Monitoring Station Data Quality Analysis Method

Hongyi Ren[1], Zhigang Huang[1](✉), Rui Li[1], and Tiantian Yang[2]

[1] School of Electronic Information Engineering, Beihang University, Beijing, China
[2] Technical Center of Air Traffic Management Bureau of CAAC, Beijing, China

Abstract. The Beidou augmentation system currently mainly provides services for stationary users and low-dynamic users. With the continuous development of autonomous driving technology, it will provide services for L4 autonomous driving users in the future. Autonomous driving users are highly dynamic users, and the service is closely related to the user's life safety, so it is necessary to ensure the accuracy and continuity of the monitoring station data. In the existing Beidou augmentation system reference station construction and acceptance technical specifications, the three indicators of multipath error, cycle slip ratio and observation data availability can be used to evaluate the data quality of the receiver. These three indicators can only reflect the correctness and completeness of the data, but cannot reflect the continuity of the data. In order to reflect the ability of monitoring stations to work continuously, this paper proposes data continuity indicators to evaluate the data quality of monitoring station receivers, which can provide new ideas for the subsequent evaluation of the reliability of monitoring station data.

Keywords: Ground-based augmentation system · Monitoring station · Data quality · Data continuity

1 Introduction

Since the Beidou Satellite Navigation System (BDS) officially provided services, China has begun to build Ground-based Augmentation System (GAS). On June 23, 2020, the BDS global network was officially completed. Beidou augmentation system will also cooperate with the BDS to provide ground-based enhancement services in China. The Beidou augmentation system is a Continuously Operational Reference System (CORS) established by multi-base station network Real Time Kinematic technology (RTK). The GAS consists of a BDS augmentation station network, a communication network, a data processing and broadcasting system, etc. It provides augmentation service by wide-area augmentation products, regional augmentation products and post-processing high-precision data products. Beidou augmentation system is mainly used in precision agriculture, conventional surveying and mapping, meteorological observation and other fields. The users in these fields are static users and low dynamic users [1]. With the continuous development of autonomous driving technology, Beidou augmentation system will

C. Yang and J. Xie (Eds.): *China Satellite Navigation Conference (CSNC 2021) Proceedings*, LNEE 773, pp. 22–32, 2021.
https://doi.org/10.1007/978-981-16-3142-9_3

provide services for L4 autonomous driving users in the future. This service is closely related to the users' life safety, so the accuracy and continuity of the monitoring station data is particularly important.

At present, the International Civil Aviation Organization (ICAO) has achieved high-level autonomous driving in the aviation applications field, and can ensure high integrity and high continuity of received data. This is related to the strict standards of ICAO when building Satellite-Based Augmentation System (SBAS) monitoring stations. If Beidou augmentation system is to provide services for high dynamic users related to life safety, the accuracy and continuity of the data must be close to the SBAS stations. Since Beidou augmentation system uses CORS stations, it is necessary to compare the data quality of CORS stations and SBAS station receivers in order to evaluate whether Beidou augmentation system can provide services for high dynamic users related to life safety such as autonomous driving. Because there is currently no data from the Beidou augmentation system monitoring station, the data from the similar system, Crustal Movement Observation Network Of China (CMONOC), is selected for comparison and analysis with the data from the WAAS monitoring stations and the US CORS stations.

In the current Beidou augmentation system reference station construction and acceptance specifications, the indicators for evaluating receiver data quality are multipath error, cycle slip ratio and observation data availability [11]. The above indicators can only reflect correctness and completeness of data, not continuity. In order to reflect capability of continuous operation capability of monitoring stations, this paper introduces the data continuity indicators to evaluate the data quality of the receivers of each monitoring station, in order to provide a reference for the improvement of the construction standards of the Beidou reference stations.

2 Data Quality Analysis Indicator and Method

2.1 Multipath Error

Multipath error is the ranging error caused by non-direct navigation signal. In calculation, it needs to rely on dual-frequency observation data, combining the pseudorange observation equation and carrier phase observation equation to eliminate the influence of tropospheric and ionospheric delay. Multipath error is the main error in the ranging signal. Affected by the multipath effect, the accuracy of the pseudorange and phase observations of GNSS will drop sharply, which can lead to signal loss in severe cases. The multipath effect seriously affects the positioning and navigation accuracy, so it needs to be considered in data quality analysis [2].

