Journal of Geodesy (2026) 100:11
https://doi.org/10.1007/s00190-025-02026-5

ORIGINAL ARTICLE

Fast and precise GNSS RTK positioning with ionosphere-constrained weighted least squares quadratic programming

Xiaolong Mi[1,4] · Xingyu Chen[1,2,5] · Robert Odolinski[3] · Yuyan Wang[1] · Wu Chen[1,4] · Yunbin Yuan[2]

Received: 16 January 2025 / Accepted: 10 November 2025 / Published online: 3 February 2026

Abstract

Achieving fast and precise real-time kinematic (RTK) positioning depends on effective carrier phase integer ambiguity resolution (IAR), yet ionospheric delays pose significant challenges. Traditional approaches often employ the ionosphere-weighted model, which requires complex variance component estimation and iterative optimization procedures that are sensitive to modeling assumptions and local ionospheric dynamics. To address these limitations, we propose a novel ionosphere-constrained model that transforms ionospheric correction into a bounded optimization problem by applying empirically derived inequality constraints to double-differenced (DD) ionospheric delays within a weighted least squares quadratic programming (QP) framework. This convex formulation guarantees global optimality, provides automatic anomaly detection through boundary convergence, and derives transparent, physically interpretable bounds from historical DD delay statistics as a function of baseline length. The QP framework conditionally activates when the unconstrained ionosphere-float solution violates empirical bounds, either guiding estimation toward reasonable interior solutions or preventing large outlier delays while remaining non-intrusive under benign conditions, thereby improving the precision of float solution and accelerating IAR. Experiments on an ultra-short baseline demonstrate that refining the DD ionospheric constraint significantly enhances the formal ambiguity success rates in case of full IAR, highlighting the critical role of precise constraint intervals. Validation using data from the Hong Kong Satellite Positioning Reference Station Network with baseline lengths ranging from a few meters to 50 km shows that our method improves the formal IAR success rate by 12% to 24% compared to the ionosphere-float model and by 5% to 15% compared to the ionosphere-weighted model, achieving near 100% for baselines up to 15 km. Additionally, the ionosphere-constrained model enhances the accuracy of the float solution through DD ionosphere inequality constraints, accelerates the convergence time of the float solution, and achieves fast IAR. Furthermore, our model improves positioning accuracy by over 20% compared to the ionosphere-weighted model, owing to its effective management of DD ionospheric constraints.

Keywords Global navigation satellite system (GNSS) · Real-time kinematic (RTK) · Quadratic programming · Ionosphere-constrained model · Ionosphere-weighted model

1 Introduction

Real-time kinematic (RTK) and network RTK (NRTK) are widely recognized as reliable and precise positioning technologies (Cai et al. 2011; Landau et al. 2007; Teunissen and Khodabandeh 2015), offering centimeter-to-millimeter accuracy for applications (Grejner-Brzezinska et al. 2005; Moore et al. 2008; Rizos 2007; Xu 2012) such as surveying, navigation, and geodesy. The premise of RTK's high-precision positioning capabilities lies in the successful integer ambiguity resolution (IAR) within a short timeframe (Parkins 2011; Teunissen 2004). IAR is the process of resolving the unknown integer number of wavelengths in the carrier

✉ Wu Chen
wu.chen@polyu.edu.hk

1 Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

2 State Key Laboratory of Precision Geodesy, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, China

3 School of Surveying, University of Otago, Dunedin, New Zealand

4 Research Centre for Low-Altitude Economy (RCLAE), The Hong Kong Polytechnic University, Kowloon, Hong Kong

5 University of Chinese Academy of Sciences, Beijing, China

phase measurements, which is essential for achieving precise positioning. For baselines shorter than, say, 10 km, the strong spatial correlation of ionospheric delay between stations allows the ionospheric effects to be safely ignored, enabling the ionosphere-fixed model (Odijk 2000b, 2000c). This model assumes that the ionospheric delay is negligible, facilitating fast IAR and thus high-precision positioning services. However, for baselines above 10 km, the ionospheric delays might need to be estimated as unknown parameters to maintain the reliability of parameter estimation (Odijk et al. 2000; Teunissen 1998). In such cases, achieving IAR quickly becomes challenging because the model's strength becomes too weak due to the need for satellite-by-satellite ionospheric delay estimation, resulting in the ionosphere-float model (Mi et al. 2019b; Teunissen 1997). Therefore, effective management of ionospheric delay is crucial for achieving fast and reliable IAR in RTK applications. This challenge underscores the need for innovative approaches to handle ionospheric variability, especially over longer baselines where traditional models may find it difficult to achieve fast and reliable IAR.

The challenge with the ionosphere-float model is that the float solution has low accuracy, characterized by poor variance and covariance matrices of the float solution (Li et al. 2014; Odijk 2000a). This inaccuracy arises because the model treats ionospheric delays as unknowns, which introduces additional uncertainty into the positioning solution (Psychas et al. 2019). A viable approach to address this issue is introducing external ionospheric delays as a pseudo-observation with a priori values, forming the ionosphere-weighted model (Mi et al. 2023; Teunissen 1998). This model allows for more accurate ionospheric effect modeling over medium to long baselines by incorporating additional information about the ionosphere (Odijk et al. 2000). The implementation of the ionosphere-weighted model in RTK positioning relies on real-time ionospheric corrections, such as those provided by the global ionosphere map (GIM) and regional reference networks. These corrections help mitigate the ionospheric effects by providing estimates of the ionospheric delay that can be used to adjust the positioning calculations. Alternatively, the differential ionospheric delay can be assumed to be zero, simplifying the model but potentially reducing accuracy. However, the successful implementation of the ionosphere-weighted model requires complex variance component estimation and iterative optimization procedures to accurately determine the stochastic model of ionospheric pseudo-observations (Mi et al. 2019b). These procedures are highly sensitive to modeling assumptions and rapidly changing local ionospheric dynamics. Least squares variance component estimation can be employed to quantitatively determine this stochastic model and achieve theoretical optimization (Amiri-Simkooei and Teunissen 2009; Teunissen and Amiri-Simkooei 2008). Teunissen and Khodabandeh (2021) investigated the mean square error (MSE) condition of weighted ionospheric delays to identify the optimal weighting strategy, yielding performances that surpass those of empirical models. Nevertheless, the complex stochastic modeling required by the ionosphere-weighted approach often leads to unstable weight solutions and suboptimal IAR performance, particularly under challenging ionospheric conditions. This highlights the ongoing need for advancements in ionospheric modeling and correction techniques to improve the robustness and accuracy of RTK systems.

