ORIGINAL ARTICLE

Open Access



Crowdsourcing RTK: a new GNSS positioning framework for building spatial high-resolution atmospheric maps based on massive vehicle GNSS data

Hongjin Xu^{1,2}, Xingyu Chen^{1,2}, Jikun Ou¹ and Yunbin Yuan^{1*}

Abstract

High-guality spatial atmospheric delay correction information is essential for achieving fast integer ambiguity resolution (AR) in precise positioning. However, traditional real-time precise positioning frameworks (i.e., NRTK and PPP-RTK) depend on spatial low-resolution atmospheric delay correction through the expensive and sparsely distributed CORS network. This results in limited public appeal. With the mass production of autonomous driving vehicles. more cost-effective and widespread data sources can be explored to create spatial high-resolution atmospheric maps. In this study, we propose a new GNSS positioning framework that relies on dual base stations, massive vehicle GNSS data, and crowdsourced atmospheric delay correction maps (CAM). The map is easily produced and updated by vehicles equipped with GNSS receivers in a crowd-sourced way. Specifically, the map consists of between-station single-differenced ionospheric and tropospheric delays. We introduce the whole framework of CAM initialization for individual vehicles, on-cloud CAM maintenance, and CAM-augmented user-end positioning. The map data are collected and preprocessed in vehicles. Then, the crowdsourced data are uploaded to a cloud server. The massive data from multiple vehicles are merged in the cloud to update the CAM in time. Finally, the CAM will augment the user positioning performance. This framework forms a beneficial cycle where the CAM's spatial resolution and the user positioning performance mutually improve each other. We validate the performance of the proposed framework in real-world experiments and the applied potency at different spatial scales. We highlight that this framework is a reliable and practical positioning solution that meets the requirements of ubiquitous high-precision positioning.

Keywords New GNSS positioning framework, Spatial high-resolution atmospheric delay correction, Crowdsourced atmospheric delay correction maps, Crowdsourced ionosphere, Crowdsourced troposphere, Ubiquitous

*Correspondence:

Yunbin Yuan

yybgps@asch.whigg.ac.cn

 ¹ State Key Laboratory of Geodesy and Earth's Dynamics, Innovation Academy for Precision Measurement Science and Technology (APM), Chinese Academy of Sciences (CAS), Wuhan 430071, China
 ² College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China

Introduction

Precise positioning services using Global Navigation Satellite Systems (GNSS) have become an indispensable part of our daily lives. This technology is widely used in various fields, such as disaster monitoring, autonomous driving, the Internet of Things, and public pedestrian navigation. It provides users with graded positioning services with accuracies ranging from millimeters to meters (Li & Chen, 2022; Liu et al., 2020; Leick et al., 2015; Li et al., 2022; Qin et al., 2022; Yang et al., 2021).



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Currently, two methods are predominantly used to provide real-time dynamic centimeter-level GNSS positioning services-traditional network RTK (Dingfa et al. Han, 1997; Hu et al., 2003; Vollath et al., 2000; Huang et al. 2022) and the recently developed PPP-RTK (Li et al., 2023; Teunissen & Khodabandeh, 2015; Wabbena et al., 2005; Zhang et al., 2022a, 2022b) technology. Both of these methods rely on Continuous Operational Reference System (CORS) base stations for high-precision positioning, which come with high construction and maintenance costs and are mainly funded by national or local governments. Moreover, the distribution of CORS base stations is relatively sparse, which means that both positioning methods require interpolation to obtain atmospheric delay correction data, which can be affected by the accuracy of the ionosphere during its active period. As a result, the interpolation accuracy can decline sharply, leading to poor positioning performance. These issues collectively lead to high costs and low user experience when using these services.

Autonomous vehicles have become more common and are now being mass-produced. These vehicles are often equipped with GNSS receivers to provide timing services and centimeter-level GNSS positioning services (Cao et al., 2022; Li et al., 2021; Qin et al., 2021; Tang et al., 2022), which provide more potential data sources for precise GNSS positioning. It is noted that the density of the road network is much greater than that of the CORS network. There is a clear imbalance in the effective use of on-board GNSS data on roads. For example, the city of Wuhan has more than 4 million privately owned cars (Security, 2022). During low and peak periods of vehicle travel, approximately 0.8-3.2 million vehicles drive on the road. The density of CORS stations is much less than one ten thousandth of the density of these vehicles. This imbalance shows that we urgently need more intelligent methods to improve the quality and coverage of precise GNSS positioning services while reducing related costs to meet the growing demand for mass precise positioning applications.

Crowdsourcing is a collaborative model that has proven highly effective in various domains, such as autonomous driving (Qin et al., 2021), map creation (Haklay, 2010; Wang et al., 2023), and earthquake warning (Minson et al., 2015). This approach is an innovative way to significantly reduce the expenses associated with gathering professional surveying data by leveraging nonprofessional crowdsourced surveying data. It provides real-time information that is both reliable and cost-effective (Liu et al., 2020).

The purpose of this study is to improve the costeffectiveness and efficiency of positioning services by overcoming the high cost and sparse distribution of traditional CORS base stations and the imbalanced effective utilization of onboard GNSS data on roads. To achieve this goal, we propose a new GNSS positioning framework that uses public GNSS observables to establish a spatial high-resolution atmospheric delay correction map. We also propose the Crowdsourced Atmospheric delay correction Map (CAM), which includes Crowdsourced Ionospheric and Tropospheric delay correction Maps (CIM and CTM). This positioning framework has several unique characteristics. First, it only requires two base stations. Based on this, a set of quality control methods is designed to obtain highprecision atmospheric information at specific times and places. Second, it gathers atmospheric information from the general public and constructs the CAM. Finally, the CAM will further augment the user positioning performance. With the application of this positioning framework, the accuracy of the CAM and the user location performance will mutually improve each other, forming a virtuous cycle that will meet the real-time precise positioning needs of autonomous driving users.

System overview

As illustrated in Fig. 1, our framework relies on collaboration among users as well as between users and the cloud, primarily comprising the following three parts. The first part involves the initialization of atmospheric information on the user end, where high-precision and reliable ionospheric and tropospheric delay correction is generated. The users solve three baseline parameters and validate their position and ambiguity closure using dual base stations. Subsequently, they upload the validated position along with atmospheric information to the cloud-server end.