When calculating the multipath error, it is necessary to combine the pseudorange observation equation and the carrier phase observation equation, and use the dual-frequency observation data to calculate, as shown in Eq. (1.1) [3].

$$\begin{cases} MP_1 = \rho_1 - [(f_1^2 + f_2^2)/(f_1^2 - f_2^2)]\varphi_1 + [2f_2^2/(f_1^2 - f_2^2)]\varphi_2 \\ MP_2 = \rho_2 - [2f_2^2/(f_1^2 - f_2^2)]\varphi_1 + [(f_1^2 + f_2^2)/(f_1^2 - f_2^2)]\varphi_2 \end{cases} \tag{1.1}$$

In Eq. (1.1), MP_1 and MP_2 are calculation amount of k_1 and k_2 frequencies including multipath error and whole-cycle ambiguity information; ρ_1 and ρ_2 are k_1 and k_2

frequencies pseudorange observation, the unit is meter; f_1 and f_2 are the frequencies of navigation signal k_1 and k_2 frequencies carrier, the unit is megahertz; φ_1 and φ_2 are k_1 and k_2 frequencies carrier phase observation, the unit is meter.

Multipath error need to be calculated using sliding windows. When the same satellite is continuously observed and there is no cycle slip, the combined ambiguity parameter will not change, and the multipath error can be obtained by calculating according to Eq. (1.2) between multiple epochs without cycle slip [4].

$$\overline{MP_k} = \sqrt{[1/(N_{sw} - 1)] \sum_{i=1}^{N_{sw}} \left[MP_k(t_i) - (\sum_{i=1}^{N_{sw}} MP_k(t_i)/N_{sw}) \right]^2} \tag{1.2}$$

In Eq. (1.2), $\overline{MP_k}$ is the evaluation value of multipath error observed by the receiver at frequency k; N_{sw} is the number of epochs in the sliding window, and the default is 50; $MP_k(t_i)$ is the calculation amount of multipath error and whole-cycle ambiguity information observed by the receiver of epoch t_i at k frequency.

Since the data of WAAS monitoring station is not smoothed, it needs to be smoothed. The sliding window of the smoother is generally set between 20–100 epochs, and the smoothing Equation is shown in Eq. (1.3) [5].

$$\rho_{s,k} = (1/M)\rho_k + [(M-1)/M]\left[\rho_{s,k-1} + \lambda(\varphi_k - \varphi_{k-1})\right] \tag{1.3}$$

In Eq. (1.3), $\rho_{s,k}$ is the carrier phase smoothing pseudorange of the kth epoch, φ_k is the carrier phase observation of the kth epoch, the unit is cycle, and M is the smoothing time constant. The larger the M, the better the smoothing effect. Its value range is 20–100 epochs (seconds). This paper uses $M = 100$.

2.2 Cycle Slip Ratio

Before calculating the cycle slip ratio, the number of cycle slips should be calculated. Cycle slip is a counting error which occurs when the receiver performs continuous carrier phase measurement, due to some reason, and leads to a whole cycle slip of the phase observation value. Cycle slip ratio refers to the ratio of the number of epochs of the complete observation value of the receiver's observation data to the number of epochs in which cycle slips occur during the observation of all observable satellites in a certain period of time.

When calculating cycle slip ratio, cycle slip detection is required to obtain the number of epochs where the cycle slip occurs. When detecting cycle slips, a combination of Melbourne-Wubbena method and ionospheric residual error method is used. If multiple satellites have cycle slips in a certain epoch, there are two methods for calculating the number of cycle slips. The first method is that the number of cycle slips is equal to the number of satellites in the epoch. The second method is that no matter how many satellites in the epoch have cycle slips, the number of cycle slips is calculated only once. The two methods will be denoted as cycle slip ratio 1 and cycle slip ratio 2 when analyzing the results below.

2.3 Observation Data Availability

Observation data availability refers to the ratio of the number of epochs containing complete observations in the receiver observation data to the number of theoretical epochs for all observable satellites in a certain period of time. The theoretical epoch refers to the epoch that may be obtained by observing all observable satellites in a period of time. This paper defines an epoch with an elevation angle greater than 15° as a theoretical epoch. The epoch with complete observations means that all observable satellites are observed within a period of time, and there are no missing epochs for all frequency pseudoranges and carrier phase observations [6]. This paper uses L1 and L2 dual-frequency observation data of each station for the statistics of observation data availability. Therefore, the epoch without missing observations of the L1 and L2 dual-frequency pseudoranges and carrier phases at the monitoring stations is defined as the complete observation epoch.