In the ionosphere-float model, the differential ionospheric delay is normally estimated as a time-varying parameter over multiple epochs. This approach acknowledges the dynamic nature of the ionosphere, which can change rapidly and unpredictably (Li et al. 2022; Odijk 2000c). Given that the differential ionospheric delay is bounded, a promising approach is to use quadratic programming (QP) (Frank and Wolfe 1956) to constrain the ionospheric delay and find the optimal solution through optimization algorithms. QP is an optimization technique that involves minimizing or maximizing a quadratic objective function subject to linear constraints (Momoh et al. 1994). It extends linear programming by accommodating more complex, nonlinear relationships in the objective function while maintaining linear constraints (Jiang and Ralph 2000). This makes QP particularly suitable for problems where the relationships between variables are not strictly linear, as is often the case in ionospheric modeling. The QP method fundamentally differs from the traditional ionosphere-weighted method by replacing stochastic weight estimation with explicit inequality bounds derived from historical statistics. This constraint-driven framework eliminates the need for complex variance component estimation and iterative tuning, making ionospheric modeling more transparent, stable, and computationally efficient.

QP has wide applications in various fields, including portfolio management (Best and Kale 2000), resource allocation (Bretthauer and Shetty 1997), and engineering design problems (Fesanghary et al. 2008). Several algorithms, such as interior-point methods (Altman and Gondzio 1999) and sequential quadratic programming (SQP) (Boggs and Tolle 1995), can solve QP problems. These algorithms are designed to efficiently handle the computational complexity associated with QP, making them suitable for real-time applications (Fan and Zhang 1998). The complexity of solving these problems can increase with the number of variables, especially for large-scale problems. Recent research has focused on developing efficient methods for solving specific types of QP problems (Cimini and Bemporad 2017; Luo et al. 2010), such as those with low-rank non-convexity or box constraints. QP has been employed to address various challenges in GNSS domain, including fault detection and exclusion (FDE) in

receiver autonomous integrity monitoring (RAIM) (Yang and Sun 2020a, 2020b). However, in high-precision GNSS positioning, there is no QP application yet. This presents an opportunity for further exploration and development of QP-based methods to enhance the accuracy and reliability of GNSS positioning systems.

The primary goal of our research is to introduce a novel ionosphere-constrained model that transforms ionospheric delay management from stochastic weight optimization into bounded constraint optimization, leveraging QP to impose empirically derived inequality constraints on double-differenced (DD) ionospheric delays. In this contribution, we first formulate the traditional single-differenced (SD) RTK models applicable to short, medium, and long baselines. We then propose a new ionosphere-constrained model using weighted least squares quadratic programming, detailing its functional model and specific implementation process. Following this, we analyze the critical importance of the accuracy of ionospheric constraints for successful IAR. Finally, the positioning performance of our proposed model is compared to traditional models.

2 Single-differenced RTK models

In this section, we will first introduce the SD GNSS RTK models with different ionospheric constraints that form the basis of this study. As the starting point of developing the SD GNSS RTK models, SD GNSS code and phase observations are given as follows

$$\begin{aligned} p^s_{1r,\,j} &= \rho^s_{1r} + \tau^s_{1r} + dt_{1r} + \mu_j I^s_{1r} + d_{1r,\,j} + \varepsilon^s_{p,\,1r,\,j} \\ \phi^s_{1r,\,j} &= \rho^s_{1r} + \tau^s_{1r} + dt_{1r} - \mu_j I^s_{1r} + \lambda_j N^s_{1r,\,j} + \delta_{1r,\,j} + \varepsilon^s_{\phi,\,1r,\,j} \end{aligned} \tag{1}$$

where s, j, and r represent satellite, frequency, and receiver, respectively. The 1 in the subscript represents the receiver at the reference station, and $(\cdot)_{1r} = (\cdot)_r - (\cdot)_1$ is the between-receiver SD notation. $p^s_{1r,\,j}$ and $\phi^s_{1r,\,j}$ are the SD code and phase observations. SD satellite-receiver range, receiver clock, tropospheric delay, ionospheric delay, receiver code bias, receiver phase bias, and phase ambiguity are described by ρ^s_{1r},dt_{1r},τ^s_{1r},I^s_{1r},$d_{1r,\,j}$,$\delta_{1r,\,j}$, and $N^s_{1r,\,j}$, respectively. $\mu_j = \lambda_j^2 \Big/ \lambda_1^2$ is the coefficient of ionospheric delay where λ_j is the wavelength. $\varepsilon^s_{p,\,1r,\,j}$ and $\varepsilon^s_{\phi,\,1r,\,j}$ represent the SD code and phase observation noise and mis-modeled random effects. It should be noted that Eq. (1) represents a rank-deficient system, which can be solved by the S-system theory (Odijk et al. 2017).