In the second part, the cloud gathers atmospheric information uploaded by the general users, creates the CAM, and initializes the on-cloud atmospheric map.

In the third part, the CAM augments user positioning and is subsequently refined. A new positioning request is made by the user, and the cloud-server end provides the atmospheric information within the grid corresponding to the approximate position of the user. Based on this precise atmospheric delay correction information, the user end utilizes an ionosphere- and troposphereweighted model to achieve rapid and precise positioning. After the atmospheric information is validated through the ambiguity closure test, it is sent back to the cloudserver end for real-time updating of the CAM.

This iterative process creates a positive feedback mechanism that continuously improves the accuracy and spatial resolution of the CAM, maintaining the high precision of user positioning.



Fig. 1 Illustration of the system structure. The system consists of three parts. The first part is the atmosphere initialization for individual vehicles. The second part is the atmosphere initialization for the cloud-sever end. The third part is the CAM augmenting user positioning and CAM refinement

Atmospheric delay initialization for individual vehicles

Single-baseline RTK functional model

In our positioning framework, we add tropospheric constraints to the traditional long-baseline ionosphereweighted single-differenced (SD) RTK function model (Odolinski et al., 2015) to solve the three baselines (two base stations and one user) separately. The specific form is given as:

$$P_{ru,j}^{s_{*}} = -\kappa_{u}^{s_{*}T} \Delta \mathbf{p}_{ru} + m_{u}^{s_{*}} T_{ru} + dt_{ru,P_{j}}^{*} + \mu_{j_{*}} I_{ru}^{s_{*}}$$

$$L_{ru,j}^{s_{*}} = -\kappa_{u}^{s_{*}T} \Delta \mathbf{p}_{ru} + m_{u}^{s_{*}} T_{ru} + dt_{ru,L_{j}}^{*} - \mu_{j_{*}} I_{ru}^{s_{*}} + \lambda_{j_{*}} N_{ru,j_{*}}^{1s_{*}}$$

$$I_{ru}^{s_{*}} = \left(I_{ru}^{s_{*}}\right)_{prior}$$

$$T_{ru} = (T_{ru})_{prior}$$

where $P_{ru,j}^{s_*}$ and $L_{ru,j}^{s_*}$ represent the SD code and phase observables, respectively. $s_* = 1_*, 2_*, \ldots, m_*$ denotes the tracked satellites; * represents the tracked systems, with G for GPS, E for Galileo, and B for BDS; u and r represent the receivers of the baseline; $j_* = 1_*, \ldots, f_*$ denotes the tracked frequencies; $\kappa_u^{s_*}$ is the line-of-sight unit vector; and $\Delta \mathbf{p}_{ru}$ denotes the coordinate increment of receiver r. dt_{ru,P_j}^* and dt_{ru,L_j}^* represent receiver clock error terms that absorb pseudorange hardware and phase hardware delays, respectively. $I_{ru}^{s_*}$ represents the relative slant ionospheric delay, with a corresponding coefficient $\mu_{j_*} = \lambda_j^2/\lambda_{1}^2$, λ_j is the wavelength for frequency j_* ; and T_{ru} and $m_u^{s_*}$ represent the relative wet Zenith Tropospheric Delay (ZTD) and corresponding mapping function, respectively. $N_{ru,j_*}^{1s_*}$ represents double-differenced integer ambiguities, and $(I_{ru}^{s_*})_{prior}$ and $(T_{ru})_{prior}$ represent the priors of $I_{ru}^{s_*}$ and T_{ru} , respectively.

Single-baseline RTK stochastic models

The variance–covariance (VCV) matrix for the combined GPS/Galileo/BDS ionosphere-weighted and troposphere-weighted model is given as

$$\overline{Q}_{yy} = blkdiag\left(Q_{yy}^G, Q_{I_{ru}}^G, Q_{yy}^E, Q_{I_{ru}}^E, Q_{yy}^B, Q_{I_{ru}}^B, \sigma_{T_{ru}}^2\right)$$
(2)

where $\sigma_{T_{ru}}^2$ is the variance of $(T_{ru})_{prior}$ and $Q_{I_{ru}}^*$ represents the single-system VCV matrix of $(I_{ru}^{s_*})_{prior}$. Q_{yy}^* represents the single-system VCV matrix of the SD code and phase observables and is given as

$$Q_{yy}^* = blkdiag\left(C_p^*, C_\phi^*\right) \otimes \left(D_n^T D_n \otimes W_{m_*}^{-1}\right)$$
(3)

where the code and phase observable a priori variance factors are given by submatrices $C_P^* = diag(\sigma_{P_{1,*}}^2, \ldots, \sigma_{P_{f_*}}^2)$ and $C_L^* = diag(\sigma_{L_{1,*}}^2, \ldots, \sigma_{L_{f_*}}^2)$, respectively, and *blkdiag* and *diag* denote the block diagonal and diagonal matrix, respectively. Here, we assume no cross-correlation between the code and phase or between frequencies. Furthermore, \otimes is the Kronecker product, D_n^T is the between-station SD operator and $W_{m_*}^{-1}$ is the elevation-dependent weighting function.

This subsection describes the stochastic models for the SD code and phase observables, and later subsections describe the stochastic models for the SD ionospheric and tropospheric delays under different conditions.

Generating atmospheric information

with an ambiguity-fixed solution for individual vehicles

During the initialization process, no prior atmospheric information is provided by the cloud server. For individual vehicles,

$$(I_{ru}^{s_*})_{prior} = 0$$

$$(T_{ru})_{prior} = 0$$

$$(4)$$

The corresponding stochastic model is given as

$$Q_{I_{ru}}^{*} = diag \left(\sigma_{I_{ru}}^{2} \otimes W_{m_{*}}^{-1} \right)$$

$$\sigma_{T_{ru}}^{2} = const$$
(5)

where $\sigma_{I_{ru}^{s*}}^2$ is the SD variance factor of the slant ionospheric delays that can be modeled as a function of the baseline length. $\sigma_{T_{ru}}^2$ is a larger constant, which is equivalent to white noise estimation. We utilize (1)-(5) to obtain the three baseline solutions constructed by the two base stations and the user and attempt to fix the ambiguity.