There are two methods for calculating observation data availability. The first method is to consider the number of visible satellites in the calculation, and the second method is to not consider the number of visible satellites in the calculation. This paper uses the first method when calculating observation data availability. The observation data availability can be calculated using Eq. (1.4).

$$DI_S = \left(\sum_{j=1}^{n} C^j \Big/ \sum_{j=1}^{n} D^j \right) \times 100\% \tag{1.4}$$

In Eq. (1.4), DI_s is the available rate of observation data; C^j is the number of epochs in which the jth satellite has a complete observation value during the observation period; D^j is the total number of theoretical epochs of the jth satellite during the observation period; n is the total number of satellites observed in the observation period, this paper takes $n = 31$.

3 Data Sources

Observation data comes from 15 monitoring stations, 8 WAAS stations (zau1, zkc1, zme1, zmp1, zfw1, zab2, zdv1 and zob1), 3 US CORS stations (alja, alla and nvpo) and 4 CMONOC stations (GSTS, JSYC, SXGX and YNTC). The location of the WAAS monitoring station (blue icon), the US CORS station (red icon) and the CMONOC station (purple icon) are shown in Fig. 1. The signal-to-noise ratio of each station is only related to the stability of receiver, and has nothing to do with environmental factors. Select continuous observational data from January 1, 2015 to January 30, 2015 for data quality analysis, and data sampling interval is 1 s. Among the 15 selected monitoring stations, 8 WAAS monitoring stations use NOV WAASGII receivers; 3 US CORS stations use LEICA receivers; 4 CMONOC stations use TRIMBLE receivers, GSTS, SXGX and YNTC stations use TRIMBLE NETR8 receiver, and JSYC station uses TRIMBLE NETR9 receiver.

In addition, the lower the elevation angle of the satellite, the worse the quality of the signal received by the receiver. Therefore, when evaluating receiver data quality, each

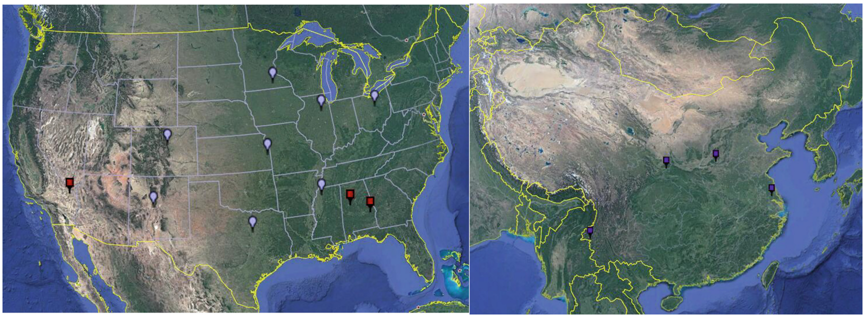

Fig. 1. Location of WAAS monitoring stations, US CORS stations and CMONOC stations

satellite should have a minimum elevation angle, and data below the minimum elevation angle should be excluded [7]. In practical applications, the elevation gate is generally limited to 5–15°. This article sets the minimum elevation angle to 15° when evaluating data quality.

4 Result Analysis

4.1 Multipath Error

Calculate the 95% quantile of the multipath error of all satellites in the one-month data from WAAS monitoring station, US CORS station and CMONOC station. The results are shown in Table 1. The unit of multipath error is meter.

From the statistical results, it can be seen that after smoothing the data from WAAS stations using Eq. (1.3), the multipath errors between WAAS stations and US CORS stations are relatively small, while the multipath errors of CMONOC stations are relatively large. The reason may be that the environmental interference of the station site of CMONOC is relatively large, while the environmental interference of the WAAS station and the US CORS station site is relatively small.

4.2 Cycle Slip Ratio

The cycle slip ratio is calculated on the data of all satellites of WAAS station, US CORS station and CMONOC station for one month. The cycle slip detection method is a combination of Melbourne-Wubbena method and ionospheric residual error method. the results are shown in Table 2.

It can be seen from the statistical results that cycle slip ratio calculated by two methods is not much different. What's more, the cycle slip ratio of WAAS stations is between 1500–2200, US CORS stations is between 2400–3000, and CMONOC stations is between 1000–2800. It shows that the consistency of the cycle slip ratio of the WAAS station and the US CORS station is relatively strong, while the consistency of the cycle slip ratio of the CMONOC station is weak.