2.1 Ionosphere-float variant

When the distance between receivers is significant, it is important to consider ionospheric and tropospheric delays. The conventional approach for addressing tropospheric delay involves dividing it into dry $(\tau_d)^s_{1r}$ and wet τ_r components, with the dry delay corrected by a model, and the wet delay estimated using an elevation-dependent mapping function m^s_r(Hadas et al. 2017). In the context of the SD GNSS RTK equation, there are three types of rank deficiencies (Mi et al. 2019a; Odolinski et al. 2015). By selecting the appropriate S-basis, a full-rank ionosphere-float model can be constructed. The first type of rank deficiency occurs between the columns of the receiver clock and the receiver code/phase biases, which can be resolved by fixing the receiver code bias on the first frequency. The second type of rank deficiency is between the columns of the receiver phase bias and phase ambiguity, which can be resolved by fixing the ambiguities. The third type of rank deficiency involves the receiver clocks, receiver code/phase biases, and ionospheric delay, which can be resolved by fixing the receiver code bias on the second frequency. In this case, the full-rank ionosphere-float model can be given as

$$\begin{aligned} \tilde{p}^s_{1r,\,j} &= \rho^s_{1r} + m^s_r \tau_{1r} + d\tilde{t}_{1r} + \mu_j \tilde{I}^s_{1r} + \tilde{d}_{1r,\,j} + \varepsilon^s_{p,\,1r,\,j} \\ \tilde{\phi}^s_{1r,\,j} &= \rho^s_{1r} + m^s_r \tau_{1r} + d\tilde{t}_{1r} - \mu_j \tilde{I}^s_{1r} + \lambda_j N^{1s}_{1r,\,j} + \tilde{\delta}_{1r,\,1} + \tilde{\delta}_{1r,\,j} + \varepsilon^s_{\phi,\,1r,\,j} \end{aligned} \tag{2}$$

where $\tilde{p}^s_{1r,\,j} = p^s_{1r,\,j} - (\tau_d)^s_{1r}$ and $\tilde{\phi}^{1s}_{1r,\,j} = \phi^s_{1r,\,j} - (\tau_d)^s_{1r}$. $d\tilde{t}_{1r} = dt_{1r} + d_{1r,\,IF}$ is the estimated SD between-receiver clock biased by the ionosphere-free receiver code bias where $d_{1r,\,IF} = \frac{\mu_2}{\mu_2-\mu_1} d_{1r,\,1} - \frac{\mu_1}{\mu_2-\mu_1} d_{1r,\,2}$, $\tilde{d}_{1r,\,j} = d_{1r,\,j} - d_{1r,\,IF} - \mu_j d_{1r,\,GF}$ is the estimated SD between-receiver differential code bias (DCB) with $j \geq 3$, i.e., it is only estimable on the third frequency and beyond. $\tilde{\delta}_{1r,\,1} = \delta_{1r,\,1} - d_{1r,\,IF} + \mu_j d_{1r,\,GF} + \lambda_j N^1_{1r,\,1}$ and $\tilde{\delta}_{1r,\,j} = \delta_{1r,\,j} - \delta_{1r,\,1} + \lambda_j N^1_{1r,\,j} - \lambda_j N^1_{1r,\,1}$ are the SD between-receiver differential phase bias (DPB) of the first frequency and the second frequency and above where $d_{1r,\,GF} = \frac{1}{\mu_2-\mu_1}(d_{1r,\,2} - d_{1r,\,1})$. $\tilde{I}^s_{1r} = I^s_{1r} + d_{1r,\,GF}$ is the SD between-receiver ionospheric delay biased by geometry-free (GF) receiver code bias. The ionosphere-float model does not impose any constraints on the ionosphere and is usually used for long baselines. Since the estimation of ionosphere parameters leads to weak model strength, it is difficult to quickly achieve IAR, and thus, it is difficult to achieve high-precision positioning in a short time span.

2.2 Ionosphere-fixed variant

If the baseline length is short enough, then we can safely assume that the ionospheric and tropospheric delays are

negligible. In this case, the third type of rank deficiency mentioned above, that is, the rank deficiency between the columns of the receiver clocks, receiver code/phase biases, and ionospheric delay, no longer exists. After solving the first two types of rank deficiency, the full-rank ionosphere-fixed model can be obtained as follows

$$\begin{aligned}\tilde{p}^s_{1r,j} &= \rho^s_{1r} + d\tilde{\tilde{t}}_{1r} + \tilde{\tilde{d}}_{1r,j} + \varepsilon^s_{p,1r,j}\\ \tilde{\phi}^s_{1r,j} &= \rho^s_{1r} + d\tilde{\tilde{t}}_{1r} + \lambda_j N^{1s}_{1r,j} + \tilde{\tilde{\delta}}_{1r,1} + \tilde{\tilde{\delta}}_{1r,j} + \varepsilon^s_{\phi,1r,j}\end{aligned} \tag{3}$$

where $d\tilde{\tilde{t}}_{1r} = dt_{1r} + d_{1r,1}$ is the SD between-receiver clock and $\tilde{\tilde{d}}_{1r,j} = d_{1r,j} - d_{1r,1}$ is the SD between-receiver DCB with $j \geq 2$.$\tilde{\delta}_{1r,1} = \delta_{1r,1} - d_{1r,1} + \lambda_j N^1_{1r,1}$ is the SD between-receiver DPB of the first frequency and $\tilde{\delta}_{1r,j} = \delta_{1r,j} - \delta_{1r,1} + \lambda_j N^1_{1r,j} - \lambda_j N^1_{1r,1}$ is the SD between-receiver DPB on frequency j. The ionosphere-fixed model is widely used in fast and high-precision positioning because it no longer needs to estimate ionospheric parameters, the model strength is enhanced, and IAR becomes easier.