Atmospheric quality control based on the ambiguity closure test

Once the ambiguities are fixed, the mutual difference between the user's positions calculated by the dual base stations and the ambiguity closure of the three baselines are tested. The expression of the ambiguity closure test is given as

$$\Delta \nabla N_{r_1 u, j}^{1s_*} + \Delta \nabla N_{u r_2, j}^{1s_*} + \Delta \nabla N_{r_2 r_1, j}^{1s_*} = 0$$
(6)

where r_1 , r_2 , and u represent the dual base stations and the user, respectively.

If the test is passed, the user's position, along with ionospheric and tropospheric information (including the adjustment value and posterior variance), is uploaded to the cloud server end.

Positioning after individual vehicle initialization

After the initialization for individual users, two methods for accessing atmospheric information to augment positioning are available. The first method is to employ the user's own historical atmospheric information weighting via random walk estimation,

where t_{k-1} represents the previous epoch, and the corresponding stochastic model is given as

$$Q_{I_{ru}}^{*} = diag\left(\sigma_{I_{ru}}^{2}\right)$$

$$\sigma_{I_{ru}}^{2} = \left(\sigma_{I_{ru}}^{2}\right)_{t_{k-1}} + prn_{I_{ru}}^{2} \cdot \Delta t$$

$$\sigma_{T_{ru}}^{2} = \left(\sigma_{T_{ru}}^{2}\right)_{t_{k-1}} + prn_{T_{ru}}^{2} \cdot \Delta t$$
(8)

where $(\sigma_{I_{ru}}^2)_{t_{k-1}}$ and $(\sigma_{T_{ru}}^2)_{t_{k-1}}$ represent the posterior SD slant ionospheric delays and relative wet ZTD of the previous epoch, $prn_{I_{ru}}^{s_*}$ and $prn_{T_{ru}}$ are the corresponding process noises, and Δt is the time interval between epochs.

The second method utilizes atmospheric information downloaded from the cloud-sever end to weight:

where $(I_{ru}^{s_*})_{t_{k-1}}^{cloud}$ and $(T_{ru})_{t_{k-1}}^{cloud}$ represent the SD slant ionospheric delays and relative wet ZTD returned from the cloud, respectively. The corresponding random model is:

$$Q_{I_{ru}}^{*} = diag \left(\sigma_{I_{ru}^{*}}^{2}, \dots, \sigma_{I_{ru}^{*}}^{2} \right)$$

$$\sigma_{I_{ru}}^{2} = \left(\sigma_{I_{ru}^{*}}^{2} \right)_{t_{k-1}}^{cloud} + prn_{I_{ru}^{*}}^{2} \cdot \Delta t$$

$$\sigma_{I_{ru}}^{2} = \left(\sigma_{I_{ru}}^{2} \right)_{t_{k-1}}^{cloud} + prn_{I_{ru}}^{2} \cdot \Delta t$$

(10)

where $(\sigma_{I_{ru}}^2)_{t_{k-1}}^{cloud}$ and $(\sigma_{T_{ru}}^2)_{t_{k-1}}^{cloud}$ represent the posterior SD slant ionospheric delays and relative wet ZTD returned from the cloud server end, respectively.

Users can achieve a single-epoch ambiguity-fixed solution and obtain centimeter-level positioning accuracy by using (1)-(3) and (7)-(8) or (1)-(3) and (9)-(10).

CAM initialization for the cloud-sever end Grid division and interpolation

As shown in Fig. 2, to create an atmospheric map for a certain period, we divide the service area into grids of a certain size, such as $1 \text{ km} \times 1 \text{ km}$, which can be adjusted according to the actual requirements. The atmospheric information provided by the user is comprehensively processed, and the information within each grid is

obtained. However, we may not be able to obtain all the atmospheric delay information within the grid initially. Therefore, we use interpolation algorithms such as LIM (Han, 1997) to fill in the gaps and interpolate all the atmospheric information within each grid based on the existing data. This helps us create the initial CAM and complete the initialization process.

Storage and communication of the CAM

Taking Wuhan city as an example, the city spans 132.12 km from east to west and 153.97 km from north to south. It can be roughly divided into a grid of 133×154 (1 km \times 1 km) squares, each of which has four float attributes: the values and variances of the single-differenced ionospheric delay and single-differenced tropospheric delay. The storage requirement for individual grid data is 16 bytes. At the same time, approximately 20 observable satellites contribute to the creation of the CAM, with a data size of approximately 12 MB. These data are stored on the cloud-server end.

The user end and the cloud-server end communicate bidirectionally to exchange information, enabling the server to provide the atmospheric delay correction information as real-time feedback to the users at the grid's approximate location.



Fig. 2 Initialization of cloud atmospheric information

Initialization time

The initialization process can be completed rapidly in practice. First, in terms of spatial distribution, users near the base station can easily obtain an ambiguity-fixed solution (Zhu et al., 2023). This means that atmospheric information in that area can be readily initialized and then extrapolated to provide priors for other newly added users. As the system iteration proceeds, a dense atmospheric map can quickly be constructed.

Second, in terms of temporal distribution, for vehicle users, low-traffic periods typically occur in the morning when ionospheric activity tends to be stable, and the ground truth of the SD ionospheric delay is often near zero, making it easier to perform AR and generate reliable atmospheric delay information.

CAM-augmented user positioning and CAM refinement

New users download atmospheric delay correction information corresponding to their approximate positions within the grid from the cloud-server end. By utilizing Eqs. (1)-(3), (9) and (10) to solve the long-baseline RTK equations with atmospheric constraints, a single-epoch ambiguity-fixed solution can be obtained. Subsequently, the atmospheric information, validated through ambiguity closure checks, is uploaded back to the cloud.

During system iteration, the crowdsourced ionospheric map and crowdsourced tropospheric map will undergo stepwise refinement, leading to the establishment of a more detailed atmospheric delay correction map. Figure 3 illustrates this process.