Table 1. 95% quantile multipath error of L1 and L2 frequency at each station for one month

Station name	Raw data (L1)	After smoothing (L1)	Raw data (L2)	After smoothing (L2)
zau1	1.0180	0.1554	0.8133	0.1768
zkc1	0.8003	0.1286	0.6320	0.1364
zme1	0.9448	0.1539	0.7734	0.1567
zmp1	0.9834	0.1570	0.7409	0.1472
zfw1	1.0773	0.1744	0.9143	0.1876
zab2	0.8955	0.1511	0.7298	0.1417
zdv1	0.8819	0.1451	0.7803	0.1524
zob1	0.7475	0.1433	0.6274	0.1526
nvpo	0.1112	–	0.1373	–
alja	0.1489	–	0.1965	–
alla	0.1741	–	0.1996	–
GSTS	0.6042	–	0.2825	–
SXGX	0.4665	–	0.4328	–
YNTC	0.3572	–	0.2897	–
JSYC	0.3605	–	0.3143	–

4.3 Observation Data Availability

Statistics on the observation data availability for one month of data at all stations are performed. The results are shown in Table 3. The epoch unit is second.

It can be seen from the statistical results that the observation data availability at 6 of the 8 WAAS stations reached 0.999. The observation data availability from two of the three US CORS stations reached 0.999. The observation data availability from the four CMONOC stations has not reached 0.99, which was the lowest.

5 Data Continuity Analysis

In the existing Beidou augmentation system standards, multipath error, cycle slip ratio and observation data availability can be used to evaluate data quality of station receiver. Analyzing the definitions of the three indicators, it is found that the multipath error essentially reflects receiver thermal noise. The cycle slip ratio reflects the ratio of the number of complete epochs of the data to the number of cycle slip epochs. The more cycle slips, the more epochs the receiver has error, but this indicator cannot reflect the time when the data cycle slips. The availability of observational data reflects whether the single epoch data of each satellite is complete, it does not consider the correctness of the data. These indicators cannot reflect the continuity of data, but autonomous driving users need to ensure the correctness and continuity of the received data, so that the Beidou augmentation system can monitor its integrity to ensure the life safety of autonomous

Table 2. Statistics of the one-month original data cycle slip ratio of each station

Station name	Cycle slip ratio 1	Cycle slip ratio 2
zau1	2166	2183
zkc1	1815	1882
zme1	1554	1644
zmp1	1585	1688
zfw1	1788	1814
zab2	1998	2018
zdv1	2135	2157
zob1	1727	1861
nvpo	2488	2502
alja	2961	2991
alla	2977	2997
GSTS	1072	1078
SXGX	1251	1286
YNTC	2475	2491
JSYC	2681	2708

driving users. In order to reflect observation data continuity and the continuous operation capability of each station, this paper proposes a data continuity indicator to evaluate receiver data quality. The following will use this indicator to compare the difference in data quality between WAAS stations, US CORS stations and CMONOC stations.

5.1 Data Continuity Definition

In the existing *GPS Standard Positioning Service Performance Standards*, the concept of spatial signal continuity and outage is defined [8]. The paper draws on this definition and defines data continuity and interruption as follows: Data continuity refers to the situation where there is no interruption in the observation arc, which is measured by the indicator of data continuity probability. Interruption means that the number of available satellites in an epoch is less than 5 due to cycle slip or too few visible satellites, or a certain epoch dual-frequency pseudorange and carrier phase data are missing. When calculating data continuity probability, the sliding window method is generally used.

Assuming that the data time period is $[t_{start}, t_{end}]$, the user's sampling interval is T, and the sliding window time is denoted as t_{op}, the total number of satellites observed in the observation period is n, this paper takes $n = 31$, the calculation equation of the data continuity probability Con is [9]:

Table 3. Statistics of the availability of observation data for one month at each station