2.3 Ionosphere-weighted variant

In high-precision positioning applications, the scenario often faced is that the baseline length is between 10 and 100 km. Ignoring the influence of ionospheric delay and using an ionosphere-fixed model will lead to biased parameter estimation. Estimating ionospheric delay and using an ionosphere-float model will reduce the model strength and make it difficult to achieve fast IAR. An alternative approach is introducing external ionospheric delay corrections to enhance the model. There are many ways to introduce ionospheric delay corrections, but these corrections are not accurate enough to completely ignore the ionospheric delay as in the ionosphere-fixed model. Therefore, the introduced ionospheric delay corrections are usually treated in a stochastic rather than a deterministic way, i.e., their uncertainties are propagated into the stochastic model. The full-rank ionosphere-weighted model can be given as follows

$$\begin{aligned}\tilde{p}^s_{1r,j} &= \rho^s_{1r} + m^s_r\tau_{1r} + d\tilde{\tilde{t}}_{1r} + \mu_j I^s_{1r} + \tilde{\tilde{d}}_{1r,j} + \varepsilon^s_{p,1r,j}\\ \tilde{\phi}^s_{1r,j} &= \rho^s_{1r} + m^s_r\tau_{1r} + d\tilde{\tilde{t}}_{1r} - \mu_j I^s_{1r} + \lambda_j N^{1s}_{1r,j} + \tilde{\tilde{\delta}}_{1r,1} + \tilde{\tilde{\delta}}_{1r,j} + \varepsilon^s_{\phi,1r,j}\\ \overline{I}^s_{1r} &= I^s_{1r} + \varepsilon^s_{I,1r,j}\end{aligned} \tag{4}$$

where $\overline{I}^s_{1r}$ is the SD ionospheric pseudo-observables. The ionosphere-weighted model accelerates IAR to a certain extent and has been well applied in the fast and precise positioning for medium baselines. However, it is always a challenge to reasonably determine the weights of the ionospheric pseudo-observations, and unreasonable weights will lead to biased parameter estimation. The ionosphere-fixed and ionosphere-float models can be regarded as two special forms of the ionosphere-weighted model (Odijk 2000b). On the one hand, when the weight of the given ionospheric pseudo-observation is large enough, the model can be considered equivalent to the ionosphere-fixed model; on the other hand, when the weight is small enough, the model can be considered equivalent to the ionosphere-float model.

The limitations of existing models highlight the need for a new approach that can provide the computational efficiency of bounded constraints while maintaining the physical interpretability of ionospheric effects. This motivates the development of our ionosphere-constrained model, which transforms the ionospheric modeling problem from stochastic estimation into bounded optimization.

3 Ionosphere-constrained model with weighted least squares quadratic programming

The ionosphere-weighted model requires external ionospheric correction information, and it is challenging to build an accurate stochastic model. The common terms of the ionospheric delay between stations are eliminated, leaving only their non-common terms after the SD process, and they are not randomly distributed but bounded. This provides us with a new idea for solving the problem of fast and precise positioning of the medium baselines, that is, to impose bounded constraints on the ionospheric delay. Let us return to the ionosphere-float model in Eq. (2) again. We can see that the SD ionospheric delay absorbs the ionosphere-free receiver code bias. If we want to directly constrain the SD ionospheric delay, then a priori information on the ionosphere-free receiver code bias is necessary. However, it is worth noting that the ionospheric delay of each satellite contains the same ionosphere-free receiver code bias, which suggests that we can choose a reference satellite to constrain the DD ionospheric delay. DD ionospheric delay can be bounded with zero as the mean. The ionosphere-constrained model can be given as follows

$$\begin{aligned}\tilde{p}^s_{1r,j} &= \rho^s_{1r} + m^s_r\tau_{1r} + d\tilde{t}_{1r} + \mu_j \tilde{I}^s_{1r} + \tilde{d}_{1r,j} + \varepsilon^s_{p,1r,j}\\ \tilde{\phi}^s_{1r,j} &= \rho^s_{1r} + m^s_r\tau_{1r} + d\tilde{t}_{1r} - \mu_j \tilde{I}^s_{1r} + \lambda_j N^{1s}_{1r,j} + \tilde{\delta}_{1r,1} + \tilde{\delta}_{1r,j} + \varepsilon^s_{\phi,1r,j}\\ &\left|\tilde{I}^{1s}_{1r}\right| \leq I_{TH}\end{aligned} \tag{5}$$

where $\tilde{I}^{1s}_{1r} = I^s_{1r} - I^1_{1r}$ is the DD ionosphere delay and $|\cdot|$ represents the absolute value operator symbol. I_{TH} is the threshold of DD ionospheric delay and is obtained by analyzing historical GNSS data. In our processing, we consistently select the satellite with the highest elevation angle ($> 15°$) at the first epoch as the reference satellite. Through

systematic analysis using different high-elevation satellites as reference satellites across the Hong Kong SatRef network, we found that the fitted slope and intercept parameters remain stable within measurement noise levels. The 15° elevation threshold effectively eliminates systematic boundary parameter variations from low-elevation multipath and atmospheric errors. While this empirical validation demonstrates practical robustness under our operational conditions, we acknowledge that rigorous mathematical proof of theoretical invariance remains for future investigation.

The ionosphere-constrained approach offers significant advantages over traditional methods. Since the boundaries are determined through straightforward statistical analysis of historical DD delay ranges, this model eliminates the complex stochastic model construction and iterative weight optimization required by ionospheric pseudo-observations in Eq. (4).

