirical baseline shift correction methods, which using the empirical double broken line to correct the velocity, the present method real-time estimated and corrected the transient baseline shifts, it is more objective and accurate. In additional, the data solution using the broadcast ephemeris and stand-alone data, it is easy for real-time application.

References

1. McComb HE, Ruge AC, Neumann F (1943) The determination of true ground motion by integration of strong-motion records: a symposium. Bull Seismol Soc Am 33:1–63
2. Graizer VM (1979) Determination of the true ground displacement by using strong motion records. Izv Earth Phys 25:26–29
3. Chiu H (1997) Stable baseline correction of digital strong-motion data. Bull Seism Soc Am 87:932–944
4. Zhu L (2003) Recovering permanent displacements from seismic records of the June 9, 1994 Bolivia deep earthquake. Geophys Res Lett 30:1740. doi:10.1029/2003GL017302
5. Elósegui P, Davis JL, Oberlander D, Baena R, Ekströ G (2006) Accuracy of high-rate GPS for seismology. Geophys Res Lett 33:L11308. doi:10.1029/2006GL026065
6. Genrich JF, Bock Y (2006) Instantaneous geodetic positioning with 10–50 Hz GPS measurements: noise characteristics and implicationsfor monitoring networks. J Geophys Res 111:B03403. doi:10.1029/2005JB003617
7. Larson KM, Bilich A, Axelrad P (2007) Improving the precision of high-rate GPS. J Geophys Res 112:B05422. doi:10.1029/2006JB004367
8. Iwan W, Moser M, Peng C (1985) Some observations on strong-motion earthquake measurement using a digital accelerograph. Bull Seism Soc Am 75:1225–1246
9. Boore DM (2001) Effect of baseline corrections on displacement and response spectra for several recordings of the 1999 Chi-Chi, Taiwan, earthquake. Bull Seism Soc Am 91:1199–1211
10. Graizer V (2006) Tilts in strong ground motion. Bull Seism Soc Am 96:2090–2106
11. Wu Y, Wu C (2007) Approximate recovery of coseismic deformation from Taiwan strong-motion records. J Seismol 11:159–170
12. Wang R, Schurr B, Milkereit C, Shao Zh, Jin M (2011) An improved automatic scheme for empirical baseline correction of digital strong-motion records. Bull Seism Soc Am 101:2029–2044
13. Bock Y, Melgar D, Crowell BW (2011) Real-time strong-motion broadband displacements from collocated GPS and accelerometers. Bull Seism Soc Am 101:2904–2925
14. Wang R, Parolai S, Ge M, Jin MP, Walter TR, Zschau J (2013) The 2011 Mw 9.0 Tohoku earthquake: comparison of GPS and strong-motion data. Bull Seism Soc Am. doi:10.1785/0120110264
15. Tu R, Wang R, Ge M, Walter TR, Ramatschi M, Milkereit C, Bindi D, Dahm T (2013a) Cost effective monitoring of broadband strong ground motion related to earthquakes, landslides or

volcanic activities by joint use of a single-frequency GPS and a MEMS-type accelerometer. Geophys Res Lett. doi:10.1002/grl.50653
16. Geng J, Bock Y, Melgar D, Crowell BW, Haase JS (2013) A new seismogeodetic approach applied to GPS and accelerometer observations of the 2011 Brawly seismic swarm: implications for earthquake early warning. Geophys Geosyst, Geochem. doi:10.1002/ggge.20144
17. Li X, Ge M, Zhang Y, Wang R, Klotz J, Wicket J (2013) Tightly integrated processing of high-rate GPS and strong-motion data: Application to earthquake early warning. Geophys J Int. doi:10.1093/gji/ggt249
18. Zumberge JF, Heflin MB, Jefferson DC, Watkins M (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102:5005–5017
19. Bisnath S., Gao Y (2007) Current state of precise point positioning and future prospects and limitations. Observing our changing earth, international association of Geodesy Symposia 133, Springer-Verlag. Berlin Heidelberg
20. Dach R, Hugentobler U, Fridez P, Michael M (eds) (2007) Bernese GPS software version 5.0. Astronomical Institute, University of Bern
21. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solut 5(2):12–28. doi:10.1007/PL00012883
22. Tu R, Ge M, Wang R, Walter TR (2013) Real-time seismic data processing: a new algorithm for tight integration of high-rate GPS and strong-motion records, demonstrated for experimental data and the 2010 Baja California earthquake (Mw 7.2). J Seismol. doi:10.1007/s10950-013-9408-x
23. Smyth A, Wu M (2006) Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring. Mech Syst Signal Process 21:706–723

Chapter 20
Effect of Separation of Navigation Satellite Antenna Inter-Frequency Phase Centers on T_{GD} Parameter

Jiaxing Liu, Shaojun Bi, Jinjun Zheng and Jun Xie

Abstract The geometric separation among the navigation satellite antenna phase centers at different frequencies affects the accuracy of T_{GD} measurement. According to the measurement and utilization method of T_{GD} parameter, this paper establishes the relational model between the parameter and the satellite inter-frequency phase center separation. Whereas the relative motion of the satellite, the monitor station and the sun has a performance of complex diversity, this paper obtains an approximate expression of T_{GD} estimating error about phase center separation, along with its effect on user ranging accuracy. Afterwards in the condition of single-station monitoring and double-station monitoring, the maximum T_{GD} error is respectively deduced based on the worst geometry between the satellite and the station(s). Simulations result in numerical values and variation trends of the maximum T_{GD} error with phase center separation. The work could be used in theoretical analysis and estimation of navigation accuracy, and also in design and manufacture of navigation satellites and monitor stations.

Keywords Inter-frequency phase center separation · T_{GD} estimating error · User ranging accuracy · Single-station monitoring · Double-station monitoring

20.1 Introduction

In order to improve the user's positioning and timing accuracy, as well as to obtain ionosphere characters via a satellite navigation system, the method of measuring the satellite inter-frequency channel delay bias called T_{GD} parameter is widely discussed. However there is a lack of research on its error sources and their

J. Liu (✉) · S. Bi · J. Zheng · J. Xie
Institute of Spacecraft System Engineering, CAST, Beijing 100094, China
e-mail: wyliujiaxing@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_20,

corresponding effects. Due to the frequency dispersion of the transmitting channel, there is geometric separation to a certain extent among the navigation satellite antenna phase centers at different frequencies, which reduces the accuracy of T_{GD} measurement. Moreover, affected by the ambient temperature and pressure, the position of the on-orbit antenna phase center changes relative to prelaunching, and moves ceaselessly with the satellite attitude, resulting in high difficulty for the separation to be corrected accurately. Thus it is significant to investigate the effect of the trait of satellite antenna phase centers on T_{GD} parameter.

20.2 Measurement and Utilization of T_{GD} Parameter

The measurement ranging from the satellite obtained by the ground station is

$$\rho_i = R_i/c + \delta t_r - (\delta t_s - \tau_i) + A \cdot \mathrm{TEC}/(cf_i^2) + d_{trop} + \tau_{ri} \tag{20.1}$$

where ρ_i is the pseudorange measurement at f_i by the ground station (GS), R_i is the distance between dot C_i to dot C_{MS}, δt_r is the bias of the GS clock relative to the navigation system clock, δt_s is the bias of the satellite clock relative to the system clock, τ_i is the delay of the f_i navigation signal in the satellite transmitting channel, A is a constant of 40.3 m·Hz^2 (e/m^2)$^{-1}$, TEC is the total electron content (e/m^2) in the unit square along the transmission path from satellite to GS, c is the light speed, d_{trop} is the troposphere delay, τ_{ri} is the delay of the f_i navigation signal in the GS receiving channel, and $i = 1, 2$. According to formula (20.1), expression on the differential group delay $(\tau_1 - \tau_2)$ of the satellite inter-frequency signal is

$$(\rho_1 - \rho_2) = (R_1 - R_2)/c + (\tau_1 - \tau_2) + \kappa \cdot \mathrm{TEC} + (\tau_{r1} - \tau_{r2}) \tag{20.2}$$

where κ is a constant expressed by

$$\kappa = \frac{40.3}{c} \cdot \frac{f_2^2 - f_1^2}{f_1^2 f_2^2}. \tag{20.3}$$

The total electron content (TEC) could be modelled by the polynomial below:

$$\mathrm{TEC} = s(\mathrm{el}) \cdot \mathrm{TEC}_v = s(\mathrm{el}) \cdot \left(c_1 + c_2\phi + c_3\lambda_{cr} + c_4\phi^2 + c_5\lambda_{cr}^2 + c_6\phi\lambda_{cr}\right) \tag{20.4}$$

where s(el) is the obliquity factor as a function of satellite elevation angle el (abbreviated to s in the following paper), TEC_v is the vertical TEC, φ and λ_{cr} are the ionospheric intercept latitude and longitude in the sun-fixed co-rotating reference frame, c_1, …, c_6 are the ionosphere model coefficients.

Estimation methods [1, 2] can be used to gain the differential group delay bias combined by those from the satellite on orbit and the receiver. If the receiver

differential delay bias is known, then the satellite differential group delay bias can be obtained in the case of superposition for the satellite antenna phase centers

$$(\rho_1 - \rho_2) - (\tau_{r1} - \tau_{r2}) = (\tau_1 - \tau_2) + \kappa s\left(c_1 + c_2\phi + c_3\lambda_{cr} + c_4\phi^2 + c_5\lambda_{cr}^2 + c_6\phi\lambda_{cr}\right). \tag{20.5}$$

The receiver delay bias in formula (20.5) can be measured by a signal generator, a frequency converter and a vector network analyzer. When a certain satellite is observed, formula (20.5) contains seven unknown variables, namely the differential group delay bias and the ionosphere model coefficients. After observing for a typical period of 5 h, Least Square method can be used.

Observation equations based on formula (20.5) is shortened by

$$y = \boldsymbol{H}x \tag{20.6}$$

where

$$\begin{cases} y = \begin{bmatrix} ((\rho_1 - \rho_2) - (\tau_{r1} - \tau_{r2}))_1 \\ \vdots \\ ((\rho_1 - \rho_2) - (\tau_{r1} - \tau_{r2}))_n \end{bmatrix}_{n\times 1} \\ x = \begin{bmatrix} (\tau_1 - \tau_2) & c_1 & \cdots & c_6 \end{bmatrix}^{\mathrm{T}}_{7\times 1} \\ \boldsymbol{H} = \begin{bmatrix} 1 & \kappa s & \kappa s\phi_1 & \kappa s\lambda_{cr1} & \kappa s\phi_1^2 & \kappa s\lambda_{cr1}^2 & \kappa s\phi_1\lambda_{cr1} \\ 1 & \kappa s & \kappa s\phi_2 & \kappa s\lambda_{cr2} & \kappa s\phi_2^2 & \kappa s\lambda_{cr2}^2 & \kappa s\phi_2\lambda_{cr2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \kappa s & \kappa s\phi_n & \kappa s\lambda_{cr}n & \kappa s\phi_n^2 & \kappa s\lambda_{cr}n^2 & \kappa s\phi_2\lambda_{cr}n \end{bmatrix}_{n\times 7} \end{cases} \tag{20.7}$$

The errors of the observations are mainly the inter-channel noise error differences and multipath error differences of the receiver at the monitoring station at different epochs, and perform the stochastic characteristic of independent and identical distribution. Thus the solution is

$$x = \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}y = \boldsymbol{B}y \tag{20.8}$$

According to the definition of the differential group delay bias T_{GD} [3], it can be obtained as following:

$$T_{GD} = \frac{\tau_1 - \tau_2}{1 - \gamma} = \frac{\mathbf{x}_1}{1 - \gamma} \tag{20.9}$$

where $\boldsymbol{x}_1$ is the first element of vector $\boldsymbol{x}$, and γ is a constant,

$$\gamma = \frac{f_1^2}{f_2^2}. \tag{20.10}$$

According to the test results of the GPS Block-I satellites [2, 4], although T_{GD} can be obtained from the prelaunching calibration data, still needs T_{GD} to be estimated accurately for the satellites on orbit, considering the in-orbit ambient temperature variations and aging of the on-board devices.

A navigation user receiver extracts T_{GD} parameter and satellite clock correction parameters from navigation messages, and calculates the bias Δt_{sv} of the satellite clock relative to the system clock by clock parameters. The user who corrects the ionosphere delay by f_1 and f_2 double-frequency signals can use Δt_{sv} as satellite clock bias to correct the pseudoranges directly. While the user with an f_1 or f_2 single-frequency signal needs to correct the satellite clock bias for measurements on f_1 and f_2 respectively by following formulae

$$\begin{cases} (\Delta t_{sv})_1 = \Delta t_{sv} - T_{GD} \\ (\Delta t_{sv})_2 = \Delta t_{sv} - \gamma T_{GD} \end{cases}. \tag{20.11}$$

20.3 Estimation Error of T_{GD} Parameter

T_{GD} parameter reflects the relation of delays for signals at different frequencies in navigation satellite transmitting channels. There is geometric separation among the navigation satellite antenna phase centers at different frequencies, which affects estimating precision for T_{GD}. It can be proven from formula (20.7) and (20.8) that

$$\boldsymbol{BI}_{n\times 1} = \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}[\,1 \quad 1 \quad \cdots \quad 1\,]^{\mathrm{T}}_{n\times 1} = [\,1 \quad 0 \quad \cdots \quad 0\,]^{\mathrm{T}}_{7\times 1}. \tag{20.12}$$

If there is no superposition for the satellite antenna phase centers at different frequencies, then

$$(\rho_1 - \rho_2) - (\tau_{r1} - \tau_{r2}) - (R_1 - R_2)/c = (\tau_1 - \tau_2) + \kappa s\left(c_1 + c_2\phi + c_3\lambda_{cr} + c_4\phi^2 + c_5\lambda_{cr}^2 + c_6\phi\lambda_{cr}\right). \tag{20.13}$$

Observation equations based on formula (20.13) can be rewritten as

$$y' = \boldsymbol{H}x' \tag{20.14}$$

where

$$y' = y + e_y = y + [\,-(R_1 - R_2)_1/c \quad \cdots \quad -(R_1 - R_2)_n/c\,]^{\mathrm{T}}_{n\times 1}. \tag{20.15}$$

Thus the solution error of the unknown vector caused by insuperposition for phase centers is

$$e_x = x' - x = \boldsymbol{B}\boldsymbol{e}_y. \tag{20.16}$$

Consequently estimating error of T_{GD} caused by insuperposition can be deduced according to formula (20.9) and (20.16)

$$e_{T_{\mathrm{GD}}} = \left|\frac{(\boldsymbol{e}_x)_1}{1-\gamma}\right| = \left|\frac{(\boldsymbol{B}\boldsymbol{e}_y)_1}{1-\gamma}\right| = \left|\frac{\boldsymbol{B}_{11}(\boldsymbol{e}_y)_1 + \cdots + \boldsymbol{B}_{1n}(\boldsymbol{e}_y)_n}{1-\gamma}\right| \tag{20.17}$$

It can be seen that the estimating error for T_{GD} not only depends on the two frequencies, but also relates to the sun-fixed coordinate of the ionospheric pierce dot, the satellite elevation for the ground station, the satellite attitude, and geometric separation of the satellite inter-frequency phase centers at all the observation epochs. In a short satellite observation arc,

$$(e_y)_1 \approx (e_y)_2 \approx \cdots \approx (e_y)_n \tag{20.18}$$

Formula (20.17) can be shortened to an approximate expression shown below with the help of formula (20.18) and (20.12). This expression of T_{GD} estimating error is irrelative to the sun-fixed coordinate of the pierce at all the observing epochs, and relative to spatial geometry of the satellite and the ground station as well as the satellite phase center separation.

$$c \cdot e_{T_{\mathrm{GD}}} \approx \left|\frac{(\boldsymbol{B}_{11} + \cdots + \boldsymbol{B}_{1n}) \cdot c \cdot (e_y)_n}{1-\gamma}\right| = \left|\frac{R_2 - R_1}{1-\gamma}\right| \tag{20.19}$$

The double-frequency user's ranging error is not concerned with the estimation error of T_{GD} parameter, but concerned with the estimation error of the satellite clock bias. According to formula (20.11), the respective error of the f_1 or f_2 single-frequency user's ranging measurement caused by T_{GD} error is

$$\begin{cases} \Delta\rho_1 = c \cdot e_{T_{\mathrm{GD}}} \\ \Delta\rho_2 = c \cdot \gamma e_{T_{\mathrm{GD}}} \end{cases}. \tag{20.20}$$

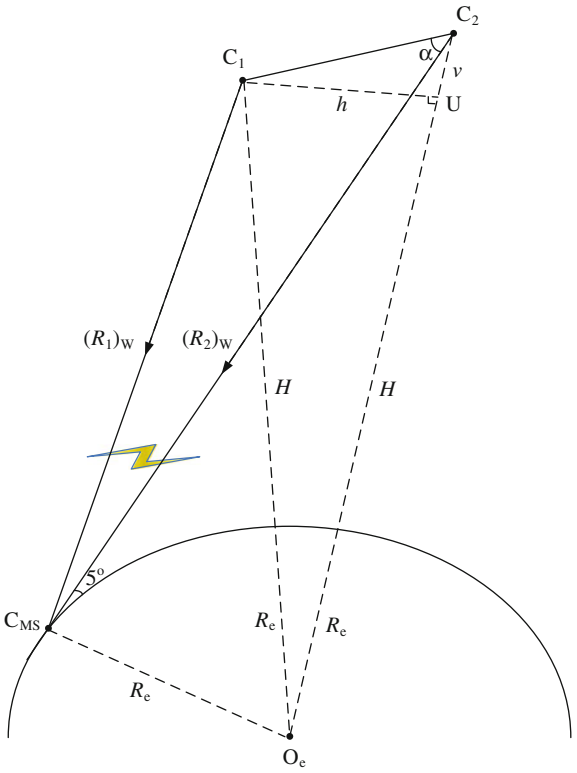


Fig. 20.1 Geometric relationship with maximum T_{GD} error (single-station monitoring)

20.4 Maximum Error Under Single-Station Monitoring

In the following spatial scene, insuperposition for satellite phase centers affects the T_{GD} error ultimately: phase center C_{MS} of the monitor antenna falls in plane $C_1C_2O_e$ (with O_e being the geometric center of the Earth), the elevation of the satellite phase center C_2 at frequency f_2 is the minimum elevation in view for the ground station (set as 5°), the elevation of the satellite phase center C_1 at frequency f_1 for the monitor station is not less than that of C_2, and the height from C_1 to ground is not larger than that for C_2, as shown in Fig. 20.1. $(R_1)_W$ and $(R_2)_W$ are defined respectively by the distance from C_{MS} to C_1 and C_2. Dot U is the projection of dot C_1 on line C_2O_e. By defining H as the height from C_1 to ground, the height from U to ground is approximately H. $|C_1U|$ is the horizontal distance between dot C_1 and C_2, namely the horizontally separative distance between phase centers, symbolized as h; $|C_2U|$ is the vertical distance between dot C_1 and C_2, namely the vertically separative distance between phase centers, symbolized as v.

According to geometric relationship, we have

$$\begin{cases} \alpha = \arctan\left(\dfrac{h}{v}\right) - \arcsin\left(\dfrac{R_e}{R_e + H + v}\sin 95^\circ\right) \\ (R_2)_W = R_e \cos 95^\circ + \sqrt{R_e^2 \cos^2 95^\circ + H^2 + 2R_e H} \\ (R_1)_W = \sqrt{(h^2 + v^2) + R_2^2 - 2\sqrt{h^2 + v^2} R_2 \cos\alpha} \end{cases} \tag{20.21}$$

where R_e is the radius of the Earth. Then it can be deduced that

$$c \cdot (e_{T_{GD}})_{max} = \left|\frac{(R_2)_W - (R_1)_W}{1 - \gamma}\right|. \tag{20.22}$$

In the RAD document supplied by COMPASS (Beidou Global Satellite Navigation System) to GPS and Galileo, the navigation signals are designed with 3 carriers [6], including B1 band centered at 1575.42 MHz, B2 band centered at 1191.795 MHz with B2a signal centered at 1176.45 MHz and B2b signal centered at 1207.14 MHz, and B3 band centered at 1268.52 MHz. Next it is analyzed with reference to B3 how the insuperposition for phase centers affects the T_{GD} for B1 and B3 and that for B2a and B3. In the simulation and analysis in this section and afterwards, MEO satellite is considered with orbit height 21528 km, and the radius of the Earth R_e is set at 6378 km.

The worst $(T_{GD})_{B1\text{-}B3}$ is shown in Fig. 20.2, where the curves from the top down correspond respectively to h from 5 to 0 cm with descending step 1 cm, and the horizontal axis shows the variation of the vertically separative distance v from 0 to 5 cm (not representing the actual COMPASS satellite phase centers). It can be seen that when the phase centers are close, T_{GD} error increases almost linearly with the increase of the vertically separative distance between phase centers.

The most significant effect of insuperposition of B2a and B3 phase centers on $(T_{GD})_{B2a\text{-}B3}$ is shown in Fig. 20.3, where the curves from the top down correspond to h from 1 to 0 cm with descending step 0.2 cm.

20.5 Maximum Error Under Double-Station Monitoring

If two or even more monitor stations are participated in estimation of T_{GD}, the maximum estimating error can be reduced significantly. This paper makes an analysis under double-station monitoring. Considering the TEC in the paths from the satellite to the two monitor stations (with the antenna phase centers at C_{MS1}

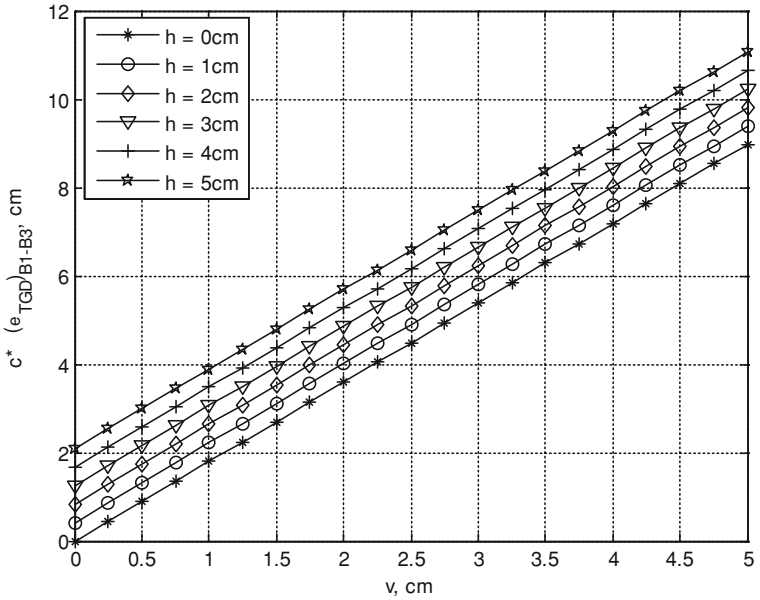


Fig. 20.2 The most significant effect of separation on $(T_{GD})_{B1\text{-}B3}$ (single-station monitoring)

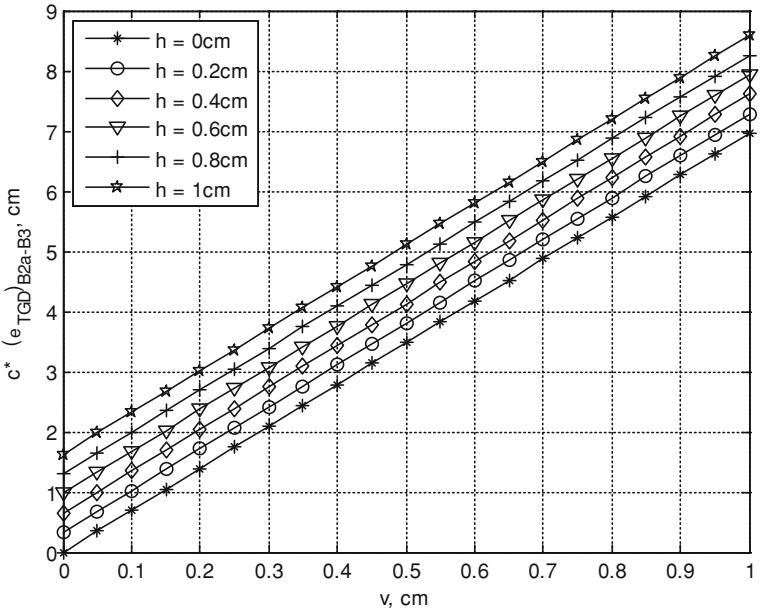


Fig. 20.3 The most significant effect of separation on $(T_{GD})_{B2a\text{-}B3}$ (single-station monitoring)

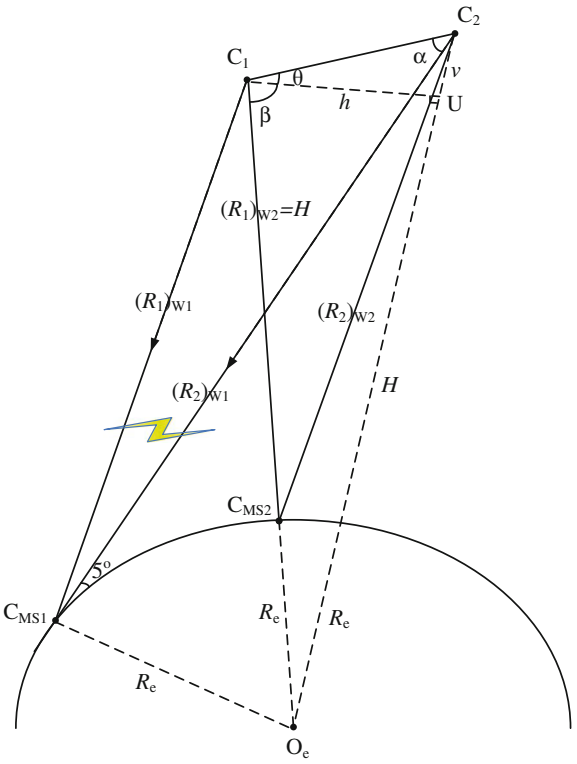


Fig. 20.4 Geometric relationship with maximum T_{GD} error (double-station monitoring)

and C_{MS2} respectively) are different, two T_{GD} results namely T_{GD1} and T_{GD2} from the two stations respectively can deduce T_{GD}

$$T_{GD} = \frac{T_{GD1} + T_{GD2}}{2}. \tag{20.23}$$

The corresponding error is

$$c \cdot e_{T_{GD}} = \left| \frac{\left((R_2)_1 - (R_1)_1\right) + \left((R_2)_2 - (R_1)_2\right)}{2(1-\gamma)} \right| \tag{20.24}$$

where $(R_1)_{W1}$ and $(R_2)_{W1}$ are the distances from C_{MS1} to C_1 and C_2 $(R_1)_{W2}$ and $(R_2)_{W2}$ are the distances from C_{MS2} to C_1 and C_2 respectively.

In order to reduce the estimation error better, C_{MS1} and C_{MS2} should be located at two sides of C_1O_e (or C_2O_e). The worst condition is considered herein, when C_{MS1} is located at C_{MS} described in the foregoing section, and C_{MS2} is located at the intersection of C_1O_e and the Earth surface, as shown in Fig. 20.4.

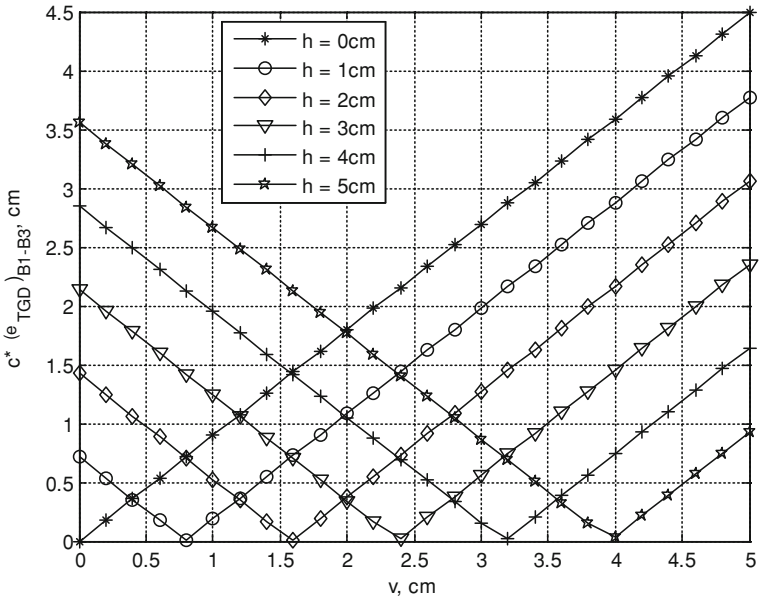


Fig. 20.5 The most significant effect of separation on $(T_{GD})_{B1\text{-}B3}$ (double-station monitoring)

From the relationship in Fig. 20.4, the formulae can be deduced as below:

$$\begin{cases} \alpha = \arctan\left(\dfrac{h}{v}\right) - \arcsin\left(\dfrac{R_e}{R_e + H + v}\sin 95^\circ\right) \\ (R_2)_{W1} = R_e \cos 95^\circ + \sqrt{R_e^2 \cos^2 95^\circ + H^2 + 2R_e H} \\ (R_1)_{W1} = \sqrt{(h^2 + v^2) + R_2^2 - 2\sqrt{h^2 + v^2} R_2 \cos\alpha} \\ \theta = \arctan\left(\dfrac{v}{h}\right) \\ \beta = \arcsin\left(\dfrac{h}{H + e}\right) \\ (R_1)_{W2} = H \\ (R_2)_{W2} = \sqrt{(h^2 + v^2) + H^2 - 2\sqrt{h^2 + v^2} H \cos(\theta + \beta)} \end{cases} \tag{20.25}$$

.

Substitute the formula above into formula (20.24), obtaining

$$(e_{T_{GD}})_{max} = \left| \frac{\left((R_2)_{W1} - (R_1)_{W1}\right) + \left((R_2)_{W2} - (R_1)_{W2}\right)}{2(1 - \gamma)} \right|. \tag{20.26}$$

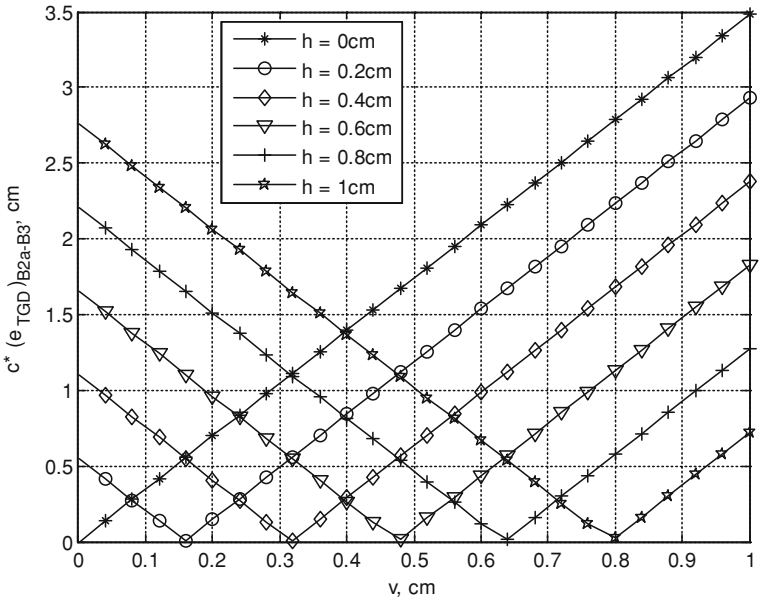


Fig. 20.6 The most significant effect of separation on $(T_{GD})_{B2a\text{-}B3}$ (double-station monitoring)

The worst $(T_{GD})_{B1\text{-}B3}$ under double-station monitoring is shown in Fig. 20.5. Comparing with the single-station monitoring, the maximum error for $(T_{GD})_{B1\text{-}B3}$ reduces largely. For example, when the horizontally separative distance is 5 cm and the vertically separative distance is 5 cm, the maximum error reduces from 11.1 cm with one station to 0.9 cm with two stations.

The worst $(T_{GD})_{B2a\text{-}B3}$ under double-station monitoring is shown in Fig. 20.6, which has a similar variation rule to that for $(T_{GD})_{B1\text{-}B3}$.

20.6 Conclusions

The geometric separation among the navigation satellite antenna phase centers at different frequencies affects the accurate measurement of T_{GD}, and further deteriorates the ranging error for single-frequency users. This paper analyzes the maximal effect of separation of satellite phase centers on T_{GD} error under single- and double-station monitoring. By the simulation of the navigation signals planned by COMPASS, it is found that double-station monitoring can greatly reduce the T_{GD} error caused by insuperposition of the phase centers comparing with single-station monitoring, so two or even more monitor stations should be managed to participate in the measurement, in order to ensure the ranging accuracy for single-frequency users.

References

1. Lanyi G (1986) Total ionospheric electron content calibration using SERIES GPS satellite data. TDA progress report 42–85, Jet Propulsion Laboratory, Pasadena, California, USA, January–March 1986
2. Dahlke SR, Coco DS, Coker C (1988) Effect of GPS system biases on differential group delay measurements. In: Proceedings of institute of navigation satellite division's international technical meeting, Colorado Springs, Colombia, USA
3. GPS Joint Program Office (2006) Navstar global positioning system IS-GPS-200 revision D. California, USA, March 2006
4. Coco DS, Coker C, Dahlke SR et al (1991) Variability of GPS satellite differential group delay biases. IEEE Trans Aerosp Electron Syst 27(6):931–938
5. Dai W, Jiao W, Jia X (2009) Application research for compass navigation satellite interfrequency bias correction terms. J Zhengzhou Inst Surv Map 26(5):367–374
6. Tan S, Zhou B, Guo S, Liu Z (2011) Research on COMPASS navigation signals of China. Chin Space Sci Technol V31(4):9–14

Chapter 21
Application of Independently Estimated DCB and Ionospheric TEC in Single-Frequency PPP

Wenfeng Nie, Wusheng Hu, Shuguo Pan, Shengli Wang, Xuhui Jin and Bin Wang

Abstract Precise Point Positioning technology is one of the most interesting research topics within the GNSS positioning and navigation community in the last decade for a number of reasons: simplified operation, cost-effectiveness and no base stations required. With the availability of dual-frequency receivers and precise GPS products, PPP has demonstrated capable of providing centimeter to decimeter level point positioning accuracy globally. As for single-frequency PPP, the accuracy deceases, particularly in the height component. One dominant factor for this degradation is the effect of unmodeled ionospheric delay. This paper investigates PPP technique using single-frequency data as single-frequency receivers are the most widely used tools for tracking, navigation and geo-referencing. Since the unmodeled ionospheric error is the biggest error source for a single-frequency receiver based on PPP system, an ionospheric error mitigation method named POLY model based on GPS regional reference station network is introduced for the reason that it can provide independent DCB and ionospheric TEC. In addition, three other different methods, as the Klobuchar model, the GIM model and the GRAPHIC algorithm, are assessed and their accuracy compared. Numerical results show that the POLY method offers better performance than GIM method, which is able to provide approximately 0.5 m point accuracy for static positioning over 24 h observation period.

Keywords DCB · TEC · Ionospheric error · Single-frequency PPP

W. Nie (✉) · W. Hu · X. Jin · B. Wang
School of Transportation of Southeast University, Nanjing, China
e-mail: wenfengnie@126.com

S. Pan · S. Wang
Instrument Science and Engineering Institute, Nanjing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_21,

21.1 Introduction

The availability of precise satellite orbit and clock products has pushed the precise point positioning (PPP) technology to a new era, whereby users can obtain high level point positioning accuracy on a global scale. Without the necessity for base stations, PPP uses un-differenced dual-frequency pseudorange and carrier phase observations from a single receiver coupled with precise satellite orbit and clock corrections. Researches have demonstrated that centimeter level for static applications and decimeter level for kinematic applications are capable [1, 2]. Taking the advantage of simplified operation, cost-effectiveness, and improved positioning accuracy, PPP technique will receive durative interests within the Global Navigation Satellite System (GNSS) community.

Since it is estimated that 75 % of all Global Positioning System (GPS) receivers used in the global are single-frequency receivers [3], single frequency PPP will be of great interest to a broad range of applications, such as deformation monitoring, precise satellite orbit determination, land resources survey and so on. As a dispersive medium to GPS signals, the ionosphere is generally recognized as the biggest error source in single frequency PPP after the use of satellite orbit and clock corrections [4]. Currently, only several meters level point positioning accuracy can be obtained because the ionospheric delay cannot be mitigated effectively using single-frequency measurements. In this regard, how to mitigate the effect of ionosphere is the key to achieve high level point positioning accuracy for single-frequency receiver users [4–7].

There exist a number of different ionospheric error mitigation methods for single-frequency GPS users. Klobuchar model with the eight ionospheric coefficients broadcast as part of the navigation message is perhaps the simplest and most widely used method. Though the method can be used in real-time, it can only mitigate approximately 50 % Root Mean Square (RMS) of the ionospheric range delay [8]. By the processing of the method, a couple of meters point positioning accuracy can be obtained [7]. Alternatively, single frequency users can utilize the ionosphere products, such as the Global Ionosphere Maps (GIM), provided by the International GNSS Service (IGS) or other organizations to correct the ionospheric delay [9]. Reference [7] show that GIM method can provide meters point positioning accuracy when the code measurements are used only, which performs better than Klobuchar method. However, the low spatial and temporal resolution and significant latency impose restrictions on the use of the GIM method. Another mitigation method known as GRoup And PHase Ionospheric Correction (GRAPHIC) algorithm is based on the single-frequency ionosphere-free linear combination of pseudorange and carrier phase measurements at the same frequency [10, 11]. As reported in [5], both GRAPHIC and GIM method are capable of providing sub-meter to meter level point positioning accuracy. However, the limitation is that the noise level of the algorithm is amplified by the code measurement noise.

In order to overcome the drawbacks of the methods mentioned above, an ionospheric error mitigation method named POLYnomial (POLY) model based on

GPS regional reference station network, also known as "Continually Operating Reference Stations (CORS)", is introduced in this paper. The advantage of the method is obvious. First, Differential Code Biases (DCBs) of the satellite and receiver, which act as the key factor on ionospheric delay determination, can be estimated using the whole day observation datasets. Second, with the access of the datasets from CORS system, the POLY model can be achieved to offer high spatial and temporal resolution and provide even positioning accuracy. But in the case of using GIM method, due to the reason that the IGS tracking network used to create the ionospheric products is unevenly distributed [12], the point positioning accuracy performs good in dense areas while bad in sparse areas, let alone the latency of 11 days for the final ionospheric products.

This paper investigates the ionospheric error mitigation methods related to the single frequency PPP. Firstly, the single frequency PPP observation equation is presented. And the four mitigation methods mentioned above are described in detail. Secondly, due to the fact that the broadcast and precise satellite clock corrections are consistent with the ionosphere-free combined satellite DCBs [13], a description of applying satellite DCBs to single frequency PPP model is given. Thirdly, at first, satellite DCBs are estimated by POLY model and compared with the DCBs provided by the Center for Orbit Determination in Europe (CODE). Then, absolute point positioning at seven JiangSu CORS (JSCORS, in Jiangsu province, China) sites with accurately known WGS-84 coordinates is carried out using the above indicated mitigation methods. Numerical results show that the POLY method offers better performance than GIM method, which is able to provide approximately 0.5 m point positioning accuracy for static application. Moreover, the GRAPHIC method performs the best unexpectedly though its noise is amplified by the code measurement noise. Corresponding explanations are shown in the analysis. Based on the performance of the four different methods in Sect. 21.4, conclusions are drawn in Sect. 21.5.

21.2 Ionosphere Mitigation for Single-Frequency PPP

The un-differenced code and carrier phase observation equations for single frequency PPP can be written and expressed as follows [14]:

$$\begin{aligned} \phi_i &= \rho + c(dt_r - dt^s) + T - \frac{I_0}{f_i^2} + \lambda_i N_i + hd_{r,i} - hd_i^s + pb_{r,i} - pb_i^s + \varepsilon_{\phi i} \\ P_i &= \rho + c(dt_r - dt^s) + T + \frac{I_0}{f_i^2} + HD_{r,i} - HD_i^s + \varepsilon_{Pi} \end{aligned} \tag{21.1}$$

where, the subscript $i(i = 1, 2)$ indicates band L_1 or L_2, r, s is respectively index for receiver and satellite, P_i, ϕ_i is respectively the measured pseudorange and carrier phase on L_i (m); ρ is the true geometric range (m); c is the vacuum speed of

light (m/s); dt^s is the satellite clock error (s); dt_r is the receiver clock error (s); T is the tropospheric delay (m); I_0 is related to the ionosphere Total Electron Content (TEC), when consider the first order of ionospheric delay only, $I_0 = 40.28\ TEC$ (TECU,TECU $= 10^{16}$ el/m^2); f_i is the frequency on L_i (Hz); λ_i is the wavelength on L_i (m/cycle); N_i is the phase ambiguity on L_i (cycle); $hd_{r,i}\left(HD_{r,i}\right), hd_i^s\left(HD_i^s\right)$ is respectively hardware delay of receiver and satellite on carrier phase (pseudorange) on L_i (m); $pb_{r,i}, pb_i^s$ is respectively the initial phase bias of receiver and satellite on L_i (m); $\varepsilon(\cdot)$ is the noise including residual multipath (m). It is noted that the orbit error, relativistic effects and phase windup are not presented in (1) because they can be removed by the use of precise IGS products or models. The tropospheric delay can be corrected at decimeter even centimeter level using existing models and meteorological measurements. As for the satellite clock error dt^s, it can also be corrected by the broadcast navigation message or precise satellite clock products from IGS [4]. But at the same time, the ionosphere-free linear combination of satellite DCBs should be applied and calibrated because both the broadcast and precise satellite clock corrections are consistent with the linear combined satellite DCBs.

The satellite DCBs can be found from the broadcast group delay message, as well as from the GIM in IONosphere map Exchange (IONEX) format file. In addition, the satellite DCBs can be independently estimated by regional POLY model [15]. The three different DCBs will be applied in this research as shown in the following and more details about them will be presented in Sect. 21.3.

As a result, the ionospheric effect becomes the dominant error source in the undifferenced code and phase measurements and they should be mitigated as precise as possible. Thus, strategies using the four different methods mentioned above can be applied:

- Using broadcast Klobuchar model with broadcast DCBs;
- Using GIM method with DCBs provided by CODE;
- Using GRAPHIC algorithm with DCBs provided by CODE;
- Using POLY model with independently estimated DCBs.

As the first method, Klobuchar model is perhaps the simplest and most widely used method for single-frequency receiver users. The purpose of the method is to yield an RMS correction of at least 50 % with a minimum requirement of coefficients and user computational time [8]. Therefore, only eight coefficients are sent as part of the satellite navigation message. In order to improve the accuracy of the model, CODE has been providing post-fit Klobuchar ionospheric coefficients that fit well with the GIM data estimated by CODE since mid-July, 2000. Results show that the post-fit Klobuchar style can achieve about 2 m point positioning accuracy when using the code measurements only, which is more consistent than the broadcast Klobuchar model [7]. As well, CODE has been estimating predicted Klobuchar-style coefficients. As demonstrated in [7], the improvement of predicted style is not as significant as the post-fit style. In this paper, the broadcast Klobuchar model is used for its typicality.

The second method for single-frequency GPS users to correct for the ionospheric error is by using the GIM provided by IGS or other organizations [9]. Since 1998, IGS has provided IONEX format files, which contain a set of satellite and receiver DCBs, as well as 13 global ionosphere maps [16]. In each map, the vertical TEC (VTEC) is given for grid points both in space and time. In detail, the temporal resolution is 2 h while the spatial grid interval is 2.5° north-south and 5° east-west. So far, the accuracy level provided by GIM is from 2 to 8 TECU, in which 1 TECU corresponds to 0.163 m on L1 frequency [7]. Note that there exists a latency of 11 days for the final GIM products. Reference [5] shows that submeter positioning accuracy is obtainable for the GIM methods using both code and carrier phase measurements in single-frequency PPP, which performs better than the broadcast Klobuchar model.

The GRAPHIC method, which is based on the single frequency ionosphere-free linear combination of the code and carrier phase measurements on the same frequency, takes advantage of the fact that the ionosphere delays the code and advances the carrier phase at the same magnitude but opposite in sign [10, 11]. Upon adding and averaging the code and carrier phase range, a combined measurement is obtained

$$\begin{aligned}\frac{P_i+\phi_i}{2} &= \rho + c(dt_r - dt^s) + T + \frac{\lambda_i N_i}{2} + \frac{\varepsilon_{Pi}+\varepsilon_{\phi i}}{2} \\ &\quad + \frac{\left(hd_{r,i} - hd_i^s\right) + \left(HD_{r,i} - HD_i^s\right)}{2} + \frac{pb_{r,i} - pb_i^s}{2}\end{aligned} \tag{21.2}$$

that no longer depends on the ionospheric delay and exhibits a noise level with standard deviation:

$$\sigma_{combine} = \frac{1}{2}\sqrt{\sigma^2(\varepsilon_{pi}) + \sigma^2(\varepsilon_{\varphi i})} \approx \frac{1}{2}\sigma(\varepsilon_{pi}). \tag{21.3}$$

Due to the reason that the bandwidth limitations keep the carrier phase observations typically 100 times more precise than the code, the noise affecting the GRAPHIC method is mainly dominated by the code measurement error, as shown in (21.3). Compared with the carrier phase measurement, the noise level of the combined measurement is amplified. Therefore, a long observation time is required for the float ambiguity to converge. As shown in [5], a few centimeters accuracy can be attained after 12–24 h observation period, which is comparable with the GIM method. But for short observation time, the GIM method performs better [5]. For the uniqueness of GRAPHIC method that the first-order ionospheric error is eliminated, the performance of the method is assessed and compared in this investigation.

The forth method is the POLY model, which is a representation of the regional fitting method for VTEC model based on the dual-frequency observations from CORS system [17]. Other similar methods include the trigonometric series model [18] and the low-degree spherical harmonics function model [9]. In general, the

carrier-phased code measurements are used because they are unambiguous [6]. Then, DCBs of the satellite and receiver are treated as constant values and estimated along with the coefficients of the fitting model [15]. As a result, regional ionosphere maps provided by POLY model, which are similar to the global ionosphere maps but in a small scale, become available. Due to the reason that datasets are more sufficient and stations are evenly distributed, the ionospheric products provided by the POLY model have higher spatial and temporal resolution. The effect of the POLY model for single-frequency PPP will be presented and compared in Sect. 21.4.

21.3 Differential Code Biases (DCBs)

Within the satellite as well as the receiver hardware, the codes transmission at two L-band frequencies (L1 and L2) are carefully synchronized so that they can broadcast simultaneously. However, absolute simultaneity is not possible, so the time difference between the transmitted times at the two frequencies is called DCB of the satellite and receiver [19]. As reported in [6], the DCBs are the key factor which cause the severest effects on ionospheric delay determination and must be reduced as much as possible. If ignored, the TEC measurement may result in an error of about ±3 and ±10 ns (1 ns is correspondent to 2.852 TECU) respectively [20].

Satellite DCBs can be found from the GIM in IONEX format, as well as from the broadcast navigation message. The broadcast satellite T_{gd}, also known as "SV group delay differential" are determined by the satellite manufacturer in a laboratory calibration before the satellites are launched [7]. Since April, 1999, the T_{gd} values are updated in order to fit well with the DCB estimated by some IGS organizations, like Jet Propulsion Laboratory (JPL). In addition, satellite DCBs can also be estimated using datasets from a regional GPS reference station network [21]. Researches have shown there is little difference between the released DCBs from GIM and the estimated DCBs by POLY model [22].

It is well-known that the offset of the broadcast and precise satellite clock corrections are always referred to the ionosphere-free linear combination of the L1 and L2 frequency. The ionosphere-free linear combination observation equations for PPP model can be expressed as follows:

$$\begin{aligned} \phi_3 &= \rho + c(dt_r + dt^s) + T + \lambda_1 N_3 + hd_{r,3} - hd_3^s + pb_{r,3} - pb_3^s + \varepsilon_{\phi 3} \\ P_3 &= \rho + c(dt_r + dt^s) + T + HD_{r,3} - HD_3^s + \varepsilon_{P3} \end{aligned} \tag{21.4}$$

where

$$\begin{aligned} \phi_3 &= \tfrac{f_1^2}{f_1^2 - f_2^2}\phi_1 \tfrac{f_2^2}{f_1^2 - f_2^2}\phi_2 & P_3 &= \tfrac{f_1^2}{f_1^2 - f_2^2}P_1 \tfrac{f_2^2}{f_1^2 - f_2^2}P_2 \\ HD_{r,3} &= \tfrac{f_1^2}{f_1^2 - f_2^2}HD_{r,1} - \tfrac{f_2^2}{f_1^2 - f_2^2}HD_{r,2} & HD_3^s &= \tfrac{f_1^2}{f_1^2 - f_2^2}HD_1^s - \tfrac{f_2^2}{f_1^2 - f_2^2}HD_2^s \end{aligned}. \tag{21.5}$$

Assume that $d_t^{s'}$ is the satellite clock correction, while d_t^s is the true satellite clock correction. There exists a relationship between $d_t^{s'}$ and d_t^s, as follows:

$$c \cdot dt^{s'} = c \cdot dt^s + HD_3^s. \tag{21.6}$$

Note that the combined receiver pseudorange DCB ($HD_{r,3}$) is absorbed in the receiver clock correction. And the errors that remain constant in the carrier phase equation are absorbed in the combined ambiguity, as show in the following equation. Thus, the combined ambiguity N_3' becomes float number.

$$\lambda_1 N_3' = \lambda_1 N_3 + hd_{r,3} - hd_3^s + pb_{r,3} - pb_3^s - HD_{r,3} - HD_3^s \tag{21.7}$$

As for the single frequency PPP, the satellite clock corrections which are consistent with combined satellite DCBs are introduced into the un-differenced code and carrier phase observation equation. Hence, the combined satellite DCBs should be calibrated. Combined with (21.1) and (21.6), the single frequency PPP model (for L1 observation) can be written as:

$$\begin{aligned} P_1 =& \rho + c\left(dt_r + dt^{s'}\right) + T + \frac{I_0}{f_1^2} + \frac{f_2^2}{f_1^2 - f_2^2}\left(HD_1^s - HD_2^s\right) + \varepsilon_{P1} \\ \phi_1 =& \rho + c\left(dt_r + dt^{s'}\right) + T - \frac{I_0}{f_1^2} + \lambda_1 N_1' + \varepsilon_{\phi 1} \end{aligned} \tag{21.8}$$

where, N_1' is the combined ambiguity which is also a float number.

It is noted that the satellite DCBs, as HD_1^s, HD_2^s, are referred to the P1 and P2 pseudorange observations. For some civilian GPS receivers that do not output P1 code but C/A code instead, the P1-C1 DCBs should be used to correct data from these receivers in order to achieve full consistency with the P1-P2 data and the satellite clock corrections [23]. Alternatively, the "cc2noncc.f" conversion utility can be used to transform RINEX observation files that have C/A code measurements to be compatible with P1 and P2 observables [7]. The satellite P1-C1 biases are approximately of the order of 0–2 ns and can be obtained from the CODE website [7, 23].

21.4 Numerical Results and Analysis

21.4.1 DCB Comparison

One of the key technologies in the process of single-frequency PPP using POLY method is the determination of satellite DCBs. The method described in [15, 22] is adopted in this investigation. Datasets are collected from BBPX station, one of the JSCORS network stations. The observation time is from 1st July to 10th July, a

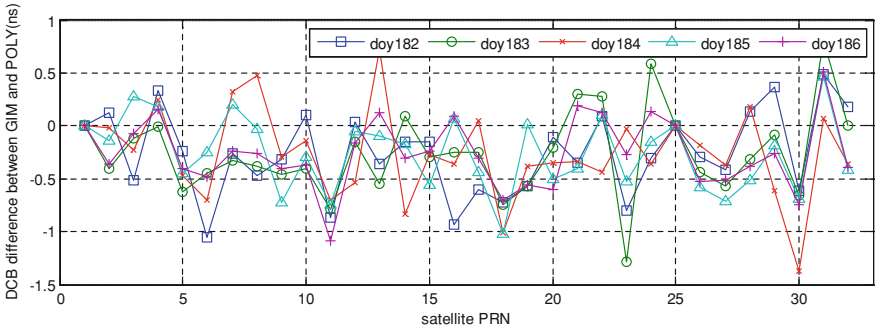


Fig. 21.1 The DCB differences between GIM and POLY

total of 10 days in 2010, with sampling interval of 30 s and cut-off angle of 15°. The DCB differences between GIM released and POLY determined are shown in Fig. 21.1, in which shows results of the former 5 days.

As shown in Fig. 21.1, most of the differences are between –1 and 0.5 ns except for the difference of PRN30 in doy184. Two reasons that may cause this: first, the data used in GIM are collected on a global scale, hence redundant observation is more sufficient, second, according to the restriction used to separate the DCB of satellite and receiver, the sum of satellite DCBs are restricted to be zero [9, 15, 22]. But, in this experiment, PRN01 and PRN25 are not observed during the ten days. Thus, the two DCBs of PRN01 and PRN25 are allotted to the other satellites, which can lead to some differences compared with GIM. Furthermore, the 10 days' results show that the mean standard deviation of the satellite DCBs is 0.3 ns, which indicates that satellite DCBs are quite stable in several days. In fact, though satellite DCBs vary from satellite to satellite, to some degrees, they can be treated as constant values respectively during some time.

In addition, the P1-P2 DCBs provided by GIM and determined by POLY, as well as the broadcast satellite group delay T_{gd} on 2nd, July 2010 are presented in Table 21.1. The Klobuchar method is abbreviated as "Klob".

From the table we can see that there exist quite large differences between DCB determined by GIM or POLY and the broadcast group delay. The biggest difference can reach to 27 ns as shown in the case of PRN23. According to the table, we can conclude that the updated group delays still do not fit well with the estimated ones. To a great extent, the group delay provided in the navigation file is not incorrect, or at least not accurate enough [7].

21.4.2 Single-Frequency PPP Results

In the experiment, datasets are collected on 19th, November 2010. And the regional POLY model is established using seven selected JSCORS GPS stations. They are BBPX, BBSN, BTGU, BTXU, BTXH, BTYX and BTRD stations.

Table 21.1 DCB comparison (ns)

PRN	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
POLY	5.3	−3.0	−1.8	0.9	−2.6	1.2	−3.3	−2.2	−4.1	1.7	2.0	1.4	−0.1	0.8	0.6
GIM	5.3	−2.9	−1.5	1.0	−2.7	0.7	−3.8	−3.0	−3.8	0.4	2.2	1.1	−0.5	0.7	0.9
Klob.	−17.2	−4.7	−6.1	−8.8	−5.1	−12.4	−3.7	−5.6	−2.8	−11.6	−11.6	−17.1	−8.8	−9.8	−9.8
PRN	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**26**	**27**	**28**	**29**	**30**	**31**	**32**
POLY	1.1	1.2	3.6	−0.8	1.9	5.8	7.3	−5.0	−1.7	−3.2	0.9	0.2	−0.4	2.9	−4.0
GIM	0.5	0.1	3.1	−0.9	1.5	5.6	6.6	−5.2	−1.8	−3.7	0.4	1.5	−1.9	3.4	−5.0
Klob.	−12.4	−17.1	−14.9	−7.9	−12.1	−17.7	−20.0	−1.4	−6.1	−4.2	−10.7	−8.8	−7.9	−13.0	−3.3

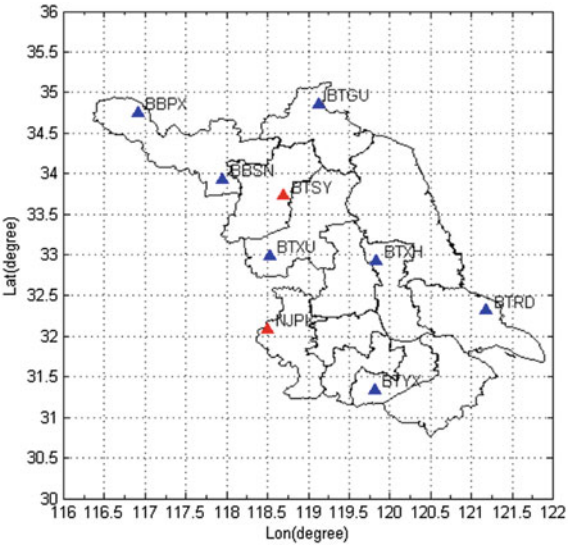


Fig. 21.2 The location of the seven JSCORS stations (the *blue* marked stations are used to establish POLY model while the *red* marked stations, as BTSY and NJPK, are used to check the accuracy of the model)

The location of these stations is shown in Fig. 21.2. The sampling interval is 30 s and cut-off angle is 15°. These stations are equipped with dual frequency geodetic quality LEICA receivers with LEIAX1202GG antenna. Ionosphere condition is very quite on that day with an Ap index of 2. For the purpose of the experiment, only L1 code and carrier phase measurements are used for single frequency PPP.

In order to check the accuracy of the POLY model, datasets from two stations over a period of 24 h observation are used. They are the stations of BTSY and NJPK. In detail, the results of BTSY stand for the accuracy inside the network while NJPK represent the accuracy outside the network. Hence, the results of the two stations are presented in Figs. 21.3 and 21.4.

Figures 21.3 and 21.4 show error values of the east, north and height components based on the four ionospheric error mitigation processing methods.

As we can see from the figures, the errors in the height component perform the worst compared with the east and north component. This phenomenon also exists in the results of dual-frequency PPP because the atmosphere delay errors along the propagation path of the GPS signals are more difficult to eliminate. As predicted, the Klobuchar model has the least accurate positioning results, particularly in the east and height component. The goal of the Klobuchar model may explain this result. In fact, the Klobuchar model is designed to yield an rms ionospheric correction of at least 50 % using a minimum of coefficients and user computational time [8]. It can be inferred that the ionosphere error corrections using the eight coefficients are not effective enough so only meters level accuracy can be obtained.

Besides, the POLY and GIM model can provide comparable positioning accuracy in horizontal component. Both east and north component of the POLY and GIM model can be convergent to 5 cm. But in the height component, POLY method performs better than the GIM. As shown in the figures, the height

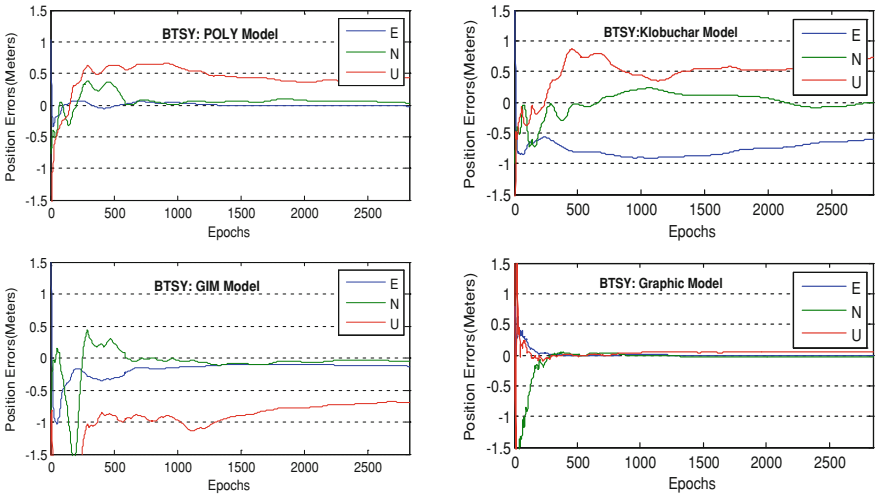


Fig. 21.3 Position errors with respect to the accurately known coordinates for BTSY station

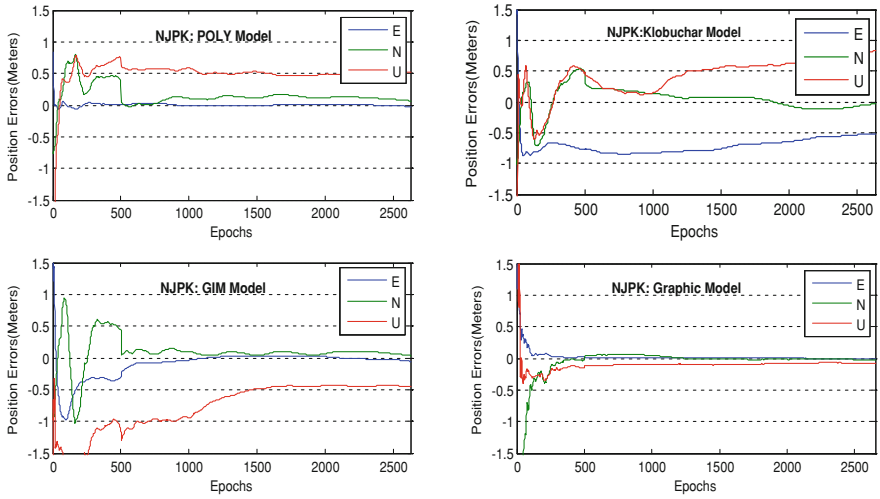


Fig. 21.4 Position errors with respect to the accurately known coordinates for NJPK station

component of POLY method can be convergent to about 0.5 m within 60 epochs while the GIM cannot match the performance even using the 24 h observation data. As we know, the final GIM has a temporal resolution of 2 h and a spatial resolution of 5° in longitude and 2.5° in latitude. At grid points, GIM has an accuracy level of about 2 TECU. In most cases, the ionospheric TEC grids are required to interpolate to the ionospheric pierce points for satellites observed at different azimuth and elevation angles so that the accuracy decreases [7]. What's

Table 21.2 The positioning bias for the four mitigation methods based on BTSY and NJPK

	Bias (m, BTSY)			Bias (m, NJPK)		
	E	N	U	E	N	U
Klob	0.786	0.045	0.577	0.713	0.073	0.507
GIM	0.140	0.031	0.856	0.036	0.108	0.655
GRHC	0.002	0.013	0.040	0.006	0.004	0.093
POLY	0.001	0.074	0.477	0.003	0.125	0.526

Table 21.3 The positioning accuracy for the four mitigation methods based on BTSY and NJPK

	RMS (m, BTSY)			RMS (m, NJPK)		
	E	N	U	E	N	U
Klob	0.792	0.118	0.589	0.721	0.166	0.548
GIM	0.153	0.085	0.865	0.097	0.154	0.703
GRHC	0.099	0.246	0.128	0.023	0.138	0.240
POLY	0.023	0.100	0.487	0.011	0.156	0.529

more, the IGS stations used for the ionosphere products is not evenly distributed [17]. As shown in the IONEX file, only five stations, as BJFS, WUHN, SHAO, LHAZ, TWTF are used in China. Therefore, due to the low spatial and temporal resolution, the GIM method is less efficient to mitigate the ionospheric delay while the POLY method can achieve 0.5 m level positioning accuracy.

In addition, the method GRAPHIC performs the best unexpectedly among the four methods, no matter for the horizontal or vertical components. In more detail, the horizontal components can be convergent to 1 cm and the height component is convergent to 1 dm for the GRAPHIC method. As for the time required for the solutions to convergent, the POLY model performs better than GRAPHIC method in the beginning 120 epoches. This is because the noise level of the GRAPHIC method is enlarged by the code measurements, and this has impact on the initial portion of the solutions. But after the ambiguities are stabilized, the GRAPHIC method follows the more precise carrier phase measurement.

Tables 21.2 and 21.3 show the positioning bias and accuracy of the four mitigation methods for station BTSY and NJPK over 24 h observation period. The Klobuchar method is abbreviated as "Klob" and the GRAPHIC method is abbreviated in the table as "GRHC". Note that the positioning results are counted after the initial time of 3 h in the table.

It can be clearly seen from the tables that the GRAPHIC method has the highest accuracy, and the POLY method takes the second place while as predicted, the Klobuchar method gives the least accurate positioning solution. In detail, there exist little difference in the north component for the four methods. Similarly, except for the least accuracy method Klobuchar method, little difference exists in the east component for the other three methods. Klobuchar method has the least accuracy in the east component. The biggest difference exists in the height

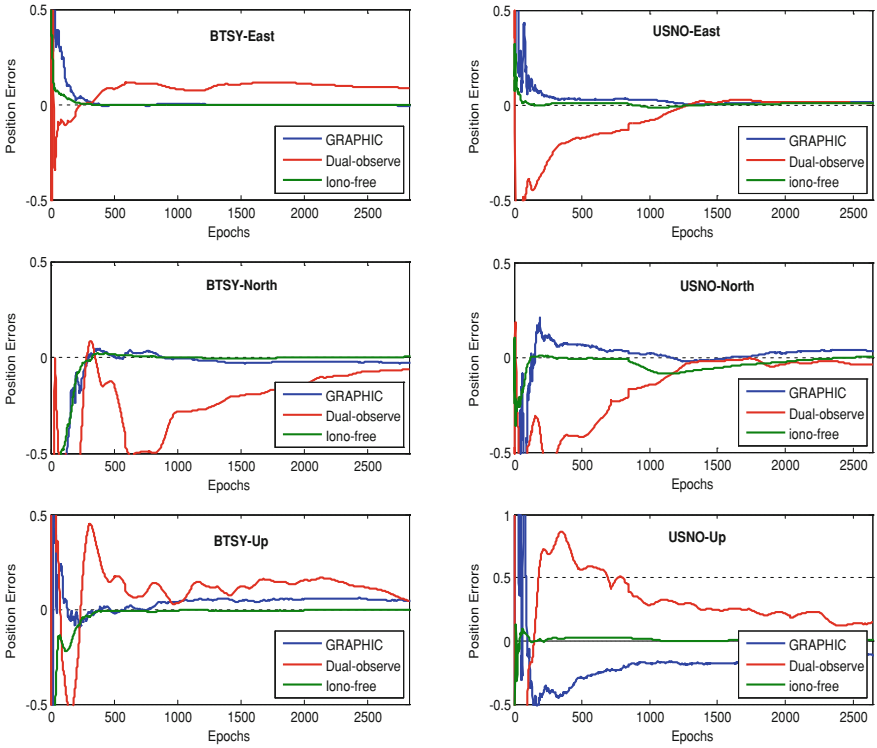


Fig. 21.5 Position errors in each component for BTSY and USNO station

component. It is obvious that the accuracy of GRAPHIC method is the highest among the four methods. Hence, to a certain extent, the accuracy of the height component determines the point accuracy of the mitigation methods.

The noise level of the GRAPHIC algorithm determines its ambiguity solution requiring a long period time to converge. For the reason that the POLY method can be implemented for real-time applications, efforts are made to find out the reasons that POLY method performs lower accuracy than GRAPHIC. In this regard, two more strategies for PPP test are assessed and compared. One is the direct apply of the ionospheric TEC derived from the carrier-smoothed dual-frequency code measurements without modelling, as the "Dual-observe" method. The other is the classic ionosphere-free combination PPP using dual-frequency observations, as the "iono-free" method. To give an overall explanation, datasets from USNO, one of the IGS tracking stations, are added. Meanwhile, the satellite DCBs are used from the GIM IONEX format file in the case of USNO solution in order to avoid the influence of the satellite DCBs. Therefore, the results of the station BTSY from JSCORS and USNO from IGS are shown in Figs. 21.5. Note that in each component, the three different strategies are compared.

As given in Fig. 21.5, the "iono-free" method performs the best on matter in the aspect of accuracy or time to converge. The performance of GRAPHIC algorithm can match "iono-free" method in the horizontal component though GRAPHIC requires a little longer time to converge. In the height component, the GRAPHIC algorithm has lower accuracy than "iono-free" method. The performance is reasonable in view of that the GRAPHIC algorithm only uses single-frequency data.

In the case of the "Dual-observe" method, it has the lowest accuracy and longest time to converge in each component among the three methods. As reported in [7], the TEC derived from the dual-frequency GPS observations is regarded as "true values". And in most cases, the phase-smoothed code observations are used to extract ionosphere delay [6], as do in this investigation. In theory, if the extracted TEC information is accurate enough, the performance of "Dual-observe" method will be better than "GRAPHIC" method, even the "iono-free" method. Because on one hand, the noise of the latter two methods are amplified by a combination of code and phase measurements. On the other hand, the single-frequency and dual-frequency linear combination only consider the first order of the ionosphere range delay. But in practice, the extracted TEC is derived from the phase-smoothed code measurements whose noise is still large. Therefore, the TEC correction is not accurate enough. In addition, the POLY model is established according to the derived TEC from seven stations, as shown in Fig. 21.2. Hence, the ionospheric delay corrections provided by POLY contain model errors, which weaken the effect of the ionospheric mitigation. In view of these, the POLY method performs less accurate than the GRAPHIC method.

21.5 Conclusions

The performances of the four different method show that the GRAPHIC method has the highest accuracy, which can provide decimeter level static point positioning accuracy during a 24 h observation period. And the POLY method takes the second place. An accuracy of approximately 0.5 m is obtainable by the POLY method, which is better than the GIM method. The GIM method however, can provide sub-meter level static positioning accuracy. Among the four methods, the widely used Klobuchar model is the lowest accuracy solution which can provide about one meter point positioning accuracy.

Compared with the other three mitigation methods, the POLY method can not only provides independently estimated and rather accurate satellite DCBs, but also the reliable ionospheric TEC information. But due to the reason that the TEC information extracted using the carrier-smoothed code measurements is not accurate enough, together with the influence of model errors, the accuracy of the POLY method performs lower than the GRAPHIC method. To improve the accuracy of the POLY method and broaden its applications to real-time process, in the future, efforts will be made from the two aspects, one is the improvement of the

noise level of the measurements for extracting TEC information and the other is the decrease of the model errors. Besides, the solar activity has large impacts on TEC performances. Consequently, more datasets on different ionospheric conditions are needed to analyze.

Acknowledgments This work was supported by Key Projects in the National Science & Technology Pillar Program during the Twelfth Five-year Plan Period.(2012BAJ23B01). The authors are very grateful to the anonymous reviewers for their constructive comments and suggestions.

References

1. Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102(B3):5005–5017
2. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solutions 5(2):12–28
3. Wyllie S, Zhang K, Talbot N (2006) An analysis of the temporal correlation of the ionospheric bias affecting GPS carrier phase observations. In: Proceedings of IGNSS symposium. Holiday Inn Surfers Paradise, Australia, pp 17–21 July (79)
4. Chen KZ, Gao Y (2005) Real-time precise point positioning using single frequency data. In: Proceedings of ION GNSS-2005, pp 1514–1523
5. Choy S, Zhang K, Silcock D (2008) An evaluation of various ionospheric error mitigation methods used in single frequency PPP. J Global Position Syst 7(1):62–71
6. Yuan YB, Huo XL, Ou JK (2007) Models and methods for precise determination of ionospheric delay using GPS. Prog Nat Sci 17(2):187–196
7. Øvstedal O (2002) Absolute positioning with single-frequency GPS receivers. GPS Solutions 5(4):33–44
8. Klobuchar JA (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst 3:325–331
9. Schaer S (1999) Mapping and predicting the Earth's ionosphere using the Global Positioning System. Geod Geophys Arb Schweiz 59:59
10. Yunck TP (1993) Coping with the atmosphere and ionosphere in precise satellite and ground positioning. Geophys Monogr Ser 73:1–16
11. Montenbruck O (2003) Kinematic GPS positioning of LEO satellites using ionosphere-free single frequency measurements. Aerosp Sci Technol 7(5):396–405
12. Fedrizzi M, Eurico R, Paula D, Ivan JK, Richard BL, Marcelo CS, Komjathy A (2002) Innovation-mapping the low-latitude ionosphere with GPS. GPS World 13(2):41–47
13. Kouba J (2009) A guide to using International GNSS Service (IGS) products. Unpublished http://igscb.jpl.nasa.gov/igscb/resource/pubs/UsingIGSProductsVer21.pdf
14. Kleusberg A, Teunissen PJG (1996) GPS for Geodesy. In: Lecture notes in earth sciences. Springer, Berlin, p 60
15. Zhang HP, Shi C, Tang WM (2008) United solution to polynomial VTEC modeling and DCB analysis using ground-based GPS observations. Geomat Inf Sci Wuhan Univ (in Chinese) 33(8):805–809
16. Schaer S, Werner G, Joachim F (1998) IONEX: the ionosphere map exchange format version 1. In: Proceedings of the IGS AC workshop, Darmstadt, Germany, vol 9, no 11
17. Komjathy A (1997) Global ionospheric total electron content mapping using the Global Positioning System. PhD dissertation, University of New Brunswick

18. Yuan YB, Ou JK (2004) A generalized trigonometric series function model for determining ionospheric delay. Prog Nat Sci 14(11):1010–1014
19. Coco DS, Clayton C, Scott RD, James RC (1991) Variability of GPS satellite differential group delay biases. IEEE Trans Aerosp Electron Syst 27(6):931–938
20. Wilson B, Anthony M (1994) Extracting ionospheric measurements from GPS in the presence of anti-spoofing. In: Proceedings of the 7th international technical meeting of the satellite division of the institute of navigation, ION GPS-94, Alexandria, VA. vol. 2, pp. 1599–1608, 20–23 September
21. Gao Y, Liao XQ, Liu ZZ (2002) Ionosphere modelling using carrier smoothed ionosphere observations from a regional GPS network. Geomatica 56(2):97–106
22. Zhang HP, Han WH, Huang L, Geng CJ (2012) Modeling global ionospheric delay with IGS ground-based GNSS observations. Geomat Inf Sci Wuhan Univ 37(10):1186–1189
23. Jefferson DC, Michael BH, Ronald JM (2001) Examining the C1-P1 pseudorange bias. GPS Solutions 4(4):25–30

Chapter 22
A New Strategy on Precise Clock Combination of IGS Analysis Centers

Kang-kang Chen and Tian-he Xu

Abstract Precise clock correction is the important basic data for estimating troposphere parameters, precise point positioning (PPP) and satellite precise orbit determination. In this paper, a strict and reliable clock combination approach is proposed. The main idea is as follows: the compatibility corrections of satellite orbit, station coordinates and geocenter offset are applied to Analysis Center (AC) clock solutions firstly. It aims at maintaining the consistency of all IGS combined products. The individual AC clock solutions' offset and drift parameters with respect to the broadcast clock solutions are estimated using robust least square method. Then individual AC solutions are aligned to a common reference time-frame. The initial weights of clock combination are determined by posteriori mean square error of parameters estimate. Using robust estimation with high-breakpoint of pollution rate resists gross error. At last the final clock combination is obtained based on the weighted average method by iteratively calculating equivalent weights. The results show that combined clock solution obtained by the proposed method has commendable consistency with IGS final clock solution. The RMS of the discrepancy between combination clock and IGS final clock is about 0.03–0.05 ns, which is significantly better than the results of AC clock solutions. Combination clock correction is also more stable and reliable. Finally, static PPP test proves that the combination clock maintains the consistency of orbit and clock at millimeter level.

K. Chen (✉)
School of Geology Engineering and Surveying, Chang'an University, Xi'an, Shanxi, China
e-mail: chenkkuser@126.com

T. Xu
State Key Laboratory of Geo-information Engineering, Xi'an, Shanxi, China

T. Xu
State Key Laboratory of Astronautic and Dynamics, Xi'an, Shanxi, China

T. Xu
State Key Laboratory of Geodesy and Earth's Dynamics, Wuhan, Hubei, China

T. Xu
Xian Research Institute of Surveying and Mapping, Xi'an, Shanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_22,

Keywords Clock combination · Reference time-frame · Robust estimation · High-breakpoint of pollution rate

22.1 Introduction

Since 1993, IGS have begun to study and provide unified precise GNSS satellite orbit, clock correction, station coordinates etc. though combining Analysis Center's solution [1]. The main motivation of clock combination for IGS is the realization of the goals of the IGS/BIPM timing project. Besides, precise clock correction is the important basic data for troposphere parameter estimation, precision positioning and satellite precise orbit determination. There is an increasing need for station clocks and reliable high sampling rate satellite clocks which promotes the development of research on precise clock combination methods [2]. Currently, at least nine IGS analysis centers provide orbit/clock in different sampling intervals. Typically, orbits/clocks are combined with one or two days delay after the last submission of AC products [1]. The first step of the clock combination is the alignment of the individual AC solutions to the IGS reference frame. The AC clock solutions have to be corrected for maintaining the consistency between orbit and clock as well as the consistency between SINEX solution and clock. The second step is that the individual AC solutions are aligned to a common reference time-frame, which eliminates the effects of different reference clocks and speed up the convergence of the clock combination. The third step of the combination is to use the high-breakpoint of pollution rate robust for outlier detection and rejection iteratively, for reference clock jumps in the individual AC solution. It accounts for satellite, station and epoch biases for each of the individual AC clock solutions. The IGS orbit/clock combination and evaluation is performed weekly [1–4].

Since IGS product combination was originally developed, many improvements have been made by the Analysis Center Coordinator (ACC). However, the basic method of the combination, the L1-norm, was not changed all the time [3]. The development of modern data processing provides a lot of new thoughts and methods for IGS clock combination. In this paper, robust least squares as the method for calculating the transformation parameters of time reference frame. The robust iterative with high-breakpoint of pollution rate is proposed as the method of calculating equivalent weight. Then we can get combination clocks based on the weighted average of AC clock solutions. The results show that combined clock solution obtained by the proposed method has commendable consistency with IGS final clock solution. The RMS of the discrepancy between combination clock and IGS final clock is $0.03 \sim 0.05$ ns, which is significantly better than those of AC clock solutions. Combination clock is also more stable and reliable. Finally, static PPP test also proves that combination clock can maintain the consistency between orbit and clock at millimeter level.

22.2 Principles and Algorithm Description of Clock Combination

It's precise and rigorous data processing for the combination clock based on several analysis center solutions. First, the consistency of orbit and clock as well as SINEX and clock should be improved, which is very important for those users who want to perform precise point positioning [3]. Secondly, the individual AC solutions are aligned to a common reference time-frame which will eliminate the effects of different reference clocks and speed up the convergence of the clock combination [2]. The initial weights of clock combination are determined by posteriori mean square error of parameters estimate. Finally, robust estimation with high-breakpoint of pollution rate is used to iteratively calculate equivalent weight, detect and correct reference clock jumps in the individual AC solution.

22.2.1 Consistency Correction of Combination Products

The station/satellite combination clocks are made consistency with the current IGS orbit/ERP and the SINEX solution. Specifically, the following compatibility corrections are applied to AC clock solutions prior to the clock combination to maintain orbit/clock as well as SINEX/clock consistency:

$$\begin{aligned} \Delta clk_{SAT} &= ((X_{AC} - X_{IGS}) \cdot X_{AC})/R_{SAT}/c \\ \Delta clk_{STA} &= ((Y_{AC} - Y_{IGS} - DY) \cdot Y_{AC})/R_{STA}/c \end{aligned} \tag{22.1}$$

where Δclk_{SAT} and Δclk_{STA} are the consistency corrections applied to AC station and satellite clock solutions respectively; X_{AC} and X_{IGS} are the coordinates of AC orbits and IGS combination orbit respectively, R_{SAT} and R_{STA} are the satellite and station radius vector respectively, c is the velocity of light, Y_{AC} is the AC station coordinate solution, Y_{IGS} is the IGS combination station coordinate solution, DY is the geocenter offset of the AC SINEX station solutions with respect to the IGS SINEX solutions [2, 3].

Station coordinate solution and geocenter offset can be directly obtained from SINX files. Satellite orbit coordinates need to be interpolated, so that orbits have the same sampling interval with combination clocks. Usually orbits' sampling interval is 15 min. However, the sampling interval of combination clocks and AC clocks is 30 s or 5 min. So we should select an exact interpolation algorithm for orbits interpolation. Chebyshev polynomial is a classical method of GPS satellite orbits standardization. The accuracy of this method is pretty good that nine-order interpolation accuracy is within one millimeter [5]. In this paper, the amount of nodes is $m > n + 1$ in order to fit more excellent coefficient [6]. Considering the computation and accuracy, we select 12 nodes and 9 order Chebyshev polynomial fitting algorithm for satellite precise orbit interpolation.

22.2.2 Unification of Time Reference Frame for AC Clock Solutions

Since AC clock solutions, in general, refer to different reference clocks, significant inconsistency errors can be introduced, in particular in the areas with solution gaps, or when satellite and/or station is missing from AC clock solutions. The AC reference clock errors for some ACs can exceed 1 ns. The clock combination accounts for these AC reference clock differences before the combination [2]. It is important that the alignment of the combined clocks to a chosen reference time. Currently broadcast clocks are used for this purpose but this can easily be changed [2, 3]. The clocks can be accurately fitted by an offset and a drift using robust least squares instead of the L1-norm estimation. At the same time the posteriori weighted mean square error of this fitting was used for the clock weighting.

L1-norm estimation minimizes the sum of the absolute value of residuals which only has the robustness but do not have minimum variance. A reliable initial is more important for robust estimation [7]. In the multi-dimensional case, there are two ways of solving such extreme problems. One is the linear programming method, the other one is the iteratively reweighted least squares (IRLS) method [8, 9]. Robust least squares estimates for the distribution of certain pollution still have minimum variance and unbiased statistical properties. The method has simple mathematical model and relatively simple calculation [8].

In this paper, the robust least squares method is used to fit transformation parameters of the time reference frame that AC clock solutions are relative to the broadcast clocks. Broadcast clocks can be calculated directly using clock coefficient in the navigation ephemeris file. Its sampling interval should be consistent with the combination clock. Some satellite broadcast clocks may appear clock jumps, which will have influence on the fitting of the transformation parameters. High-order differential method is used to detect the satellite clock jumps. Considering the number of satellites with clock jump is less and it is difficult to determine the size of the clock jump value accurately, we eliminate these satellites before the combination.

The error equation of robust least squares parameter estimation is as follows:

$$V = BX - L\,, \quad P_0 = I. \tag{22.2}$$

The clock solutions of AC all satellites throughout the day are aligned to the chosen reference time through fitting an offset and a drift. Due to that initial residual could not be calculated accurately, equivalent weight also couldn't be determined accurately. So we should get the weight through iteratively calculating [8]. The residuals of the first adjustment are determined by the least squares estimate. Because the least squares estimation is insensitive to gross errors. In the same condition, the absolute value of the statistics of error testing or the weight function of robust estimation reduces. At the same level of significance, the possibility of accepting false values in gross error detection or robust estimation

also increases. In order to reasonably reflect the size of the observation error and improve the resistance of robust estimation, we should select the variance factor statistic with good stability and high reliability as far as possible [10]. We use the residuals of the first adjustment by median method to calculate mean square error factor σ_0.

$$\sigma_0 = \frac{med(\left|\sqrt{p_i} \cdot v_i\right|)}{0.6745}. \tag{22.3}$$

The accuracy of broadcast clocks is poor, and AC satellite clocks also exist outliers. In order to resist these gross errors effectively and ensure the efficiency of parameter estimation, the calculation of equivalent weight is divided into two steps. First, we eliminate gross errors using strong eliminated weight function for obtaining better parameter initial value. Then we iteratively calculate equivalent weight using weak eliminated weight function which will improve the efficiency of parameter estimation. The strong eliminated weight function is as follows:

$$\bar{p}_i = \begin{cases} p_i & \left|v_i'\right| \le c_0 \\ 0 & \left|v_i'\right| > c_0 \end{cases}. \tag{22.4}$$

We choose Huber weight function as the weak eliminated weight function.

$$\bar{p}_i^k \begin{cases} \bar{p}_i^{k-1} & \left|v_i'^{k-1}\right| \le c_1 \\ \bar{p}_i^{k-1} \cdot \frac{c_1}{\left|v_i^{k-1}\right|} & \left|v_i' k - 1\right| > c_1 \end{cases} \tag{22.5}$$

The weight function is structured based on the student residual statistic. v_i' is the standardized residual [10]. The value of c_0 is 1.0 and c_1 is 1.2.