Fig. 3 CAM refinement and user positioning with CAM augmentation

Experimental results and discussion

We first describe the experimental setup in "Experimental setup" section. Then, "Quality of vehicle GNSS data" and "Assessment of the atmospheric quality generated by long-baseline dynamic users" sections evaluate the various stages of the initialization process, including the data quality, position accuracy, AR, and atmospheric solution accuracy. In "Dynamic baseline (40–300 km) generation of the CAM" section, we extend the baseline to 300 km to explore the spatial applicability of the framework. In the "Simulation experiment based on CAM-augmented positioning" and "Multi-vehicle based augmentation in a real environment" sections, we conduct simulation and multi-vehicle based experiments on CAM-based localization after initialization.

Experimental setup

The data were gathered on May 12th, 2023, in Wuhan, covering the urban area from the Second to the Fourth Ring Road. The vehicle test lasted approximately 6 h, and the test trajectory and base station distribution are displayed in Fig. 4. The distance between base stations ranged from 40 to 50 km, and the distance between the base stations and the test trajectory ranged between 0–60 km. We used the traditional network RTK algorithm to calculate the trajectory and obtain the ground truth. During the experiment, we selected dual base stations to create a dynamic triangulation network with onboard users for dynamic long-baseline RTK

positioning experiments. Specifically, we chose UBAS and UBHH as the dual base stations for the new positioning framework as the basis for further experiments.

The specific configuration of the test equipment and scenarios is shown in Fig. 5. During the test, we used two GNSS receivers of varying cost—the geodetic receiver Trimble Net-R9 (R9) and the low-cost receiver Septentrio Mosaic X5 (MSX5). The data sampling rate was set to 1 Hz, and we validated the new GNSS positioning framework proposed in this paper using dual-frequency GPS and Galileo data. The test covered multiple typical scenarios, such as elevated roads with broad views, urban roads, and road sections obstructed by trees, tall buildings, and elevated bridges.

The strategies for primary data processing are listed in Table 1. During the initialization process of the system, the study considered only the continuity of ambiguity between epochs, which was estimated as a random walk process. The atmospheric parameters were given a zero mean a priori and were estimated as white noise due to the poor quality of dynamic data compared to static data. This strategy resets the atmospheric information at each epoch. Once the ambiguity closure test, based on dual base stations, is passed, the atmospheric delay at that location is uploaded to the cloud server end to complete the initialization of the CAM. Afterward, users within the service area use the atmospheric delay from the cloud as a prior for localization. If the initialization is not completed and the atmospheric parameters are estimated as



Fig. 4 Vehicle experiment test trajectory and base station distribution



Fig. 5 Test equipment (top) and typical test scenarios (bottom)

Table 1 Data processing strategies used in this study

ltem	Initialization		Initialized		
Baseline length	<100 km	>100 km	<100 km	>100 km	
Precise ephemeris	No	Yes	No	Yes	
lonospheric prior	$\Delta l = 0$		$\Delta l = \Delta l_{cloud}$		
Tropospheric prior	$\Delta T = 0$		$\Delta T = \Delta T_{cloud}$		
Frequency	GPS L1&L2 GAL E1&				
Estimator	Kalman filter				
Constraints between epochs	Ambiguity Ambiguity and atmosphere				
Elevation mask	15°				
SNR mask	35 dBHz				

a random walk process, the atmospheric delay calculated from the subsequent-epoch solutions will become less accurate if the estimation of the atmospheric delay in the first epoch deviates significantly from the actual value.

Quality of vehicle GNSS data

This subsection elaborates on satellite visibility, PDOP, and multipath error to assess the feasibility of creating a crowdsourced atmospheric map.

The number of satellites that are available and their corresponding PDOP values have a significant impact on the position and AR performance. During a test conducted with elevation cutoff angles of 15° and 40°, 97.67% and 95% of the periods, respectively, had 10 or more visible satellites, as shown in the first row of Fig. 6. In cases where fewer satellites are visible, there may be severe GNSS electromagnetic signal occlusion under the

elevated bridge, but at the same time, there will be a wide field of view on the elevated bridge. The city's ring roads, elevated roads, and other wide roads provide a good spatial distribution of road networks and GNSS observation conditions. This means that in most cases, we will have sufficient satellites for positioning, which creates favorable conditions for establishing a CAM.

Based on the first two rows of Fig. 6, as the elevation cutoff angle increases, the number of available satellites decreases gradually, while the corresponding PDOP increases. Table 2 also shows the HDOP and VDOP values of multiple GNSSs at elevation cutoff angles ranging from 15° to 40°. Notably, as the satellite elevation increases, the DOP value gradually increases. At an elevation cutoff angle of 40°, the HDOP is still only 1.07, whereas the VDOP is 2.85. Thus, at higher elevation cut-off angles, the positioning accuracy of the GNSS in the



Fig. 6 The first and second lines show the numbers of visible satellites and PDOP at an elevation cutoff angle of 15° to 40° for the rover station, respectively. The third line is the multipath error of the base station and the rover station

Table 2	Percentage of vis	ible satellites (NSat \geq	10) and average DOP	at elevation cutoff angles of 15°	' to 40° for rover stations
---------	-------------------	------------------------------	---------------------	-----------------------------------	-----------------------------

Elevation mask	15°	20°	25°	30°	35°	40 °
Ratio (NSat≥10)	97.67%	97.32%	97.12%	96.74%	95.85%	95.00%
PDOP	1.45	1.58	1.82	2.21	2.58	3.03
HDOP	0.69	0.72	0.77	0.87	0.95	1.07
VDOP	1.27	1.41	1.66	2.03	2.41	2.85

horizontal direction is less affected, and the impact in the vertical direction is more significant.

In Fig. 6, the third row displays the multipath errors of the R9 and MSX5 receivers and two base stations. The data suggest that the multipath error of static base stations is much lower than that of dynamic users. In dynamic environments, the multipath errors for R9 and MSX5 are approximately the same. This indicates that both R9 and MSX5 can be used to establish a CAM.

It is important to understand that in situations where the availability of GNSS signals is limited, such as in tunnels or in areas with high multipath effects such as urban canyons, even short-baseline RTK positioning can be difficult to achieve. The new GNSS positioning framework that we propose is an exclusive satellite navigation system positioning method, and its highest positioning performance is that of short-baseline RTK positioning. Therefore, the discussion of this study is limited to scenarios where GNSS signals are available and of good quality and where short-baseline RTK positioning can be performed. Situations in which GNSS signals are completely blocked, the signal quality is low, or short-baseline RTK positioning cannot be completed are not within the scope of this study.