3.1 General form of quadratic programming

Once the boundary of a reasonable ionospheric delay is determined, the next step is to find the optimal solution under the inequality constraints. QP is a mathematical optimization problem characterized by a quadratic objective function and linear constraints, which can be used to solve the optimal problem of the ionosphere-constrained model. Next, we will give the general form of QP and then derive the weighted least squares (WLS) QP and its solution suitable for solving the problem in this paper.

The general form of a QP problem can be expressed as follows

$$\min_x \frac{1}{2} x^T Q x + c^T x \tag{6}$$

which is subject to the following conditions

$$\begin{aligned} Ax &\le b \\ Cx &= d \end{aligned} \tag{7}$$

where x is the estimated unknowns. Q must be a symmetric positive semidefinite matrix representing the quadratic coefficients to ensure that it is convex, which means that for any vector z, the quadratic form satisfies the following conditions

$$z^T Q z \ge 0 \tag{8}$$

Convexity of the objective function is important for two reasons. First, a convex function has the property that any local minimum is also a global minimum, which ensures that the optimization algorithm can reliably find the global optimum. In addition, convex functions are easier to analyze mathematically, simplifying the derivation of optimality conditions and convergence proofs for the optimization algorithm. A and C represent the matrices containing the coefficients of the inequality constraints and equality constraints, respectively, while b and d represent the upper bound of the inequality constraints and the right side of the equality constraints, respectively.

3.2 Weighted least squares quadratic programming

The core of the GNSS positioning solution is WLS, so the ionosphere-constrained model is no exception. The WLS problem aims to minimize the weighted sum of squared residuals, which can be formulated as a QP problem. Consider a weighted least squares $y = Hx$, where y is the vector of observed values and H is the design matrix corresponding to the coefficients of the estimated unknowns. Let W be a diagonal matrix of weighting functions. The WLS objective function can be written as

$$\min_x (y - Hx)^T W (y - Hx) \tag{9}$$

Then, expanding and simplifying the objective function, the following will be obtained:

$$(y - Hx)^T W (y - Hx) = y^T W y - 2 y^T W H x + x^T H^T W H x \tag{10}$$

Since $y^T W y$ is a constant, it can be omitted from the optimization problem. Therefore, the WLS problem can be transformed into the following QP problem:

$$\min_x \frac{1}{2} x^T (H^T W H) x - (H^T W y) x \tag{11}$$

which is also subject to Eq. (7). In addition, $Q = H^T W H$ and $c = -H^T W y$ correspond to the general QP problem. As mentioned earlier, Q being convex is crucial for solving QP problems. Therefore, we will next prove that Q in WLS is convex. To achieve this, Q will be shown to be positive semi definite, that is, to satisfy the condition of Eq. (8). Substituting $Q = H^T W H$ into Eq. (8) we get

$$z^T Q z = z^T (H^T W H) z = (Hz)^T W (Hz) \tag{12}$$

Since W usually represents the variance–covariance matrix of the observations, it must be a symmetric positive definite matrix. Thus, for any nonzero vector u

$$u^T W u \ge 0 \tag{13}$$

Let $u = Hz$, then we have

$$(Hz)^T W (Hz) \ge 0 \text{ for any non-zero } z \tag{14}$$

Therefore, Q in WLS is positive semidefinite, which proves that the quadratic form is convex.

In the proposed method, each DD ionospheric delay parameter is constrained within a physically reasonable interval, i.e., a lower and upper bound. These bounds are determined based on statistical analysis (see Eq. (5)) and are referred to as the ionospheric delay thresholds I_{TH}.

Mathematically, these thresholds enter the QP algorithm as inequality constraints. Specifically, if I_i denotes the DD ionospheric delay for the $i-th$ satellite, then for each epoch and each satellite, we impose

$$l_l \leq I_i \leq t_i \tag{15}$$

where l_i and t_i are the lower and upper bounds for I_i.

In the QP formulation (Eqs. (6)–(7)), these constraints are incorporated as follows

$$l \leq A_{\text{ion}} x \leq t \tag{16}$$

where A_{ion} is a selection matrix extracting the ionospheric components and l,t are the vectors of lower and upper bounds for all satellites. It should be noted that there are no equality constraints imposed on the ionospheric delays in our QP model; all ionospheric constraints are inequality constraints. This ensures that the estimated ionospheric delays remain physically reasonable, while allowing the QP solver to find the optimal solution within these bounds. The method of determining these thresholds will be discussed later.

3.3 Active set method in weighted least squares quadratic programming

Several methods are available to solve WLS QP problems with equality and inequality constraints, including interior-point methods (Altman and Gondzio 1999), gradient projection methods (Serafini et al. 2005), and active set method (Best and Chakravarti 1990). The active set method offers a distinct advantage by iteratively refining the set of active constraints, treating them as equalities, and solving the resulting equality-constrained quadratic programming problem (Boland 1997). It efficiently handles the sparsity and structure of large-scale problems, incorporates both equality and inequality constraints, and systematically converges to the optimal solution (Coleman and Hulbert 1989). This makes it a robust and reliable method for precise data fitting and modeling in GNSS, where accuracy and feasibility are paramount.