An offset and a drift can be accurately estimated using the above method. At the same time the posteriori mean square error of this fitting was used for the clock weighting. The weighted average clock correction for each satellite and epoch is then computed using the AC clock corrections and weights. A new set of alignment parameters (clock offset and drift) between the weighted clock average and each AC is estimated (one set of parameters for all satellites). Every AC clock solution is then realigned using these new parameters. The weight of AC clock correction is recalculated using the new posteriori mean square error. The accuracy and reliability of the new reference value is better than the broadcast clock correction, so we give a looser critical value for the strong eliminated weight function. The value of c_0 is 3.0, c_1 is the same as the above. The final clock correction is represented as clk'_{AC}. The weight of AC clock correction is represented as W_{AC}, which is as the initial weight of subsequent clock combination.

22.2.3 Clock Combination Based on Robust Estimation with High-Breakpoint of Pollution Rate

In view of that there are outliers in AC clock solution. We iteratively calculate equivalent weight using robust estimation with high-breakpoint of pollution rate. The combination clock correction is calculated based on weighted average. In this paper, we structure robust estimation with high-breakpoint of pollution rate based on the idea of double-elimination combination method. The clock combination belongs to one dimensional parameter estimation. The median of all the AC clock correction of the same epoch can provide a reliable initial value. The resistance of median is strong that breakpoint of pollution rate can reach 50 %. But its efficiency is low. The efficiency is only 64 % of LS estimation when the data is normal distribution. We choose median as the iterative initial value, which uses its strong resistance and avoids its low efficiency [8].

$$\hat{X}_0 = med(L_i) \tag{22.6}$$

where L_i is the individual AC clock correction, $\hat{X}_0$ is the initial solution of combination clock. The difference between AC clock solution and the initial solution is the initial residual.

$$v_i = L_i - \hat{X}_0 \tag{22.7}$$

We use the initial residual by median method to calculate mean square error factor σ_0. It's the same as Eq. 22.3. W_{AC} is used as the initial weight of AC clock solution. The strong eliminated weight function is same as Eq. 22.4, but the threshold value should be looser. The median as the initial value is reliable which can detect gross errors accurately. The value of c_0 can be loosen to 6.0.

The equivalent weight achieved through last step will be as the initial weight of next iterative calculation. We choose IGG-III as the iterative weight function [11, 12].

$$\bar{p}_i' = \begin{cases} \bar{p}_i & \left|v_i'\right| \le k_0 \\ \bar{p}_i \frac{k_0}{\left|v_i'\right|}\left(\frac{k_1 - \left|v_i'\right|}{k_1 - k_0}\right)^2 & k_0 < \left|v_i'\right| \le k_1 \\ 0 & \left|v_i'\right| > k_1 \end{cases} \tag{22.8}$$

where the value of k_0 can be chosen as 2.5 and k_1 as 6.0. After this step, the equivalent weight is represented as W_{AC}'. Finally, the combination clocks can be calculated as follows:

$$clk_{com} = \frac{\sum_{ac}^{Nac} W_{AC}' \cdot clk_{AC}'}{\sum_{ac}^{Nac} W_{AC}'} \tag{22.9}$$

where *Nac* is the total number of ACs. The RMS of individual AC with respect to the combination clock is calculated as follows:

$$RMS_{AC} = \sqrt{\frac{\sum_{sat}^{Nsat_{AC}} \sum_{i}^{Nclk_{AC}^{sat}} (clk_{com}^{sat} - clk_{AC}^{'sat})_i^2}{(\sum_{sat}^{Nsat_{AC}} Nclk_{AC}^{sat}) - 2}} \tag{22.10}$$

where $Nsat_{AC}$ is the total satellite number of AC, $Nclk_{AC}^{sat}$ is the total epoch number of an AC satellite.

22.3 Calculation and Analysis

To prove the accuracy and reliability of the method proposed in this paper, we selected 7 days' data from 128 to 134 of annual day in 2011 (GPS week 1635) for calculation and analysis. The sampling interval of combination clock is 30 s for satellite and 5 min for station separately with four ACs COD, EMR, ESA and MIT. The initial weight of AC clock solution plays a key role on the accuracy of combination clock. In order to find a better weight scheme, we have designed and performed seven kinds of weighting method as follows:

Case 1: The initial weight of AC clock solution is determined by the posteriori mean square error of fitting parameter.

Case 2: Each AC has the equal weight.

Case 3: The initial weight of AC each satellite clock is determined by the RMS separately. The RMS is calculated by individual AC clock solution with respect to the mean of the clock solutions of AC.

Case 4: The initial weight of AC clock solution can be appropriately determined according to the historical accuracy (see Fig. 22.1).

Case 5: The initial weight of AC clock correction is determined by the mean square error of clock correction provided in the clock file.

Case 6: The initial weight of AC clock correction is determined by the absolute deviation of AC clock solution with respect to the mean of the clock solutions of ACs at each epoch.

Case 7: The initial weight of AC clock correction is determined by the absolute deviation of AC clock solution with respect to the median of the clock solutions of ACs at each epoch.

The Table 22.1 shows week's average RMS of seven weighting scheme for the combination clock solution relative to the IGS final products. It can be concluded that scheme 4 is the best for satellite clock combination, scheme 7 is the best for station clock combination. The whole clock combination accuracy of scheme 4 is

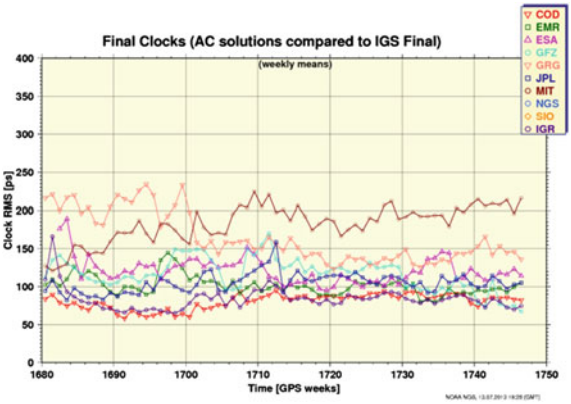


Fig. 22.1 RMS of AC clock solutions

Table 22.1 Comparison of clock combination accuracy for 7 weighting schemes (*unit* ps)

Method	CLKRMS	RMSSAT	RMSSTA
1	40.9	36.5	51.8
2	41.9	37.5	52.5
3	40.2	35.8	50.9
4	40.0	34.3	52.8
5	42.2	36.4	54.5
6	42.0	38.0	50.5
7	45.6	43.6	48.7

the best, scheme 3 and scheme 1 can also provide better results. The scheme 4 is difficult to perform because of its poor operability and applicability, which need to determine the weight according to the RMS showing in Fig. 22.1 artificially. The accuracy of scheme 1 and scheme 3 are both excellent. Besides, the calculation formula is also strict and clear. In this paper, we choose scheme 1 as the weighting method for the initial weight.

In order to evaluate the accuracy and reliability of the combination clock, we carried on the detailed comparison and analysis with the IGS final product. Figure 22.2 shows the RMS that the clock corrections of each satellite of AC relative to the IGS final. The abbreviation com refers to the combination solution. Figure 22.3 shows the RMS that the clock corrections of each satellite of analysis center relative to the combination solution. Figure 22.4 shows the RMS of satellite clock compared with the IGS final and combination solution respectively for every day. Figure 22.5 shows the RMS of station clock compared with the IGS final and combination solution respectively for every day. Table 22.2 lists the week's average RMS for satellite clock, station clock and the whole clock of analysis center relative to the IGS final separately. Table 22.3 lists the week's average RMS for satellite clock, station clock and the whole clock of AC relative to the combination solution separately. Finally, we examined the continuity of combination

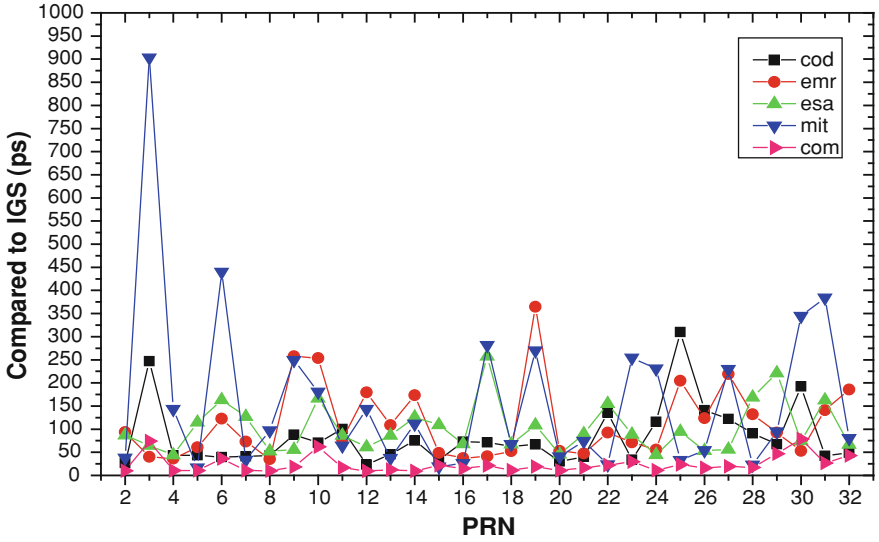


Fig. 22.2 RMS of each satellite with respect to the IGS final

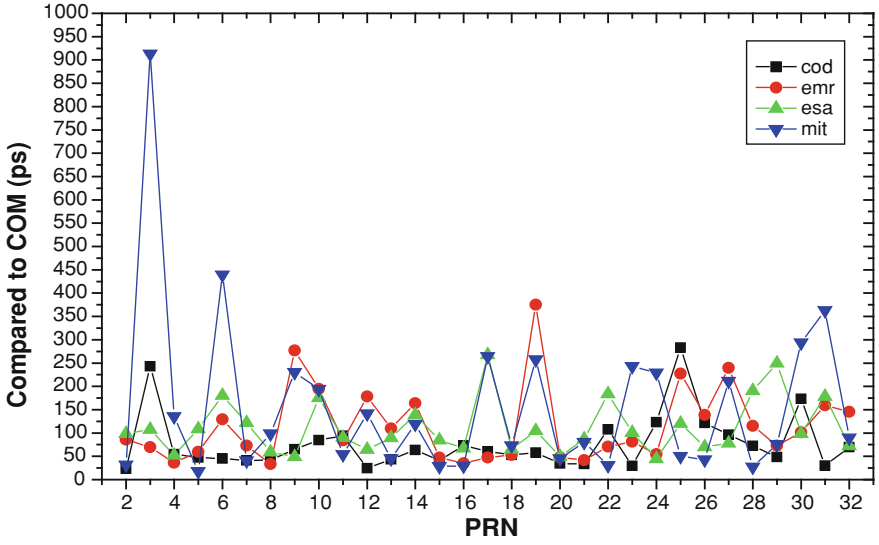


Fig. 22.3 RMS of each satellite with respect to the combination clock

clock corrections between day and day. It is difficult to show all the statistical results of the continuity of combination clock corrections in one figure. So we only give some statistical results as a representative here. Figure 22.6 is a variation of four satellites seven days all of the epoch combination clock corrections.

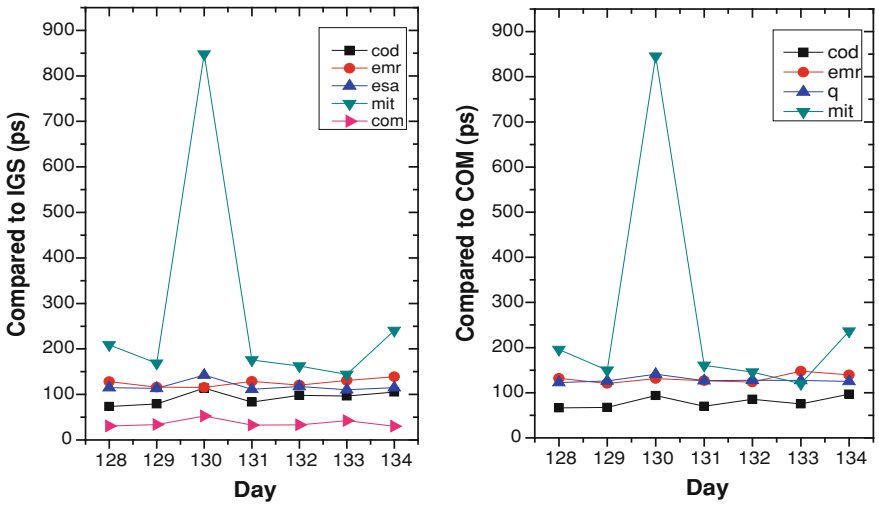


Fig. 22.4 RMS of everyday satellite clock with respect to IGS and COM separately

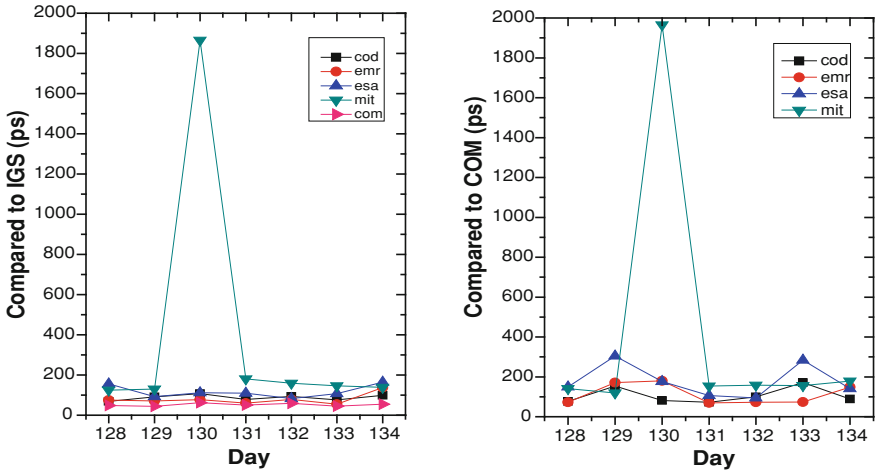


Fig. 22.5 RMS of everyday station clock with respect to IGS and COM separately

Table 22.2 The mean RMS of AC 7 days' clock solution with respect to IGS (*unit* ps)

AC	CLKRMS	RMSSAT	RMSSTA
COD	91.7	92.7	87.6
EMR	120.2	125.5	79.1
ESA	118.4	117.9	117.8
MIT	300.1	278.4	392.1
COM	40.9	36.5	51.8

Table 22.3 The mean RMS of AC 7 days' clock solution with respect to COM (*unit* ps)

AC	CLKRMS	RMSSAT	RMSSTA
COD	87.8	79.3	106.4
EMR	130.1	131.6	112.6
ESA	140.6	128.1	179.4
MIT	306.8	264.7	410.4

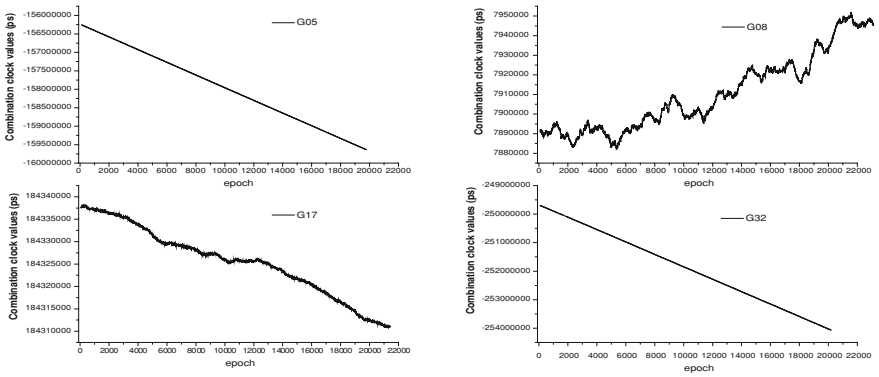


Fig. 22.6 The variation of combination clock in 7 days

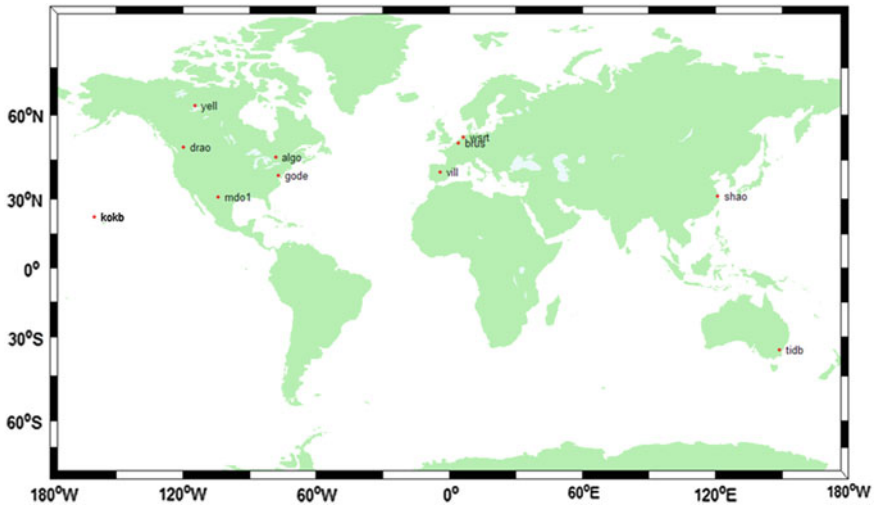


Fig. 22.7 The distribution of sites used

One way to evaluate the combination clock/orbit product is to use it in precise point positioning to analyze data from a single receiver [3]. We have selected eleven sites in global coverage (see Fig. 22.7) and one week period to perform

Table 22.4 Mean daily repeatability (mm) of PPP with different orbits/clocks

Product	period	North (mm)	East (mm)	Up (mm)
IGS	128–134	1.2	2.2	4.0
COM	128–134	1.5	3.1	4.2
EMR	128–134	2.7	4.8	8.0
MIT	128–134	2.6	5.8	6.2

Table 22.5 Mean ITRF position differences (mm) of PPP with different orbits/clocks

Product	period	X	Y	Z	3D
IGS	128–134	13.6	13.0	18.7	26.5
COM	128–134	13.1	13.8	17.3	25.7
EMR	128–134	12.1	14.3	18.0	26.0
MIT	128–134	11.2	14.0	19.0	26.1

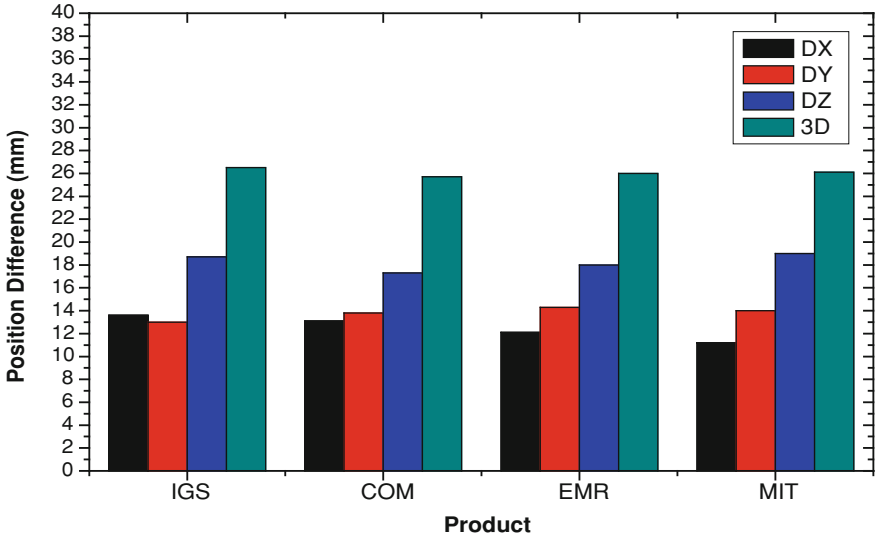


Fig. 22.8 IGS08 position differences (mm) of PPP with different orbits/clocks

such an evaluation by using the Bernese software. We also have carried on the static PPP test using analysis centers EMR, MIT and the IGS final product as a comparison. Table 22.4 shows the average daily repeatability of PPP. Table 22.5 and Fig. 22.8 both show the average position differences of PPP with respect to IGS weekly SINX solution [13].

As we can see from the above results:

(1) Scheme 1 and scheme 3 can both be used as the method of determining the initial weight of analysis center clock solution. The RMS of combination

clock relative to IGS final is in the 0.03–0.05 ns. The RMS of satellite clock is better than the station clock.

(2) The RMS of combination clock is obviously better than AC clock. The results of AC clock relative to IGS final relative to combination clock are almost the same. There are obvious outliers in clock corrections of MIT on the third day, which is consistent with the statistic results of IGS. The clock solution of MIT is not adopted for calculating combination clock on this day.

(3) Combination solution is continuous between day and day. The clock solutions of PRN05 and PRN32 satellite are with good linear variation. The clock drift rate of clock solutions of PRN08 and PRN17 satellite is obvious. The combination clock solutions of four satellites are all continuous, without clock jumps, which is consistent with the IGS final solution.

(4) The daily repeatability of PPP using combination clock is within 5 mm, which is a quite impressive achievement. The average daily repeatability of PPP of combination clock is better than analysis center EMR and MIT, but is slightly worse than IGS. The best repeatability is obtained on the North direction at 2 mm. The Up direction is the worst at 4–8 mm. The difference of average position differences of PPP of four kinds of product is very small. The accuracy of the X and Y direction is similar and Z direction is slightly worse. It is fully demonstrated by the results of static PPP test that the combination clock solutions have high accuracy and high reliability.

22.4 Conclusion

Precise clock correction is an extremely important product of IGS products. It is also the important basic data for PPP and satellite precise orbit determination. Using the combination method proposed in this paper, we can obtain a combination clock solution which is reliable and consistent with the IGS final clock correction. The RMS of the discrepancy between combination clock and IGS final clock is about 0.03–0.05 ns, which is significantly better than the AC clock solution. Currently, Chinese Beidou satellite navigation system (BDS) has been officially put into operation and opened to civil. Some BDS analysis centers will be build and provide all kinds of products including clock correction. This study will be significant for assessing the accuracy, reliability of individual BDS AC clock correction and providing a unified, high-precision, high-reliability clock products.

Acknowledgments This work was supported by the State key laboratory of geodesy and geodynamics open funded projects (SKLGED2013-4-2-EZ) and the national natural science fund project (41174008) and the Open Foundation of State Key Laboratory of Astronautic and Dynamics (2014ADL-DW0101).

References

1. Kouba J, Mireault Y, Lahaye F (1995, 1994) IGS orbit/clock combination and evaluation, Appendix I of the analysis coordinator report, International GPS service for geodynamics (IGS) 1994 annual report, pp 70–94
2. Kouba J, Springer Tim (2001) New IGS station and satellite clock combination. GPS Solution 4(4):31–36
3. Springer TA, Zumberge JF, Kouba J (1998) The IGS analysis products and consistency of the combined solutions, IGS analysis center workshop, Darmstadt, 9–11 Feb 1998
4. Kouba J, Mireault Y (1997) Analysis coordinator report international GPS service for geodynamics (IGS) 1996 annual report, pp 55–100
5. Wei EH, Chai H (2006) Comparative study on the interpolation method of GPS precise ephemeris, GNSS World of China. 5:13–15 (In Chinese)
6. Jiang Nan, Xu TH, Xu Yan (2012) An improved method for determination of GOCE orbital velocity of the geometric method, Bulletin of Surveying and Mapping. 11:7–10 (In Chinese)
7. Qiu WN (2003) Selecting weight iteration method with reliable initial values. Geomatics and Information Science of Wuhan University 8:452–454 (In Chinese)
8. Zhou JW, Huang YC, Yang YX, Ou JK (1997) Robust least square method. Huazhong University of Science and Technology Press, Wuhan (In Chinese)
9. Lauknes TR, Zebker HA, Larsen Y (2010) InSAR deformation time series using an L1-Norm small-baseline approach. IEEE Trans Geosci Remote Sens 114:2097–2109
10. Yang YX, Wu FM (2006) Robust estimation equivalent weight function with variable threshold. J Geomatics Sci Technol 10:317–320 (In Chinese)
11. Yang YX (1993) Robust estimation theory and it's applications. Bayi Publishing House, Beijing (In Chinese)
12. Ou JK (1996) An design of three-step robust solutions. Acta Geodaetica et Cartographica Sinica, 8:173–179 (In Chinese)
13. Kouba J, Heroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solution 5(2):12–28

Chapter 23
GNSS Dynamic PPP Based on Additional Priori Coordinate and Epoch Constraints Information

Nan Zang, Qin Zhang, Shijun Li, Guanwen Huang and Shuangcheng Zhang

Abstract With real-time communication transmission technology development and global GNSS gradually increasing the number of continuous tracking station, No base station to support real-time Precise Point Positioning (RTPPP) technology gradually been widespread concern and applications. Especially in some type of slowly varying time monitoring and early warning of geological disasters. Disasters such as landslides, mine subsidence, earthquake monitoring. In order to avoid divergent positioning results, real-time precise point positioning technology in general can be sequential adjustment method solver. Since real-time PPP technology is a dynamic PPP positioning, coordinates are not generally adding constraints but choosing single epoch estimation. Considering some the motion characteristics of low moving carriers, the carrier has known more precise priori coordinates before positioning itself. At the same time the carrier at a certain fixed length of time (e.g. 2 h) can be considered stationary. The a priori information can be used to help improve the positioning accuracy of the carrier real-time PPP. Based on this starting point, in this paper, A real-time precise point positioning (PPP) technology based on priori coordinate and epoch constraints information. And use the Chang'an University Compass Analysis Center provides real-time BDS orbit and clock products. Combined with BDS monitoring data for experimental verification. Demonstrated the effectiveness of the algorithm and accuracy.

Keywords Real-time · Precise point positioning · Priori coordinate · Epoch constraint

N. Zang (✉) · Q. Zhang · S. Li · G. Huang · S. Zhang
College of Geology Engineering and Geomantic, Chang'an University, Yanta Road 126
710054 Xi'an, People's Republic of China
e-mail: zang6050@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_23,

23.1 Introduction

With the American GPS, the Russian GLONASS, European GALILEO, China's Beidou satellite navigation system satellite navigation system gradually established and perfected, GNSS global satellite navigation system has been progressively oriented the development of multi-constellation portfolio compatible with the direction of cooperation. As the precise ephemeris and precise clock have continuously improved provided by IGS, precise point positioning technology (Precise Point Positioning, referred to as PPP) becomes today's global a hotspot in the field of satellite navigation and positioning currently. Among GNSS sophisticated process software such as Bernese GIPSY and Panda have been achieved the precise point positioning function [1, 2], in addition, the Canada University of Calgary precise point positioning software named P3 has also been widely used [3]. Meanwhile, many domestic research scholars has also conducted in-depth research on PPP technology, and has made many research results [4–8]. Currently, the post-processing accuracy of static PPP reaches mm–cm, dynamic positioning accuracy is cm–dm, due to its significant economic and flexibility, it has been widely used [9, 10]. Real-time PPP technology is dynamic PPP, positioning coordinates are not generally with additional constraints but choose single epoch estimated. For dynamic precise point positioning, because it has been severely affected by measurement noise, measurement error, environmental factors, and highly requires the accuracy of satellite orbit and clock errors and other reasons, its positioning accuracy can often just gets cm-dm-level [11], so it is difficult to meet the needs of the high positioning accuracy requirements of the dynamic measurement.

This paper do some research on the real-time dynamic PPP study for regional graded type disaster, mainly including the monitoring of landslide, mine subsidence monitoring, dam deformation monitoring and earthquake monitoring. Currently those monitoring work has been done through designing GNSS monitoring network, the stability of reference point have effect on the monitoring results, and can only obtain the relative deformation, other than absolute deformation. Therefore, this paper will consider taking the dynamic GNSS PPP technology into these disasters monitoring work. The accuracy of traditional single epoch dynamic PPP is low, which can not meet the need of highly precise disaster monitoring. Taking into account the low rate features of the above—mentioned types of graded type carriers, and the carrier itself knows more precise priori coordinates before positioning, while the carrier is considered stationary in a fixed length of time as the landslide within two hours is considered stationary, these priori information helps to improve the positioning accuracy of the real-time PPP. Proposed GNSS dynamic PPP algorithm based additional priori coordinates and epoch constraints for this information. Firstly, the processing data is divided into fixed length of time, and considered stationary in the arcs, in the arc we use the static PPP solution for epoch solution, and then the priori variance of the first epoch of each arc is constraint with priori coordinate and epoch constraints. Study results showed:

Dynamic precise point positioning accuracy of the proposed method compared to conventional dynamic PPP solver has greatly improved, this algorithm is especially suitable for low dynamic real-time monitoring of graded type disaster.

23.2 Precise Point Positioning Conventional Mathematical Model

23.2.1 Precise Point Positioning Observation Equation

Precise point positioning technique uses of post-precise orbit and precise satellite clock error products, it uses the un-differenced positioning processing with the phase and pseudorange observations obtained by the dual-frequency receivers, in general we use a combination of the elimination of the ionosphere (LC combination) observation equation, The mathematical model is as follows [1–3]:

$$\mathbf{P}_{\mathrm{IF}} = \frac{\mathrm{f}_1^2 \cdot \mathbf{P}_1 - \mathrm{f}_2^2 \cdot \mathbf{P}_2}{\mathrm{f}_1^2 - \mathrm{f}_2^2} = \boldsymbol{\rho} + \mathrm{c}\mathbf{dt} + \mathbf{d}_{\mathrm{trop}} + \boldsymbol{\varepsilon}(P_{\mathrm{IF}}) \tag{23.1}$$

$$\boldsymbol{\Phi}_{\mathrm{IF}} = \frac{\mathrm{f}_1^2 \cdot \boldsymbol{\Phi}_1 - \mathrm{f}_2^2 \cdot \boldsymbol{\Phi}_2}{\mathrm{f}_1^2 - \mathrm{f}_2^2} = \boldsymbol{\rho} + \mathrm{c}\mathbf{dt} + \mathbf{d}_{\mathrm{trop}} + \frac{\mathrm{cf}_1\mathbf{N}_1 - \mathrm{cf}_2\mathbf{N}_2}{\mathrm{f}_1^2 - \mathrm{f}_2^2} + \boldsymbol{\varepsilon}(\Phi_{\mathrm{IF}}) \tag{23.2}$$

In the equations, $\mathbf{P}_i$ is the code observation loading on carrier $\mathbf{L}_i$; $\boldsymbol{\Phi}_i$ is the carrier phase observation; $\boldsymbol{\rho}$ is the distance between satellite and station; c is lightspeed; $\mathbf{dt}$ is the clock error of receivers ; $\mathbf{d}_{\mathrm{trop}}$ is the troposphere delay; f_i is the frequency of $\mathbf{L}_i$; $\mathbf{N}_i$ is the integer ambiguity of $\mathbf{L}_i$; $\boldsymbol{\varepsilon}(.)$ is measurement noise.

23.2.2 Parameter Estimation Method-Sequential Least Squares Parameter Adjustment

The paper uses the parameter estimation method is the sequential least squares method. The main thought of the method is as follows [12]: Let the k-1 epoch observation is denoted as $\underset{n_{k-1}}{L_{k-1}}$, its weight matrix is P_{k-1} ,Then the error equation is:

$$V = B\hat{x} - l \tag{23.3}$$

$$V = B_1x_1 + B_2x_2 + B_3x_3 - l \tag{23.4}$$

In the equation: $\hat{x} = [\,x_1 \quad x_2 \quad x_3\,]^T$, $B = [\,B_1 \quad B_2 \quad B_3\,]$, $x_1 = t$ is the clock error of receivers, $x_2 = [\,x \quad y \quad z\,]^T$ is the coordinate of receivers, $x_3 = [\,\mathrm{ZTD} \quad \mathrm{N}\,]^T$, ZTD is the troposphere delay, N is the integer ambiguity.

We obtain the k-1 epoch of the normal equation

$$\begin{bmatrix} N_{11}^{k-1} & N_{12}^{k-1} & N_{13}^{k-1} \\ N_{21}^{k-1} & N_{22}^{k-1} & N_{23}^{k-1} \\ N_{31}^{k-1} & N_{32}^{k-1} & N_{33}^{k-1} \end{bmatrix} \begin{bmatrix} x_1^{k-1} \\ x_2^{k-1} \\ x_3^{k-1} \end{bmatrix} = B_{k-1}^T P_{k-1} l_{k-1} \tag{23.5}$$

$$N_{ij}^{k-1} = B_j^T P_{ij}^{k-1} B_i \ \ (i,j=1,2,3) N_{\hat{X}_{k-1}} = \begin{bmatrix} N_{11}^{k-1} & N_{12}^{k-1} & N_{13}^{k-1} \\ N_{21}^{k-1} & N_{22}^{k-1} & N_{23}^{k-1} \\ N_{31}^{k-1} & N_{32}^{k-1} & N_{33}^{k-1} \end{bmatrix}$$

In the equation:

After adjustment:

$$\begin{bmatrix} x_1^{k-1} \\ x_2^{k-1} \\ x_3^{k-1} \end{bmatrix} = \begin{bmatrix} Q_{11}^{k-1} & Q_{12}^{k-1} & Q_{13}^{k-1} \\ Q_{21}^{k-1} & Q_{22}^{k-1} & Q_{23}^{k-1} \\ Q_{31}^{k-1} & Q_{32}^{k-1} & Q_{33}^{k-1} \end{bmatrix} B_{k-1}^T P_{k-1} l_{k-1} \tag{23.6}$$

In equation:

$$\hat{x}_{k-1} = \begin{bmatrix} x_1^{k-1} \\ x_2^{k-1} \\ x_3^{k-1} \end{bmatrix}, \ Q_{\hat{X}_{k-1}} = \begin{bmatrix} Q_{11}^{k-1} & Q_{12}^{k-1} & Q_{13}^{k-1} \\ Q_{21}^{k-1} & Q_{22}^{k-1} & Q_{23}^{k-1} \\ Q_{31}^{k-1} & Q_{32}^{k-1} & Q_{33}^{k-1} \end{bmatrix} = N_{\hat{X}_{k-1}}^{-1},$$

In brief:

$$\hat{x}_{k-1} = Q_{\hat{X}_{k-1}} B_{k-1}^T P_{k-1} l_{k-1} \tag{23.7}$$

In equation $Q_{\hat{X}_{k-1}} (= Q_{\hat{X}_{k-1}\hat{X}_{k-1}})$ is the covariance matrix of $\hat{X}_{k-1}$ (that is $\hat{x}_{k-1}$), Therefore

$$Q_{\hat{X}_{k-1}} = (B_{k-1}^T P_{k-1} B_{k-1})^{-1} \tag{23.8}$$

$$Q_{\hat{X}_{k-1}}^{-1} = B_{k-1}^T P_{k-1} B_{k-1} \tag{23.9}$$

$\hat{x}_{k-1}$ represents the value of $\hat{x}$ obtained by adjustment of k-1 epoch.

Put the equation of the last epoch with the k epoch equation joint solving, from two observations as a whole adjustment, can be composed of normal equation:

$$\left(Q_{\hat{X}_{k-1}}^{-1} + B_k^T P_k B_k\right)\hat{x} - \left(B_{k-1}^T P_{k-1} l_k + B_k^T P_k l_k\right) = 0 \tag{23.10}$$

The solution is:

$$\hat{x} = \left(Q_{\hat{X}_{k-1}}^{-1} + B_k^T P_k B_k\right)^{-1} \left(B_{k-1}^T P_{k-1} l_{k-1} + B_k^T P_k l_k\right) \tag{23.11}$$

Equations (23.3)–(23.11) is the sequential adjustment solution formula.

23.3 PPP Algorithm Based on Additional Priori Coordinate and Epoch Constraints

23.3.1 Dynamic PPP Based on Additional Priori Coordinate

The receiver clock error in dynamic PPP parameter is a dynamic parameter, can not be inherited and constrained epoch by epoch, therefore, the parameters need to be reduced, which is equivalent to make its variance–covariance matrix plus with a larger value and the specific changes are as follows:

$$Q_{11_{\text{new}}}^{k-1} = Q_{11}^{k-1} + 10^{10} \tag{23.12}$$

Q_{22}^{k-1} is the covariance matrix of the k-1 epoch coordinates (x, y, z), $Q_{22}^{k-1} = \begin{bmatrix} \sigma_{x^{k-1}}^2 & & \\ & \sigma_{y^{k-1}}^2 & \\ & & \sigma_{z^{k-1}}^2 \end{bmatrix}$. The main change of method of PPP based on additional priori coordinate is giving the priori variance Q_{22}^{k-1} a constraint Q_{c}^{k-1} to obtain a new coordinates priori variance $Q_{22_{\text{new}}}^{k-1}$

$$Q_{\text{c}}^{k-1} = \begin{bmatrix} \sigma_c^2 & & \\ & \sigma_c^2 & \\ & & \sigma_c^2 \end{bmatrix} \tag{23.13}$$

$$Q_{22_{\text{new}}}^{k-1} = Q_{22}^{k-1} + Q_c^{k-1} \tag{23.14}$$

Put $Q_{22_{\text{new}}}^{k-1}$ into Eqs. (23.10) and (23.11), we can obtain the k epoch parameter estimation by sequential least squares adjustment.

23.3.2 Dynamic PPP Based on Additional Epoch Constraints

The observation information is divided into one hour interval as an arc. The coordinates of the first epoch of each arc segment should be initialized, does not inherit any information on last arc segment. In the arc should use conventional static precise point positioning for location results.

Do the same type (23.12) the receiver clock error change into dynamic parameters by reduction parameter. Give the diagonal elements of Q_{22}^{k-1} plus a larger value, the increased matrix form as follows:

$$Q_{\text{d}}^{k-1} = \beta \begin{bmatrix} 10^{10} & & \\ & 10^{10} & \\ & & 10^{10} \end{bmatrix} \tag{23.15}$$

In the equation: β is a control variable $\beta = \begin{cases} 0 & \text{In arc epoch} \\ 1 & \text{others} \end{cases}$

$$Q_{22_{\text{new}}}^{k-1} = Q_{22}^{k-1} + Q_{\text{d}}^{k-1} \tag{23.16}$$

Put $Q_{22_{\text{new}}}^{k-1}$ into Eqs. (23.10) and (23.11) we can obtain the k epoch parameter estimation by sequential least squares adjustment.

23.3.3 Dynamic PPP Based on Additional Priori Coordinates and Epoch Constraints Information

Combination with the above two, divided the observation information into one hour intervals as an arc. At the first epoch of each arc should constrain the priori variance of coordinates. But in the arc it should use the conventional static precise point positioning to get the positioning results.

Do the same type (23.12) the receiver clock error change into dynamic parameters by reduction parameter and make Q_{22}^{k-1} plus constraint Q_{e}^{k-1}.

$$Q_{\text{e}}^{k-1} = \beta \begin{bmatrix} \sigma_e^2 & & \\ & \sigma_e^2 & \\ & & \sigma_e^2 \end{bmatrix} \tag{23.17}$$

In the equation β is consistent with the definition in the previous section.

$$Q_{22_{\text{new}}}^{k-1} = Q_{22}^{k-1} + Q_{\text{e}}^{k-1} \tag{23.18}$$

Put $Q_{22_{\text{new}}}^{k-1}$ into Eqs. (23.10) and (23.11) we can obtain the k epoch parameter estimation by sequential least squares adjustment.

23.3.4 Algorithm Implementation Process

For GNSS real-time kinematic precise point positioning function in multi-system, it mainly uses ionosphere-free PPP model for positioning solution. Specific implementation process as shown (Fig. 23.1):

Firstly, the ionosphere-free combination observation equation is used to eliminate the ionosphere first order item error, and then uses the model to correct the troposphere, tides, tidal pole shift, the antenna phase center offset, phase turn around and other errors, and use the parameter estimates to further correct troposphere residual error, Finally, uses sequential least squares parameter estimation method based on additional priori coordinates and epoch constraints to get the station coordinates.

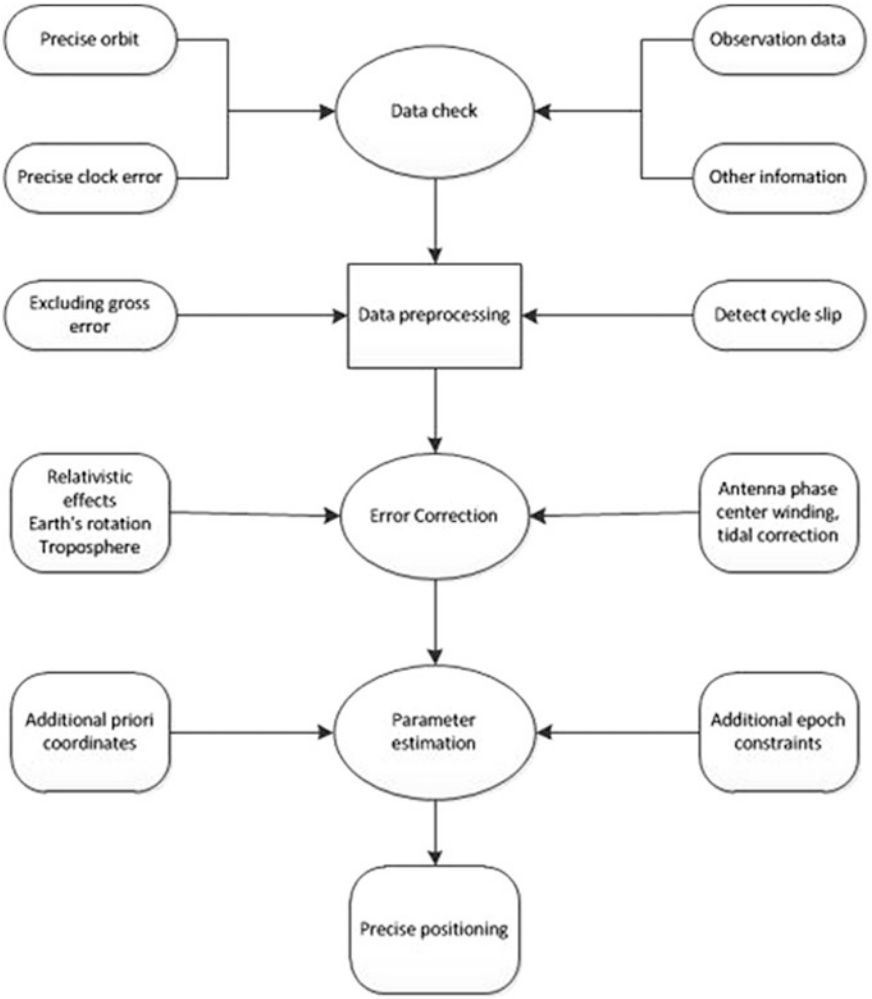


Fig. 23.1 PPP based on additional priori coordinate and epoch constraints information

23.4 Results Analysis

23.4.1 Precise Point Positioning Based on GPS Test

In order to verify the accuracy and effectiveness of the algorithm, we select the observation data of February 9, 2013 nain, tskb and thu3 3 IGS stations, data sampling interval is 30 s. The selection of IGS precise ephemeris is 15 min and 30 s sampling rate precise clock error file to solve. In this paper, we use the dynamic method for static data.

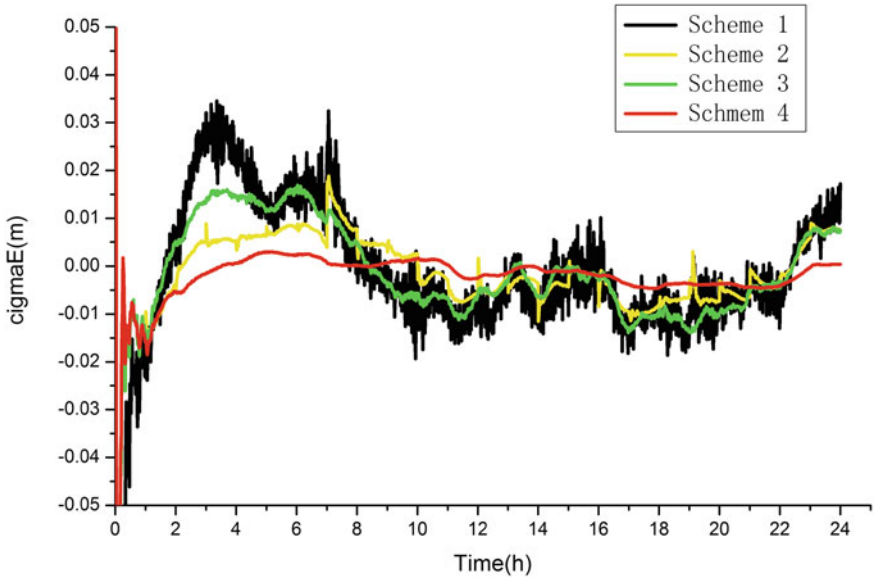


Fig. 23.2 The E orientation residuals of Thu3 station

Designed 4 programs for comparison and analysis as follows:

Scheme 1: The traditional dual-frequency dynamic precise point positioning method of solver.
Scheme 2: Dynamic PPP based on additional priori coordinate of solver.
Scheme 3: Dynamic PPP based on additional epoch constraints of solver.
Scheme 4: Dynamic PPP based on additional priori coordinate and epoch constraints information of solver.

In order to compare the results of the four schemes obtain. The results obtained from four programs were compared with the true value, Figs. 23.2, 23.3 and 23.4 represents about Thu3 station residuals in ENU direction, the average precision statistic of three stations and convergence time Statistic are shown in Fig. 23.5, Tables 23.1 and 23.2.

It can be seen from Figs. 23.2, 23.3 and 23.4: the positioning accuracy and convergence rate of Scheme 2, Scheme 3 and Scheme 4 which are due to the effective use of the priori coordinate and epoch constraints information of graded type carriers are more improved than dynamic precise point positioning. And the results of Scheme 4 are significantly improve data convergence times, and accuracy after convergence. Figure 23.5 is about GPS data for statistical accuracy of the three stations, the dynamic precise point positioning of improved Scheme 3 and Scheme 4 is far superior to the traditional dynamic precise point positioning algorithm. Table 23.1 is a description of the accuracy of the four schemes: the plane accuracy of scheme 1 is 2.1 cm and elevation accuracy is 4.5 cm; the plane accuracy of scheme 2 is 1.2 cm and elevation accuracy is 3.5 cm; the plane accuracy of scheme 3 is 1.3 cm

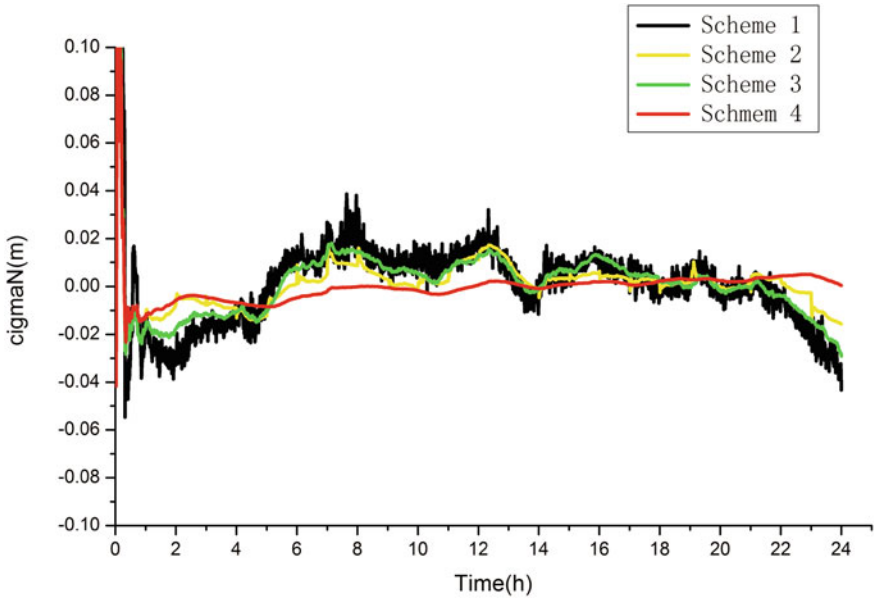


Fig. 23.3 The N orientation residuals of Thu3 station

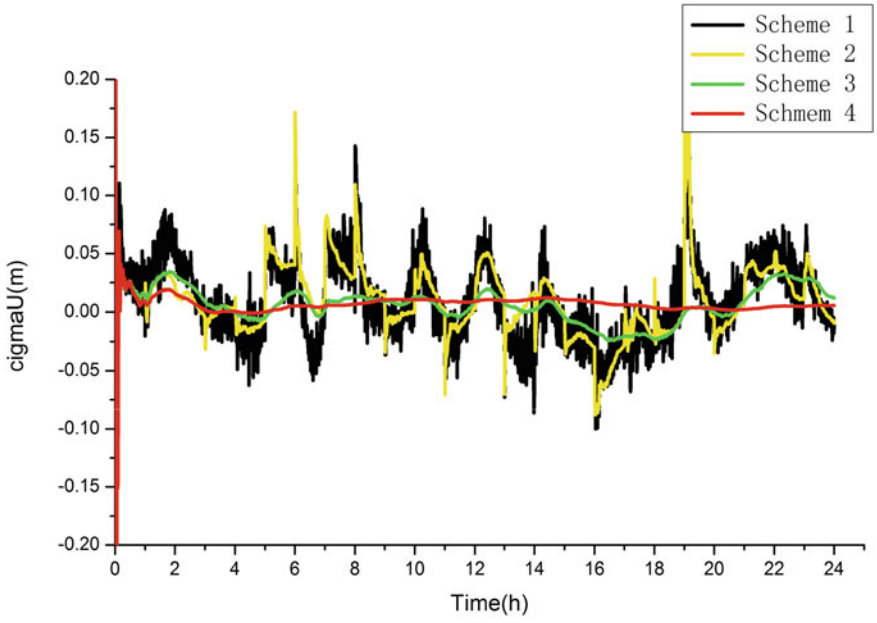


Fig. 23.4 The U orientation residuals of Thu3 station

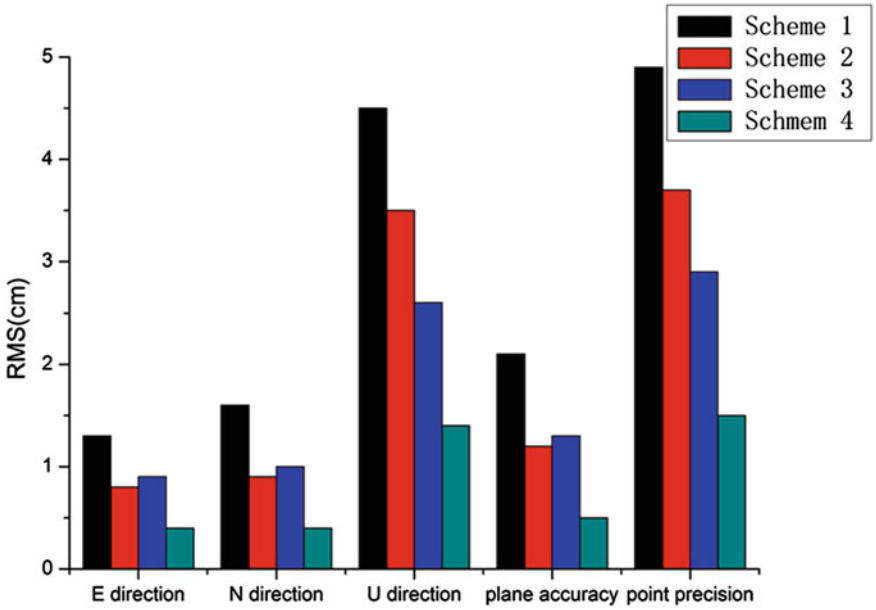


Fig. 23.5 The ENU direction average accuracy indicators statistics of three stations

Table 23.1 Three stations average accuracy statistics

Average RMS of three stations(cm)	Scheme 1	Scheme 2	Scheme 3	Scheme 4
E direction	1.3	0.8	0.9	0.4
N direction	1.6	0.9	1.0	0.4
U direction	4.5	3.5	2.6	1.4
Plane accuracy	2.1	1.2	1.3	0.5
Point precision	4.9	3.7	2.9	1.5

Table 23.2 Three stations average convergence time statistics

Average convergence time of three stations (cm)	Scheme 1 (h)	Scheme 2 (min)	Scheme 3	Scheme 4
<5	8.2	8.5	8.5 min	8.5 min
<3	No reach	16.5	16.5 min	16.5 min
<1	No reach	No reach	1.5 h	1.25 h

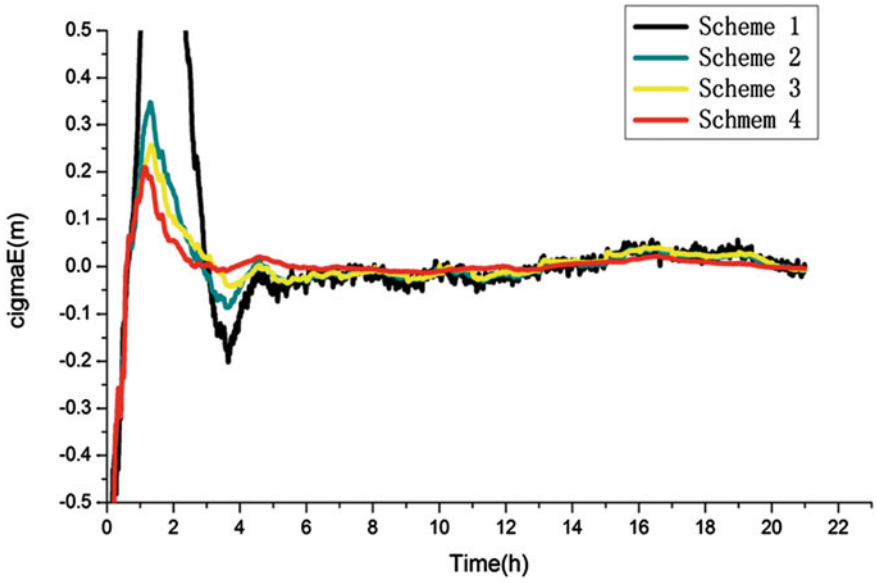


Fig. 23.6 The E orientation residuals of Cha1 station

and elevation accuracy is 2.6 cm; the plane accuracy of scheme 4 is 0.5 cm and elevation accuracy is 1.4 cm; Table 23.2 shows the convergence time of scheme 4 has significantly improved compared with other schemes. The results show that the plane accuracy of scheme 4 can reach millimeter and it is far superior to scheme 1. The accuracy of the elevation direction has greatly improved.

23.4.2 Precise Point Positioning Based on BDS Test

In order to verify the accuracy and effectiveness of the algorithm for BDS, we select the observation data of February 23, 2013 cha1, bjf1, gua1, kun1, lha1 and xia1 6 stations, data sampling interval is 30 seconds. The selection of BDS precise ephemeris is 15 min and 30 s sampling rate BDS precise clock error file to solve. In this paper, we use the dynamic method for static data. The solutions adopted in line with the above scheme. In order to compare the results of the four schemes obtain. The results obtained from four programs were compared with the true value, Figs. 23.6, 23.7 and 23.8 represents about cha1 station residuals in ENU direction, the average precision statistic of six stations and convergence time statistic are shown in Fig. 23.9, and Table 23.3.

It can be seen from Fig. 23.6, 23.7 and 23.8: the positioning accuracy and convergence rate of Scheme 2, Scheme 3 and Scheme 4 which are due to the effective use of the priori coordinate and epoch constraints information of graded type carriers are more improved than dynamic precise point positioning. And the results of

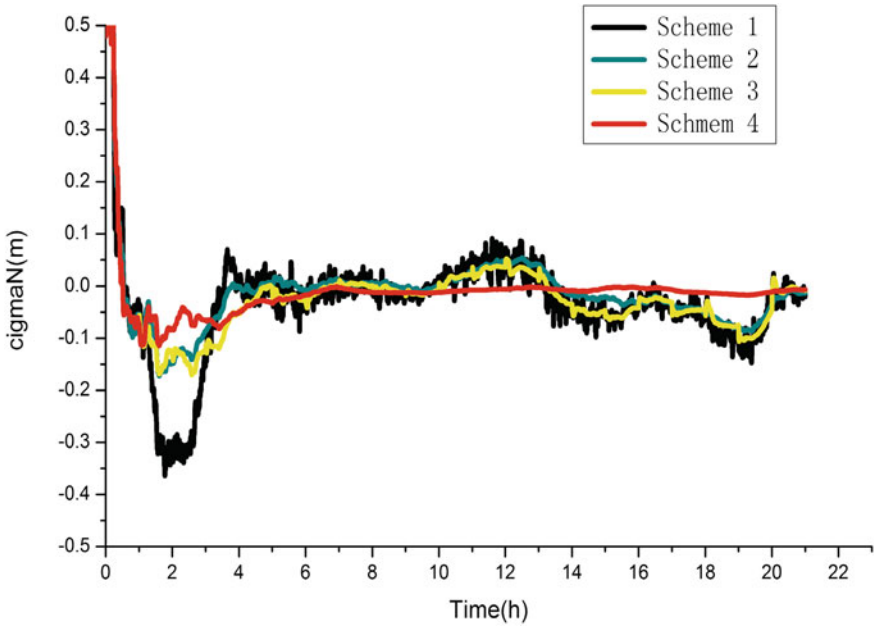


Fig. 23.7 The N orientation residuals of Cha1 station

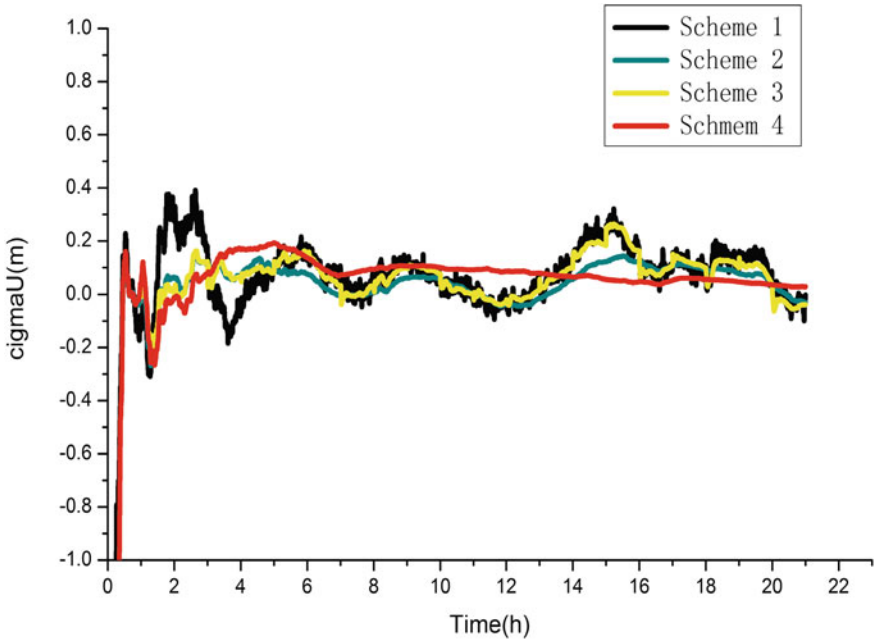


Fig. 23.8 The U orientation residuals of Cha1 station

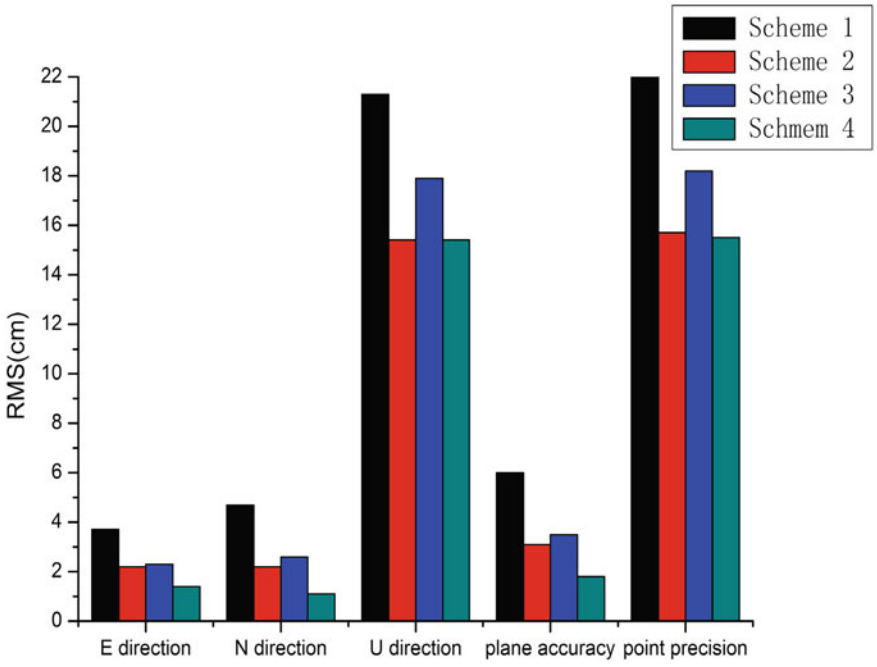


Fig. 23.9 The ENU direction average accuracy indicators Statistics of 6 stations

Table 23.3 Six stations average accuracy statistic

Average RMS of six stations (cm)	Scheme 1	Scheme 2	Scheme 3	Scheme 4
E direction	3.7	2.2	2.3	1.4
N direction	4.7	2.2	2.6	1.1
U direction	21.3	15.4	17.9	15.4
Plane accuracy	6.0	3.1	3.5	1.8
Point precision	22.1	15.7	18.2	15.5

Scheme 4 are significantly improve data convergence times, and accuracy after convergence. Figure 23.9 is about BDS data for statistical accuracy of the six stations, the dynamic precise point positioning of improved Scheme 3 and Scheme 4 is far superior to the traditional dynamic precise point positioning algorithm. Table 23.3 is a description of the accuracy of the four schemes: the plane accuracy of scheme 1 is 6.0 cm and elevation accuracy is 21.3 cm; the plane accuracy of scheme 2 is 3.1 cm and elevation accuracy is 15.4 cm; the plane accuracy of scheme 3 is 3.5 cm and elevation accuracy is 17.9 cm; the plane accuracy of scheme 4 is 1.8 cm and elevation accuracy is 15.4 cm; The results show that the plane accuracy of scheme 4 is far superior to scheme 1. The accuracy of the elevation direction is also far superior to scheme 1.

23.5 Conclusions and Recommendations

For the known priori coordinates and epoch constraints information of some graded type carriers, this paper proposes a dynamic GNSS PPP technology based on additional priori coordinate and epoch constraints which can improve the accuracy of dynamic positioning. This method has the advantage of reducing convergence time and improving the positioning accuracy. Additional priori information is using priori known more precise coordinate information of carriers so that the positioning accuracy is improved. Additional epoch constraints information is that the graded type carriers are considered stationary in one hour arc. Using static precise point positioning to solve can reduce convergence time and improve the positioning accuracy. Combining the advantages of both, the additional priori coordinates and epoch constraints precise point positioning can quickly converge and accuracy after convergence can be greatly improved compared to traditional dynamic precise point positioning. This algorithm is taking into the slow degeneration of the carrier, can be applied to real-time slow variations in disasters such as landslide monitoring, mine subsidence monitoring, dam deformation monitoring and earthquake monitoring work.

Acknowledgments Geographic Information Engineering National Key Laboratory Fund (SKLGIE2013-Z-2-1), State Key Laboratory of Aerospace Dynamics Fund (2013ADL-DW0103), National Natural Science Foundation Youth Foundation (41304033, 41104019,41274005), second navigation significant special issue (GFZX0301040308), China geological Survey geological survey project (1212011220186).

References

1. Zumberge JF (1999) Automated GPS data analysis service. GPS Solut 2(3):76–78
2. Hugentobler U, Dach R, Fridez P (2005) Document of Bernese of software version 5.0. University of Bern, Bern
3. Gao Y (20047) Performance analysis of precise point positioning using real-time orbit and clock products. J Global Position Syst 3(1):95–100
4. Liu J, Ye S (2002) GPS precise point positioning using undifferenced phase observation. Geomat Inform Sci Wuhan Univ 3(27):234–239
5. Zhang X, Liu J, Rene F (2006) Application of precise point positioning in airborne survey. Geomat Inform Sci Wuhan Univ 1(31):19–22
6. Han B, Ou J (2003) Precise point positioning based on undifferenced GPS data. Geomat Inform Sci Wuhan Univ 4(28):409–412
7. Hu C, Chen W, Gao S et al (2005) Data processing for GPS precise point positioning.Trans Nanjing Univ Aeronaut Astronaut 22(2):43–50
8. Zhang X, Li X, Guo P et al (2008) Realization and precision analysis of single-frequency precise point positioning software. Geomat Inform Sci Wuhan Univ 33(8):783–787
9. Huang G (2008) Research on GPS precise point positioning and adjustment of high precision GPS baseline network and the Realization of its software. Master's degree thesis, Chang'an University, Xian

10. Nie J, Yang Y, Wu F (2010) An algorithm of dynamic precise point positioning based on modified particle filtering. Acta Geodaetica Cartogr Sin 39(4):338–343
11. Zheng Z et al (2009) Analysis of factors affecting convergence time and measure in GPS precise point positioning. J Geodesy Geodyn 2009(5):107–111
12. Yang Y (2006) Adaptive navigation and kinematic positioning. Cartographica Press, Beijing

Chapter 24
Model and Data Driven Partial Ambiguity Resolution for Multi-Constellation GNSS

Yanqing Hou and Sandra Verhagen

Abstract The goal of carrier-phase ambiguity resolution is to exploit that the carrier-phase observations start to act as very precise pseudoranges. With the development of modern GNSS (GPS, BDS, Galileo, Glonass), more than 30 satellites are visible, however, it might be impossible to reliably fix all the ambiguities due to the computation time. Additionally, due to high measurement noise or residual atmosphere delays in case of longer baselines, the observation model is not strong enough, which makes it impossible to fix all the ambiguities. Therefore Partial Ambiguity Resolution (PAR) becomes more and more essential for real-time precise positioning. In this contribution, a Model and Data driven PAR (MD-PAR) strategy is proposed, and implemented in two different ways. The performance of MD-PAR is assessed using a simulation study by the probability of correct subset fixing, the subset size, and the Root Mean Square (RMS) of the baseline solution. Furthermore, MD-PAR is compared with the classic strategy, which uses only model information. The analysis and simulation results both suggest that the new strategies have better performance than the classic strategy.

Keywords Partial ambiguity resolution · Model and data driven · Multi-GNSS · Success rate · Criterion · Ratio test

24.1 Introduction

The carrier-phase observations start to act as very precise pseudoranges once the unknown integer number cycles has been resolved. With the upcoming BeiDou Satellite System and Galileo, the number of operational GNSS satellites and signals will be increasing dramatically in the near future, such that by 2020 we may

Y. Hou (✉) · S. Verhagen
Geoscience and Remote Sensing, Delft University of Technology,
Stevinweg 1 2628 CN Delft, Netherlands
e-mail: h.yanqing@tudelft.nl

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_24,

have 30 visible satellites and more than 100 observables available at once. This may necessitate ambiguity resolution of the full set being not possible, due to the required computation time. Besides, in case of longer baselines, the observation model is not strong enough due to high measurement noise or residual atmosphere delays, which makes it impossible to fix all the ambiguities. Therefore PAR becomes more and more essential for real-time precise positioning.

Many strategies for the PAR have been proposed in literature. These strategies can be generally divided into three levels and two categories as follows.

1. Satellite level

The strategies in this level choose subsets of satellites. This is based on the assumption that the highly noisy ambiguities are from several certain satellites, which hinder the reliable ambiguity fixing. Thus the reliable ambiguity fixing may become immediately possible after removing these satellites. The strategies in this level include the elevation ordering strategy [1, 2], SNR ordering strategy [3]. The visible satellites are ordered by elevation or by SNR, and the lower satellites or the noisier satellites are removed if Full Ambiguity Resolution (FAR) is impossible. The lower satellites are usually noisier, and their weights are already very low if the elevation weighting is applied in the stochastic model. In this case, removing lower satellites or noisier satellites may not contribute much to the precision of the float ambiguities.

2. Frequency level

The strategies in this level first make widelane ambiguities, which are easier to resolve due to their longer wavelength. If the widelane and narrow lane ambiguities as a full set can not be resolved, only the widelane ambiguities will be resolved as a subset [4, 5]. The disadvantage is that the widelane ambiguities can not contribute much to the fix solution if narrow lane ambiguities are not resolved.

3. Ambiguity level

The strategies in this level choose directly subsets of ambiguities. They either choose subsets of the original ambiguities or the decorrelated ambiguities [6, 7]. In order to reach a minimum success rate threshold, the Success Rate Criterion (SRC) strategy [6] removes the least precise conditional float ambiguities which were ordered during the Z-transformation of Least-squares AMBiguity Decorrelation Adjustment (LAMBDA) [8, 9]. The integrity related strategy [7] allows incorrectly fixed ambiguities, as long as the integrity requirement is still met. The strategy uses a routine to build the subsets being considered, which is time consuming. In order to reduce the influence of the bias among the float ambiguities, the partial decorrelation strategy [10, 11] does partial Z-transformation of the float ambiguities.

The two categories are:

1. Model driven

In this category, the subsets are chosen without taking any measurement. They choose the subsets based on either the variance-covariance (vc-) matrix of the float

ambiguity or satellite elevation. The strategies that fall in this category include the SRC strategy [6], elevation ordering strategy [1, 2], ADOP ordering strategy [3], and integrity related strategy [7].

2. Data driven

In this category, the subsets are chosen with measurements being involved. The strategies that fall in this category include the SNR ordering strategy [3], consistency checking strategy [4, 12–15]. The data driven strategies apply validation steps after the ambiguities are fixed, where the Fix Failure rate Ratio Test (FFRT) [16] or consistency checking are used as validation methods. For instance, the widelane and consistency checking strategy [4] works as follows. After the widelane ambiguities are resolved, they are tested by FFRT. If the best candidate does not pass the ratio test, the ambiguities that have the same values in the best and second best candidates will be collected as a subset. We refer this process as consistency checking strategy, which were implemented in many different ways [13–15]. This strategy is mostly based on experience, and therefore has less control on the quality of the fixed ambiguities.

The model driven strategies can select the subsets before the time-consuming integer search, and guarantee a high success rate. However, without the validation step, the incorrectly fixed ambiguities could be accepted as correct ones, which would introduce bias to the baseline solution.

The validation step applied by the data driven strategies will largely avoid the incorrectly fixed ambiguities. The probability of accepting the incorrectly ambiguities can be controlled by the failure rate of FFRT. However, the existing data driven strategies do not consider the rejected ambiguities, which may be correctly fixed if a smaller subset is selected.

In order to take better advantage of both two types of strategies, we propose a model and data driven strategy MD-PAR, which also considers the rejected ambiguities by the FFRT. The MD-PAR, as its name suggests, includes two part: the model driven part, and the data driven part. Here we use the classic strategy SRC as the model driven part, and the ratio test as the data driven validation step. Two different implementations are used to select smaller subsets from the rejected ambiguities by FFRT.

Moreover, the strategy should be independent of system(s), and applicable to single- or multi-frequencies. Thus we do not take into account the widelane strategies. The SRC strategy, as the classic model driven strategies, and the consistency checking strategy, as a data driven strategy, will be considered in this study.

The structure of this paper is as follows. Section 24.2 describes the model driven part and the data driven part of MD-PAR. Section 24.2.1 describes the SRC strategy. Section 24.2.2 briefly describes the fix failure rate ratio test. Section 24.2.3 describes the two implementations of MD-PAR, which are Two-step SRC strategy and SRC with consistency checking strategy. Section 24.3 presents the result of a simulation study on MD-PAR strategy, and compared with the classic strategy. Section 24.4 concludes the main results and presents the future work.

24.2 Methodology

The MD-PAR has model driven part and data driven part. The model driven part uses SRC strategy; the data driven part uses FFRT validation and two different implementations to select the smaller subsets. SRC strategy, FFRT validation and the two implementations are described as follows.

24.2.1 SRC

A GNSS observation model can be put in the following linearized equation

$$y = Aa + Bb + e, \tag{24.1}$$

where y is the vector of code and carrier observations; a is the vector of unknown ambiguities; b is the vector of baseline components, which possibly include residual atmospheric delays; e is the vector of measurement noise, which is assumed to have a zero-mean Gaussian normal distribution; A and B are the design matrices for the ambiguity and baseline components respectively. As a result of Best Linear Unbiased Estimate (BLUE), the real valued estimates of a, b and their vc-matrices are

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}, \begin{bmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{bmatrix}. \tag{24.2}$$

Normally, LAMBDA transformation is used to decorrelate the float ambiguities, which can be written as

$$\hat{z} = Z^T \hat{a}, \tag{24.3}$$

where Z^T needs to fulfill three conditions [17]: (1). its entries are all integers; (2). it should be invertible; (3). the entries of its inverse matrix are all integers. The construction of Z^T can be found in [8]. The SRC strategy chooses to fix a subset of $\hat{z}$, such that a minimum success rate threshold is still obtained, based on evaluating the integer bootstrapping success rate [6]. The subset of the ambiguities is denoted as $\hat{z}_P$, which is selected as follows.

The vc-matrix of the decorrelated ambiguities $\hat{z}$ is decomposed as

$$Q_{\hat{z}\hat{z}} = L^T D L. \tag{24.4}$$

The entries of L and D satisfy:

$$\begin{aligned} l_{j,i} &= \sigma_{\hat{z}_i \hat{z}_{j|J}} \sigma_{\hat{z}_{j|J}}^{-2}; \\ d_j &= \sigma_{\hat{z}_{j|J}}^2, \end{aligned} \tag{24.5}$$

where $\sigma_{\hat{z}_{j|J}}$ represents the conditional standard deviation of $\hat{z}_j$, conditioned on the set of fixed ambiguities $J = \{j+1, j+2, \ldots, n\}$; n is the total number of ambiguities, and $\sigma_{\hat{z}_{n|n}} = \sigma_{\hat{z}_n}$ is of the most precise float ambiguity. The entries of D are sequentially ordered after each decorrelation step such that $\sigma_{\hat{z}_{i|J}} \geq \sigma_{\hat{z}_{k|J}}$, for any $i < k \leq j$.

The integer bootstrapping success rate is calculated as

$$P_{s|i} = \prod_{j=i}^{n} \left(2\Phi \left(\frac{1}{2\sigma_{\hat{z}_{j|J}}} \right) - 1 \right), \tag{24.6}$$

where $\Phi(x)$ is the cumulative distribution function of standard Gaussian normal distribution. By setting a criterion of the success rate, we can select a subset of float ambiguities. In practice, however, we usually set a criterion of failure rate P_0, instead of a criterion of success rate P_{s0} (i.e., $P_{s0} + P_0 = 1$). Correspondingly, this failure rate criterion is a maximum allowable failure rate. In the remaining of this paper, we will use failure rate criterion instead of success rate criterion.

The SRC strategy is described in Fig. 24.1. First the LAMBDA Z-transformation is applied, such that the conditional standard deviations are known, see Eqs. (24.4) and (24.5). Note that the LAMBDA Z-transformation is implemented such that the ambiguities are automatically ordered according to their conditional variance, in this case with $\hat{z}_n$ being the most precise ambiguity, and $\hat{z}_{j|J}$ the most precise ambiguity conditioned on the ambiguities $J = \{j+1, j+2, \ldots, n\}$.

Next, we calculate the bootstrapping success rate of the full set, if the success rate is less than the success rate criterion, then remove the least precise ambiguity and calculate the success rate. Repeat this process until the success rate of a subset is larger or equal to the success rate criterion, or end up with a null subset. After $\hat{z}_P$ is fixed to $\breve{z}_P$, it is then used to get a fixed baseline solution

$$\breve{b}_P = \hat{b} - Q_{\hat{b}\hat{z}_P} Q_{\hat{z}_P \hat{z}_P}^{-1} (\hat{z}_P - \breve{z}_P), \tag{24.7}$$

with $Q_{\hat{b}\hat{z}_P}$ and $Q_{\hat{z}_P \hat{z}_P}$ being calculated as

$$\begin{aligned} Q_{\hat{b}\hat{z}_P} &= Q_{\hat{b}\hat{a}} Z_P, \\ Q_{\hat{z}_P \hat{z}_P} &= Z_P^T Q_{\hat{a}\hat{a}} Z_P, \end{aligned} \tag{24.8}$$

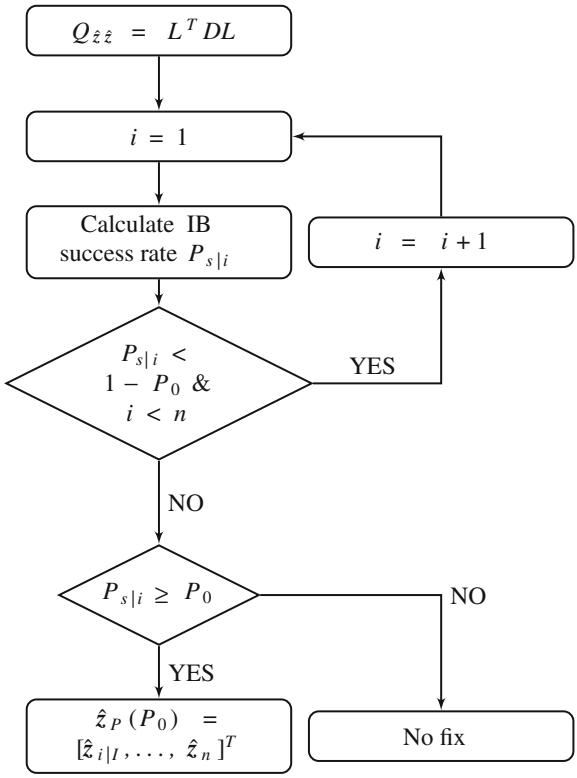


Fig. 24.1 The flow diagram of SRC strategy, with n the size of full set, P_0 the maximum allowable failure rate

where Z_P contains the corresponding columns of the transformation matrix Z, with Z_P satisfies

$$\hat{z}_P = Z_P^T \hat{a}. \tag{24.9}$$

If the fixed ambiguities can be considered deterministic (i.e. the success rate is very close to 1), the vc-matrix of the fixed baseline solution is obtained as

$$Q_{\check{b}_P\check{b}_P} = Q_{\hat{b}_P\hat{b}_P} - Q_{\hat{b}\hat{z}_P} Q^{-1}_{\hat{z}_P\hat{z}_P} Q_{\hat{z}_P\hat{b}}. \tag{24.10}$$

The advantages of SRC are threefold: (1) the real measurement is not needed during this process, since the subset can be determined using only the vc-matrix of the float ambiguities; (2) the Z-transformation is only performed once, which makes the strategy not very time consuming; (3) the subset is selected by bootstrapping success rate, but fixed by ILS, so the real success rate is higher than the criterion we set; thus, the FFRT with the failure rate equal to P_0 is obsolete, since real failure rate will be lower or equal to the failure rate criterion we set.

24.2.2 Fix Failure Rate Ratio Test

In the data driven strategy, once $\hat{z}_P$ is fixed to integer vector $\breve{z}_P$, an acceptance test is usually applied before fixing the baseline. Here we use the Fixed Failure rate Ratio Test (FFRT), which is defined as:

$$\text{accept } \breve{z}_P, \text{ if } RT = \frac{||\hat{z}_P - \breve{z}_P||^2_{Q_{\hat{z}_P\hat{z}_P}}}{||\hat{z}_P - \breve{z}'_P||^2_{Q_{\hat{z}_P\hat{z}_P}}} \leq \mu, \tag{24.11}$$

where $||.||^2_Q$ is weighted squared norm, i.e., $||x||^2_Q = x^T Q^{-1} x$; $\breve{z}_P$, $\breve{z}'_P$ are the candidates that have the shortest and second shortest weighted distances to $\hat{z}_P$; μ is a critical value selected by a given failure rate P_0. The value of μ given P_0 can be looked up from a table generated with Monte Carlo simulation by [16].

24.2.3 Implementations of Smaller Subsets Selection

If the fixed ambiguities is not accepted by FFRT, a smaller subset is selected. Two possible implementations are presented here, based on the criteria that they should be generally applicable and computationally efficient. Moreover, it is expected that they can provide a significantly small RMS of baseline, as will be validated in Sect. 24.3.

24.2.3.1 Two-Step SRC With Ratio Test

Assuming that the maximum allowable failure rate is equal to P_0, it is explained in [16] that the FFRT can then be applied as an acceptance test such that this requirement is met. However, in case of a weak model this may lead to a very low acceptance probability (fix rate) if we try to fix the full set. On the other hand, when applying the SRC strategy with P_0, the subset might become very small, which does not provide a small RMS of the baseline solution. Therefore, it is proposed to start with a subset selected with the SRC strategy, but with a somewhat larger maximum allowable failure rate, $P'_0 > P_0$. This subset will be larger or equal to the one obtained with P_0. In order to still guarantee a failure rate of P_0, the FFRT with P_0 is applied. The advantage is that there can be a considerable acceptance probability for this larger subset without compromising on the failure rate requirement. At the same time, there is also a probability that this subset is rejected. In that case, a smaller subset is selected by applying SRC with P_0. The two step procedure will be called TSRC. The flow diagram of this strategy is shown in Fig. 24.2.

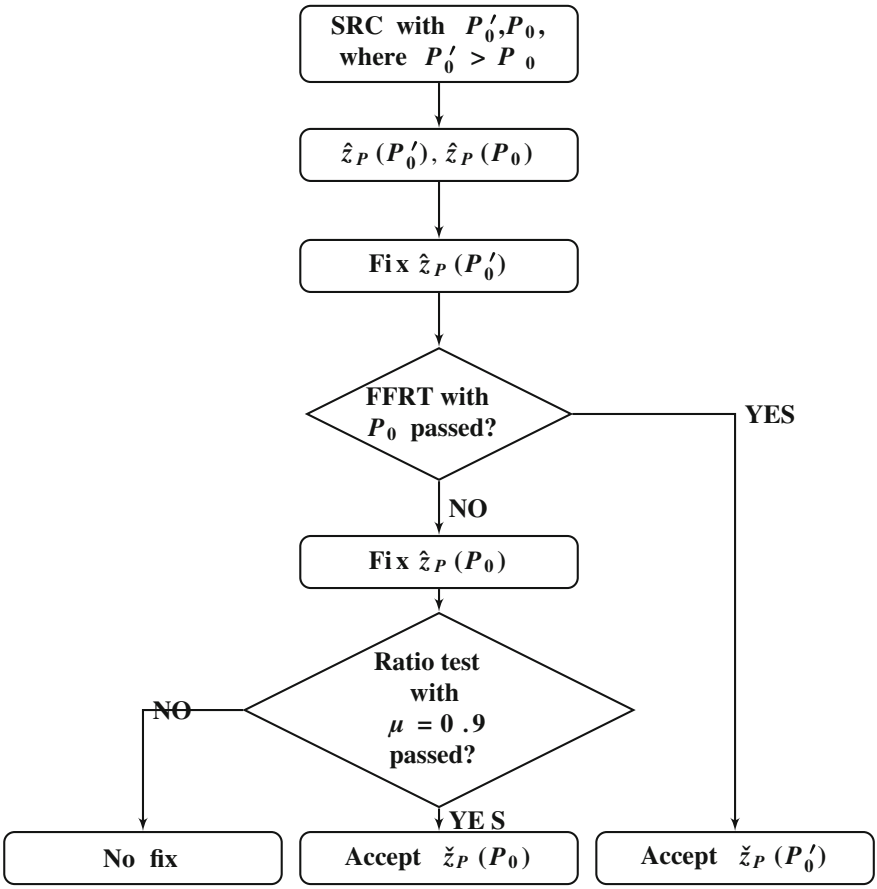


Fig. 24.2 The flow diagram of the TSRC strategy, with P_0', P_0 the initial and the maximum allowable failure rates

In summary, the idea is to start with a larger initial subset, and if the corresponding fixed subset is rejected with FFRT, a second smaller subset is fixed by applying SRC with P_0. However, a problem with this approach is that the actual failure rate of this two-step procedure can be larger than P_0. This is because the float solutions for which the second subset is fixed are conditional float solutions: conditioned on being rejected by the FFRT applied to initial subsets, and therefore are not normally distributed. Figure 24.3 shows an example for an initial subset with size 2. This initial subset is normally distributed; the simulated samples in the figure are colored to indicate whether they are accepted (dark) or rejected (light grey). In the latter case, the smaller subset in the next step of TSRC contains only the second ambiguity (along vertical axis), which has a distribution as shown on the left of the figure.