Performance analysis of dynamic long-baseline positioning

In this section, we first focus on position accuracy, which is of utmost concern to the general public. We plot two types of receivers to show the difference in position between epochs 442,500 and 444,500, with a sequence of 2000s. Figure 7 shows the epochs that match well; the position difference calculated near epoch 443,100 is relatively large and is not shown. We evaluated the positions of two receivers under long-baseline dynamic positioning and found that both types of receivers, which have different costs, exhibited slight errors compared with the ground truth. Taking the dynamic long baselines constructed by R9, MSX5, and UBAS as examples, the positioning accuracies (STDs) in the three ENU directions are (1.1, 1.18, 4.84) cm and (0.91, 1.06, 4.74) cm, respectively, as shown in the first row of Fig. 7. The third row of



Fig. 7 Differences in dynamic long-baseline positioning. The first two lines show differences between the long-baseline solutions and network RTK of two rover stations and dual base stations. The third line calculates the mutual differences between the same rover stations from the baselines of the dual base stations

Fig. 7 shows that for a single user, the position difference of the baseline solutions of the dual base stations is relatively small. The mutual difference accuracies of the two types of receivers are (0.89, 0.83, 2.18) cm and (1.02, 1.07, 3.19) cm, respectively. The difference in the single-baseline solutions between R9 and MSX5 is also small, indicating that both types of receivers perform consistently in long-baseline dynamic positioning.

In addition, by comparing the first and third lines of Fig. 7, we observe that the dynamic long-baseline solution provides better consistency in the vertical direction than does the traditional network RTK solution. Although the network RTK user-end model has fewer parameters to estimate, its redundancy is the same as that of the atmosphere-weighted long baseline model. Moreover, its user-end model loses accuracy information when utilizing atmospheric information provided by the server, and the interpolation error of the atmospheric delay is absorbed by the position, resulting in slightly poorer consistency compared to the long baseline atmosphere-weighted model.

This experiment verified the reliability of the dynamic long-baseline positioning of dual base stations. The results further support the potential of using massive vehicle user data to generate atmospheric information. It is reasonable for us to use the mutual difference in the same user's position calculated by dual base stations as the basis for establishing a dynamic reference network. This process will aid in AR and the creation of high-quality atmospheric maps.

Performance of dynamic long-baseline AR

Correct AR is a prerequisite for achieving high-precision positioning at the centimeter level for the GNSS. We plotted the position difference, ADOP, and ratio sequences of the two types of receivers between epochs 442,500 and 442,800 with a sequence of 300 s in Fig. 8. The graph shows that there are three significant fluctuations in the position differences. The ADOP and ratio indicators show significant changes in the first two areas, while the fluctuations are insignificant in the third area. This makes it challenging for dynamic positioning users to determine whether the ambiguity is correctly fixed based on the theoretical indicators of the ADOP and ratio. For this reason, we introduced dual base stations. This allows us to test the user positioning accuracy directly through mutual position differences. Additionally, the ambiguity parameter in the equation of the phase



Fig. 8 Theoretical indicators of the AR and position differences

observables is closely correlated with the ionospheric parameters. Incorrect AR will result in corresponding errors in the ionospheric parameters. Therefore, we examine the ambiguity closure of the dynamic triangulation network to determine whether the AR is correct and then identify and eliminate erroneous ionospheric parameter estimations.

In traditional network RTK or PPP-RTK, the server side relies on the correlation between ambiguity and ionospheric delay to refine ionospheric information for initialization. Additionally, the user end uses interpolated atmospheric delay correction to achieve fast integer ambiguity resolution without contributing to the server's atmospheric delay refinement. In contrast, the new GNSS positioning framework enables the user to use the server's atmospheric delay information to achieve instantaneous integer ambiguity resolution while also contributing to the refinement of the server's atmospheric delay. This approach allows us to provide users with atmospheric correction services that offer higher spatial resolution and greater accuracy than interpolated atmospheric correction by CORS stations.

Assessment of the atmospheric quality generated by long-baseline dynamic users

We chose UBAS as the base station, while R9 and MSX5 were selected as rover stations. We then calculated

ambiguity-fixed baseline solutions for UBAS-R9 and UBAS-MSX5 separately. Afterward, we differenced the baseline solutions for UBAS-R9 and UBAS-MSX5, which included the position, single-differenced ionospheric delay, and single-differenced tropospheric delay. The results are shown in Fig. 9.

First, the position estimation of dynamic long baselines can converge steadily with centimeter-level accuracy. Second, the ionospheric mutual differences corresponding to most satellites are within 2 cm, but there are still a few cases that exceed this range. This is because the current criterion used to determine the fixed solution of the ionospheric delay relies only on the theoretical index (ratio), which is not entirely reliable. We will further provide more accurate criteria to determine whether the ionospheric delay was correctly estimated. Finally, the mutual differences in the tropospheric delay show a relatively stable trend, which is maintained at the millimeter level. This is due to the correlation between the troposphere and each observed satellite, which has good observability and high redundancy.

In addition, in the static scenario of nearly 200 s between epochs 43,500 and 43,700, there is a smaller difference between the position and atmospheric parameters, which can be effectively attributed to typical vehicle scenarios: users are waiting at traffic lights or experiencing traffic congestion, which provides favorable



Fig. 9 Consistency between the single-differenced ionosphere and troposphere under various dynamic conditions



Fig. 10 Single-differenced ionospheric mutual differences calculated by UBAS-R9 and UBAS-MSX5, with only AR (left), the constraint of dual base stations obtaining user position mutual differences (middle), and the ambiguity closure constraint formed by dual base stations and users (right)

observation conditions for obtaining high-precision position and atmospheric delay information.