Station name	Complete epochs	Theoretical epochs	Observation data availability
zau1	16836660	16837630	0.999942
zkc1	17026189	17028967	0.999837
zme1	16769180	16773937	0.999716
zmp1	16883339	16910956	0.998367
zfw1	17160542	17163652	0.999819
zab2	17190557	17192056	0.999913
zdv1	17037398	17039217	0.999893
zob1	16555748	16594309	0.997676
nvpo	19598451	19749312	0.992361
alja	19202394	19216828	0.999249
alla	19160275	19167356	0.999631
GSTS	14501292	19137082	0.757759
SXGX	15695267	17255934	0.909558
YNTC	16320075	19277386	0.846592
JSYC	19237450	19562306	0.983394

$$Con = \left\{ \sum_{j=1}^{n} \sum_{t=t_{start},inc=T}^{t_{end}-t_{op}+1} \left[\prod_{k=t,inc=T}^{t+t_{op}-1} bool(Statu_j(k)) \right] \right\} / \left[\sum_{j=1}^{n} \sum_{t=t_{start},inc=T}^{t_{end}-t_{op}+1} bool(Statu_j(t)) \right] \quad (1.5)$$

When using data continuity probability to calculate Eq. (1.5), it is need to calculate the number of available satellites for each epoch. First, count the satellites with an elevation angle greater than 15° and complete observations for each epoch, and take $bool(Statu_j(k))$ function value of these satellites as 1, and $bool(Statu_j(k))$ function value of other satellites as 0, where k is the epoch time and j is the satellite number. After that, perform cycle slip detection on the observation data of all satellites, record the satellites with cycle slips in each epoch, and set $bool(Statu_j(k))$ function value of these satellites to 0. Finally, count the number of available satellites for each epoch excluding the satellites that have occurred cycle slips. If the number of available satellites in the epoch is greater than 5, $bool(Statu_j(k))$ function value of the satellites that occurred cycle slips is set to 1. Otherwise, $bool(Statu_j(k))$ function value of all satellites in this epoch is set to 0. When calculating the $bool(Statu_j(t))$ function value of the denominator, it is need to count the elevation angle of each satellite in each epoch. When the elevation angle of the satellite j in the tth epoch is greater than 15°, the $bool(Statu_j(t))$ function is set to 1, otherwise it is set to 0.

It can be seen from Eq. (1.5) that the selection of sliding window time t_{op} is very important for calculating data continuity probability. At present, the methods used for positioning mainly include Precise Point Positioning (PPP) and network RTK. The convergence time of PPP is too long to meet the real-time requirements of automatic driving. The network RTK technology uses multiple reference stations deployed on the ground to form a continuous GPS reference station network, comprehensively utilizes the observation information of each reference station, and establishes an accurate error model to generate a Virtual Reference Station (VRS) near the user that does not exist physically [10]. The VRS broadcast compensation value of power generation to user receiver at 15 s intervals. In order to generate an accurate ionospheric compensation value, it is necessary to ensure that the 15 s continuous dual-frequency observation data of each station is accurate, so the sliding window time t_{op} is selected as 15 s. If dual-frequency observation data of the reference station is abnormal within 15 s, the data during this period cannot be used to generate the correct ionospheric compensation value, so that value of $bool(Statu_j(k))$ function for the 15 s is set to 0. If there is no abnormality in the dual-frequency observation data of the reference station within 15 s, value of $bool(Statu_j(k))$ function for the 15 s is set to 1.

6 Result Analysis

Take the sampling interval T as 1 s, and the sliding window time t_{op} as 15 s. Statistics on the continuity probability of one-month data of all monitoring stations, the statistical results are shown in Table 4. The epoch unit is second. From the definition of the data continuity probability, when the sliding window time t_{op} is 1 s and receiver data achieves ideal reception, data continuity probability will be equal to observation data availability. Therefore, data continuity probability is less than or equal to observation data availability. Statistics on the difference between the two indicators, the D-value is shown in Table 4. The larger the difference between observation data availability and data continuity probability, the higher probability of data interruption and the worse data quality.

It can be seen from the statistical results that the data continuity probabilities of 8 WAAS stations and 3 US CORS stations all exceed 0.99. However, the data continuity probability of the 4 CMONOC stations did not reach 0.99. The result shows that the continuity of WAAS station and US CORS station is good, and the continuity of CMONOC station is poor. It can be seen from the difference between the maximum and minimum data continuity probability that 8 WAAS stations are less than 0.5%, 3 US CORS stations are less than 0.8%, and 4 CMONOC stations reach 27%. It shows that the consistency of the receiver performance of WAAS station is the strongest, and the consistency of the receiver performance of CMONOC is weak. What's more, from the results of D-value, it can be seen that the WAAS station and the US CORS station are between 0.11% and 0.47%, while the CMONOC station is between 0.40% and 5.72%. It shows that the data of the station of CMONOC is more prone to interruption, while the data of the WAAS monitoring station and the US CORS station are not easily interrupted.