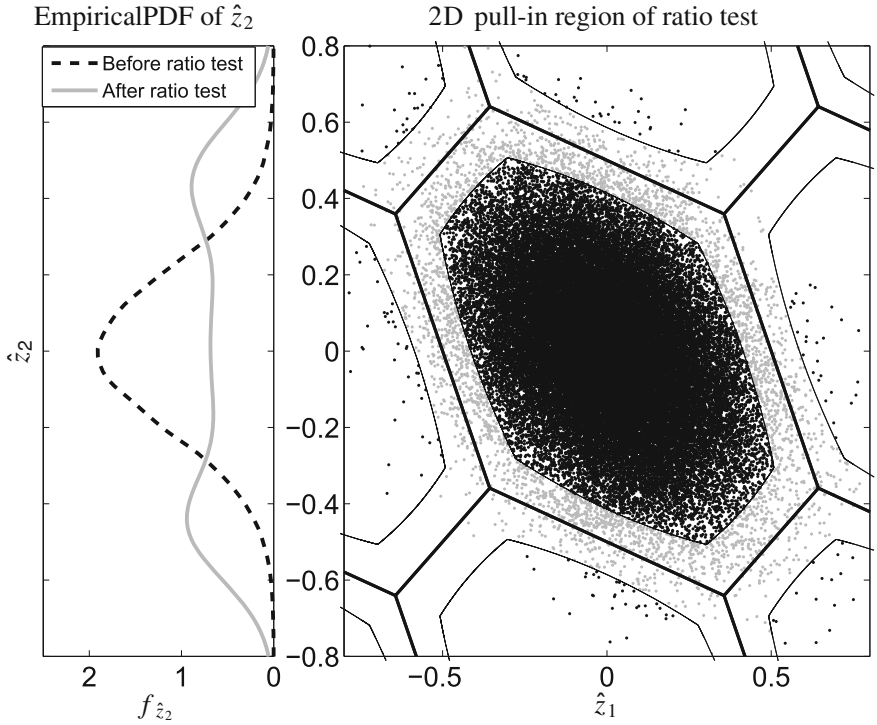


Fig. 24.3 The two dimensional example of pull-in region. In the *right panel*, the dark area is the acceptance region by FFRT, the *grey* area is the rejection region. In the *left panel*, the *dotted curve* represents the empirical probability density function (PDF) of the float ambiguity $\hat{z}_2$, the *solid curve* represents the conditional PDF of the float ambiguity $\hat{z}_2$ being rejected by FFRT

The larger the initial subset, the larger the actual failure rate is expected to be. This means that P_0' should not be chosen too large. Since it can be seen in Fig. 24.3 that for the second subset a lot of probability mass is close to the boundaries of the pull-in region, the SRC strategy in this second step is combined with a ratio test with a high constant critical value of 0.9 (based on empirical results), such that the float solutions close to the boundary of pull-in region are rejected. Recall that applying the FFRT does not make sense, as it will always be accepted as we mentioned in Sect. 24.2.1.

The choice of P_0' is a trade-off between a lower failure rate and a bigger subset. In Sect. 24.3 we will investigate what will be a good choice for P_0'.

24.2.3.2 SRC With Consistency Checking

The SRC with Consistency checking (SRC-C) strategy first selects an initial subset by SRC with a given failure rate P_0', then the ILS result of the subset is tested by FFRT. If FFRT rejects the subset, then we select the entries that are identical in the

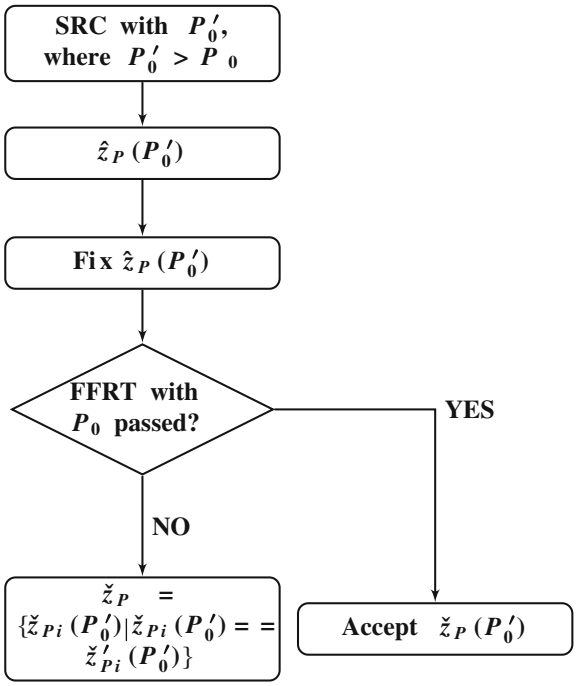


Fig. 24.4 The flow diagram of the SRC-C strategy, with P_0', P_0 the initial and the second given failure rates. $\breve{z}_{Pi}(P_0')$ and $\breve{z}'_{Pi}(P_0')$ are respectively the ith entries of the best and second best candidates $\breve{z}_P(P_0')$ and $\breve{z}'_P(P_0')$

best and second best candidates as the subset. This selection is referred to as consistency checking in this study. The consistency checking was used in [4] to resolve widelane ambiguities. But here we use it to resolve the LAMBDA transformed ambiguities. Comparing with TSRC strategy, SRC-C does ILS search only once, and gives comparable failure rate and subset size as TSRC. The flow diagram of this strategy is shown in Fig. 24.4.

24.3 Numerical Experiment

In order to compare the performance of the two implementations of MD-PAR and SRC, a numerical experiment was carried out at Perth, Australia, in 22nd of November, 2013. We simulated 24 h dual-frequency measurements of GPS and BDS, with the interval of adjacent epoch being 30 min, and the number of samples at each epoch being 10^5. The number of visible satellites was around 15–22 with a cutoff angle of 10°. The single-epoch AR was considered, since the main contribution of PAR in case of weaker models should be faster convergence of the baseline solution. The ionosphere weighted model [18, 19] and troposphere float model were used in double differencing (DD) observation. The elevation dependent weighting [20, 21] was applied in the stochastic model. The setup of the

Table 24.1 Setup of the simulation model

Date	Location	System	Cutoff	σ_ϕ (m)	σ_ρ (m)	σ_i (m)	Samples
22-11-2013	lat = −32.004° lon = 115.895° h = 24.003 m	GPS (L1, L2) BDS (B1, B2) GPS + BDS	10	0.003	0.1	0.01–0.1	10^5

model is shown in Table 24.1. The standard deviations of the phase (σ_ϕ), code (σ_ρ), and ionosphere corrections (σ_i) are the values for zenith directions, and are considered as the same for GPS and BDS. In each epoch, we considered GPS only, BDS only and GPS + BDS systems, and different standard deviations of ionosphere delay. As an example, the simulation with $\sigma_i = 0.02$ m is presented. We compared the three strategies introduced in the previous sections, by means of empirical fix rate, empirical failure rate, empirical subset size, and RMS of baseline in the subsections as follows.

24.3.1 Fix Rate and Failure Rate

For each epoch, the empirical fix rate P_{fix} and failure rate P_f are calculated as

$$\begin{aligned} P_{fix} &= \frac{N_{fix}}{N} = \frac{N_s + N_f}{N}, \\ P_f &= \frac{N_f}{N}, \\ N &= N_s + N_f + N_u, \end{aligned} \tag{24.12}$$

where N is the number of samples in each epoch, N_s, N_f and N_u are the numbers of correctly fixed samples, incorrectly fixed samples and unfixed samples.

For the specific scenario with $P_0' = 0.01, P_0 = 0.001$, the fix rate and failure rate are shown in Fig. 24.5. The upper panels show the fix rate, and the lower panels show the failure rate; the left, middle and right panels show the simulation scenarios of GPS only, BDS only and GPS + BDS systems. As shown in the figure, the fix rate of all three strategies is higher than 0.999, the failure rate of all three strategies is mostly below P_0. The failure rate of SRC is smaller than that of TSRC and SRC-C, which is expectable, since we set a higher initial failure rate for the other two strategies, i.e., $P_0' > P_0$. In some epochs, the failure rate of TSRC and SRC-C are slightly higher (max. 4×10^{-4}) than P_0, which shows that the failure rate of TSRC and SRC-C is not strictly controlled. However, with a sufficiently low initial failure rate P_0', the empirical failure rate of TSRC and SRC-C will be close to the required P_0.

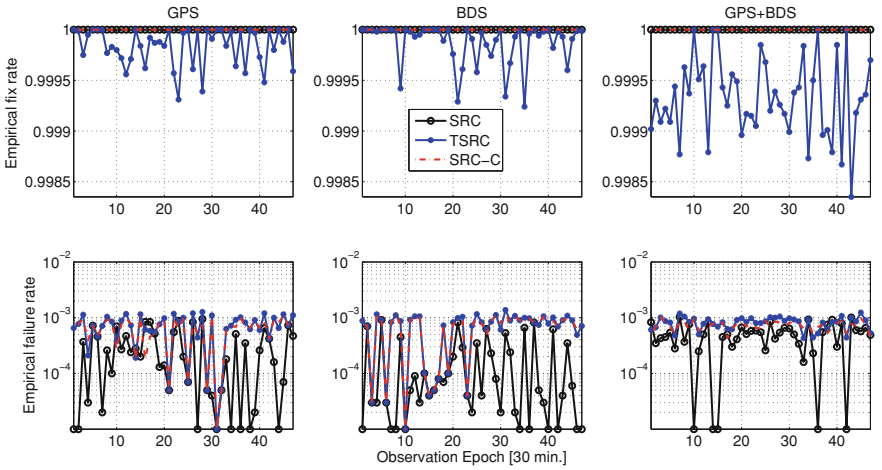


Fig. 24.5 Fix rate and failure rate of the three strategies. The *upper panels* show the fix rate, the *lower panels* show the failure rate. The *left*, *middle* and *right panels* show the results of GPS only, BDS only and GPS + BDS systems

From the above result we conclude that the failure rate of TSRC and SRC-C strategies will be sufficiently close to the required failure rate P_0, if we set a sufficiently low initial failure rate P_0' for them.

24.3.2 Subset Size

We know that the bigger the subset size is, the more contribution is expected to the baseline precision. To compare the subset sizes selected by the three strategies at each epoch, we calculate the subset size ratio, which is

$$\bar{r}_s = \frac{1}{N_{fix}} \sum_{k=1}^{N_{fix}} \frac{ns(k)}{n}, \tag{24.13}$$

where $ns(k)$ is the subset size of the kth sample, and n is the full set size. Figure 24.6 shows the full set sizes and the subset size ratios given by the three strategies in each epoch. As expected, the subset sizes given by TSRC and SRC-C are larger than the subset sizes given by SRC. This is because we set the failure rate P_0' for TSRC and SRC-C higher than P_0. But the difference is marginal, because P_0' for TSRC and SRC-C is close to P_0, which already obtains rather small subsets initially.

The above result shows that the subset sizes given by TSRC and SRC-C are larger than that given by SRC, but the difference will be marginal if the P_0' is too low.

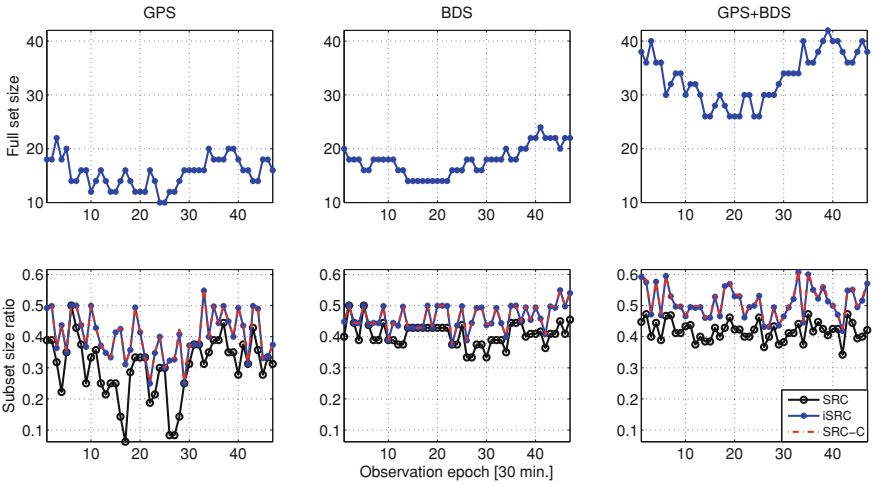


Fig. 24.6 Subset sizes given by the three strategies. The *upper panels* show the full set size, the *lower panels* show the subset size ratio

24.3.3 RMS of Baseline Solution

The mean RMS of baseline solution over all samples is calculated as

$$\begin{aligned} \overline{RMS} &= \frac{1}{N_{fix}} \sum_{k=1}^{N_{fix}} RMS(k), \\ RMS(k) &= \sqrt{\sigma_x^2(k) + \sigma_y^2(k) + \sigma_z^2(k)} \end{aligned} \tag{24.14}$$

where $\sigma_x^2(k)$, $\sigma_y^2(k)$ and $\sigma_z^2(k)$ are the variance of baseline coordinates taken from $Q_{\check{b}_P\check{b}_P}$ of the kth sample.

The mean RMS of the baseline given by the float solution and the three PAR strategies are shown in Fig. 24.7. Accordingly with the subset sizes, the mean RMS of the baseline given by the three PAR strategies is smaller than that given by the float solution. Furthermore, TSRC and SRC-C have even smaller mean RMS than SRC. Especially when SRC gives a very small subset or null subset, TSRC and SRC-C can give a considerable large subset, and thereby a better improvement of baseline precision. This is one of the main advantage of MD-PAR, that when SRC is not applicable, we fix a smaller subset instead of fixing nothing. For instance in Figs. 24.6 and 24.7, in the GPS only system, at epoch 26 and 27, the subset sizes given by SRC are quite small, and the mean RMS given by SRC overlay with that of the float solution, while the subset sizes given by TSRC and SRC-C are large and the mean RMS given by TSRC and SRC-C is smaller than that of the float solution.

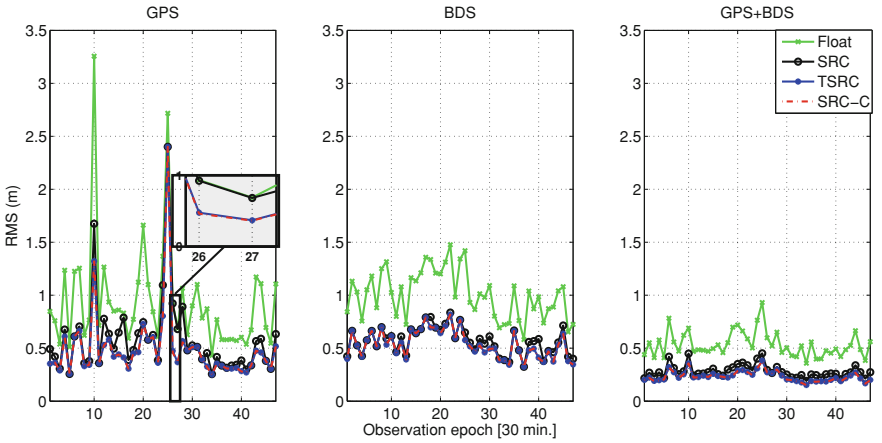


Fig. 24.7 The mean RMS of baseline given by the float solution and the three PAR strategies

Table 24.2 The 95 % upper limit of the baseline errors ϵ [m], i.e., $P(|b_e - b| < \in) = 95\ \%$

ϵ (m)	Float	SRC	TSRC	SRC-C
GPS	3.3	1.5	1.4	1.4
BDS	2.3	1.3	1.3	1.3
GPS+BDS	1.4	0.7	0.6	0.6

The above result shows that the mean RMS given by TSRC and SRC-C is smaller than that given by SRC, especially when SRC fixes small or null subsets.

In addition, the error of the baseline length at the first epoch is shown here as a direct impression of RMS. The error of the baseline length is calculated as

$$|b_e - b| = \sqrt{(b_e - b)^T (b_e - b)}, \tag{24.15}$$

where b is the truth of baseline coordinate vector, b_e is the estimated baseline coordinate vector. b_e represents the fixed solutions given by SRC, TSRC, SRC-C, or the float solution.

For each strategy, the 95 % confidence level (denoted as ϵ in the table) of the baseline error is shown in Table 24.2, and the probability of the baseline error that are less than 1 m is shown in Table 24.3. Comparing the rows of Table 24.2, we found GPS + BDS system has smaller ϵ than the single systems. PAR has most significant improvement (around 1 m) in the single GNSS systems. Comparing the columns of Table 24.2, we found that all three PAR strategies have smaller ϵ than the float solution, and TSRC and SRC-C have even smaller ones. In very few cases (less than 10^{-4}), the errors given by the three PAR strategies are larger than the float solution, which is controlled by the failure rate we set.

Table 24.3 The probability of the baseline errors less than 1 m

$P(\lvert b_e - b\rvert < 1\text{ m})$ (%)	Float	SRC	TSRC	SRC-C
GPS	44.9	82.0	82.9	82.8
BDS	59.1	86.4	88.1	88.0
GPS+BDS	83.8	99.6	99.8	99.8

As shown in Table 24.3, PAR has larger probabilities of small baseline error than the float solution. Moreover, TSRC and SRC-C have better performance on reducing the baseline errors than SRC.

Although only single-epoch results are shown here, the results in Tables 24.2 and 24.3 suggest that the PAR strategies can significantly reduce the convergence time of the baseline solution.

As mentioned in Sect. 24.2.1, if we fix a subset of Z-transformed ambiguities, the original ambiguities will be real valued. Thus, in multi-epoch case, we do not fix and hold the original ambiguities. Instead, we use Kalman filter to estimate only the float original ambiguities and float baseline solution. If PAR is applied, the precision of the float original ambiguities and the float baseline will be higher than where PAR is not applied, as shown in Tables 24.2 and 24.3. These float values are used in the next epoch as the initial state, which will make the convergence time shorter.

The above result shows that in the single-epoch case, the baseline error can be significantly reduced by PAR. Moreover, TSRC and SRC-C reduced the baseline error more significantly than SRC. Consequently we can expect a shorter convergence time of the baseline solution from PAR, in the multi-epoch case.

24.3.4 Choose a Better Initial Failure Rate P_0'

By the above comparisons, we have a brief conclusion that with the setup parameters $P_0' = 0.01,\ P_0 = 0.001$, the empirical failure rate of all three PAR strategies is below the threshold we set. The subset sizes given by the TSRC and SRC-C strategies are larger than that given by SRC, but the difference is marginal. As we mentioned before, this is because P_0' is small, which gives small subsets to TSRC and SRC-C initially. If we raise up P_0', however, the empirical failure rate of TSRC and SRC-C are expected to be high too. Thus the choice of P_0' is a trade-off between a lower failure rate and a bigger subset. In order to choose a proper P_0', we simulated with different values of P_0', and compared the results by average failure rate $P_{f,ave}$ over all epochs, and the average subset size ratio over all epochs, which are calculated as

$$\begin{aligned} P_{f,ave} &= \frac{1}{nep}\sum_{ep=1}^{nep} P_f, \\ \bar{r}_{s,ave} &= \frac{1}{nep}\sum_{ep=1}^{nep} \bar{r}_s, \end{aligned} \tag{24.16}$$

Table 24.4 The average empirical failure rate over all epochs, and the average subset size ratio over all epochs

		$P_{f,ave}(\times 10^{-3})$			$\bar{r}_{s,ave}$		
		SRC	TSRC	SRC-C	SRC	TSRC	SRC-C
GPS	$P_0' = 0.010$	0.3	0.7	0.7	0.30	0.40	0.40
	$P_0' = 0.025$	0.3	0.8	0.9	0.30	0.43	0.44
	$P_0' = 0.050$	0.3	0.8	1.7	0.30	0.42	0.47
BDS	$P_0' = 0.010$	0.2	0.7	0.7	0.41	0.46	0.46
	$P_0' = 0.025$	0.2	0.8	1.0	0.41	0.50	0.51
	$P_0' = 0.050$	0.2	0.8	2.3	0.41	0.50	0.55
GPS + BDS	$P_0' = 0.010$	0.5	0.8	0.7	0.42	0.51	0.51
	$P_0' = 0.025$	0.5	0.7	1.1	0.42	0.57	0.59
	$P_0' = 0.050$	0.5	0.7	2.9	0.42	0.59	0.68

The maximum allowable failure rate is set to 1×10^{-3}

where *nep* is the number of simulated epochs.

Comparing the columns in Table 24.4, we can see that for TSRC and SRC-C, the larger the P_0' we set, the higher the average failure rate $P_{f,ave}$ is, and the larger the subsets they obtain. With $P_0' = 0.01$, and $P_0' = 0.025$, $P_{f,ave}$ is always below 0.001, whereas with $P_0' = 0.05$, $P_{f,ave}$, $P_{f,ave}$ of SRC-C is higher than 0.001. The reason is that SRC-C does not give control over the failure rate. Thus we need to carefully choose the value of P_0' for SRC-C strategy. We can conclude that SRC-C strategy works only when the initial subset is selected with a sufficiently low failure rate. For the three different P_0's considered here, the best choice is $P_0' = 0.025$, since the three strategies have both sufficiently low failure rate and big subsets in this case. Besides, SRC-C strategies obtain a bit larger subsets than TSRC, but TSRC has more control over the failure rate than SRC-C.

Different scenarios with $\sigma_i = 0.01$ to 0.1 m were also simulated, and the relative performance of the three strategies with different values of σ_i is similar. Due to space limitation, we presented only the scenario with $\sigma_i = 0.02$ m. However, the conclusion is general when comparing the PAR strategies.

24.4 Conclusion

In this paper, we proposed the MD-PAR strategy and two implementations: TSRC and SRC-C, and compared their performance with the classic model driven PAR strategy SRC by a simulation. The empirical fix rate, empirical success rate, RMS of baseline solution, and the probability of baseline errors are used as the criteria to compare their performances. The analysis results and the simulation results both suggest that the model and data driven strategies TSRC and SRC-C have better performances than the model driven strategy SRC. They give larger subsets, and thereby smaller RMS of baseline solution, and smaller baseline errors than the

SRC strategy. The TSRC strategy has more control over the final failure rate, while the SRC-C strategy uses less time to get the fixed subset. The value of the initial failure rate used in TSRC and SRC-C, however, should be carefully chosen such that SRC-C strategy will not give too large final failure rate.

In the future, experimental validation with real data will be done. Also multi-epoch processing will be considered.

Acknowledgements The authors would like to acknowledge China Scholarship Council (CSC) for supporting the first author's Ph.D. studies at the Delft University of Technology, Netherlands, and thank for the comments from Prof. Peter Teunissen, and the discussion with Lei Wang and Dr. Zishen Li.

References

1. Mowlam A (2004) Baseline precision results using triple frequency partial ambiguity sets. In: Procedings of ION GNSS-2004, Long Beach CA, pp 2509–2518
2. Takasu T, Yasuda A (2010) Kalman-filter-based integer ambiguity resolution strategy for long-baseline RTK with ionosphere and troposphere estimation. In: Proceedings of ION GNSS 2010, Portland, Oregon, pp 161–171
3. Parkins A (2011) Increasing GNSS RTK availability with a new single-epoch batch partial ambiguity resolution algorithm. GPS Solutions 15(4):391–402
4. Dai L, Eslinger D, Sharpe T (2007) Innovative algorithms to improve long range RTK reliability and availability. In: Proceedings of ION NTM 2007, San Diego CA, pp 860–872
5. Yanming F, Li B (2008) Three carrier ambiguity resolution: generalised problems, models, methods and performance analysis using semi-generated triple frequency GPS data. In: Proceedings of ION GNSS 2008, Savannah GA, pp 2831–2840
6. Teunissen PJG, Joosten P, Tiberius CCJM (1999) Geometry-free ambiguity success rates in case of partial fixing. In: Proceedings of ION national technical meeting 1999 and 19th biennal guidance test symposium, San Diego CA, pp 201–207
7. Khanafseh S, Pervan B (2010) New approach for calculating position domain integrity risk for cycle resolution in carrier phase navigation systems. Trans Aerosp Electron Syst 46(1):296–307
8. Teunissen PJG (1993) Least-squares estimation of the integer GPS ambiguities. In: Invited lecture, section IV "Theory and Methodology", at the general meeting of the international association of Geodesy, Beijing, China, pp 1–16
9. Teunissen PJG (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geodesy 1–2:65–82
10. Henkel P (2009) Geometry-free linear combinations for Galileo. Acta Astronaut 65(9–10):1487–1499
11. Henkel P, Günther C (2010) Partial integer decorrelation: optimum trade-off between variance reduction and bias amplification. J Geodesy 84(1):51–63
12. Dai L, Eslinger DJ, Sharpe RT, Hatch RR (2011) Partial search carrier-phase integer ambiguity resolution. U.S. Patent 7,961,143 B2
13. Vollath U, Doucet KD (2009) GNSS signal processing with partial fixing of ambiguities. U.S. patent 7,538,721 B2
14. Vollath U (2010) Generalized partial fixing. U.S. patent 2010/0253575 A1
15. Lawrence DG (2009) A new method for partial ambiguity resolution. In: Proceedings of ION ITM 2009, Anaheim CA, pp 652–663
16. Teunissen PJG, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. Survey Review 151(April):138–151

17. Teunissen PJG (1995) The invertible GPS ambiguity transformations. Manuscripta Geodaetica 20:489–497
18. Teunissen PJG (1997) The geometry-free GPS ambiguity search space with a weighted ionosphere. J Geodesy 71(6):370–383
19. Odijk D (2000) Stochastic modelling of the ionosphere for fast GPS ambiguity resolution. Geodesy Beyond 2000:387–392
20. Jin, X.X., Jong, C.D.d.: Relationship between satellite elevation and precision of GPS code observations. Journal of Navigation **49**(2) (1996) 253–265
21. Teunissen PJG (1997) GPS ambiguity resolution: impact of time correlation, cross-correlation and satellite elevation dependence. Stud Geophys Geod 41(2):181–195

Chapter 25
Prediction of Polar Motion Based on Combination of Weighted Least-Squares and Autoregressive Moving Average

Zhangzhen Sun, Tianhe Xu, Yijun Mo and Chao Xiong

Abstract High accurate prediction of polar motion has important significant and useful value for high-precision space navigation and positioning. In this paper, the weighted least-squares (WLS) is proposed to use to predict different span of polar motion, as the cycles and trends have the characteristic of time-varying in the observational data of polar motion. Autoregressive Moving Average (ARMA) can be applied to fit the residuals of polar motion as it can be regard as a smooth, zero-mean sequence. The LS + AR model, the LS + ARMA model, the WLS + AR model and the WLS + ARMA model are used to predict the different span of polar motion, the results show that, the application of weighted least-squares can improve the polar motion prediction accuracy effectively. And the WLS + ARMA model is equal to WLS + AR model, and in some days WLS + ARMA model is better than WLS + AR model.

Keywords Polar motion · Weighted least-squares · Autoregressive moving average · Prediction

Z. Sun (✉) · Y. Mo · C. Xiong
Aerors Inc, Xi'an, Shanxi, China
e-mail: sunzhangzhen@126.com

T. Xu
State Key Laboratory of Geo-information Engineering, Xi'an, Shanxi, China

T. Xu
State Key Laboratory of Astronautic and Dynamics, Xi'an, Shanxi, China

T. Xu
State Key Laboratory of Geodesy and Earth's Dynamics, Wuhan, Hubei, China

T. Xu
Xian Research Institute of Surveying and Mapping, Xi'an, Shanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_25,

25.1 Introduction

The Earth's rotation movement characterizes the situation of the whole earth movement, as well as the interaction between the Earth's various spheres of the Earth's core, mantle, crust and atmosphere [1]. It can be described by Earth Orientation Parameters (EOPs), which includes three parts: the Nutation-precession parameters, the Polar Motion (PM) parameters and the Universal Time (UT1-UTC) or the Length of day (LOD). The PM is an important parameter to characterize the movement of Earth, which is used to describe the instantaneous movement of the earth's rotation axis in the body and the PM on the earth surface is changing slowly. EOP is the necessary parameters to achieve mutual conversion of the celestial reference frame and earth reference frame, and it is very important for high-precision space navigation and positioning. Modern measurement techniques such as VLBI, SLR, GPS and DORIS can provide people with high precision and high spatial and temporal resolution of EOPs. However, due to the complexity of data processing, it is difficult to access to these parameters in real-time. In order to meet the needs of the space navigation and positioning, high-precision prediction for EOPs is urgent. For the prediction of PM, a large number of predict models have been proposed [2–8]. At present, there are two major prediction methods for PM: linear and nonlinear models. The linear models include: least square extrapolation (LS), combination of least square extrapolation and auto-regressive (LS + AR) [5–7], combination of least square extrapolation and auto-covariance (LS + AC) [9] et al. The nonlinear models include: Sequence of threshold auto-regressive model (TAR), Artificial Neural Network (ANN) [10–12], Fuzzy inference [13] and so on.

In the above methods, whether the independent method or the combined method, LS model should be performed based on the deterministic cycles and trends. In fact, they have the characteristic of time-varying in the observational data of the PM [14, 15]. The closer the observational data is near to the prediction point, the greater the impact on the prediction is. When using the least-squares to estimate the model parameters, it regards the influence of historical data without considering such time-varying characteristics. In other words, the same impact or contribution of the recent data and the past data on the prediction value is adopted, which is obviously unreasonable. Based on this, the weighted least-squares combination of AR model and the weighted least-squares combination of ARMA model is proposed to predict the PM in this paper. The results show that the weighted least-squares can intensify the impact of recent data and improve the prediction accuracy effectively; the WLS + ARMA model is equal to WLS + AR model in prediction accuracy. For x-component of PM, AR model a litter better in the near real-time prediction and ARMA model performs better in the long-team prediction. For y-component of PM, ARMA model a litter better in the near real-time prediction and AR model performs better in the long-team prediction.

25.2 Models

25.2.1 WLS Model

The weighted least-squares (WLS) method considers a weight matrix when estimated parameters based on LS. In this paper, the WLS with combination of autoregressive model (WLS + AR) or Autoregressive Moving Average model (WLS + ARMA) is proposed to predict PM. The weight element of the recent data should be chosen bigger in the weight matrix.

The LS is used to fit the series of PM considering the linear and periodic terms, which includes Chandler wobble, annual and semi-annual periodic terms. The fitting equation of LS model can be expressed as:

$$\begin{aligned} \mathrm{f(t)} = \mathrm{a}_0 + a_1 t + B_1 \cos(\frac{2\pi t}{R_1}) + B_2 \sin(\frac{2\pi t}{R_1}) + C_1 \cos(\frac{2\pi t}{R_2}) \\ + C_2 \sin(\frac{2\pi t}{R_2}) + D_1 \cos(\frac{2\pi t}{R_3}) + D_2 \sin(\frac{2\pi t}{R_3}) + \cdots \end{aligned} \tag{25.1}$$

where a_0 is the constant term, a_1 is the linear term, B_i, C_i and D_i are the coefficients for periodic terms, R_i is the corresponding periodic, t is the time of UTC. The estimator by using the WLS for parameter solution can be written as:

$$X = (B^T P B)^{-1} B^T P L \tag{25.2}$$

where

$$\boldsymbol{X} = [\ a_0 \quad a_1 \quad B_1 \quad B_2 \quad C_1 \quad C_2 \quad D_1 \quad D_2 \quad \cdots\ ] \tag{25.3}$$

X is the estimated parameter vector; B is the coefficient matrix, L is observation vector of PM time series, P is the weight matrix. A diagonal matrix is considered for P in this paper. In order to strength the characteristic of time-varying in the observational data of the PM according to the principle that the closer the observational data is near to the prediction point, the greater the impact on the prediction is, the weight matrix P can be expressed as:

$$\boldsymbol{P} = \begin{bmatrix} P_1 & & \\ & P_2 & \\ & & P_3 \end{bmatrix}^e \tag{25.4}$$

where P_1, P_2 and P_3 are the diagonal matrix with diagonal elements as $\frac{1}{3}, \frac{1}{2}, 1$ respectively, and e is the power.

25.2.2 AR Model

AR model is the description of the relationship between a random series $z_t(t = 1, 2, \cdots, N)$ before time t and the white noise of time t. Its expression can be written as following:

$$z_t = \sum_{i=1}^{p} \phi_i z_{t-i} + a_t \tag{25.5}$$

where $\phi_1, \phi_2, \cdots, \phi_p$ are the autoregressive coefficients, and a_t is the white noise with zero mean, p is the model order. The above equation denoted by AR (p) is well-known as the AR model of the order p.

The AR model requires the time series stationary random, which should meet the steady, normal and zero mean conditions. The key of using the AR model is how to determine the order p. Usually there are three methods for the determination of p, the final prediction error criterion, the information criterion and the delivery function criterion. In practice, these 3 methods are almost equivalent. In this paper, the final prediction error criterion is adopted.

The final prediction error criterion can be expressed as:

$$\text{FPE(p)} = P_M(n + p + 1)/(n - p - 1) \tag{25.6}$$

where

$$P_M = \frac{1}{n-p} \sum_{t=p+1}^{n} (z_t - \sum_{j=1}^{p} \phi_j z_{t-j})^2 \tag{25.7}$$

The value p that minimums the FPE (p) is chosen as the order of AR model. The model coefficients $\phi_1, \phi_2, \cdots, \phi_p$ can be determined by solving the Yule-Walker equations.

25.2.3 ARMA Model

ARMA is the combination of Autoregressive model and moving average model. It can be expressed as:

$$z_t = \sum_{i=1}^{p} \phi_i z_{t-i} + a_t + \sum_{j=1}^{q} \theta_j a_{t-j} \tag{25.8}$$

where: $\phi_1, \phi_2, \cdots, \phi_p$ are the autoregressive coefficients, $a_t, a_{t-1}, \cdots, a_{t-q}$ are the white noises, $\theta_1, \theta_2, \cdots, \theta_p$ are the Moving Average coefficients, p is the

autoregressive model order, q is the moving average model order. p and q can be determinate by partial autocorrelation coefficients and correlation coefficients.

25.2.4 Error Analysis

In order to evaluate the prediction accuracy, we use the mean absolute error (MAE) standards. It can be expressed as following:

$$E_i = P_i - O_i \tag{25.9}$$

$$MAE_j = \frac{1}{n}\sum_{i=1}^{n}(|E_i|) \tag{25.10}$$

where P_i is the predicted value of the i-th prediction, O_i is the corresponding observation value, E_i is real error (Assumed the observation value as true value), n is the total prediction number.

25.3 Calculation and Analysis

25.3.1 Data Description

In this paper, the time series EOP 08C04 file provided by the IERS is used for calculation and analysis (http://hpiers.obspm.fr/eoppc/eop/). The sampling interval of these data is one day. The time series contains from January 1962 (37665 MJD) to now. The PM, the LOD and the UT1-UTC et al. are included. Firstly, the LS or WLS is used to fit the PM observations according to Eqs. (25.1) and (25.2) for obtaining the model parameters and removing the cycle and trend terms. Then, AR model or ARMA model is used to predict the residuals with combination of LS or WLS extrapolation to generate PM prediction values.

25.3.2 Data Processing and Analysis

In order to test the validation of the proposed method 4 cases of computation scheme are applied to predict PM as followings:

Case 1: Prediction of PM based on LS+AR model;
Case 2: Prediction of PM based on LS+ARMA model;
Case 3: Prediction of PM based on WLS+AR model;
Case 4: Prediction of PM based on WLS+ARMA model.

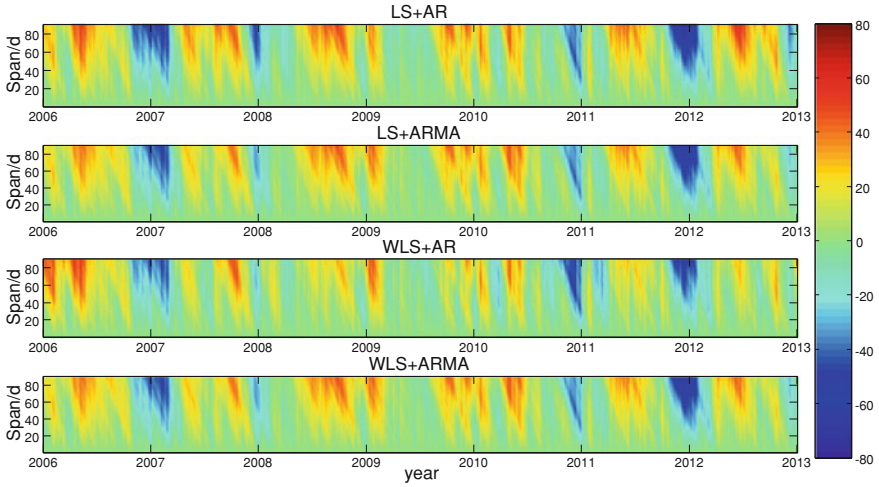


Fig. 25.1 Prediction errors for x-component of PM (*unit* mas)

The selection of basic sequence length of the PM has a great impact on the prediction accuracy. Research shows that it is high accuracy can be obtained when the basic sequence length of the polar motion is chosen as 10-year based on LS + AR model [16]. Therefore, the basic sequence length of least squares and weighted least squares is 10-year in this paper. Forecast accuracy statistics from January 1, 2006 to January 1, 2013, with the span of 1–90 days of the forecast are calculated and compared. The optimal weighted power e can be dynamically determined based on the principle of minimum prediction error. The true error of 4 cases is shown in Figs. 25.1, 25.2.

From Figs. 25.1, 25.2, we can see that variation of the prediction error shows almost the same trends based on 4 cases. The prediction error of WLS + AR model or WLS + ARMA model is lower than those of to LS +A R model and LS + ARMA model. Table 25.1 shows statistical prediction accuracy of each prediction span. From Table 25.1 we can see that the weighted least squares can improves the prediction accuracy of PM effectively.

Figure 25.3 shows the MAE of 4 schemes for different prediction span. From Fig. 25.3 and Table 25.1, we can see that the LS + ARMA model has higher prediction accuracy than LS + AR model, especially in x-component of PM. The WLS + AR model and WLS + ARMA model almost have the equal prediction accuracy. For x-component of PM, the WLS + AR model is a little better in the near real-time prediction, and in long-term prediction, the WLS + ARMA model is better. For y-component of PM, the WLS + ARMA model performs better in the near real-time prediction, and in long-term prediction, the WLS + AR model is better.

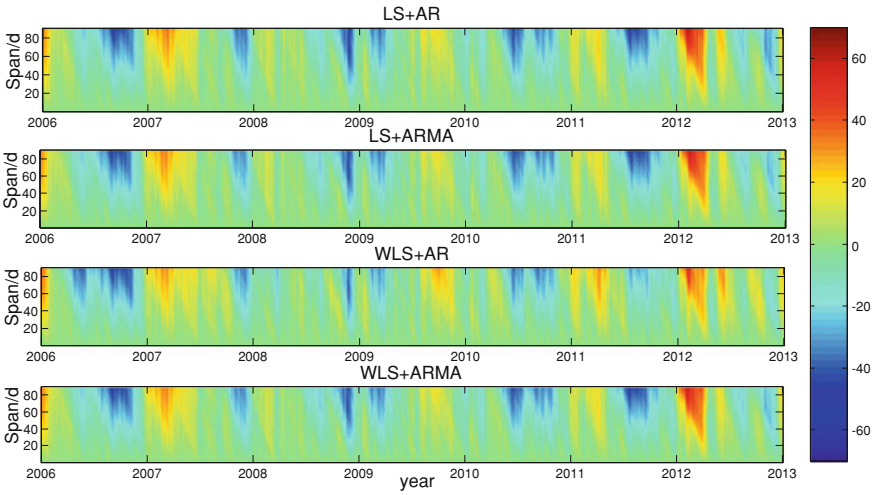


Fig. 25.2 Prediction errors for y-component of PM (in mas)

Table 25.1 Statistical result of prediction accuracy

Span	X-pole/mas				Y-pole/mas			
	LSAR	LSARMA	WLSAR	WLSARMA	LSAR	LSARMA	WLSAR	WLSARMA
1	0.2432	0.2419	0.2412	0.2414	0.1945	0.1935	0.1945	0.1938
2	0.5678	0.5659	0.5606	0.5620	0.4117	0.4097	0.4107	0.4087
3	0.9334	0.9290	0.9196	0.9226	0.6492	0.6465	0.6450	0.6424
4	1.2976	1.2897	1.2740	1.2778	0.8722	0.8694	0.8648	0.8625
5	1.6560	1.6429	1.6183	1.6218	1.0819	1.0787	1.0710	1.0684
6	2.0022	1.9794	1.9483	1.9509	1.2732	1.2710	1.2592	1.2583
7	2.3230	2.2909	2.2529	2.2511	1.4529	1.4543	1.4376	1.4407
8	2.6334	2.5865	2.5370	2.5356	1.6316	1.6363	1.6139	1.6197
9	2.93841	2.8762	2.8172	2.8123	1.8101	1.8145	1.7895	1.7936
10	3.23498	3.1577	3.0880	3.0841	1.9855	1.9873	1.9603	1.967
20	6.17212	5.8851	5.6282	5.5944	3.8049	3.8141	3.6754	3.7149
30	9.26791	8.6813	8.1233	8.0390	5.7603	5.7807	5.5547	5.6372
40	12.1863	11.2846	10.4609	10.3048	7.7368	7.7597	7.5078	7.6257
50	15.1113	13.7349	12.5618	12.3568	9.8438	9.8328	9.4575	9.6069
60	17.9577	16.1041	14.4788	14.2427	11.9618	11.8727	11.3189	11.478
70	20.6023	18.3175	16.4027	16.1424	14.0626	13.8771	13.1431	13.2991
80	23.1079	20.3469	18.0776	17.8137	16.2265	15.9181	14.8541	15.0035
90	25.4553	22.1963	19.4088	19.1581	18.3340	17.9377	16.3743	16.5478

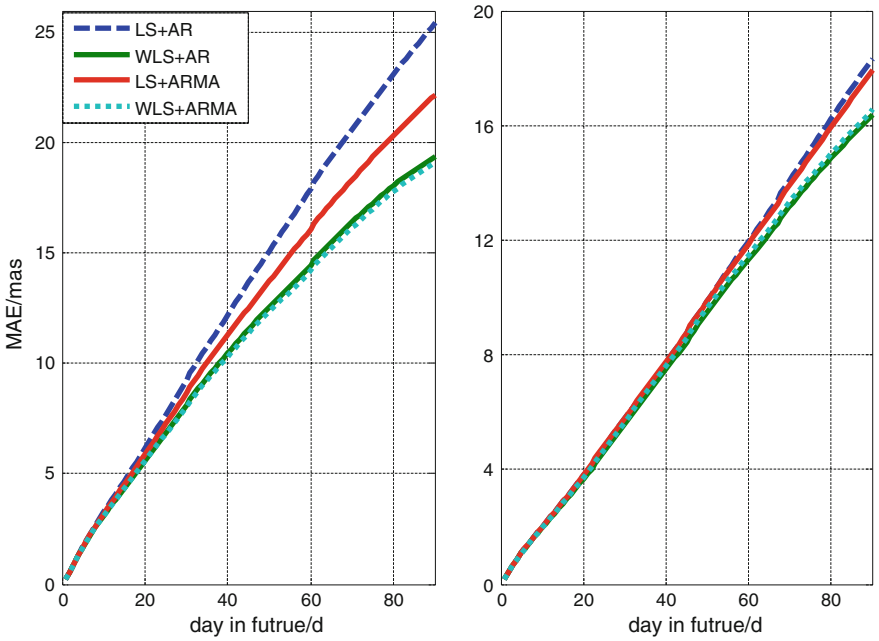


Fig. 25.3 MAE for different prediction span

25.4 Conclusions

Due to the influence of the periodic term and linear term in the PM data, the closer the observational data is near to the prediction point, the greater the impact on the prediction is. A weighted least-squares is developed to fit the observation data in order to highlight the influence of recent data. The AR model and ARMA model are used to predict the residual time series. The results show that the LS + ARMA model has higher prediction accuracy than LS + AR model for x and y component of PM. The WLS + AR model and WLS + ARMA model almost have the same prediction accuracy with minor difference in the near real-time and long-term prediction. The WLS + AR model and WLS + ARMA model has higher prediction accuracy than LS + AR model and LS + ARMA model, which can be a good choice for PM prediction.

Acknowledgments This work was supported by Natural Science Foundation of China (41174008) and the State Key Laboratory of Geodesy and Geodynamics Open Funded Projects (SKLGED2013-4-2-EZ) and the Open Foundation of State Key Laboratory of Astronautic and Dynamics (2014ADL-DW0101).

References

1. Zheng DW, Yu NH (1996) Earth rotation and it's relations to geophysical phenomena: the changes of length of the day. Adv Geophys 11(2):81–101 (In Chinese)
2. Schuh H, Ulrich M, Egger D et al (2002) Prediction of Earth orientation parameters by artificial neural networks. J Geodesy 76:247–258
3. Akyilmaz O, Kutterer H (2004) Prediction of Earth rotation parameters by fuzzy inference systems. J Geodesy 78:82–93
4. Schuh H, Nagel S, Seitz T (2001) Linear drift and periodic variations observed in long time series of polar motion. J Geodesy 74:701–710
5. Kosek W, Kalarus M, Niedzielski T (2008) Forecasting of the Earth orientation parameters—comparison of different algorithms. In: Capitaine N (ed) Proceedings of the journèes 2007, systèmes de référence spatio-temporels "The celestial reference frame for the future". Observatoire de Paris Systèmes de Référence Temps-Espace UMR8630/CNRS, Paris, pp 155–158
6. Kosek W (2010a) Future improvements in EOP prediction. In: Proceedings of the IAG 2009, "Geodesy for planet earth", 31 August–4 September, 2009, Buenos Aires, Argentina (accepted)
7. Kosek W (2010b) Causes of prediction errors of pole coordinates data. In: Proceedings of the 6th Orlov's conference, "The study of the Earth as a planet by methods of geophysics, geodesy and astronomy", 22–24 June 2009. MAO NAS of Ukraine, Kiev, pp 96–103
8. Kalarus M, Schch H, Kozek W et al (2010) Achievements of Earth orientation parameters prediction comparison campaign. J Geodesy 84:587–596
9. Kosek W, McCarthy DD, Luzum BJ (1998) Possible improvement of Earth orientation forecast using autocovariance prediction procedures. J Geodesy 72:189–199
10. Liao D, Wang Q, Zhou Y, Liao X, Huang C (2012) Long-term prediction of the Earth orientation parameters by the artificial neural network technique. J Geodyn, GEOD-1110, 6
11. Schuh H, Ulrich M, Egger D et al (2002) Prediction of Earth orientation parameters by artificial neural networks. J Geodesy 76:247–258
12. Wang QJ (2007) Studies on the prediction of Earth's variable rotation by artificial neural networks. Shanghai Astronomical observatory, Chinese Academy of Sciences, Shanghai (In Chinese)
13. Akyilmaz O, Kutterer H (2004) Prediction of Earth rotation parameters by fuzzy inference systems. J Geodesy 78:82–93
14. Zhang H, Wang QJ, Zhu JJ et al (2011) Joint model of weighted least-squares and AR in prediction of polar motion. Prog Astron 29(3): 343–352 (In Chinese)
15. Sun Z, Xu T (2012) Prediction of Earth rotation parameters based on an improved weighted least-squares and AR model. Geodesy Geodyn 3(3): 57–64
16. Xu XQ, Zhou YH (2010) High precision prediction method of Earth orientation parameters. J Spacecraft TT&C Technol 29(2): 70–76 (In Chinese)

Chapter 26
Ambiguity Acceptance Testing: A Comparison of the Ratio Test and Difference Test

Lei Wang, Sandra Verhagen and Yanming Feng

Abstract Integer ambiguity resolution is an indispensable procedure for all high precision GNSS applications. The correctness of the estimated integer ambiguities is the key to achieving highly reliable positioning, but the solution cannot be validated with classical hypothesis testing methods. The integer aperture estimation theory unifies all existing ambiguity validation tests and provides a new prospective to review existing methods, which enables us to have a better understanding on the ambiguity validation problem. This contribution analyses two simple but efficient ambiguity validation test methods, ratio test and difference test, from three aspects: acceptance region, probability basis and numerical results. The major contribution of this paper can be summarized as: (1) The ratio test acceptance region is overlap of ellipsoids while the difference test acceptance region is overlap of half-spaces. (2) The probability basis of these two popular tests is firstly analyzed. The difference test is an approximation to optimal integer aperture, while the ratio test follows an exponential relationship in probability. (3) The limitations of the two tests are firstly identified. The two tests may under-evaluate the failure risk if the model is not strong enough or the float ambiguities fall in particular region. (4) Extensive numerical results are used to compare the performance of these two tests. The simulation results show the ratio test outperforms the difference test in some models while difference test performs better in other models. Particularly in the medium baseline kinematic model, the difference tests outperforms the ratio test, the superiority is independent on frequency number, observation noise, satellite geometry, while it depends on success rate and failure rate tolerance. Smaller failure rate leads to larger performance discrepancy.

L. Wang (✉) · Y. Feng
Science and Engineering Faculty, Queensland University of Technology, Queensland, Australia
e-mail: L62.wang@qut.edu.au

S. Verhagen
Delft University of Technology, Delft, The Netherlands

L. Wang
School of Geodesy and Geomatics, Wuhan University, Wuhan, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_26,

Keywords Ratio test · Difference test · Ambiguity acceptance test · Ambiguity validation · Integer aperture estimation

26.1 Introduction

Integer ambiguity resolution is an indispensable procedure for high precision GNSS applications. After resolving the integer ambiguity parameters, the carrier phase measurements can be used as high precision distance measurements, and thus improve the coordinate parameter estimation. However, integer ambiguity resolution is very risky because if the ambiguity parameters are fixed to incorrect integers, an unexpected large error would be introduced to the positioning solution without notice and consequently the unnoticed bias will contaminate the positioning results. In order to avoid the unacceptable error caused by the incorrectly fixed ambiguities, an additional validation procedure has to be carried out, which aims at rejecting potentially wrongly fixed integer parameters. This procedure is called ambiguity acceptance testing or ambiguity validation.

Many ambiguity acceptance tests have been proposed and examined extensively. Most of the ambiguity acceptance test methods construct test statistics based on the squared Euclidean norm between the estimated real-valued ambiguity parameters and the integer ambiguity parameter candidates. The ratio test [4] and the difference test [20] are the two most popular tests among these test methods, because they are simple and efficient.

The standard carrier-phase based GNSS positioning model can be solved in three steps:

(1) Applying the standard least-squares or Kalman filtering without considering the integer nature of the ambiguity parameters. Corresponding estimated ambiguity parameters and real-valued parameters are denoted as $\hat{a}$ and $\hat{b}$. As the GNSS positioning model is a linear model, the estimated parameters $\hat{a}$ and $\hat{b}$ follow multivariate normal distribution, its vc-matrix is denoted as:

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \begin{pmatrix} Q_{\hat{a}\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}\hat{b}} \end{pmatrix} \tag{26.1}$$

The integer nature of parameter vector $\hat{a}$ was considered in this step. Therefore, the estimated parameters $\hat{b}$ and $\hat{a}$ are referred to as the 'float solution'.

(2) Mapping the real-valued ambiguity parameter $\hat{a}$ to integers, which is also known as the 'integer ambiguity resolution'. The integer ambiguity resolution procedure includes ambiguity estimation and ambiguity validation. The ambiguity estimation is to map the float ambiguity $\hat{a}$ to an integer. The ambiguity validation is to decide whether to accept the fixed 'integer candidate'.

The ambiguity estimation procedure can be expressed as:

$$\breve{a} = z \Leftrightarrow \hat{a} \in S_z, z \in \mathbb{Z}^n \tag{26.2}$$

where S_z is the so-called 'pull-in region' centered at integer vector z. n is the dimension of the ambiguity vector. A float ambiguity vector $\hat{a}$ falling in a certain pull-in region S_z will be mapped to its center z by the integer estimator. With a given $Q_{\hat{a}\hat{a}}$, the pull-in region is defined by the integer estimator, thus different integer estimators may map the same $\hat{a}$ to different integers $\breve{a}$. Among all admissible integer estimator, integer least-squares (ILS) is the optimal integer estimator [14]. It is because the ILS can provide the highest probability to fix to the correct integer estimation.

The reliability of the fixed solution can be evaluated by the ambiguity residual and its Euclidean norm. The ambiguity residual is defined as the difference between the real-valued vector $\hat{a}$ and the 'best integer candidate' $\breve{a}$ [15, 23]. The ambiguity residual and its squared Euclidean norm can be expressed as:

$$\begin{aligned} &\breve{\epsilon} = \hat{a} - \breve{a}, \hat{a} \in \mathbb{R}^n, \breve{a} \in \mathbb{Z}^n \\ &\|\hat{a} - \breve{a}\|^2_{Q_{\hat{a}\hat{a}}} = (\hat{a} - \breve{a})^T Q^{-1}_{\hat{a}\hat{a}} (\hat{a} - \breve{a}) \end{aligned} \tag{26.3}$$

As mentioned, the ambiguity acceptance test have to be applied to ensure the correctness of the fixed integer ambiguities $\breve{a}$. If the fixed ambiguity parameter $\breve{a}$ is rejected, the float solution $\hat{a}$ will be adopted as the final solution. If the fixed ambiguity parameter $\breve{a}$ is accepted, the $\breve{a}$ can be used to update $\hat{b}$. The details of ambiguity acceptance testing will be discussed in the following sections.

(3) Updating real-valued parameters $\hat{b}$. If the 'best integer candidate' $\breve{a}$ is accepted by ambiguity validation procedure, the real-valued parameters $\hat{b}$ can be improved based on the fixed integer ambiguity parameters $\breve{a}$.

In this paper, two popular ambiguity acceptance tests named ratio test and difference test are compared from the definition, the acceptance region, the probability basis and the performance aspects. The similarity and difference between these two tests will be compared.

26.2 The Definition of the Ratio Test and the Difference Test

The ratio test and the difference test are the two simplest, but most popular ambiguity acceptance tests. This section, the concept of these two tests as well as their acceptance region are compared.

26.2.1 The Ratio Test and Its Acceptance Region

The ratio test is defined as [1, 5, 6, 8]:

$$\frac{\|\hat{a}-\breve{a}\|^2_{Q_{\hat{a}\hat{a}}}}{\|\hat{a}-\breve{a}_2\|^2_{Q_{\hat{a}\hat{a}}}} \leq \mu_R \tag{26.4}$$

where $\breve{a}$ and $\breve{a}_2$ are known as the 'best integer candidate' and the 'second best integer candidate' respectively. The 'best' and 'second best' is defined in terms of the Euclidean norm (see Eq. 26.3). The best integer candidate is defined as $\|\hat{a}-z\|^2_{Q_{\hat{a}\hat{a}}} = min, z \in \mathbb{Z}^n$, and similarly, second best candidate has the second smallest distance.

The float ambiguity $\hat{a}$ and its variance-covariance (vc-) matrix $Q_{\hat{a}\hat{a}}$ can be obtained by the standard least-squares. The best integer candidate $\breve{a}$ and second best integer candidate $\breve{a}$ can be obtained from the integer least-squares estimator. The details of the integer least-squares search procedure have been extensively discussed [2, 3, 9]. μ_R is the threshold for the ratio test and its value varies in the interval [0,1].

According to the integer aperture estimation theory, the ratio test is a member of the class of integer aperture estimators, denoted as 'RTIA' [16, 17, 21, 22]. Similar to the class of integer estimators, all integer aperture estimators have their own pull-in region, which is preferably called 'acceptance region' in the ambiguity acceptance testing. The acceptance region of the ratio test is defined as [21, 22]:

$$\Omega_{0,R} = \{x \in S_0 | \left\| x + \frac{\mu_R}{1-\mu_R} z \right\|^2_{Q_{\hat{a}\hat{a}}} \leq \frac{\mu_R}{(1-\mu_R)^2} \|z\|^2_{Q_{\hat{a}\hat{a}}}, \forall z \in \mathbb{Z}^n \backslash \{0\}\} \tag{26.5}$$

where $\Omega_{0,R}$ means the acceptance region for ratio test centered at the integer vector $\{0\}$. The best integer candidate $\breve{a}$ is assumed to be the 0-vector and z is a non-zero n-dimensional integer vector. The Eq. (26.5) is transformed from the definition of the ratio test (see Eq. 26.4), only when the $\hat{a}$ meets all these constraints, the fixed integer candidate $\breve{a}$ can be accepted.

The Eq. (26.5) can be geometrically interpreted as an overlap of many hyper-ellipsoids. $\forall z \in \mathbb{Z}^n \backslash \{0\}$, there is a corresponding hyper-ellipsoid centered at $-\frac{\mu_R}{1-\mu_R} z$ and its size is governed by $\frac{\mu_R}{(1-\mu_R)^2} \|z\|^2_{Q_{\hat{a}\hat{a}}}$. These hyper-ellipsoids can be uniquely described by four factors: shape, size, orientation and center. The Eq. (26.5) shows that the center and size of hyper-ellipsoids are controlled by both μ_R and z. The center of each hyper-ellipsoid is determined by z scaled by $-\frac{\mu_R}{1-\mu_R}$; larger μ_R lead to larger $\frac{\mu_R}{1-\mu_R}$. Similarly, the hyper-ellipsoid size is determined by $\|z\|^2_{Q_{\hat{a}\hat{a}}}$ scaled by $\frac{\mu_R}{(1-\mu_R)^2}$; larger μ_R also lead to a larger size. The other two factors, shape and orientation are determined by the eigenvalues and eigenvectors of $Q_{\hat{a}\hat{a}}$. Consequently, all ellipsoids have exactly the same eccentricity (or flattening) and

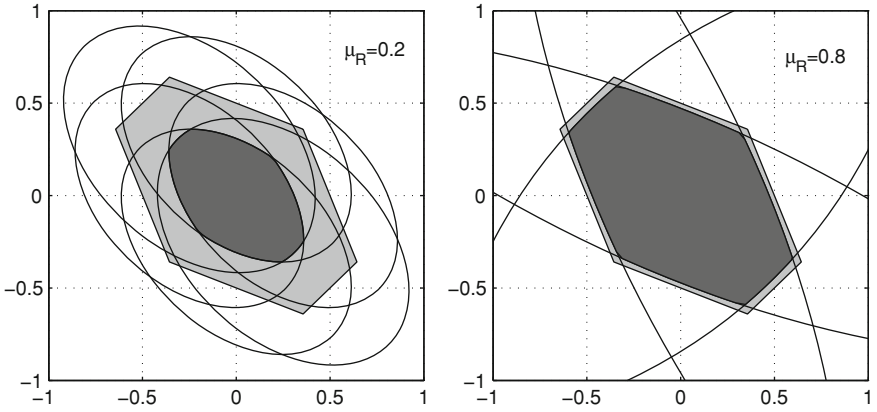


Fig. 26.1 A 2-D example of ratio test acceptance region with different threshold. The *left* panel presents the construction of ratio acceptance region with threshold $\mu_R = 0.2$, while the threshold in the *right* panel is 0.8. The *light gray* region is corresponding integer least-*squares* pull-in region

orientation. A $Q_{\hat{a}\hat{a}}$ with a larger condition number corresponds to more elongated hyper-ellipsoids and with a smaller condition number corresponds to more sphere-like hyper-ellipsoids. The ratio test acceptance region is bounded by the surface of these ellipsoids.

A 2-dimensional example of the ratio test acceptance region with different μ_R is presented in Fig. 26.1. The left panel shows the case $\mu_R = 0.2$, the ellipses are more centralized and have a smaller size. The curvature radius on the acceptance boundary is small and consequently, the acceptance region is more ellipse-like. The right panel shows ratio test acceptance region when $\mu_R = 0.8$. All ellipses have a larger size and their centers are more spread out. The corresponding acceptance region boundaries have a large curvature radius and become more straight-line-like. As a result, the whole acceptance region is more hexagon-like. According to above analysis, the ratio test is very sensitive to its threshold, thus determining its threshold reasonably is very critical for the ratio test.

26.2.2 The Difference Test and Its Acceptance Region

The difference test is defined as [20]:

$$\|\hat{a} - \check{a}_2\|^2_{Q_{\hat{a}\hat{a}}} - \|\hat{a} - \check{a}\|^2_{Q_{\hat{a}\hat{a}}} \geq \mu_D \tag{26.6}$$

The difference test uses the difference of two distances rather than the ratio to test the closeness between $\hat{a}$ and $\check{a}$. μ_D is the threshold for the difference test. The difference test is an integer aperture estimator as well, denoted as 'DTIA'

[17, 21, 22]. The corresponding acceptance region for the difference test can be given as [21, 22]:

$$\Omega_{0,D} = \{x \in S_0 | x^T Q_{\hat{a}\hat{a}}^{-1} z \leq \frac{1}{2}(\|z\|^2_{Q_{\hat{a}\hat{a}}} - \mu_D), \forall z \in \mathbb{Z}^n \backslash \{0\}\} \tag{26.7}$$

where $x^T Q_{\hat{a}\hat{a}}^{-1} z$ means the orthogonal projection of vector x on z direction in $Q_{\hat{a}\hat{a}}$ spanned space. With given $Q_{\hat{a}\hat{a}}$, the Eq. (26.7) shows the intersection of a series of half spaces. The acceptance region of the difference test coincides with the ILS pull-in region when $\mu_D = 0$. Another extreme case is $\mu_D = \frac{1}{2} min(\|\check{a} - \check{a}_2\|^2_{Q_{\hat{a}\hat{a}}})$ [10]. The orthogonal projector is not only used in difference test acceptance region construction, but also used in integer aperture least-squares (IALS) and projector test (PTIA) acceptance region construction [21].

The boundaries of the difference test acceptance region can be expressed as:

$$x^T Q_{\hat{a}\hat{a}}^{-1} z = \frac{1}{2}(\|z\|^2_{Q_{\hat{a}\hat{a}}} - \mu_D), \forall z \in \mathbb{Z}^n \backslash \{0\} \tag{26.8}$$

For a given integer vector z, the corresponding boundary is a subspace orthogonal to the vector z, which can be proven as follows:

$$\left.\begin{aligned} x_1^T Q_{\hat{a}\hat{a}}^{-1} z &= \frac{1}{2}(\|z\|^2_{Q_{\hat{a}\hat{a}}} - \mu_D) \\ x_2^T Q_{\hat{a}\hat{a}}^{-1} z &= \frac{1}{2}(\|z\|^2_{Q_{\hat{a}\hat{a}}} - \mu_D) \end{aligned}\right\} \\ \Rightarrow x_1^T Q_{\hat{a}\hat{a}}^{-1} z = x_2^T Q_{\hat{a}\hat{a}}^{-1} z \\ \Leftrightarrow (x_1 - x_2)^T Q_{\hat{a}\hat{a}}^{-1} z = 0; \tag{26.9}$$

The last equation indicates that the arbitrary vector $x_1 - x_2$ in the boundary subspace is perpendicular to a particular vector z, thus the whole boundary subspace is perpendicular to the integer vector z. The point $\frac{1}{2}(1 - \frac{\mu}{\|z\|^2_{Q_{\hat{a}\hat{a}}}})z$ is always on the boundary of DTIA acceptance region. The difference test acceptance region for a 2-D example is illustrated in Fig. 26.2. The figure illustrates that different μ_D does not change the orientation of acceptance region boundary, because the orientation of the half-spaces is defined by eigenvectors of $Q_{\hat{a}\hat{a}}$.

26.3 Probability Analysis of Ratio Test and Difference Test

The ambiguity acceptance testing is a hypothesis testing problem. Typical hypothesis testing problems can be solved by constructing a test statistic following a well-known distribution, and the threshold can be determined with the help of the

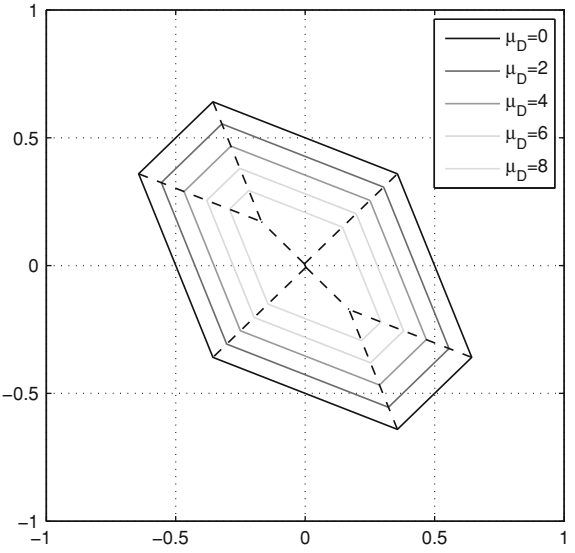


Fig. 26.2 A 2-D difference test acceptance region example with different thresholds. The boundaries of the acceptance region with respect to different μ_D is presented in different color and the *dash line* is the boundaries correspond to different second best integer candidate $\breve{a}_2$

knowledge about the distribution. For the ambiguity acceptance test problem, the construction of the test statistic attracted more attention while the distribution is less studied. This section analyzes the probability distribution foundation of ratio test and difference test, and the limitations of these tests are identified based on a probability analysis.

26.3.1 The Probability Basis of Ambiguity Validation

Early investigation on the ratio test assumed that the squared Euclidean norm $\|\hat{a} - \breve{a}\|^2_{Q_{\hat{a}\hat{a}}}$ follows a χ^2-distribution, and the ratio of the two squared Euclidean norms follow the F-distribution [5, 6]. However, this assumption is not correct. The first reason is that both numerator and denominator are function of $\hat{a}$, thus they are inevitably correlated. Another reason is that the stochastic characteristics of the fixed ambiguity parameter vector $\breve{a}$ is not considered in the assumption. The squared Euclidean norm $\|\hat{a} - \breve{a}\|^2_{Q_{\hat{a}\hat{a}}}$ follows a χ^2-distribution only when the $\breve{a}$ is deterministic, however, this is not true in general. The stochastic properties of $\breve{a}$ can be described by a probability mass function (PMF) [15, 21, 23].

The float ambiguity parameter $\hat{a}$ follows a multivariate normal distribution and its probability density function (PDF) can be given as:

$$f_{\hat{a}}(x) = \frac{1}{\sqrt{detQ_{\hat{a}\hat{a}}}(2\pi)^{\frac{1}{2}n}} exp\{-\frac{1}{2}\|x - a\|^2_{Q_{\hat{a}\hat{a}}}\} \tag{26.10}$$

where $detQ_{\hat{a}\hat{a}}$ means determinant of the matrix $Q_{\hat{a}\hat{a}}$. The symbol '$exp\{\cdot\}$' means the exponential function $e^{\{\cdot\}}$. The PDF of the ambiguity residual $\check{\epsilon}$ can be given as a infinite sum of $f_{\hat{a}}(x+z)$ because all integer vectors z can be viewed as potential true ambiguities. The definition of the ambiguity residual PDF can be expressed as [15, 23]:

$$f_{\check{\epsilon}}(x) = \sum_{z \in Z^n} f_{\hat{a}}(x+z) s_0(x), x \in \mathbb{R}^n, s_z(x) = \begin{cases} 1 & \text{if } x \in S_z \\ 0 & \text{otherwise} \end{cases} \tag{26.11}$$

Based on the ambiguity residual distribution, a theoretical optimal integer aperture estimator is derived, which is known as optimal integer aperture estimator (OIA) [18, 21]. The OIA can always achieves maximum success rate with a given failure rate constraint. With modified Neyman-Pearson lemma, the constrained maximum success-rate problem can be converted to a 'likelihood ratio' form [11, 18]. The definition of OIA is given as:

$$\begin{aligned} \Omega_{0,O} =& \{x \in S_0 | \frac{f_{\hat{a}}(x)}{f_{\check{\epsilon}}(x)} \geq \mu_O\} \\ \Leftrightarrow \Omega_{0,O} =& \{x \in S_0 | \frac{f_{\hat{a}}(x)}{\Sigma_{z \in \mathbb{Z}^n} f_{\hat{a}}(x+z)} \geq \mu_O\} \\ \Leftrightarrow \Omega_{0,O} =& \{x \in S_0 | \frac{f_{\hat{a}}(x)}{f_{\hat{a}}(x) + f_{\hat{a}}(x-\check{a}_2) + f_{\hat{a}}(x-\check{a}_3) + \ldots} \geq \mu_O\} \end{aligned} \tag{26.12}$$

In the equation, $\check{a}$ is assumed to be the integer vector 0, S_0 is corresponding ILS pull-in region.

26.3.2 Probability Analysis for the Difference Test

The difference test can be expressed with $f_{\hat{a}}(x)$ and $f_{\hat{a}}(x-\check{a}_2)$ using the following transformation:

$$\begin{aligned} & \|x-\check{a}_2\|^2_{Q_{\hat{a}\hat{a}}} - \|x\|^2_{Q_{\hat{a}\hat{a}}} \geq \mu_D, \forall x \sim N(0, Q_{\hat{a}\hat{a}}), x \in S_0 \\ \Leftrightarrow & exp\{-\frac{1}{2}(\|x-\check{a}_2\|^2_{Q_{\hat{a}\hat{a}}} - \|x\|^2_{Q_{\hat{a}\hat{a}}})\} \leq exp\{-\frac{1}{2}\mu_D\} \\ \Leftrightarrow & \frac{exp\{-\frac{1}{2}\|x-\check{a}_2\|^2_{Q_{\hat{a}\hat{a}}}\}}{exp\{-\frac{1}{2}\|x\|^2_{Q_{\hat{a}\hat{a}}}\}} \leq exp\{-\frac{1}{2}\mu_D\} \\ \Leftrightarrow & \frac{f_{\hat{a}}(x-\check{a}_2)}{f_{\hat{a}}(x)} \leq exp\{-\frac{1}{2}\mu_D\} \\ \Leftrightarrow & \frac{f_{\hat{a}}(x)}{f_{\hat{a}}(x) + f_{\hat{a}}(x-\check{a}_2)} \geq \mu'_D, \mu'_D = \frac{1}{exp\{-\frac{1}{2}\mu_D\}+1} \end{aligned} \tag{26.13}$$

After the transformation, the difference test becomes a truncated version of OIA. For the test statistics of DTIA and OIA, it holds that: $\frac{f_{\hat{a}}(x)}{f_{\hat{a}}(x)+f_{\hat{a}}(x-\check{a}_2)} < \frac{f_{\hat{a}}(x)}{f_{\check{\epsilon}}(x)}$, thus the OIA test statistic $\frac{f_{\hat{a}}(x)}{f_{\check{\epsilon}}(x)}$ can be viewed as a upper bound of DTIA test statistic $\frac{f_{\hat{a}}(x)}{f_{\hat{a}}(x)+f_{\hat{a}}(x-\check{a}a_2)}$. Performance of DTIA depends on the closeness between the two test statistics, when the difference becomes ignorable, the DTIA become a closed form of OIA.

26.3.3 Probability Analysis for the Ratio Test

Similarly, the ratio test can be expressed with $f_{\hat{a}}(x)$ and $f_{\hat{a}}(x-\check{a}_2)$ as well:

$$\begin{aligned}
&\frac{\|x\|^2_{Q_{\hat{a}\hat{a}}}}{\|x-\check{a}_2\|^2_{Q_{\hat{a}\hat{a}}}} \leq \mu_R, \forall x \sim N(0, Q_{\hat{a}\hat{a}}), x \in S_0 \\
\Leftrightarrow &\|x\|^2_{Q_{\hat{a}\hat{a}}} \leq \mu_R \|x-\check{a}_2\|^2_{Q_{\hat{a}\hat{a}}} \\
\Leftrightarrow &exp\{-\frac{1}{2}\|x\|^2_{Q_{\hat{a}\hat{a}}}\} \geq exp\{-\frac{1}{2}\mu_R\|x-\check{a}_2\|^2_{Q_{\hat{a}\hat{a}}}\} \\
\Leftrightarrow &f_{\hat{a}}(x) \geq C^{(\mu_R-1)}(f_{\hat{a}}(x-\check{a}_2))^{\mu_R}, C = \sqrt{detQ_{\hat{a}\hat{a}}}(2\pi)^{\frac{1}{2}n}
\end{aligned} \tag{26.14}$$

The last formula in the Eq. (26.14) shows that the ratio test does not have a similar form like OIA. The ratio test discriminates $f_{\hat{a}}(x)$ and $f_{\hat{a}}(x-\check{a}_2)$ with an exponential relationship, the non-linear relationship makes the ratio test performance more unpredictable.

26.3.4 The Limitations of the Ratio Test and the Difference Test

According to the above analysis, the difference test can be viewed as an approximation of the OIA, while the ratio test does not have a similar form like the OIA. The OIA considers all possible integer candidates, while the ratio test and the difference test only consider $\check{a}$ and $\check{a}_2$, which makes these two tests sub-optimal.

In the ratio test and the difference test statistics, the second best integer candidate $\check{a}_2$ is incorporated as the only alternative integer candidate. If $\hat{a}$ is 'close' enough to $\check{a}$ and 'far' enough from the alternative candidate $\check{a}_2$, both discrimination tests will accept $\check{a}$. However, this conclusion does not always make sense. The definition of $f_{\check{\epsilon}}(x)$ (see Eq. 26.11) shows all integer candidates $z \in \mathbb{Z}^n$ have a probability contribution to $f_{\check{\epsilon}}(x)$, or we can say all integer candidates $z \in \mathbb{Z}^n$ are

potential true integer ambiguity vectors, but with different likelihoods. Most of them have extremely low probability, so that they are 'almost' impossible to be true. Only the integer candidates with considerable probability contribution need to be considered in ambiguity validation. According to the multivariate normal distribution PDF (see Eq. 26.10), $f_{\hat{a}}(x)$ is a monotone decreasing function of $\|x-z\|^2_{Q_{\hat{a}\check{a}}}$, thus smaller $\|x-z\|^2_{Q_{\hat{a}\check{a}}}$ means more likely to be true. Particularly, the ratio test and the difference test only incorporate the best candidate $\check{a}$ and the second best candidate $\check{a}_2$. There is no doubt that $\check{a}$ and $\check{a}_2$ are the two most likely true integer candidates, but only considering these two candidates in ambiguity acceptance testing still might be not enough. There are two reasons:

(1) Under-evaluation of the failure risk in the weak models. Generally, the best candidate $\check{a}$ and the second best candidate $\check{a}_2$ are two adjacent integer candidates. If the model is strong enough, the probability contribution from the non-adjacent candidates are small enough to be ignored, thus the following equation holds true:

$$f_{\hat{\epsilon}}(x) \approx f_{\hat{a}}(x) + \Sigma f_{\hat{a}}(x+\check{a}\prime), \check{a}\prime \in \mathbb{Z}^n \tag{26.15}$$

where $\check{a}\prime$ pre all integer candidate adjacent to $\check{a}$. If the model is not strong enough, the probability contribution from non-adjacent candidates cannot be ignored anymore, but the ratio test and the difference test still cannot consider these 'non-adjacent candidates'. Consequently, the failure risk is under-evaluated and the decision would be over-optimistic. The 1-dimensional case is demonstrated in Fig. 26.3. The left panel shows the strong model case ($\sigma = 0.2$), Eq. (26.15) holds true, and the ratio test and difference test can be applied safely. While the right panel shows the weak model case ($\sigma = 0.6$), the impact of the non-adjacent integer candidates ($f_{\hat{a}}(x-2)$ and $f_{\hat{a}}(x+2)$) becomes significant. While the ratio test and the difference test do not consider $f_{\hat{a}}(x-2)$ and $f_{\hat{a}}(x+2)$, and thus the ratio test and difference test performance would be degraded in this case.

(2) Ignoring the impact of $\check{a}_3$ in strong models. The ratio test and the difference test can be overly optimistic even in a strong models. Every integer vector have $2^n - 1$ pairs of adjacent integer vectors [21, 26], all of them are potential $\check{a}_2$. Which one is the real $\check{a}_2$ depends on the location of $\hat{a}$. For higher than 1-dimensional case, $\|\hat{a}-\check{a}_2\|^2_{Q_{\hat{a}\check{a}}}$ may equal to $\|\hat{a}-\check{a}_3\|^2_{Q_{\hat{a}\check{a}}}$ for both strong models and weak models. In this case, $f_{\hat{a}}(x-\check{a}_3) \approx f_{\hat{a}}(x-\check{a}_2)$. However, the ratio test and difference test cannot incorporate $f_{\hat{a}}(x-\check{a}_3)$, thus the failure risk is under-evaluated again. The 2-dimensional example is demonstrated in Fig. 26.4. The acceptance region of the ratio test and difference test does not coincide with the OIA acceptance region around the dash-line, which is because $f_{\hat{a}}(x-\check{a}_2) = f_{\hat{a}}(x-\check{a}_3)$ on the dashed line. This scenario may occur in both strong model and weak model scenarios.

According to above analysis, the simple form of the ratio test and the difference test makes them sub-optimal in two aspects. Fortunately, the first aspect can be

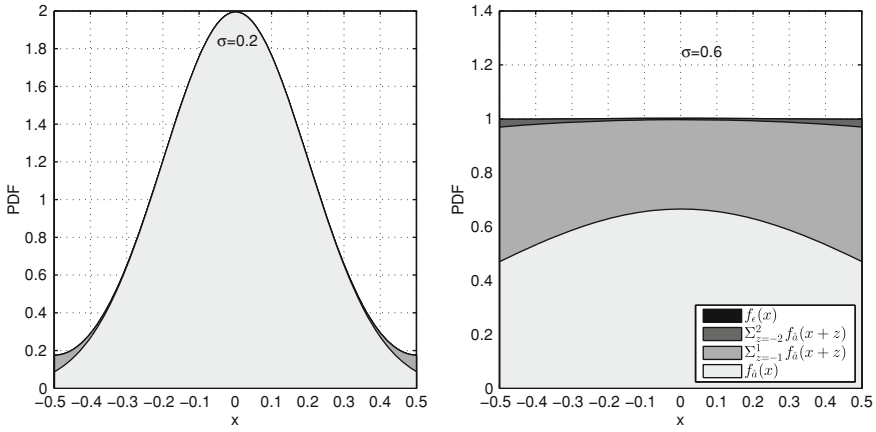


Fig. 26.3 1-D demonstration of ambiguity residual distribution $f_{\check{\epsilon}}(x)$. The *left* panel shows the ambiguity residual distribution $f_{\check{\epsilon}}(x)$ in a strong model ($\sigma = 0.2$). It is clear that $f_{\check{\epsilon}}(x) \approx f_{\check{a}}(x) + f_{\check{a}}(x + \check{a}_2)$ holds in the strong model. The *right* panel shows the ambiguity residual distribution $f_{\check{\epsilon}}(x)$ in a weak model ($\sigma = 0.6$). In this case, $f_{\check{\epsilon}}(x) \approx \Sigma^2_{z=-2} f_{\check{a}}(x+z)$ and the impact of non-adjacent integer $f_{\check{a}}(x+2)$ and $f_{\check{a}}(x-2)$ is not ignorable any longer

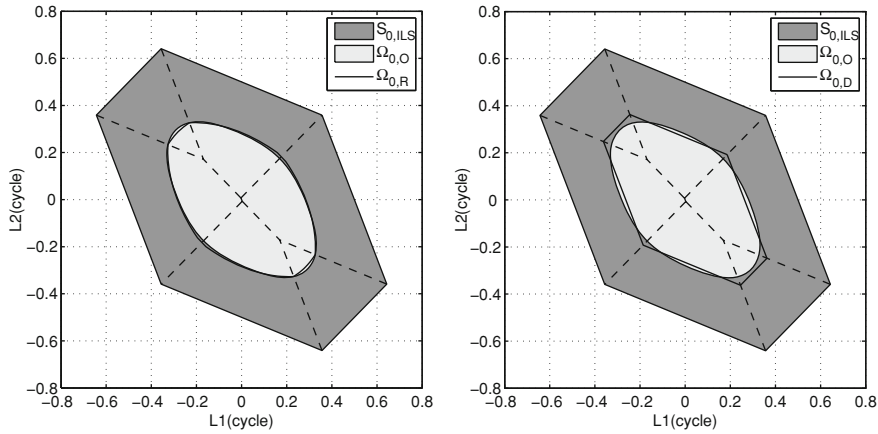


Fig. 26.4 A 2-D example of the acceptance region comparison between the ratio test, the difference test and the OIA. The fixed failure rate approach is applied to determine the threshold of the ratio test, the difference test and the OIA. The OIA acceptance region is used as a reference acceptance region, the acceptance region for the ratio test (*left* panel) and the difference test (*right* panel) are compared

largely overcome with the fixed failure rate approach. The threshold for the tests is then model-driven, such that its value is adapted to guarantee a maximum allowable failure rate [19, 24, 25].

However, the impact of $\breve{a}_3$ cannot be mitigated with the fixed failure rate approach, because the impact depends on the location of $\hat{a}$; different $\hat{a}$ may suffer different impacts, and the location of $\hat{a}$ cannot be controlled. Fortunately, the impact of $\breve{a}_3$ on the ratio test and difference test performance is usually very limited, and the ratio test and difference test still perform close to optimal in most cases.

26.4 The Numerical Comparison Between the Ratio Test and the Difference Test

In addition to the theoretical analysis, a numerical comparison between the ratio test and the difference test is carried out based on simulation data in this section. The numerical analysis includes two parts, an overall comparison and a geometry-based model comparison. The overall comparison aims at finding out the behavior difference in different GNSS models, which enables us to make clear the behavior of the two tests. In addition to the overall comparison, another comparison is made to take a closer look at the geometry-based model. It is because the geometry-based model is widely adopted in kinematic GNSS data processing. As only strong model is worth to attempt to fix the ambiguities, the geometry-based model comparison only considers model with high success rate. The success rate is used as a quality indicator and the fixed failure rate approach is employed to make these two tests comparable. The fixed failure rate ratio test (FF-ratio) has been extensively investigated [19, 24, 25], while the fixed failure rate difference test have not drawn much attention in existing studies.

26.4.1 Overall Comparison

The challenge of the overall comparison is how to set up models with different model strength. In this simulation, the model is described from three aspects: the function model, the stochastic model and the system configuration.

The function model refers to how to model the relationship between observations and unknown parameters. Both real-valued parameters and integer parameters are involved in the mixed integer model. The integer ambiguity parameters are identical for all models, but the real-valued parameters can involve different parameters, generally including geometry parameters, troposphere parameters and ionosphere parameters. Based on different parametrization, different models are generated. The geometry information in the observation model can be parametrized as a function of coordinate components (geometry-based model) or geometric distances (geometry-free model) [13]. If the coordinates are precisely known, the geometry parameters can be absent from the model, which is known as

Table 26.1 Simulation configuration

Configuration items	Default value	Variation range
GPS frequency (MHz)	1575.42, 1227.60, 1176.45	N/A
Beidou frequency(MHz)	1561.098, 1268.52, 1207.14	N/A
GPS obs. std	0.30 m, 0.003 m	0.01–5 m, 0.001–0.05 m
Beidou obs. std	0.60 m, 0.005 m	N/A
UD ion. std	0.02 m	0.01–0.5 m
Elevation cutoff	15°	N/A
Troposphere MF	GMF	N/A
Weighting strategy	Equal weight	N/A
Solution type	EBE	EBE/Time-averaged
Location	Brisbane (27°S, 153°E)	90N–90S, 180E–180W
Visible satNum	7, 9, 16	4–11, 4–10, 4–18
Time	0:00:00 22/04/2013	N/A
Failure rate tolerance	0.05	N/A
Sample number	100,000	N/A

geometry-fixed (CORS network) model. Troposphere information in the observation model can be modeled by a zenith tropospheric delay (ZTD) parameter with mapping functions (troposphere-float model) or by assuming the residual troposphere delays to be absent or known (troposphere-fixed model). The residual ionosphere delays are estimated on satellite basis and prior information is adopted as pseudo-observations to strengthen the model, where the variance of the pseudo-observations must be taken into account (ionosphere-weighted model). The two extreme cases, ionosphere-fixed model and ionosphere-float model are considered as well [12].