The first step of the active set method is initialization, that is, choosing a feasible initial point x_0 that satisfies all the constraints

$$\begin{aligned} Ax_0 &\leq b \\ Cx_0 &= d \end{aligned} \tag{17}$$

The feasible initial point x_0 of the ionosphere-constrained model is obtained by assuming the DD ionospheric delay is zero. In this case, a model like the ionosphere-fixed model can be constructed as follows

$$\begin{aligned} \tilde{p}^s_{1r,\,j} &= \rho^s_{1r} + m^s_r \tau_{1r} + d\tilde{t}_{1r} + \mu_j \tilde{I}^1_{1r} + \tilde{d}_{1r,\,j} + \varepsilon^s_{p,\,1r,\,j} \\ \tilde{\phi}^s_{1r,\,j} &= \rho^s_{1r} + m^s_r \tau_{1r} + d\tilde{t}_{1r} - \mu_j \tilde{I}^1_{1r} + \lambda_j N^{1s}_{1r,\,j} + \tilde{\delta}_{1r,\,1} + \tilde{\delta}_{1r,\,j} + \varepsilon^s_{\phi,\,1r,\,j} \end{aligned} \tag{18}$$

where only one ionospheric parameter $\tilde{I}^1_{1r}$ is estimated. The solution of the WLS of the above model can be used as the feasible initial point, which satisfies all the above constraints. Please note that the ionosphere-fixed model (Eq. (3)) cannot give an initial solution because the parameters of the ionosphere-fixed model and the ionosphere-constrained model do not have the same interpretation. Then, we initialize the active set S_0 so that it includes all equality constraints and inequality constraints that are active at x_0.

The second step is to solve the equality-constrained QP, that is, in each iteration, solve the quadratic programming problem with only the active constraints as equalities considered. The equality-constrained QP is as follows

$$\begin{aligned} &\min_x \frac{1}{2} x^T (H^T W H) x - (H^T W y) x \\ &s.t.\, A_S x = b_S \end{aligned} \tag{19}$$

where A_S and b_S are the matrix of active constraints and the vector of corresponding right-hand sides, respectively. Next, solve the Karush–Kuhn–Tucker (KKT) system to obtain the currently active set

$$\begin{pmatrix} H^T W H & A_S^T \\ A_S & 0 \end{pmatrix} \begin{pmatrix} x \\ \alpha \end{pmatrix} = \begin{pmatrix} H^T W y \\ b_S \end{pmatrix} \tag{20}$$

where α are the Lagrange multipliers. The standard linear algebra techniques can be used to get the solutions to both x and α.

The third step is to check optimality, which consists of computing the Lagrange multipliers α of the active constraints and checking the sign of each Lagrange multiplier. If all $\alpha_i \geq 0$, the current solution x is optimal. However, if any $\alpha_i < 0$, indicating that the current solution is not optimal, the algorithm identifies the constraint associated with the largest negative multiplier, denoted as $\alpha_{\min}$.

The fourth step is to update the active set by removing the constraint corresponding to the largest negative Lagrange multiplier $\alpha_{\min}$ from the active set S, resulting in a new active set. Next, all inequality constraints at the current solution x should be evaluated, which involves calculating the values of $Ax-b$. If any inequality constraints are found to be violated, i.e.,$Ax_i > b_i$, the constraint with the most violations (the one

with the largest value of $Ax_i - b_i$) is added to the active set S.

The fifth step is to repeat steps 2–4 until convergence. The algorithm constructs the Lagrangian function using the new active set and solves the KKT conditions to obtain a new solution and the corresponding Lagrange multipliers. The optimality of the new solution is then checked, and the active set is updated accordingly. This iterative process continues until a convergence criterion is met, such as $\left\|x^{k+1} - x^k\right\| < \varepsilon$ or when the KKT conditions are satisfied, where k and ε represent the number of iterations and the set threshold, respectively. Through this systematic approach, the active set method effectively navigates the feasible region to find the optimal solution to the quadratic programming problem.

The effectiveness of the ionospheric inequality constraints in the QP framework depends critically on the choice of constraint bounds. These bounds are empirically derived from statistical analysis of historical DD delay variability as a function of baseline length, ensuring they capture realistic ionospheric behavior rather than arbitrary limits. The QP framework operates through conditional activation with three distinct operational scenarios. First, when the unconstrained ionosphere-float solution violates the empirical bounds but the QP optimization converges to an interior solution within the constraints, the method successfully guides the estimation toward more precise, physically reasonable parameter estimates that enhance float solution accuracy. Second, when the unconstrained solution significantly violates the bounds and the QP solution converges directly to the constraint boundaries, this indicates the presence of large outliers or challenging ionospheric conditions requiring adaptive boundary management. Third, when the unconstrained solution naturally satisfies the empirical bounds, the QP method reduces to the standard ionosphere-float solution with minimal computational overhead.

The conditional activation mechanism operates through three distinct scenarios: systematic stabilization when unconstrained solutions violate empirical bounds, automatic anomaly detection when solutions converge to boundaries, and non-intrusive behavior when natural bounds are satisfied. This adaptive framework provides both computational efficiency and quality control capabilities.

In addition, the QP framework is used to obtain the optimal constrained float solution, including real-valued ambiguities and ionospheric delays. However, the integer nature of carrier phase ambiguities is not resolved in the QP step. Therefore, after obtaining the float solution, IAR using the LAMBDA method (Teunissen 1995) or equivalent integer least squares algorithms is needed to perform IAR. The QP-based constraint optimization enhances the precision and reliability of the float ambiguities when needed, which in turn improves the success rate and robustness of subsequent IAR.

4 The role of DD ionospheric constraints

To elucidate the influence of DD ionospheric constraints on GNSS parameter estimation, we employ an ultra-short baseline configuration with the proposed ionosphere-constrained model. This setup is strategically chosen to validate the assumption that the DD ionospheric delay is negligible, thereby allowing a direct comparison between the estimated parameters and their true values. An epoch-by-epoch estimation strategy is utilized, which ideally results in a Gaussian distribution of the estimated DD ionospheric delays when no constraints are applied.