Figure 10 further illustrates the agreement in the single-differenced ionosphere delay between R9 and MSX5 under different judgment criteria. It can be observed that position constraints do not have a significant effect on the localization and removal of ionospheric gross errors, but ambiguity closure can effectively eliminate satellites with ionospheric solution errors. Taking satellite G08 as an example, combined with the data in Table 3, the percentage of epochs that passed the ambiguity closure test (i.e., the success rate of ambiguity resolution, calculated as the number of epochs that passed the ambiguity closure test divided by the fixed number of epochs) between MSX5 and the dual base stations is 66.42%; in contrast, the percentage for R9 is only 20.83%, which is almost not closed around 443,500 epochs. Therefore, there is a significant difference in the estimated ionospheric delay values around this period. After screening for ambiguity closure, the standard deviation (STD) of

Satellite PRN	MSX5 L1/E1	MSX5 L2/E5b	Success Rate	R9 L1/E1	R9 L2/E5b	Success rate
G01	_	_	_	_	_	_
G03	1323/1770	1341/1765	75.36%	1257/1803	1204/1803	68.25%
G07	1135/1765	1123/1765	63.97%	1481/1803	1470/1803	81.84%
G08	755/1142	754/1130	66.42%	167/1253	355/1253	20.83%
G14	1601/1808	1577/1765	88.95%	1704/1817	1684/1803	93.59%
G17	1219/1765	1210/1765	68.81%	1076/1803	1083/1803	59.87%
G19	93/788	77/746	11.06%	53/752	51/671	7.32%
G21	1315/1765	1318/1765	74.59%	1521/1803	1480/1803	83.22%
G30	1170/1765	1144/1765	65.55%	1428/1803	1376/1803	77.76%
E13	-	-	-	-	-	-
E21	1579/1609	1574/1609	97.98%	1679/1779	1540/1754	91.09%
E26	1574/1609	1569/1609	97.67%	1608/1754	1499/1754	88.57%
E27	1251/1619	1250/1609	77.48%	1094/1754	1072/1754	61.74%
E33	348/1114	344/1091	31.38%	209/1097	233/1072	20.39%

 Table 3
 Percentage passed the ambiguity closure test within 2000 epochs



Fig. 11 Effect of different single-differenced ionospheric noise on ambiguity resolution

the single-differenced ionospheric parameters calculated by R9 and MSX5 is 0.74 cm.

To demonstrate how the single-differenced ionosphere delay affects ambiguity resolution, additional simulation experiments are included. An ultrashort baseline comprising CUTA and CUT0, which was collected at Curtin University on January 2, 2023 (with an ionospheric ground truth of 0), was chosen as an example. Figure 11 illustrates the impact of different levels of single-differenced ionospheric noise on the ADOP. Based on prior experience, an ADOP of less than 0.12 cycles corresponds to a success rate of ambiguity resolution greater than 99.9% (Odijk & Teunissen, 2008). A single-differenced ionospheric error of 5 cm is required for ambiguity resolution, while the ionospheric accuracy generated by our dynamic long baseline reaches 0.74 cm. Such high-precision ionospheric information can assist in instantaneous ambiguity resolution for almost any long baseline.

Thus far, we have generated a centimeter-level betweenstation single-differenced ionospheric delay and a

SAT PRN	MSX5-40 km (%)	R9-40 km (%)	MSX5-200 km (%)	R9-200 km (%)	MSX5-300 km (%)	R9-300 km (%)	Elevation (deg)
G03	75.36	68.25	54.56	49.60	41.52	48.04	25
G07	63.97	81.84	48.39	67.84	32.54	40.47	35
G08	66.42	20.83	13.38	9.86	0.56	6.97	15
G14	88.95	93.59	67.51	82.62	45.65	56.58	55
G17	68.81	59.87	51.70	45.34	29.27	40.48	35
G19	11.06	7.32	7.82	2.76	2.69	5.26	15
G21	74.59	83.22	56.40	77.27	26.92	41.66	35
G30	65.55	77.76	48.27	65.42	23.86	36.06	50
E21	97.98	91.09	84.47	88.22	49.25	66.86	75
E26	97.67	88.57	72.28	81.73	48.87	60.58	70
E27	77.48	61.74	49.19	47.16	30.91	40.05	25
E33	31.38	20.39	36.33	28.58	28.01	32.17	15

Table 4 🛛	Ambiguity	resolution	success	rates fo	r dv	vnamic l	long	baselines	from	40 tc	300	km
						/						



Fig. 12 Statistics of ambiguity resolution success rates for dynamic long baselines from 40 to 300 km

millimeter-level between-station single-differenced tropospheric delay using vehicle dynamic data.

Dynamic baseline (40-300 km) generation of the CAM

This section investigates the feasibility of generating a CAM using dynamic long baseline data of over 100 km. We achieved this by increasing the distance between base stations and dynamic users. We selected base station data that were collected at distances of 200 km and 300 km from the motion trajectory and used precise ephemeris data for long-baseline solutions. We evaluated the most crucial factor of correct AR, as illustrated in Fig. 12 and Table 4.

The data presented in Fig. 12 reveal two clear trends. First, as the satellite elevation increases, the visibility of the satellite improves significantly, resulting in an increase in the success rate of ambiguity resolution. Second, as the baseline length increases, the success rate generally decreases, although there are some exceptions for the G19 and E33 satellites at low elevations. These two trends suggest that while fixing the dynamic long baseline ambiguities, it is essential to prioritize selecting data with high satellite elevations. This would increase the accuracy and stability of atmospheric parameter solutions.

When the baseline distances are set at 40 km and 200 km and the elevation is 25 degrees or greater, the success rate is greater than 50%. In contrast, for a 300 km baseline, within the same elevation range, the success rate varies from 25% to 50%. Importantly, the success rates for

the low-cost receiver MSX5 and the geodetic receiver R9 are similar, which provides valuable reference for popularizing device selection for the general public in practical applications.

The core of the new framework is that it accurately calculates the atmospheric parameters of a low-cost user dynamic long baseline to complete system initialization. By utilizing the correlation between ambiguity and the ionospheric delay, we can transform this problem into that of answering two key questions: whether the ambiguity can be resolved and whether it is resolved correctly. Figure 12 shows that a dynamic long baseline of less than 300 km has the possibility of correct AR, and the success rate exceeds 25% at a cutoff elevation of 25°. Although this percentage may seem relatively low for individual dynamic users, there is enough redundant information for overall mass users. By processing and merging data effectively, we can enhance the overall quality and coverage of the CAM.