The stochastic model is used to describe the randomness of the observations. In this simulation study, the statistical correlation between different observation types is simply ignored, only observation noise (including code and carrier phase) and ionosphere noise are considered. Different prior noise strategies and their combination is tested, details are given in Table 26.1.

The system configuration refers to frequency number, constellation, processing strategy etc., which are beyond mathematical modeling but still affect the model strength. The single frequency, dual frequency and (future) triple frequency case are all considered in the simulations. The GPS only, Beidou only and GPS + Beidou constellations are considered. Considering different satellite geometries, several locations with different visible satellite numbers are picked out from a global scaled grid. Both epoch by epoch (EBE) and time averaged model are considered in this simulation.

All above factors are organized as four categories: satellite geometry impact, stochastic model, ionosphere variance and time averaged model. Only relevant configuration changes when discussing a particular factor, the remaining configuration factors are kept to their 'default value' as shown in the corresponding column in Table 26.1. The variation range of particular factors are listed in the

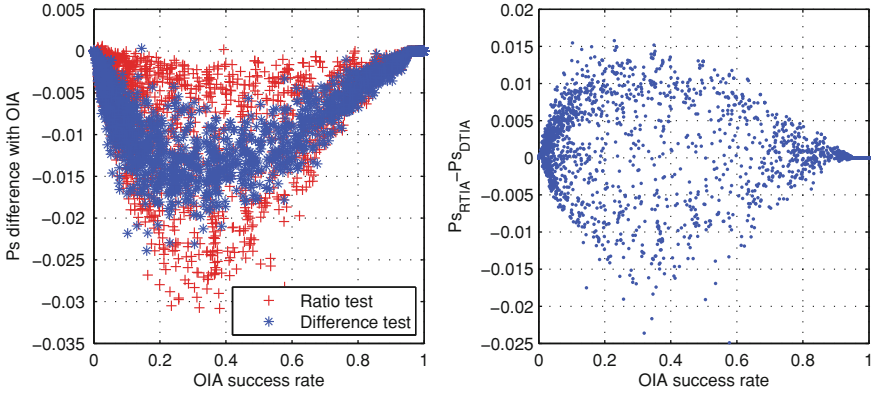


Fig. 26.5 A comparison of the ratio test and the difference test in terms of success rate. Approximately 5,000 samples with different geometry model, ionosphere model, troposphere model, frequency number, observation noise, constellation are presented in this figure. The *left* panel presents the success rate difference between OIA and ratio test, difference test. The *right* panel presents success rate difference between ratio test and difference test. In this particular case, failure rate tolerance is 5 %

'Variation Range' column in Table 26.1. Thus, the simulation scheme only reflects one particular factor impact on the ratio test and the difference test performance in each experiment. Overall, about 5,000 different scenarios with different model combinations are generated. These models are used to represent all possible scenarios in GNSS data process.

Examples with different models or configurations are generated in this experiment. The fixed failure rate approach is applied to the ratio test and the difference test to determine their threshold. The Monte-Carlo simulation method is used to calculate their success rate and failure rate. 100,000 samples are generated for each models. Finally, the success rate (percentage of samples being correctly accepted by the test) is used as a quality indicator to evaluate the performance of the tests.

The simulation results are presented in Fig. 26.5. The left panel shows the success rate difference between the OIA and the ratio test, the difference test. The OIA achieves the maximum success rate with the fixed failure rate approach, and consequently the success rate difference between OIA, ratio test and difference test can be used as a measure of optimality. The figure shows that the performance of the ratio test and the difference test are comparable with OIA, although these two tests are sub-optimal according the probability analysis. The ratio test and the difference test have similar behavior, but the success rate of the ratio test is more disperse. The right panel shows the success rate difference between the ratio test and the difference test. Apparently, there is no clear preference between these two tests in model level.

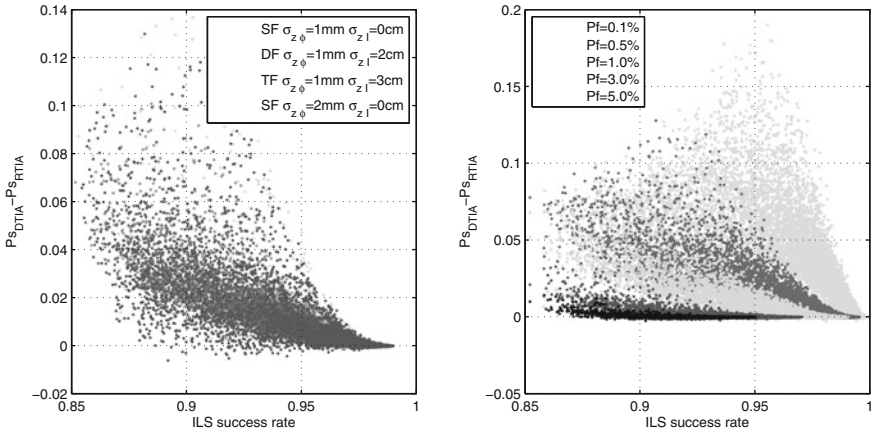


Fig. 26.6 The ratio test and the difference test performance comparison in the strong model. The *left* panel shows success rate difference of these two tests in different scenarios. Corresponding failure rate tolerance for all scenarios are 1.0 %. The *right* panel shows success rate difference of the two tests with different failure rate tolerance in dual frequency, $\sigma_{z\phi} = 1$ mm, $\sigma_{zI} = 2$ cm model

26.4.2 Geometry-Based Model Comparison

According to the overall comparison, both the ratio test and the difference test perform close to optimal, and there is no clear preference for these two tests. The next step is to investigate the most relevant model in GNSS positioning applications.

In the second simulation, the geometry-based, ionosphere-weighted and troposphere-fixed model is focused in this section, because this model is commonly used in medium-distance kinematic GNSS data process. In this simulation, the ionosphere standard deviation in zenith direction is assumed to be smaller than 3 cm, which is equivalent to less than 40 km baseline. An exponential elevation dependent model [7] is used to incorporate higher noise and larger multipath errors at lower elevations case. In order to consider the satellite geometry variation, samples from different locations and time slots are simulated. An global covered 15° by 15° grid network is simulated and 24 h observations with 30 min interval on 22 April 2013 is generated. There are 12672 group observation equations in total, only the strong model (integer bootstrapping success rate higher than 85 %) is selected and investigated in the simulation. Four group of the experiments are carried out to compare the impact of frequency number, observation noise, ionosphere variance on the ratio test and difference test.

The simulation results are presented in Fig. 26.6. The left panel shows the success rate difference between the ratio test and the difference test with different models and the same failure rate tolerance (1.0 %). The results show that the difference test outperforms ratio test in these particular model and the performance

seems independent from the frequency number, observation noise and ionosphere variance. The right panel shows the success rate difference between the two tests with different failure rate tolerances and same underlying model. The configuration used in the right panel is dual frequency, $\sigma_{z\phi} = 1$ mm and $\sigma_{zI} = 2$ cm. The simulation results indicate that the success difference between these two tests depends on failure rate tolerance. Smaller failure rate tolerance leads to a larger discrepancy between the success rates. The success rate discrepancy between the ratio test and the difference test reaches up to 20 % when the failure rate tolerance is 0.1 %, while the maximum success rate discrepancy is about 8 % for the $P_f = 1\%$ case. The overall comparison uses 5 % as failure rate tolerance, and the corresponding success rate difference is less than 2 %, which agrees well with the conclusion from this figure.

26.5 Concluding Remarks

This paper compared two popular ambiguity acceptance tests from four aspects: the definition, the acceptance region, the probability basis and their performance. Based on the probability analysis, the limitations of these two tests are identified. In addition to theoretical analysis, extensive numerical comparison are made based on the simulation results. The major findings of this paper are summarized as:

(1) The acceptance region for the ratio test and the difference test have different characteristics. The ratio test acceptance region is an overlap of many ellipsoids. The ratio test threshold can control both size and shape of the ratio test acceptance region. While the difference test acceptance region is constituted by several half spaces which are perpendicular to integer candidate vectors. The difference test threshold variation does not affect the orientation of half spaces.
(2) Both the difference test and the ratio test can be expressed with the normal distribution probability density function. The difference test can be viewed as a truncated version of the OIA. Especially when the model is strong enough, the difference test becomes a closed form of OIA. While the ratio test can be expressed as a exponential relationship between two probability density functions. The non-linear probability relationship in the ratio test makes the ratio test performance analysis more difficult.
(3) According to the probability analysis, two limitations of the ratio test and the difference test are identified. The first limitation is that the ratio test and the difference test may under-evaluate failure risk in weak models. In the weak model, the probability of non-adjacent integer candidates to ambiguity residual distribution $f\check{\epsilon}(x)$ cannot be ignored, while these two tests can only consider the best integer candidate $\check{a}$ and the second best integer candidate $\check{a}_2$. Consequently, the failure risk from non-adjacent integers are improperly ignored. This limitation can be mitigated by applying fixed failure rate

approach. The second limitation is the impact of $\check{a}_3$ is ignored in both strong model and weak model. Due to the non-uniqueness of $\check{a}_2$, it is possible that $\hat{a}$ have same distances to $\check{a}_2$ and $\check{a}_3$ in both strong model and weak model. In this case, probability contribution of $f_{\hat{a}(x+\check{a}_3)}$ is the same as $f_{\hat{a}(x+\check{a}_2)}$, but the ratio test and the difference test can only consider $\check{a}_2$, the $\check{a}_3$ is simply ignored although they are as important as $\check{a}_2$ sometimes. This limitation comes along with their definition and cannot be eliminated or mitigated. Only considering $\check{a}$ and $\check{a}_2$ is not always enough for ambiguity validation purpose, that is the reason why the ratio test and the difference test are sub-optimal ambiguity validation methods.

(4) From the numerical simulation results, the ratio test and the difference test have similar performance in all models with fixed failure rate approach. Both of them performs very well in all tested model. In some models, the ratio test outperforms the difference test, while in other models the difference test achieves higher success rate. There is no clear preference which one is better in the overall comparison.

(5) Particularly for the medium distance kinematic positioning model, the difference test outperforms the ratio test with the fixed failure rate approach. The superiority is more significant with a small failure rate tolerance. Another finding is that the superiority does not depends on satellite geometry, frequency number, observation noise, ionosphere variance etc. According to the simulation results, the difference test is recommended in kinematic GNSS data processing.

Acknowledgments This work is financially supported by Australia cooperative research center for spatial information (CRC-SI) project 1.01 'New carrier phase processing strategies for achieving precise and reliable multi-satellite, multi-frequency GNSS/RNSS positioning in Australia'. Large scale simulation in this research is supported by QUT High performance computing facilities. This work is finalized during the first author staying at TU Delft with assistant professor Sandra Verhagen as his host. Valuable comments from Professor Peter J. G. Teunissen in Curtin University improve the quality of the manuscript significantly. All above support is gratefully acknowledged.

References

1. Abidin HZ (1993) Computational and geometrical aspects of on-the-fly ambiguity resolution. Technical Report 164, UNB
2. Chang X, Yang X, Zhou T (2005) mLAMBDA: a modified LAMBDA method for integer least-squares estimation. J Geodesy 79(9):552–565
3. De Jonge P, Tiberius C (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects. Technical Report 12, Delft Geodetic Computing Centre, Delft
4. Euler H-J, Goad CC (1991) On optimal filtering of gps dual frequency observations without using orbit information. J Geodesy 65(2):130–143
5. Frei E, Beutler G (1990) Rapid static positioning based on the fast ambiguity resolution approach FARA: theory and first results. Manuscripta geod 15(6):325–356
6. Han S (1997) Quality-control issues relating to instantaneous ambiguity resolution for real-time GPS kinematic positioning. J Geodesy 71(6):351–361

7. Jin X, Jong CDD, de Jong CD (1996) Relationship between satellite elevation and precision of GPS code observations. J Navig 49(02):253–265
8. Landau H, Euler H.-J (1992) On-the-fly ambiguity resolution for precise differential positioning. In: Proceedings of ION GPS 1992, pp 607–613
9. Li B, Verhagen S, Teunissen P (2013) GNSS integer ambiguity estimation and evaluation: LAMBDA and Ps-LAMBDA. In: Proceedings on China satellite navigation conference (CSNC) 2013, vol 244. Springer-Verlag, Berlin, pp 291–301
10. Li T, Wang J (2014) Analysis of the upper bounds for the integer ambiguity validation statistics. GPS solutions 18(1):85–94, Springer
11. Neyman J, Pearson ES (1933) On the problem of the most efficient tests of statistical hypotheses. Philos Trans R Soc Lond Ser A 231:289–337 Containing papers of a mathematical or physical character
12. Odijk D (2000) Stochastic modelling of the ionosphere for fast GPS ambiguity resolution. In: Geodesy beyond 2000, Springer, p 387–392
13. Odijk D (2010) What does geometry-based and geometry-free mean in the context of GNSS. Inside GNSS 3(2):22–27
14. Teunissen PJG (1999) An optimality property of the integer least-squares estimator. J Geodesy 73(11):587–593
15. Teunissen PJG (2002) The parameter distributions of the integer GPS model. J Geodesy 76(1):41–48
16. Teunissen PJG (2003) A carrier phase ambiguity estimator with easy-to-evaluate fail rate. Artif Satell 38(3):89–96
17. Teunissen PJG (2003) Towards a unified theory of gnss ambiguity resolution. J Global Position Syst 2(1):1–12
18. Teunissen PJG (2005) GNSS ambiguity resolution with optimally controlled failure-rate. Artif Satell 40(4):219–227
19. Teunissen PJG, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. Surv Rev 41(312):138–151
20. Tiberius C, De Jonge P (1995) Fast positioning using the LAMBDA method. In: Proceedings of 4th international conference on differential satellite systems, Citeseer, p 1–8
21. Verhagen S (2005) The GNSS integer ambiguities: estimation and validation. PhD thesis, Delft University of Technology
22. Verhagen S, Teunissen PJG (2006) New global navigation satellite system ambiguity resolution method compared to existing approaches. J Guid Control Dyn 29(4):981–991
23. Verhagen S, Teunissen PJG (2006) On the probability density function of the GNSS ambiguity residuals. GPS Solutions 10(1):21–28
24. Verhagen S, Teunissen PJG (2013) The ratio test for future GNSS ambiguity resolution. GPS solutions 17(4):535–548, Springer
25. Wang L, Feng Y (2013) Fixed failure rate ambiguity validation methods for GPS and COMPASS. In: Proceedings China satellite navigation conference (CSNC) 2013, vol 2. Springer-Verlag,Berlin, pp 396–415
26. Xu P (2006) Voronoi cells, probabilistic bounds, and hypothesis testing in mixed integer linear models. IEEE Trans Inf Theory 52(7):3122–3138

Chapter 27
A Multi-Step Multi-Order Numerical Difference Method for Traveling Ionospheric Disturbances Detection

Long Tang and Xiaohong Zhang

Abstract In this paper, we developed a multi-step multi-order numerical difference method to detect the Traveling ionospheric disturbance (TID). This method can eliminate TEC trend effectively and has a wide span of detectable periods. In this study, the principle of the multi-step multi-order numerical difference method is presented firstly. Then, we execute three experiments with the simulated data and real data to test the feasibility of the method: the first experiment using the simulated TID signal without the TEC trend, the second one with real TEC observations but simulated TID signals and the last one having real TEC observations and TID waves. The experimental results demonstrate (1) detected TID wave has the same frequency compared to the true signal and the ratio of amplitude between them is also consistent with the theoretical result; (2) distinct to first-order method, the second-order difference process can eliminate trend term effectively even when TEC varies drastically; (3) this method can detrend TEC and extract TID signals simultaneously and can detect multiple disturbances with various periods.

Keywords GPS · Total electron content · Traveling ionospheric disturbance · Multi-step multi-order numerical difference method

27.1 Introduction

Traveling ionospheric disturbance (TID) is wave-like fluctuations of the plasma that propagate through the ionosphere at a wide range of velocities and frequencies [1]. It is important to monitor ionospheric disturbance and clarify their characteristics for

L. Tang (✉) · X. Zhang
School of Geodesy and Geomatics, Wuhan University, Wuhan, China
e-mail: ltang@whu.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_27,

high precision GPS positioning and researching on some geophysical events. On the one hand, TIDs constitute the major part of irregularities in the plasma that can cause decrease on the performance of precise position and navigation [2]; on the other hand, the TID signals induced by some geophysical events, such as earthquakes, tsunamis and underground nuclear explosions can be used to research the events in turn [3–5]. Consequently, TIDs research is a very meaningful work.

Since the 1990s, GPS total electron content (TEC) data has become the most widely used observation technique on TIDs research due to the high spatial and temporal resolution [6]. The foundational work is to adequately detect the disturbances in GPS TEC data. Kotake et al. [7] and Tsugawa et al. [8] obtained the ionospheric disturbance series by subtracting the 1-h running average from the time sequence of GPS TEC data. Ding et al. [9] and Wang et al. [10] removed the TEC background trends using 2–3 order polynomial fits with 2–3 h sliding windows. Katamzi et al. [11] used the Savitzky-Golay filter to extract the disturbances series and the duration of sliding window were also more than 1 h. These methods can acquire the TEC disturbances effectively but generally need long duration to eliminate the background trends, which is time-consuming for lager amounts of data and not suitable for real-time process.

Hernández-Pajares et al. [12] employed a simple numerical difference method with only a couple of observations to extract TIDs series that is beneficial to real-time application. However, this method can not effectively detrend the TEC series when it is varied drastically (generally with lower elevation angle). Furthermore, some TIDs can be 'neglected' due to the limited detectible period scope. In order to solve these problems, we developed a multi-step multi-order numerical difference method to detect the TIDs in this paper.

27.2 TID Detection Method

As it is well know, the geometry-free combination of the dual frequency GPS carrier phases can be used to calculate the ionospheric TEC

$$L_4 = k \cdot s + b \tag{27.1}$$

where L_4 is geometry-free combination with unit of meter; s represents the ionospheric TEC with unit of TECU; k is the conversion factor between TEC and observation ($k \approx 0.105$ m/TECU); b is an unknown constant bias for each different satellite-receiver continuous arch of data, including carrier phase ambiguity and hardware delays. Although Eq. (27.1) cannot acquire absolute value of TEC at a particular time, it can capture the TEC variation over time with high precision, which is important for TIDs detection.

Hernández-Pajares et al. [12] applied a simple numerical difference method to extract TIDs. The procedure is show in Eq. (27.2):

$$\Delta s(t) = s(t) - 0.5(s(t+\tau) + s(t-\tau)) \tag{27.2}$$

where $\Delta s(t)$ is the first-order numerical difference of TEC series; t is the observation epoch and τ is the time step. Generally, Eq. (27.2) can effectively eliminate the TEC trends for data with higher elevation cut-off angle, such as 40°. However, the redundant trends are significant in disturbances series obtained by lower elevation angle data, which is also important for some studies such as ionospheric tsunami monitoring. To further remove the trends and apply the lower elevation angle data, multi-order numerical difference is presented:

$$\Delta^n s(t) = \Delta^{n-1} s(t) - 0.5(\Delta^{n-1} s(t+\tau) + \Delta^{n-1} s(t-\tau)) \tag{27.3}$$

where n is the magnitude of difference order. In the next section, we will use experiments to demonstrate the multi-order process can further remove the residual trends in the lower elevation angle data.

For a given TID with frequency f, a complex exponential function can be used to express the wave $d(t)$:

$$d(t) = A_0 \exp(j2\pi f t) \tag{27.4}$$

where A_0 is the TID amplitude and $j = \sqrt{-1}$is the complex unit. According to Eq. (27.2), the trends of TEC are deducted, then:

$$\Delta s(t) = d(t) - 0.5(d(t+\tau) + d(t-\tau)) = 2\sin^2(\pi f \tau) d(t). \tag{27.5}$$

Similarly, we obtain the expression between $d(t)$ and $\Delta^2 s(t)$:

$$\Delta^2 s(t) = 4\sin^4(\pi f \tau) d(t) \tag{27.6}$$

Accordingly, the expression between $d(t)$ and $\Delta^n s(t)$ is as follow:

$$\Delta^n s(t) = 2^n \sin^{2n}(\pi f \tau) d(t) \tag{27.7}$$

According to Eq. (27.7), the observed disturbances all have the same frequency of f compared to the $d(t)$, suggesting that the equation is able to capture the disturbance waves correctly.

Now, we discuss the effective detectible period scope. To define the augmentation factor R, we use the ratio between the observed and true amplitudes of the TID wave.

$$R = 2^n \sin^{2n}\left(\frac{\pi}{T}\tau\right) \tag{27.8}$$

where $T = 1/f$ is the period of the wave. The method is very effective when R is larger than 1, whereas the sensibility will degrade if it is smaller than 0.4. For a given τ, the relationship between R and T for second-order numerical difference is shown in Fig. 27.1.

As seen from Fig. 27.1, the detectable span of disturbance period is closely associated with the time step and is generally limited for a specific single time

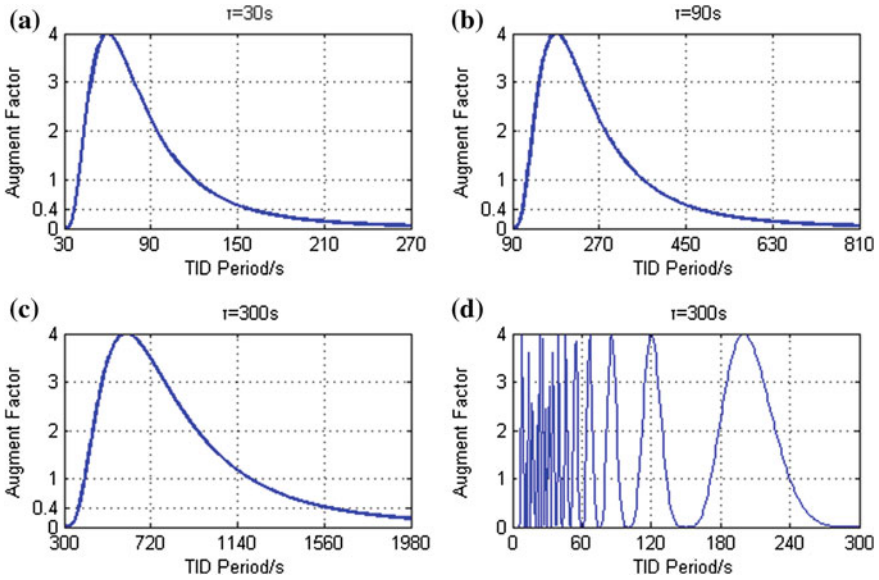


Fig. 27.1 The relationship between augmentation factor and TID period for different time steps

step. For example, a span with an augmentation factor larger than 0.4 is only 40 to 160 s (this is the similar when the magnitude of difference order is set to 1) for $\tau = 30$ s, which may overlook other important TIDs. To solve this problem, we use multiple time steps to replace the single time step to extend the scope of the detectable period. According to Fig. 27.1, the ranges for augmentation factors larger than 0.4 where $\tau = 30$ s, $\tau = 90$ s and $\tau = 300$ s are 0.7–2.6, 1.8–7.8 and 6.0–26.3 min, respectively. Therefore, the combined range is 0.7–26.6 min after applying the multi-step process, which is enough for the detection of most TIDs induced by acoustic-gravity waves. Furthermore, majority of period during this span have augmentation factor larger than 1, which is sensitive for weaker signals. It is noted that for TID with period less than the time step, such as $\tau = 300$ s in Fig. 27.1d, the bandwidths are very narrow and can be negligible.

In general, the magnitude of difference order is set as 2, or 3 with shorter time step, because too high order will increase the workload and amplify the background noise when the time step is larger such as 300 s.

27.3 Experiments and Analysis

In this section, we conduct three experiments with the simulated data and real data to test the feasibility of the multi-step multi-order method for TID detection. The experimental programs are showed in Table 27.1.

Table 27.1 Experimental programs to test the feasibility of the multi-step multi-order numerical difference method

Experiments	Data	
	Background trends	TID signal
1	–	Simulated
2	Real	Simulated
3	Real	Real

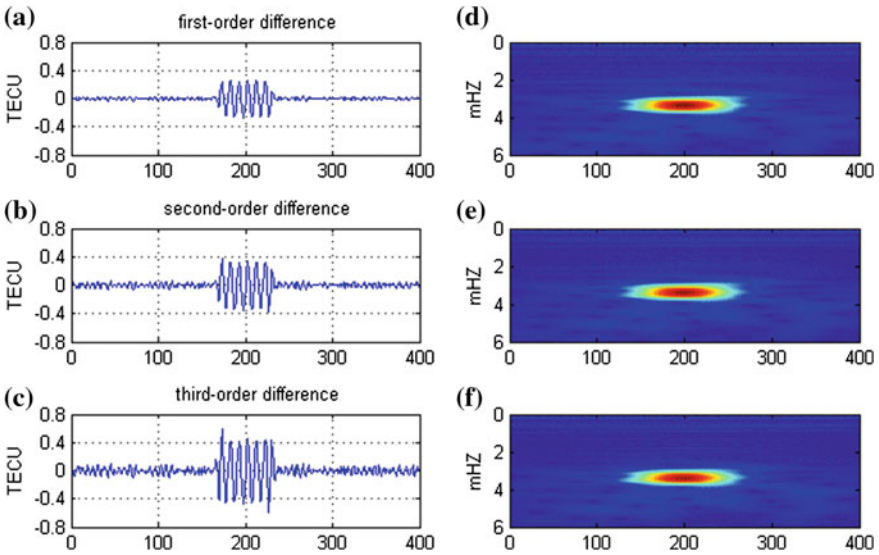


Fig. 27.2 The detection results for simulated TID signal in experiment one. The left panels are disturbances time series and the right panels are corresponding time-frequency diagrams

27.3.1 Experiment One

In experiment one, we use the detection method to process the simulated TID signal without TEC background trends. There are total 400 epochs and the sampling interval is 30 s in accordance with the standard GPS observation data. The data are composed of TID signal with period of 5 min (frequency is 3.3 mHZ) and amplitude of 0.20 TECU during 170–230 epochs, plus random noise with amplitude less than 0.02 TECU during all the epochs. According to Fig. 27.1, the time step is set as 90 s.

Figure 27.2 shows the processing results: panels (a)–(c) are disturbances time series by first-order, second-order and third-order differences, respectively; panels (d)–(f) are corresponding time-frequency diagrams. As seen from the figure, there are significant disturbances during 170–230 epochs in the time series and the amplitudes of disturbances for panel (a) to panel (c) are approximately 0.25 TECU, 0.33 TECU and 0.43 TECU, respectively. According to Eq. (27.8), the augmentation factors are 1.31, 1.71 and 2.24 for different orders. Then the true

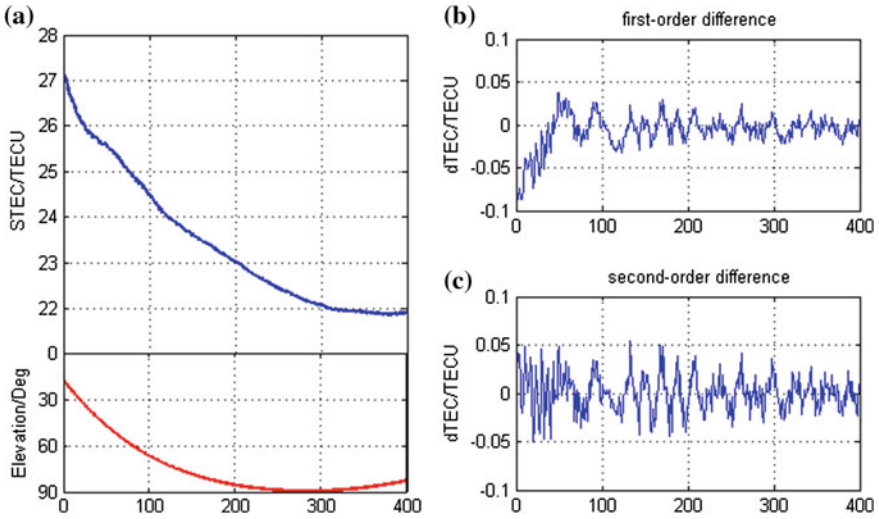


Fig. 27.3 The real STEC time series (panel **a**) and corresponding disturbances time series by first-order difference process (panel **b**) and second-order difference process (panel **c**)

amplitudes are 0.19 TECU, 0.19 TECU and 0.19 TECU, which is approximately equal to the simulated signal. The centre frequencies of the three time-frequency diagrams are all 3.3 mHZ and similar to the primitive signal. These results are consistent with theoretical results indicating the validity of detection method.

27.3.2 Experiment Two

In experiment two, we use the real observed GPS TEC data without TID waves to check the ability of the detection method to eliminate the background trends firstly; then two simulated TID signals with different periods are exerted on the primitive TEC series for the purpose of verifying the capacity of the detection method to capture the TID signals with background trends.

Figure 27.3a shows an arc of slant TEC (STEC) data of one GPS station obtained by the phase-levelling code method [13]. The satellite elevation angles are approximately varied from 20 to 90°. Corresponding disturbances time series by first-order difference with time step 300 s are given in Fig. 27.3b. The results show that residual trends are significantly in epochs with lower elevation angles that the TEC changes drastically. Then we change the strategy to second-order difference and the residual trends are eliminated (see Fig. 27.3c), suggesting that the second-order difference process can effectively eliminate trend term even when TEC varies drastically compared to the first-order difference method.

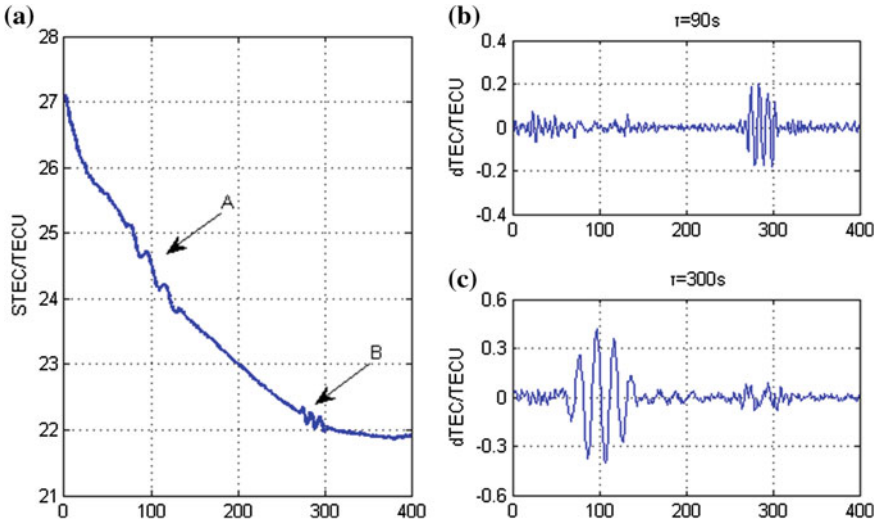


Fig. 27.4 The real STEC time series plus simulated TID signals (panel **a**) and corresponding disturbances time series by second-order difference process with time step 90 s (panel **b**) and 300 s (panel **c**)

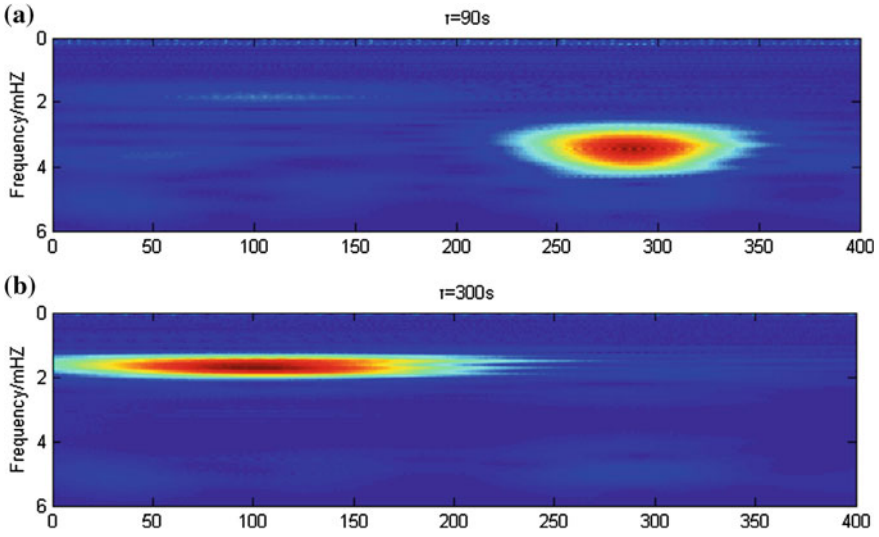


Fig. 27.5 The time-frequency diagrams of disturbances time series by second-order difference process with time step 90 s (panel **a**) and time step 300 s (panel **b**)

Two simulated TID signals are exerted on the primitive STEC time series (see Fig. 27.4a): TID 'A' is during 60–140 epochs with period 10 min (frequency is 1.67 mHZ) and amplitude 0.1 TECU, while TID 'B' is during 270–300 epochs

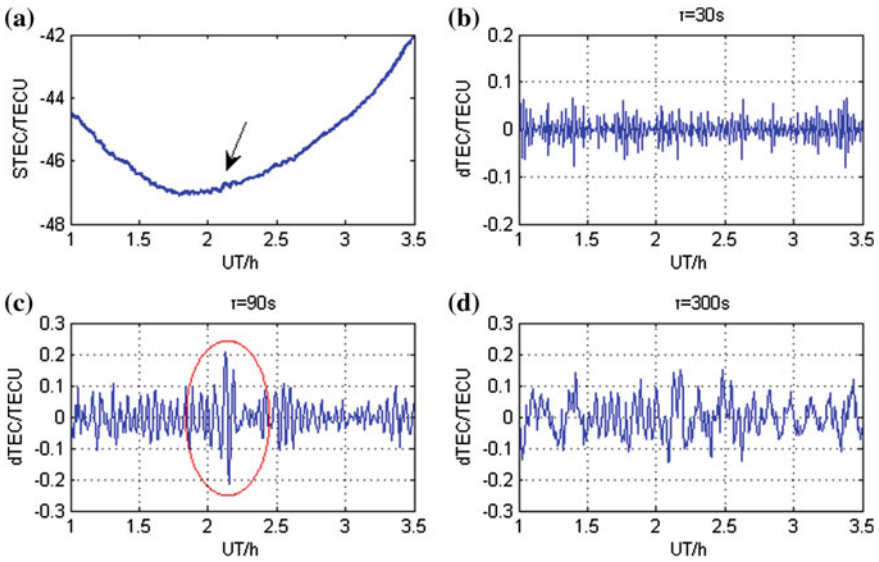


Fig. 27.6 The real STEC time series with TID signal (panel **a**) and corresponding disturbances time series by second-order difference process with time step 30 s (panel **b**), time step 90 s (panel **c**) and 300 s (panel **d**)

with period 5 min and amplitude 0.1 TECU. Figure 27.4b shows the disturbances time series by second-order difference process with time step 90 s. Obviously, there is a TID signal during 270–300 epochs with amplitude approximately 0.18 TECU and period 5 min (frequency 3.3 mHZ showing in Fig. 27.5a). According to Eq. (27.8), the augmentation factor is 1.7 and true amplitude is 0.1 TECU indicating the detected signal is TID 'B'. Similarly, Figs. 27.4c and 27.5b show the TID 'A' can also be captured when the time step is set to 300 s. These results suggest that the detection method can detrend background TEC and extract TIDs signal simultaneously, and the multi-step process can capture multiple TID signals with different periods while the single-step method may ignore TID signal with period outside the detectible scope.

27.3.3 Experiment Three

In experiment three, we use the multi-step multi-order numerical difference method to detect the real TID signal. The magnitude of difference order is set as 2. Figure 27.6a shows the arc of STEC data measured by GPS PRN21 on IGS station AIRA in DOY 145, 2009. The arrow denotes there is a TID signal around 2:00 UT but the period of the TID is unknown in advance. So, the multi-step process is necessary.

The disturbances time series extracted by multi-order numerical difference method are displayed in Fig. 27.6b–d with time step 30, 90 and 300 s, respectively. As can be seen, TEC background trends are eliminated effectively by all the time steps whereas only the strategy with time step 90 s captures the TID signal(see the ellipse in Fig. 27.6c). If the time step was simply set as single value such as 300 s [12] or 30 s, the TID signal wound be overlooked. Similar to previous experiments, the process of real TID signal illustrate the multi-step multi-order numerical difference method can effectively omit the background trends and capture the signal correctly.

27.4 Summary

This study develops a multi-step multi-order numerical difference method to detect TID signals using GPS observations. According to theoretical analysis, extracted TID signals have the same frequency consistent with the primitive signal, suggesting that the detection method is able to capture the disturbance waves correctly. The multi-step process has a wider period spectrum for TID detection compared to the previous single-step method, which is enough for the detection of most TIDs induced by acoustic-gravity waves. Furthermore, majority of period during this span have augmentation factor larger than 1, which is sensitive for weaker signals. To test the feasibility of the method, we conduct three experiments with the simulated data and real data. The experimental results are in accordance with theoretical analysis and show the multi-order difference process can effectively eliminate trend term even when TEC varies drastically compared to the previous first-order difference method.

The theoretical and experimental results demonstrate the capacity of the multi-step multi-order numerical difference method to detect TID signals. The method needs only a couple of observations to extract TIDs series that can significantly reduce the time when process large mounts of data, which is suitable for real-time application. In the future, it can be used to real-time monitor some geophysical events such as tsunami.

Acknowledgment This study was supported by National 973 Project China (Grant No. 2013CB733301) and National Natural Science Foundation of China (Grant No. 41074024, No. 41204030).

References

1. Tsugawa T, Kotake N, Otsuka Y et al (2007) Medium-scale traveling ionospheric disturbances observed by GPS receiver network in Japan: a short review. GPS Solutions 11(2):139–144
2. Lejeune S, Wautelet G, Warnant R (2012) Ionospheric effects on relative positioning within GPS dense network. GPS Solutions 16(1):105–116

3. Liu JY, Tsai HF, Lin CH et al (2010) Coseismic ionospheric disturbances triggered by the Chi-Chi earthquake. J Geophys Res 115:A8303
4. Artru J, Ducic V, Kanamori H et al (2005) Ionospheric detection of gravity waves induced by tsunamis. Geophys J Int 160:840–848
5. Yang YM, Garrison JL, Lee SC (2012) Ionospheric disturbances observed coincident with the 2006 and 2009 North Korean underground nuclear tests. Geophys Res Lett 39:L02103
6. Beutler G, Rothacher M, Schaer S et al (1999) The international GPS service (IGS): an interdisciplinary service in support of earth sciences. Adv Space Res 23(4):631–653
7. Kotake N, Otsuka Y, Ogawa T et al (2007) Statistical study of medium-scale traveling ionospheric disturbances observed with the GPS networks in Southern California. Earth Planets Space 59:95–102
8. Tsugawa T, Otsuka Y, Coster A et al (2007). Medium-scale traveling ionospheric disturbances detected with dense and wide TEC maps over North America. Geophys Res Lett 34:L22101
9. Ding F, Wan W, Ning B et al (2012) Two-dimensional imaging of large-scale traveling ionospheric disturbances over China based on GPS data. J Geophys Res 117:A08318
10. Wang M, Ding F, Wan W et al (2007) Monitoring global traveling ionospheric disturbances using the worldwide GPS network during the October 2003 storms. Earth Planets Space 59:407–419
11. Katamzi Z, Smith N, Mitchell C et al (2011) Statistical analysis of travelling ionospheric disturbances using TEC observations from geostationary satellites. J Atmos Solar Terr Phys 74:64–78
12. Hernández-Pajares M, Juan M, Sanz J et al (2006). Medium-scale traveling ionospheric disturbances affecting GPS measurements: spatial and temporal analysis. J Geophys Res 111:A07S11
13. Mannucci AJ, Wilson BD, Yuan DN et al (1998) A global mapping technique for GPS-derived ionospheric total electron content measurements. Radio Sci 33(3):565–582

Chapter 28
Extraction of Ocean Tide Semidiurnal Constituents' Vertical Displacement Parameters Based on GPS PPP and Harmonic Analysis Method

Hong Zhao, Qin Zhang, Ming Zhang, Shuangcheng Zhang, Wei Qu, Rui Tu and Zhi Liu

Abstract Using GPS data of 4 CORS stations in Hong Kong 2012, four ocean tide semidiurnal constituents' vertical displacement parameters of every station were calculated. Firstly, 4 stations' displacement time series were extracted by GPS PPP algorithm. Next, combined with harmonic analysis method, the vertical displacement parameters of semidiurnal constituents were obtained. Simultaneously, in order to evaluate accuracy and reliability of ocean loading displacement parameters determined by PPP technology, we compared the results with ocean loading displacement parameters provided by HAMTIDE11A.2011 global ocean tide model refined by local tidal data of eastern and southern sea of China and the difference was less than 1 mm. Therefore, PPP techniques can be used for measuring ocean tide loading displacement parameters, and it has implications for precision modeling of tides and high-precision geodetic measurements.

Keywords GPS technology · PPP · Harmonic analysis · Ocean loading displacement parameters

28.1 Introduction

Earth's tidal force originated by the sun and the moon produces periodic deformation of the solid Earth and causes sea load change periodically, which is called as ocean tide loading. Many scholars have studied that ocean tide loading has

H. Zhao (✉) · Q. Zhang · M. Zhang · S. Zhang · W. Qu
College of Geology Engineering and Geomatics, Chang'an University, Xi'an 710054, China
e-mail: zhaohong710@163.com

R. Tu
GFZ German Reasearch Centre for Geosciences, Potsdam, Germany

Z. Liu
China JK Institute of Engineering Investigation and Design, Xi'an 710054, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_28,

impact on the geophysical field observations. Zhou Xuhua researched high-precision geodetic loading correction, discovered ocean tide loading effects cannot be ignored, especially on high-precision observations, it is necessary to establish the appropriate loading correction model [1]; Zhou Jiangcun studied the effect of ocean tide loading on GPS baselines and had the result of ocean tide loading effect on stations is mm, in the vertical displacement of the coastal stations the effect reached centimeter level [2]; Zhao Hong investigated ocean tide loading effect on GPS PPP, founded the affect on GPS positioning of coastal stations is 4–5 times of inland area [3].

Ocean loading effect correction depends on tidal wave parameters provided by global ocean models, and precision altimetry data is the foundation of global ocean models. Although satellite altimetry can offer more precise characteristics of the ocean tides in given area, but due to the special geographical coastal area shoreline structures and complex seabed topography, thus the accuracy and resolution of global ocean models obtained by satellite altimetry are limited and cannot meet the desired specifications IERS accuracy. At present, scientists are seeking the latest geodetic means to directly measure displacement of ocean tide constituents which accuracy is better than the model estimates. In the recent years, many scholars use gravity measurements, VLBI, GPS and other technologies to research the effects of ocean tide loading [4].

Compared with gravity measurements and VLBI technique, GPS technology in ocean loading effects research has advantage of global coverage, all-weather, low cost et al. M. Vergnolle used DGPS to get the displacement parameters of ocean tide loading in France [5], Yuan Linguo researched OTL effect based on GPS and obtained the ocean tide constituents' displacement parameters of CORS stations in HONG KONG [6]. These scholars use GPS technology to get the OTL displacement parameters and compared with the latest global ocean model. Thus using GPS technology to get OTL displacement parameters is feasible.

The accuracy of GPS precise point positioning has reached millimeter-level, and the effect of ocean tide loading on stations is centimeter level, so using the high accuracy GPS technology to get the long time series of coordinate values of stations is one of the best means to compute the displacement parameters of ocean tide constituents. But GPS precise point positioning technology can be affected by various factors such as error and noise, it is necessary to use harmonic analysis method to correct and denoising. This paper used PPP technology and harmonic analysis method to determine the vertical displacement parameters of ocean tide semidiurnal constituents, and compared with global ocean tide model refined by high accuracy local tide data of China East and South Sea, then discussed the feasibility and reliability of using GPS technology to monitor ocean loading displacement parameters.

28.2 Principle of Harmonic Analysis Method to Extract the Parameters of Tide Constituents

Harmonic analysis method is based on the value of the time series of station coordinates acquired by GPS precise point positioning technology and ocean tide constituents' period computed by Fast Fourier transform on coordinate time series, then establish a polynomial model, using least squares to solve ocean loading displacement parameters. Polynomial model as shown in Formula (28.1):

$$y(t) = b_0 + b_1 t + \sum_{k=1}^{N} A_k \cos(\delta_k t) + B_k \sin(\delta_k t) \tag{28.1}$$

$$\begin{cases} A_k = \hat{a}'_k + \hat{a}'_{-k} \\ B_k = \hat{a}'_k - \hat{a}'_{-k} \end{cases} \tag{28.2}$$

$y(t)$ is time series of station coordinates, k represents the ocean tide constituents, b_0, b_1, A_k, B_k is the unknown parameters, δ_k is the angular velocity of each tidal waves, $\delta_k = 2\pi f u_k$, fu_k is frequency of each tide constituent and known value (provided by FFT).

The unknown parameter b_0, b_1, A_k, B_k in Formula (28.1) solved by least squares fit, and then substitute the result into Eq. (28.2) so we can get parameters $\hat{a}'_k, \hat{a}'_{-k}$. But should consider the impact of two factors:

(1) The observation noise, using colored noise or white noise model to be correct $\hat{a}'_k, \hat{a}'_{-k}$. Correction formula is as follows:

$$\begin{aligned} \hat{a}_k &= \hat{a}'_k + \Delta_{noise} \\ \hat{a}_{-k} &= \hat{a}'_{-k} + \Delta_{noise} \end{aligned} \tag{28.3}$$

(2) Effects of modulation effects, Effects of the station for the satellite modem on station is should be considered, f_k, u_k are impact on the amplitude and phase, respectively, formula is as follows:

$$\hat{a}_k e^{i\delta_k t} = f_k a_k e^{i\delta_k t + i u_k} \tag{28.4}$$

Computing Formula (28.4) in two steps, first is to solve each satellite astronomical for each tidal waves' node phase modulation, amplitude modulation node, that's f_k, u_k, (this paper select the correction factors of middle time of the time series), second is to compute a_k, a_{-k}.

Through Formula (28.5) we can get the final ocean loading displacement parameters (tidal wave amplitude and phase values):

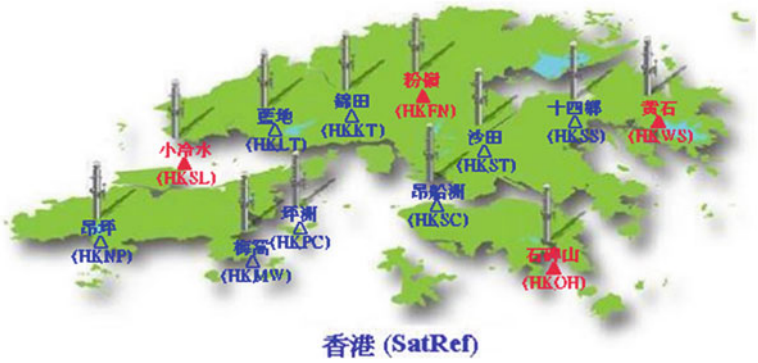


Fig. 28.1 Distribution of HONG KONG CORS stations

$$\begin{aligned} L_k &= |a_k| + |a_{-k}| \\ \theta_k &= \frac{ang(a_k) + ang(a_{-k})}{2} \bmod 180 \\ g_k &= v_k - ang(a_k) + \theta_k \end{aligned} \tag{28.5}$$

L_k, g_k is final amplitude and phase of tide constituent, v_k is sum of Astronomy phase and phase modulation.

28.3 Data Process and Discussion

28.3.1 Data Selection

This paper uses data of 2012 of 4 CORS in Hong Kong—HKOH, HKWS, HKSL, HKFN, Data sampling interval is 5 s. Station distribution is shown in Fig. 28.1.

Figure 28.2 shows contour of the sum of the vertical OTL displacement amplitude predicted using the NAO99b model of the eight major diurnal and semi-diurnal constituents ($M_2, S_2N_2, K_2, K_1, O_1, P_1, Q_1$). From Fig. 28.2 we can know the OTL effect on Hong Kong is up to several centimeters, it is necessary to research ocean tide loading in Hong Kong.

28.3.2 Data Processing

Data were processed by Precise Point Positioning software package developed by the research group which static accuracy is mm, dynamic positioning accuracy is

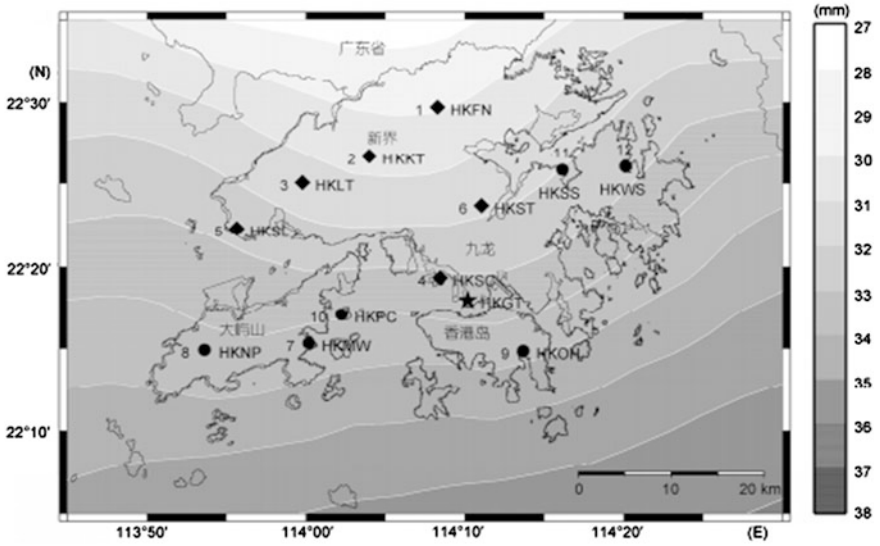


Fig. 28.2 Contour map of ocean loading displacement amplitude of NAO99b model of eight major tidal waves in Hong Kong [4]

1–2 cm [7] (Data processing has been added the troposphere, tidal, and other corrections, OTL correction is not added in order to obtain the real effects of ocean loading). Extracting tidal parameters is as follows:

The first step, PPP solving the displacement time series of stations timely;
The second step, in order to deduct frequency noise, low pass filter is carried out on displacement time series, then Fast Fourier transform is used to get period of ocean tide constituents.
The third step is to use of harmonic analysis, least squares and solve the correction coefficient.
The fourth step is decomposing the coefficients to compute amplitude and phase.

Through the above steps, we can get the amplitude and phase of each tide constituent. Regard amplitude provided by Hamtide.2011a global ocean tide model which refined by high accuracy local tide data of China East and South Sea as reference values, we can analysis feasibility of using GPS technology and harmonic analysis to determine ocean loading displacement parameters.

28.3.2.1 Result and Discussion

After the first step on 4 CORS data of 2012 respectively, coordinate time series of each station can be got. In order to obtain the time-series variation in the vertical direction (actual ocean tide effect), we should exclude abnormal points (excluding

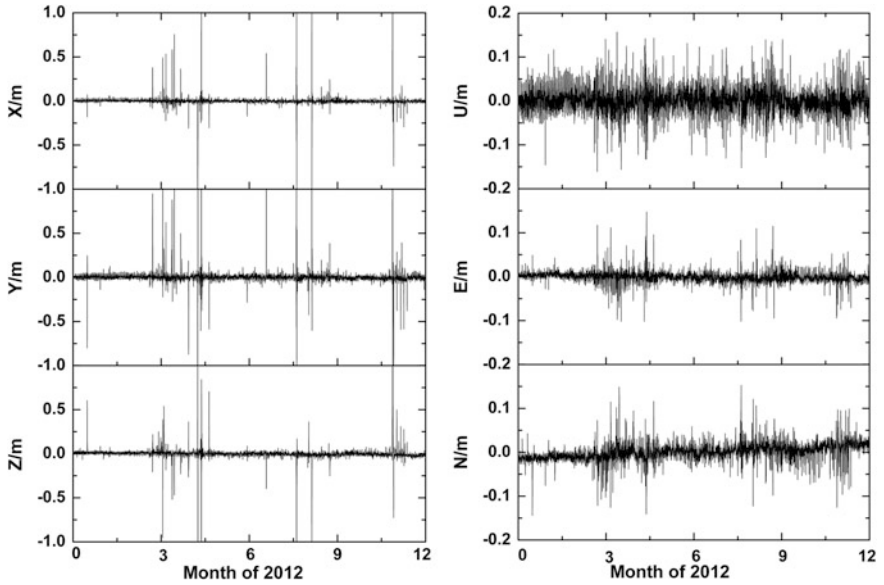


Fig. 28.3 Original and processed time series values of HKFN station

outliers larger error), coordinate conversion and other operations. Original and processed time series are shown in Fig. 28.3, we take HKFN station as an example. Compared observed ocean tide with predicted vertical OTL displacement amplitude computed by the refinement Hamtide.2011a model [3]. Ocean tide loading comparison of measured and modeled effects is shown in Fig. 28.4.

Measured and modeled effects of ocean tide loading comparison of 4 CORS of the first 40–60 days of 2012 is shown in Fig. 28.4. It can be seen that: measured ocean tide loading by PPP technology is bigger than predicted by ocean tide model. Generating the phenomena is because PPP is affected by multipath effects and various errors. We can also see that is observed ocean tide loading effect which has diurnal and semi-diurnal period changes. To separate the different cycles of the tide constituent would require the use of spectral analysis method.

28.3.2.2 Spectrum Analysis to Detect Tide Constituents' Period

Four CORS stations' vertical OTL effect values got by GPS technology were processed by two-step: firstly in order to deduct frequency noise, low pass filter is carried out on displacement time series; secondly Fast Fourier transform is used to get period of ocean tide constituents. Simultaneously, Fast Fourier transform is also used of predicted OTL value which using the refinement Hamtide.2011a model of the eight major diurnal and semi-diurnal constituents. Observed and

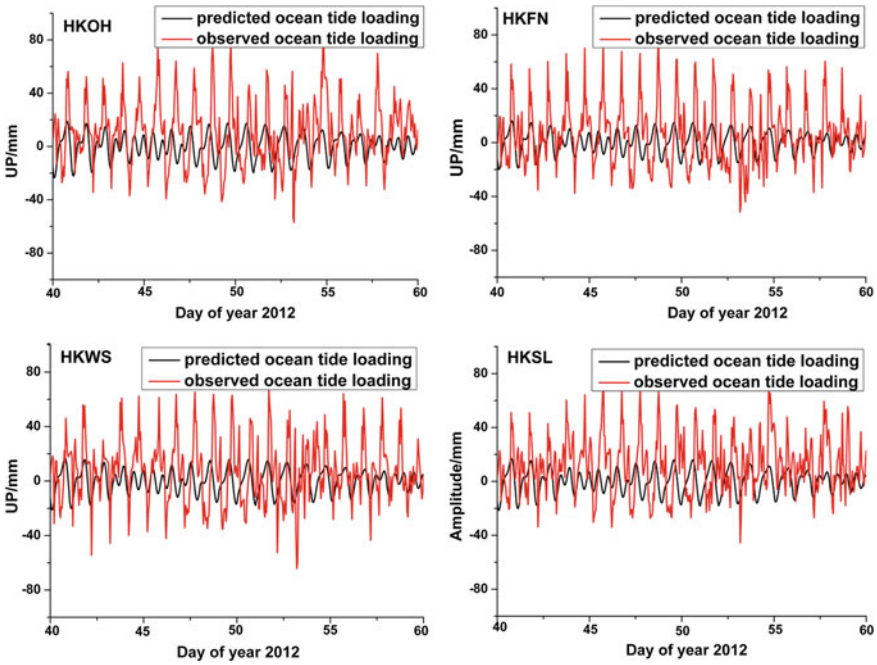


Fig. 28.4 Measured and modeled effects of ocean tide loading comparison

predicted tidal amplitude spectrum analysis result comparison chart of each station is shown in Fig. 28.5.

It is seen from Fig. 28.5 that GPS OTL vertical displacement and model predicted OTL vertical displacement through Fast Fourier transform can detect period and amplitude of each ocean tide constituent. And period and amplitude through above is consistent (Period detected by FFT and fixed period of tide constituent are listed in Table 28.1). There is minor difference between tide constituents with 12 h period. But there is big difference between tide constituents with 24 h period. The reason of why arises big difference is that solving process cannot deduct the satellite orbit errors and multipath influence of 24 h period. There is an effective way of sidereal day filter to eliminate bigger influence produced by various errors.

28.3.2.3 Harmonic Analysis to Compute Amplitude of Tide Constituent

We substitute 4 stations' GPS OTL displacements and period by FFT of each semi-diurnal constituent into Formula (28.1), after correction by Formula (28.3) and Formula (28.4), through Formula (28.5) we can get vertical displacement amplitude of each semi-diurnal constituent. The GPS OTL amplitude with

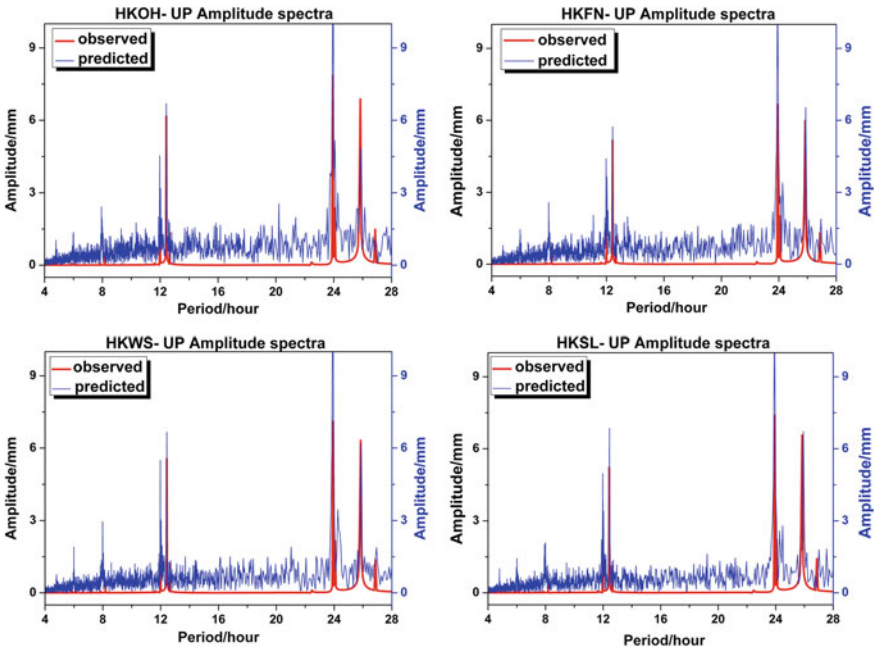


Fig. 28.5 Observed and modeled tidal amplitude spectrum analysis result comparison chart

Table 28.1 Tidal wave period detected by spectrum analysis compared with the tidal wave of natural period

Constituents	Natural period(h)	Period detected by FFT(h)
M2	12.42	12.42
S2	12.00	12.01
N2	12.06	12.06
K2	11.98	11.99

refinement model amplitude is compared and results listed in Table 28.2. We take M2 constituent as example to show difference between refinement model and GPS-derived parameters in Fig. 28.6.

It is seen from Table 28.2 and Fig. 28.6 that after using spectrum analysis and harmonic analysis with various corrections on displacement time series of stations derived by GPS PPP technique, we can get amplitude of each semi-diurnal tide constituent .By comparing with refinement Hamtide.2011a model, we have found that the difference between them is less than 1 mm. It demonstrated that PPP-derived OTL parameters have a good consistent with high accuracy ocean tide model.

Table 28.2 Comparison of GPS vertical displacement with the refinement model of semi-diurnal constituents

Station	M_2/mm		S_2/mm		N_2/mm		K_2/mm	
	Model	Harmonic analysis	Model	Harmonic analysis	Model	Harmonic analysis	Model	Harmonic analysis
HKOH	6.67	6.66	1.90	2.48	1.41	1.43	0.58	0.75
HKWS	6.00	6.55	1.7	1.19	1.3	1.24	0.50	0.81
HKSL	5.64	5.74	1.53	1.60	1.18	1.28	0.46	0.71
HKFN	5.58	5.63	1.54	1.40	1.2	1.66	0.44	0.76

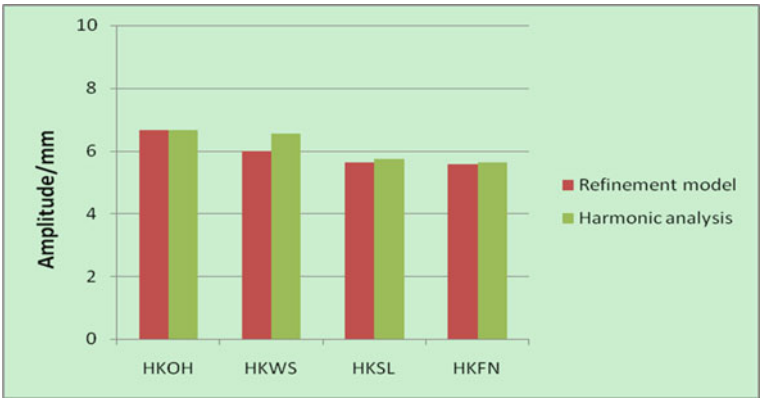


Fig. 28.6 GPS vertical displacement parameters and refinement model comparison of M2 constituent

28.4 Conclusions

For the purpose of making a thorough inquiry about feasibility and reliability of using GPS technology to monitor ocean loading displacement parameters. Data of 2012 of 4 CORS in Hong Kong are selected. This paper used PPP technology and harmonic analysis method to determine the vertical displacement parameters of semi-diurnal ocean tide constituents, and compared with global ocean tide model refined by high accuracy local tide data of China East and South Sea. We have found that the difference between GPS OTL displacement estimates and displacement parameters provided by refinement global ocean tide model is sub-millimeter. The study demonstrated that GPS is capable of estimating the OTL displacement and provide a new way for precision geodetic measurements.

References

1. Zhou XH, Wu B, Li J et al (2001) Ocean tide displacement corrections in high precision geodesy. Acta Geodaet Et Cartograph Sin 30(4):327–330
2. Zhou JC, Sun HP (2005) Effect of ocean tide loading on GPS baseline measurement. J Geodesy Geodyn 25(4):27–32
3. Zhao H, Zhang Q, Huang GW et al (2012) Effect of ocean tide loading on GPS precise positioning based on different ocean tide models. J Geodesy Geodyn 32(5):108–112
4. Yuan LG, Ding XL, Sun HP et al (2010) Determination of ocean tide loading displacements in Hong Kong using GPS technique. Earth Sci 40(6):699–714
5. Vergnolle M, Bouin M-N, Morel L (2008) GPS estimates of ocean tide loading in NW-France: determination of ocean tide loading constituents and comparison with a recent ocean tide model. Geophys J Int 173:444–458. doi:10.1111/j.1365-246X.2008.03734.x
6. Yuan LG, Ding XL, Zhong P (2009) Estimates of ocean tide loading displacements and its impact on position time series in Hong Kong using a dense continuous GPS network. J Geod 83:999–1015. doi:10.1007/s00190-009-0319-0
7. Rui T, Maorong G, Hongping Z, Guanwen H (2012) The realization and convergence analysis of combined PPP based on raw observation. Elsevier. http://dx.doi.org/10.1016/j.asr.2013.03.005

Chapter 29
Estimate DCB of BDS Satellites Based on the Observations of GPS/BDS

Chuanjun Wei, Qin Zhang, Lihong Fan, Shuangcheng Zhang, Guanwen Huang and Kejie Chen

Abstract In single frequency precise point positioning, satellite's DCBs is one of the main factors impacting the positioning accuracy. In the consideration of BDS navigation system has not completed and the limited distribution of regional monitoring network, these adverse condition make it is hard to estimate the DCBs accuracy. So put forward using GPS rice data to estimate the DCB of BDS satellite by make a global ionosphere model using the observations of GPS and BDS. In this paper the strategy of estimate the DCB of BDS satellite is discussed in detail. Further analysis showed that the stability of BDS satellite's DCBs is about 0.15 ns at 1 month. In order to further validate the precision of DCB of BDS satellites from this strategy, used BDS single frequency observations to position with the correct of BDS satellite's DCB or without it. Compared with no DCBs correction, single-frequency user's positioning accuracy is improved from 47% to 63% by using DCBs correction from the algorithm.

Keywords BDS · GPS · Combined Solution · Ionosphere · Differential code bias

29.1 Introduction

GNSS satellites and receivers transmitting and receiving navigation signals of different frequencies by different channels, but the time when the navigation signal generated by the different channels are not entirely consistent. The time delay

C. Wei (✉) · Q. Zhang · L. Fan · S. Zhang · G. Huang
College of Geology Engineering and Geomatics, Chang'an University,
Xi'an 710054, China
e-mail: chuanjunw2013@163.com

K. Chen
GFZ German Reasearch Centre for Geosciences, Potsdam, Germany

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_29,

between different frequencies is called Differential Code Bias (DCB). Its size in a few nanoseconds to tens of nanoseconds. It is the biggest error source of using navigation observations to calculate the Total Electron Content (TEC) of the ionosphere. Meanwhile is one of the error sources navigation and positioning by single-frequency [1]. So obtain accurate DCB parameter is very critical for improving accuracy of calculate the total electron content of the ionosphere and the single-frequency user positioning.

Currently internationally recognized GPS satellite DCB is provided by the IGS (International GNSS Service) ionospheric group, is the result of comprehensive processing by four joint ionosphere analysis center (IAAC) [2,3]. Most of them calculate satellites' DCB parameters is through make a global ionospheric model. Because the BeiDou navigation system has not been completed,so the satellites and tracking stations is less, can not make a good global ionospheric modeling by a single system and thus a good estimate of BeiDou satellite DCB parameters. At present, the research on the BeiDou satellite DCB is mostly based on global grid ionospheric model that send by CODE [4]. This will inevitably introduces the error of interpolaion and reducing the solver accuracy. Li Zishen et al. propose Two-step method of determining differential code biases in GNSS/Compass satellites-IGGDCB, This method is based a small number of stations can achieve a accurate estimates of the Beidou satellite DCB [5]. Taking into account of the GPS system has a global distribution of satellite constellations and a large number of continuous stations, So put forward using GPS rice data to estimate the DCB of BDS satellite by make a global ionosphere model using the observations of GPS and BDS.

This paper describes the methods and principles of combined GPS/BDS observations to make a global ionospheric modeling firstly, then Shows the statistical results of calculated GPS, Beidou satellite DCB parameter, Finally, Compass real data from the PPP pseudo-single-frequency dynamic positioning solver, The results are displayed that using the BDS satellites DCB parameter correction in a single-frequency pseudo range Beidou positioning, the positioning accuracy has improved significantly.

29.2 Combine GPS/BDS Establish Global Ionospheric Model

29.2.1 Spherical Harmonics Ionospheric Modeling Method

The foundation equation make ionospheric model using GNSS observations is:

$$p^i_{j,4} = p^i_{n,j} - p^i_{m,j} = \frac{40.28 d_{TEC}}{f_n^2} - \frac{40.28 d_{TEC}}{f_m^2} - (b_n - b_m) \tag{29.1}$$

$$L_{j,4}^{i} = L_{n,j}^{i} - L_{m,j}^{i} = \frac{40.28 d_{TEC}}{f_m^2} - \frac{40.28 d_{TEC}}{f_n^2} - (B_n - B_m) \tag{29.2}$$

In which, $P_{j,4}^{i}$, $L_{j,4}^{i}$ represent the value of two different frequencies n, m pseudo range and phase no geometric combination. d_{TEC} represent the value of STEC, f_n, f_m represent the value of two different frequencies n, m. In this paper, Using the observation data of compass B1, B2 frequency to model ionospheric. Their frequencies are 1561.098 and 1207.14 MHZ. b_n, b_m represent DCB of n, m frequency, B_n, B_m represent integer ambiguity of frequency n and m.

Taking into account the accuracy of the Beidou system code observations in m level and phase observations at mm-level accuracy, using L4 ambiguity of continuous arcs to calculate the approximate ionospheric delay of sight direction. The method is as follows: assuming in a no cycle slips continuous arc, the total number of observation epochs is T, then,

$$L_4^{n,m} = \frac{40.28 d_{TEC}}{f_m^2} - \frac{40.28 d_{TEC}}{f_n^2} + \frac{\sum_{i=1}^{T} \left[L_4^{n,m}(i) + p_4^{n,m}(i) \right] \sin e_i}{\sum_{i}^{T} \sin e_i} + (b_n + b_m) \tag{29.3}$$

$$STEC_{12} = \frac{f_1^2 f_2^2}{40.28\left(f_1^2 - f_2^2\right)} L_4^{1,2} = 8.99322 L_4^{1,2} \tag{29.4}$$

in which, e_i is the Satellite elevation angle of epoch i. $STEC_{12}$ is the ionospheric delay value of sight direction. In this paper, using single-layer model and the spherical harmonic function to fit the global ionospheric electron content [6]. Its expression is as follows:

$$STEC(\beta, s) = \sum_{n=0}^{n_{\max}} \sum_{m=0}^{n} \frac{p_{nm}(\sin \beta)(a_{nm} \cos ms + b_{nm} \sin ms)}{M} + DCB_s + DCB_r \tag{29.5}$$

For all stations observed GPS and Compass data according to Eq. (29.5) establish equations, since the equations of satellite and receiver DCB DCB coefficients as the satellite and the receiver can not be directly isolated DCB, the focus of benchmark constraints used for this purpose, and even if DCB of all GPS satellites and zero [3], because it is a dual system modeling, still need to make all of the Beidou satellite DCB sum to zero, this constraint does not affect the ionosphere model parameters [3].

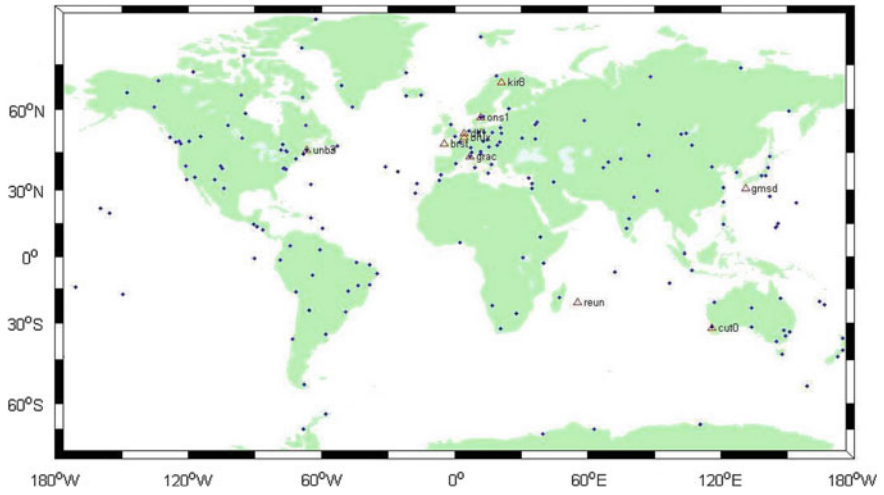


Fig. 29.1 The GPS and BDS stations distribution used in modeling

29.2.2 Data Processing

This paper selected GPS global uniformly distributed more than 200 observation stations and MEGX IGS monitoring network provided/BDS dual system observatory data, station distribution shown in Fig. 29.1, which marked the point of the triangle GPS/BDS dual system monitoring station, a total of 11 stations, dot labeled single system for GPS monitoring stations, there are 191 stations on the 2013 mid-volume days 45–69 days of GPS/BDS global ionospheric data were modeled and calculated solution Beidou satellite DCB parameters.

Combined GPS/BDS global ionospheric data flow modeling algorithm shown in Fig. 29.2.

29.3 DCB Solver Accuracy Analysis

29.3.1 DCB Stability Analysis Solver

Modeling results Shows the in the same GPS satellite Beidou satellite DCB DCB and valuation, and the current GPS satellites recognized internationally CODE DCB provides the most accurate valuation, calculated according to the above processing strategy in mid-2013 Integrated Day 45–69 days, 25 days of the GPS satellites and the GPS satellites DCB DCB CODE released to the true value and the difference between the statistical average rms deviation obtained by the calculation shown in Fig. 29.3.

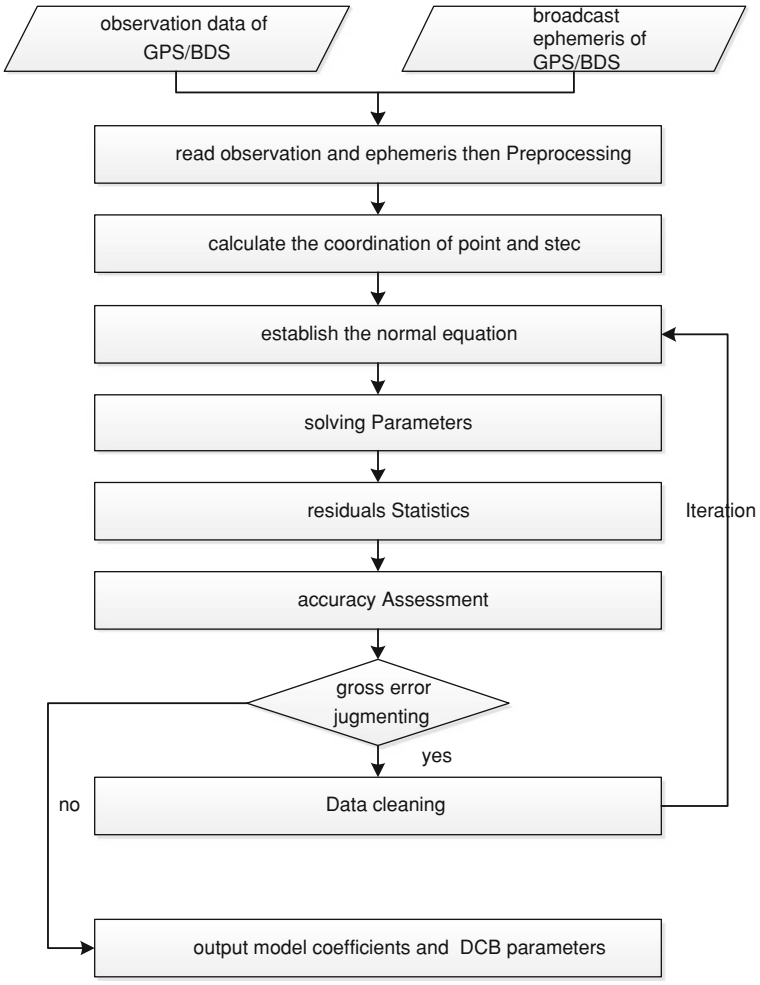


Fig. 29.2 Flow chart of global ionospheric modeling combined GPS/BDS observations

As can be seen from Fig. 29.3 that in addition to the 27th satellite abnormal, the remaining 31 GPS satellites DCB and CDOE released differences are worth about 0.05 ns, statistical difference between the rms value of less than 0.04 ns, while viewing the IGS organization announced DCB value within that period of 27 satellites, which solver in error is 1.6 ns, and the value of IGS organization announced several IGS analysis is based on the value of the global center weighted according to certain methods get, the 27th of its satellite solving the error is much larger than the other satellites in the same period of about 0.1 ns error value, which shows 27 satellites simultaneously IGS analysis centers within each segment calculated DCB quite different. So you can think of GPS satellites DCB accuracy

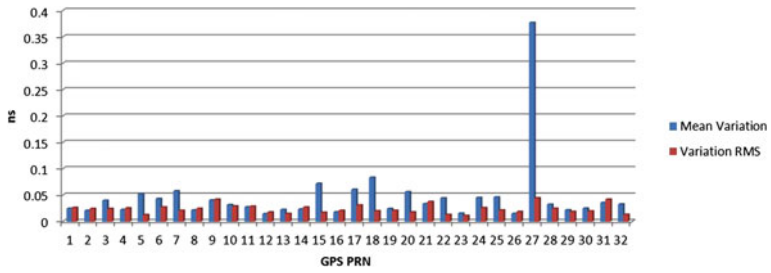


Fig. 29.3 GPS satellite's DCB difference between computed and CODE

of the calculation is about 0.1 ns, this will demonstrate the feasibility of the Beidou satellite DCB calculated using the method.

Since there is no authoritative publication Beidou satellite DCB values, and can not be directly assessed the accuracy of the Beidou satellite DCB, and temperature control equipment, such as satellite-borne very precise, usually in a certain period of time (e.g. 1 month) considers its satellite hardware delay is unchanged. Thus, the stability of the satellite frequency deviation between the valuation accuracy of satellite frequency deviation between the valuation parameters can react to a certain extent [7].

Table 29.1 shows the mean and standard mid-2013 Integrated Day 45–69 days Beidou satellite DCB parameters calculated difference, except for the 1st and 4th satellites, the standard deviation of the remaining satellites DCB parameters are better than 0.15 ns, As 1st and 4th are GEO satellite, only two stations can form an effective puncture its information (in addition to three GEO satellites are three stations can be effectively observations), so the solution DCB parameters calculated standard deviation worse than other satellites.

As a comparison, given the stability of the situation GPS satellite DCB and calculated parameters of the model solution released simultaneously in paragraph CODE GPS satellite DCB parameters. The results are shown in Fig. 29.4.

Can be seen from Fig. 29.4 solver GPS satellite hardware delay deviation STD value most in less than 0.05 ns, and CODE satellite hardware delay release of a considerable deviation stability, and solving the resulting stability than Beidou GPS satellite DCB satellite stability is poor, mainly due to the calculation method of the GPS satellites and CODE DCB parameters are used more than 200 globally distributed data stations, stations its uniform distribution and structure makes perfect constellation of GPS satellites solver DCB parameters are more stable, and only a dozen observatories all Beidou system, and uneven distribution. Figure 29.5 is a partial value of 25 days Beidou DCB time-series variation of the calculated figure, it can be seen from the figure the change is relatively stable.

Since GEO geostationary satellites, location and GEO satellite signal station formed puncture site basically unchanged, thus the total electric ion content they observed zenith ionospheric VTEC (Vertial Total Electron Content) only by the

Table 29.1 The statistical values of BDS satellite's DCB from day 45–69 in 2013

BDS satellites	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C 14
Mean	14.09	4.81	4.08	2.38	0.28	0.56	3.66	2.36	−5.51	−5.06	−6.76	−4.79	−6.52	−3.53
Standard deviation	0.25	0.14	0.11	0.28	0.13	0.06	0.13	0.05	0.08	0.14	0.09	0.12	0.09	0.12

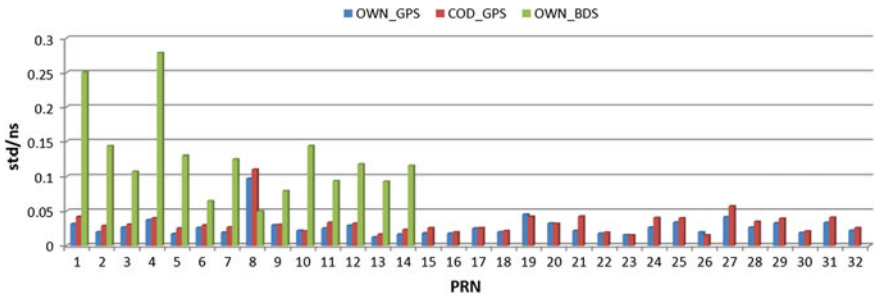


Fig. 29.4 The statistical values of STD for GPS and BDS satellite's DCB through 25 days

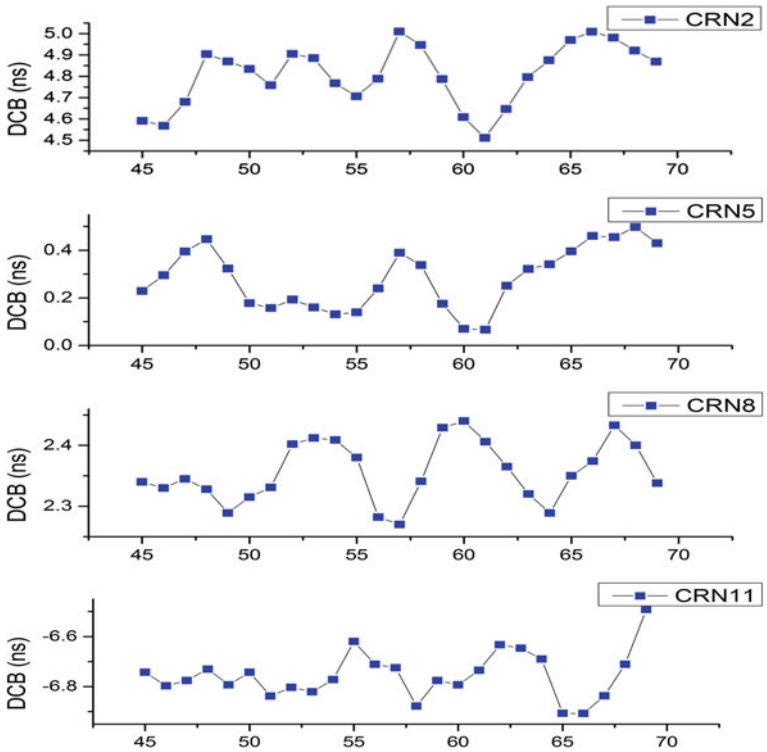


Fig. 29.5 25 days DCB varation of partial BDS satellites

time changes. Figure 29.6 is a change in the value of the point VTEC mid-2013 Integrated Day 41 days GMSD station on the puncture site (N27.58, E128.76) Beidou 2 satellites (GEO satellites) formed after the deduction of the satellite and the receiver DCB get as comparative calculation based on interpolation grid model ionosphere CODE announced the puncture site throughout the day to get the VTEC, their trends and differences shown in Fig. 29.6.

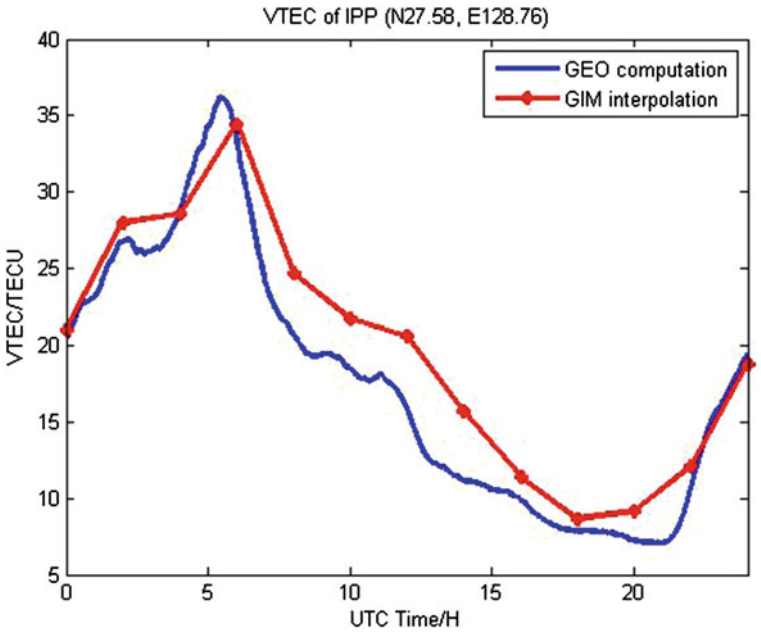


Fig. 29.6 The comparation of ionospheric delay between GEO computation and GIM interpolation

As can be seen from Fig. 29.6, to obtain the value of VTEC day the puncture point was followed by the dual-band Found GEO satellite and receiver DCB charged with the size according to the value of VTEC global ionospheric grid model and the interpolated variations day the basic trend is the same, indicating that the reliability of the Beidou satellite DCB parameter count to get to know, but you can see VTEC values calculated from the measured values obtained by GEO has changed the time on the more detailed, according to global ionospheric grid model because the size of the space needed for each interpolation and temporal interpolation to further reduce the grid ionospheric electron content model accuracy, good performance can not puncture point and the value of VTEC characteristics change with time.