Figure 1 presents the DD ionospheric delay estimates with float ambiguities across different constraint scenarios, ranging from no constraints to constraints as stringent as 0.5 cm. Each color in the figure corresponds to a different satellite, providing a comprehensive view of the constraint effects. The subplot in the upper left demonstrates that the absolute DD ionospheric delay from an unconstrained least squares solution remains within 30 cm. This finding indicates that any ionospheric constraint exceeding this threshold is ineffective, as shown by the identical results of the 50 cm constraint and the unconstrained scenario in the lower left subplot. This observation highlights the importance of selecting appropriate constraint levels. Constraints that are too lenient fail to influence the solution, rendering them ineffective. Conversely, as constraints become tighter, the DD ionospheric estimates converge toward the true values of zero. This convergence not only improves the accuracy of the ionospheric delay estimates but also positively impacts the estimation of other critical parameters, such as position and float ambiguities, which are of paramount importance in GNSS applications.

To further evaluate the impact of ionospheric constraints on the error distribution, Fig. 2 displays both the error distribution and the empirical cumulative distribution function (ECDF) of DD ionospheric delays under different constraints. The aggregated data from all satellites reveal that, in the absence of effective constraints, the DD ionospheric delay follows a normal distribution, as confirmed by its ECDF. In addition, ineffective ionospheric constraints result in a perfect S-curve (50 cm constraint), similar to the case with no constraints, again highlighting the importance of reasonable constraints. However, when constraints are effectively applied, the distribution of errors deviates from normality. For example, with a 10 cm constraint, the error distribution resembles a truncated normal distribution, with the truncation boundary defined by the constraint. This change is reflected in the ECDF, where the characteristic S-curve of a normal distribution transitions toward a linear form. As constraints become more stringent, the ECDF approaches a straight line, indicating a uniform distribution. This transformation in error distribution underscores the efficacy of

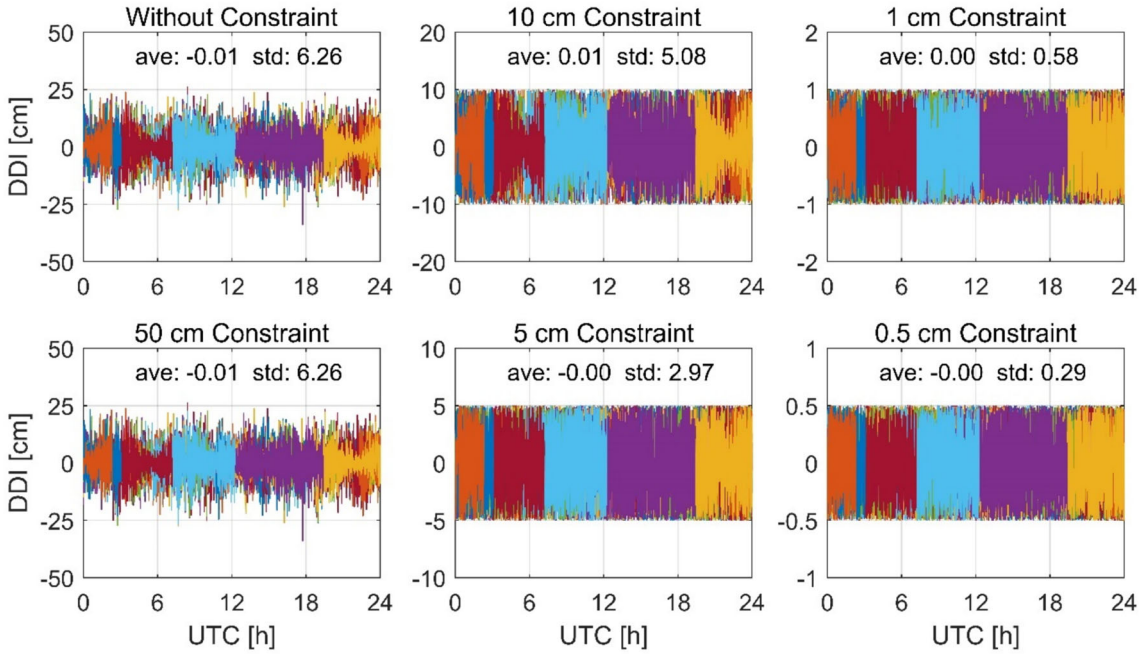

Fig. 1 Estimation of DD ionospheric delay under different ionospheric constraints