Page 15 of 18

Simulation experiment based on CAM-augmented positioning

The experiments described above verify the initialization of the system. The results show that 12 satellites continuously observed by dynamic users can be combined with dual base stations ambiguity closure tests to generate high-precision atmospheric delay correction information. This section assesses whether the generated atmospheric information can be used to improve the positioning performance and assesses the ambiguity resolution of users for a single-epoch solution. The primary experimental settings are shown in Table 5. Other parameters, such as the elevation and SNR masks, remain the same as those in Table 1.

Figure 13 shows the position, ADOP, and ratio sequences of UBHH-R9 under different atmospheric constraints. The results indicate that simultaneously constraining ionospheric and tropospheric delays is the most effective approach. This suggests that our proposed framework can utilize low-cost receivers for generating atmospheric products, enabling fast integer ambiguity

Table 5 Main data processing strategies

Items	Scheme 1	Scheme 2	Scheme 3	Scheme 4		
Estimator	Single-epoch least squares					
Constraints between epochs	None					
Base station	UBHH (approximately 40 km away from the user)					
Generation device	MSX5					
Augmentation device	Net-R9					
Atmospheric augmentation	lon and trop	lon	Trop	None		



Fig. 13 Single-epoch ambiguity-fixed position differences, ratio, and ADOP sequences of UBHH-R9 for four schemes



Table 6 Main strategies for multi-vehicle based augmentation

Item	Strategies
Dual base stations	Base1: HB01 (approximately 30 km); Base 2: HB05 (approximately 40 km)
Generation device	Car 2: MSX5
Augmentation device	Car 1: NovAtel PwrPak7
Ground truth (base)	JFNG (within 5 km)
Ground truth (rover)	Car 1: SPAN-SE + iMAR-FASA

resolution in precise positioning. Constraining only the ionosphere also yields good positioning augmentation, while there is almost no significant gain when only the tropospheric delay is constrained.

We observed noteworthy fluctuations in the position and ADOP sequences, particularly at approximately 443,800 s, coinciding with the transition of the vehicle's dynamic conditions from static to dynamic. This phenomenon was particularly pronounced when employing only the ionospheric constraint strategy. The observed instability can be attributed to the smaller number of prior parameters in this scheme compared to that of scheme one, leading to challenges when transitioning between epochs. Hence, in dynamic environments, it becomes imperative to meticulously consider and optimize the configuration of prior parameters to ensure the stability of the position solution.

Multi-vehicle based augmentation in a real environment

We further conduct multi-vehicle experiments to evaluate the enhanced positioning effects of atmospheric information generated by different vehicles. The experimental setup is consistent with real-world applications.

The data were collected on March 13, 2024. The vehicle trajectory was within 5 km of the reference station JFNG, and the ground-truth trajectory was provided by short-baseline RTK/INS. Dual base stations HB01 and HB05, located 30 km and 40 km away from the vehicle, respectively, were selected. The experimental configuration is shown in Fig. 14, and the strategy settings are detailed in Table 6. We continue to compare the four atmospheric augmentation schemes from "Simulation experiment based on CAM-augmented positioning" section.



Fig. 15 Single-epoch ambiguity-fixed position differences, ratios, and ADOP sequences of the HB01-PwrPak7 for the four schemes

The correct AR is determined by validating the atmospheric information through the ambiguity closure test in the dynamic triangulation network constructed by MSX5, HB01, and HB05. The corresponding atmospheric information is then used to enhance the positioning of Car 1's NovAtel PwrPak7. Figure 15 indicates that the atmospheric information we generated can effectively assist in resolving ambiguities in long-baseline RTK positioning, thus achieving centimeter-level positioning.

Conclusion and future work

This paper presents a novel GNSS positioning framework based on massive vehicle GNSS data. The framework aims to establish a spatial high-resolution CAM to improve the performance of GNSS positioning in complex atmospheric environments. To ensure the correctness of the ambiguity resolution and accurate ionospheric delay generation for dynamic long baselines, we introduce dual base stations. The CAM augments positioning and CAM generation with dynamic long baselines ranging from 40 to 300 km verify the positioning performance of the framework and applied potency at different spatial scales. Overall, this framework offers a viable solution for generating low-cost, high-resolution spatial atmospheric correction information within a local area.

Despite the highly promising results, various issues will deserve attention to further improve the applications and performance of the new framework. These include the generation of crowdsourced undifferenced atmospheric delay correction maps, which are expected to provide unprecedented spatial high-resolution atmospheric delay correction information for ionospheric and tropospheric monitoring research. Concerning the high storage requirements for atmospheric maps on the cloud server end, in the future, compression algorithms such as compressive sensing (Candès et al., 2006) will be used to efficiently store historical atmospheric information on the cloud server end, which will provide data for further scientific research. Finally, the criteria for grid partitioning need to be further refined using techniques such as grid clustering (Schikuta, 1996).

Acknowledgements

The authors would like to thank IGS for providing precise ephemeris products and CHCNAV for providing the base station data.

Author contributions

HX proposed the new GNSS positioning framework and conducted experiments and analyses under the supervision of JO and YY. XC provided the trajectories of the network RTK solutions. All authors were involved in writing the paper, conducting the literature review, and discussing the results.

Funding

This research is funded by the National Key R&D Program of China (NO. 2022YFB3903903) and the National Natural Science Foundation of China (NO. 41974008, NO. 42074045).

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

Yunbin Yuan is an editorial board member for Satellite Navigation and was not involved in the editorial review or decision to publish this article. All authors declare that they have no competing interests.