29.3.2 Compass Single Frequency PPP Application

In order to verify the validity and accuracy of the Beidou satellite DCB seeking methods and parameters, using five observations in mid-2013, Taipei bucket receiver Integrated Day 41 days in different areas, the use of self-developed Chang'an University Compass Single Frequency Precise Point Beidou positioning software for single-frequency dynamic pseudorange PPP trials and dual-band

Table 29.2 Dynamic pseudo range PPP positioning error for BDS B1 of two projects

Station	BJF1	CENT	CUT0	SIGP	XIAN
No DCB corrections/m	5.74	5.76	5.65	5.37	5.49
DCB corrections/m	2.71	2.31	2.08	2.83	2.12
Magnitude improvement	52.78 %	59.86 %	63.14 %	47.22 %	61.44 %

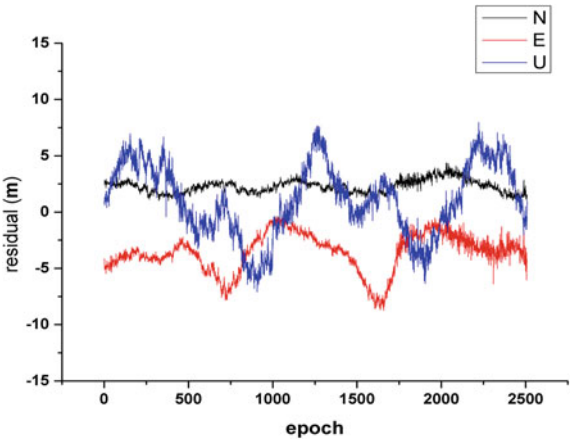


Fig. 29.7 The positioning error without DCB correction

network solution with the GPS coordinates (less than 5 cm) the accuracy of the analysis as the true value. Test is divided into two programs, the first scenario is correct DCB without Beidou satellite pseudo-range single-frequency dynamic positioning PPP, the second program was conducted in accordance with the Beidou satellite correction solver DCB parameters obtained from the single-frequency dynamic pseudo PPP positioning. Two test programs using the same precision Beidou satellite orbit and clock error precision grid ionospheric correction products. Which Compass precise orbit and clock error products offered by Compass calculated Analysis Center, Chang'an University, ionospheric delay correction is to correct the ionospheric grid according to documents released CODE.

In this paper, a full day of dynamic statistical accuracy (RMS), statistical results of the two programs shown in Table 29.2:

As can be seen from Table 29.2, after adding Compass DCB correct single-frequency dynamic PPP point pseudo range error reduced 2.08–3.53 m, improving the margin of 47.22–63.14 %, positioning accuracy is significantly improved. To further compare the coordinate difference between the two series of the program solves for the given station 3 CENT direction positioning day residuals series, Figs. 29.7 and 29.8.

From Fig. 29.7, can be seen in Fig. 29.8, Beidou DCB was added after correction, a series of coordinates of solved more stable positioning accuracy is also improved significantly.

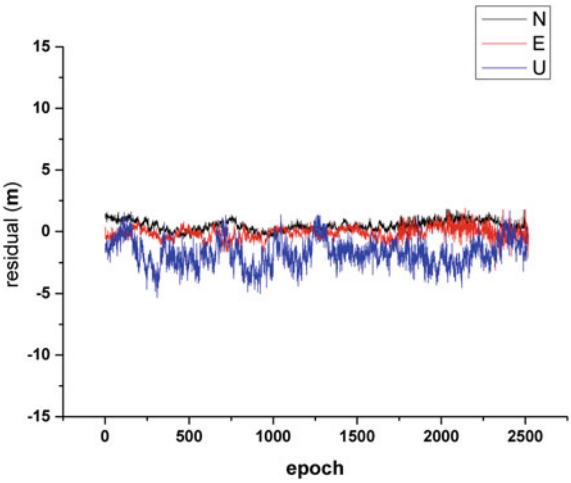


Fig. 29.8 The positioning error with DCB correct

29.4 Result

For the current constellation of the Big Dipper and the number of unfinished global network of monitoring stations less uneven distribution problem, a combined GPS/BDS global ionospheric modeling approach to estimating method Beidou satellite DCB parameter estimates obtained for 25 consecutive stability is within days of satellite DCB 0.15 ns, while differences in GPS satellite DCB parameters obtained in CODE's about 0.05 ns, indicating the feasibility of using this method to estimate the Beidou satellite DCB. Further the calculated Beidou DCB was added to the single-frequency dynamic Compass PPP pseudorange positioning, positioning accuracy of the results show that not added added ratio correction DCB DCB correction improves 2–3 m, to improve the margin of 47–63 %.

Due to the characteristics GEO geostationary satellites, ground receivers corresponding puncture point basically unchanged by the ionospheric delay associated only with time, when our country has a large number of cases under the uniform distribution of Compass stations can be directly calculated based on monitoring VTEC cases forming station and GEO satellites to monitor the puncture site ionospheric electron density and changes in our country. In the case of a GEO satellite DCB high precision, the ionospheric delay values by dual frequency than the measured value obtained by the calculation of the global grid ionospheric model interpolated value has a higher degree of confidence and can be very better reflect the advantages of ionospheric VTEC changes over time.

References

1. Dai W, Jiao W, Jia X (2009) Application research for compass navigation satellite inter frequency bias correction terms. J Geomatics Sci Technol 26(5):367–369
2. Geng C, Zhang H, Zhai C (2009) Real time estimation of DCB using Kalman filters. Geomatics Inf Sci Wuhan Univ 34(11):1309–1322
3. Zhang H, Han W, Huang L (2012) Modeling global ionospheric delay with IGS ground-based GNSS observations. Geomatics Inf Sci Wuhan Univ 37(10):1187–1189
4. Wu X, Ping J, Liu L (2011) Hardware delay solution of regional satellite navigation system. Geomatics Inf Sci Wuhan Univ 36(10):1218–1221
5. Li Z (2013) Study on the mitigation of ionosphere delay and the monitoring of global ionospheric TEC based on GNSS/compass. University of Chinese Academy of Sciences, China
6. Schaer S (1999) Mapping and predicting the Earth's ionosphere using the Global Positioning System. Astronomical Institute of the University of Bern, Switzerland, 1999
7. Wilson BD, Mannucci AJ (1993) Instrumental biases in ionospheric measurement detrived from GPS data. In: Proceedings of the ION GPS-93, Salt Lake City, Utah, USA, Sept 22–24

Chapter 30
Regional Ionospheric TEC Gradients Estimation Using a Single GNSS Receiver

Cheng Wang, Jiexian Wang and Yu Morton

Abstract This paper presents an algorithm to estimate regional ionospheric total electron content (TEC) and TEC gradients. In this study, the ionospheric vertical TEC (VTEC) at an ionosphere piercing point (IPP) in a local area is represented by contributions from the VTEC at other IPPs in the local area and spatial gradients of TEC along both latitudinal and longitudinal directions. By differencing the TEC between each pair of observable IPPs, linear equations of local TEC gradients can be established if only first order and second order TEC spatial derivatives in the TEC representation are retained. The TEC gradients obtained from these linear equations are then combined with VTEC of observable IPPs to generate regional TEC maps. The algorithm is tested using a single GNSS receiver located on the campus of Tongji University over a 24-h period. The TEC values obtained from the algorithm are compared with that from the polynomial-based regional ionospheric TEC model constructed using data from 13 stations in Yangtze River Delta area and also with estimations generated from the IGS global ionosphere maps (GIM). The results show that the average difference between TEC values generated from a single receiver using the TEC gradient-based algorithm and from 13 stations using the polynomial-based regional TEC model is around 1 TECU, and the standard deviation of the differences is about 1.5 TECU. Also, the average difference between TEC values calculated from the proposed TEC algorithm and from GIM is about 0.5 TECU, and the standard deviation of the differences is nearly 1.6 TECU. The results demonstrate that the proposed TEC gradient-based method has the potential to produce accurate regional TEC maps for precise positioning and for ionospheric monitoring using measurements from only a single receiver. The algorithm can be

C. Wang (✉) · J. Wang
College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China
e-mail: acex@foxmail.com

C. Wang · Y. Morton
Department of Electrical and Computer Engineering, Miami University, Oxford, USA
e-mail: mortonyt@muohio.edu

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_30,

easily modified to incorporate other multi-GNSS measurements such as GLONASS, Beidou System and Galileo to further improve precision.

Keywords Ionosphere · Total electron content · Gradient · Model

30.1 Introduction

Ionosphere TEC is an important parameter for both satellite navigation and scientific studies of the ionosphere and space weather. TEC determines the first order ionosphere induced range error, which is a dominant error source in GNSS generated navigation solutions. Existing models such as the Klobuchar model [1], the International Reference Ionosphere (IRI) model [2], the Bent model [3], and the NeQuick model [4] are only suited for scientific analysis of the general trend of the ionosphere and are limited by their accuracy in practical applications. Measurement-based regional and global TEC estimation have limited precision and spatial and temporal resolution or require densely populated GNSS stations to capture detailed ionosphere structures [5–15].

Previous work by C. Wang and Y. Morton presented an approach to precisely estimate ionosphere delay correction using a sparse network of receivers in a relatively small high latitude local region [16]. The algorithm treats the VTEC above a local area as contributions from the VTEC values of all observable IPPs and from TEC gradients in the latitudinal and longitudinal directions. Differencing pseudorange measurement equations between each pair of observable IPPs eliminates the VTEC contribution from the local zenith, leaving TEC spatial gradients to be the dominant variables. This paper uses the same concept presented in the conference paper [16] and applies it to single receiver measurements collected on the campus of Tongji University. The results are compared with GIM and with estimations using polynomial-based regional TEC model and data collected in the Yangtze River Delta area between N30.5°–N32° latitude and E120°–E122.5° longitude over a relatively fine grid of 0.1° latitude and 0.2° longitude.

The remaining paper is organized as the following. Section 30.2 summarizes the detailed methodology of the TEC gradient estimation and polynomial model. Section 30.3 presents the TEC estimation using a regional ionospheric TEC model and using two variations of the gradient-based approaches over Yangtze River Delta area. Section 30.4 summarizes the work and future directions.

30.2 Methodology

The gradient-based TEC estimation methodology was presented in the conference paper [16, 17] and will be summarized below for the sake of completeness. It uses a dual-frequency GPS receiver generated code and carrier phase observables. The carrier phase "smoothed" code measurements can be expressed as the following:

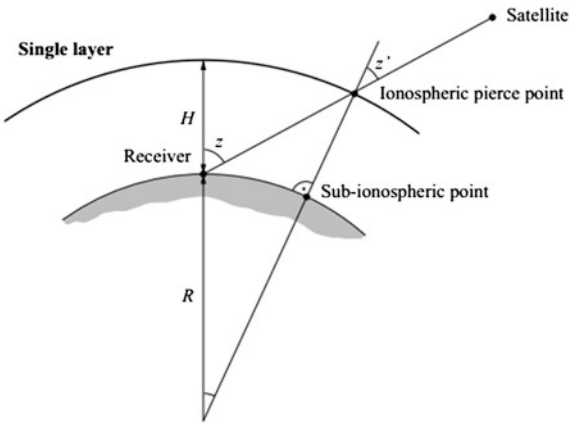


Fig. 30.1 Single thin-shell ionosphere model [18, 19]

$$P_f = \rho_0 + c(\Delta t_r - \Delta t_s) + T + I_f + c\left(b_{r,f} + b_{s,f}\right) + \varepsilon_f \tag{30.1}$$

where the subscript f indicates the frequency dependency of the terms; P is the carrier phase-smoothed code measurements; ρ_0 is the true geometric range between the receiver and a satellite; c is the speed of light; Δt_r and Δt_s are receiver and satellite clock error with respect to GPS time respectively; T is the tropospheric delay; I is the ionospheric delay; b_r and b_s are the receiver and satellite hardware delays respectively; and ε contains the multipath effect, measurement noise, and other error sources. By differencing the carrier-smoothed code phase measurements between L1 and L2 signals, we eliminate all non-dispersive terms in Eq. (30.1):

$$\Delta P = P_1 - P_2 = I_1 - I_2 + c(b_{r1} - b_{r2} + b_{s1} - b_{s2}) + \Delta\varepsilon_{12} \tag{30.2}$$

Denote the difference between the receiver L1 and L2 hardware delays and between the satellite L1 and L2 hardware delays as the receiver and satellite differential code bias (*DCB*) respectively and apply the thin shell approximation of the ionosphere (Fig. 30.1) to approximate the ionosphere delay as the product of a VTEC term and a mapping function, Eq. (30.2) can be re-written as the following:

$$\Delta P = \frac{40.3\left(f_2^2 - f_1^2\right)}{f_1^2 f_2^2} * mf(z) * VTEC + c(DCB_r + DCB_s) \tag{30.3}$$

where f_1 and f_2 are the carrier frequencies of L1 and L2 signals respectively; $mf()$ is the ionospheric mapping function which depends on the zenith distance z at the station (z is the satellite inclination angle at the receiver as shown in Fig. 30.1); *VTEC* is the VTEC at the IPP, and the noise term is ignored. Equation (30.3) can be further re-arranged in this form:

$$\frac{1}{g}\Delta P = VTEC + \frac{c}{g}(DCB_r + DCB_s), g = \frac{40.3\left(f_2^2 - f_1^2\right)}{f_1^2 f_2^2} * mf(z) \tag{30.4}$$

In this study, a modified ionosphere single layer mapping function is used [20]:

$$mf(z) = \frac{1}{cos\, z'} \text{ with } sin\, z' = \frac{R}{R+H}\sin(\alpha z) \tag{30.5}$$

where z' is the zenith distance of the satellite at the IPP; R is the mean earth radius; H is the ionospheric single layer height; α is a correction factor. In this study, $\alpha = 0.9782$ is used which is obtained by performing best fit of (30.5) with respect to the JPL extended slab model (ESM) mapping function.

Prior ionospheric TEC model based on polynomial functions uses Eq. (30.6) to calculate VTEC in the regional ionosphere at different latitude and longitude.

$$VTEC = \sum_{i=0}^{n}\sum_{j=0}^{m} E_{ij}(\varphi - \varphi_0)^i (s - s_0)^j \,, s - s_0 = (\lambda - \lambda_0) + (t - t_0) * \frac{\pi}{12} \tag{30.6}$$

where n and m are polynomial degrees of the model; E_{ij} are the polynomial term coefficients; φ and λ are latitude and longitude of the IPP, respectively; φ_0, λ_0 are latitude and longitude of the expansion point, which is typically the central location in the local area; t is the observation time in hours; t_0 is the central time of the model interval.

In this study, the VTEC at the IPP is represented as the following:

$$\begin{aligned} VTEC_i = {} & VTEC_0 + (\lambda_i - \lambda_o)f'_\lambda + (\phi_i - \phi_0)f'_\varphi + (\lambda_i - \lambda_o)^2 f''_{\lambda\lambda} \\ & + 2(\lambda_i - \lambda_o)(\phi_i - \phi_0)f''_{\lambda\varphi} + (\phi_i - \phi_0)^2 f''_{\varphi\varphi} \end{aligned} \tag{30.7}$$

where $VTEC_0$ is the VTEC at the central location in the local area; $f'_\lambda, f'_\varphi, f''_{\lambda\lambda}, f''_{\lambda\varphi}, f''_{\varphi\varphi}$ are the first order and second order VTEC derivatives along latitude and longitude respectively. Substituting (30.7) into (30.4):

$$\begin{aligned} \frac{1}{g_i}\Delta P_i = {} & VTEC_0 + (\lambda_i - \lambda_o)f'_\lambda + (\phi_i - \phi_0)f'_\varphi + (\lambda_i - \lambda_o)^2 f''_{\lambda\lambda} \\ & + 2(\lambda_i - \lambda_o)(\phi_i - \phi_0)f''_{\lambda\varphi} + (\phi_i - \phi_0)^2 f''_{\varphi\varphi} + \frac{c}{g_i}(DCB_R + DCB_s^i) \end{aligned} \tag{30.8}$$

where the subscript i identifies the IPP associated with a signal path connecting a pair of satellite and receiver; DCB_R and DCB_s^i are the receiver and satellite DCB associated with the IPP i.

The difference between a pair of IPPs i and j is:

$$\begin{aligned}\frac{1}{g_i}\Delta P_i - \frac{1}{g_j}\Delta P_j = &(\lambda_i - \lambda_j)f'_\lambda + (\phi_i - \phi_j)f'_\varphi + (\lambda_i - \lambda_j)(\lambda_i + \lambda_j - 2\lambda_0)f''_{\lambda\lambda} \\ &+ 2(\lambda_i\phi_i - \lambda_j\phi_j - (\lambda_i - \lambda_j)\phi_0 - (\phi_i - \phi_j)\lambda_0)f''_{\lambda\varphi} \\ &+ (\phi_i - \phi_j)(\phi_i + \phi_j - 2\phi_0)f''_{\varphi\varphi} + c\left(\frac{1}{g_i} - \frac{1}{g_j}\right)DCB_R \\ &+ c(\frac{1}{g_i}DCB_s^i - \frac{1}{g_j}DCB_s^j)\end{aligned} \tag{30.9}$$

ΔP can be computed from GPS receiver measurements, while g, λ, and ϕ can be estimated using known receiver position and satellite ephemeris. Assuming that there are a total of n IPPs at a given epoch for receivers in a local network, there are a total of $n*(n-1)/2$ pairs of IPPs, and hence $n*(n-1)/2$ independent equations. The unknowns in Eq. (30.9) are $f'_\lambda, f'_\varphi, f''_{\lambda\lambda}, f''_{\lambda\varphi}, f''_{\varphi\varphi}$ and the DCBs of satellites and receivers.

Time domain linear interpolations are used to constrain the gradient temporal variations in the solution. If the time window used to constrain the solution is 1 h, to estimate the TEC gradients over a 24-h period, there will be 25 pairs of gradients at each of the hour on the day and the first hour on the next day. The gradients at the other times can be evaluated using linear combinations of these 25 sets of gradients.

GPS satellite DCB values can be acquired from IGS GIM, and DCB values for receivers can be calculated by IONOLAB-BIAS [21] which is a receiver DCB estimation algorithm that has been used at some IGS and non-IGS stations using GIM TEC and satellites DCB values. These satellite and receiver DCB values can be used to evaluate the results computed by Eq. (30.9).

Using the TEC spatial gradients and VTEC estimation at each IPP, the VTEC for any location (λ, ϕ) in the local area can be calculated as the following:

$$\begin{aligned}VTEC = \frac{1}{n}\sum_{k=1}^{n}(&VTEC_k + (\lambda - \lambda_k)f'_\lambda + (\phi - \phi_k)f'_\varphi + (\lambda - \lambda_k)^2 f''_{\lambda\lambda} \\ &+ 2(\lambda - \lambda_k)(\phi - \phi_k)f''_{\lambda\varphi} + (\phi - \phi_k)^2 f''_{\varphi\varphi})\end{aligned} \tag{30.10}$$

where n is the total number of visible satellites at a given epoch; $VTEC_k$ is the VTEC at the k IPP. $VTEC_k$ could be calculated by the formula (30.4), because the DCBs of satellites and receivers are all known.

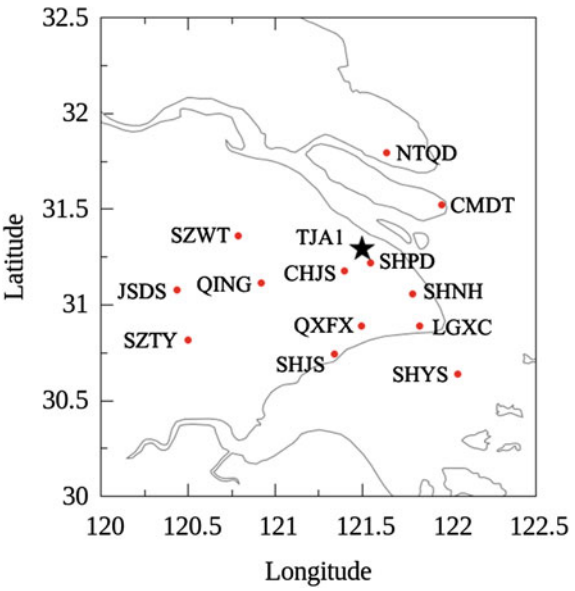


Fig. 30.2 Location of the stations used in the study

30.3 Results and Discussions

30.3.1 Polynominal-Based Regional TEC Model Results

The regional ionospheric TEC modeling based on polynomial functions (30.6) presented in Sect. 30.2. is applied to data collected on January 21, 2013 from 13 reference stations in Yangtze River Delta area as shown in Fig. 30.2. The 13 stations are CHJS, CMDT, JSDS, LGXC, NTQD, QING, QXFX, SHJS, SHNH, SHPD, SHYS, SZTY, SZWT. And the station TJA1 with star tag is a GNSS receiver located on the campus of Tongji University, China. The TEC gradients estimation algorithm presented above is applied to data on the same day from this GNSS receiver.

Table 30.1 shows the standard deviations of the regional polynomial-based ionospheric TEC model over Yangtze River Delta for polynomials with n and m ranging from 1 to 5. Table 30.2 shows the percentage of absolute value of residuals in each interval within (0,1) (1,2) (2,3) (3,4), and (4,5) TECU. Table 30.2 also listed the percentage of the absolute residuals within of 3 TECU.

According to Table 30.1, the standard deviations for different n and m values are very similar at around 1.8 TECU. It shows that for a small area like the Yangtze River Delta area, the precision of the overall regional ionospheric TEC model is not seriously affected by the degree of polynomials. Table 30.2 shows that there are large differences among the percentage of absolute residuals in different residual intervals for different polynomial degrees. The smallest and largest percentages of absolute residuals within 3 TECU are about 82 and 92 %

Table 30.1 The standard deviation of TEC based on different polynomial degrees

n\m	1	2	3	4	5
1	1.828	1.827	1.827	1.827	1.827
2	1.802	1.800	1.798	1.798	1.798
3	1.797	1.794	1.792	1.792	1.792
4	1.796	1.794	1.792	1.791	1.794
5	1.795	1.792	1.790	1.790	1.792

Table 30.2 Percentage of absolute TEC residual based on polynomial models

(n, m)\TEC residual (TECU)	(0, 1)	(1, 2)	(2, 3)	(3, 4)	(4, 5)	<=3 TECU
1, 1	51.76	24.41	9.94	4.81	2.82	86.12
1, 2	47.70	22.34	16.44	4.84	2.92	86.48
1, 3	49.72	20.63	17.75	4.02	2.56	88.10
1, 4	40.85	24.46	16.87	5.77	2.92	82.18
1, 5	41.90	24.28	15.98	5.37	2.70	82.16
2, 1	52.45	24.94	9.77	4.66	2.05	87.16
2, 2	52.16	25.05	9.82	4.76	2.18	87.03
2, 3	52.08	25.21	9.80	4.70	2.17	87.09
2, 4	51.97	25.39	9.75	4.68	2.12	87.11
2, 5	49.39	29.84	12.86	3.01	1.42	92.08
3, 1	51.40	24.60	11.08	3.99	1.85	87.08
3, 2	52.12	24.37	10.70	4.44	1.95	87.19
3, 3	53.26	24.64	9.63	4.45	2.06	87.53
3, 4	45.26	26.09	12.54	4.64	2.38	83.89
3, 5	52.12	29.82	9.98	2.34	1.96	91.92
4, 1	49.58	25.47	13.86	5.70	2.24	88.91
4, 2	39.74	33.70	19.23	2.75	2.20	92.67
4, 3	53.34	24.53	9.79	4.42	2.08	87.66
4, 4	52.25	24.81	10.07	4.33	2.08	87.13
4, 5	53.15	24.79	9.35	4.46	2.19	87.29
5, 1	44.27	26.18	13.23	4.86	2.36	83.68
5, 2	50.60	25.08	10.92	4.49	1.92	86.60
5, 3	53.30	24.59	9.68	4.46	2.05	87.57
5, 4	53.23	24.68	9.51	4.54	2.09	87.41
5, 5	53.30	24.45	9.41	4.43	2.30	87.16

respectively. There are three cases generated less than 3 TECU residual beyond 90 %. The corresponding n and m values are 2 and 5, 3 and 5, 4 and 2, respectively. The most optimal setting is $n = 3$ and $m = 5$ which has 52.12 % of absolute residuals under 1 TECU.

Table 30.3 Proportions of absolute residual of the two approaches in different interval

Approach\TEC residual (TECU)	(0, 1)	(1, 2)	(2, 3)	(3, 4)	(4, 5)	<=3
1	78.25	17.41	3.35	0.51	0.14	99.01
2	78.38	17.27	3.53	0.62	0.07	99.18

30.3.2 TEC Gradient-Based Model Results

The receiver DCBs can be estimated along with the regional ionospheric TEC gradients estimation from (30.9). Also, it can be calculated by IONOLAB-BIAS independently. Two approaches are taken in this study. In the first approach, the receiver DCB and TEC gradients are estimated simultaneously. In the second approach, the receiver DCB is determined by IONOLAB-BIAS first, then substituted into (30.9) as a known parameter to estimate the TEC gradients. The standard deviation values of these two approaches are 0.4830 and 0.4865 TECU respectively. Also, the percentages of absolute residual of the two approaches within different TEC intervals are listed in Table 30.3.

30.3.3 Comparison with GIM and Regional TEC Model

The ionospheric TEC maps over Yangtze River Delta area by the TEC gradient-based approaches from (30.10) are evaluated against GIM and the regional polynomial-based TEC model presented above. The mean, standard deviation (STD), and root mean square errors (RMS) among the estimated VTEC by gradients, GIM results, and regional TEC model are computed as the following:

$$\Delta TEC_{mean} = \frac{1}{m}\sum_{i=0}^{m} r_i \tag{30.11}$$

$$\Delta TEC_{std} = \sqrt{\frac{1}{m}\sum_{i=0}^{m} (r_i - \bar{r})^2} \tag{30.12}$$

$$\Delta TEC_{rms} = \sqrt{\frac{1}{m}\sum_{i=0}^{m} r_i^2} \tag{30.13}$$

where r_i is the difference of VTEC values; $\bar{r}$ is the average difference r_i; m is the total number of samples.

Table 30.4 shows the statistical performance of the TEC gradient-based solution. In the Table 30.4, "App1" refers to the first gradient-based approach which treats the receiver DCB as an unknown parameter, "App2" is the second approach where the receiver DCB is calculated by IONOLAB-BIAS, and "Poly" refers to the regional polynomial-based TEC model.

Table 30.4 Statistical performance of VTEC estimation

Approaches\statistics	Mean	STD	RMS
App1 versus poly	1.090	1.479	1.837
App1 versus GIM	0.543	1.578	1.669
App2 versus poly	0.633	1.475	1.605
App2 versus GIM	0.087	1.581	1.583
Poly versus GIM	−0.547	1.510	1.606

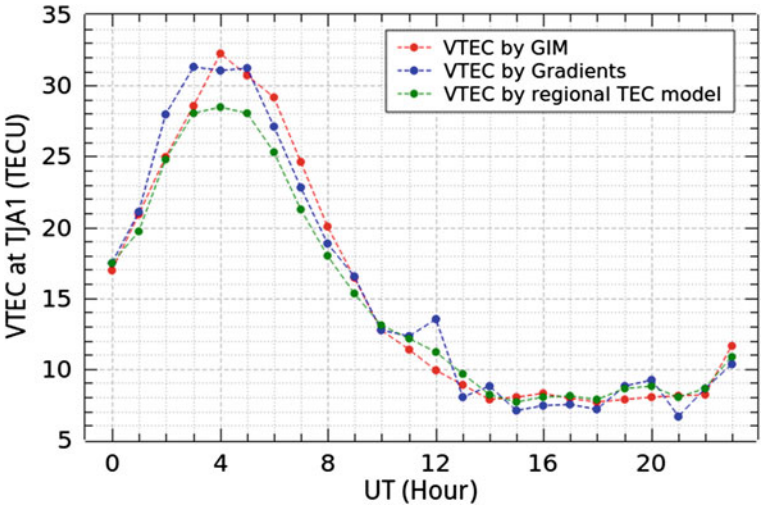


Fig. 30.3 VTEC at TJA1 estimated from GIM, Poly, and App2

The statistical performance of App2 appears to be better than that of App1. The smallest RMS between App2 and GIM is about 1.58 TECU. A possible explanation maybe that the receiver DCB estimated by App1 includes model errors such as contributions from omission of high order remainders. These errors are carried into the VTEC estimation using Eq. (30.4).

VTEC at station TJA1 is also calculated using App2. The results are compared with that of GIM and Poly. Figure 30.3 plots the estimated VTEC at TJA1 and hourly GIM results are shown as red dots, and the hourly TEC values by Poly are shown as green dots. Most of the data points show excellent agreement among App2, Poly, and GIM. The largest difference is about 4 TECU at 4-hour UTC between Poly and GIM. The largest difference between App2 and Poly is around 3 TECU at 2-, 3-, 5-h UTC.

30.4 Conclusions and Future Work

This paper presented a TEC gradient-based algorithm to estimate regional ionospheric TEC gradients and TEC using a single GPS receiver located on campus of Tongji University in China. The performance of the TEC gradient-based algorithm was assessed by comparing the estimated TEC with that provided by regional polynomial-based TEC model and GIM. 13 GPS stations located in the Yantz Delta area are used to generate the regional polynomial-based TEC map. The results show that the proposed method generates TEC map in close agreement that of regional TEC model and GIM. The standard deviation of the TEC differences between that generated by the single receiver using the TEC gradient approach and by 13 stations using the polynomial-based regional TEC model is about 1.6 TECU. It demonstrates that the TEC map generated by regional ionospheric TEC gradients using just a single GNSS receiver could have a relative high precision.

In the future, we will incorporate other multi-GNSS measurements including GLONASS, BDS, and Galileo. More evaluations of the algorithm using stations at different geographical locations and different space weather conditions will also be performed.

References

1. Klobuchar JA (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst AES-23(3):325–331
2. Rawer K, Bilitza D, Ramakrishnan S (1978) Goals and status of the international reference ionosphere. Rev Geophys 16:177–181
3. Bent RB, Llewellyn SK, Schmid PE (1972) A highly successful empirical model for the worldwide ionospheric electron density profile. DBA Systems, Melbourne
4. Dudeney (1978) An improved model of the variation of electron concentration with height in the ionosphere. J Atmos Terr Phys 40(2):195–203
5. De Franceschi G, Zolesi Bruno (1998) Regional ionospheric mapping and modelling over Antarctica. Ann Geofis 41:813–818
6. Sakai, Takeyasu, Matsunaga et al (2007) Mitigating ionospheric threat using a dense monitoring network. In: Proceedings of the 20th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2007), Fort Worth, Texas, Sept 2007, pp 927–938
7. Jensen ABO, Ovstedal O, Grinde G (2008) Development of a regional ionosphere model for Norway. In: Proceedings of the 2008 national technical meeting of the Institute of Navigation, San Diego, CA, Jan 2008, pp 893–902
8. Lejeune S, Wautelet G, Warnant R (2012) Ionospheric effects on relative positioning within a dense GPS network. GPS Solutions 16(1):105–116
9. Coster AJ, Gaposchkin EM, Thornton LE (1992) Real-time ionospheric monitoring system using GPS navigation. J Inst Navig 39(2):191–204
10. Komjathy A (1997) Global ionospheric total electron content mapping using the global positioning system Ph.D Dissertation, Department of geodesy and geomatics engineering technical report No. 188, University of New Brunswick, Fredericton, p 248

11. Gao Y, Heroux P, Kouba J (1994) Estimation of GPS receiver and satellite L1/L2 signal delay biases using data from CACS. In: Proceeding of KIS-94, Banff, Canada, 30 Aug–2 Sept 1994
12. Skone S (1998) Wide area ionosphere grid modeling in the auroral region. Ph.D Dissertation, The University of Calgary, UCGE report number 20123, Calgary
13. Liao X (2000) Carrier phase based ionosphere recovery over a regional area GPS network. UCGE reports number 20143, University of Calgary, Calgary, p 120
14. Schaer S (1999) Mapping and predicting the earth's ionosphere using the global positioning system. Ph.D Dissertation, Astronomical Institute, University of Berne, Switzerland, p 205
15. Walker JK (1989) Spherical cap harmonic modeling of high latitude magnetic activity and equivalent sources with sparse observations. J Atmos Terr Phys 51(2):67–80
16. Wang C, Morton Y (2013) Ionosphere TEC and TEC gradients estimation using a regional GNSS network. In: Proceedings of the ION GNSS, Nashville, Sept 2013
17. Bourne H, Morton Y (2013) GPS receiver ionosphere error correction based on spatial gradients and IGS satellite DCBs. In: Proceedings of the ION Pacific PNT, Honolulu, HI, Apr 2013
18. Global Ionosphere Maps [EB/OL] (2012). http://aiuws.unibe.ch/ionosphere
19. User manual of the Bernese GPS Software Version 5.0 (2007)
20. Schaer S (1999) Mapping and predicting the earth's ionosphere using the global positioning system. Ph.D Dissertation, Astronomical Institute, University of Berne, Switzerland, p 205
21. Arikan F, Nayir H, Sezen U, Arikan O (2008) Estimation of single station interfrequency receiver bias using GPS-TEC. Radio Sci 43(RS4004):1–13

Chapter 31
A Inter Satellite Link Algorithm of Satellite Navigation System Based on the Integration

Ying Hou, Zhi-chao Mi, Wei-bo Yu and Da-wei Niu

Abstract Over the years satellite networks has played an more and more important role, and there has been the trend to build Inter-Satellite Links (ISLs) in the satellite networks as the ISLs can provide high network flexibility and reliability. Integration is one of the important characteristics of the navigation system, and the network topology of the inter satellite has great influence on it, so it's necessary to improve the performance of navigation system through appropriate inter satellite link planning. Aiming at this problem, this paper firstly established the link planning principles, and then the visible coverage relationship is proposed for a satellite link allocation algorithm, which is the foundation of the network topology formation. The topology gives preference to the data transmission between satellite and ground, the simulation results show that this network topology has an advantage in hops and delay performance.

Keywords Satellite navigation system · Satellite-earth integration · ISLs · Algorithm

31.1 Introduction

With the growing demand for satellite applications, satellite navigation systems have become an important strategic national resource base [1]. Establishing inter-satellite links in the navigation system will provide the system continuous inter-satellite observations, improve the survivability and autonomous operation capability of the system. The formation of a satellite network has become an inevitable trend in the development of satellite navigation.

Y. Hou (✉) · Z. Mi · W. Yu · D. Niu
Institute of Communication Engineering, PLA University of Science and Technology, Nanjing, China
e-mail: m.y.sky@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_31,

Inter-Satellite Links (ISLs) can be divided into two categories: Intra-Layer ISLs with same orbital altitude and Inter-Layer ISLs with different orbital altitude, and the former can be divided into Intra-Plane ISLs and Inter-Plane ISLs. Due to the limited number of satellites equipped with antennas, usually less than the number of satellites can be established ISLs, so satellite operators have a choice when you want to build the chain in its visible satellites. The inter-satellite network design must consider the ISLs allocation problem.

To carry out research work abroad for ISLs early, the U.S. GPS BLOCK IIR [2] navigation system is currently in orbit navigation satellite constellation to achieve the only links inter the systems. In the study of theory, paper [3] proposed the allocation algorithm based on maximum connectivity bicyclic satellite network link, paper [4] aimed at optimizing the network topology Lu hailed as the goal, using the graph theory to solve links distribution network of ISLs and dynamic switching problems. Some domestic scholars have also conducted a preliminary ISLs research and analysis, paper [5] proposed building a constellation of ISLs standards and optimized design, paper [6] discusses several classic constellation design methods, paper [7] proposed the ISLs allocation algorithm taking both ranging and data communication of navigation system into account.

However, most of the existing studies are oriented to optimize the performance of the communications network, which has a great difference in the demand and use with the navigation system. In the navigation system, most of the cost of power is used to broadcast navigation signals, the research for communications system cannot be directly applied to navigation system. However, the research on navigation satellites has mostly focused on the inter-satellite data communications performance optimization, ignoring the data transfer performance between satellites and ground stations.

Transmission of information in navigation system mainly includes three aspects: satellite and ground stations, the satellite and satellites, and ground stations and user goals. For navigation systems, inter-satellite information exchange demand is not strong, the most important is the transmission of link manages information between ground stations and the satellites. The design of ISLs must pay attention to the problem of transmission and relay between the satellites and the ground stations, which is the focus of this study.

Satellite networks and terrestrial mobile networks have big difference, the satellite have predictable trajectories, constellation dynamic topology can be precalculated under certain rules. As navigation system providing services to a large number of user navigation open system requiring high degree of reliability,and completing the chain and routing on the satellites will increase the complexity of implementation, so that the ISLs planning can be controlled by the ground stations to increase the reliability of navigation system.

In this paper, based on a typical walker-type navigation constellation, we focused on the integration of satellite navigation systems, proposed the ISLs designing algorithm applied to the satellite-earth communications as well as finished the simulation and analysis, the results shows that this algorithm has a certain reference value.

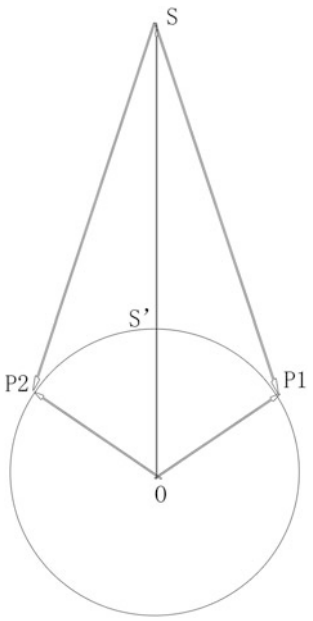


Fig. 31.1 The ground cover for satellite S

31.2 Planning Principles

In order to ensure the normal operation of the system, navigation constellations should meet certain requirements, including the visibility between satellites and ground station, and the visibility between satellites.

The visibility of ground station and navigation system depends on the satellite coverage characteristics on the earth surface. The visibility of point A on the surface to satellite S is defined like this, if point A is in the coverage of satellite S, then S is visible to A [8]. At a time, the coverage of satellite S to the earth is shown as Fig. 31.1.

When satellites are visible to ground stations, they can communicate through satellite-earth links directly, otherwise, they have to communicate with the relay of ISLs.

There are three possible visible relationships between satellites, always visible, intermittent visible, and always invisible. The premise condition for two satellites to establish ISLs is that they are visible. The satellites in the same orbital plane own fixed relative position relationship so that they can build permanent ISLs. However, whether the satellites in the different orbital plane can build ISLs is closely related with the change of the relative position, the always visible satellites can establish fixed topology planning while the intermittently visible satellites have to establish dynamic topology planning. Taking the navigation system functions and requirements into account, to meet the needs of integration in the navigation system, following principles of ISLs planning are established:

1. The ground planning principles. All ISLs planning is finished by the ground station, and then periodically uploaded to the satellite, so that the satellite link state is completely transparent to the ground, and the uncontrollable factors are reduced, the reliability of navigation system is improved;
2. The minimum hop principle. Large number of hops will bring larger transmission delay and high bit error rate, reduce the quality of service navigation system. With limited ISLs resources, another principle is to make the transmission hop number between satellites and ground stations as small as possible;
3. The priority landing principle. Compared to the inter satellite communication, data transmission between ground stations has higher reliability and controllability. Therefore, ISLs planning mainly aims at the invisible satellite with the ground station. When the satellite and the ground station is visible, the data directly go through the satellite-earth transmission link without the transfer of ISLs;
4. The periodic principle. Navigation system has periodicity, the change of the network topology is repeated, and the planning calculations need to have results and circulation with the system operation according to the periodic topology, to ensure simplest computational process;
5. The topology snapshot principle. The topological snapshot transform continuous changes of topological structure of the network into a series of discrete static network topology, if we ignore the link length change, in every time interval satellite topology is fixed, thus the dynamic characteristics of satellite networks can be ignored, the routing algorithm of inter satellite links is simplified, so that we can use some mature static routing algorithms such as Dijkstra shortest path algorithm;
6. The least switch principle. As the link switch will increase satellite antenna operating frequency and affect the quality of communication, thus in the planning of inter satellite links, the number of switch should be as less as possible.

31.3 Design Algorithm

The input of design algorithm is a series of fixed visible relationships obtained according to the operation rules of navigation system, including visible relationships between satellite and ground station and inter satellite, which can be produced by the STK satellite simulation software. The output of the algorithm is the link allocation scheme of a computation period navigation system.

31.3.1 Mathematical Description

The relationship is described by the two 0-1 matrix V_1 and V_2. $V_1 = [v_{ai}]$ is used to describe the visible relationships between satellite and ground station, a is the

ground station node of system while i is the satellite node of system. $V_2 = [v_{ij}]$ is used to describe the visible relationships between satellites, i and j are both satellite nodes of system. Matrix elements $v = 1$ shows the two nodes are visible, while $v = 0$ shows they are not. Navigation system inter satellite links allocation scheme is described by 0-1 matrix $L = [l_{ij}]$, i and j are both satellite nodes of system. Matrix elements $l_{ij} = 1$ shows building ISLs between i and j, while $l_{ij} = 0$ shows the opposite.

Assuming that the antenna equipped with the same number k of all satellites, and one antenna points to the ground station for satellite-earth communication, two antennas point to the intra-plane satellites to build permanent ISL, other antennas point to the inter-plane satellites. So that k should be a integer as least as 3, every satellite could build no more than $k - 1$ ISLs. Satellite the antenna has a two-way connection characteristics, $l_{ij} = 1$ shows the transceiver antenna of i and j is paired at the same time, that is to say, when and only when the receiving antenna of j is paired with the transmitting antenna of i, the receiving antenna of i is paired with the transmitting antenna of j. By definition, V_1, V_2 and L are all symmetric matrix. The planning problem inter satellite links can be described as: given two visual relation matrix V_1 and V_2, integer k, and meet the following conditions: V_1 and V_2 are both symmetric matrix, $k > 3$. Inter satellite link design algorithm aims at getting the link assignment matrix L which meets the following conditions:

1. In link allocation matrix L, the number of 1 in each line or column is not more than $k - 1$, constraining the number of inter satellite links cannot be greater than the number of antennas that can be matched;
2. For any $l_{ij} = 1$, accordingly in V2 $v_{ij} = 1$, shows the satellite visible is the premise to establish inter satellite links.

31.3.2 Algorithm Process

In order to solve the abstract mathematics questions in Sect. 31.3.1, the algorithm finds the optimal link allocation through iterative processing mode solution. The main idea is: in order to guarantee the minimum data transmission hops from satellite to ground station, when satellite is visible to the ground station, they communicated directly, need 1 hop; when satellite is not visible to the ground station, firstly consider the intra-orbit chaining satellite with stable state, choose the one which is visible to the ground station to relay, need 2 hops; when none of the two satellites is visible to the ground station, choose the inter-orbit satellite which is visible to the ground station to relay, also need 2 hops; when all the above are not satisfied, according to Dijkstra shortest path algorithm to select the appropriate communication path. Link planning algorithm in this paper is based on the minimize hop to completion link allocation optimization, chain building process is shown as follows:

Step 1: Greedy link building
Input: satellite visible relationship matrix V_2

Output: greedy link assignment matrix L_1

Operation: Read the satellite visible relationship matrix V_2, for all $v_{ij} = 1$, put $l_{ij} = 1$, in all can satisfy the premise of inter satellite visibility, greedy to establish inter satellite link.
Step 2: Satellite-earth visibility screening
Input: satellite and ground station visible relation matrix V_1, greedy link assignment matrix L_1

Output: screening link assignment matrix L_2

Operation: read satellite and ground station visible relation matrix V_1, when $v_{ai} = 0$, if $v_{aj} = 0$, put $l_{ij} = 0$. That selected satellites which can communicate directly with the ground station to build ISL when the satellite and the ground station is not visible at the time, and remove the links to satellites which are not visible to the ground station.
Step 3: Two-way selection optimization
Input: screening link assignment matrix L_2

Output: double optimized link assignment matrix L_3

Operation: read screening link assignment matrix L_2, when both l_{ij} and l_{ji} equal to 1, retain it, otherwise put it to 0. That is according to the two-way selection conditions of inter satellite links for screening, ensure the link at the same time meet requirements of two satellites, and carry out the demolition that cannot meet the double of inter satellite links.
Step 4: Quantitative restrictions
Input: k, double optimized link assignment matrix L_3

Output: link allocation matrix L

Operation: check each column of double optimized link assignment matrix L_3 to find whether there are more than k-1 elements that equal to 1. In the guarantee of the whole network connectivity, put the exceeded parts to 0. That removed the number of links larger than the available antenna.

The whole chain building process contains 4 steps, greedy link building, satellite-earth visibility screening, two-way selection optimization and quantitative restrictions. After building chains through greedy way, according to constraints remove links sequentially, finally get link allocation matrix required.

31.4 Simulation Analysis

31.4.1 Simulation Environment

Since the 1960s in last century, people have put forward a lot of constellation design method. Practice has proved, three track surface constellation is the trend of the design of satellite navigation system of the next generation. The navigation system

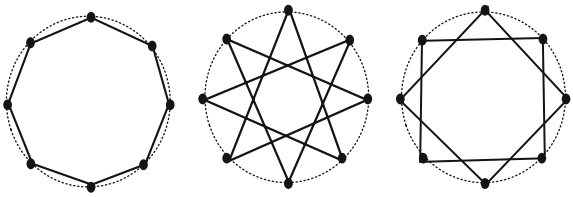


Fig. 31.2 Topologies of intra-orbit ISLs

in this paper includes navigation constellation and the ground station of two parts. Navigation constellation using constellation model proposed in literature [9], 24 MEO satellites evenly distributed in the 3 rail surface, forming standard Walker 24/3/2 constellation, orbit height is 21,500 km, orbital inclination is 55, satellite antenna number is 4, the cycle time of the system is about 12 h 50 min. The 8 satellites in X orbital plane ($x = 1, 2, 3$) are clockwise labeled as $Mx1, Mx2, \ldots, Mx8$. According to the actual situation in China we set up multiple ground stations, set the ground station set $E = \{\text{station } 1, \text{ station } 2, \ldots, \text{ station } N\}$. Due to the cable link between ground stations, there are high reliability and controllability of data transmission, so simulation analysis does not consider the data transmission between ground stations.

Typically, inter satellite link topology in Walker constellation is the four link model, each satellite separately with the two satellites in same orbital plane, and two satellites in adjacent orbit plane to establish inter satellite links, network link allocation method has been applied successfully in the iridium system satellite network [10]. In this navigation constellation, the satellites in same orbit has fixed relative position, due to the blocking effect of the earth, the two satellites with phase difference of 180° are always not visible, otherwise the other satellites are always visible. Take each MEO satellite can build two ISLS with intra-plane satellites for example. Considering the network symmetry and other factors, there are mainly 3 kinds as shown in Fig. 31.2.

In the figure, the third kind of topology is not fully connected, it has two closed inter satellite link rings which are unable to transmit data between each other, in this tracks when only a satellite and the ground station is visible, four satellites in another ring cannot communicate with ground station, obviously not suitable for practical use. This paper only discussed the first ring topology and second kinds of anise star topology.

31.4.2 Simulation Results

31.4.2.1 The Intra-Plane ISLs Set

According to the spatial symmetry of the Walker constellation, geometric characteristics of any satellite can represent all the stars in the constellation, so just discuss the parameters of the first M11 satellite, the relevant parameters of the

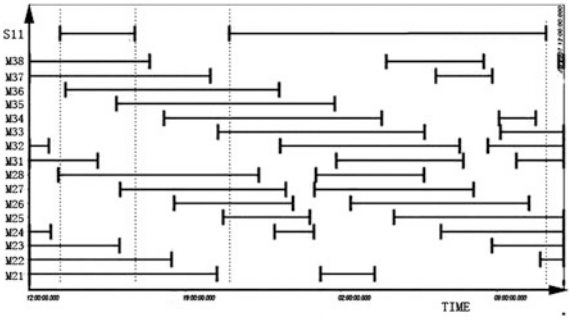


Fig. 31.3 Visibility between $S_{11,}$ other satellites and E

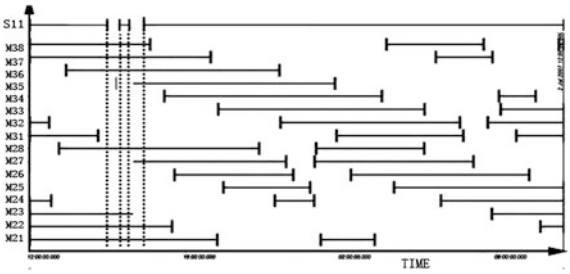


Fig. 31.4 Visibility between $S'_{11,}$ other satellites and E

Table 31.1 The probability to be visible for the earth when M11 is not

Satellites	M12	M13	M14	M15	M16	M17	M18
Visible probability (%)	20	40	57	59	55	38	19

other seven stars can be derived. In order to analyze the effect of the inter satellite link topology for satellite integration performance, define the ring topology intra-plane link building satellite set S_i,the anise star topology intra-plane link building satellite set S'_i, take M11 for example, then $S_{11} = \{M11, M12, M18\}$, $S'_{11} = \{M11, M14, M16\}$. The simulation time is set to 24 h, Figs. 31.3 and 31.4 show in a simulation period, the visible relation between set E and set $S_{11,}$ S'_{11} and the other satellites.

The chart shows that for the ring topology set S_{11}, the three satellites are at close range and the visible interval overlap a lot, so that there is large blind area between S_{11} and E. In the blind area, we cannot find a satellite that is always visible to ground station, that is to say, for the ring topology, we cannot achieve full coverage to ground stations within two hops. Meanwhile, for the anise star topology set $S'_{11,}$ in the blind area from S11 to E, there are six satellites that are visible to E, M21, M22, M28, M36, M37, M38. Through the analysis and calculation can draw, when the M11 is not visible to the ground station, the visible probability of other satellites to ground station is shown in Table 31.1.

Among the six satellites that can build links, the ring topology chooses the two that has the smaller probability, M12 and M18. And the anise star topology

Table 31.2 The inter-plane ISLs based on the integration algorithm

Orbit 1–2	Orbit 1–3	Orbit 2–3
M11–M28	M12–M35	M23–M32
M13–M22	M14–M37	M24–M31
M15–M21	M16–M36	M25–M33
M17–M27	M18–M38	M26–M34

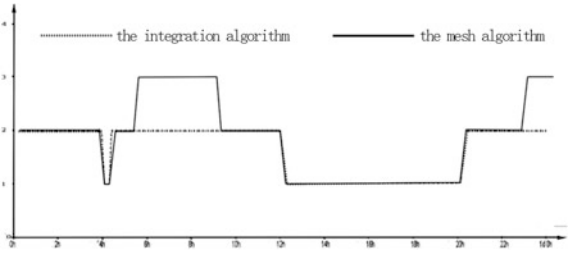


Fig. 31.5 The hops of two topologies algorithms

chooses the two that has the largest probability, M14 and M16. Based on the above consideration, from the point of view of a better guarantee satellite integration perspective, this paper choose the anise star topology as the intra-plane ISLs topology.

31.4.2.2 Analysis of ISLs Topology

Compares the satellite-earth communication performance and characteristics with the two different link allocation method through simulation, the first one is the integration algorithm proposed in this paper, the second one is the mesh algorithm which is similar to the iridium system.

Analyze and calculate according to constellation operation data, the integration algorithm forms links allocation as shown in Table 31.2.

Take the situation from M11 to ground station as an example, in a simulation cycle, observe the transmission of data. For the integration algorithm, the probability of one hop is about 43 %, the probability of two hops through intra-plane satellites is about 52 %, the probability of two hops through inter-plane satellite is about 5 %. For the mesh algorithm, the probability of one hop is also about 43 %, the probability of two hops through intra-plane satellites is about 30 %, the probability of two hops through inter-plane satellite is about 17 %, the probability of three hops through intra-plane satellites or inter-plane satellites is about 10 %. Figure 31.5 shows the hops of two algorithms, obvious the integration algorithm performs better, completing full coverage from satellites to earth within two hops.

In order to illustrate convenient, the simulation experiment will only signal propagation delay as the performance index link allocation algorithm, real-time extension only related to the distance. Figure 31.6 lists the data transmission time delay between satellite and ground with the two kinds of algorithm. As can be

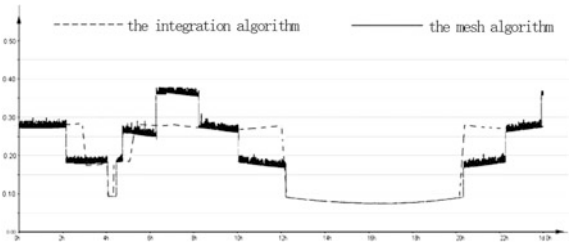


Fig. 31.6 The delay of two topologies algorithms

seen, compared to the mesh algorithm, the integration algorithm proposed in this paper to ensure the satellite integration performance with smaller transmission delay.

31.5 Conclusion

The existing ISLs research tend to optimize the communication performance between satellites as the goal, rarely consider the navigation constellation integration requirements. Based on the performance optimization of communication between satellite and ground station, this paper put forward a kind of link planning algorithm for satellite navigation system that can realize the data transmission hop count minimizes between satellite and ground station with better coverage. In this paper, a typical navigation system of algorithms are simulated, the simulation results show that in most of the time the constellation operation period, the hop and end-to-end transmission delay performance of integration algorithm in this paper is better than the mesh algorithm. So this algorithm can be conveniently applied to satellite navigation system of inter satellite network.

In addition, since the Beidou satellite navigation system of our country is mixed constellation composed of different types of navigation satellite, in the next step of the work, we will according to this situation for further study.

References

1. Luders RD (1961) Satellite networks for continuous zonal coverage. J Am Rocket Soc 31:179–184
2. Rajan JA (2002) Highlights of GPS IIR: autonomous navigation. In: ION 58th annual meeting, Albuquerque
3. Harathi K, Krishna P, Newman-Wolfe RE et al (1993) A fast link assignment algorithm for satellite communication networks. In: 12th annual international phoenix conference on computers and communications, Tempe, 23–26 March 1993
4. Noakes MD, Cain JB, Nieto JW et al (1993) An adaptive link assignment algorithm for dynamically changing topologies. IEEE Trans Commun 41(5):694–706
5. Lo MW (1999) Satellite constellation design. IEEE Comput Sci Eng 21(1):58–67

6. Wang Z-y, Wang P, Gu X-m (2006) Research on design of permanent inter-satellite-links in satellite networks. J Commun 27(8):129–133
7. Shi L-y, Xiang W, Tang X-m (2011) A link assignment algorithm applicable to crosslink ranging and data exchange for satellite navigation system. J Astronaut 32(9):1971–1977
8. Liu H, Zhang N (2000) Research on constellation visibility for global navigation satellite system. Syst Eng Electron 22(5):23–25
9. Shu-sen T (2008) Development and thought of compass navigation satellite system. J Astronaut 29(2):391–396
10. Stephen RP, Richard AR, Carl EF et al (1999) An operational and performance overview of the Iridium low earth orbit satellite system. In: Communications Surveys & Tutorials, Wright Patterson AFB

Chapter 32
Ill-Posedness Analysis of Ambiguity Resolution Using BDS-2 and Its Fusion with GPS

Wang Gao, Chengfa Gao, Shuguo Pan, Shengli Wang and Xiaofei Shi

Abstract Different from GPS or GLONASS, BDS (BeiDou Navigation Satellite System) consists of not only MEO satellites but GEO and IGSO satellites, both of which are high-orbit satellites. Now BDS has finished its second step of the all three steps, and can provide regional navigation and positioning service. BDS-2 now mainly consists of 5 GEO, 5 IGSO and 4 MEO satellites. For most places of China, the GEO satellites always gather in a south part of the visible satellites sky plot. The IGSO satellites are also often seen in the south part. These lead into poor satellite geometry. Besides this, the running angular velocities of GEO and IGSO satellites are much smaller compared with MEO satellites. This results in that the geometric relationship between the satellite and the station changes slowly, so that there exists strong correlation between the observation equations from adjacent epochs. The poor satellite geometry and the small velocities make the ambiguity resolution (AR) of BDS-2 be a more serious ill-posed problem. In this paper the ill-posedness of the AR of BDS-2 was analyzed. And on this basis, some different combinations of GEO, IGSO and MEO satellites of BDS-2 were used in the AR experiments to reveal the characteristic of ill-posedness. AR experiment of GPS was also carried out at the same time for comparison with BDS-2. At last, the

W. Gao (✉) · C. Gao · X. Shi
School of Transportation, Southeast University, 2 Sipailou, Nanjing 210096, China
e-mail: gaowang_1990@126.com; 1300317284@qq.com

C. Gao
e-mail: gaochfa@163.com

X. Shi
e-mail: shixiaofei.love@163.com

S. Pan · S. Wang
School of Instrument Science and Engineering, Southeast University, 2 Sipailou, Nanjing 210096, China
e-mail: panshuguo@gmail.com

S. Wang
e-mail: victory_wsl@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_32,

fusion AR experiment of BDS and GPS was done. From these experiments we concluded that BDS-2 AR which mainly relied on high-orbit satellites is a serious ill-posedness problem, so it takes much more time than GPS. The fusion with GPS will make the problem change much better.

Keywords BDS-2 · Ambiguity resolution · Ill-posed · Fusion with GPS

32.1 Introduction

BeiDou Navigation Satellite System (BDS) is one of the Global Navigation Satellite Systems (GNSS) in the world, which can offer various services including surveying, timing, navigation of transportation and so on [1, 2]. Since December 2012, BDS has finished its second step of the all three steps [1, 3] (so called BDS-2 in this paper), and now it can provide regional navigation and positioning service for Asian-Pacific region. Different from GPS, GLONASS and other existing GNSS systems, the BDS-2 includes not only the frequently-used MEO satellites, but also includes 5 GEO (geostationary orbit) satellites and 5 IGSO (inclined geosynchronous satellite orbit) satellites in its constellation [4, 5]. Both of them are high-orbit satellites and their orbital altitude is about 36,000 km, while the orbital altitude of BDS MEO satellites is just about 21,150 km [6]. The five GEO satellites lie over the equator, which means to almost all the places in China, the five satellites are always in the south part on the sky plot of visible satellites. Besides they are nearly in a straight line on the plot, in which they are not far from each other. IGSO satellites run over both the sides of equator equally. So to the most places especially the north part of China, the IGSO satellites are also commonly seen in the south part on the sky plot [7]. The four MEO satellites run around the earth in quick speed, so they cannot be observed for long time like the GEO and IGSO satellites. So most visible satellites of BDS-2 in China are distributed in the south half of sky plot. This causes the poor geometry of visible satellites.

As we all know the GEO satellites are almost static whose angular velocities are nearly zero. Although the IGSO satellites are not static, but due to their high orbits, their angular velocities are much smaller than that of MEO satellites. GNSS carrier phase high-precision positioning model is generally based on the theory of integer least squares. The parameters are usually estimated by the following three steps [8]: the float solution of parameters, the search of integer ambiguity and the fixed solution of ambiguity. At the initial period of AR, due to the existence of unknown coordinates vector and ambiguity parameters, the solution model is usually ill-posed. That is to say the ambiguity parameters cannot be solved in short time. The general method to solving the problem is to make the geometric relationship between satellites and stations change by long-term and multi-epoch observations. This change depends mainly on the variation of the angles between the stations

and satellites. The rate of the change depends on the satellite orbit altitude and operating characteristics.

In this paper, based on the operating characteristics of different satellites, the ill-posedness in AR of BDS-2 was analyzed. And then some different combinations of GEO, IGSO and MEO satellites of BDS-2 were used in the AR experiments to verify the characteristic of ill-posedness. AR experiment of GPS was also carried out at the same time for comparison. At last the fusion AR experiment of BDS and GPS was done to test the AR effect of multiple-constellation fusion.

32.2 Model of Carrier Double-Difference Ambiguity Resolution

The same as GPS, the model of BDS-2 double-difference (DD) carrier phase observables can be written as:

$$\lambda\Delta\nabla\varphi_{A,B}^{p,q}(t_i) = \Delta\nabla R_{A,B}^{p,q}(t_i) + \lambda\Delta\nabla N_{A,B}^{p,q} + \delta_{A,B}^{p,q}(t_i) \tag{32.1}$$

$$\delta_{A,B}^{p,q}(t_i) = -\Delta\nabla I_{A,B}^{p,q}(t_i) + \Delta\nabla T_{A,B}^{p,q}(t_i) + \Delta\nabla\delta m_{A,B}^{p,q}(t_i) + \Delta\nabla\varepsilon_{A,B}^{p,q}(t_i) \tag{32.2}$$

where $\Delta\nabla X_{A,B}^{p,q}(t_i)$ is double-difference value of X from receiver A, B and satellite p, q at the time t_i λ; is nominal carrier phase wavelength and φ is carrier phase observables; R is geometric range from a receiver and a satellite; N is integer ambiguity; I and T are dispersive atmospheric (ionospheric) delay and non-dispersive atmospheric delay, and in the later part of this paper, they will be neglected due to the very short baseline; m is multipath on carrier; ε is noise on carrier and other errors. Equation (32.1) can be further written as [9]:

$$L = Aa + Bb + e,\ P \tag{32.3}$$

where, L is the given GNSS data vector consists of DD carrier phase observables and modelled error; e is the noise vector; a and b represent the unknown parameter vectors (baseline vector and ambiguity vector respectively); A and B are the corresponding coefficient matrix which are assumed known; P is the weight matrix.

The procedure for solving the model like Eq. (32.3) can be divided into three steps. In the first step, float solution of ambiguity parameters and the co-factor matrix can be gotten with the standard least-squares adjustment as follows [9]:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = N^{-1}W, \qquad \begin{bmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{bmatrix} = N^{-1} \tag{32.4}$$

$$\begin{cases} N = (A\ B)^T P(A\ B) \\ W = (A\ B)^T PL \end{cases} \tag{32.5}$$

where the N represents the normal matrix; W represents the constant term matrix.

The second and the third step are to search the various combinations of integer ambiguities and fix the integer ambiguity respectively according to the float solution and the co-factor matrix. A widely used method in the two later steps is the Least-Square Ambiguity Decorrelation Adjustment (LAMBDA) method [10, 11]. This method was adopted in ambiguity searching and fixing in the next part of the paper.

To a linear observational equation system like Eq. (32.4), the condition number is often used as the measure index of the ill-posedness in the system [12]. For Eq. (32.4), it can be defined as follows:

$$\mathrm{cond}(N) = \| N^{-1} \| \quad \| N \| \tag{32.6}$$

where $\mathrm{cond}(N)$ is the condition number of N; $\|\bullet\|$ represents the operator of the matrix spectrum norm (as called 2-norm). The influence of condition number on the calculation system can be seen in [12, 13], which will not be detailed in this paper.

32.3 Analysis of Ill-Posedness in BDS-2 Ambiguity Resolution

The running angular velocity of GEO, IGSO satellite from BDS-2 and MEO satellite from GPS are shown in Fig. 32.1. It is needed to note that the MEO satellites from BDS have the similar orbital altitude and operating characteristic, so the angular velocities of two kinds MEO satellites are nearly the same in values.

As can be seen from Fig 32.1, angular velocity of GEO satellites are nearly zero due to its earth-synchronous character. And the angular velocity of IGSO satellites is just about 30–50 % of that of MEO satellites. This is consistent with their orbital altitude and operating characteristics. If the satellite has larger running angular velocity, the geometric relationship between the satellite and the station changes much more in the process of ambiguity resolution within the same observation time [6]. So the correlation between observations from adjacent epochs is mainly decided by the satellite running angular velocity. Due to the existence of ambiguity parameters, the AR system is usually ill-posed, so it generally requires observations from some epochs for a long period. There are five GEO satellites and five IGSO satellites in the current BDS-2, while only four MEO satellites around the earth. At most time in China, the visible satellites we can observe mainly consists of GEO and IGSO satellites. So it can be inferred that there is more serious ill-posedness in AR of BDS-2 compared GPS or GLONASS [10].

In the Sect. 32.1, we also mentioned that we often got poor satellite geometry if we just received the signals from the current BDS-2. Different from the influence of angular velocity, the satellite geometry mainly influence the correlation between different satellites in a certain epoch. In DD positioning calculation, the satellite

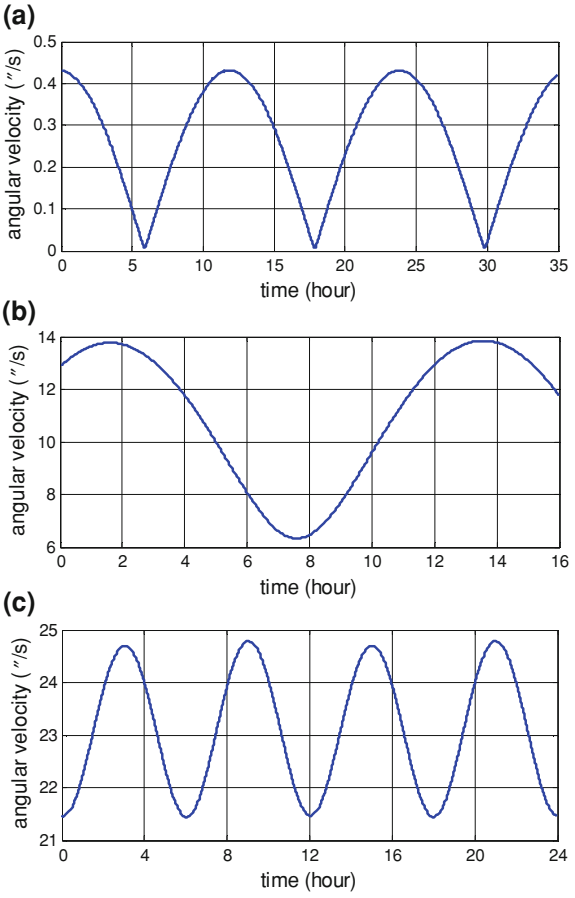


Fig. 32.1 The running angular velocity of satellites **a** angular velocity of GEO satellites **b** angular velocity of IGSO satellites **c** angular velocity of MEO satellites

geometry will become worse if two non-reference satellites are too close or satellites gather in an uneven area on the sky plot. Condition number with ambiguity fixed (ambiguity parameters are known) or DOP can be used to measure this influence. The poor satellite geometry of course will cause the ill-posedness worse [14].

We just want to analyse the ill-posedness of BDS-2 AR, so in order to avoid the influence of other factors (e.g. atmospheric error, multipath effects) we used a short-distance (about 3 m) baseline data including BDS and GPS observations to do some experiments. The date was collected on Nov 18th 2013 in Southeast University, Nanjing and the sample interval is 1 s. In order to do these experiments fully, we chose 1,000 continuous epochs, during which we could observe 4 GEO, 4 IGSO and 2 MEO satellites from BDS and 7 MEO satellites from GPS. The sky plot of visible satellites can be seen in Fig. 32.2.

So as to reflect the different influence from different satellites, we do the following experiments which are listed in Table 32.1 (Experiment 7 is in Part 4):

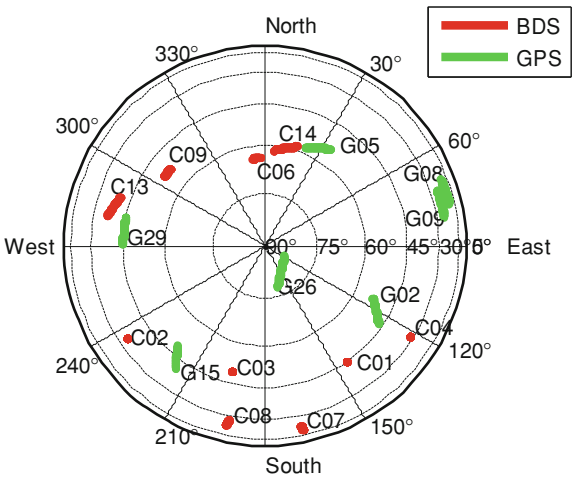


Fig. 32.2 The sky plot of BDS/GPS visible satellites

Table 32.1 Experiments details with different satellites combination

	Satellites number	Satellites combination	Satellites detail
1	7	4GEO+3IGSO	C01 C02 C03 C04 C06 C07 C08
2	7	3GEO+4IGSO	C01 C02 C03 C06 C07 C08 C09
3	7	3GEO+2IGSO+2MEO	C01 C02 C03 C06 C07 C13 C14
4	7	2GEO+3IGSO+2MEO	C01 C02 C06 C07 C08 C13 C14
5	10	4GEO+4IGSO+2MEO	C01 C02 C03 C04 C06 C07 C08 C09 C13 C14
6	7	7MEO	G29 G15 G26 G05 G08 G09 G02
7	14	4GEO+3IGSO+7MEO	C01 C02 C03 C04 C06 C07 C08 G29 G15 G26 G05 G08 G09 G02

Because we just want to research the ill-posedness caused by the satellite geometry and running angular velocity, so we do some processing with the DD carrier observations. We get the real and reliable ambiguities and station coordinates by processing the long-time observations. So we make each DD observation has the same precision with a random noise ranging from −1 to 1 cm according to Eqs. (32.1) and (32.3). In all the experiments, we use the ionosphere-free observations combined of the basis dual frequency observations, and the different DD observations are treated as the same weight.

In each experiment we calculate the float ambiguities bias (**FloatA**), the condition number at last with sequential adjustment using 1,000 epochs (**CondSA**), the mean PDOP (**MPDOP**), the mean condition number of the 1,000 epochs with ambiguity fixed (**MCondFix**, that is to say the ambiguities are not treated as unknown parameter, so this index just reflect the instantaneous situation of relative distribution of satellites), the fixed time with LAMBDA method (**FixTime**). The

Table 32.2 Calculation results of different experiments

Experiment no.	Time/s (FloatA < 1)	Time/s (FloatA < 0.5)	Fixtime /s	Cond SA	MCond-fix	MPD-OP
1	None	None	197	5.16e + 9	109.65	1.88
2	403	959	116	3.72e + 7	115.77	2.11
3	355	591	52	1.24e + 7	73.19	1.66
4	309	702	43	7.91e + 6	210.80	1.59
5	182	672	23	1.69e + 7	89.95	1.11
6	207	340	32	3.83e + 5	6.03	2.29
7	225	349	12	1.10e + 6	9.27	0.78

Note 'None' means that the FloatA cannot be within ±1 or ±0.5 in the 1,000 epochs

criteria that the ambiguities are fixed successfully is that the ambiguities are the same with the known correct ambiguities and the Ratio is always larger than 2 from the fixed epoch.

The results are listed in Table 32.2:

In the table FloatA < 1 and FloatA < 0.5 means that the absolute value of FloatA is smaller than 1 or 0.5, and the later results also meet the corresponding condition. The time of FloatA < 1 or FloatA < 0.5 indicates the convergence speed of float ambiguity solution. As we have mentioned, the observations has been processed to the same precision. So the AR difference among the several experiments is mainly caused by the different satellite operating characteristics and different satellite geometry.

In the experiment 1–4, the satellites with larger running angular velocity were added gradually to replace the satellites with smaller running angular velocity. As we can see the AR effect becomes better and better on the whole. This is because the larger running angular velocity will make the geometric relationship between the satellite and the station change more quickly, so the ill-posedness can be reduced more quickly [6, 12]. It is important to note that the 'MCondFix' may reflect the DD observation structure better compared with MPDOP. The reason is that MCondFix is gotten from the DD model and it can reflect the correlation among different DD observation equations, while the PDOP just reflect the open degree of satellite distribution. For example when two non-reference satellites are quite close to each other on the sky plot, the MCondFix will reflect the phenomenon while the PDOP cannot reflect that. From Table 32.2 we can see that the MCondFix of GPS is much better than that of BDS-2. Actually it can be also reflected in Fig. 32.2, where we can see the GPS satellites have much more even distribution.

Experiment 5 includes all the visible satellites of BDS-2. We can see that the more satellites, the better the AR effect will be. Actually when there are more satellites even the GEO satellites, the correlation among the unknown ambiguities will be reduced because they provide more redundant observations to their common unknown parameter i.e. the baseline vector. But the addition of GEO satellites will make the condition number larger like the change from Experiment 3 to 5.

From Experiment 6 we can see that the AR effect of GPS is much better compared with Experiment 1, 2, 3, 4, in which the number of satellites is all seven. The reason is that the GPS satellites are all MEO satellites and they have better satellite geometry.

In order to reflect the difference of AR effect in detail, the results of three representative experiments (Experiment 1, 3, 6) are shown in Figs. 32.3, 32.4 and 32.5.

Figure 32.3a–c show the float ambiguities bias; Fig. 32.4 shows the condition number of the normal matrix with ambiguity estimated; Fig. 32.5 shows the Ratio during the search of ambiguity.

The numbers of satellites used in the three experiments are all the seven. The main difference lies in the composition of GEO, IGSO and MEO satellites. As we can see there is significant difference among the float ambiguities bias, especially in Experiment 1 in which the float ambiguities bias are still over 10 cycles after 500 epochs. In Fig. 32.4, the condition number indicates the ill-posedness of the whole calculation system including baseline vector and ambiguities. We can see Experiment 6 consisting of all GPS satellites has the smallest condition number during the calculating process of the three, while Experiment 1 has the largest. This is corresponding with the float ambiguities bias and the ambiguity fixed effect. From Fig. 32.5 we can also see Experiment 6 has the best Ratio of the three experiments especially at beginning of the search, and Experiment 1 has the worst one.

32.4 Fusion of BDS-2 and GPS Ambiguity Resolution

From part 3, we analyzed that the current BDS-2 has much stronger ill-posedness in AR process especially when its MEO satellites cannot be observed. At the present stage, we can use BDS-2 and GPS satellites together to the ambiguity resolution adjustment, as they have the common unknown parameters, the baseline vector or others. There are some exact differences between the two systems, such as the time system, the coordinate reference system and so on. The time system can be unified according [15]. There is just a little difference between the two coordinate reference systems so in the DD model, the baseline vector can be seen the same for the two system. It is better to choose their own reference satellites respectively. This can avoid some complex error terms like the error caused by different channels.

We combined the seven high-orbit satellites from Experiment 1 and the seven GPS satellites from Experiment 6 together to calculate ambiguities (as the Experiment 7 in Table 32.1). The float ambiguities bias can be seen in Fig. 32.6; the condition number with ambiguity estimated and the Ratio can be seen in Figs. 32.4 and 32.5 respectively. The related results are listed in Table 32.2 (the last line of the table).

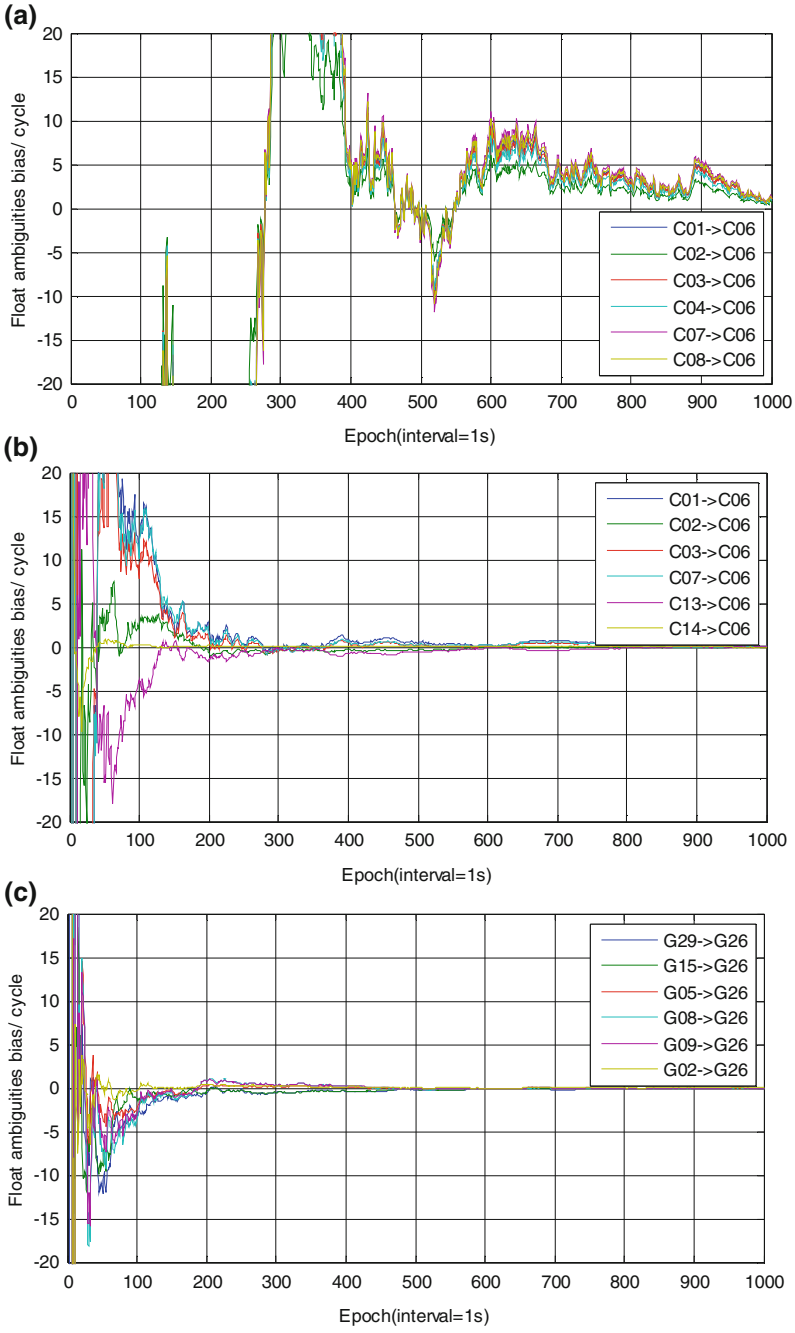


Fig. 32.3 Float ambiguities bias of Experiment 1, 3, 6 **a** Float ambiguities bias of Experiment 1. **b** Float ambiguities bias of Experiment 3. **c** Float ambiguities bias of Experiment 6

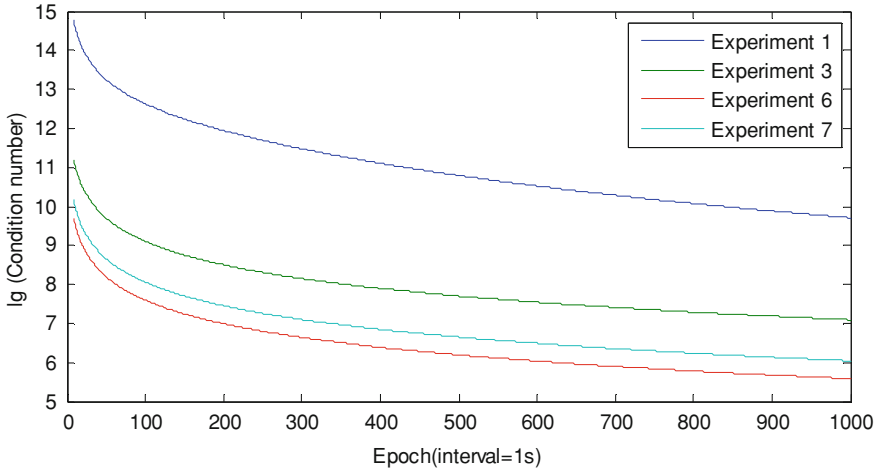


Fig. 32.4 Condition number of the calculation system of different experiments

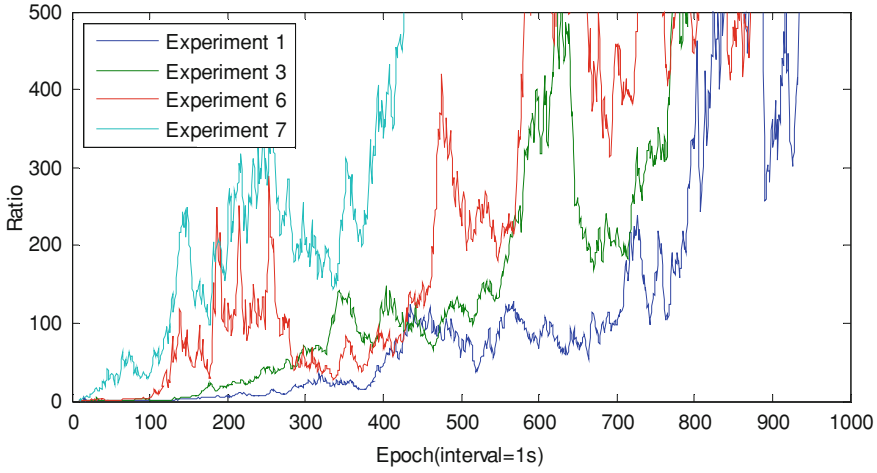


Fig. 32.5 Ratio during the ambiguity search of different experiments

From Fig. 32.6, we can see the float ambiguities change much better than that of Experiment 1. In Fig. 32.4 the condition number also become much smaller than that of Experiment 1 and 3. Due to the introduction of high-orbit satellites, the condition number is slightly larger than GPS alone. But due to the increase of satellites number, the redundant observations increase. So the correlation among the unknown ambiguities is reduced. This is also can be seen from Table 32.2,

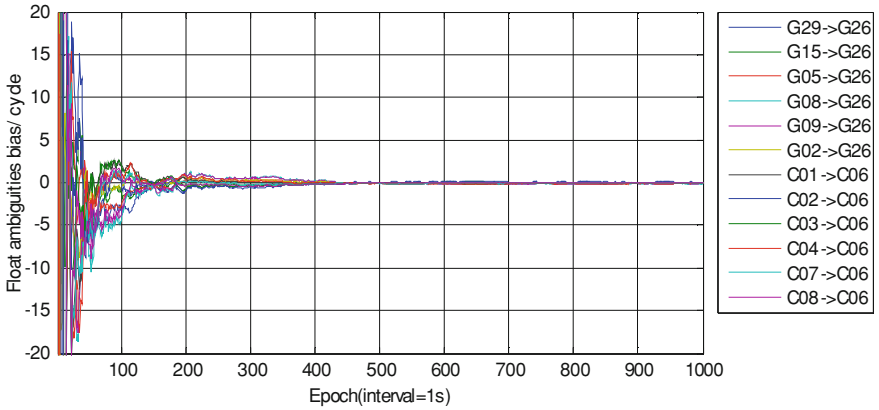


Fig. 32.6 Float ambiguities bias of Experiment 7 (Fusion of BDS-2 and GPS)

where the ambiguity fixed time is largely decreased to only 12. This is also better than that of GPS alone or the BDS-2 with ten satellites (Experiment 5). The Ratio is also larger than other experiments in Fig. 32.5.

32.5 Conclusions and Advice

The ill-posedness of the AR of BDS-2 was discussed in the paper, and some experiments with different satellites were carried out to analyze the problem in details. In the current BDS-2, ten high-orbit satellites are the main components and the visible satellites we can often observe in China are also them. High-orbit satellites especially the GEO satellites have the much smaller running angular velocity, so the geometric relationship between the satellite and the station changes slowly. In DD model, this will cause the equations from adjacent epochs have more serious correlation. Besides, the GEO and IGSO satellites mainly gathered in the south part on the sky plot in China especially the northern regions. When the four earth-around MEO satellites cannot be observed, it often comes with the poor satellite geometry. The main two factors cause the current BDS-2 AR as a more serious ill-posed problem compared with GPS. The experiments carried out in the paper have also verified this. The fused ambiguity resolution of BDS-2 and GPS can reach much better effect. So to the current BDS-2 especially the situation only high-orbit satellites can be observed, it is better to fuse GPS and BDS-2 together to DD ambiguity resolution if possible.

Acknowledgments This work is supported by the Key Projects in the National Science and Technology Pillar Program during the Twelfth Five-year Plan Period (No. 2012BAJ23B01). The authors are very grateful to the anonymous reviewers for their constructive comments and suggestions.