In summary, compared with the observational data availability indicator, the data continuity indicator considers the influence of cycle slips, the number of satellites available in a single epoch, and the ionospheric compensation value generated by VRS in network RTK every 15 s. Besides, the data continuity probability can be used to compare the differences in the data quality of the three types of monitoring stations. Therefore, when

Table 4. Data continuity and D-value of each station

Station name	Continuous sliding windows	Theoretical epochs	Data continuity probability (/s)	D-value
zau1	16811507	16835110	0.998598	0.001344
zkc1	16978428	17026419	0.997181	0.002656
zme1	16692489	16771473	0.995291	0.004425
zmp1	16801227	16908422	0.993660	0.004707
zfw1	17133409	17161034	0.998390	0.001429
zab2	17164660	17189438	0.998559	0.001354
zdv1	16984698	17036669	0.996949	0.002944
zob1	16524978	16593385	0.995877	0.001799
nvpo	19559835	19746162	0.990564	0.001797
alja	19178607	19213986	0.998159	0.001090
alla	19134760	19164553	0.998445	0.001186
GSTS	13405111	19133722	0.700601	0.057158
SXGX	15584279	17252686	0.903296	0.006262
YNTC	16240570	19274474	0.842595	0.003997
JSYC	19043588	19558778	0.973659	0.009735

evaluating the data quality of the monitoring station receiver, data continuity indicators should be added.

7 Conclusion

This paper uses data of three types of monitoring stations to carry out statistical analysis on the three indicators proposed in Beidou augmentation system reference station construction and acceptance specifications for evaluating quality of receiver data. It is found that the original indicators can only reflect the correctness and completeness of the monitoring station data, but cannot evaluate the ability to continue working. Therefore, this paper proposes a data continuity indicator for assessing continuous working capacity of monitoring station, and carries out a statistical analysis of the indicator. The results show that this indicator can reflect the difference between three types of monitoring stations and the consistency of receiver performance. Therefore, in order to provide services to highly dynamic users related to life safety, data continuity indicators should be added to the existing augmentation system specifications.

The deterioration of data quality is related to cycle slips, which are affected by factors such as multipath, interference, satellite power drop or increase, receiver reliability, and user movement. Because the terrain near each monitoring station and the electromagnetic interference received are different, it is necessary to conduct a specific analysis based on the situation of the monitoring station itself.

Finally, thanks to China Seismological Bureau's CMONOC data for providing data support for the creation of this paper.

References

1. China Satellite Navigation Office: BeiDou Navigation Satellite System Ground-based Augmentation Service Interface Control Document (Version 1.0) (July 2020). http://www.beidou.gov.cn/xt/gfxz/202008/P020200803362071652972.pdf
2. Li, Y., Li, J., Pan, L., et al.: Quality assessment of the Beidou-3 new signal B1C and B2a observation data. Adv. Earth Sci. **33**(11), 1161–1168 (2018)
3. Pan, L., Cai, C.: Software design and implementation for Beidou data quality analysis. Eng. Surv. Mapp. **23**(10), 67–71 (2004)
4. Feng, X., Jin, G., Fan, J., et al.: Experimentation and analysis of multipath effect in Pseudorange measurement of GNSS receiver. Mod. Electron. Tech. **5**, 77–81 (2013)
5. Xie, G.: Principles of GPS and Receiver Design, p. 93. Electronic Industry Press, Beijing (2009)
6. Ding, Y.: Research on Evaluation Methods for Multi-GNSS Data Quality and Integrity. Southeast University, Nanjing (2017)
7. Lu, Y.: Principle and Implementation Technology of Beidou/GPS Dual Mode Software Receiver, pp. 313–315. Electronic Industry Press, Beijing (2016)
8. U.S. Department of Defense: Global Positioning System Standard Positioning Service Performance Standard (2008)
9. Liu, S., Jia, X.: Algorithm and evaluation models of GNSS signal-in-space continuity. J. Navig. Position. **4**(01), 98–102 (2016)
10. Song, K., Zhou, S., Zhang, X.: The working principle of network RTK technology. Haihe Water Resour. **02**, 61–62 (2007)
11. Liu, Z., Gong, Y., Zhou, L.: Impact of China's high speed train window glass on GNSS signals and positioning performance. Satell. Navig. **1**(1), 1–16 (2020). https://doi.org/10.1186/s43020-020-00013-z