Received: 22 January 2024 Accepted: 9 April 2024 Published online: 13 May 2024

References

- Candès, E. J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory, 52*(2), 489–509.
- Cao, S., Lu, X., & Shen, S. (2022). GVINS: Tightly coupled GNSS-visual-inertial fusion for smooth and consistent state estimation. *IEEE Transactions on Robotics*, 38(4), 2004–2021.
- Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703.
- Han, S. (1997). Carrier phase-based long-range GPS kinematic positioning. UNSW Sydney.
- Huang, D., Feng, W., Li, J., Gong, X., & Li, S. (2022). Grid VRS technique of real time high-precision location-based service. Acta Geodaetica et Cartographica Sinica, 51(8), 1717–1724.
- Hu, G., Khoo, H., Goh, P., & Law, C. (2003). Development and assessment of GPS virtual reference stations for RTK positioning. *Journal of Geodesy*, 77, 292–302.
- Leick, A., Rapoport, L., & Tatarnikov, D. (2015). GPS satellite surveying. Wiley.
- Li, B., & Chen, G. (2022). Precise cooperative positioning for vehicles with GNSS and INS integration. *Acta Geodaetica Et Cartographica Sinica*, *51*(8), 1708–1715.
- Liu, J., Guo, W., Guo, C., Gao, K., & Cui, J. (2020). Rethinking ubiquitous mapping in the intelligent age. Acta Geodaetica Et Cartographica Sinica, 49(4), 403–406.
- Li, X., Han, J., Li, X., Huang, J., Shen, Z., & Wu, Z. (2023). A grid-based ionospheric weighted method for PPP-RTK with diverse network scales and ionospheric activity levels. *GPS Solutions*, 27(4), 191.
- Li, X., Huang, J., Li, X., Shen, Z., Han, J., Li, L., & Wang, B. (2022). Review of PPP-RTK: Achievements, challenges, and opportunities. *Satellite Navigation*, 3(1), 28.
- Li, X., Wang, X., Liao, J., Li, X., Li, S., & Lyu, H. (2021). Semi-tightly coupled integration of multi-GNSS PPP and S-VINS for precise positioning in GNSSchallenged environments. *Satellite Navigation*, 2, 1–14.
- Minson, S. E., Brooks, B. A., Glennie, C. L., Murray, J. R., Langbein, J. O., Owen, S. E., Heaton, T. H., Iannucci, R. A., & Hauser, D. L. (2015). Crowdsourced earthquake early warning. *Science Advances*, 1(3), e1500036.
- Odijk, D., & Teunissen, P. (2008). ADOP in closed form for a hierarchy of multifrequency single-baseline GNSS models. *Journal of Geodesy*, 82, 473–492.
- Odolinski, R., Teunissen, P., & Odijk, D. (2015). Combined GPS+ BDS for short to long baseline RTK positioning. *Measurement Science and Technology*, *26*(4), 045801.
- Qin, T., Zheng, Y., Chen, T., Chen, Y., & Su, Q. (2021). A light-weight semantic map for visual localization towards autonomous driving. In 2021 IEEE international conference on robotics and automation (ICRA).
- Schikuta, E. (1996). Grid-clustering: An efficient hierarchical clustering method for very large data sets. In *Proceedings of 13th international conference on pattern recognition*.
- Security, T. M. B. o. t. M. o. P. (2022). In the first half of 2022, the number of motor vehicles in China reached 406 million, with 492 million drivers and over 10 million new energy vehicles. Retrieved 11.08 from https://weibo.com/ttart icle/x/m/show/id/2309404788286563680815?_wb_client_=1

- Tang, H., Zhang, T., Niu, X., Fan, J., & Liu, J. (2022). IC-GVINS: A robust, real-time, INS-centric GNSS-visual-inertial navigation system for wheeled robot. arXiv:2204.04962.
- Teunissen, P., & Khodabandeh, A. (2015). Review and principles of PPP-RTK methods. *Journal of Geodesy*, *89*(3), 217–240.
- Vollath, U., Buecherl, A., Landau, H., Pagels, C., & Wagner, B. (2000). Multi-base RTK positioning using virtual reference stations. In *Proceedings of the 13th international technical meeting of the satellite division of the institute of navigation (ION GPS 2000).*
- Wabbena, G., Schmitz, M., & Bagge, A. (2005). PPP-RTK: Precise point positioning using state-space representation in RTK networks. In Proceedings of the 18th international technical meeting of the satellite division of the Institute of navigation (ION GNSS 2005).
- Wang, Y., Kuang, J., Liu, T., Niu, X., & Liu, J. (2023). CrowdMagMap: Crowdsourcing based magnetic map construction for shopping mall. *IEEE Internet of Things Journal*.
- Yang, Y., Vang, C., & Ren, X. (2021). PNT intelligent services. Acta Geodaetica Et Cartographica Sinica, 50(8), 1006–1013.
- Zhang, B., Hou, P., & Odolinski, R. (2022a). PPP-RTK: From common-view to allin-view GNSS networks. *Journal of Geodesy*, 96(12), 102.
- Zhang, Q., Bai, Z., Huang, G., Du, Y., & Wang, D. (2022). Review of GNSS landslide monitoring and early warning. Acta Geodaetica Et Cartographica Sinica, 51(10), 1985.
- Zhang, X., Ren, X., Chen, J., Zuo, X., Mei, D., & Liu, W. (2022b). Investigating GNSS PPP–RTK with external ionospheric constraints. *Satellite Navigation*, 3(1), 1–13.
- Zhu, F., Chen, X., Ma, L., Liu, W., & Zhang, X. (2023). S2L-RTK: Temporal ionospheric modeling for RTK baselines varying from short to long. GPS Solutions, 27(3), 156.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Hongjin Xu received a B.S. degree from the China University of Geosciences (Wuhan) in 2018. He is currently a Ph.D. candidate at the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences. His research focuses on integrated navigation theory and applications.

Xingyu Chen is working toward a Ph.D. at the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, China. His research focuses on multi-GNSS network RTK theory and applications.

Jikun Ou received M.S. and Ph.D. degrees from the Institute of Geodesy and Geophysics (IGG), Chinese Academy of Sciences (CAS), in 1982 and 1994. He is a research professor and a doctoral supervisor at the State Key Laboratory of Geodesy and Earth Dynamics of the Innovation Academy for Precision Measurement Science and Technology, CAS. His studies include the control of data quality, detection of gross errors, and precise orbit determination for LEO and GNSS satellites.

Yunbin Yuan is a professor and the director of the GNSS Application and Research Group, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences. His current research interests include GNSS-based spatial environmental monitoring and analysis, high-precision GNSS satellite navigation and positioning, GNSS in orbit-determination applications, and integrated navigation.