References

1. Yang Y, Li J, Xu J et al (2011) Contribution of the compass satellite navigation system to global PNT users. Chin Sci Bull 56(26):2813–2819
2. Yang Y (2010) Progress, contribution and challenges of Compass/Beidou satellite navigation system. Acta Geodaetica Cartogr Sin 39(1):1–6
3. Ran C (2012) Development of the BeiDou navigation satellite system. Global navigation satellite systems. In :Report of the joint workshop of the national academy of engineering and the chinese academy of engineering, Washington
4. Zhao Q, Guo J, Li M, et al (2013) Initial results of precise orbit and clock determination for COMPASS navigation satellite system[J]. J Geodesy 87(5): 475–486
5. Zhou S, Cao Y, Zhou J et al (2012) Positioning accuracy assessment for the 4GEO/5IGSO/2MEO constellation of compass. Sci China Phys Mech Astron 55(12):2290–2299
6. Gao W, Gao C, Pan S et al (2013) The Analysis of ill posedness in GNSS high-precision differential positioning. In: Proceedings of the China satellite navigation conference (CSNC 2013). Springer, Berlin, pp 311–321
7. He H, Li J, Yang Y, et al (2013) Performance assessment of single-and dual-frequency BeiDou/GPS single-epoch kinematic positioning. GPS Solutions http://link.springer.com/article/10.1007/s10291-013-0339-3
8. Guo Q, Hu Z (2007) The effect of ill-condition equations of GPS rapid positioning on the GPS baseline solution. Sci Survey Mapp 32(2):42–43
9. Teunissen P (2007) Influence of ambiguity precision on the success rate of GNSS integer ambiguity bootstrapping. J Geodesy 81(5):351–358
10. Teunissen P (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geodesy 70(1–2):65–82
11. Teunissen P, De J, Tiberius C (1997) Performance of the LAMBDA method for fast GPS ambiguity resolution. Navigation 44(3):373–383
12. Wang S, Wang Q, Gao W et al (2013) Analysis and valuation of ill-condition in baseline solution of GNSS multi-system. J Southeast Univ (Natural Science Edition) 43(4):753–757
13. Lu X, Feng Z, Liu J (2007) Analysis theory on ill-conditioned system with application in surveying. Surveying and Mapping Press, Beijing, pp 7–15
14. Li B, Shen Y, Feng Y (2010) Fast GNSS ambiguity resolution as an ill-posed problem. J Geodesy 84(11):683–698
15. Deng C, Tang W, Liu J et al (2013) Reliable single-epoch ambiguity resolution for short baselines using combined GPS/BeiDou system. GPS Solutions, http://link.springer.com/article/10.1007/s10291-013-0337-5

Chapter 33
Extraction and Application of Un-differenced Atmospheric Delays with Un-combined Precise Point Positioning Technique

Denghui Wang, Chengfa Gao, Shuguo Pan, Yang Yang and Shengli Wang

Abstract A new method for the extraction and application of un-differenced atmospheric delays with un-combined precise point positioning technique, is proposed based on the regional continuous operational reference system (CORS) network. The estimable state vector, which lets in the receiver clock offset, the ionospheric slant delay and zenith tropospheric wet delay (ZWD), is estimated epoch by epoch at every reference station. In addition, the least squares and the ionospheric regional polynomial model is also used to separate the DCB and the ionospheric delay. For improving the convergence time of the PPP, a new PPP algorithm, which was suited for single-frequency and dual-frequency receivers, was put forward. The algorithm took full use of the interpolation of regional atmospheric delay value as the quasi-observable, and the un-combined satellite-different observations were used to obtain the precise positioning results. Nevertheless, the method mentioned above is investigated and examined on the IGS tracking stations. The results show that all the satellites' DCB are solved and compared with the products published by CODE, the disparity between them is mostly not greater than 0.2 ns. Furthermore, with the aim of the real-time fast

D. Wang (✉) · C. Gao · Y. Yang
School of Transportation, Southeast University, 2 Sipailou, Nanjing 210096, China
e-mail: owenxun@gmail.com

C. Gao
e-mail: gaochfa@163.com

Y. Yang
e-mail: johnny2014@sina.com

S. Pan · S. Wang
School of Instrument Science and Engineering, Southeast University, 2 Sipailou, Nanjing 210096, China
e-mail: panshuguo@gmail.com

S. Wang
e-mail: victory_wsl@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_33,

precise point positioning based on the sparse regional CORS, results from American CORS data set is introduced and discussed. The positioning accuracy can reach 10 cm in the plane within 30 min for the single-frequency PPP user and the plane position result in 1–2 cm also can be accomplished after the convergence of the filter. For dual-frequency receiver, it takes approximately 10 min to get positions with accuracy better than 10 cm at the three directions for the dual-frequency PPP user. On the other hands, due to its epoch-by-epoch nature, the method proposed above is also suited for dynamic precise point positioning.

Keywords Un-differences atmospheric delays · Un-combined PPP · DCB · Ionosphere-weighted model

33.1 Introduction

Real-time precise point positioning technique has achieved more and more attention due to its convenience, global coverage and unified coordinate frame. But for now, real-time PPP usually takes about 30 min in order to obtain positions with accuracy better than 10 cm [1, 2]. At present, the research shows that an accuracy level of centimeter can be achieved for both combined and un-combined PPP models [3]. And the atmospheric delay effects, which are the most important factor which restricts the further improvement of the accuracy and the convergence time of precise point positioning users, should be accurately estimated [4, 5].

To extract the un-differenced atmospheric delay and improve the PPP user convergence time, a new method is proposed based on the regional continuous operational reference system (CORS) network. The unknown parameter, which included the receiver clock offset, the ionospheric slant delay and zenith tropospheric wet delay (ZWD), was estimated epoch by epoch at every reference station. And the un-combined PPP technology was also employed in the method. Due to the effect of the receiver and satellite differential code bias (DCB), the least squares and the ionospheric regional polynomial model is used to separate the DCB and the ionospheric delay. Moreover, the temporal and spatial constraints imposed on ionospheric parameters are also presented in the least square equation. For improving the convergence time of the PPP, the true ionospheric delay and the zenith tropospheric wet delay at each reference station is used to interpolate the user's atmospheric delay. The ionosphere-weighted model is introduced into users' precise point positioning, taking full use of high-precision ionospheric delay information offered by regional sites. The new method can effectively improve the convergence time and positioning accuracy of single-frequency or dual-frequency PPP user.

Nevertheless, the method mentioned above is investigated and examined on the IGS tracking stations. The results indicate that all the satellites' DCB are solved and compared with the products published by CODE, the disparity between them

is generally not greater than 0.2 ns. Furthermore, with the aim of the real-time fast precise point positioning based on the regional CORS, results from American CORS data set is presented and discussed. For dual-frequency receiver, it took approximately 10 min to get positions with accuracy better than 10 cm and more gentle result and faster convergence time could be achieved. What is more; the positioning accuracy could reach 10 cm in the horizontal direction within 30 min for the single-frequency PPP user. And after the convergence of the filter, the horizontal position result in 1–2 cm also could be accomplished. An order of magnitude was improved compared with the positioning result in 10–20 cm using UofC model or the GIM model. On the other hands, due to its epoch-by-epoch nature, the method proposed above is also suitable for dynamic precise point positioning.

In this paper, the main principle and background is introduced in this section. The un-combined observation and DCB parameter is analyzed in Sect. 33.2. In Sect. 33.3, the Kalman filter model of the extraction of un-differenced atmospheric delays is discussed in detail. We explain the least square equation with the constraint conditions for the purpose of separating the ionospheric delay and DCB. At the end of this section, the Kalman filter model utilizing the interpolation of ionospheric delay is proposed and described. In Sect. 33.4, the test and the result, including the comparison between the estimated satellite's DCB result and the CODE product, the difference between the ionospheric delay with un-combined precise point positioning technique and the GIM product, and the effect of this approach on the regional precise point positioning, is shown and discussed. Finally, the concluding remarks are described in Sect. 3.3.5.

33.2 The Observation Equations

For GNSS data processing, the four original measurements from satellite*s* to receiver *k*, un-differenced observation equations could be written as:

$$\begin{aligned} P^s_{j,k} &= \rho^s_k + c\delta t_k - c\delta t^s + T^s_k + \alpha_j I^s_k + O^s_k + d_{k,P_j} - D^s_{P_j} + d^s_{k,others} + \varepsilon^s_{k,P_j} \\ \Phi^s_{j,k} &= \rho^s_k + c\delta t_k - c\delta t^s + T^s_k - \alpha_j I^s_k + O^s_k + b_{k,\Phi_j} - b^s_{\Phi_j} + d^s_{k,others} + \lambda_j N_j + \varepsilon^s_{k,\Phi_j} \end{aligned} \tag{33.1}$$

where:*j* is the frequency number; $P^s_{j,k}$ is pseudorange measurement (m); $\Phi^s_{j,k}$ is carrier-phase measurement (m); ρ^s_k is geometric range between satellite and receiver antenna phase center (m). Due to the high-precision coordinates of the reference stations, the distance is exactly known. δt_k and $c\delta t^s$ is the receiver and satellite clock offset respectively, T^s_k represent troposphere delays. In this study, the meteorological inputted parameters are composed of temperature and pressure from global pressure and temperature model (GPT) and the partial pressure of water vapor from Unb3 model. Global mapping function (GMF) is used to be the

mapping function of the priori troposphere model, which the zenith troposphere dry delay is computed by Saastamoinen model. I_k^s represents ionospheric delays, which is related to the signal frequency. α_j is the frequency scaling, set as: $\alpha_j = f_1^2/f_j^2$, i.e., $\alpha_1 = f_1^2/f_1^2$ for L1 frequency and $\alpha_2 = f_1^2/f_2^2$ for L2 frequency. O_k^s is satellite orbital error (m); d_{k,P_j}, $D_{P_j}^s$ are receiver and satellite DCB between pseudoranges. It's important to note that the bias is related to the frequency. $d_{k,others}^s$ is the other error which influences the measurement, including tidal correction, plate motions error, multipath delay etc. b_{k,Φ_j}, $b_{\Phi_j}^s$ are receiver and satellite carrier phase fractional bias at different frequencies. λ_j is carrier wavelength and N_j is carrier-phase ambiguity (cycle) at frequency j; ε_{k,P_j}^s, ε_{k,Φ_j}^s is measurement errors of carrier-phase and pseudorange.

33.2.1 The Ionosphere-Free Observation Model

In the traditional PPP method, the ionosphere-free observation model is generally used to eliminate the effects of first-order ionospheric delay. And the residual higher order terms of ionospheric delay is usually less than 1 %. Due to error propagation function, it also amplifies measurement noise to the original measurement nonetheless. Although the residual higher order terms of ionospheric delay is usually less than 1 % of the total amount of ionospheric delay, but it can result in tens of centimeters of equivalent ranging error in the solar maximum. The ionosphere-free observation model could be written as:

$$\begin{aligned} P_{if,k}^s &= \rho_k^s + c\delta t_k - c\delta t^s + T_k^s + h_k \sin E_k^s + O_k^s + d_{k,P_{if}} + m_{k,P}^s - D_{P_{if}}^s + d_{k,others}^s + \varepsilon_{k,P_{if}}^s \\ \Phi_{if,k}^s &= \rho_k^s + c\delta t_k - c\delta t^s + T_k^s + h_k \sin E_k^s + O_k^s + b_{k,\Phi_{if}} + m_{k,\Phi}^s - b_{\Phi_{if}}^s + d_{k,others}^s + \lambda_{if} N_{if} + \varepsilon_{k,\Phi_{if}}^s \end{aligned} \tag{33.2}$$

From Eq. (33.2), $d_{k,P_{if}}$, $D_{P_{if}}^s$ are receiver and satellite ionosphere-free DCB. In the PPP processing, the satellite clock bias, released by IGS correction products, is used frequently. It is worth noting that $D_{P_{if}}^s$ in Eq. (33.2) can be eliminated by IGS released satellite clock bias products.

33.2.2 The Influence of Hardware Delay on Un-combined Model

The un-combined model is put forward by M.O. Keshin, A.Q. Le etc. at Delft University of Technology. The four kinds of un-differenced GNSS measurements are used to reflect the influence of all kinds of errors and atmospheric changes.

In general, however, for the un-combined model, the unbiased hardware delay item is divided into three parts; Firstly, the satellite ionosphere-free DCB, which related to the satellite, can be eliminated by the satellite clock bias released by IGS. Secondly, the receiver ionosphere-free DCB, which is irrelevant to the frequency, would be absorbed through the receiver clock bias. Thirdly, the rest of hardware delay, related to the frequency in data processing, is absorbed by each satellite ionospheric delay. It can be described as:

$$c\delta t_k' = c\delta t_k + d_{k,P_{if}}, c\delta t_{IGS}^s = c\delta t^s + D_{P_{if}}^s, I_k^{s\prime} = I_k^s + \frac{f_2^2}{f_1^2 - f_2^2}(B_k - B^s) \quad (33.3)$$

where: $c\delta t_{IGS}^s$ and I_k^s denote the satellite clock bias released by IGS and the true ionospheric delay between satellite and receiver respectively [6, 7]. I_k^s' denotes the unknown parameter which would be estimated in the equation. B_k, B^s is receiver and satellite differential code bias between P1 and P2, they both read: $B_k = d_{k,P_2} - d_{k,P_1}$, $B^s = D_{P_2}^s - D_{P_1}^s$.

It's important to note that the types of receiver should be considered carefully according to different types of measurement capture technology. The IGS tracking station tested in the paper can receive the precision code on both L1 and L2 frequency. And the American CORS receivers can receiver coarse/acquisition (C/A) code on L1 frequency instead of P1.

33.2.3 The Atmosphere Model

The troposphere

The troposphere delay could be obtained from the empirical troposphere models, for example, Hopfield, Saastamoinen and UNB3. These models give the similar value when the satellite's elevation angle is upon 20 degree. In this study, troposphere delay would be divided into two parts: dry component parts and wet component parts. The dry component parts is computed by Saastamoinen model, with the meteorological inputted parameters, composed of temperature and pressure from global pressure and temperature model (GPT) and the partial pressure of water vapour from Unb3 model. Global mapping function (GMF) is used to be the mapping function of the troposphere model. The correction accuracy can reach 5 to 8 mm. For the tropospheric wet component parts, the zenith wet delay is estimated as an unknown parameter in the characteristics of a random process.

After the un-differenced tropospheric wet delay is obtained, the interpolation of the zenith wet delay is applied to remain the residual tropospheric zenith delay for PPP users.

The ionosphere

For the purpose of comparing the ionospheric delay with un-combined PPP technique, the global ionosphere maps model (GIM) is mentioned in the paper. The GIM model could not provide real-time ionosphere parameters and the accuracy is not good enough, which influences the single-frequency user getting high-accuracy results. In the extraction of the ionospheric delay, the slant ionospheric delay of each satellite is also estimated as an unknown parameter in the characteristics of a random process [8, 9].

In the paper, polynomial curve-fitting model is used to separate out the true ionospheric delay and the hardware delay. The ionosphere-weighted model is used to give a reasonable weight of the ionospheric interpolation.

33.3 The Extraction and Application of Un-differenced Atmospheric Delays

In this section, we introduce a new method for extraction and application of un-differenced atmospheric delays with un-combined PPP technique. The observation model and the dynamic model are introduced below in the adopted Kalman filter.

33.3.1 Atmospheric Delay Estimated

Using the known coordinates of the reference stations, the estimable state vector, which includes the receiver clock offset, the ionospheric slant delay and zenith tropospheric wet delay (ZWD), is estimated epoch by epoch at every reference station [10]. In case of the receiver collect phase and code observations from n satellites at epoch i, the extended Kalman filter is written as:

$$\begin{cases} X_i = \Phi_{i,i-1}X_{i-1} + W_i & E(W_i) = 0, \mathrm{Cov}(W_i) = Q_i \\ L_i = B_i X_i + V_i & E(V_i) = 0, \mathrm{Cov}(V_i) = R_i, \mathrm{Cov}(V, W) = 0 \end{cases} \tag{33.4}$$

The vector of observables L_i identifies for four UD measurements. The design matrix B_i could be expressed as:

$$B_i = \begin{bmatrix} B_{MF_w} & B_{\delta t'_k} & B_I & 0 & 0 \\ B_{MF_w} & B_{\delta t'_k} & \frac{f_1^2}{f_2^2}B_I & 0 & 0 \\ B_{MF_w} & B_{\delta t'_k} & -B_I & B_{N_1} & 0 \\ B_{MF_w} & B_{\delta t'_k} & -\frac{f_1^2}{f_2^2}B_I & 0 & B_{N_2} \end{bmatrix} \tag{33.5}$$

where:

$$B_{MF_w} = \begin{bmatrix} MF_w(\theta^1) & \cdots & MF_w(\theta^n) \end{bmatrix}^T, B_{\delta t'_k} = \overbrace{\begin{bmatrix} -c & \cdots & -c \end{bmatrix}}^{n\times 1}{}^T$$

$$B_I = \mathrm{diag}(\overbrace{\begin{matrix} 1 & \cdots & 1 \end{matrix}}^{n\times n}), B_{N_1} = \mathrm{diag}(\overbrace{\begin{matrix} \lambda_1 & \cdots & \lambda_1 \end{matrix}}^{n\times n}), B_{N_2} = \mathrm{diag}(\overbrace{\begin{matrix} \lambda_2 & \cdots & \lambda_2 \end{matrix}}^{n\times n})$$

$MF_w(\theta^1)$ is the mapping function of ZWDs with the satellite elevation θ.

The vector of unknown parameters is split into a temporal part and a non-temporal part. The temporal part includes zenith tropospheric wet delay, receiver clock offsets and n-dimensional state vector of ionospheric slant delay and the non-temporal part consists of the original UD ambiguities:

$$X_i = \begin{bmatrix} ZTD_{w,k} & \delta t'_k & I_k^{s\prime} & N_1^s & N_2^s \end{bmatrix}^T, (s = 1 \cdots n) \tag{33.6}$$

$\Phi_{i,\ i-1}$ is the $(3n + 2) \times (3n + 2)$ state transfer matrix, set as:

$$\Phi_{i,i+1} = \mathrm{diag}(\Phi_{ZTD_{w,k}}, \Phi_{\delta t'_k}, \Phi_{I_k^{s\prime}}, \Phi_{N_1^s}, \Phi_{N_2^s}) \tag{33.7}$$

where: $\Phi_{ZTD_{w,k}} = \exp(-\frac{\Delta t}{\tau_{ZTD_{w,k}}})$, $\Phi_{\delta t'_k} = 0$, $\Phi_{I_k^{j\prime}} = \exp(-\frac{\Delta t}{\tau_{I_k^{j\prime}}})$

$$\Phi_{N_1^j} = \Phi_{N_2^j} = E_n.$$

Q_i is the $(3n + 2) \times (3n + 2)$ dynamic noise correlation matrix, set as:

$$Q_i = \mathrm{diag}(Q_{ZWD_k}, Q_{\delta t'_k}, Q_{I_k^{s\prime}}, Q_{N_1^s}, Q_{N_2^s}). \tag{33.8}$$

In the Eq. (33.8), the dynamic model for ZWDs, receiver clock offset, ionospheric slant delay is different rely on its varied dynamic characteristics. For ZWDs parameters, it is considered as random walk: $Q_{ZWD_k} = q_{ZWD}\Delta t$, $q_{ZWD} = 1 \sim 9\,\mathrm{cm}^2/\mathrm{h}$. Δt is the time interval between epochs. The white noise process is simple and effective to describe the random process of receiver clock offset. It is the equivalent of that the clock offset could be eliminated in the double difference relative positioning: $Q_{\delta t'_k} = E_n \otimes (q_{\delta t'_k} \cdot q_{\delta t'_k})$. For ionospheric slant delay parameters, it is also considered as random walk related to the satellite zenith angle z' on the position of ionospheric puncture: $Q_{I_k^{s\prime}} = q_{Ion}\Delta t/\cos(z')$, $q_{Ion} = 0.1\ \mathrm{m}^2/\mathrm{s}$. The ambiguity is regarded as an invariant parameters: $Q_{N_1^s} = Q_{N_2^s} = E_n \otimes 10^{-16}\Delta t$.

33.3.2 Separation of DCBs

As mentioned in Sect. 33.2 of this paper, the adoption of the satellite clock products issued by IGS and the un-combined measurements would lead to the offset of the

ionospheric delay value. And it would influence the PPP processing of the users. So it is necessary to separate the DCBs to get the true ionospheric slant delay.

For CORS, the regional VTEC polynomial curve-fitting model is established and the temporal and spatial constraints imposed on ionospheric parameters are also presented in the least square equation. Each satellite compared to the reference station's location, need to compute its ionospheric puncture point position (IPP) in the average ionospheric height. The ionosphere thin layer model, assumed that the ionosphere height is an average of 350 km (generally set to 300–400 km), is used for regional CORS [11]:

$$\begin{aligned} I_k^s &= I_{VTEC}^s / \cos(z') \\ I_k^s &= a_{00} + a_{10}dB + a_{01}dS + a_{11}dBdS + a_{20}dB^2 + a_{21}dB^2dS \\ dS &= dL - dL_0 + (t - t_0) \end{aligned} \tag{33.9}$$

where: I_k^s is true ionospheric delay which is mentioned in Eq. 33.4; I_{VTEC}^s is the satellite vertical ionospheric delay; a_{00}, a_{10}, a_{01}, a_{11}, a_{20}, a_{21} are coefficients of polynomial curve-fitting model [12]; dB, dL is difference of latitude and longitude between IPP and station's location; t is the epoch time and t_0 is reference time.

This article assumes that the above polynomial model parameters are related to time subsection and the interval length is set to 1 h. What is more, the constraints between each time subsection should be presented in the least square equation. The benchmark that the sum of all satellite's DCB the mean should be zero is also considered.

33.3.3 Regional Enhanced PPP

For the purpose of improving the convergence time of precise point positioning, the un-differenced atmospheric delay would be used as the predicted quasi-observable. Two steps need to be completed: Firstly, the interpolation of the atmospheric delay needs to be done to generate the ZWD of the user's location and the ionospheric slant delay of each satellite. Secondly, the Kalman filter should be adopted, utilizing the interpolation of ionospheric delay.

In the Network RTK technology, many scholars have done deep research on the implementation methods. Different correction terms of users' location and influences on positioning can be generated when different interpolation methods utilized. In the paper, the linear combination model (LCM) is transplanted on PPP processing [13, 14].

Instead of using the ionosphere-free observations and GIM model, the un-combined satellite-different observations are used to obtain the precise positioning results, which the ionosphere-weighted model is also applied [15]. In the PPP algorithms proposed next, the receiver DCB correction and clock offset can be removed based on the difference between satellites and the satellites' DCB

correction, which is estimated by all multiple reference stations, could be sent to the users. And if the satellites' DCB is not corrected, it still can be absorbed by the ambiguity at different frequencies.

Therefore, the vector of unknown parameters can be determined by utilizing the interpolation of atmospheric delay. The interpolation of the zenith wet delay is applied to remain the residual tropospheric zenith delay for PPP users, proving the pseudo observation equation. Assuming that the user's receiver could obtain n satellites' interpolation of slant ionospheric delay, we can add $n + 1$ quasi-observable equation into the vector of observables. The vector of observations L_i could be written as:

$$L_i = \begin{bmatrix} \left[\nabla P_1^{1,r}(i) \quad \cdots \quad \nabla P_1^{n,r}(i)\right]^T, \left[\nabla P_2^{1,r}(i) \quad \cdots \quad \nabla P_2^{n,r}(i)\right]^T \\ \left[\nabla \Phi_1^{1,r}(i) \quad \cdots \quad \nabla \Phi_1^{n,r}(i)\right]^T, \left[\nabla \Phi_2^{1,r}(i) \quad \cdots \quad \nabla \Phi_2^{n,r}(i)\right]^T \\ \left[B_{STEC}^1 \quad \cdots \quad B_{STEC}^n \quad B_{ZTD}^1\right]^T \end{bmatrix} \tag{33.10}$$

where:∇ is single-difference between satellites and B_{STEC}^s denotes the each satellites' interpolation of slant ionospheric delay. The dimensions of its observation matrix are $(5n - 3) \times 1$ for two frequencies. And it is also suitable for single-frequency receivers, which the dimensions are $(3n - 1) \times 1$, due to the lacking of carrier-phase and pseudorange at frequency 2. The vector of unknown parameters could be described as:

$$X_i = \left[\delta x \quad \delta y \quad \delta z \quad zwd \quad I_k^s \quad \Delta N_1^{s,r} \quad \Delta N_2^{s,r}\right]^T, (s = 1 \cdots n) \tag{33.11}$$

where: δx, δy, δz is receiver's position vector, I_k^s is the $n \times 1$ satellites' slant ionospheric delay vector and $\Delta N_1^{s,r}$, $\Delta N_2^{s,r}$ is the $(n - 1) \times 1$ ambiguity vector at different frequencies. The initial noise matrix R_i can be written as

$$R_i = \mathrm{diag}(\sigma_{P_1}^2, \sigma_{P_2}^2, \sigma_{\phi_1}^2, \sigma_{\phi_2}^2, \sigma_{ion}^2, \sigma_{zwd}^2) \tag{33.12}$$

In the paper, the observable noise is set as:

$$\sigma_{P_1} = \sigma_{P_2} = 0.3\,m,\ \sigma_{\Phi_1} = \sigma_{\Phi_2} = 0.003\,\mathrm{m}$$

The determination of σ_{ion} is chosen as 0.1 m, σ_{zwd} is set as 0.03 m, according to the interpolation accuracy of the ionospheric delay which would be shown in Sect. 33.4.

33.4 Test Results and Analysis

DCB precision analysis

As mentioned in this paper, satellite differential code bias precision is analyzed. The experiment selected the 10 IGS tracking stations on 3, 10, 17, 21, June, 2013. The four days data, sampling interval of 30 s, is used to calculate the DCB valuation accuracy. The information on each site is as follows (Fig. 33.1):

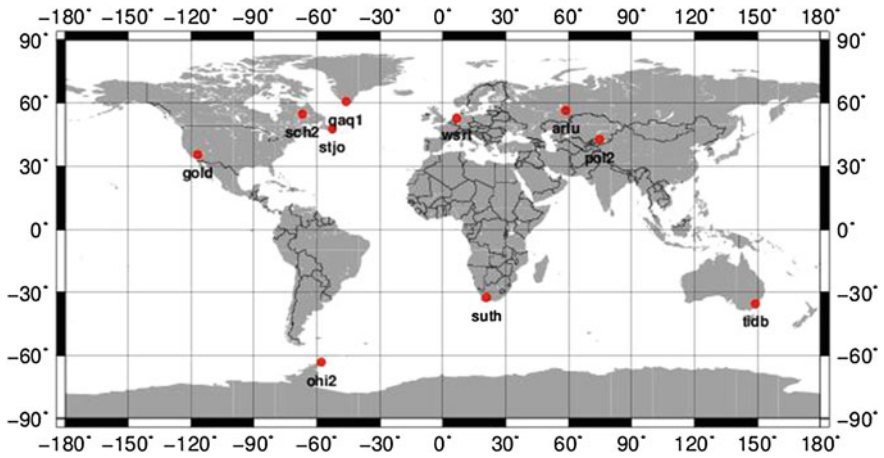


Fig. 33.1 The distribution of 10 IGS tracking stations

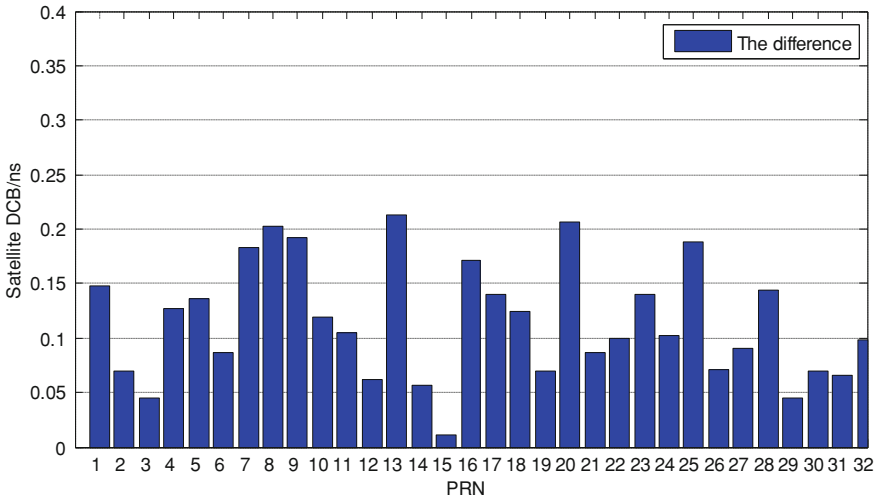


Fig. 33.2 The comparison between the new method and CODE

More than 100 globally distributed IGS stations with dual-frequency GPS observation data is used by CODE, and the spherical harmonic function is used to fit the global ionosphere in the estimation of each station and satellite DCB. Due to its strong spatial structure distribution and high reliability, the DCB product released by CODE can be thought of as true value. So the comparison between the result estimated by our method and the one by CODE is done as follows:

In the figure above: all the satellites' DCB is compared with the products published by CODE, the disparity between them is mostly not greater than 0.2 ns.

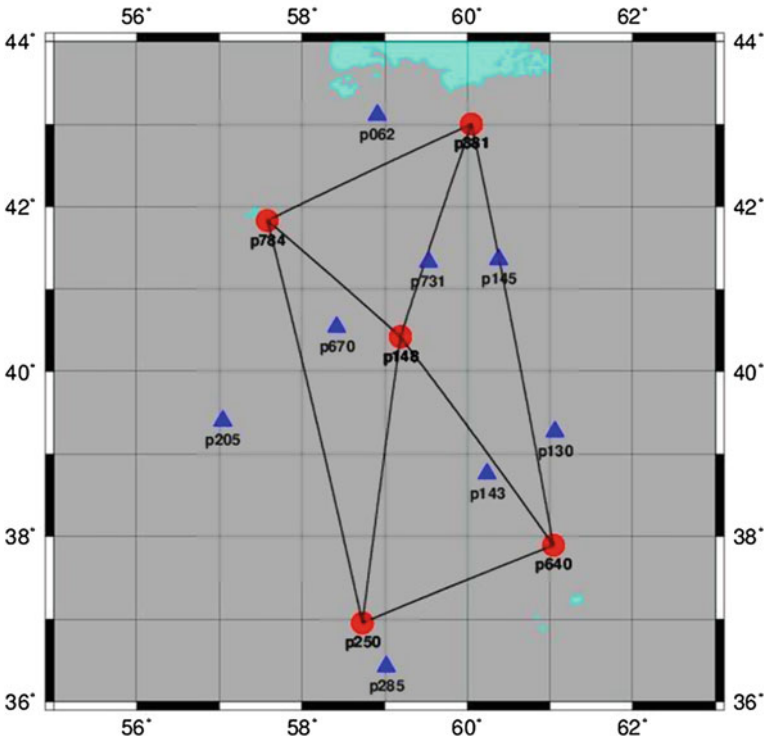


Fig. 33.3 American GPS reference network

Due to the part of the error term correction in the PPP processing is not high enough on the test, the higher precision result of DCB can be gotten if the further elaboration is made (Fig. 33.2).

Regional enhanced PPP

A 24 h data set observed in American CORS in 8 August 2013 was used to test the performance of the proposed strategy for regional enhanced PPP. The test network consisted of five and eight reference stations are used as the PPP users' receivers in the experiment, as shown in Fig. 33.3. The distance between the reference stations is from 207 to 681 km. The sampling interval of the data is also 30 s (Fig. 33.4).

The effect of interpolation on atmospheric delay is a major inhibitor to the convergence time of the PPP users. In Fig. 33.5, we depicted two lines: one is the difference between the interpolation of the slant ionospheric delay from the reference stations and the true value estimated by un-combined PPP technology, and the other is the satellite elevation.

The discrepancy can be seen obviously and indicates that the regional predicted model has an acceptable accuracy within almost ten centimeters of error and the difference is related to the satellite elevation. The smaller the satellite elevation is, the bigger the difference could be obtained. On the other hand, the jitter could be

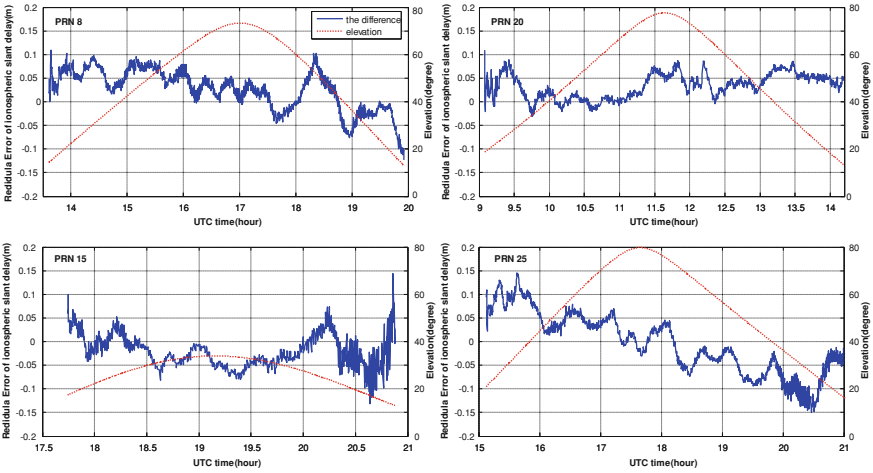


Fig. 33.4 The residual errors of several satellites differences in the method

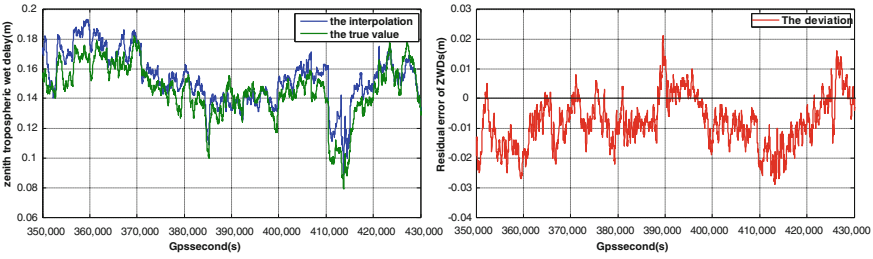


Fig. 33.5 The comparison of the effect of ZWD

observed at 20 to 21 o'clock UTC (12 to 13 o'clock local time) in figure due to strenuous exercise of the ionosphere. Therefore, The determination of the observable noise about the interpolation of slant ionospheric delay is set as 0.1 m.

In Fig. 33.5, the comparison has been done among the true zenith troposphere wet delay and the interpolation. And the different value figure is also shown in the figure. In the sparse multiple reference station networks, the residual error of ZWD range between 3 cm. The effort of the interpolation has a good performance on improving the accuracy of troposphere delay instead of the empirical model.

With the help of prior ionospheric delay and ZWD at the PPP users' location, the algorithm mentioned above could be investigated and tested. Figure 33.6 shows the single-frequency positioning result by our method:

In the Fig. 33.6, the positioning accuracy can reach 10 cm in 30 min or less. Due to the accuracy of the ionospheric delay interpolated by the sparse regional

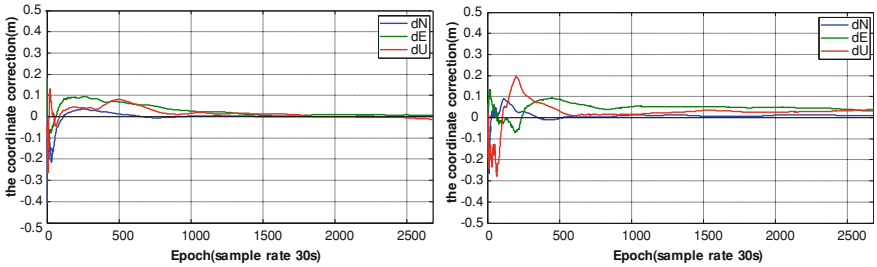


Fig. 33.6 The effect of the PPP algorithm on single-frequency receivers

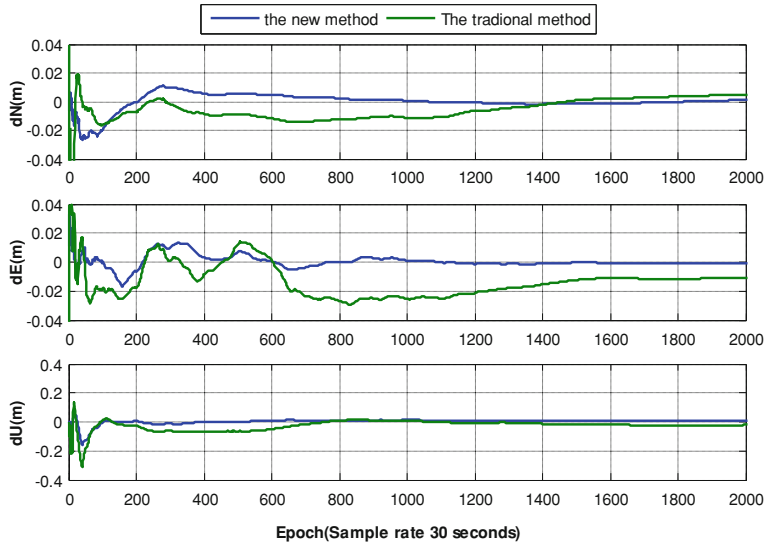


Fig. 33.7 The effect of PPP algorithm on dual-frequency receivers

CORS, the precision in the east eastward is not high enough and the further improvement of the precision is restricted. But it also improves the accuracy within 3–5 cm after filtering convergence by at least an order of magnitude compared with the conventional single frequency method. An order of magnitude was improved compared with the positioning result in 10–20 cm using UofC model or the GIM model.

For dual frequency receiver, the addition of the ionosphere constraint can weaken ill-posed characteristic of the normal equation, achieving the goal of accelerating convergence time. Two lines are depicted in the Fig. 33.7 in the three directions, one is the new method utilizing the interpolation of atmospheric delay and the other is the traditional method using the ionosphere-free model. The effect is shown in the Fig. 33.7.

As shown in the Fig. 33.7, the proposed strategy for dual-frequency PPP seems to work well on the convergence time. It only takes about 10 min to obtain positions with accuracy better than 10 cm in the three directions. And the addition of the quasi-observable equation has good effects on the vertical direction and makes jitter more smoothly in the plane even though the accuracy of the ionospheric delay interpolated is only 10 cm. However the performance of the proposed shows that the stability and high-precision of the positioning results could be achieved compared to the traditional one. And the effects on other stations have a similar behavior, which are not listed here.

33.5 Concluding Remarks

This study focuses on improving the accuracy and the convergence time of PPP users based on the sparse regional CORS. In this contribution, we propose a new approach for the extraction and application of un-differenced atmospheric delays and the main characteristics of the paper is summarized as below:

(1) The un-combined PPP technology is suitable for generating the un-differenced atmospheric delays compared with the ionosphere-free observation model. The Kalman filter is adopted and the parameter setting needs more consideration and research so that more stable and reliable ionospheric delay could be obtained.
(2) With our method, the estimated satellite DCBs are reliable. Compared with the products published by CODE, the disparity between them is mostly not greater than 0.2 ns. The true ionospheric delays are obtained and the comparison between our method and the others would be done to ensure the precision of them.
(3) The implementation methods in the network RTK technology are transplanted in our method. The precision of interpolation algorithm restricted the convergence time and position precision of rover PPP. The results showed that the positioning accuracy can reach 10 cm in the plane within 30 min for the single frequency PPP user and takes about 10 min to obtain positions with accuracy better than 10 cm for the dual-frequency PPP user. An order of magnitude was improved compared with the single-frequency positioning result in 10–20 cm using UofC model or the GIM model. It is noted that the distance between the sites is far away, and the better results could be achieved by more intensive distribution of the reference stations.

Acknowledgments This work is supported by the Key Projects in the National Science & Technology Pillar Program during the Twelfth Five-year Plan Period.(2012BAJ23B01). The authors are very grateful to the anonymous reviewers for their constructive comments and suggestions.

References

1. Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102:5005–5017
2. Liu JN, Ye S (2002) R.GPS precise point positioning using undifferenced phase observation. Geomatics Inf Sci Wuhan Univ 27(3):234–240
3. Zhang XH, Zuo X, Li P (2013) Mathematic model and performance comparison between ionosphere-free combined and uncombined precise point positioning. Geomatics Inf Sci Wuhan Univ 38(5):561–565
4. Zhang BC, Ou JK, Yuan YB et al (2011) Calibration of slant total electron content and satellite-receiver's differential code biases with uncombined precise point positioning technique. Acta Geodaetica Cartogr Sin 40(4):447–453
5. Tu R, Zhang Q, Huang GW et al (2010) On ionosphere-delay processing methods for single-frequency precise-point positioning. J Geodesy and Geodyn 30(2):99–103
6. Zhang H, Gao Z, Ge M et al (2013) On the convergence of ionospheric constrained precise point positioning (IC-PPP) based on undifferential uncombined raw GNSS observations. Sensors 13(11):15708–15725
7. Zhang BC, Ou JK, Yuan YB et al (2012) Extraction of line-of-sight ionospheric observables from GPS data using precise point positioning. Sci China Earth Sci 55(11):1919–1928
8. Klobuchar JA (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst 23(3):325–331
9. Schaer S (1999) Mapping and predicting the Earth's ionosphere using the global positioning system. University of Berne, Berne
10. Zhang B (2009) Determination of un-differenced atmospheric delays for network-based RTK. In: Proceedings of the 22nd international technical meeting of the satellite division of the institute of navigation (ION GNSS 2009), vol 2001. pp 2727–2738
11. Zhang XH, Li XX, Guo F et al (2008) Realization and precision analysis of single frequency precise point positioning software. Geomatics Inf Sci Wuhan Univ 33:783–787
12. Li X (2012) Improving real-time PPP ambiguity resolution with ionospheric characteristic consideration. In: Proceedings of the 25th international technical meeting of the satellite division of the institute of navigation. Nashville, pp 17–21
13. Wei Li, Pengfei C, Jinzhong B (2011) Regional ionosphere delays calibration and accuracy assessment based on uncombined precise point positioning. Geomatics Inf Sci Wuhan Univ 36(10):1200–1203
14. Jiang WP, Zou X, Tang WM (2012) A new kind of real-time PPP method for GPS single-frequency receiver using CORS network. Chin J Geophys 55(5):1549–1556
15. Wang D, Gao C, Pan S (2013) Single-epoch integer ambiguity resolution for long-baseline RTK with ionosphere and troposphere. In: China satellite navigation conference (CSNC), Wuhan, pp 125–137

Chapter 34
Latest GNSS Results of the CMONOC Network and Its Application in Earthquake Monitoring

Junping Chen, Yize Zhang, Yibing Xie, Weijie Tan, Sainan Yang and Bin Wu

Abstract Crustal Movement Observation Network Of China (CMONOC) is a fundamental facility which has wide-range applications in diverse areas with high precision and high spatial resolution. CMONOC network consists of GNSS, VLBI, SLR and gravity observing stations. The observation data contributes to the GNSS meteorology, plate tectonics and earthquake monitoring etc. Shanghai Astronomical Observatory (SHAO) is one of the data centers of the CMONOC network. We developed the CMONOC GNSS data processing platform and provide precise products supporting CMONOC applications. In this paper, we present the latest results from the routine data analysis. Besides the conventional results of GNSS orbits, clocks, ERPs and station coordinates etc., we present for the first time the following results: (1) regional velocity field for the 247 continuous stations, which is derived based on integrated data analysis of the CMONOC and IGS network; (2) earthquake monitoring by ionosphere disturbance analysis, which is realized by analyzing the ionosphere TEC changes time series using the quartile method; (3) earthquake monitoring by instantaneous station motion analysis, which is determined by differential 1 Hz phase observation in real-time.

Keywords CMONOC · GGDAA · Velocity field · Earthquake · Ionosphere

J. Chen (✉) · Y. Zhang · Y. Xie · W. Tan · S. Yang · B. Wu
Shanghai Astronomical Observatory, Chinese Academy of Science, Shanghai, People's Republic of China
e-mail: junping@shao.ac.cn

Y. Zhang · Y. Xie
College of Surveying and Geo-Informatics, Tongji University, Shanghai, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_34,

34.1 Introduction

Crustal Movement Observation Network Of China (CMONOC) is in operation since 1999. It is a fundamental facility which has wide-range applications in diverse areas with high precision and high spatial resolution. It consists of 260 GNSS stations, several Very Long Baseline Radio Interferometry (VLBI) and Satellite Ranging (SLR) and precise leveling and gravity stations. Observation data of CMONOC make an important contribution in the field of research concerning the GNSS meteorology [1], plate tectonics and earthquake monitoring [2].

As one of the data center of the CMONOC, Shanghai Astronomical Observatory (SHAO) operates the routine GNSS data analysis since June 2011 [3]. Data analysis implements similar strategies as it used in our IGS routine data analysis [4]. In the CMONOC routine analysis, we combine the observations from the IGS and CMONOC network, and all solutions are aligned to the ITRF2008 frame. Figure 34.1 shows the number of stations processed in our routine job, where around 300 stations are processed per day with around 80 are IGS sites. Based on daily solutions, precise orbits, clocks, troposphere and coordinates etc. are derived [5]. Besides these precise products, we present for the first time the following new results: (1) regional velocity field for the 247 continuous stations, which is derived based on integrated data analysis of the CMONOC and IGS network; (2) earthquake monitoring by ionosphere disturbance analysis, which is realized by analyzing the ionosphere TEC changes time series using the quartile method; (3) earthquake monitoring by instantaneous station motion analysis, which is determined by differential 1 Hz phase observation in real-time.

34.2 Velocity Field of the CMONOC GNSS Stations

We start the daily solution since January 2011, which includes 252 stations till December 2013. To obtain velocity field of the continental China, we combine all the daily solutions using the software Quasi-observation Combination Analysis (QOCA, http://gipsy.jpl.nasa.gov/qoca) [6]. The strategy of data combination follows the following 3 steps: (1) generation of the daily solutions, which are under the ITRF2008 frame; (2) data combination using the package of analyze_tseri of QOCA to get velocity of continental China for the 247 stations with data interval longer than 1 year; (3) velocity transformation from global ITRF2008 frame to the stable Eurasia plate (EURA).

In the data combination using QOCA, we carry out data editing for the time series jumps due to known events, including coseismic offsets due to the Japan earthquake in 2011 and Lushan earthquake in 2013; nonlinear variations of some stations as YNXP, YNJD and YNSM in Yunnan province. Annual/semiannual periodical motion is considered for all sites. In additional we perform the principal component analysis (PCA) to remove the common mode error from each raw time

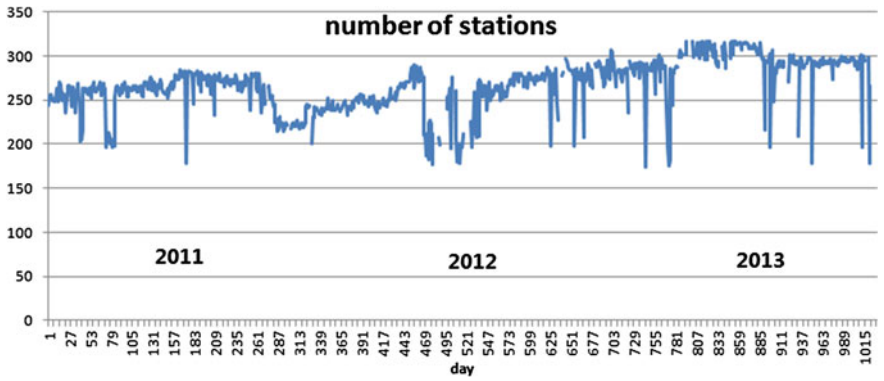


Fig. 34.1 Number of stations processed at the SHAO, where around 80 stations are from the IGS global tracking network. Data analysis is in daily batch with latency of 1 day, and stations with latency of more than 1 day in data transfer are not included in the daily solution

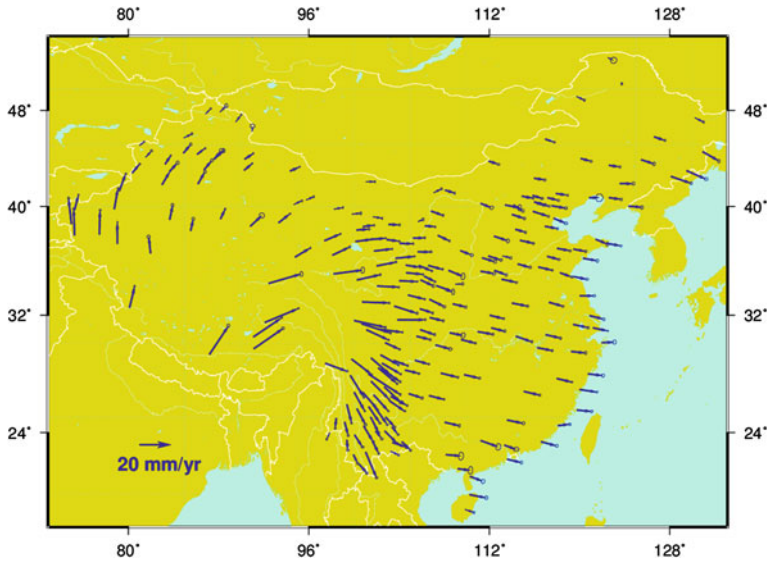


Fig. 34.2 CMONOC GPS result of the horizontal velocity field of continental China respect to EURA. The arrows represent the velocities. The ellipses represent the 95 % confidence level

series [7]. For the velocity transformation, we follow the NNR-NUVEL model and select 11 reference stations [8] to calculate the Euler vector of the EURA plate. Weighted least squares algorithm is applied in the estimation of the Euler vectors through an iterative process [9]. The resulting Euler vectors are (54.30°N, 100.30°W), and 0.253°/M.a. The horizontal velocity field of continental China respect to EURA is shown in Fig. 34.2, where 247 stations are included.

34.3 Ionospheric Disturbance Monitoring Using the CMONOC Network

Ionosphere abnormal disturbances were observed for the first time in the 1965 Alaska earthquake [10], and the analysis of the seismic ionospheric disturbances is developed since then. The comparison between the instantaneous TEC and statistical quantity of Total Electron Content (TEC) reflects such disturbances [11]. Statistical quantity of TEC could be selected according to the quartile definition, where the first quartile (designated Q1) splits off the lowest 25 % of data from the highest 75 %; and the second quartile (designated Q2) cuts data set in half; the third quartile (designated Q3) splits off the highest 25 % of data from the lowest 75 %. The difference between the Q3 and Q1 quartiles is called the interquartile range (IQR). We define in this paper the upper quartile L1 as the sum of the Q2 and IQR and the lower quartile L2 as the difference between Q2 and IQR. Ionospheric disturbances is detected when the instantaneous TEC is not within the range defined by the upper and lower quartiles.

According to [11], the TEC quartiles at current epoch is calculated using a sample of 15 TECs, where the TECs at same epoch of successive 15 days are used. We set up the ionospheric disturbances monitoring system based on this strategy and the system run in the following way: (1) calculation of the Vertical TEC (VTEC) using precise models, e.g. the Global Ionospheric Mapping (GIM) of the IGS; (2) calculation the real-time VTEC and VTEC quartiles of current epoch; (3) comparison between real-time VTEC and VTEC quartiles to check for ionospheric disturbances.

Taking the Lushan (Sichuan Province) M_S7.0 earthquake on April 20, 2013 as an example. We use the IGS GIM data to calculate the VTEC over the epicenter (30.3° N, 103.0° E), as shown in Fig. 34.3. From Fig. 34.3, VTEC disturbance is obvious on April 16–21 with observed VTEC below the VTEC lower bounds (L2), especially on the day before earthquake. The VTEC gradually return to normal and stable after April 22. According to the data provided by the national center for space science, Chinese academy of sciences, the geomagnetic disturbance coefficient Kp and Dst index are at lower levels, which confirms ionospheric disturbance during this earthquake is not caused by geomagnetic storm or similar physical environment.

To further illustrate the ionospheric VTEC abnormalities during this earthquake, we check the results of two CMONOC stations near the epicenter, namely SCDF (30.978°N,101.122°E) and YNDC (26.108°N, 103.181°E). The real-time VTEC and VTEC quartiles time series are shown in Fig. 34.4. We observe that VTECs are below the lower bounds (L2) during the same period as for the epicenter.

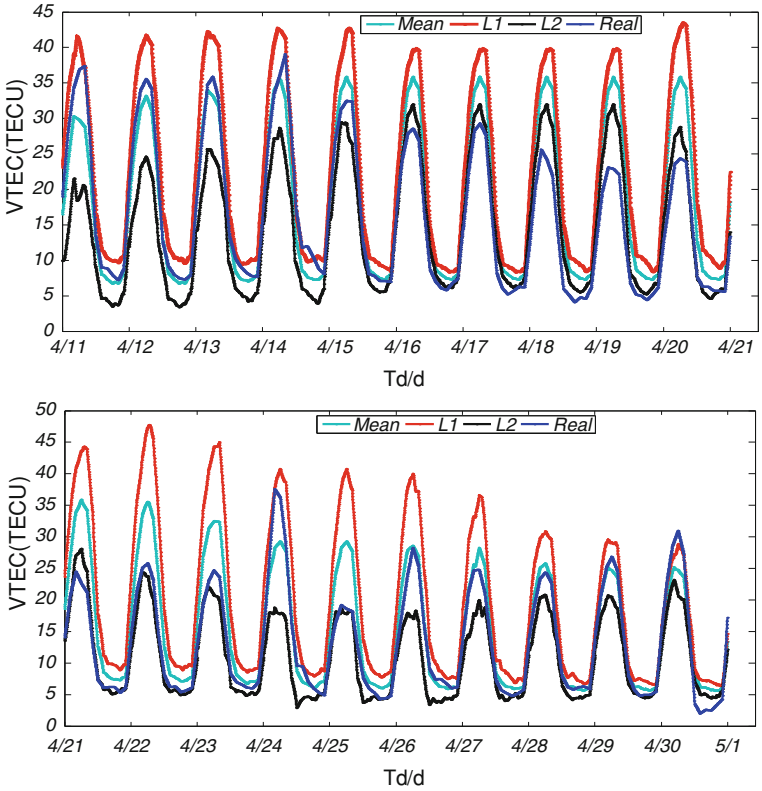


Fig. 34.3 Time series of VTEC over the epicenter of Lushan earthquake, where "mean" refers to the VTEC median over the same epoch of successive 15 days and "real" refers to real-time VTEC observed at each epoch

34.4 Coseismic Monitoring Using the CMONOC Network

With the fast development of high frequency GNSS technology, GNSS data plays an important role in instantaneous station motion monitoring and earthquake source rupture processing. Traditional GNSS high-frequency data processing is based on the Precise Point Positioning (PPP) or double difference baseline solution to obtain the instantaneous site displacements. However, the station position estimates are affected by unknown error resources and are limited by the capability of ambiguity fixing. Thus these strategies are still challenging for earthquake monitoring [12]. To overcome the limitations of the position solutions, the strategy of so-called instantaneous site velocity solution [13] is developed by using the GNSS Doppler observations.

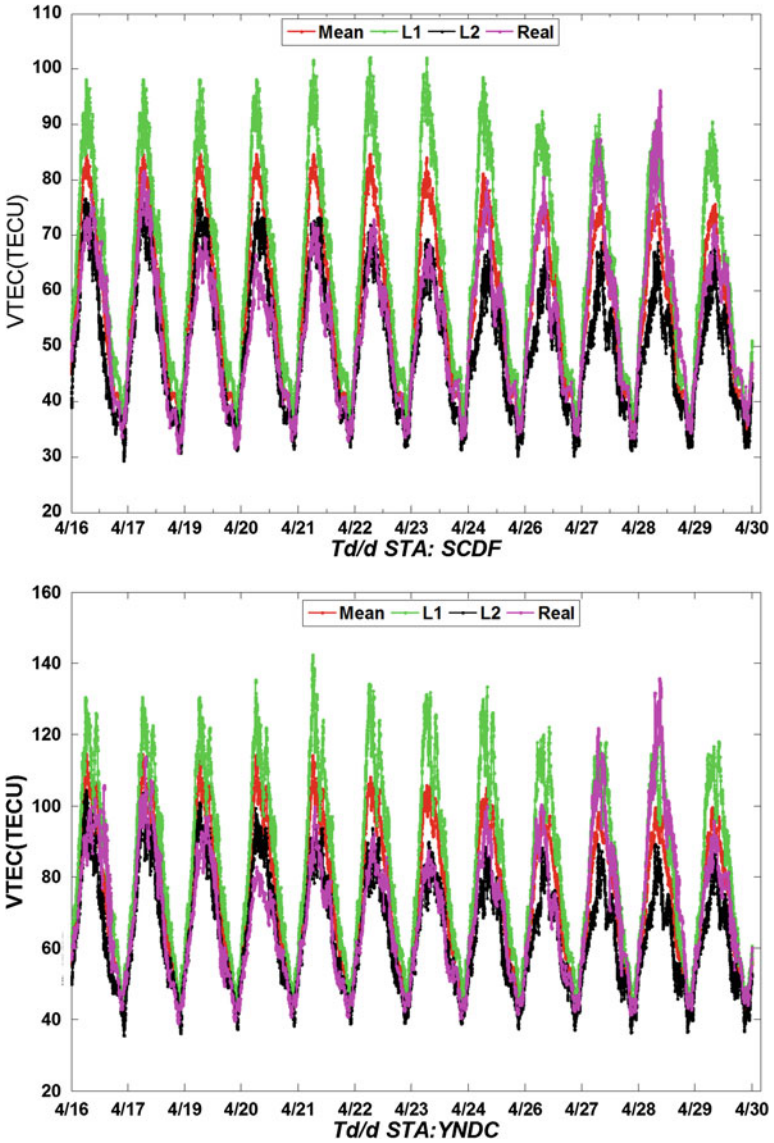


Fig. 34.4 Time series of VTEC for 4 stations during the Lushan earthquake, where "mean" refers to the VTEC median over the same epoch of successive 15 days and "real" refers to real-time VTEC observed at each epoch

Applying this strategy, we set up the coseismic monitoring system for the continuous station in the CMONOC network and the system runs in the following way: (1) retrieval of real-time high rate streams or files from internet protocols; (2)

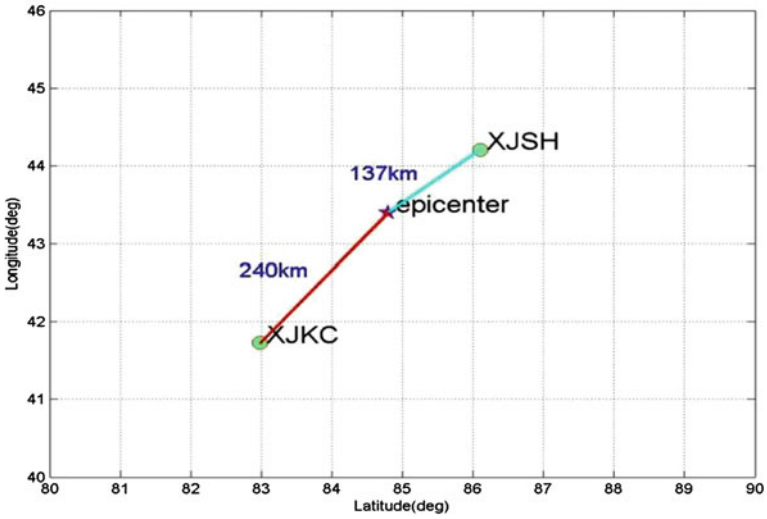


Fig. 34.5 Plot of the epicenter and selected stations

calculation of the Doppler observations from the phase observations, and estimation of station instantaneous velocity based on Doppler observation; (3) check for coseismic offsets based on the instantaneous solution.

Taking the Xinjiang Ms6.6 earthquake on June 30, 2012 as an example. The earthquake happened at 05:07:31 (UTC 21:07:31 on June 29), and the stations XJSH (137 km away from epicenter, sampling 1 s) and XJKC (240 km away from epicenter, sampling 1 s) tracks the coseismic displacement. Figure 34.5 illustrates the location of the epicenter and two stations. Using the above mentioned strategy, we analysis the instantaneous velocity and coseismic offsets as shown in Figs. 34.6 and 34.7. The east components of the station velocity exhibit pure white noise feature before the earthquake. The estimated instantaneous velocity and coseismic offsets show obvious changes when the seismic waves reach these stations, where maximum amplitude of instantaneous velocity reaches 16 mm/s for XJSH and 13 mm/s for XJKC, and the maximum coseismic offset reaches 11 mm for XJSH and 10 mm XJKC, respectively. The shaking effects continue for around 20 s, the stations become static and the estimated instantaneous velocity restore to white noise after that. We observe that the epoch that seismic wave arrives at the station XJSH is around 31 s ahead of the time it reaches XJKC. Based on the time difference, we calculate the seismic waves speed as 3.32 km/s.

The high rate data of the CMONOC network is current archived in file with latency of one day, and the coseismic monitoring is currently performed in post-processing mode. With the stations being upgraded to real-time capability, the system will support real-time coseismic monitoring.

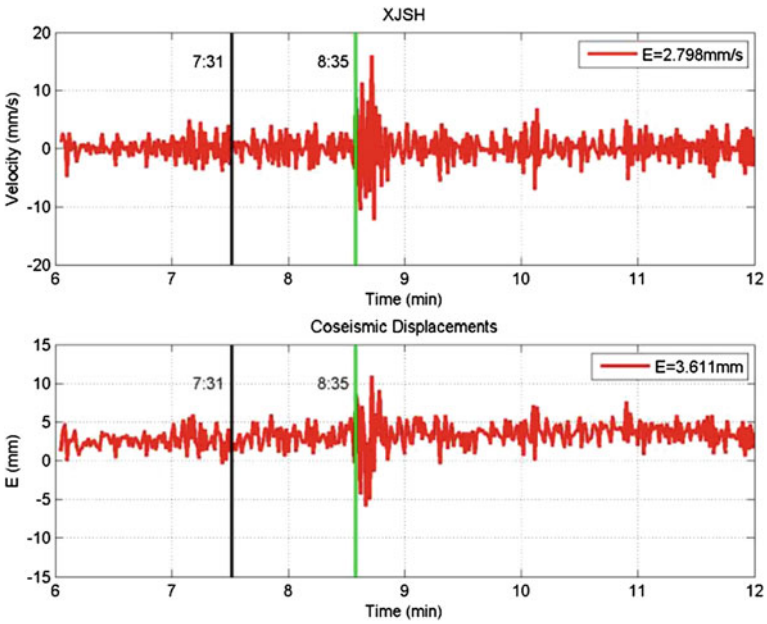


Fig. 34.6 Velocity and coseismic displacement time series of station XJSH, 150 km away from epicenter

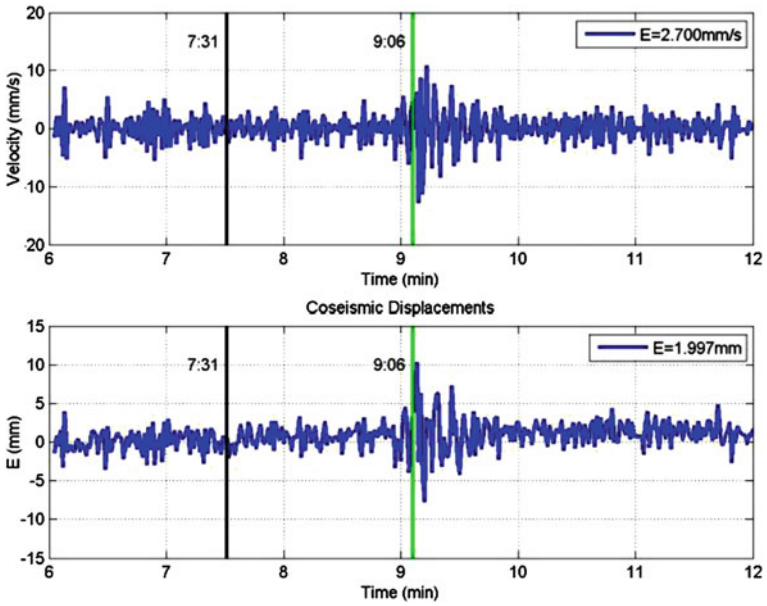


Fig. 34.7 Velocity and coseismic displacement time series of station XJKC, 250 km away from epicenter

34.5 Conclusion

We present the latest results from the CMONOC data analysis center at the SHAO. On the basis of the routine GNSS data analysis of the CMONOC and IGS sites, we present three new achievements: (1) the velocity field of continuous stations, which is released for the first time based on the combination of 3 years' daily solutions; (2) quartile based VTEC monitoring system, which is applied to the ionospheric disturbance monitoring during the Lushan earthquake; and (3) the instantaneous velocity strategy based coseismic monitoring system, with which the instantaneous velocity changes and coseismic offsets are monitored during the 2012 Xinjiang earthquake. The application of CMONOC data in earthquake monitoring could be realized in real-time with the network upgraded to real-time.

Acknowledgment This paper is supported by the 100 Talents Programme of The Chinese Academy of Sciences, the National High Technology Research and Development Program of China (Grant No. 2013AA122402), the National Natural Science Foundation of China (NSFC) (Grant No. 11273046 and 40974018), and the Shanghai Committee of Science and Technology (Grant No. 12DZ2273300,13PJ1409900).

References

1. Wang Y, Liu L, Liang H, Ding K et al (2006) Research on precipitable water vapor in plateau and plain areas with GPS technique. J Geodesy Geodyn 26(1):88–92 (in Chinese)
2. Liu J, Shi C, Xu C, Jiang W (2001) Present day crustal movement speed field of China continent block using local repeated GPS network. Geomatics Inf Sci 26(3):189–195 (Journal of Wuhan University, in Chinese)
3. Junping C, Bin W, Xiaogong H, Haojun L (2012) Shanghai astronomical observatory CMONOC data analysis center. In: The 3rd China satellite navigation academic conference electronic proceedings-S08 satellite navigation model and method (in Chinese)
4. Junping C, Bin W, Xiaogong H, Haojun L (2012) SHA: the GNSS analysis center at SHAO. In: Proceedings of the 3rd China Satellite Navigation Conference, CSNC 2012, revised selected papers. Lecture Notes in Electrical Engineering, vol 160 LNEE, pp 213–221
5. Shanghai Observatory GNSS Analysis Center. http://www.shao.ac.cn/shao_gnss_ac
6. Dong D, Herring TA, King RW (1998) Estimating regional deformation from a combination of space and terrestrial geodetic data. J Geodyn 72:200–214
7. Danan D, Peng F et al (2006) Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis. J Geophys Res, 111:B03405
8. Min W, Zhengkang S, Zhijun N et al (2003) The crustal movement of and activity block model. Sci China 33:21–32 (in Chinese)
9. Ye S, Huang C (2000) Astrogeodynamics. Shandong Science and Technology Press, Shandong (in Chinese)
10. Barnes RA, Leonard RS (1965) Observations of ionospheric disturbance following the Alaska earthquake. J Geophys Res 70(9):1250–1253

11. Zhao Y, Zhang X, Liu J (2010) The ionospheric electron content disturbance analysis before and after the WenChun earthquake. Prog Geophys 25(4):447–453 (in Chinese)
12. Zhang X, Guo B (2013) Real-time tracking the instantaneous movement of crust during earthquake with a stand-alone GPS receiver. Chin J Geophys 56(2):1928–1936 (in Chinese)
13. Colosimo G, Crespi M, Mazzoni A (2011) Real-time GPS seismology with a stand-alone receiver: a preliminary feasibility demonstration. J Geophys Res 116:B11302. doi:10.1029/2010JB0079

Chapter 35
A Study of Analytical Solar Radiation Pressure Modeling for BeiDou Navigation Satellites Based on RayTracing Method

Weidong Feng, Xiangyu Guo, Hongxing Qiu, Junli Zhang and Kaifeng Dong

Abstract Based on the characteristics of the Beidou navigation satellites, this paper introduce a RayTracking analytical solar radiation pressure modeling method which scatters the incident sunlights into a beam array, tracks the solar pressure of every discrete beam generated by reflection of satellite surface, superimposes the radiation pressures of all the discrete beams and gives an accurate solar radiation pressure model with analytical expressions using Fourier analysis. Considering the mutual occlusions of different parts of the satellite surface and the effect of sunlight beams reflected repeatedly by the satellite surface, this method can be used in solar radiation pressure modeling of satellites with complex surface structures. This paper also analyzes the effects of characteristics of satellite attitude and optical properties of the surface, then establishes an analytical solar radiation pressure model for BeiDou IGSO navigation satellites. With a unified form of expressions in the states of dynamic yaw control and zero yaw control of satellite attitude, this model can be used in orbital determination analysis in the period of crossing dynamic yaw and zero yaw control.

Keywords Navigation satellite · Solar radiation pressure · Light tracing

35.1 Introduction

To the high-orbit and medium-orbit satellites such as navigation satellites, solar radiation pressure, which is the largest perturbation except the gravity of Earth, sun and moon, is the key factor affecting the dynamical model accuracy of navigation

W. Feng (✉) · X. Guo (✉) · H. Qiu · J. Zhang · K. Dong
BSIR, No. 5117 Mailbox, Haidian, Beijing, China
e-mail: weberfeng2000@163.com

X. Guo
e-mail: guoxiangyu2000@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_35,

satellites. After GPS was built, researchers have been dedicated to investigate the solar radiation pressure model acting on GPS satellites. Some relevant models has been established, such as ROCK4, ROCK42, T20, ECOM, GSPM04 and GSPM10 [1, 3, 4–5]. The solar radiation pressure which is a complex perturbation acting on the spacecraft surface, was affected by many different factors, such as satellite attitude, surface structure and optical property. Compared with the GPS satellites, BeiDou satellites have great discrepancies. The related research shows that using GPS solar radiation pressure model directly in precise orbit determination of BeiDou satellite, there will be a great orbit error. Therefore, to build the special solar radiation pressure model aiming at the characteristics of Beidou navigation satellite is crucial.

Both analytical modeling and empirical modeling are usually used in solar radiation pressure modeling. Analytical solar radiation pressure modeling is building the corresponding mathematical model based on solar radiation pressure mechanism, satellite attitude, satellite structure and optical properties. On the other hand, in the process of empirical solar radiation pressure modeling, mathematical model is built through parameter estimation and identification based on measured data of a long period. As introduced above, empirical solar radiation pressure modeling which requires a large amount of measured data, is developed based on the priori analytical solar radiation pressure model. Therefore, as a prerequisite, analytical solar radiation pressure model needs to be developed firstly.

This paper introduce a RayTracking analytical solar radiation pressure modeling method which scatters the incident sunlight into beam array, tracks the radiation pressure of every discrete beam generated by satellite surface reflection, superimposes the pressures of all the discrete beams and gives an accurate solar radiation pressure model with analytical expression using Fourier analysis. Then, we establish an analytical solar radiation model for Beidou IGSO navigation satellite, analyze and compare the solar radiation pressure characteristics in the state of dynamic yaw control and zero yaw control of satellite attitude.

35.2 RayTracing Solar Radiation Pressure Modeling

As Fig. 35.1 shows, RayTracing solar radiation pressure modeling scatters the solar flux into the plane pixel array vertical to the direction of sunlight incidence in which one pixel represents one incident beam [6], traces and calculates the track of every incident beam through multiple specular reflections by satellite surface, until this beam disappears in space or its intensity decreases to the set threshold value. Then, the accumulation of the solar radiation pressure generated by every reflection is the pressure of this beam acting on satellite surface. Furthermore, the total solar radiation pressure on the satellite surface is generated by the pressure accumulation of all the beams.

As Fig. 35.2 shows, when the satellite surface is irradiated by sunlights, reflected lights include specular and diffuse reflection [3].

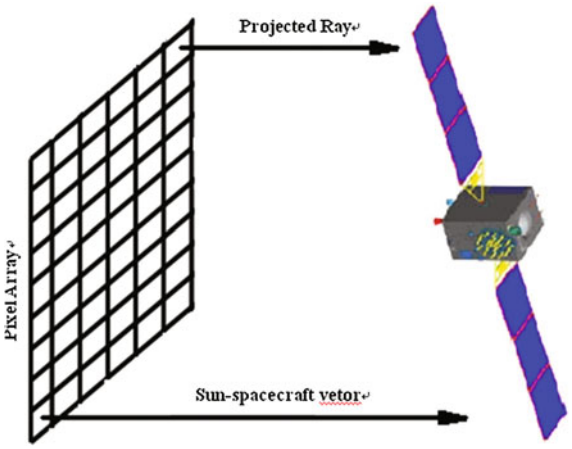


Fig. 35.1 Schematic: RayTracing solar radiation pressure modeling

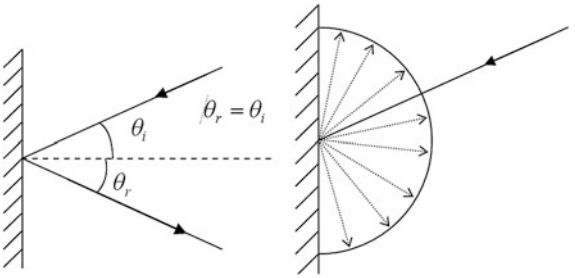


Fig. 35.2 Diagram: specular reflection and diffuse reflection

Normal and tangential components of solar radiation pressure of specular reflection on the satellite surface are as follows.

$$a_{f,D} = -\left(\frac{A \cdot E}{c}\right)(1 + \mu v)\cos^2\theta \tag{35.1}$$

$$a_{f,s} = -\left(\frac{A \cdot E}{c}\right)(1 - \mu v)\sin\theta\cos\theta \tag{35.2}$$

In these equations, A represents infinitesimal area of satellite surface, E represents solar radiation intensity, c represents light velocity, μ represents infinitesimal mirror coefficient of satellite surface, v represents infinitesimal reflection coefficient of satellite surface, θ represents beam angle of incidence.

Diffuse reflection which can be approximated as Lambert reflection, only produces radiation pressure on the normal direction of satellite surface, its expression is as follow.

$$a_{f,D} = -\frac{2}{3}\left(\frac{A \cdot E}{c}\right)(1 - \mu)v\cos\theta \tag{35.3}$$

The solar radiation pressure modeling based on RayTracing which is convenient to consider the relationship of block among various components and the second reflected light beam, can be applied to arbitrary satellite with complex surface structure.

35.3 BeiDou IGSO Satellite Solar Radiation Pressure Model

Based on the RayTracing solar radiation pressure model described in Sect. 35.2, an adaptive solar radiation pressure model aimed at the characteristics of BeiDou IGSO satellites is established in this chapter.

35.3.1 Satellite Yaw Attitude Modeling

As an important factor affecting solar radiation pressure, satellite attitude model must be established before solar radiation pressure modeling. To ensure adequate power supply for satellite and avoid the equipment installed on the west board of satellite body being irradiated, dynamic yaw attitude control mode is used for GPS navigation satellite. In this mode, satellite yaw angle varies dynamically with solar orientation to make the sun be in the XOZ plane of satellite body and the east board of the satellite be irradiated all the time [2]. On the other hand, this mode will cause 2 times of nearly 180-degree yaw attitude maneuver during per revolution of satellite when the angle between the incident direction of sun and the orbital plane is small. To avoid this kind of attitude maneuver, new attitude control modes are adopted in the second-generation BeiDou navigation satellites. Zero yaw control mode is adopted for GEO satellite (satellite yaw angle remains at zero). At the same time, an attitude control mode which combine dynamic yaw control with zero yaw control is adopted for MEO and IGSO satellite. Its principle was introduced as follow: dynamic yaw control mode is adopted when the angle between the incident direction of sunlight and the orbital plane is greater than 4°; zero yaw control mode is adopted when the angle between the incident direction of sunlight and the orbital plane is less than 4° and the yaw angle is less than 5°. According to the constraints above, the yaw angle of satellite in the state of dynamic yaw attitude control is as follow.

$$\begin{cases} tg\beta = \dfrac{tgE_s}{\cos A_s} \\ \cos A_s \cos E_s \cos \beta \geq 0 \end{cases} \tag{35.4}$$

In this equation, β is yaw angle of satellite, Es is the angle between direction vector of satellite-sun and orbital plane, As is the angle between the orbital plane projection of direction vector of satellite-sun and the vector of satellite-geocenter.

35.3.2 Solar Radiation Pressure Model of Dynamic Yaw

The established solar radiation pressure model is described in the satellite body coordinate system. In the state of dynamic yaw control, the sun is always in the XOZ plane of the satellite body coordinate system; Meanwhile, the east plane of the satellite is towards sun all the time, the south, north and west plane can't be irradiated by sun, the plane oriented to ground and opposite to ground are irradiated by sun alternately. In the satellite body coordinate system, the only one degree of freedom in the solar direction is the angle with + Z-axis β. The solar radiation pressure model of BeiDou IGSO satellite in the state of dynamic yaw control is established with RayTracing method through which the solar radiation pressure values of different angle β is calculated and expanded into a Fourier series of angle β (the small item is eliminated).

$$\begin{cases} F_x = -0.082 - 16.871\sin\beta - 0.071\sin(2\beta - 1.772) \\ \qquad + 0.092\sin 3\beta + 0.113\sin 5\beta \\ F_y = 0.0 \\ F_z = -17.004\cos\beta + 0.134\cos 5\beta \end{cases} \tag{35.5}$$

In these equations, F_x, F_y, F_z represent triaxial components of solar radiation pressure in the satellite body coordinate system individually, unit is 1.0D-5 N.

According to forms of expression of GPS satellite solar radiation pressure models T20 and T30, decomposing the equations above into solar radiation pressure components in directions of sun-satellite, x-axis and z-axis, we can get the equations as follow.

$$\begin{cases} F_{xt} = -0.082 + 0.0665\sin\beta - 0.071\sin(2\beta - 1.772) \\ \qquad + 0.092\sin 3\beta + 0.113\sin 5\beta \\ F_{zt} = -0.0665\cos\beta + 0.134\cos 5\beta \\ F_d = -16.9375 \end{cases} \tag{35.6}$$

In the equations above, Fd represents the solar radiation pressure component along the direction vector of satellite-sun; Fxt and Fzt represent the solar radiation pressure component along the X-axis and Z-axis of satellite body, the unit is 1.0D-5 N.

The solar radiation pressure expression of BeiDou IGSO satellite described in this paper is similar to T20, T30 model of GPS. However, there are great discrepancies of solar radiation pressure values between these models. To facilitate

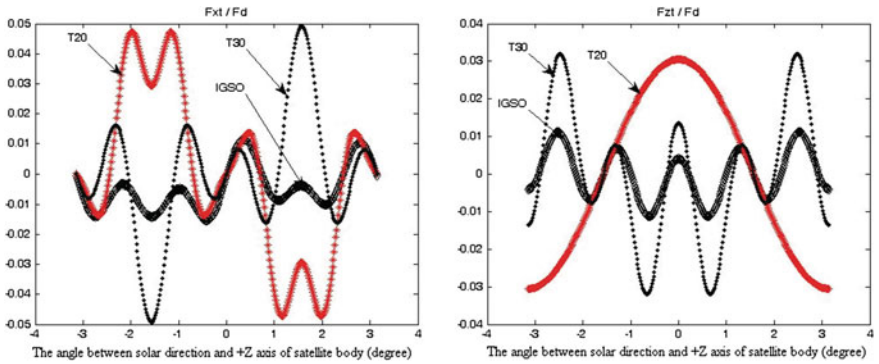


Fig. 35.3 *Curve* parameter comparison of the three different solar radiation pressure models

comparison, we take F_{xt}/F_d, F_{zt}/F_d as comparison parameters. Figure 35.3 shows the F_{xt}/F_d, F_{zt}/F_d curve of these 3 models. From the figure, it is easy to discover obvious discrepancies are existent among these models: the solar radiation pressure proportion of the cycle terms of T20, T30 in the X-axis and Z-axis direction to the solar direction reach 5, 3 %, while this proportion for Beidou IGSO satellite is 1 %.

35.3.3 Solar Radiation Pressure Model of Zero Yaw

In the state of zero yaw, satellite yaw angle is zero, the sun no longer maintains in the XOZ plane of the satellite body coordinate system, so there is an angle exsiting with XOZ plane. Therefore, there are two degrees of freedom in the solar direction variation: the angle α with the XOZ plane of satellite body; the angle β between projection of direction vector of satellite-sun in the XOZ plane of satellite body and +Z axis. Considering the influence of β, α, the analytical solar radiation pressure model is established. The triaxial components of solar radiation pressure in the satellite body are as follows.

$$\begin{cases} F_x = A_x + B_x \cos\alpha \\ A_x = 0.0411 + 9.0331\sin\beta - 0.0411\cos 2\beta \\ \qquad - 0.0851\sin 3\beta - 0.1051\sin 5\beta \\ B_x = -0.1226 - 25.9853\sin\beta + 0.1102\cos 2\beta \\ \qquad + 0.1767\sin 3\beta + 0.2182\sin 5\beta + 0.0477\sin 9\beta \end{cases} \tag{35.7}$$

$$\begin{cases} F_y = A_y + B_y \sin\alpha \\ A_y = 0 \\ B_y = -10.6042 + 0.0840\sin\beta + 0.135\cos 2\beta \\ \qquad + 0.1934\cos 4\beta + 0.0461\cos 8\beta \end{cases} \tag{35.8}$$

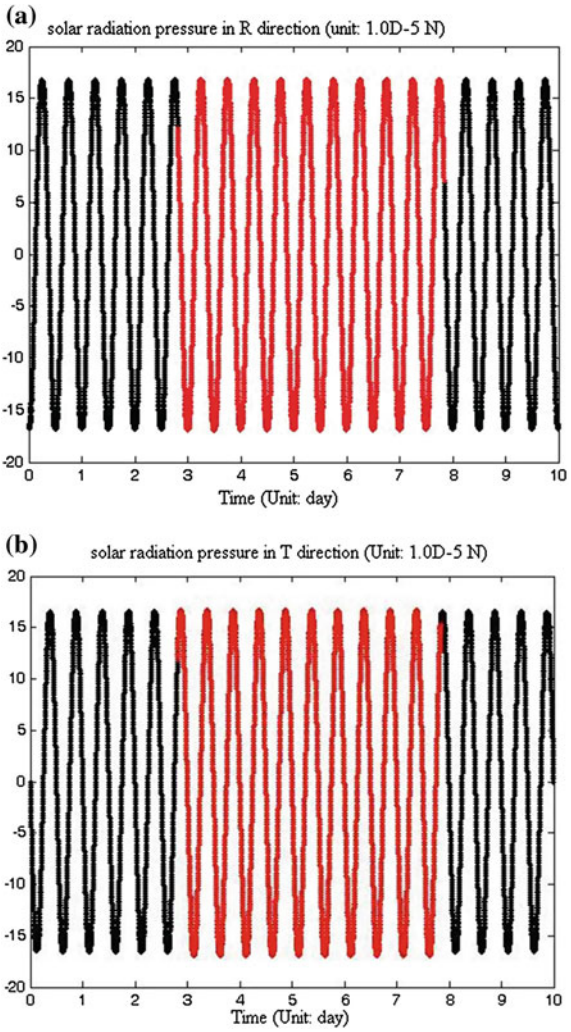


Fig. 35.4 *Curve* solar radiation pressure variation during conversion period between dynamic yaw and zero yaw

$$\begin{cases} F_z = A_z + B_z \cos\alpha \\ A_z = 9.1816\cos\beta - 0.0411\sin 2\beta - 0.1245\cos 5\beta \\ B_z = -0.0328 - 26.2664\cos\beta - 0.0309\cos 2\beta \\ \qquad + 0.0853\sin 2\beta + 0.2583\cos 5\beta + 0.0417\cos 7\beta \\ \qquad + 0.0586\cos 9\beta \end{cases} \tag{35.9}$$

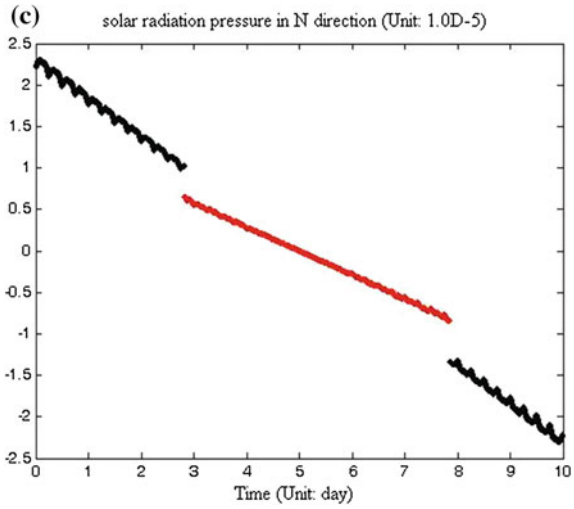


Fig. 35.4 continued

In the state of dynamic yaw, $\alpha = 0$. Therefore, dynamic yaw solar radiation pressure model introduced above is a special case of zero yaw solar radiation pressure model which can also be used in the mode of dynamic yaw.

35.4 The Solar Radiation Pressure Variation in the Conversion Period from Dynamic Yaw to Zero Yaw

The variation law of satellite yaw angle and solar radiation pressure are different between dynamic yaw control mode and zero yaw control mode. Figure 35.4 shows solar radiation pressure variation in the R\T\N direction during the conversion period between dynamic yaw control and zero yaw control of BeiDou IGSO satellite calculated by the model in this paper (the angle between sun and orbital plane varies from $-8°$ to $8°$). It is easy to see the solar radiation pressure in N direction changes obviously from Fig. 35.4.

To make further explanations of solar radiation pressure variation in the state of dynamic yaw control and zero yaw control, the solar radiation pressure of satellite is stimulated and calculated individually when the satellite maintains dynamic yaw control mode and zero yaw control mode. Figure 35.5 shows the mutual difference of solar radiation pressure in these two different modes. The periodical variation is the main manifestation in R and T direction, whereas linear variation is the main manifestation in N direction.

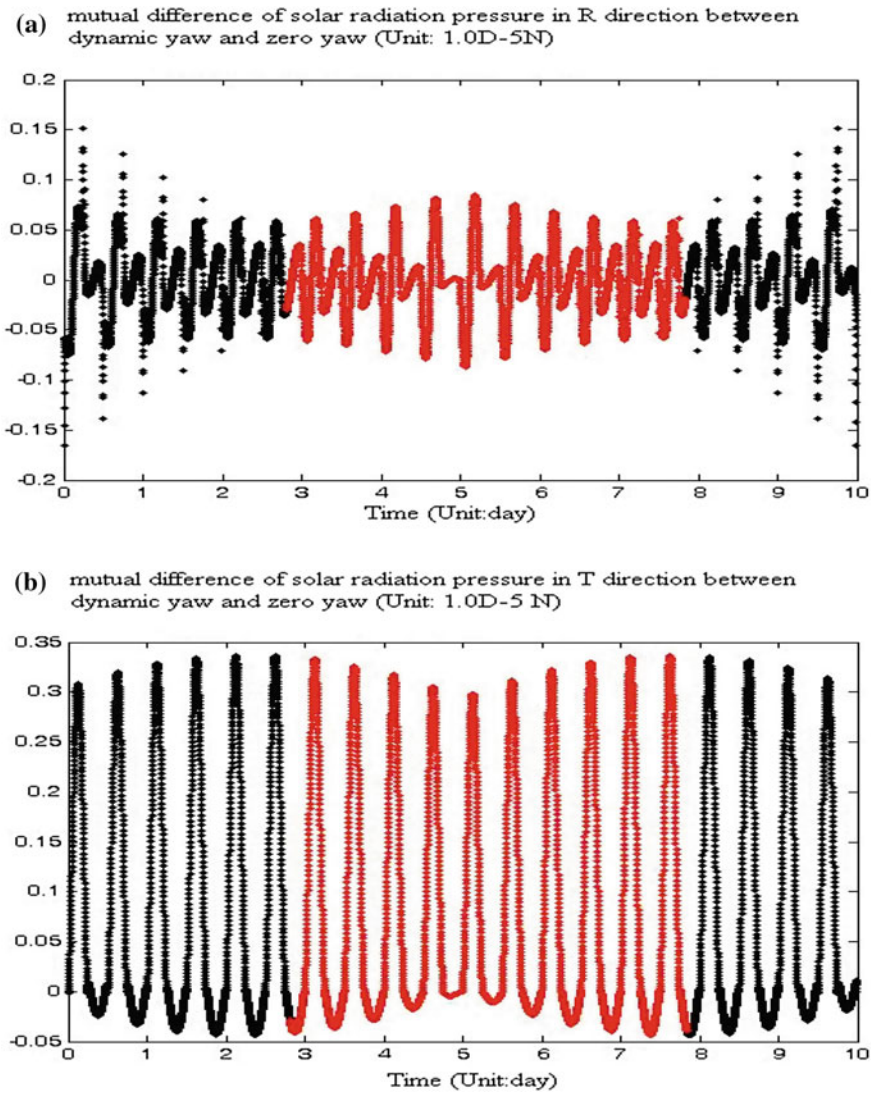


Fig. 35.5 *Curve* mutual differences of solar radiation pressure between the states of dynamic yaw and zero yaw

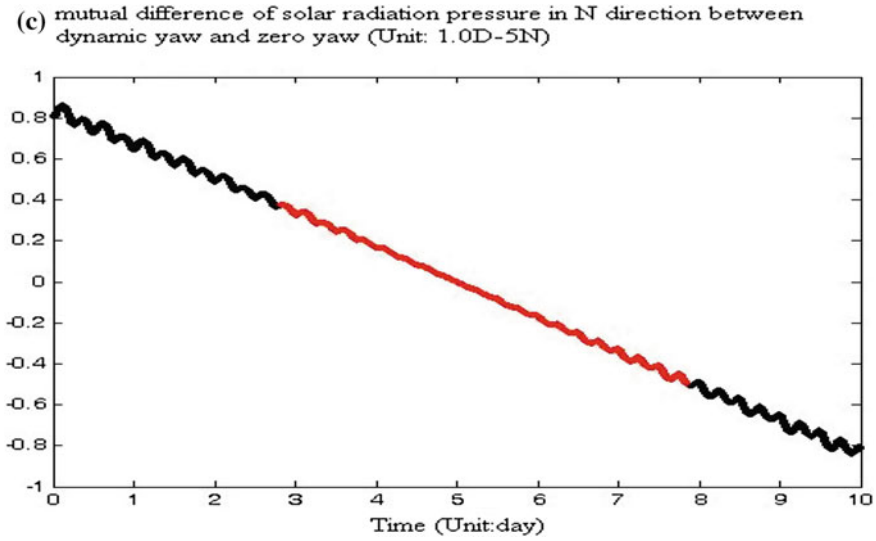


Fig. 35.5 continued

35.5 Conclusion

Because of the great discrepancies of solar radiation pressure between BeiDou satellite and GPS satellite, the special accurate solar radiation pressure model should be established aiming at the characteristics of Beidou satellite. With RayTracing method, this paper establish the analytical solar radiation model for BeiDou IGSO satellites with which solar radiation pressure characteristic variation is analyzed in the mode of dynamic yaw control and zero yaw control. Furthermore, this model provides a priori model for precise orbit determination of BeiDou IGSO satellite. Only considering the solar radiation pressure generated by solar direct irradiation in this paper, it is necessary to take further considerations of the influence of multiple factors for this model, such as atmospheric radiation, thermal radiation of satellite body and radiation of satellite antennas. Meanwhile, long-arc tracking data are necessary to make analysis and improvement of this model.

References

1. Laurent F (2009) A study of solar radiation pressure acting on GPS Satellites. Doctoral dissertation of the University of Texas
2. Kouba J (2009) A simplified yaw-attitude model for eclipsing GPS satellites. GPS Solut 13:1–12
3. McMahon JW (2011) An analytical theory for the perturbative effect of solar radiation pressure on natural and artificial satellites. Doctoral dissertation of the University of Colorado

4. El-Saftawy MI (1998) The effect of direct solar radiation pressure on a spacecraft of complex shape. Astrophys Space Sci 259:151–171
5. Webb CE (2007) Radiation force modeling for ICESat precision orbit determination. Doctoral dissertation of the University of Texas
6. Ziebart M (2001) Analytical solar radiation pressure modelling for GLONASS using a pixel array. J Geodesy 75:587–599

Chapter 36
Integration of GNSS and Seismic Data for Earthquake Early Warning: A Case Study on the 2011 Mw 9.0 Tohoku-Oki Earthquake

Bofeng Guo, Xiaohong Zhang and Xingxing Li

Abstract Earthquake early warning (EEW), which is considered to be a pragmatic and viable way to reduce the damage and casualties during a large earthquake, relies on the accurate estimation of broadband displacements and the capability of rapidly detection of the first arrival wave (P-wave). Real-time high-rate GNSS is a reliable tool to directly capture displacements including static offsets and dynamic motions at the near field, which does not suffer from the clip, rotation and tilt problems as the traditional seismic instrument does. However, due to the large high-frequency noise and the low sampling rates of GNSS measurements, it is hard to pick up the P-wave arrival accurately in GNSS-derived displacement history. To overcome this problem, the combination of high-rate GNSS and collocated accelerometers shows promise as a more reliable and effective way, because accelerometers perform very well with high precision in the high-frequency range. In this study, we investigate the method by using collocated GNSS and accelerometers for EEW. We first introduce a new approach, namely the temporal point positioning (TPP) method, which could directly obtain coseismic displacement with a single GNSS receiver in real-time. The TPP method overcome the convergence problem of precise point positioning (PPP), and also avoids the integration process of the Variometric approach. And then we apply a multi-rate Kalman filter to fuse GNSS-derived coseismic displacement with collocated accelerometer data for attaining integrated displacements with a high precision and reliability. Finally, we detect the arrival time of P-wave and determine the earthquake magnitude from the integrated results. The performance of collocated GNSS and accelerometers is validated using data from GEONET (1 Hz GPS) and K-NET/KiK-Net (100 Hz accelerometer) stations, with the collocated distance less than 2 km, in the near field of the Mw 9.0

B. Guo (✉) · X. Zhang · X. Li
School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan 430079 Hubei, China
e-mail: guobofeng@whu.edu.cn

X. Li
German Research Centre for Geosciences (GFZ), Telegrafenberg 14473 Potsdam, Germany

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_36,

Tohoku-Oki earthquake occurred on March 11, 2011. Using the broadband displacements derived by the above method, we detect the arrival time of P-wave with a mean value 0.15 s offset different from the USGS reference values, and the estimated magnitude is Mw 9.06, which is achievable within 2–3 min after earthquake initiation.

Keywords Temporal point positioning · Earthquake early warning · Real-time high-rate GNSS · Accelerometer · Integrated displacement

36.1 Introduction

Earthquake early warning (EEW), which is an effective way for the earthquake emergency preparedness and earthquake disaster reduction, relies on the capability of rapidly detection of the first arrival wave (P-wave) and the accurate estimation of broadband displacements [1, 2]. Traditionally, EEW depends on seismic instruments [3–5], because broadband seismometers and strong-motion accelerometers measure velocity and acceleration with great accuracy, making them particularly effective at detecting ground motions [6]. The problem with seismic instruments is that they are prone to saturation, rotation and tilt during large earthquakes, and hence the earthquake-induced displacements through single integration or double integration of the observed waveforms are not reliable in real-time [7], which leads to underestimation of the earthquake's true size in the initial hours, such as the Sumatra earthquake [8] and the Tohoku-Oki earthquake [9]. Real-time high-rate GNSS, unlike seismic instruments, directly capture surface displacements at the sub-centimeter level with regard to an international terrestrial reference frame, meaning that they could provide reliable estimates of displacements during large magnitude events [1, 8]. However, due to the large high-frequency noise and the low sampling rates of GNSS measurements, it is hard to pick up the arrival time of P-wave accurately in GNSS-derived displacement history [10]. The complementary character of GNSS and seismic sensors are well recognized and the combination of them can be mutually beneficial in both sensitivity and large dynamic displacements [11].

In this study, we first introduce a new approach, namely the temporal point positioning (TPP) method, which could directly obtain coseismic (static and dynamic) displacements in real-time with a single GNSS receiver, then we apply a multi-rate Kalman filter to fuse GNSS-derived coseismic displacement with collocated accelerometer data for attaining displacements with a high precision and reliability, and further we detect the arrival time of P-wave, locate the epicenter and determine the magnitude. Finally, the performance of collocated GNSS and accelerometers was validated by using data from GEONET (1 Hz GPS) and K-NET/KiK-Net (100 Hz accelerometer) stations, with the collocated distance less

than 2 km, in the near field of the Mw 9.0 Tohoku-Oki earthquake occurred on March 11, 2011.

36.2 Estimation of GNSS Displacements Based on a Single Receiver

In the seismological applications, the displacement (position variation) of one station due to an earthquake is more curial information for EEW and source inversion. Currently, there are two strategies for a single GNSS receiver to obtain displacements between two appointed epochs. One strategy is that absolute position series are firstly calculated based on PPP [12], and then the displacement is obtained by subtracting one epoch position from the other epoch position [1, 11, 13–16]. The weakness of this approach is a long convergence period of ambiguity fixing. The other strategy is that velocities are determined at first, and then integrated together from one epoch to another epoch [17, 18]. In contrast with PPP, the latter one no longer needs a long convergence period since ambiguities are eliminated using the time difference of phase observations. However, this strategy has to face a problem that the integration process from velocities to displacements leads to a serious accumulated drift as long as the velocities contain few biases.

A novel method, namely TPP, is developed by Li et al. [19] to directly estimate coseismic displacement of one epoch relative to the chosen epoch with known coordinates. The model can be expressed as [19],

$$B^s_{r,j}(t_0) + t_r(t_0) + T^s_r(t_0) = l^s_{r,j}(t_0) + \mathbf{u}^s_r(t_0)\cdot x(t_0) + o^s(t_0) + t^s(t_0) + I^s_{r,j}(t_0) - \varepsilon^s_{r,j}(t_0) \tag{36.1}$$

$$\mathbf{u}^s_r(t_n)\cdot x(t_n) - t_r(t_n) - T^s_r(t_n) = -l^s_{r,j}(t_n) - t^s(t_n) - o^s(t_n) + B^s_{r,j}(t_n) - I^s_{r,j}(t_n) + \varepsilon^s_{r,j}(t_n) \tag{36.2}$$

$$B^s_{r,j}(t_n) = B^s_{r,j}(t_0) \tag{36.3}$$

$$\mathbf{u}^s_r(t_n)\cdot x(t_n) - \Delta t_r(t_0,t_n) = \mathbf{u}^s_r(t_0)\cdot x(t_0) - \Delta I^s_{r,j}(t_0,t_n) - \Delta t^s(t_0,t_n) - \Delta o^s(t_0,t_n) - \Delta I^s_{r,j}(t_0,t_n) + \Delta T^s_{r,j}(t_0,t_n) + \Delta\varepsilon^s_{r,j}(t_0,t_n) \tag{36.4}$$

where, s stands for a satellite, r stands for the GPS receiver, j stands for the frequency of L1 or L2 carrier phase, $l^s_{r,j}$ is "observed minus computed" phase observations; u^s_r is the receiver-to-satellite unit direction vector; x denotes receiver position; o^s denotes satellite orbit error; $T^s_r, I^s_{r,j}$, denote tropospheric and

ionospheric delay; t^s, t_r are satellite and receiver clock errors; $B^s_{r,j}$ is the real-valued phase ambiguity; $\varepsilon^s_{r,j}$ are measurement noise of carrier phase.

Equation (36.4) is final model of TPP, which is deduced from Eqs. (36.1) to (36.3). From the view of positioning, the core of the TPP model is that the estimated ambiguities $B^s_{r,j}$ is fixed from the epoch t_0 to t_n, and the well-known receiver position $x(t_0)$ could be determined before the earthquake. With the satellite ephemeris and a certain prior models aided, all errors in Eq. (36.4) are corrected. Therefore, the receiver position $x(t_n)$ can be accurately resolved and the associated displacement is obtained. From the view of position variation (displacement), Eq. (36.4) is in the same form as the time-differenced equation of phase observations under the condition that there are no cycle slips in observation data. After all errors corrected, the displacement of an arbitrary epoch t_n with respect to the chosen epoch t_0 could be directly retrieved, which avoids the integration process of the Variometric approach described in Colosimo et al. [17]. Worth noting that different from the Variometric approach, the TPP method is based on the known accurate position at the chosen epoch, which is an important factor for achieving high-accuracy displacements [19].

36.3 Integrated Displacements Using a Kalman Filter

In order to estimate very high-rate accurate and reliable displacements in real time, a multi-rate Kalman filter to fuse GNSS-derived coseismic displacement with collocated accelerometer data is applied [10, 20].

In the integrated procedure, the discrete state equation can be expressed as [10],

$$x_k = \Phi_{k,k-1}x_{k-1} + B_{k,k-1}a_{k-1} + w_k \quad w_k \sim N(0, Q_{w_k}) \tag{36.5}$$

$$\Phi_{k,k-1} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \qquad B_{k,k-1} = \begin{bmatrix} \tau^2/2 \\ \tau \end{bmatrix}$$

where $x_k = [d_k \quad v_k]^T$ represents the state vector, $\Phi_{k,k-1}$ is a transition matrix, $B_{k,\ k\text{-}1}$ is an input matrix, τ is the sampling interval of accelerometer. $a_{k\text{-}1}$ is the raw accelerometer data as the system inputs. w_k is the state noise with the covariance matrix Q_{wk}, which can be expressed as a function of Kalman filter process noise,

$$Q_{w_k} = \begin{bmatrix} \tau^3/3 \cdot \sigma_a^2 & \tau^2/2 \cdot \sigma_a^2 \\ \tau^2/2 \cdot \sigma_a^2 & \tau \cdot \sigma_a^2 \end{bmatrix} \tag{36.6}$$

where σ_a^2 is the acceleration variance as a system dynamic noise.

The observational equation for the Kalman filter can be written as

$$l_k = A_k x_k + e_k \quad e_k \sim N(0, R_{e_k}) \tag{36.7}$$

$$R_{e_k} = \frac{r}{\tau_{GNSS}}$$

where l_k represents the observation (here is the GNSS displacement), $A_k = [1 \quad 0]^T$ is the design matrix, and e_k is the observation noise with R_{e_k} as its covariance matrix. The value of R_{e_k} is r/τ_{GNSS}, r denotes the GNSS displacement noise, and τ_{GNSS} is the GNSS sampling interval.

The integrated results can be solved by the time update of Eq. (36.5) and the measurement update of Eq. (36.7) through a standard Kalman filter. It is noted that the time update is executed at every accelerometer sampling, and the measurement update with GNSS displacement is applied at every GNSS epoch.

36.4 Application to the 2011 Mw 9.0 Tohoku-Oki Earthquake

The Tohoku-Oki earthquake, occurred at 05:46:39 GPST on 11 March, 2011 in Japan with the moment magnitude 9.0, are well recorded by both accelerometer stations and high-rate GPS receivers. Following we will use the data to evaluate the performance of integrated displacements derived by the proposed method above.

Figure 36.1 shows a map of collocated GPS and accelerometer stations in the near-field areas within about 500 km from the epicenter (38.322°N, 142.369°E) as determined by the U. S. Geological Survey (abbr. USGS). GPS and accelerometer station pairs are chosen with distance less than 2 km of each other, which are considered effectively collocated [21]. GPS data are collected from GEONET stations with 1 Hz sample rate, and 100 Hz accelerometer data of K-Net and KiK-Net stations are provided by NIED.

36.4.1 TPP Analysis Using Only GPS Data

To evaluate the accuracy of the TPP method, we firstly processed all the GPS data collected at about 80 collocated stations before the earthquake. The results were converted to North/East/Up component and were compared with the zero displacement of truth. The TPP displacements of station 0550 for 15 min interval from 05:20:00 to 05:35:00 (GPST) are shown in Fig. 36.2. The red lines are the results, with consideration of three main error components mentioned in Sect. 36.2, using precise ephemeris, the ionosphere-free linear combination observation and the

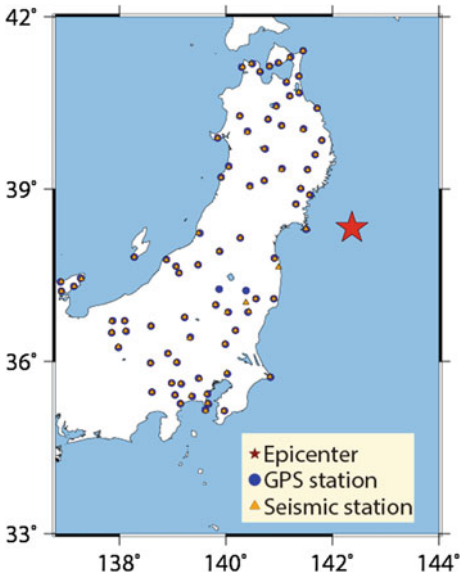


Fig. 36.1 Location of epicenter of the Tohoku-Oki earthquake (38.322°N, 142.369°E) and the distribution map of collocated GPS and accelerometer stations. The *red star* stands for the epicenter. The *blue circles* stand for GPS sites. The dark *yellow triangles* represent accelerometer sites

accurate initial coordinates (PPP_POPC_LC solution). Meanwhile, we draw other displacements with different processing strategies to reflect the impact of those errors on TPP displacements. The blue line shows the results using precise ephemeris, the ionosphere-free linear combination observation and the coordinates derived from the standard position positioning (SPP) with 0.7 m error for North, −0.1 m error for East, and 3.0 m error for Up component (SPP_POPC_LC solution). The black lines show the displacements using broadcast ephemeris, the ionosphere-free linear combination observation and the SPP-derived initial position (SPP_BOBC_LC solution). The results in cyan line are obtained by using broadcast ephemeris, the L1 carrier phase measurement and the SPP-derived initial position (SPP_BOBC_L1 solution).

From Fig. 36.2, the PPP_POPC_LC solution is close with the zero truth line with no obvious drift. After 15 min, the displacements are merely −1.0, −0.6, and 2.5 cm for north/east/up components, respectively. Compared with the PPP_POPC_LC solution, the SPP_POPC_LC solution has a visible drift, and the drift values for 15 min are 6.6, 22.0, and 18.3 cm in north/east/up component. The SPP_BOBC_LC solution becomes more fluctuant with long-term variations due to the imprecise satellite orbit and clock correction, and the biases are several decimetres in the horizontal and vertical component. The ionospheric delay seems to be a big error source of TPP displacements, which could lead to about 1 m bias in the east and up component by comparison of the SPP_BOBC_LC solution and SPP_BOBC_L1 solution. From all four results, it is demonstrated that the PPP_POPC_LC solution is the most precise, which satisfies the demands of seismic monitoring and EEW. Considering that real-time precise orbit and precise clock corrections are available online via the IGS real-time pilot project, and the

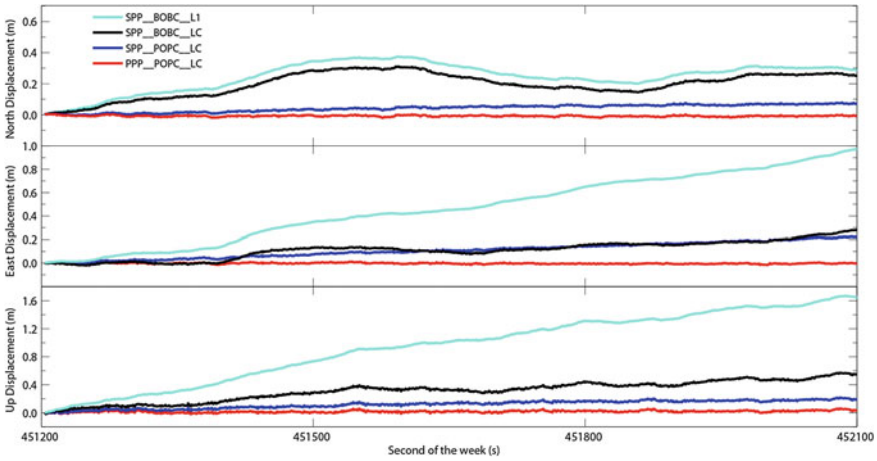


Fig. 36.2 Displacements of station 0550 (38.3012°N, 141.5009°E) derived from TPP method. From *top* to *bottom* are the results in North/East/Up components

Table 36.1 Root mean squares of the drift errors at about 80 sites

RMS (min)	North (cm)	East (cm)	Up (cm)
15	1.05	0.84	1.92
10	0.92	1.02	1.67
5	0.89	0.78	2.08
3	0.78	0.88	1.75
1	0.68	0.39	1.53

accurate site coordinates is usually well known, and thus the PPP_POPC_LC solution is accessible in real time during an earthquake. Here, we mainly focus on the PPP_POPC_LC solution in this study.

For the static period, TPP-derived displacements of all 80 sites at one certain epoch are compared with zero to calculate the root mean squares (RMS) as an indication of the TPP displacement accuracy. We respectively calculate the RMS values at different duration time (1, 3, 5, 10 and 15 min). The results are summarized in Table 36.1. The RMSs of the PPP_POPC_LC solution are stable from 1 to 15 min, and the values in north and east are almost within 1 cm, and in up with 2 cm, which is comparable to the accuracy of PPP after convergence period.

36.4.2 Integrated Displacements Analysis

In order to assess the performance of the integrated displacements, we compare them with GPS-only results derived from TPP and seismic-only waveforms derived from double integration of uncorrected acceleration data. As an example,

the nearest station pair is selected, which is K-NET station MYG011 (38.3052°N, 141.5044°E) and the GEONET station 0550 (38.3012°N, 141.5009°E), being separated by 560 m and located 75.7 km from the epicentre of the Tohoku-Oki earthquake. We adopt the approach stated in Sect. 39.3 to combine raw 100 Hz accelerations with 1 Hz TPP-derived displacements. The GPS displacement error is set to 1.0 cm for north/east and 2.0 cm for up component based on the above statistical RMS values. The acceleration noise is determined from windows (2–3 s) of pre-event accelerometer data or introduced from the empirical value, such as the K-NET acceleration precision is 0.015 cm/s^2 [21]. Trifunac and Todorovska [22] demonstrated the rotation and tilt of accelerometer is a main source of acceleration noise during shaking, and thus we enlarge the pre-event acceleration noise by 1,000 times when the acceleration record value exceeds a certain times of the pre-event acceleration precision [10]. In addition, the acceleration changes rapidly during a large earthquake, and sometimes the change value in the sample interval (such as 0.01 s) could reach to several cm/s^2 [23]. In such case, the system dynamics variance should be adjusted larger to fit the rapid state change.

Figure 36.3 shows the comparison of the integrated displacements and GPS-only/seismic-only displacements for the collocated stations MYG011/0550, and three sub-figures depict displacements for north/east/up components from the left to right. The uncorrected seismic displacements trace out the dynamic motions, but the latter portion of displacements drifts with a linear or parabolic trend. GPS-only solution is not drift in the displacement history and clearly depicts peak displacements and permanent offsets, which is demonstrated that the GPS-only displacements are good at low frequency band. However, the discrepancy of GPS-only displacement from the integrated displacements depicts a high-frequency noise of GPS-only results because of the low precision of GPS.

Compared to the GPS-only solution, the integrated displacements are remarkably improved by acceleration data in aspect of dynamical displacement precision, which can be used to detect P-wave arrival. We adopt the STA/LTA ratio method to automatically locate the arrival time of P-wave just as the same as done by Li et al. [2]. The window length of STA is 0.2 s, the window length of LTA is 2 s, and meanwhile the threshold of STA/LTA ratio is 10. As an example, Fig. 36.4 shows the first 20 s of the integrated results for station MYG011/0550 in up component and its corresponding STA/LTA ratio values. It is clearly shown that the STA/LTA ratio dramatically increases over the pre-set threshold when P-wave appears, and the detected the arrival time of P-wave is 14.08 s after the origin time. Meanwhile, we do the same procedure for other four collocated station pairs (MYG001/0172, MYG003/0914, IWT013/0169, FKS001/0038) with the distance away the epicenter less than 150 km to obtain the P-wave arrival time, and the detected results are 17.33, 18.80, 20.01, 21.23 s, respectively, with a mean value 0.15 s offset different from the theoretical P-wave travel times calculated by TauP Toolkit.

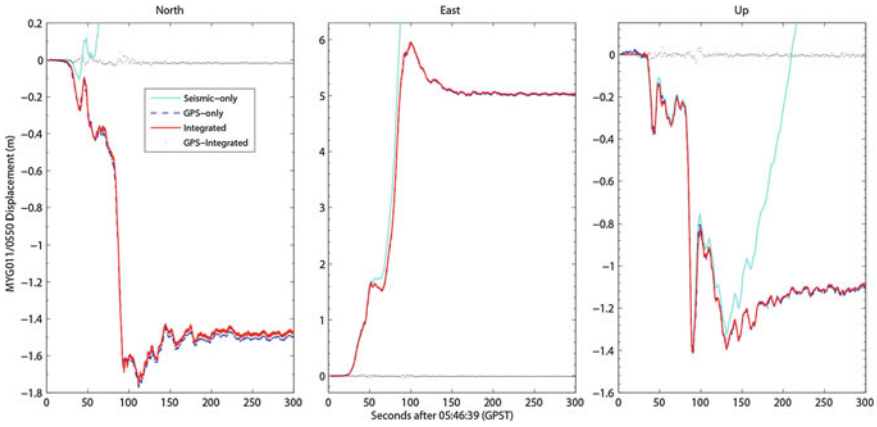


Fig. 36.3 Comparison of seismic-only, GPS-only, and integrated displacements on the collocated MYG011 and 0550 stations for 300 s from the origin of the Tohoku-Oki earthquake. The *red line* depicts the integrated displacement (100 Hz), the *cyan line* depicts the uncorrected seismic-only displacement (100 Hz), the *blue dotted line* depicts the GPS-only displacement (1 Hz), and the *black dots* depict the discrepancy between integrated displacements and GPS-only displacements

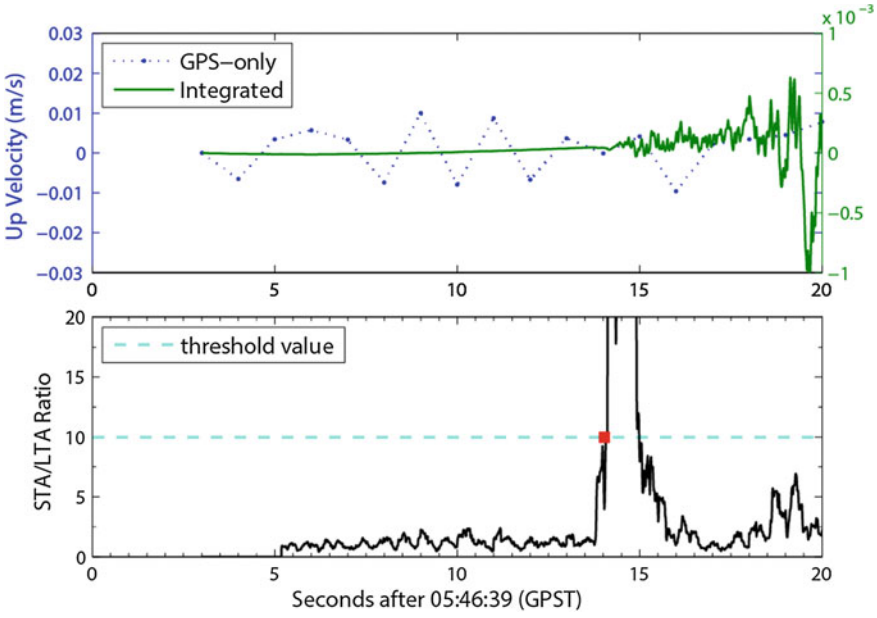


Fig. 36.4 The first 20 s of the integrated velocities in up component on the collocated MYG011/0550 stations pair from the origin time of the Tohoku-Oki earthquake. The *green line* shows the integrated velocities (100 Hz), and the *blue dotted line* shows the GPS-only velocities (1 Hz). The *black line* in the *bottom* sub-figure is STA/LTA ratio results calculated from integrated velocities, and the *red square point* denotes the first arrival time of P-wave

36.4.3 Magnitude Determination Through Attenuation Relationships of Peak Broadband Displacement

For the issuing of effective tsunami warnings and emergency response, the rapid and accurate determination of the earthquake magnitude soon after a great earthquake is necessary. In order to rapidly determine the magnitude of impending earthquakes, several magnitude estimation methods based on P-wave have been developed [4, 24], which are effective for relatively small earthquakes. But for large earthquakes, they are problematic due to saturation effects [6]. To avoid the saturation problem, moment magnitude is applied for large earthquake magnitudes. However, moment magnitude, based on slip distribution inverted, is difficult to estimate in real time [25]. Taking into account the above-mentioned considerations, here we adopt the empirical scaling relationship between magnitude and peak ground displacement (PGD) to estimate the magnitude, which is deduced by Crowell et al. [26]. The scaling relationship is expressed as a function of moment magnitude M_w and hypocentral distance R,

$$\log_{10}(PGD) = -5.013 + 1.219M_w - 0.178 \cdot \log_{10}(R) \tag{36.8}$$

We apply Eq. (36.8) to estimate the expected moment magnitude based on the integrated displacement results at every accelerometer sampling. Figure 36.5 shows the magnitude estimates as a function of time after earthquake origin for three collocated stations pairs. The pair MYG011/0550 station is the nearest from the earthquake source (source depth is defined as 29 km), the pair FKS013/0211 is located about 211 km from the source, and another pair KNG011/3065 is about 416.2 km. For the pair MYG011/0550 station, the estimated magnitude becomes stable at 115 s after the origin time of the earthquake, with a value of Mw 8.88. For the pair FKS013/0211 station, the final estimated magnitude is Mw 9.12 at 145 s. And the magnitude determined by the pair KNG011/3065 station could reach Mw 9.16 at 196 s. It is demonstrated that the moment magnitude can be correctly determined within 2–3 min after earthquake initiation. In Fig. 36.6, the recorded peak displacements distribute around the reference line, indicating that the attenuation in log of peak displacement is proportional to log of hypocentral distance. The magnitude estimates from all 80 collocated stations are well within the uncertainties of the scaling relationship (the uncertainty is 0.224 demonstrated in Crowell et al. [26]), with a mean magnitude value is 9.06. It is shown that a reliable magnitude is achievable within 2–3 min using a few near-field collocated stations.

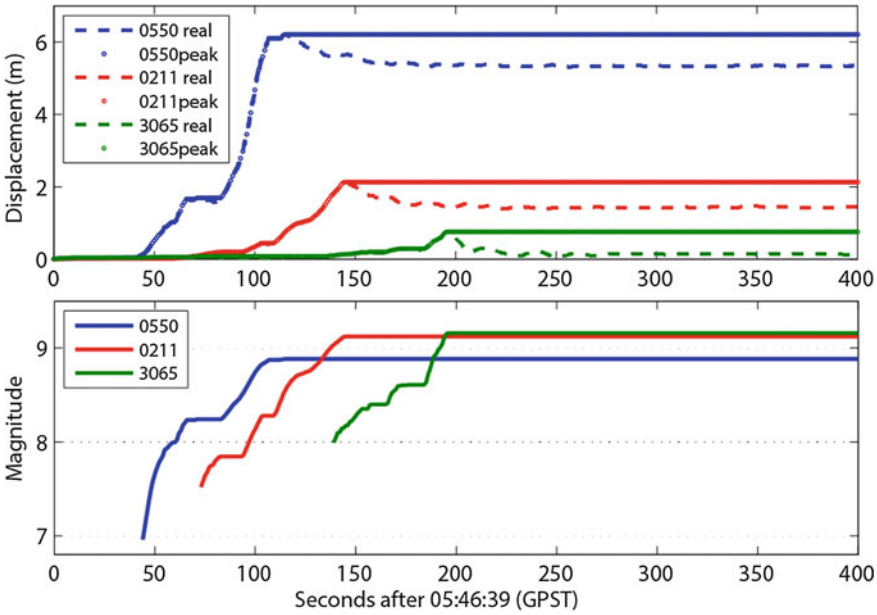


Fig. 36.5 Magnitude estimates after earthquake origin for the Tohoku-Oki earthquake. The *top* sub-figure is the 3D displacement waveform for three collocated stations pairs MYG011/0550 (*blue*), FKS013/0211 (*red*), and KNG011/3065 (*green*). The *solid line* represents the real displacement, and the *dot line* denotes the peak displacement up to the current

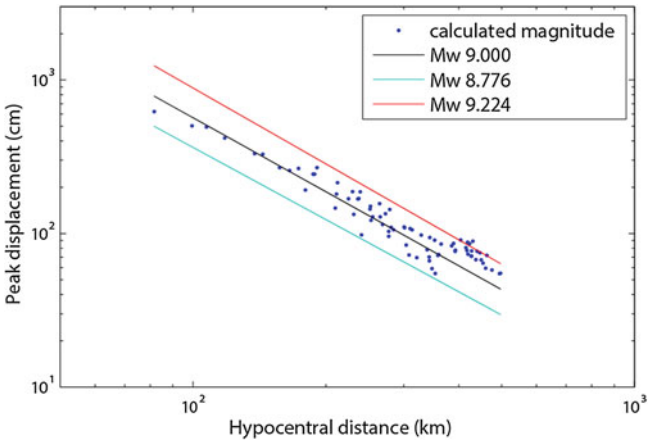


Fig. 36.6 Relation between peak 3D displacement and hypocentral distance. The *black line* represents the reference value (Mw 9.0) calculated by Eq. (36.8) with one standard deviation of 0.224 in *red line* (Mw 9.224) and *cyan line* (Mw 8.776). The recorded displacements of collocated stations (*blue dots*) scatter around the expected line of Mw 9.0

36.5 Conclusions

This study investigates the joint use of the collocated GNSS and acceleration data for earthquake early warning. The performance was validated based on the collocated GPS and acceleration data collected during the 2011 Tohoku-Oki earthquake, and here comes some conclusion as follows,

(1) The TPP method is a quite suitable way for obtaining coseismic displacement with a single GNSS receiver in real-time, which overcomes the convergence problem of PPP and avoids the integration process for the velocity estimation approach.
(2) The accuracy of TPP solution by using the precise satellite ephemeris, the ionosphere-free linear combination observation and the accurate initial coordinates are 1 cm in north and east component, and 2 cm in up component during a few to tens of minutes.
(3) A more accurate and reliable broadband displacement could be obtained by a multi-rate Kalman filter, which provides the full spectrum of the seismic motion. It is noteworthy that the acceleration noise as a system process noise should be adjusted larger to fit the rapid state change with consideration of the station dynamic status during shaking period.
(4) Based on the broadband displacements, the arrival time of P-wave can be picked up accurately with a mean value 0.15 s offset, and the magnitude for large earthquake could be estimated robustly and rapidly. For the Tohoku-Oki earthquake, the estimated magnitude is Mw 9.06, which can be determined within 2–3 min after earthquake initiation.

Acknowledgments The K-NET and KiK-Net strong-motion data for the 2011 Tohoku earthquake were provided by the National Research Institute for Earth Science and Disaster Prevention (NIED) of Japan. GEONET (GPS Earth Observation Network System) data was provided by the Geospatial Information Authority of Japan (GSI). Thanks also to the International GNSS Service (IGS) for providing GPS data of globally distributed reference stations. This study was supported by Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20130141110001), National Natural Science Foundation of China (Grant No. 41074024, No. 41204030), National 973 Project China (Grant No. 2013CB733301), the Fundamental Research Funds for the Central Universities (No.: 2012214020207) and the Surveying and Mapping Foundation Research Fund Program, National Administration of Surveying, Mapping and Geoinformation (12-02-010).

References

1. Allen RM, Ziv A (2011) Application of real-time GPS to earthquake early warning. Geophys Res Lett 38:L16310
2. Li X, Zhang X, Guo B (2013) Application of collocated GPS and seismic sensors to earthquake monitoring and early warning. Sensors 13:14261–14276

3. Aranda JE, Jimenez A, Ibarrola G, Alcantar F, Aguilar A, Inostroza M, Maldonado S (1995) Mexico City seismic alert system. Seismol Res Lett 66:42–53
4. Allen RM, Kanamori H (2003) The potential for earthquake early warning in southern California. Science 300:786–789
5. Kanamori H (2007) Real-time earthquake damage mitigation measures. In: Gasparini P, Manfredi G, Zschau J (eds) Earthquake early warning systems. Springer, Berlin, Heidelberg
6. Allen RM, Gasparini P, Kamigaichi O, Bose M (2009) The status of earthquake early warning around the world: an introductory overview. Seismol Res Lett 80:682–693
7. Wang R, Schurr B, Milkereit C, Shao Z, Jin M (2011) An improved automatic scheme for empirical baseline correction of digital strong-motion records. B Seismol Soc Am 101:2029–2044
8. Blewitt G, Kreemer C, Hammond WC, Plag HP, Stein S, Okal E (2006) Rapid determination of earthquake magnitude using GPS for tsunami warning systems. Geophys Res Lett 33:L11309
9. Ohta Y, Kobayashi T, Tsushima H, Miura S, Hino R, Takasu T, Fujimoto H, Iinuma T, Tachibana K, Demachi T, Sato T, Ohzono M, Umino N (2012) Quasi real-time fault model estimation for near-field tsunami forecasting based on RTK-GPS analysis: application to the 2011 Tohoku-Oki earthquake (Mw 9.0). J Geophys Res 117:B2311
10. Bock Y, Melgar D, Crowell BW (2011) Real-time strong-motion broadband displacements from collocated GPS and accelerometers. B Seismol Soc Am 101:2904–2925
11. Xu P, Shi C, Fang R, Liu J, Niu X, Zhang Q, Yanagidani T (2013) High-rate precise point positioning (PPP) to measure seismic wave motions: an experimental comparison of GPS PPP with inertial measurement units. J Geodesy 87:361–372
12. Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102:5005–5017
13. Kouba J (2003) Measuring seismic waves induced by large earthquakes with GPS. Stud Geophys Geodesy 47:741–755
14. Wright TJ, Houlie N, Hildyard M, Iwabuchi T (2012) Real-time, reliable magnitudes for large earthquakes from 1 Hz GPS precise point positioning: the 2011 Tohoku-Oki (Japan) earthquake. Geophys Res Lett 39:L12302
15. Li X, Ge M, Zhang X, Zhang Y, Guo B, Wang R, Klotz J, Wickert J (2013) Real-time high-rate coseismic displacement from ambiguity-fixed precise point positioning: application to earthquake early warning. Geophys Res Lett 40:295–300
16. Geng J, Bock Y, Melgar D, Crowell BW, Haase JS, Cecil H, Ida M (2013) A new seismogeodetic approach applied to GPS and accelerometer observations of the 2012 Brawley seismic swarm: implications for earthquake early warning. Geochem Geophys Geosyst 14:2124–2142
17. Colosimo G, Crespi M, Mazzoni A (2011) Real-time GPS seismology with a stand-alone receiver: a preliminary feasibility demonstration. J Geophys Res 116:B11302
18. Zhang X, Guo B (2013) Real-time tracking the instantaneous movement of crust during earthquake with a stand-alone GPS receiver. Chin J Geophys 56:1928–1936 (in Chinese)
19. Li X, Ge M, Guo B, Wickert J, Schuh H (2013) Temporal point positioning approach for real-time GNSS seismology using a single receiver. Geophys Res Lett 40:5677–5682
20. Smyth A, Wu M (2007) Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring. Mech Syst Signal Process 21:706–723
21. Emore GL, Haase JS, Choi K, Larson KM, Yamagiwa A (2007) Recovering seismic displacements through combined use of 1-Hz GPS and strong-motion accelerometers. B Seismol Soc Am 97:357–378
22. Trifunac MD, Todorovska MI (2001) A note on the usable dynamic range of accelerographs recording translation. Soil Dyn Earthq Eng 21:275–286

23. Tu R, Ge M, Wang R, Walter TR (2014) A new algorithm for tight integration of real-time GPS and strong-motion records, demonstrated on simulated, experimental, and real seismic data. J Seismol 18:151–161
24. Kanamori H (2005) Real-time seismology and earthquake damage mitigation. Annu Rev Earth Planet Sci 33:195–214
25. Fang R, Shi C, Song W, Wang G, Liu J (2013) Determination of earthquake magnitude using GPS displacement waveforms from real-time precise point positioning. Geophys J Int. doi:10.1093/gji/ggt378
26. Crowell BW, Melgar D, Bock Y, Haase JS, Geng J (2013) Earthquake magnitude scaling using seismogeodetic data. Geophys Res Lett 40:6089–6094

Chapter 37
Assessment of GPT2 Empirical Troposphere Model and Application Analysis in Precise Point Positioning

Weirong Chen, Chengfa Gao and Shuguo Pan

Abstract Precise Point Positioning (PPP) has been demonstrated to be a powerful tool in geodetic applications, such as deformation monitoring. Troposphere delay is an important error source which directly affects positioning accuracy in height direction. At the end of 2012, an improved model named Global Pressure and Temperature 2 (GPT2) was proposed. Compared with early empirical models, this new model mainly eliminates the weakness of limited spatial and temporal variability. In this study, we assess the precision of GPT2 model and apply it in PPP analysis. The analysis data of VMF1, which is produced using European Centre for Medium-Range Weather Forecasts (ECMWF), provides the nearly true value of zenith delays and mapping function for International GNSS Service (IGS) stations. Therefore a globally distributed set of 11 IGS stations is chosen to validate GPT2 model. In the case of using GPT2 as a priori model while the residual zenith delay still estimated with other unknown parameters, it would improve the zenith troposphere delay adjustments to nearly zero-mean. Therefore GPT2 is helpful to improve the efficiency in PPP data processing. However, GPT2 model is only resting upon a global 5° grid and sufficient global troposphere models are not yet available. Due to the complexity of wet zenith delay, PPP height solutions would be unsatisfactory when residual troposphere delay is not parameterized. We conclude that GPT2 is capable of predicting troposphere delay worldwide with an acceptable uncertainty. And GPT2-based PPP solution performs well only in the case of regarding residual model error as an additional unknown parameter.

Keywords Global pressure and temperature 2 · Zenith troposphere delay · Mapping function · Precise point positioning · Global navigation satellite system

W. Chen (✉) · C. Gao
School of Transportation, Southeast University, No. 2 Sipailou, Nanjing, China
e-mail: cwr@seu.edu.cn

S. Pan
School of Instrument Science and Engineering, Southeast University, No. 2 Sipailou, Nanjing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_37,

37.1 Introduction

With the help of precise satellite orbit and clock available from International GNSS service (IGS), Precise Point Positioning (PPP) [1, 2] becomes a very pragmatic tool in geodetic application, such as deformation monitoring. In satellite positioning, the Earth's neutral atmosphere causes propagation delay of electromagnetic signal. The propagation path delay, which is often called troposphere delay, could be about 2.3 m at the zenith direction and 20 m at lower receiver-to-satellite direction such as elevation angles below 10° [3]. Therefore it is an important error source for precise geodetic positioning. As the troposphere is a non-dispersive medium, combination of multiple frequency observations is no longer applicable to remove troposphere delay. It is normally modelled as the sum of two components, which are hydrostatic (dry) and water vapor (wet) delays [4, 5]. Both of them could be decomposed into a zenith delay and a mapping function. The accuracy of Global Navigation Satellite System (GNSS) positioning is greatly affected by the residual errors in modelling troposphere delay. Both zenith delay modelling and mapping function have an important role.

In the last few years, the International Association of Geodesy (IAG) recommended a new procedure to compute the zenith delay [6]. Based on this procedure, there are some most commonly used models today. Hydrostatic zenith delay could be computed from the surface pressure according to the formula of Saastamoinen [7] as given by Davis et al. [8]. And a priori zenith wet delay could also be computed provided that the local temperature and water vapor pressure is known. The instantaneous local meteorological data are always obtained from global empirical models, which approximate the spatial and temporal variability, such as Global Pressure and Temperature (GPT) model [9, 10].

Nowadays, the latest empirical model used in GNSS processing model is called GPT2 [11]. GPT2 model was proposed by Vienna University of Technology at the end of 2012. Monthly mean profiles provided by European Centre for Medium-Range Weather Forecasts (ECMWF) from 2001 to 2010 were analyzed. After the calculation of mean, annual, and semi-annual terms of meteorological data based on global grid points, the results were expressed worldwide in the form of 5° grid at mean Earth's surface heights. In addition, hydrostatic and wet mapping function coefficients could also be obtained. This updated version gives much better spatial and temporal resolution than the prior GPT model.

The main purpose of this paper is to evaluate GPT2 model, which is called the state-of-the-art empirical troposphere model. Its accuracy and suitability for precise GNSS analyses are what we care about. For evaluation, the benchmark should be defined ahead, and then the differences compared with the benchmark could be analyzed. In situ measured troposphere delay is the best choice, but locally meteorological sensor would not be always available for most sites. Hence, a ray-traced numerical weather model (NWM) can be used to retrieve required data. As for the ECMWF NWM, it provides a realistic representation of the slant delay on sub-daily basis. Although a password is needed for any data of ECWMF, the

access of Vienna Mapping Functions 1 (VMF1) requires no password. VMF1 is based on ray-tracing through ECMWF operational analysis at an elevation angle of 3.3° [12]. Both zenith delay and mapping function could be obtained from VMF1. It can act as a rigorous benchmark. For this reason, VMF1 was undertaken to assess GPT2 model in two aspects, zenith delay modelling and mapping function. In addition, GPT2 empirical model is applied in PPP data processing, which will help assess the impact of model error in precise geodetic positioning.

This paper is organized as follows. Section 37.2 details the realization process of slant troposphere delay calculation using GPT2 model. Section 37.3 presents a case study, and compares the GPT2 model values with the reference values from the VMF1 model. Section 37.4 describes the resultant position errors due to the implementation of GPT2 model. Finally, Sect. 37.5 draws the conclusions.

37.2 GPT2-Based Troposphere Delay Model

37.2.1 Zenith Troposphere Delay

The hydrostatic delay component accounts for roughly 90 % of the total troposphere delay at any site globally. It can be accurately computed a priori based on reliable surface pressure data using the formula of Saastamoinen [7] as given by Davis et al. [8]:

$$Z_h = \frac{0.0022768P}{1 - 0.00266\cos(2B) - 0.28 \times 10^{-6}H} \tag{37.1}$$

where Z_h is hydrostatic zenith delay in meters, P is barometric pressure in millibars, B is geodetic latitude in radians and H is geodetic height in meters.

On the other hand, wet zenith delay component is more difficult to model accurately and is therefore estimated as an unknown along with other geodetic quantities of interest. Although a loss of precision would be involved, the wet zenith delay could also be computed using Saastamoinen formula:

$$Z_w = \frac{0.0022768 \times \left(\frac{1255}{T} + 0.05\right) \times e}{1 - 0.00266\cos(2B) - 0.28 \times 10^{-6}H} \tag{37.2}$$

where Z_w is wet zenith delay in meters, T is temperature in Kelvin, e is water vapour pressure in millibars.

A station approximate position is readily available most of the time. Therefore the hydrostatic and wet zenith delay could be easily evaluated from (37.1) and (37.2) as long as the station meteorological data are obtained. As 1 mbar pressure error causes an a priori delay error of about 2.3 mm at sea level, it is essential to

use accurate estimates of meteorological data [13]. Hence, in the case of local meteorological sensor is not available and NWM has a large computational burden, a precise empirical meteorological model is recommended [14].

37.2.2 Mapping Function

Mapping function is a mathematical expression to describe how the slant troposphere delay varies with respect to receiver-to-satellite elevation angle. Many mapping functions were proposed in the past, such as FCULa, NMF, VMF1 and GMF [15–17]. By comparison, it could be found that there is a general form of the hydrostatic and wet mapping functions:

$$MF_{h,w} = \frac{1 + \frac{a}{1+\frac{b}{1+c}}}{\sin(\varepsilon) + \frac{a}{\sin(\varepsilon)+\frac{b}{\sin(\varepsilon)+c}}} \tag{37.3}$$

where *MF* is the mapping function, hydrostatic and wet component represented by the subscript 'h' and 'w', ε is the elevation angle, a, b and c are empirical coefficients with different values in various mapping functions. As the VMF1 is produced using ECMWF, it is recommended for any global application for its high precision. VMF1 relies on empirical equations for 'b' and 'c' coefficients of the continued fraction form, whereas the 'a' coefficients are determined from rigorously ray-traced mapping functions at an elevation angle of 3.3° [12]. Therefore, its analysis data of the ECMWF are made available with a delay. Although it is said that the latency would be less than 34 h, real-time processing is not realizable using VMF1. The 'a' coefficients should be determined in another way.

37.2.3 GPT2 Empirical Model

Strictly speaking, GPT2 is not a troposphere model. It only provides an estimate of meteorological data and the 'a' coefficients for any global grid points near the Earth surface. It provides mean values A_0 as well as annual (A_1, B_1) and semi-annual (A_2, B_2) variations based on a 5 × 5° external grid file [11]. GPT2 model could be called once per 24 h session per station using the equations as follows:

$$\begin{aligned} X = A_0 + A_1 \cos\left(\frac{doy}{365.25}2\pi\right) + B_1 \sin\left(\frac{doy}{365.25}2\pi\right) + \ldots \\ A_2 \cos\left(\frac{doy}{365.25}4\pi\right) + B_2 \sin\left(\frac{doy}{365.25}4\pi\right) \end{aligned} \tag{37.4}$$

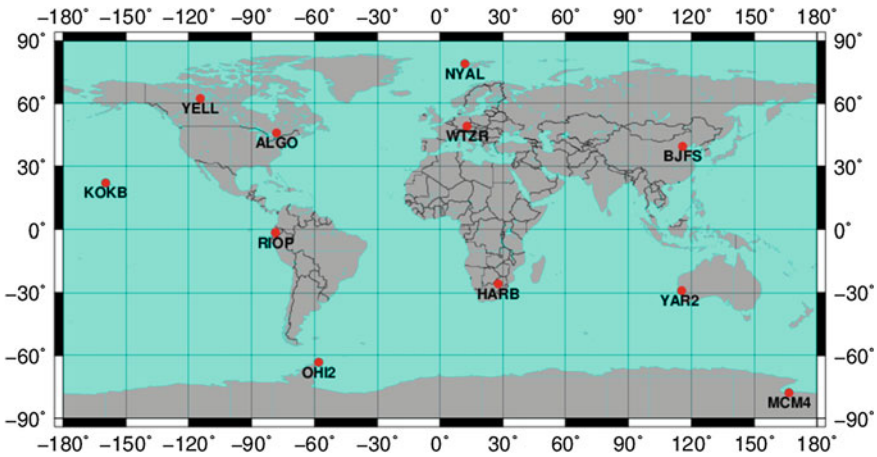


Fig. 37.1 Locations of the selected IGS stations used for model testing

where X denotes a vector of parameters, which includes barometric pressure, temperature, temperature lapse rate, specific humidity, hydrostatic and wet mapping function coefficients and geoid undulation, *doy* is the specific time of year.

After loading these parameters from an external grid ASCII-file, there is still some work to do. As all the values are provided at mean Earth's surface heights, a difference exists between station height and grid height. Therefore gradient correction should be applied to obtain station barometric pressure and temperature. In addition, when the barometric pressure is given, specific humidity could be transferred to water vapour pressure. Finally, a bilinear interpolation should be done from the four nearest grid points. With the help of GPT2, both zenith troposphere delay and mapping function could be determined using the method described before.

In this section, we detail the process of GPT2-based troposphere delay model. To obtain zenith delay, Saastamoinen formula is applied with meteorological data determined by GPT2. While in the aspect of mapping function, GPT2 is also used to generate some important coefficients. With these coefficients, VMF1 could be applied with no need for the 6-hourly analysis data of ECMWF.

37.3 A Case Study at Permanent Stations

37.3.1 Data Pre-processing

To evaluate GPT2-based troposphere delay model presented in the previous section, a globally distributed set of 11 IGS permanent station data are used. This set of 11 well-performing IGS stations (shown in Fig. 37.1), representing the polar,

Table 37.1 Positions of the stations and hydrostatic zenith delay modelling differences (cm) with respect to VMF1 during 2012

Station	Ellipsoidal latitude (degree)	Ellipsoidal height (m)	DOY since 2012					Mean	STD
			001	091	181	271	361		
NYAL	78.93	78.57	1.29	1.65	0.92	0.12	3.73	1.54	1.35
YELL	62.48	180.95	−2.02	5.08	1.38	1.43	−2.88	0.60	3.18
WTZR	49.14	666.02	−0.03	1.66	0.77	0.95	0.89	0.85	0.60
ALGO	45.96	200.93	0.38	−0.23	2.29	−2.30	−3.31	−0.63	2.22
BJFS	39.61	87.47	−0.59	−0.16	0.01	0.20	−0.81	−0.27	0.42
KOKB	22.13	1167.36	0.17	−0.37	−0.03	0.21	−0.13	−0.03	0.24
RIOP	−1.65	2817.19	0.83	0.70	0.96	0.61	0.75	0.77	0.13
HARB	−25.89	1558.08	0.17	1.25	−0.12	0.95	0.36	0.52	0.56
YAR2	−29.05	241.29	−0.05	−0.26	−1.32	−0.21	0.66	−0.24	0.71
OHI2	−63.32	32.49	−0.87	−2.22	3.63	−2.70	−4.09	−1.25	2.96
MCM4	−77.84	97.96	2.16	−1.07	−6.51	−2.29	−1.94	−1.93	3.11
Mean			0.13	0.55	0.18	−0.28	−0.62	−0.01	
STD			1.11	1.91	2.58	1.46	2.28		

mid-latitude and equatorial regions, is considered to be a sufficient data sample for this purpose.

In addition, VMF1 model are used as the nearly true value. Relevant data could be downloaded by the service at the Technical University of Vienna. VMF1 was originally site specific, but now a gridded version has also been introduced. We have chosen the site-VMF1 for the analysis, but as Kouba found negligible differences between the gridded and site-VMF1 [18], the results should be applicable to both.

The assessment is performed in two steps. First, Z_h and Z_w are computed from GPT2 derived meteorological data and the results are compared to site-VMF1 data. In the second step, the realization of GPT2-based mapping function is presented. The differences between rigorously ray-traced mapping function VMF1 and empirical mapping function GPT2 are analyzed.

37.3.2 Performance of GPT2 in Zenith Delay Modelling

The purpose of this assessment is to identify how the choice of empirical model affects the overall accuracy of zenith troposphere delay. For the empirical model, we evaluate the effect of Saastamoinen formula, in which the meteorological parameters are obtained from GPT2. The differences (in cm) for hydrostatic and wet zenith delay solutions are shown in Tables 37.1 and 37.2. As seen in Table 37.1, the modeling errors in Z_h are equal to −0.01 cm on average, and range from −1.93 cm at MCM4 up to 1.54 cm at NYAL. Although the average value of errors is nearly zero-mean, the standard deviations may reach up to 3 cm. It is

Table 37.2 Positions of the stations and wet zenith delay modeling differences (cm) with respect to VMF1 during 2012

Station	Ellipsoidal latitude (degree)	Ellipsoidal height (m)	DOY since 2012					Mean	STD
			001	091	181	271	361		
NYAL	78.93	78.57	0.13	1.18	−0.65	0.35	1.94	0.59	1.00
YELL	62.48	180.95	0.06	−4.69	−0.81	−2.52	−0.64	−1.72	1.91
WTZR	49.14	666.02	−6.27	0.32	−3.12	0.95	−1.74	−1.97	2.90
ALGO	45.96	200.93	−1.07	0.08	3.12	3.55	1.53	1.44	1.96
BJFS	39.61	87.47	−0.47	0.79	−5.97	6.66	0.14	0.23	4.49
KOKB	22.13	1167.36	11.77	12.22	10.66	14.70	4.78	10.83	3.69
RIOP	−1.65	2817.19	12.22	7.05	8.01	5.96	6.89	8.03	2.45
HARB	−25.89	1558.08	3.00	5.69	4.14	0.15	−2.64	2.07	3.32
YAR2	−29.05	241.29	0.91	5.07	0.95	4.07	−0.91	2.02	2.47
OHI2	−63.32	32.49	1.46	0.27	−0.81	2.10	−0.39	0.53	1.23
MCM4	−77.84	97.96	0.28	−0.89	0.10	−0.30	−0.63	−0.29	0.49
Mean			2.00	2.46	1.42	3.24	0.76	1.98	
STD			5.45	4.64	4.80	4.70	2.86		

because the empirical model is designed using 10 years NWM data. It represents the normal atmospheric conditions.

Compared with the hydrostatic delay, the wet delay component performs not so well. As seen in Table 37.2, the modeling errors in Z_w are equal to 1.98 cm on average, and range from −1.97 cm at WTZR up to 10.83 cm at KOKB. We should note that the wet delay component accounts for roughly 10 % of the total delay. The value is rather small with respect to the hydrostatic delay component. Even so, the means and standard deviations of differences with respect to VMF1 site solutions are larger. The results are consistent with actual situation, that the wet delay component is much more difficult to be evaluated using an empirical model. It depends on the amount of water vapor, which varies randomly and can not be easily modeled.

The GPT2-based model has been demonstrated extensively applied in several regions globally distributed, being capable of predicting total zenith delays with average uncertainties of about 2 cm under normal atmospheric conditions.

37.3.3 Performance of GPT2 in Mapping Function

We now assess the performance of mapping function named GPT2. The only differences between VMF1 and GPT2 mapping functions are the a coefficients used as input in (37.3). The bias between the two mapping functions has been multiplied by the zenith delay for each site to obtain the equivalent error in slant delay. Thus, the unitless mapping function bias could be converted to units of length. Here we use the VMF1 zenith delays from analysis data of ECMWF, which are considered accurate enough.

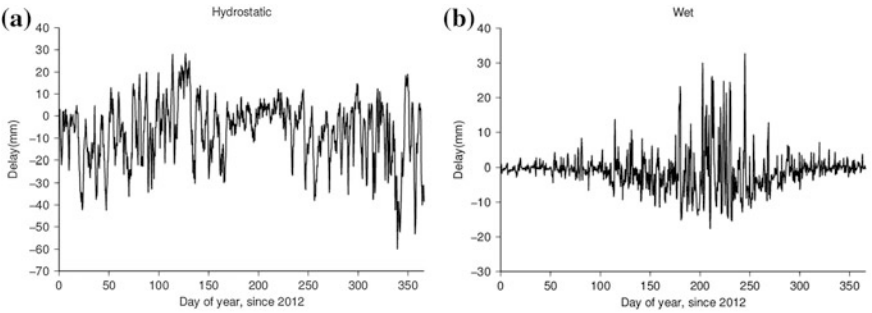


Fig. 37.2 Bias of the GPT2-based hydrostatic (*left*) and wet (*right*) mapping functions for the year 2012

Figure 37.2 shows the mapping function bias at 5° elevation angle. Both the GPT2-based hydrostatic and wet mapping functions are included for a single station BJFS during 2012. As VMF1 provides zenith delays and mapping functions at 0, 6, 12, and 18 UT everyday, the results are compared four times per day. It is possible to see that the use of GPT2 empirical coefficients would introduce a bias up to 50 mm at a low elevation angle for the hydrostatic delay. When it refers to the wet delay, as the delay value is rather smaller, the bias could keep within 10 mm most of the time.

Obviously, there is an abnormal manifestation from DOY 180 to DOY 250. During this period, the wet delay bias would increase to 30 mm. It could also be noticed that the difference of wet zenith delay is relatively large near DOY 200 in station BJFS (seen in Table 37.2). To explain this abnormal manifestation, there must be a special weather event. Through searching the internet, we find that a flash flood hit the city of Beijing on July 21, 2012. It is called the heaviest rain in six decades. And Fangshan District (where station BJFS is located in) was the most heavily-affected area. It received a record-setting 460 mm of rain. In that case, GPT2 would not be able to provide troposphere delay in high precision.

In this section, we evaluate GPT2-based troposphere delay using a globally distributed set of 11 IGS permanent stations data. And the empirical results are compared with VMF1. Since the empirical model is only determined by a mean value with an annual and semiannual variation, it could be applied in case the best accuracy is not required.

37.4 Effect of Troposphere Delay in Positioning

It is well known that the troposphere delay has a high correlation with the station height parameter for precise GNSS positioning. A rule-of-thumb proposed by MacMillan and Ma, also adopted by Boehm, said that the error in station height is equal to 1/5 the error in the slant delay at the lowest elevation angle [19, 20].

Table 37.3 Observation and parameters used in PPP processing

Parameters	Model and constraint
Observation	LC and PC with 30 s sampling rate and 5° elevation angle cutoff
Weight	Elevation-dependent weighting
Phase windup	Model correction [21]
Iono. delay	Ionosphere-free linear combination
Rec. clock	White noise
Tidal displace.	Solid earth, pole tide. ocean loading (IERS 2003)
Site coordinates	Static mode, treated as constants
Ambiguity	Estimated real-valued
Relativistic	Periodic Clock Corrections (−2*R*V/c)
PCO/PCV	IGS antenna model igs08_wwww.atx applied

To evaluate this effect on precise geodetic positioning, PPP is applied to several IGS permanent stations.

As mentioned in Sect. 37.3, the same 11 stations are selected. And 24 h 30 s GPS data in 17th June 2013 are collected from Crustal Dynamics Data Information System at NASA's Goddard Space Flight Center. At these stations, position coordinates and carrier phase ambiguities are estimated as constants, and receiver clock offsets are estimated as white noise at each epoch. Extended Kalman Filtering is used as the data processing method. A cutoff elevation angle of 5^{o} is adopted. In addition, the IGS final products are used, including satellite orbit and clock and Earth rotation parameters. Details in the processing are listed in Table 37.3, which are consistent to operational IGS data processing strategy [2, 21].

In the processing of troposphere delay, GPT2 validation is done in two ways. In the first way, GPT2-based model described in Sect. 37.2 is directly applied, including zenith delay modeling and mapping function. In the second way, as the wet zenith delay component is more difficult to model accurately, the residual is estimated as an additional unknown along with other parameters.

37.4.1 Directly Predicted Troposphere Delay Based on GPT2

We begin the assessment by using predicted GPT2-based troposphere delay directly in PPP processing. Commonly used wet troposphere delay parameter is eliminated from PPP solutions. We then compare the results with IGS solutions. It has been validated in the previous section, that GPT2 is capable of modeling the troposphere slant delay accurately under normal atmosphere conditions. Therefore, it was initially conceived that PPP processing without residual troposphere parameters could still perform well or at least not too bad.

The positioning results are unexpected. Limited by the length of this paper, positioning differences in 4 representative IGS stations with different latitudes are plotted. As seen in Fig. 37.3, a clear trend is present for the East and North

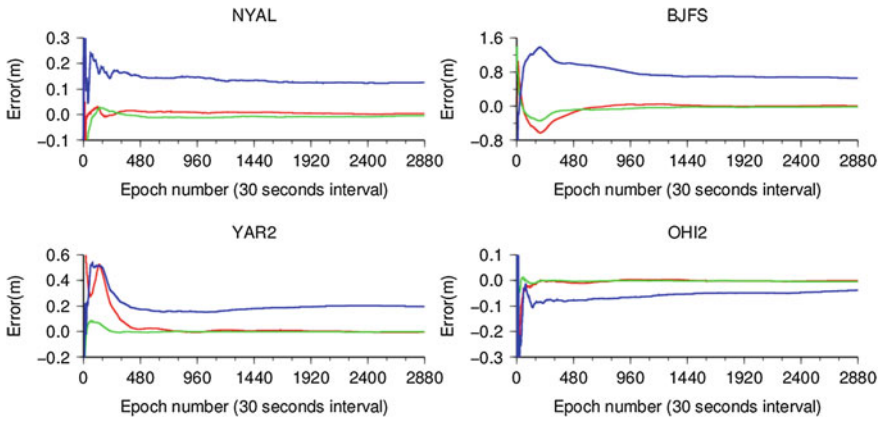


Fig. 37.3 Positioning performance for the East, North and Up components (Here the troposphere delays are taken as modeled values)

components and a bias for the Up component. The results show that, for a daily PPP solution, an accuracy of millimeter scale could be achieved in horizontal direction. But the height error is a little large, even as high as 0.7 m in station BJFS.

Overall, GPT2-based troposphere model is a spatially and temporally high-resolution troposphere model. It could largely eradicate the weaknesses of early empirical models. However, it is not yet available to use only a global model to eliminate the troposphere parameters in high precision GNSS data processing. It may cause a large height error due to the empirical troposphere model. The bias may even reach up to several decimeters.

37.4.2 Residual Estimated with GPT2 as a Priori Model

It has been validated that sufficient global model is not yet available. Existing models, even the latest GPT2 model, would introduce a height bias for troposphere model error. To correct this height positioning error, we add a residual parameter as an unknown to estimate along with other geodetic quantities. The same data processing strategy are adopted except for troposphere delay.

For comparison, the same observations are used as before and the positioning differences for East, North and Up components are plotted in Fig. 37.4. We can see that the height bias disappear at this time. All the positioning results could converge to centimeter-level accuracy within 1 h. After the 24 h positioning is totally finished, millimeter-level accuracy could be separately achieved in three directions. It should be noted that, about 2 cm positioning difference for the Up component at station BJFS still exists after 24 h. The reason has not been found. There may be something wrong in the data processing at station BJFS.

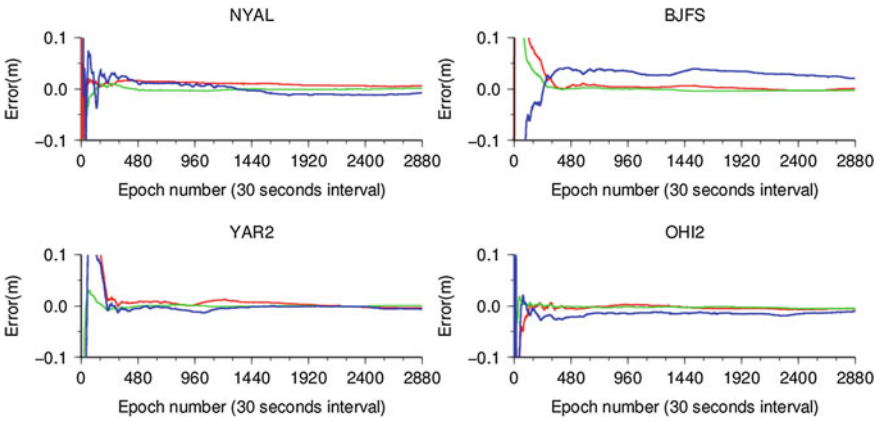


Fig. 37.4 Positioning performance for the East, North and Up components (Here the troposphere delays are taken as modeled a priori values with additional random walk residuals)

In this section, GPT2-based troposphere model is applied in PPP and the results are analyzed. It is not suitable for an empirical model, such as GPT2 model, to emulate troposphere delay variations exactly. Predominantly due to complex variations in the wet component of the troposphere delay as well as the unmodeled horizontal gradient component, a residual troposphere delay should be added as an unknown parameter for PPP.

37.5 Conclusions

This paper focuses on the analysis of troposphere delay based on global empirical models. In this study GPT2 model, which is the current state-of-the-art empirical troposphere model used for space geodetic techniques, has been assessed. A globally distributed set of IGS permanent station data is evaluated for both zenith delay modeling and mapping function. Additionally, GPT2-based troposphere model is applied in two different ways for PPP processing to identify the effect on height positioning accuracy. Based on daily observation data, the troposphere delay and positioning results are analyzed and compared. Through the measured data of reference stations the following conclusions can be drawn:

- GPT2-based troposphere model is simple formatted and easy to achieve. As an empirical model, it only requires station geodetic coordinates and the specific day of year. With the meteorological data and mapping functions coefficients obtained from GPT2, both hydrostatic and wet slant troposphere delay could be determined using relevant commonly used formulas. This approach is quite useful for practical use in the case of a priori troposphere delay is required.

- GPT2-based troposphere model could predict total zenith delays with average uncertainties of about 2 cm under normal atmospheric conditions. Compared with other empirical model, it could ensure that the troposphere parameter adjustments are more nearly zero-mean. This model is meaningful in case the best accuracy is not required, or in case measured in situ data or NWM data is not available.
- The model error in zenith direction would be magnified after multiplying a mapping function. And positioning accuracy for the Up component is directly affected by the error in slant delay. A direct application of GPT2-based troposphere model in geodetic positioning will lead to unsatisfactory height positioning accuracy. Hence, troposphere parameters can not be eliminated from space geodetic solutions so far. An additional residual troposphere parameter would bring significant precision improvements to compensate the model error. The significance of GPT2 lies in the availability of precise a priori troposphere delay, which could improve the efficiency of data processing.

References

1. Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102(B3):5005–5017
2. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solutions 5(2):12–28
3. Hu W, Gao C (2002) GPS measurement principles and applications, 1st edn. China Communications Press, Beijing, pp 96–97
4. Leick A (1995) GPS satellite surveying, 2nd edn. Wiley, New York 220
5. Xu G (2007) GPS: theory, algorithms, and applications, 2nd edn. Springer, Berlin
6. McCarthy DD, Petit G (2004) IERS conventions (2003). In: International earth rotation and reference systems service (IERS), Germany
7. Saastamoinen J (1972) Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. Geophys Monogr Ser 15:247–251
8. Davis JL, Herring TA, Shapiro II, Rogers AEE, Elgered G (1985) Geodesy by radio interferometry: effects of atmospheric modeling errors on estimates of baseline length. Radio Sci 20(6):1593–1607
9. Böhm J, Heinkelmann R, Schuh H (2007) Short note: a global model of pressure and temperature for geodetic applications. J Geodesy 81(10):679–683
10. Kouba J (2009) Testing of global pressure/temperature (GPT) model and global mapping function (GMF) in GPS analyses. J Geodesy 83(3–4):199–208
11. Lagler K, Schindelegger M, Böhm J, Krásná H, Nilsson T (2013) GPT2: empirical slant delay model for radio space geodetic techniques. Geophys Res Lett 40:1069–1073
12. Boehm J, Werl B, Schuh H (2006) Troposphere mapping functions for GPS and very long baseline interferometry from European centre for medium-range weather forecasts operational analysis data. J Geophys Res 111:B02406
13. Tregoning P, Herring TA (2006) Impact of a priori zenith hydrostatic delay errors on GPS estimates of station heights and zenith total delays. Geophys Res Lett 33:L23303
14. Penna N, Dodson A, Chen W (2001) Assessment of EGNOS tropospheric correction model. J Navig 54(1):37–55
15. Mendes VB, Prates G, Pavlis EC, Pavlis DE, Langley RB (2002) Improved mapping functions for atmospheric refraction correction in SLR. Geophys Res Lett 29(10):1414

16. Niell AE (1996) Global mapping functions for the atmosphere delay at radio wavelengths. J Geophys Res 101(B1):3227–3246
17. Böhm J, Niell A, Tregoning P, Schuh H (2006) Global mapping function (GMF): a new empirical mapping function based on numerical weather model data. Geophys Res Lett 33:L07304
18. Kouba J (2008) Implementation and testing of the gridded Vienna mapping function 1 (VMF1). J Geodesy 82(4–5):193–205
19. Urquhart L, Nievinski FG, Santos MC (2013) Assessment of troposphere mapping functions using three-dimensional ray-tracing. GPS Solutions 1–10. doi:10.1007/s10291-013-0334-8
20. Boehm J, Kouba J, Schuh H (2009) Forecast Vienna mapping functions 1 for real-time analysis of space geodetic observations. J Geodesy 83(5):397–401
21. Wu JT, Wu SC, Hajj GA (1993) Effects of antenna orientation on GPS carrier phase. Man Geodetica, 18:91–98

Chapter 38
Comparison of IRI-2012 and Rapid GIMs With GNSS-Derived TEC Over China

Yan Xiang, Yunbin Yuan and Ningbo Wang

Abstract The ionospheric delay is one of the predominant errors limiting the accuracy of Global Navigation Satellite Systems (GNSS), especially for single-frequency users. Meanwhile the IGS (International GNSS Service) Ionospheric Working Group and International Reference Ionosphere play an essential role in promoting ionospheric studies. This contribution will analyze the performance of IGS rapid Global Ionospheric Maps (GIMs) and the latest version International Reference Ionosphere model IRI-2012 aiming to provide more valid ionospheric correction in China. Three pairs of stations from CMONOC (Crustal Movement Observation Network of China) in 2013 are employed to analyze the discrepancy between rapid GIMs, IRI-2012 and GNSS-derived TEC. Results show that the rapid GIMs and IRI-2012 have distinct difference from GPS-derived TEC both in daytime and nighttime. In terms of rapid GIMs, the performance over IGS stations is generally better than those are non IGS stations, with pronouncedly better accuracy of about 20 and 50 % in daytime and nighttime respectively with respect to GNSS-derived TEC. Moreover, IRI-2012 is likely to overestimate TEC value in daytime, and underestimate TEC value in nighttime.

Keywords Rapid GIM · IRI-2012 · GNSS-derived TEC · Ionosphere

Y. Xiang (✉) · Y. Yuan · N. Wang
Institute of Geodesy and Geophysics, State key laboratory of Geodesy and Earth's dynamics, Chinese Academy of Sciences, Wuhan 430074, China
e-mail: xiangyan@asch.whigg.ac.cn

Y. Xiang · N. Wang
University of Chinese Academy of Sciences, Beijing 100049, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_38,

38.1 Introduction

Total electron content (TEC), a measure of the total number of electrons per m^2 along the line-of-sight between a receiver and a satellite transmitter, is a significant parameter of ionosphere. Ionospheric range correction is often related with TEC that is either positively or negatively proportional to the ionosphere correction according to code and phase respectively. In light of this, precise estimation of TEC is a key point to adjust ionospheric correction error.

Actually, there have existed several ionospheric models already over the last 50 years. They are generally categorized into three types [1]: (1) physical models based on ionospheric physics and chemistry. This is best illustrated by two examples like Time Dependent Ionospheric Model (TDIM) from Utah State University (USU) and the University Coupled Thermosphere—Ionosphere Model (CTIM) from University College London and Sheffield. While physical models require extensive computer resources to provide affordable run times and inputs derived from empirical models like Horizonal Wind Model (HWM); (2) semi-physical models that simplify the physical models by reducing the number of input parameters. The purpose of these models is to represent the ionosphere's spatial and temporal structure using a limited number of numerical coefficients, like Fully Analytical Ionospheric Model (FAIM) and the Parameterised Ionospheric Model (PIM); (3) empirical or semi-empirical models based on observations. The parameters are based on statistical systematic ionospheric variations from past data records usually regarding to monthly average parameters, including the Bent Model, the International Reference Ionosphere (IRI) and Global Ionospheric Maps (GIMs).

Over last decades, much attention has been devoted to ionospheric models comparison. As early as in 1991, Brown, L. D. et al. evaluated the performance of six ionospheric models as predictors of TEC, namely (1) the International Reference Ionosphere (IRI) (2) the Bent model (3) the Ionospheric Conductivity and Electron Density (ICED) model (4) the Penn State model (5) the Fully Analytic Ionospheric Model (FAIM, a modification of the Chiu model), and (6) the Damen-Hartranft model. They recommended utilizing an f0E2 from coefficients coupled with a new slab thickness developed at Boston University [2]. Komjathy A. et al. reported that both for day-time and night-time periods the IRI90 model appeared to be more accurate than the Broadcast model for low solar activity and mid-latitude conditions [3]. Validation with the external resources like Topex/Jason, digisonde and Incoherent Scatter Radar (ISR) has also been investigated by a number of authors [4–8]. Ho et al. [4] investigated global ionosphere mapping with Topex/pesion and climatological model (Bent model) and considered GIM as a new technique was able to monitor global ionospheric features in near real time. Mosert et al. [6] reported that the IRI predictions generally overestimate the TEC values by comparing with digisonde measurements.

However, rapid global ionospheric maps having been produced by IGS since 2003 [9, 10] and the latest modified version IRI-2012 [11] both drawn little attention from the ionospheric research community. After being released in 2003,

rapid global ionospheric maps with a latency of one day have achieved a similar accuracy with the final ionsopehric global ionospheric maps. Hence, users who desire an up-to-data ionospheric correction will benefit from rapid GIMs. Meanwhile IRI-2012 as an empirical model which depends on various data sets plays an important role in ionosphere prediction which is of great significance. In these respects, there is a need to further analyze and validate both models to provide more valid information to draw on.

In this study, an experimental examination of TEC from rapid GIMs and IRI-2012 with reference to TEC derived from dual-frequency GNSS observation data was done to analyze ionospheric correction effects of IRI-2012 and rapid GIMs, in an effort to serve as a reference for single-frequency GNSS receiver users to achieve better Positioning, Navigation, and Timing.

38.2 Methodology and Data

38.2.1 IRI Modeled TEC

Sponsored by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI), the International Reference Ionosphere (IRI) is a commonly used empirical model in the ionosphere community. It is an internationally recommended standard for the specification of plasma parameters in earth's ionosphere. Several steadily improved editions of the model have been released, and the latest version is IRI-2012. Once location and date are provided, IRI will offer monthly averages of the electron density, electron temperature, ion temperature, and ion composition in the altitude range from 50 to 1500 km [11]. Moreover, a great amount of available and reliable data sources for the ionospheric plasma is assimilated, such as a worldwide ionosonde network, radar data and satellite and rocket measurements. The software package was downloaded from SPDF (Space physics Data Facility) and CCMC (Community Coordinate Modeling Center).

38.2.2 Global Ionospheric Maps (GIMs) Derived TEC

GIMs have been routinely contributed by Ionospheric Associate Analysis Centers (IAACs), including CODE, ESA, JPL, and UPC. They are computing the global distribution of TEC independently with different mathematic approaches. Specifically, CODE and ESA developed the TEC into spherical harmonics expansion up to a degree and order of 15 with a 2 h resolution; linear composition of bi-cubic splines was used by JPL and three-dimension voxel with two-layer tomographic approach over individual station was used by UPC. Their products are categorized

into rapid (latency less than 24 h) and final (latency 11 days) schedules [12], with a similar level of accuracy according to Hernández-Pajares [12]. They are produced in the uniform IONEX format [13]. However, rapid products draw little attention from the ionospheric research community.

Concerning the practical usage of GIMs, E is represented as a function of geocentric latitude β, longitude λ, and universal time t. Interpolating between consecutive rotated TEC maps is a recommended method to achieve better effect. The methodology is formulated as (38.1)

$$E(\beta,\lambda,t)=\frac{T_{i+1}-t}{T_{i+1}-T_i}E_i\left(\beta,\lambda_i'\right)+\frac{t-T_i}{T_{i+1}-T_i}E_{i+1}\left(\beta,\lambda_{i+1}'\right) \tag{38.1}$$

where $T_i \le t \le T_{i+1}$, and$\lambda_i' = \lambda + (t - T_i)$.

38.2.3 TEC Derived From Dual-Frequency GNSS Data

The slant total electron content (STEC) of ionospheric pierce point (IPP) on radio path can be directly obtained by combining dual-frequency carrier phase and code-delay observations, the so-called leveling process. Once differential code biases (DCB) [14] are removed and mapping function is calibrated, precise IPP VTEC information can be acquired. The biases between them, which were defined it as internal accuracy, will reflect the performance of each ionospheric map. Besides, the accuracy of IPP TEC was discussed in.

The process of extracting the IPP VTEC involves three steps:

Step one:

Smooth code P_1 and P_2 by carrier phase, the so called "leveling carrier to code" algorithm [15]

Step two:

Model ionosphere TEC station by station to determine transmitting and receiving hardware biases [16] by subtraction of the simultaneous smoothed code as

$$\tilde{P}_{2,i}^k - \tilde{P}_{1,i}^k = \alpha\frac{sTEC}{f^2} + b_R + b_S + \varepsilon_P, \tag{38.2}$$

where α is the constant value to convert TECu (1TECU $= 10^{16}$ electrons/m^{-2}) to length units; f denotes the frequency of $\tilde{P}_{2,i}^k - \tilde{P}_{1,i}^k$; b_R and b_s refer to the receiver and satellite DCB (Different Code Biases) between P2 and P1 respectively. Note that they were treated as one unknown parameter here in single station model. In addition, general Triangle Function Series was used in the single station model [17]

Step three:

Table 38.1 Locations of three group stations

Station name	Longitude (°E)	Latitude (°N)	Geomagnetic latitude (°N)
KUNM	102.80	25.03	18.21
XIAM	118.08	24.44	17.49
LUZH	105.41	28.88	22.34
WUHN	114.36	30.53	24.00
BJFS	115.90	39.61	33.45
WUSH	79.21	41.20	35.89

Remove DCB computed in step 2 to get unbiased STEC at line-of-sight; Calibrate mapping function (a typical thin slab assumption) along the line of sight to acquire precise IPP VTEC at the zenith

$$vTEC = sTEC \cdot \cos z' = stec \cdot \sqrt{1 - \left(\frac{R}{R+H}\right)^2 \cdot \cos^2(z)}, \tag{38.3}$$

where R is the mean Earth's radius; H is the height of the thin layer; z and z' are the satellite zenith distance at the observation point and at the IPP respectively. Here the minimum elevation is 15° [18]

38.2.4 Data Set and Comparison Method

To analyze the performance of rapid GIMs and the latest vesion IRI-2012 over the IGS stations and non IGS stations, three station groups at three latitude bands are chosen to analyze in comparison with the TEC derived from dual-frequency GNSS receivers, namely KUNM versus XIAM, WUHN versus LUZH and BJFS versus WUSH as present in Table 38.1, among which KUNM, WUHN, and BJFS are IGS stations.

Data from COMNOC in the first month of 2013 is sampled to test the performance of modified edition IRI-2012 and rapid GIMs. To analyze the ionospheric correction effect in both daytime and nighttime respectively, the results have been split into 06:00–17:59 (day time) and 18:00–05:59 (night time).

For statistical analysis, the monthly ionospheric relative correction is defined as Eq. (38.4):

$$PER_n = \left(1 - \frac{|TEC_{\mathrm{model,n}} - TEC_{\mathrm{mean},n}|}{TEC_{\mathrm{mean},n}}\right) \times 100\,\%, \tag{38.4}$$

$$VTEC_{mean,n} = \frac{\sum_{S=1}^{S_n} P_S \cdot VTEC^S_{mean,n}}{\sum_{S=1}^{S_n} P_S}, \tag{38.5}$$

$$VTEC_{model,n} = \frac{\sum_{S=1}^{S_n} P_S \cdot VTEC^S_{model,n}}{\sum_{S=1}^{S_n} P_S}. \tag{38.6}$$

Here $P = \frac{1}{1+\cos^2(elev)}$; *elev* denotes the satellite elevation, since the accuracy of ionosphere estimation is limited by the mapping function [18].

38.3 Results

38.3.1 Overall Results

Figure 38.1 shows diurnal variations of TEC derived from dual-frequency observation (red dot), rapid GIMs (green dot) and IRI-2012 (blue dot) from a typical day of the 2nd, January, 2013 at the six stations. The results are ordered according to different latitude bands, namely KUNM versus XIAM, WUHN versus LUZH and BJFS versus WUSH. As can be clearly seen from Fig. 38.1, all of them generally produce a similar tendency that there is a gradual increase from the minimum TEC in the midnight then reaching the maximum around the midday before dropping to a stable value at night.

However, taking a further examination of the period from midnight to 6 a.m., IRI TEC values represented by blue dot is notably smaller than GNSS-derived TEC denoted by red dot. Conversely, it shows an opposite trend in the daytime from 8 a.m. to 18 p.m., i.e. the TEC estimation of IRI is almost twice much larger than the TEC derived from dual-frequency measurements in the daytime.

In terms of rapid GIMs, smaller biases from GNSS-derived TEC can be evidently observed comparing with IRI TEC, which means the consistency between rapid GIMs and GNSS-derived TEC is better than that of IRI-2012. Similarly, a trend is also revealed that TEC from rapid GIMs is still but marginally larger than GNSS-derived TEC in the daytime.

38.3.2 Rapid GIMs Result

The mean, std and relative percentage of biases between rapid GIMs TEC and dual-frequency TEC in the first month of 2013 is presented in Fig. 38.2. The

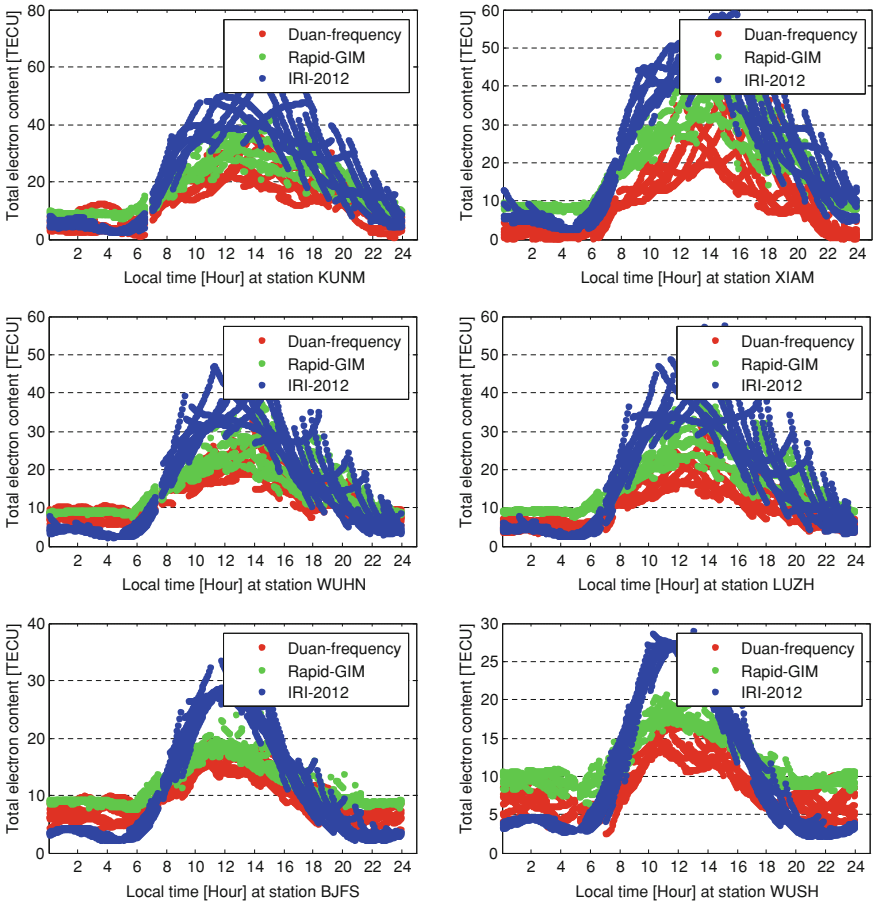


Fig. 38.1 Dural variation of TEC derived from dual-frequency observables, rapid GIMs and IRI-2012

stations in each plot are sorted by geomagnetic latitude increasing from left to right. It can be clearly seen that std decreases gradually with the increase of latitude both in daytime and nighttime. According to Fig. 38.2, concerning the mean TEC estimation, the value in daytime is generally smaller than that at night. For example, the average residual results for daytime in the range of $-1 \sim 5.0$ TECU, whereas that for nighttime within spreads ~ 6.0 TECU in spite of the fact that TEC value is much smaller in nighttime than that in daytime. It can also be evidently seen that the results of IGS stations are more reliable and valid than non IGS stations in terms of all three indexes. To be more specific, the average biases of non IGS station almost double in comparison with those of IGS stations both in daytime and nighttime. In terms of std, it is roughly 1TECU smaller of IGS

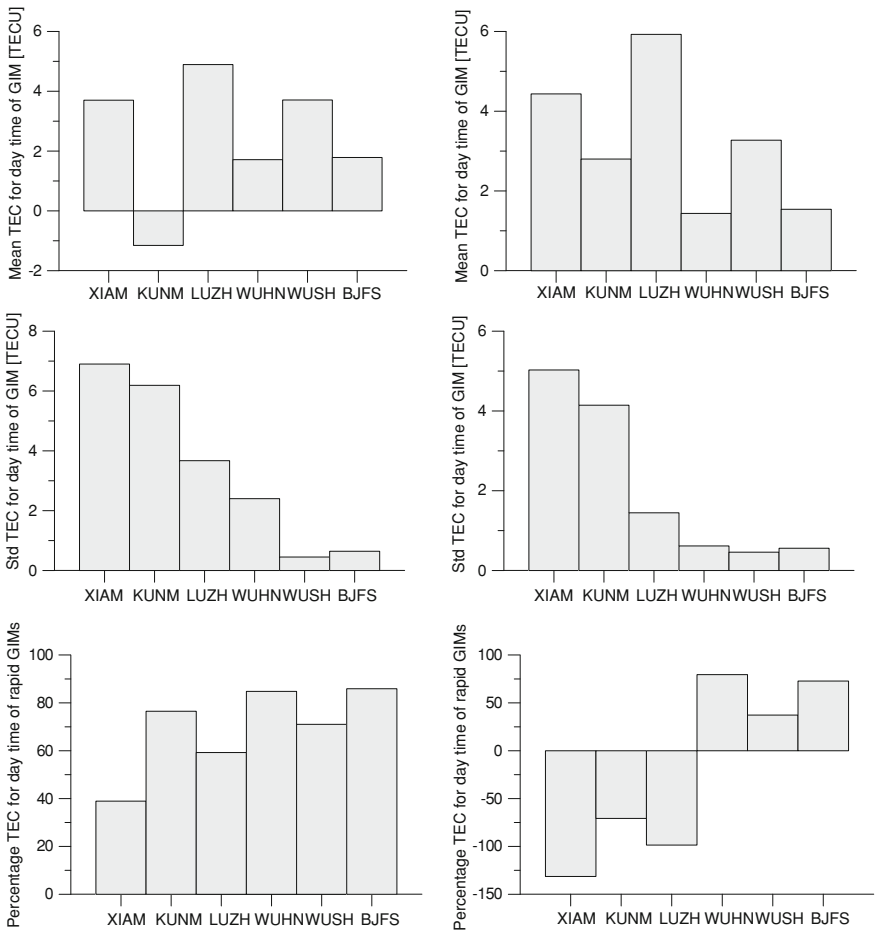


Fig. 38.2 Mean, std and relative percentage of rapid GIMs compared with dual-frequency TEC in daytime (*left*) and nighttime (*right*)

stations than non IGS stations except that of BJFS, which is slightly larger. Regarding to relative percentage, the mean ionospheric correction is approximately positive 70 %. Nevertheless, the correction in the nighttime is even negative over some stations, which means that the TEC estimations over those stations are twice larger than the GNSS-derived TEC.

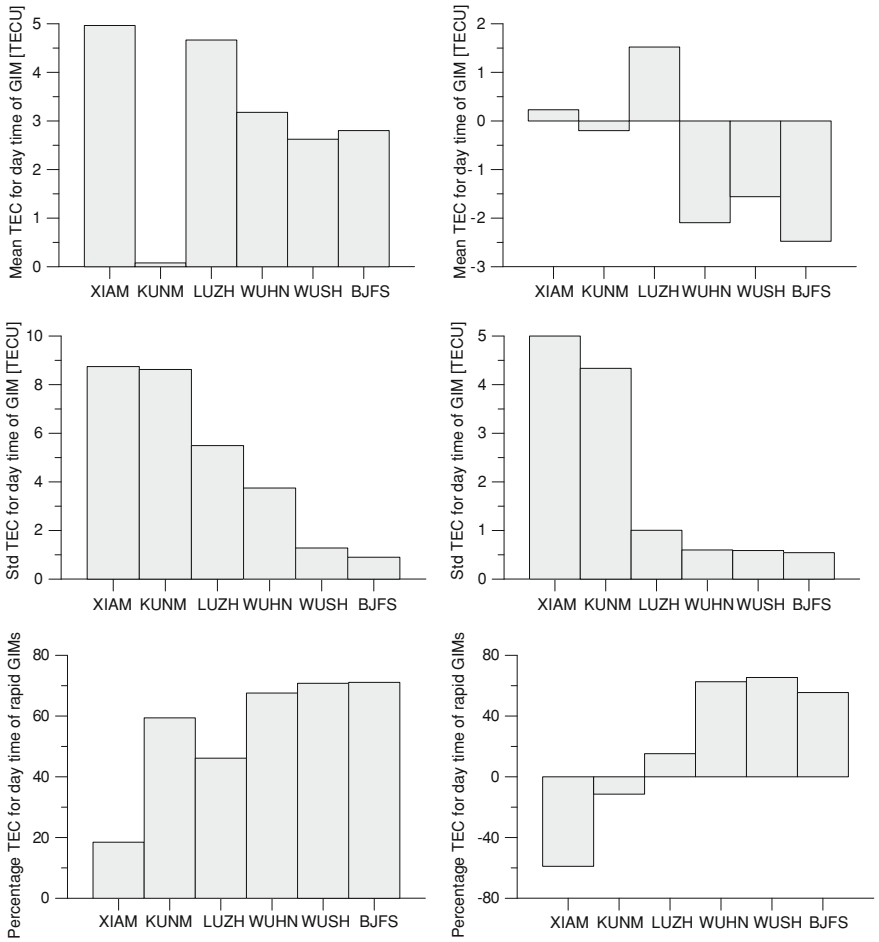


Fig. 38.3 Mean, std and relative percentage of IRI-2012 compared with dual-frequency TEC in daytime (*left*) and nighttime (*right*)

38.3.3 IRI-2012 Results

Different from GIMs which depend on the raw GNSS data, IRI shows no obvious distinction between IGS stations and non IGS stations. But there is one general similarity, according to Fig. 38.3, that the value of std decreases with the latitude bands due to the larger ionosphere TEC value in the lower latitude.

As is revealed from the mean TEC estimation of Fig. 38.2, it seems that IRI-2012 overestimates TEC value during daytime, and conversely underestimates TEC value because of positive biases in daytime and negative biases in nighttime by comparing with TEC derived from dual-frequency GNSS observation respectively. This point agrees with the analysis of Mosert et.al. [6] and Bhuyan and

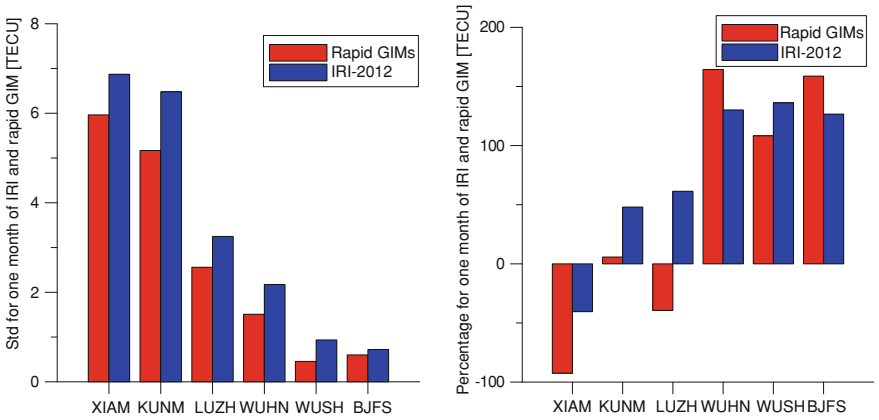


Fig. 38.4 Std and relative percentage of rapid GIMs and IRI-2012

Borah [19]. Concerning the relative percentage of ionospheric correction in daytime, it varies from roughly 20 to 70 %. By contrast, the relative percentage in nighttime is much worse than the daytime, even with negative percentage.

Overall, the std and relative percentage of rapid GIMs and IRI-2012 in comparison with GNSS-derived TEC is presented in Fig. 38.4. According to the figure, the std value is smaller than that of IRI-2012, thus it is generally considered GIM is better than IRI-2012. While in terms of relative percentage, rapid GIMs have better performance than IRI-2012 in the daytime according to Figs. 38.2 and 38.4. In light of this, it is difficult to determine which one is better in the nighttime, i.e., sometimes rapid GIMs better and sometimes IRI-2012 better. Therefore, the overall trend of relative percentage is uncertain.

38.4 Conclusion

In this work, the less well-known rapid GIMs and latest modified version IRI-2012 are investigated by comparing with the TEC derived from the dual-frequency GNSS observation in an effort to promote the wide application and serve as a reference for those who need ionospheric correction or as initial values for ionospheric tomography. The experimental data set samples are chosen from COMNOC at six stations in the first month of the year of 2013.

Results show that there is a pronounced discrepancy between rapid GIMs and IRI-2012. As far as rapid GIMs are concerned, IGS stations perform better than those are non IGS stations with better relative percentage of about 20% and 50% in the daytime and night time respectively. That is to say, the accuracy would be improved by more measurements. Nowadays, the observations are only obtained from terrestrial GNSS measurements and these reference stations are not only

globally distributed, especially over the ocean. Provided GPS occultation measurements and altimetry data are assimilated into the model, it will make a difference.

Moreover, IRI-2012 tends to overestimate the TEC in the daytime and underestimate TEC in the nighttime. But it would serve as an excellent initialization for ionospheric tomography with three dimension electron density. Generally, the std between TEC from rapid GIMs and TEC derived from dual-frequency raw GNSS data is smaller than that of IRI-2012. However, it is difficult to distinguish which one is better if relative percentage is regarded as a criterion from the analyzed data set.

Although this study used a limited database, it shows that it is possible to reflect the overall characteristic of each model by comparing with real-time dual-frequency GNSS measurements.

Acknowledgments We thank CDDIS (Crustal Dynamics Data Information System), IGS and COMNOC to allow us to use their data set to finish our research. This research was supported by National 973 (No. 2012CB825604) and 863 programs (No. 2012AA121803), China Natural Science Funds (No. 41304034, 41231064, 41104012 and 41021003), and the CAS/SAFEA International Partnership Program for Creative Research Teams (KZZD-EW-TZ-05).

References

1. Feltens J et al (2011) Comparative testing of four ionospheric models driven with GPS measurements. Radio Sci 46(RS0D12):1–11
2. Brown LD, Daniell RE, Fox MW, Klobuchar JA, Doherty PH (1991) Evaluation of six ionospheric models as predictors of total electron content. Radio Sci 26:1007
3. Komjathy A, Langley R (1996) Improvement of a global ionospheric model to provide ionospheric range error corrections for single-frequency GPS users. In: Presented at the ION 52nd annual meeting, vol. 19, Cambridge, MA, pp 21
4. Ho C, Wilson B, Mannucci A, Lindqwister U, Yuan D (1997) A comparative study of ionospheric total electron content measurements using global ionospheric maps of GPS, TOPEX radar, and the Bent model. Radio Sci 32:1499
5. Zhang M-L, Radicella SM, Shi J-K, Wang X, Wu S-Z (2006) Comparison among IRI, GPS-IGS and ionogram-derived total electron contents. Adv Space Res 37:972
6. Mosert M, Gende M, Brunini C, Ezquer R, Altadill D (2007) Comparisons of IRI TEC predictions with GPS and digisonde measurements at Ebro. Adv Space Res 39:841
7. Jee G, Lee HB, Kim YH, Chung JK, Cho J (2010) Assessment of GPS global ionosphere maps (GIM) by comparison between CODE GIM and TOPEX/Jason TEC data: ionospheric perspective. J Geophys Res Space 115(A10):A10319. doi:10.1029/2010ja015432
8. Makarevich RA, Nicolls MJ (2013) Statistical comparison of TEC derived from GPS and ISR observations at high latitudes. Radio Sci 48:441
9. Komjathy A, Hernández-Pajares M (2004) The IGS global TEC maps: present and future. In: National Radio Science Meeting, Boulder, CO, USA
10. Hernandez-Pajares M, et al (2009) The IGS VTEC maps: a reliable source of ionospheric information since 1998. J Geodesy 83:263
11. Bilitza D, McKinnell L-A, Reinisch B, Fuller-Rowell T (2011) The international reference ionosphere today and in the future. J Geodesy 85:909
12. Hernández-Pajares M (2004) IGS Ionosphere WG: an overview. In: Proceeding COST, pp 29–29

13. Schaer S (1999) Mapping and predicting the Earth's ionosphere using the global positioning system, PhD Thesis, Astronomical Institute, University of Berne
14. Yuan Y, Ou J (1999) The effects of instrumental bias in GPS observations on determining ionospheric delays and themethods of its calibration. Acta Geod Cartogr Sin 28:110
15. Dach R, Hugentobler U, Fridez P, Meindl M (2007) Bernese GPS software version 5.0. Astronomical Institute, University of Bern p 640
16. Li Z, Yuan Y, Fan L, Huo X, Hsu H (2013) Determination of the differential code bias for current BDS satellites.Geoscience and Remote Sensing, IEEE Transactions 52(7):3968–3979
17. Yuan Y, Ou J (2004) A generalized trigonometric series function model for determining ionospheric delay. Prog Nat Sci 14:1010
18. Komjathy A, Langley RB (1996) An assessment of predicted and measured ionospheric total electron content using a regional GPS Network. In: The Proceedings of the national technical meeting of the Institute of Navigation, pp 615–624
19. Bhuyan P, Borah RR (2007) TEC derived from GPS network in India and comparison with the IRI. Adv Space Res 39:830

Chapter 39
Mitigation of Ionospheric Delay in GPS/BDS Single Frequency PPP: Assessment and Application

Zishen Li, Lei Fan, Yunbin Yuan, Sandra Verhagen, Peter de Bakker, Hong Yuan and Shiming Zhong

Abstract Single-frequency (SF) Precise Point Positioning (PPP) is a promising technique for real-time positioning and navigation at sub-meter (about 0.5 m) accuracy level because of its convenience and low cost. With satellite orbit and clock error being greatly mitigated by the precise products from the International GNSS Service (IGS), ionospheric delay becomes the bottleneck of SF PPP users. There are five commonly used approaches to mitigate ionospheric delay in SF PPP: (1) broadcast ionospheric model in Global Navigation Satellite System (GNSS) navigation message; (2) global ionospheric map released by the IGS; (3) local ionospheric model generated using GNSS data from surrounding reference stations; (4) satellite based ionospheric model; (5) the parameter estimation method. Those approaches are briefly reviewed in our contribution and the performances of some classical ionospheric approaches for SF PPP are validated and compared using GPS data from two networks in China and the Netherlands respectively. Validation results show that a set of reference stations is critical for SF PPP with sub-meter positioning accuracy, especially in China. It is better to model the ionospheric delay in a satellite by satellite mode rather than an integral mode under the assumption of a thin-layer ionosphere. Comparing to GIM, the suggested approach, satellite based ionospheric model (SIM), can improve the horizontal positioning accuracy of SF PPP from 0.40 to 0.10 m in China and from 0.20 to 0.05 m in the Netherlands, while it can improve the vertical accuracy from 0.70 to 0.15 m (China) and from 0.20 to 0.10 m (the Netherlands). Furthermore, the recommended ionospheric model has been applied to GPS/BDS data for SF

Z. Li (✉) · H. Yuan
Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing 100094, China
e-mail: lizishen@aoe.ac.cn

Z. Li · L. Fan · Y. Yuan · S. Zhong
State Key Laboratory of Geodesy and Earth's Dynamics, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan 430074, China

S. Verhagen · P. de Bakker
Delft University of Technology, Delft 2600 GA, The Netherlands

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2014 Proceedings: Volume II*, Lecture Notes in Electrical Engineering 304,
DOI: 10.1007/978-3-642-54743-0_39,

PPP as well. The experiment in Beijing shows that the positioning of about 0.5 m accuracy can be achieved by single epoch SF PPP based on a reference network of about 40 km inter-station distance. The accuracy of SF PPP based on an accumulation of 10–15 min of observations in dynamic mode is about 0.04 m (horizontal) and 0.04–0.08 m (vertical) using only GPS data, while it is about 0.03 m (horizontal) and 0.03–0.06 m (vertical) by combining GPS and BDS data.

Keywords Ionospheric delay mitigation · SIM · Single frequency precise point positioning · GPS · BDS

39.1 Introduction

There is a tremendous demand for real-time positioning with sub-meter (about 0.5 m) accuracy in the modern urban management, such as traffic guidance, urban planning, emergency rescue etc. PPP is a very efficient and convenient GNSS (Global Navigation Satellite System) positioning approach since it only relies on a GNSS receiver and correction data from a reference network. The PPP technique, which was firstly introduced by Zumberge et al. [44], can achieve decimeter to centimeter level positioning accuracy by applying various corrections, such as satellite orbit, satellite clock, ionospheric delay, and tropospheric delay [3, 20, 21]. In order to mitigate the ionospheric delay, dual-frequency signals are usually required in the traditional PPP technique. However, a dual-frequency receiver is too expensive for many applications and the high cost is currently one of the major barriers for the PPP technique in many common applications. Actually, reasonable accuracy at a low cost is preferred by many potential PPP users, e.g. a positioning with about 0.5 m accuracy is usually sufficient for the lane identification [37]. Therefore, PPP technique with single frequency receiver can be a perfect solution balancing the positioning accuracy and acceptable cost.

Ionosphere delay is one of the major challenges of single frequency (SF) PPP, because it cannot be eliminated without dual-frequency signals. SF PPP can be expected to perform well only when the ionospheric delay is accurately mitigated. However, the variations and characteristics of ionospheric delay are usually difficult to be modeled. Numerous approaches for mitigating the ionospheric delay of SF PPP have been studied in previous literature [1, 4, 23, 28, 33, 34]. Generally, these approaches can be divided into five levels based on the accuracy and implement method.

(1) Broadcast Ionospheric Model (BIM): this model is distributed along satellite ephemeris in the navigation message, such as GPS Klobuchar model [17, 18], BDS Klobuchar-like model [2, 38]. Due to the simplicity of model and

limitation of updating interval, however, BIM can only achieve the correction of about 0.5–1.2 m and 2.0–2.6 m on the zenith direction at low and high ionospheric activities respectively [8, 38, 40]. Therefore, BIM is generally not sufficiently accurate for SF PPP to achieve sub-meter level positioning.

(2) Global Ionospheric Map (GIM): it is one of the most popular ionospheric products for SF PPP. The ionospheric Total Electron Content (TEC) values are represented on a global scale grid and updates every two hours [31]. Currently, the IGS released GIM is a combined products from four ionospheric associate analysis centres [15]. The nominal accuracy of GIM is about 0.30–0.80 m in the zenith direction on average, but it is even lower in areas with fewer contributing GNSS stations [29, 46].

(3) Local Ionospheric Map (LIM): it characterizes vertical ionospheric delay over a small region with dual-frequency GNSS data from local reference stations. In contrast to GIM, the local real-time GNSS data contributes to the LIM estimation and the ionospheric delay provided by LIM is usually of a better accuracy. How to select a mathematic function to represent the variation in local ionospheric delay is one of the most critical issues for LIM. Many functions have been studied, including polynomial function [6], triangle series function [12, 39] (adjusted) low order spherical harmonic function [32, 46], and spherical cap harmonic function [26]. However, the LIM is generally established under the assumption of ionospheric thin-layer, a so-called mapping function is required for converting ionospheric delay from the line-of-sight (LOS) to vertical direction. The lower the satellite elevation, the larger the error resulting from the mapping function [47].

(4) Satellite based Ionospheric Map (SIM): SIM is a new regional ionospheric delay modeling method aiming to provide high accuracy ionospheric delay. In SIM, the LOS ionospheric delay at the rover station is directly derived from the corresponding observation of surrounding reference stations without a thin-layer assumption and mapping function. It has been used to reduce dual-frequency PPP convergence time and mitigate the ionospheric delay in SF PPP [11, 28, 41, 42, 45]. This method is more straightforward; the difference of satellite elevations between different stations is not considered any more. Therefore, this approach is more suitable for regional ionosphere modeling.

(5) Parameterized Ionospheric Model (PIM): Different from the aforementioned four ionospheric modeling methods, the LOS ionospheric delay in PIM is modeled by a number of unknown parameters and estimated simultaneously with positioning, e.g. the vertical delay and two gradient components [24, 34], a time variant parameter for each satellite (GRAPHIC, Group And Phase Ionospheric Correction) [5] etc. However, prior ionosphere information is still essential for this method to improve the parameter estimation. Otherwise, the long convergence time will become unacceptable for real-time or near real-time users.

According to the above analysis, a precise ionospheric model is very essential for the sub-meter level SF PPP. In view of this, we will focus on the mitigation method of ionospheric delays for a SF PPP user, including some assessments and

applications. The following parts of this paper are organized as follows: Sect. 39.2 briefly reviews some commonly used approaches applicable for SF PPP user; Sect. 39.3 validates their performances using real GPS data from two networks in both China and the Netherlands, and provides some useful suggestions on ionospheric delay correction for SF PPP users; Sect. 39.4 attempts to apply the recommend approach identified in Sect. 39.3 in SF PPP using a preliminary GPS/BDS dataset; finally, Sect. 39.5 summarizes some major findings in this paper and future work.

39.2 Review of Some Classical Ionospheric Models for SF PPP User

The purpose of this contribution is to find a preferred ionospheric model for local SF PPP by comparing the performance of commonly used approaches. Before this comparison, some classical ionospheric models for SF PPP users is briefly reviewed in this section, including global ionospheric map, local ionospheric model and satellite based ionospheric model.

39.2.1 Global Ionospheric Map-GIM

The GIM defined in the IONosphere map EXchange format (IONEX) is a public product released by IGS and generated by combining the daily GIMs from four ionospheric associate analysis centres (IAACs), including Center for Orbit Determination in Europe (CODE; University of Berne, Switzerland), Jet Propulsion Laboratory (JPL; Pasadena, California, USA), European Space Operations Center of European Space Agency (ESOC; Darmstadt, Germany), and Technical University of Catalonia/gAGE (UPC; Barcelona, Spain) [14]. The GIM from each IAAC is evaluated by comparing the differences of TEC from itself and real GNSS data and a weight is calculated for the final combination based on these differences [16].

Different strategies have been adopted for ionospheric TEC modeling by those IAACs. The global ionospheric vertical TEC (vTEC) over a single day is represented by a series of spherical harmonic expansions up to degree and order of 15 with a 2-h temporal resolution by CODE and ESOC, whereas it is represented by a linear composition of bi-cubic splines with 1,280 spherical triangles and 15 min resolution by JPL [9, 10, 19, 32]. The current ionospheric TEC models used by CODE, JPL, and ESOC are all based on the ionospheric thin-layer assumption, and a mapping function is necessary for converting the ionospheric TEC from LOS to vertical direction. Different from those approaches used by CODE, JPL and ESOC, UPC GIM is produced by interpolating the vTEC over each ionospheric IPP. The vTEC is computed by means of a two-layer (450 and 1130 km) tomographic approach over each individual station [13, 30]. The LOS ionospheric TEC is

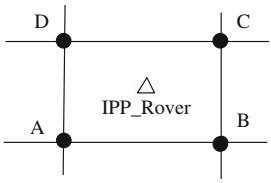


Fig. 39.1 Interpolation of rover LOS ionospheric delay using the surrounding 4 TEC values

extracted from dual-frequency data by the geometry free combination of phase-smoothed code. The differential code biases (DCB) in satellites and contributing receivers is estimated simultaneously with the global ionospheric TEC modeling.

With GIM, the TEC values are defined on a 5° (longitude) × 2.5° (latitude) grid with a 2-h temporal resolution. In order to apply GIM to SF PPP, the vTEC at one satellite IPP for the rover needs to be firstly interpolated between two consecutive maps and then mapped to the LOS ionospheric delay [31]. The interpolation and mapping method is shown by Fig. 39.1 and Eq. (39.1). The solid dots in Fig. 39.1 are four surrounding grid points defined in GIM and the triangle is the location of rover ionospheric IPP.

$$\begin{cases} I_{r,t} = VTEC_{r,t} \cdot mf(z) \cdot A \\ VTEC_{r,t} = \dfrac{t_2 - t}{t_2 - t_1} VTEC_{r,t_1} + \dfrac{t - t_1}{t_2 - t_1} VTEC_{r,t_2} \\ VTEC_{r,t_i} = (1-p) \cdot (1-q) \cdot VTEC_{A,t_i} + p \cdot (1-q) \cdot VTEC_{B,t_i} \\ \qquad\qquad + p \cdot q \cdot VTEC_{C,t_i} + q \cdot (1-p) \cdot VTEC_{D,t_i} \quad (i = 1,2) \\ p = \dfrac{\varphi_r - \varphi_A}{d\varphi}, q = \dfrac{\lambda_r - \lambda_A}{d\lambda} \end{cases} \tag{39.1}$$

where $I_{r,t}$ is the predicted ionospheric delay of one satellite at t for rover r; $VTEC_{r,t}$ is the vTEC for the rover interpolated from the GIM; $mf(z)$ is the mapping function defined as $mf(z) = (1 - \sin^2 z)^{-2}$, where z is the satellite's zenith distance at corresponding IPP; A is a constant value used to convert TEC unit to length unit, defined as $A = 40.28 \cdot 10^{16} \cdot f^{-2}$; f is the frequency on which the code and phase are used for SF PPP; t_1 and t_2 are the nearest two times at which the map is selected for interpolation, assuming $t_2 \geq t \geq t_1$; $VTEC_{r,t_1}$ and $VTEC_{r,t_2}$ are the corresponding vTEC obtained from the ionospheric map at t_1 and t_2, respectively; $VTEC_{A,t_i}$, $VTEC_{B,t_i}$, $VTEC_{C,t_i}$ and $VTEC_{D,t_i}$ are the ionospheric vTEC at grid point A, B, C and D respectively; φ_r and λ_r are the geographic latitude and solar longitude of ionospheric IPP at rover; λ_A and φ_A are the geographic latitude and solar longitude of grid point A; $d\varphi$ and $d\lambda$ are the interval of latitude and longitude in GIM, $d\varphi = 2.5°$ and $d\lambda = 5°$ for IGS released GIM; The ionospheric delay for each satellite can be individually predicted using Eq. (39.1).

The final and predicted GIMs are all released by IGS. The latency of final product is about 2 weeks and thus cannot be used in real-time or near-real-time

application. The predicted product is released 1–4 days ahead and is feasible for real-time SF PPP. With the IGS GIM product, the users are able to acquire precise ionospheric delay without setting up new reference stations for monitoring the ionosphere. In our experiment, GIM refers the IGS predicted product.

39.2.2 Local Ionospheric Model-LIM

A small-scale network equipped with dual-frequency receivers is also capable to provide precise ionospheric delay. Those dual-frequency receivers are also named reference stations. LIM is established using the LOS ionospheric TEC extracted from the raw data of reference stations with DCB correction. Generally, the variation in ionospheric vTEC over a small area (e.g. 20–80 km) is very smooth and easy to be represented by the polynomial. The mapping function is also necessary in LIM. The coefficients of polynomial ionospheric model are broadcasted to the rover along with their variances. After receiving the coefficients and its RMS of polynomial based ionospheric model, the rovers can calculate their own ionospheric delay and variances for each satellite. The polynomial ionospheric model is generally described as [6]:

$$I_{r,t} = A \cdot mf(z) \cdot \sum_{n=0}^{N} \sum_{m=0}^{M} E_{nm,t} \left(\beta_{r,t} - \beta_0\right)^n \left(s_{r,t} - s_0\right)^m \tag{39.2}$$

where, $I_{r,t}$ is the predicted ionospheric delay of one satellite at epoch t for rover r; $\beta_{r,t}$, $s_{r,t}$ are the latitude and longitude of rover ionospheric IPP, respectively; β_0 and s_0 are the latitude and longitude of the geometric center of polynomial model; N and M are the maximum orders of the polynomial model in terms of latitude and longitude respectively and E_{nm} represents the unknown coefficients of polynomial model estimated using the data from reference network. In our experiment, a second order polynomial model is adopted, i.e. $N = M = 2$.

Based on LIM, the real-time data from reference stations is contributed to local ionospheric modeling; thus, the accuracy of ionospheric delay from LIM is usually higher than that of GIM. In order to improve the temporal resolution of ionospheric modeling, the LIM is updated every epoch in our experiment, but the latency of data transmission is not considered.

39.2.3 Satellite based Ionospheric Model-SIM

Both GIM and LIM assume that the LOS ionospheric TEC is concentrated on a shell of infinitesimal thickness and a mapping function is commonly used to convert the LOS ionospheric TEC to vTEC. The mapping function is defined as an approximate trigonometric function, e.g. single-layer mapping function (SLM), or

modified SLM etc. [22, 27, 32]. However, the variation of LOS ionospheric TEC is very complex and the distribution of ionospheric density cannot be completely described by vTEC plus mapping function. Moreover, the mapping function cannot reflect the azimuth-dependent variations in ionospheric TEC. Previous study demonstrated that the modeling error resulted by the ionospheric thin-layer assumption and mapping function is about 0.05–0.20 m at different levels of ionospheric activities [7, 46, 47]. Thus, the modeling error cannot be acceptable for the ionospheric delay correction of 0.10–0.20 m accuracy level.

In order to reduce the modeling error, a satellite based ionospheric model (SIM) is developed using regional reference stations. It represents the LOS ionospheric delay on a satellite basis rather than ionospheric vTEC with mapping function. The SIM has been used by Geng et al. [11] and Zhang et al. [41] for estimating the ionospheric delay with a relatively high precision to reduce the convergence time of dual-frequency PPP. Different modeling/interpolation methods can be used in SIM, e.g. inter-stations weighted average and Kriging interpolation [28, 42, 43]. In this contribution, the ionospheric delay for the rover is interpolated according to the relative geographical locations of reference stations and rover, shown by Eq. (39.3).

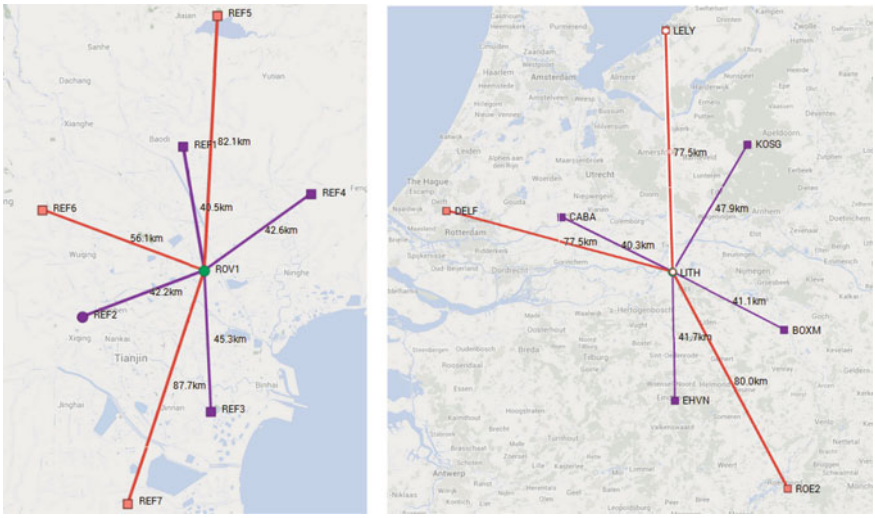
$$\begin{cases} I_{r,t}^{s} = \sum_{i=1}^{N} P_i^s I_{i,t}^s \\ P_i^s = \left(D_{i,r} \cdot \sum_{m=1}^{N} D_{m,r}^{-1} \right) \end{cases} \tag{39.3}$$

where $I_{r,t}^{s}$ is the LOS ionospheric delay correction of satellite s to rover r at time t; N is the number of surrounding reference stations; $I_{i,t}^{s}$ is the LOS ionospheric delay of satellite s to station i at epoch t; $D_{i,r}$ is the spherical distance between rover and station i; $D_{m,r}$ is the spherical distance between rover and station m.

It can be found that the SIM is based on the approximation that the elevations of one satellite at those reference and rover stations are considered the same. Thus, although SIM can avoid the modelling error caused by the mapping function, a new modelling error resulting from this approximation is introduced. Generally, the approximation is very reasonable for small-scale network, but the error caused by the approximation will become larger with the increasing of inter-station distance. So, SIM is very suitable to predict precise ionospheric delays for a rover from a small-scale network.

39.3 Assessment and Comparison of Different Ionospheric Delay Mitigation Models in SF PPP

The accuracy of ionospheric delay correction is the major barrier hindering SF PPP centimeter accuracy positioning, currently only achievable using a dual-frequency PPP technique [21, 23]. In this section, the aforementioned approaches are applied



(Latitude: 38.5^0N-40.0^0N) (Latitude: 50.5^0N-52.5^0N)

Fig. 39.2 Distribution of networks in China (*left*) and the Netherlands (*right*) in the experiment

to SF PPP data processing and their performances are assessed and compared. Some useful suggestions about ionospheric delay mitigation in SF PPP are summarized from their performance comparison between different approaches.

39.3.1 Description of GPS Data

Two GPS networks in China and the Netherlands are selected for this experiment respectively. Each network involves eight stations, the station in the center is considered as the rover and the other surrounding stations are used as reference stations. The locations and the distance between the rover and the reference stations are marked in Fig. 39.2. Both networks are located in middle latitudes (38°N–53°N), and the global ionospheric activities are at a medium level during the experimental period.

All the reference stations are divided into two groups according to the distance to the rover station. The distances in different groups are about 40 and 80 km respectively, and they are very typical for GNSS application in urban area. Using two sub-networks can reflect the variations of ionospheric modeling in different inter-station distances scenarios.

The type of antenna and receiver of selected stations at each station are listed in Table 39.1. The network in China is equipped with exactly the same antennas and receivers, whereas the network in the Netherlands uses mixed-type antennas and receivers. The experimental data was collected from 13th to 15th, Oct. 2012 with an

Table 39.1 Antenna and receiver type of receivers used for experiment

Networks	Station name	Antenna type	Receiver type	Role in validation
China	ROV1	TRM59800.00	TRIMBLE NETR8	Rover
	REF1	TRM59800.00	TRIMBLE NETR8	40 km inter-station distance of network
	REF2	TRM59800.00	TRIMBLE NETR8	
	REF3	TRM59800.00	TRIMBLE NETR8	
	REF4	TRM59800.00	TRIMBLE NETR8	
	REF5	TRM59800.00	TRIMBLE NETR8	80 km inter-station distance of network
	REF6	TRM59800.00	TRIMBLE NETR8	
	REF7	TRM59800.00	TRIMBLE NETR8	
The Netherlands	LITH	S2SPG_A1+M	Topcon ODYSSEY_E	Rover
	CABA	ASH700936C_M	ASHTECH Z-XII3	40 km inter-station distance of network
	EHVN	S2SPG_A1+M	Topcon ODYSSEY_E	
	KOSG	AOAD/M_B	LEICA GRX1200GGPRO	
	BOXM	S2X3G+C+M	Topcon ODYSSEY_E	
	DELF	TRM29659.00	TRIMBLE 4700	80 km inter-station distance of network
	LELY	S2SPG_A1+M	Topcon ODYSSEY_E	
	ROE2	S2SCR.G3+M2	Topcon ODYSSEY_E	

interval of 30 s. The software of SF PPP is developed based on RTKLIB which is an open sources program package for GNSS Positioning (http://www.rtklib.com) [35, 36]. The experiment is carried out in a simulated real-time mode and the processing steps of ionospheric delay correction at each epoch are the following:

(1) Determine the LOS ionospheric TEC and its variance for each pair of satellite and reference receiver with dual-frequency PPP method [48]. The precise satellite orbit and clock from CODE, the satellite and receiver DCB estimated using IGGDCB is adopted in the determination of LOS ionospheric TEC. The IGGDCB is a two-step method for DCB determination, which can work well with only a few ground tracking stations [25]. In practice, the DCB product can be obtained in the previous day because DCB is relatively stable over days.
(2) Predict the LOS ionospheric delay for each pair of satellite and rover receiver based on the approach of LIM and SIM respectively. The variance of corresponding ionospheric delay is also estimated based on the law of error propagation. In this step, the GIM is also introduced to calculate the LOS ionospheric delay for rover based on the method described in Sect. 39.2.1.
(3) Positioning based on SF PPP with the ionospheric delay predicted by different approaches in step (2). The predicted ionospheric delay is considered as a pseudo-observation with an estimated variance in SF PPP for the mitigation. Only the code and phase on frequency L1 from rover receiver is used and the satellite orbit, satellite clock and DCB are corrected by the product from

CODE and IGGDCB. The elevation-dependent weight is applied to the raw observation and the cut-off elevation is 10^0. Although the rover is static, the positioning is processed in a dynamic mode with Kalman filter.

(4) Repeating steps (1)–(3), the rover data could be processed epoch by epoch.

39.3.2 Accuracy of Predicted Ionospheric Delay for Rover

The LOS ionospheric delay extracted from the dual-frequency data of rover receiver can be taken as 'true' ionospheric delay and used to validate the accuracy of predicted ionospheric delay from GIM, LIM and SIM. The root mean square (RMS) is used as a measure of accuracy, which is calculated as:

$$RMS_t = \sqrt{\frac{\sum_{i=1}^{S_t} \Delta_{i,t}^2}{S_t}} \tag{39.4}$$

where, RMS_t is the accuracy of predicted ionospheric delay at epoch t; $\Delta_{i,t}$ is the difference between predicted and 'true' ionospheric delay for satellite i; S_t is the total number of visible satellites with the cut-off elevation of 10°.

The accuracy comparison between three ionospheric modeling methods is presented in Figs. 39.3 and 39.4. Figure 39.3 shows the RMS of predicted ionospheric delay in 40 km inter-station distance network and Fig. 39.4 shows the corresponding result in 80 km inter-station distance network. It can be seen that the ionospheric delay prediction using GIM has a lower accuracy in China area (RMS $<$ 3 m) than in the Netherlands (RMS $<$ 1 m).

Results in Figs. 39.3 and 39.4 also show that the accuracy of ionospheric delay prediction using SIM is much better than that of LIM and GIM in both China and the Netherlands. The accuracies of GIM and LIM in local afternoon (about 11:00–16:00, UTC) are much worse than that in other period, whereas the accuracy of SIM is basically the same during whole day. The reason is that the ionospheric activities usually reach a relatively high level in the afternoon and the variations in ionospheric delay may become more complex, and are difficult to be captured by GIM and LIM. Moreover, the error of ionospheric thin-layer assumption and mapping function will also become larger with high ionospheric activities. Unlike with GIM and LIM, the SIM has the ability of capturing this complex variations in ionospheric delay and can predict the ionospheric delay with more accuracy. It should be pointed out that GIM in China (about 1.0–2.0 m) has much poorer accuracy than in the Netherlands (about 0.40–0.80 m). It is because only 4–6 monitoring stations in China are contributed to IGS GIM generation, while more than 50 stations in Europe are used.

The mean and standard deviation of SIM, LIM and GIM accuracies during the experimental period are shown in Table 39.2. The mean accuracy of SIM and LIM

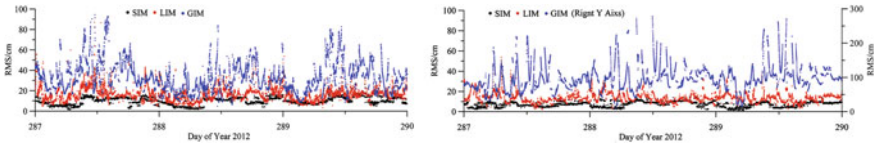


Fig. 39.3 Accuracy of ionospheric delay predicted using different approaches with an inter-station distance of 40 km in the Netherlands (*left*) and China (*right*)

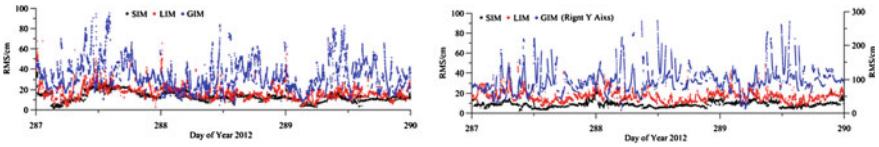


Fig. 39.4 Accuracy of ionospheric delay predicted using different approaches with an inter-station distance of 80 km in the Netherlands (*left*) and China (*right*)

Table 39.2 Mean and standard deviation of the accuracies of SIM/LIM/GIM in China and the Netherlands

Networks	SIM/cm		LIM/cm		GIM/cm
	40 km	80 km	40 km	80 km	
China	7.1 ± 2.1	8.9 ± 2.2	13.8 ± 5.6	16.3 ± 4.9	96.0 ± 34.5
The Netherlands	9.9 ± 2.8	12.5 ± 4.7	15.4 ± 6.0	17.0 ± 7.4	33.8 ± 14.0

in China is about 2–3 cm better than that in the Netherlands. This may be because the network in China is equipped with identical receivers and antennas at each station while the network in the Netherlands uses a variety of receivers and antennas. Overall, the ionospheric delay predicted by SIM achieves better than 10 cm accuracy in 40 km inter-station network and about 13 cm accuracy in 80 km inter-station network. The accuracy of LIM, which is about 15 cm (40 km) and 17 cm (80 km), is a bit lower than SIM. The GIM is not as good as the other two methods and it only achieves about 0.4 and 1.0 m accuracy level in the Netherlands and China respectively.

39.3.3 Positioning Accuracy of SF PPP

The ionospheric delay predicted with three different methods is applied to SF PPP and the positioning results are compared in this section. The dual-frequency PPP results in static mode are used as 'true' rover position. The SF PPP results based on GIM, LIM and SIM are presented in Figs. 39.5, 39.6, 39.7. The accuracy is illustrated on the eastern, northern and up components in local coordinate system.

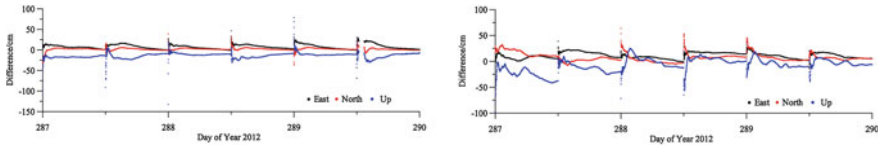


Fig. 39.5 Positioning result of SF PPP based on SIM (*left*) and LIM (*right*) with 40 km reference stations in the Netherlands

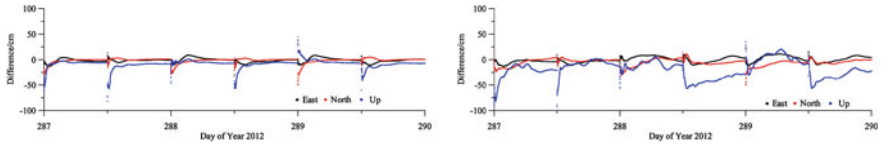


Fig. 39.6 Positioning result of SF PPP based on SIM (*left*) and LIM (*right*) with 40 km reference stations in China

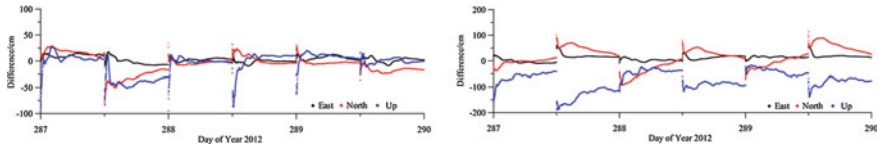


Fig. 39.7 Positioning result of SF PPP based on GIM in the Netherlands (*left*) and China (*right*) respectively

The Kalman filter in SF PPP is reset every 12 h to investigate the convergence time. Due to the limitation of space, the results based on 80 km reference stations are absent here.

It is shown by the positioning results that the initial positioning accuracy of SF PPP in each session based on SIM and LIM is almost the same (about 0.25 m), but the positioning accuracy based on SIM converges much better than that based on LIM with the data accumulation, particularly in the up component. The reason may be that the predicted ionospheric delay based on LIM is less accurate during ionospheric activity as shown by Figs. 39.3 and 39.4. The accuracy of GIM-based SF PPP can be improved significantly with the data accumulation in the Netherlands, while the improvement is not so significant in China and a systematic bias also exists in the up component. The positioning accuracy of SF PPP based on GIM in the Netherlands is much better than that in China. This result further indicates that the GIM is not very suitable for SF PPP user in China.

The average positioning accuracies of SF PPP in eastern, northern and up components in the Netherlands and China are summarized in Tables 39.3 and 39.4 respectively. The first 1 h in each session is considered as the convergence process and artificially excluded in the statistic. Compared with LIM, the SIM can improve the positioning accuracy about 20 %, especially in China area. The positioning

Table 39.3 Average positioning accuracy of SF PPP based on SIM, LIM and GIM in the Netherlands

Models		E/cm			N/cm			U/cm		
		Mean	STD	RMS	Mean	STD	RMS	Mean	STD	RMS
40 km	SIM	7.1	4.9	8.7	2.1	1.8	2.7	−13.2	4.3	13.9
	LIM	10.9	5.6	12.3	6.5	6.5	9.2	−7.0	12.7	14.5
80 km	SIM	4.7	3.9	6.1	1.9	1.6	2.5	−13.9	4.8	14.7
	LIM	7.0	5.4	8.8	3.9	5.5	6.8	−4.7	17.8	18.4
	GIM	2.8	5.9	6.6	−6.1	15.8	16.9	−2.9	18.1	18.3

Table 39.4 Average positioning accuracy of SF PPP based on SIM, LIM and GIM in China

Models		E/cm			N/cm			U/cm		
		Mean	STD	RMS	Mean	STD	RMS	Mean	STD	RMS
40 km	SIM	−1.0	3.8	3.9	−0.6	2.3	2.3	−7.0	3.4	7.7
	LIM	0.0	5.9	5.9	−4.7	5.5	7.3	−17.1	16.5	23.7
80 km	SIM	−1.7	4.7	5.0	−0.2	3.1	3.1	−6.8	5.1	8.5
	LIM	2.6	5.9	6.5	−3.9	6.9	7.9	−8.3	15.6	17.7
	GIM	9.8	8.9	13.3	15.8	36.3	39.6	−71.8	33.5	79.2

accuracy (RMS) of SF PPP in the Netherlands based on SIM is about 0.06–0.10 m and 0.13–0.15 cm in the horizontal and vertical components respectively and it is about 0.04–0.06 m and 0.07–0.09 m in China.

Applying GIM, the positioning accuracy of SF PPP is about 0.18 m in the Netherlands, whereas it is only about 0.40 m (horizontal) and 0.8 m (vertical) in China. The positioning result of SF PPP in China can hardly reach a sub-meter accuracy level with GIM ionospheric delay correction. Therefore, it is necessary to introduce some reference stations in China to provide more accurate ionospheric delay for SF PPP of sub-meter accuracy level.

In addition, comparing to the 'true' position, there is a systematic bias on the up component in all three methods. This bias may result from the residual error of ionospheric delay after applying the corrections, but it is still not confirmed so far.

39.3.4 Convergence Time of SF PPP

In addition to positioning accuracy, the convergence time is also very important for the real-time application of SF PPP. It is defined as the time to reach required accuracy level and the accuracy can be kept for at least 2 h in this experiment. The accuracy of 1.0, 0.75, 0.50 and 0.25 m in horizontal and vertical components is selected for the convergence time statistics respectively. Tables 39.5 and 39.6 show the convergence time of SF PPP based on LIM, SIM and GIM in the Netherlands and China respectively. The number shown in Tables 39.5 and 39.6 is

Table 39.5 Convergence time of SF PPP based on SIM, LIM and GIM in the Netherlands (*unit* epoch number with a measurement interval of 30s)

Distances of inter-station	Models		Accuracy			
			<100 cm	<75 cm	<50 cm	<25 cm
40 km	SIM	Horizontal	1.0	1.0	1.2	14.2
		Vertical	1.2	1.8	2.7	10.5
	LIM	Horizontal	1.0	1.0	4.3	15.9
		Vertical	1.2	1.7	7.8	28.8
80 km	SIM	Horizontal	1.0	1.0	1.0	14.5
		Vertical	1.0	1.5	2.3	18.2
	LIM	Horizontal	1.0	1.0	3.0	16.1
		Vertical	1.7	3.2	14.3	28.4
	GIM	Horizontal	1.0	2.0	7.0	19.6
		Vertical	1.7	10.3	22.8	39.2

Table 39.6 Convergence time of SF PPP based on SIM, LIM and GIM in China (*unit* epoch number)

Distances of inter-station	Models		Accuracy			
			<100 cm	<75 cm	<50 cm	<25 cm
40 km	SIM	Horizontal	1.0	1.0	1.5	23.5
		Vertical	1.0	1.2	19.2	58.3
	LIM	Horizontal	1.0	1.0	1.7	23.5
		Vertical	1.2	9.7	50.8	130.7
80 km	SIM	Horizontal	1.0	1.0	1.7	34.0
		Vertical	1.0	1.5	17.7	59.0
	LIM	Horizontal	1.0	1.0	2.0	38.5
		Vertical	1.2	2.2	56.2	114.5
	GIM	Horizontal	3.0	67.3	–	–
		Vertical	57.0	–	–	–

Note '–' means the corresponding accuracy is impossible to be achieved

the average epoch number for the convergence of all the six sessions. If the convergence time is longer than 2 h, the convergence time is regarded as infinite or accuracy requirement is not achievable.

The convergence time of SF PPP based on GIM becomes longer and longer when the required accuracy is better than 0.5 m on horizontal and vertical components and the SF PPP based on GIM even doesn't converge to the accuracy better than 0.75 m in China shown in Table 39.6. The results also show that the convergence time of SF PPP based on LIM and SIM is much shorter than that based on GIM. With the convergence time of about 1.5 min, the positioning accuracy of SF PPP based on SIM is better than 0.5 m in the Netherlands and better than 0.5 m (horizontal)/0.75 m (vertical) in China respectively. For the vertical accuracy of 0.5 m, the convergence time of SF PPP based on SIM in the Netherlands is only about 2 min, while that based on LIM is more than 4–7 min.

The convergence time of SF PPP based on LIM is about twice longer than that of SIM in term of the vertical accuracy. The superiority of SIM is more significant when a high level of accuracy is required, e.g. 0.5 or 0.25 m.

Therefore, it is further demonstrated that IGS released GIM is not feasible for sub-meter level SF PPP requirement in China. In the approach of ionospheric modeling, the error from the ionospheric thin-layer assumption and the mapping function may indeed not be ignorable and it may slow down the convergence of SF PPP very significantly, particularly in the vertical component.

39.3.5 Suggestions and Recommendations

It is feasible to use the SF PPP technique for achieving sub-meter positioning in real-time or near real-time mode. The IGS released GIM is able to aid the SF PPP achieving a positioning with 0.4–0.6 m accuracy based on 10–20 min accumulation observation in the Netherlands, whereas it is currently nearly invalid for the sub-meter SF PPP user in China. Additionally, the results also show that more Chinese stations should be considered in the GIM computation.

The performance of LIM is almost the same as that of SIM for SF PPP users requiring 0.75 m accuracy, but SIM is much better than LIM for the SF PPP users requiring 0.5 m accuracy. SF PPP based on LIM can achieve positioning accuracy better than 0.5 m with 1–2 min accumulation data. The results also indicate that the error resulting from ionospheric thin-layer and mapping function cannot be ignored directly for the precise ionospheric delay correction.

Therefore, a reference station network with 40–80 km inter-station distance is suggested to be set up for SF PPP user requiring sub-meter accuracy level in China and it can also be considered as an alternative for GIM in the Netherlands for the high precise SF PPP. SIM is recommended to be used for predicting the ionospheric delay correction at the rover rather than LIM. Based on SIM, the ionospheric delay is modeled on a satellite basis and independent of the ionospheric thin-layer assumption and mapping function.

39.4 Application of SIM in GPS/BDS Data

The Chinese BeiDou Navigation Satellite System (BDS) began to provide positioning, navigation and timing services (PNT) in the Asia-Pacific areas from the end of December, 2012 [2]. Currently, nearly 20 GPS and BDS satellites can be tracked in the Asia-Pacific region and the data of GPS and BDS can be combined together for SF PPP user. In this section, SIM, the recommend approach for ionospheric delay correction in Sect. 39.3, will be applied to SF PPP using GPS and BDS data and the positioning accuracy and convergence time will be analyzed.

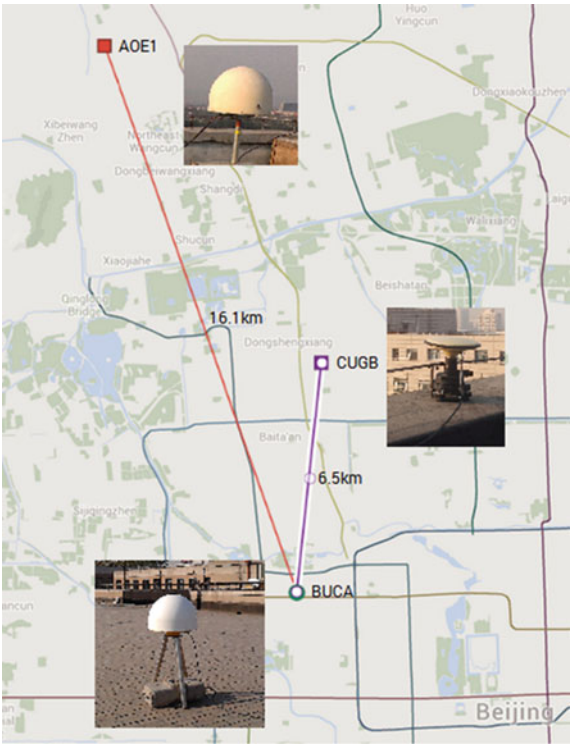


Fig. 39.8 Distribution of experimental GPS+BDS receivers located in Beijing, China

39.4.1 Description of GPS and BDS Data

Three GPS/BDS receivers were used to collect GPS and BDS data in Beijing, China on 13th Nov., 2013 for this experiment. These receivers are located in Academy of Opto-Electronics belonging to Chinese Academy of Sciences, China University of Geosciences (Beijing), and Beijing University of Civil Engineering and Architecture and named as AOE1, CUGB and BUCA respectively. Figure 39.8 shows the distribution of these stations and installed antennas. The types of antenna and receiver at AOE1 and BUCA are CC40GE and N71 M produced by CHC (http://www.chcnav.com/), whereas those at CUGB are GPS 704X and UR240-CORS-II produced by UNICORE (http://www.unicorecomm.com/).

In this experiment, BUCA receiver is selected as the rover station and the other two receivers (AOE1 and CUGB) are selected as reference stations. The length of baseline AOE1-BUCA is about 16.1 km and it is about 6.5 km for baseline CUGB-BUCA. Due to the limited number of contributed receivers, it is difficult to form a network with different inter-station distances. Thus, AOE1 and CUGB are individually considered as reference stations to provide the ionospheric delay correction for rover BUCA, so that the performance of SIM can be assessed with different inter-station distances.

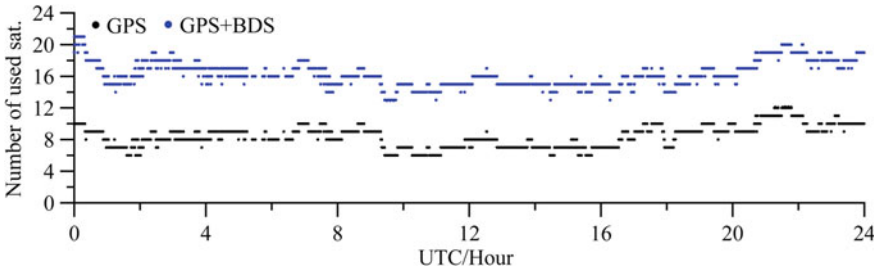


Fig. 39.9 Number of GPS and BDS satellites tracked by rover BUCA with the cut-off elevation of 10°

Currently, there are in total 14 BDS satellites in orbit consisting of five GEO satellites at the altitude of 35,786 km, five IGSO satellites at the altitude of 35,786 km with 55° inclination and four MEO satellites at the altitude of 21,528 km with 55° inclination [2]. Figure 39.9 illustrates the number of GPS and GPS + BDS satellites tracked by BUCA receiver at each epoch with the cut-off elevation of 10°. It can be seen that 7–9 GPS and 8–12 BDS satellites are observed during the whole day respectively.

The steps of data processing are the same as that described in Sect. 39.3.1, but it should be pointed out that (1) the DCB in GPS satellite is corrected using the CODE-released product and the DCB in BDS satellite and three receivers are estimated using IGGDCB; (2) the biases of receiver/satellite antenna phase center have not been calibrated in our experiment.

39.4.2 Result of GPS and GPS+BDS SF PPP

The accuracy of predicted ionospheric delay and positioning result of SF PPP based on a preliminary GPS/BDS dataset will be explored in this section, as well as the convergence time of SF PPP. Since there is only one reference station (AOE1 or CUGB), the predicted ionospheric delay at rover is actually the corresponding value obtained from each reference receiver. Assuming the ionospheric delay extracted from the dual-frequency observation of rover is true value, the mean and standard deviation of the accuracy of GPS/BDS ionospheric delay predicted from different reference stations is shown in Table 39.7. The whole day is divided into six periods with four-hour interval. It can be seen that the accuracies of ionospheric delay predicted from GPS and BDS satellites are almost the same. According to the daily characteristic of ionospheric activity, the accuracy becomes a little poorer in local afternoon (08:00:00–12:00:00, UTC). The result also shows that the accuracy of ionospheric delay predicted from AOE1 is better than that from CUGB, although the distance from rover to CUGB is much shorter than that from rover to AOE1. The reason may be that the receiver and antenna specific

Table 39.7 Mean and standard deviation of the accuracies of GPS/BDS ionospheric delay

Sessions	SIM-AOE1/cm		SIM-CUGB/cm	
	GPS[a]	BDS[a]	GPS[a]	BDS[a]
00:00:00–04:00:00	12.1 ± 7.0	16.5 ± 4.9	20.5 ± 6.6	20.1 ± 3.6
04:00:00–08:00:00	17.9 ± 6.5	15.3 ± 5.2	28.5 ± 4.3	31.0 ± 6.0
08:00:00–12:00:00	21.4 ± 7.8	9.7 ± 4.2	24.2 ± 3.8	27.7 ± 4.3
12:00:00–16:00:00	18.4 ± 4.4	11.7 ± 3.5	20.3 ± 5.1	21.8 ± 3.6
16:00:00–20:00:00	17.2 ± 6.6	16.7 ± 1.7	28.2 ± 9.0	24.7 ± 1.9
20:00:00–24:00:00	13.3 ± 5.1	16.4 ± 6.3	21.2 ± 6.3	26.5 ± 6.1
Mean	15.6 ± 5.5		24.6 ± 5.4	

[a] Number before "±" is the mean and that after "±" is the corresponding standard deviation

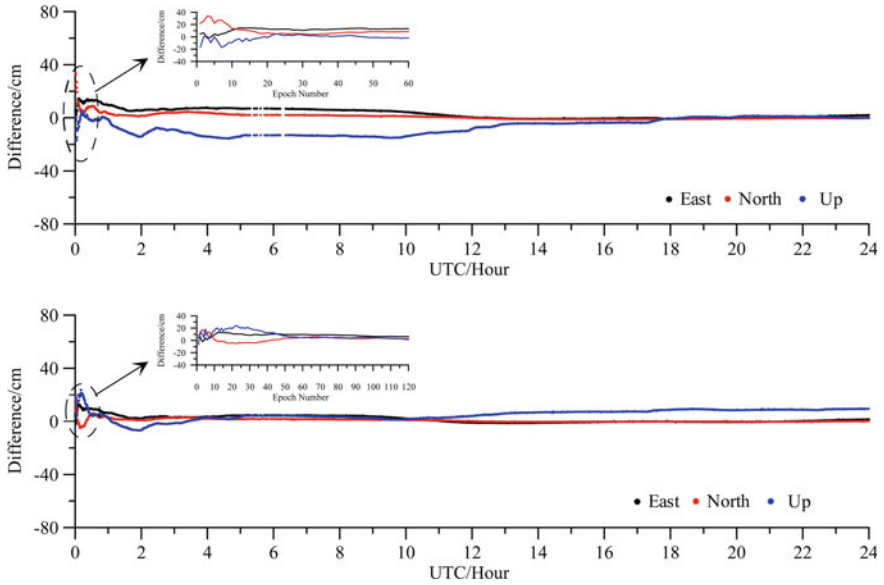


Fig. 39.10 Positioning results of SF PPP using the ionospheric delay predicted from AOE1 based on GPS (*upper*) and GPS+BDS (*lower*) respectively

biases can be completely eliminated in this case. Overall, the ionospheric delay predicted from AOE1 reach the accuracy level of about 0.16 m, while that from CUGB is about 0.25 m. This accuracy is comparable with that demonstrated by GPS data in Sect. 39.3.2.

Figure 39.10 shows the differences between estimated and 'true' coordinates of rover in eastern, northern and up components using the predicted ionospheric delay from AOE1 (inter-station distance of about 16.1 km). The upper panel shows the positioning result with GPS data, while the lower panel shows the positioning result using GPS + BDS data. The 'true' position of rover is calculated with dual-frequency PPP technique using GPS only data. Figure 39.11 shows

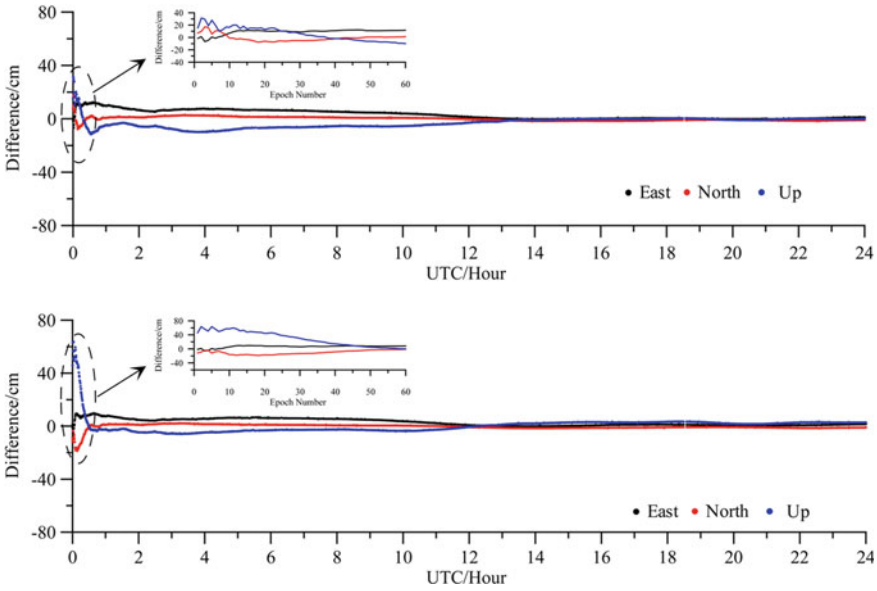


Fig. 39.11 Positioning results of SF PPP using the ionospheric delay predicted from CUGB based on GPS (*upper*) and GPS+BDS (*lower*) respectively

Table 39.8 Daily positioning accuracy of GPS/BDS SF PPP based on the ionospheric delay predicted from AOE1 and CUGB

Accuracy	SIM-AOE1/cm		SIM-CUGB/cm	
	GPS	GPS + BDS	GPS	GPS + BDS
Horizontal	4.2	2.6	4.2	3.7
Vertical	5.7	4.3	4.6	3.1

the similar positioning result based on the reference station of CUGB (inter-station distance of about 6.5 km).

Except the results shown by the lower panel of Fig. 39.11, the positioning accuracy SF PPP in eastern, northern and up components at the first epoch is better than 0.5 m when applying the ionospheric delay correction based on SIM. Compared with the positioning result based on GPS alone, the combination of GPS data BDS data could improve the instantaneous positioning accuracy of SF PPP (lower panel of Fig. 39.10 but not 39.11). This high accurate instantaneous positioning is advantageous to reduce the convergence time of SF SPP.

The positioning accuracy of SF PPP using 24-h data is given by Table 39.8. It can be seen that the horizontal and vertical accuracy of SF PPP is better than 0.05 and 0.06 m respectively. The accuracy of SF PPP with SIM ionospheric delay correction is comparable with dual-frequency PPP without ambiguity fixing.

Table 39.9 Convergence time of SF PPP based on reference station AOE1 and CUGB using GPS/BDS data (*unit* epoch number)

		50–100 cm		25–50 cm		<25 cm	
		GPS	GPS + BDS	GPS	GPS + BDS	GPS	GPS + BDS
AOE1	Horizontal	1	1	1	1	8	1
	Vertical	1	1	1	1	1	1
CUGB	Horizontal	1	1	1	1	1	1
	Vertical	1	1	1	15	6	33

The convergence time for different positioning accuracies is summarized in Table 39.9. In order to see the convergence process clearly, the positioning results during the first 2 h in Figs. 39.10 and 39.11 are also zoomed in. It can be seen that the positioning with 0.25 m horizonal accuracy can be achieved by single epoch SF PPP based on the ionospheric delay correction using SIM. In term of the result based on reference station AOE1, the convergence time can be reduced significantly when a more accurate positioning is required (e.g. better than 0.25 m). However, the result based on reference station CUGB shows that the convergence time for vertical component become longer when combining GPS and BDS data. The reason needs to be further analyzed.

Based on the here presented analysis, it suggests that SIM approach can provide the ionospheric delay at the accuracy level of 0.2 m and improve single epoch SF SPP to reach the sub-meter positioning accuracy level using GPS and BDS data. The accuracy of SF PPP result based on ionospheric delay correction from SIM in one day is about 0.05 and 0.06 m in horizontal and vertical components respectively. In order to obtain a much better ionospheric delay correction, the same antenna and receiver is suggested to be used in the reference and rover station. Experiment result also suggests that the combined GPS and BDS data can further improve the vertical accuracy of SF PPP and reduce the convergence time.

39.5 Conclusions and Future Works

The PPP technique with single frequency receiver is one of the potential approaches to achieve a sub-meter (better than 0.5 m) positioning with low cost. The ionospheric delay, as the toughest error sources in SF PPP, has to be mitigated as much as possible to realize this goal. This paper has reviewed the commonly used ionospheric delay mitigation method for single frequency user, including global ionospheric map released by IGS, local ionospheric model based on second-order polynomial and satellite based ionospheric delay model. The performances of different approaches are assessed and compared using two GPS networks from China and the Netherlands with different inter-station distances. The assessment is carried out in the following three aspects: accuracy of predicted ionospheric delay, positioning accuracy and convergence time of SF PPP.

Comparison result demonstrates that: (1) the IGS released GIM can currently not aid SF PPP achieving a positioning with sub-meter accuracy in China, whereas it is effective for SF PPP users of sub-meter accuracy level in the Netherlands; (2) A reference network surrounding the rover with 40–80 km is necessary for SF PPP in China to meet the sub-meter positioning requirement; (3) A satellite based ionospheric model (SIM) in which the ionospheric thin-layer assumption and mapping function can be avoided is suggested for SF PPP based on a regional reference network rather than the traditional ionospheric modeling method.

Following these suggestions and recommendations, the SIM has been applied to SF PPP based on a limited GPS and BDS dataset gathered in Beijing, China. Numerical result demonstrates that the SF PPP based on SIM can achieve a sub-meter positioning accuracy, even with single epoch. Compared with the result only based on GPS data, the combined GPS and BDS data can improve the accuracy about 20 % and reduce the convergence time, particularly in the vertical component.

However, the period of experimental data collection and distribution of receiver location are all at a medium level of ionospheric activities, the drawn conclusions may be more conservative for low ionospheric activities as well as being relatively short and optimistic for high ionospheric activities; thus, more experiment should be further carried out in different levels of ionospheric activates. In addition, the inconsistency of different types of antenna and receiver between rover and reference stations needs to be further analyzed.

Acknowledgments This research was partially supported by National Key Basic Research Program of China (Grant No: 2012CB825604), National Natural Science Foundation of China (Grant No: 41304034, 41231064), Beijing Natural Science Foundation (Grant No: 4144094), Scientific Cooperation between China and the Netherlands programme '*Compass, Galileo and GPS for improved ionosphere modeling*' and the State Key Laboratory of Geodesy and Earth's Dynamics (Institute of Geodesy and Geophysics, CAS) (Grant No:SKLGED2014-3-1-E). The GPS data used in the Netherlands was kindly provided by the NETPOS (the Netherlands Positioning Service) of the Dutch Kadaster. The GPS related products used in our experiment were downloaded from the IGS Global Data Center CDDIS (Crustal Dynamics Data Information System, Greenbelt, MD, USA, www.cddis.gsfc.nasa.gov) and the ftp servers of CODE (Center for Orbit Determination in Europe, Switzerland, ftp.unibe.ch). Prof. Junhuan Peng and Dr. Yanli Zheng from China University of Geosciences (Beijing) and Prof. Keliang Ding from Beijing University of Civil Engineering and Architecture provided the helps on GPS/BDS data collection in China. Thanks for valuable suggestions from Lei Wang, Yanqing Hou and Dr. Wei Yan.

References

1. Allain D, Mitchell C (2009) Ionospheric delay corrections for single-frequency GPS receivers over Europe using tomographic mapping. GPS Solut 13(2):141–151
2. BD-SIS-ICD (2012) BeiDou navigation satellite system signal in space interface control document, China Satellite Navigation Office, Beijing
3. Bisnath S, Gao Y (2008) Current state of precise point positioning and future prospects and limitations. In: Sideris MG (ed) Observing our changing earth, vol 133, pp 615–623

4. Bree RP, Tiberius CJM (2012) Real-time single-frequency precise point positioning: accuracy assessment. GPS Solut 16(2):259–266
5. Chen K, Gao Y (2005) Real-time precise point positioning using single frequency data. Paper presented at proceedings of ION GNSS 2005, Long Beach, CA
6. Coco DS et al (1991) Variability of GPS satellite differential group delay biases. IEEE Trans Aerosp Electron Syst 27(6):931–938
7. Conte J et al (2011) Accuracy assessment of the GPS-TEC calibration constants by means of a simulation technique. J Geodesy 85(10):707–714
8. Feess WA, Stephens SG (1987) Evaluation of GPS ionospheric time-delay model. IEEE Trans Aerosp Electron Syst AES-23(3):332–338
9. Feltens J et al (1998) Routine production of ionosphere TEC maps at ESOC-first results (IGS presentation). Paper presented at proceedings of the 1998 IGS AC workshop, ESOC, Darmstadt, Germany
10. Feltens J (2007) Development of a new three-dimensional mathematical ionosphere model at European Space Agency/European Space Operations Centre. Space Weather 5(S12002):1–17
11. Geng J et al (2010) Rapid re-convergences to ambiguity-fixed solutions in precise point positioning. J Geodesy 84(12):705–714
12. Georgiadiou Y (1994) Modeling the ionosphere for an active control network of GPS stations. LGR-Series-Publications of the Delft Geodetic Computing Centre, vol 7
13. Hernández-Pajares M et al (1999) New approaches in global ionospheric determination using ground GPS data. J Atmos Solar Terr Phys 61(16):1237–1247
14. Hernández-Pajares M (2004) IGS ionosphere WG status report: performance of IGS ionosphere TEC maps (Position paper)
15. Hernández-Pajares M (2006) Summary and current status of IGS ionosphere WG activities—a potential future product: global maps of effective ionospheric height. In: IGS technical workshop
16. Hernández-Pajares M et al (2009) The IGS VTEC maps: a reliable source of ionospheric information since 1998. J Geodesy 83(3):263–275
17. IS-GPS (2004) Navstar GPS space segment/navigation user interfaces (ICD-GPS-200D), Revision D. ARINC Engineering Services, LLC, El Segundo, CA
18. Klobuchar JA (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst AES-23(3):325–331
19. Komjathy A et al (2005) Automated daily processing of more than 1000 ground-based GPS receivers for studying intense ionospheric storms. Radio Sci 40(6):S6006
20. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS Solut 5(2):12–28
21. Kouba J (2009) A guide to using International GNSS Service (IGS) products. http://www.igs.org/igscb/resource/pubs/UsingIGSProductsVer21.pdf
22. Lanyi G et al (1988) A comparison of mapped and measured total ionospheric electron content using global positioning system and beacon satellite observations. Radio Sci 23(4):483–492
23. Le AQ et al (2008) Use of global and regional ionosphere maps for single-frequency precise point positioning. In: Sideris MG (ed) Observing our changing Earth international association of Geodesy symposia
24. Li X et al (2013) A method for improving uncalibrated phase delay estimation and ambiguity-fixing in real-time precise point positioning. J Geodesy 87(5):405–416
25. Li Z et al (2012) Two-step method for the determination of the differential code biases of COMPASS satellites. J Geodesy 86(11):1059–1076
26. Liu J et al (2010) Spherical cap harmonic model for mapping and predicting regional TEC. GPS Solut 15(2):109–119
27. Mannucci AJ et al (1999) GPS and ionosphere: review of radio science 1996–1999. Oxford University Press, New York
28. Odijk D et al (2014) Single-frequency PPP-RTK: theory and experimental results. In: Proceedings of the IAG general assembly, Melbourne, Australia, pp 571–578

29. Orús R et al (2002) Performance of different TEC models to provide GPS ionospheric corrections. J Atmos Solar Terr Phys 64(18):2055–2062
30. Orús R et al (2005) Improvement of global ionospheric VTEC maps by using kriging interpolation technique. J Atmos Solar Terr Phys 67(16):1598–1609
31. Schaer S et al (1998) IONEX: the IONosphere map EXchange format version 1. Paper presented at proceedings of the IGS AC Workshop, Darmstadt, Germany
32. Schaer S (1999) Mapping and predicting the earth's ionosphere using the global positioning system. Ph D thesis, Astronomical Institutes, University of Bern, Berne, Switzerland
33. Schüler T et al (2011) Precise ionosphere-free single-frequency GNSS positioning. GPS Solut 15(2):139–147
34. Shi C et al (2012) An improved approach to model ionospheric delays for single-frequency precise point positioning. Adv Space Res 49(12):1698–1708
35. Takasu T, Yasuda A (2009) Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. In: International symposium on GPS/GNSS, International Convention Center Jeju, Korea November 4–6
36. Takasu T (2009) RTKLIB: open source program package for RTK-GPS. In: FOSS4G 2009, Tokyo, Japan
37. Tiberius C et al (2011) Staying in lane: real-time single-frequency PPP on the road. In: Proceedings of inside GNSS(November/December), pp 48–53
38. Wu X et al (2012) Evaluation of COMPASS ionospheric model in GNSS positioning. Adv Space Res 51(6):959–968
39. Yuan Y, Ou J (2004) A generalized trigonometric series function model for determining ionospheric delay. Prog Nat Sci 14(11):1010–1014
40. Yuan Y et al (2008) Refining the Klobuchar ionospheric coefficients based on GPS observations. IEEE Trans Aerosp Electron Syst 44(4):1498–1510
41. Zhang B (2013) Study on the theoretical methodology and applications of precise point positioning using un-differenced and uncombined GNSS data. Graduate University of Chinese Academy of Sciences, Wuhan, China
42. Zhang H et al (2013) On the convergence of ionospheric constrained precise point positioning based on undifferential uncombined raw observation. Sensors 13:15708–15725
43. Zou X et al (2012) A new ambiguity resolution method for PPP using CORS network and its real-time realization. Paper presented at China Satellite Navigation Conference (CSNC) 2012. Springer, Berlin/Heidelberg, Germany, Guangzhou, China
44. Zumberge JF et al (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res Solid Earth 102(B3):5005–5017
45. Ding W (2012) Research on key technologies of real time precise point positioning system. Ph D thesis, University of Chinese Academy of Sciences, Wuhan, China (in Chinese)
46. Li Z (2012) Study on the mitigation of ionospheric delay and the monitoring of global ionospheric TEC based on GNSS/Compass. Ph D thesis, Institute of Geodesy and Geophysics, University of Chinese Academy of Sciences, Wuhan, China (in Chinese)
47. Wen J et al (2010) Experimental observation and statistical analysis of the vertical TEC mapping function. Chin J Geophys 53(1):22–29 (in Chinese)
48. Zhang B et al (2011) Determination of ionospheric observables with precise point positioning. Chin J Geophys 54(4):950–957 (in Chinese)