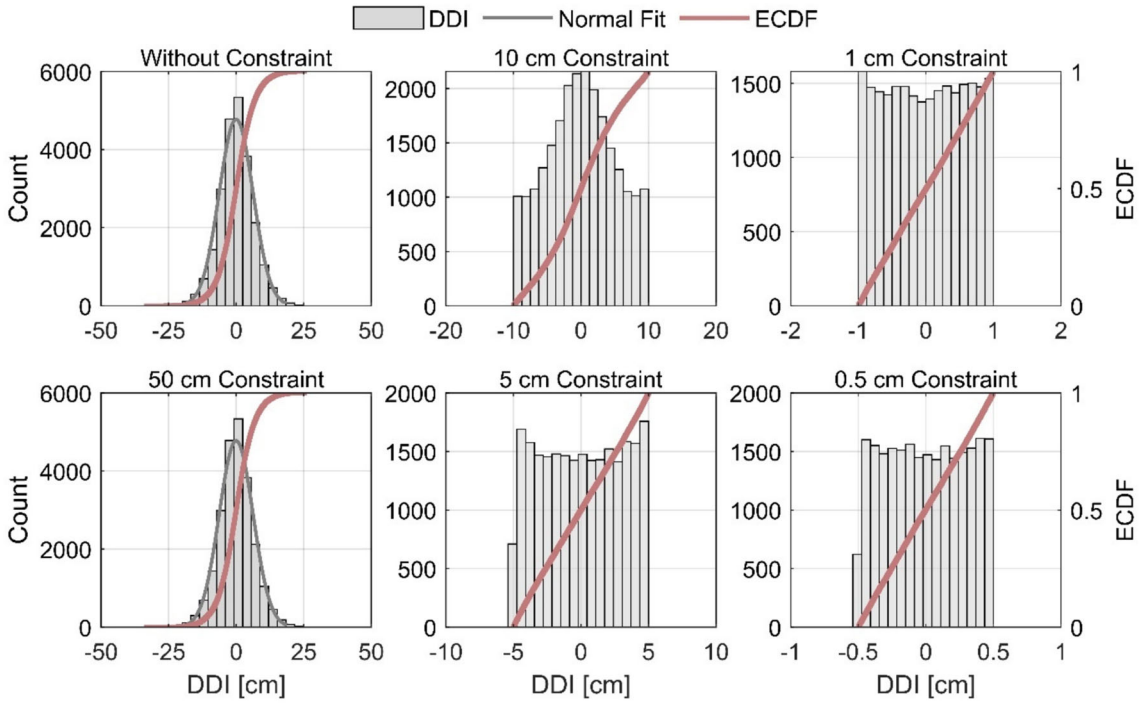

Fig. 2 Error distribution and empirical cumulative distribution function of DD ionospheric delay with different constraints

well-chosen constraints in enhancing parameter estimation. By tightening constraints, we can achieve a more uniform and predictable error distribution, which is crucial for optimizing GNSS solutions.

The primary objective of implementing ionospheric constraints in this study is to enhance the success rate of IAR, thereby fully leveraging the benefits of phase observations. Figure 3 presents the single-epoch success rate statistics under various ionospheric constraints, providing a basis for evaluating their effectiveness. Several key insights emerge from this analysis. Firstly, when ionospheric delays are estimated without any constraints in the ionosphere-float model,

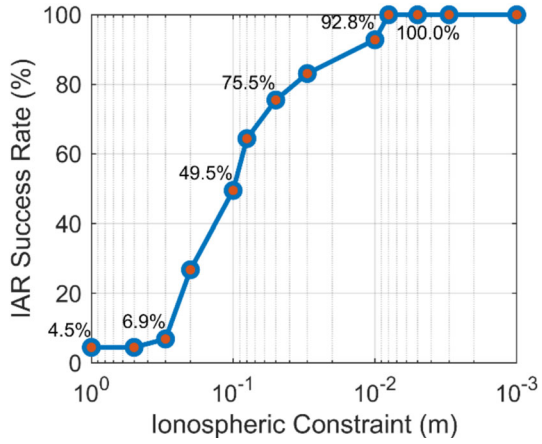

Fig. 3 Success rate of integer ambiguity resolution under different ionospheric constraints

achieving IAR becomes exceedingly challenging. Secondly, the introduction of effective ionospheric constraints confines the error in estimated ionospheric delays to a manageable range, facilitating IAR. Notably, when ionospheric constraints are applied at the decimeter level, there is a marked improvement in the success rate compared to unconstrained conditions, although the rate remains suboptimal. Thirdly, with centimeter-level ionospheric constraints, the success rate of ambiguity resolution improves significantly, reaching 92.8% when constraints are tightened to 1 cm. Lastly, achieving millimeter-level ionospheric constraints results in a 100% IAR success rate, effectively mirroring the performance of an ionosphere-fixed model. In summary, reasonable ionospheric constraints are instrumental in fixing ambiguities, and more precise ionospheric constraints are highly beneficial.

5 Performance evaluation using the Hong Kong satellite positioning reference station network

This section aims to evaluate the feasibility of the proposed ionosphere-constrained model based on weighted least squares quadratic programming in RTK positioning. We utilized the Hong Kong Satellite Positioning Reference Station Network (SatRef), which consists of 16 reference stations with baseline lengths ranging from a few kilometers to 50 km (see Fig. 4). This network provides robust data support for validating the performance of the proposed method.

5.1 Experimental setup

The experimental setup involves using dual-frequency GPS data in two batches, one for determining both the boundaries of the DD ionospheric delay for the ionosphere-constrained model and the optimal stochastic weights for the ionosphere-weighted model from November 1 to December 31, 2023, and the other for validating the performance of the ionospheric constraint model for one week from January 1 to 7, 2024. The primary objective was to evaluate the performance of the proposed ionosphere-constrained (IC) model in RTK positioning, with a focus on full IAR success rate and positioning accuracy. The ionosphere-constrained model, as described by Eq. (5), was benchmarked against the ionosphere-float (IF) model (Eq. (2)), the ionosphere-weighted (IW) model (Eq. (4)) and the ionosphere-fixed (IFixed) model (Eq. (3)). To ensure fair comparison, both the ionosphere-constrained and ionosphere-weighted models were calibrated using the same Hong Kong SatRef dataset through data-driven parameter estimation. The processing strategies for these models are detailed in Table 1, highlighting the distinct approaches each model takes in handling ionospheric delays.

All four models employed a Kalman filter, which was reinitialized every 2 h to explore the convergence behavior and stability of each model. The ionosphere was consistently modeled as white noise within the Kalman filter framework, but the models differed in their constraints on ionospheric delays. Specifically, the ionosphere-float model represents the traditional long-baseline approach, estimating ionospheric delays on a per-satellite basis without additional constraints. The ionosphere-weighted model applies data-driven stochastic weights derived from the same historical dataset using least squares variance component estimation (Amiri-Simkooei et al. 2009; Teunissen and Amiri-Simkooei 2008; Mi et al. 2019b). In contrast, the ionosphere-constrained model imposes empirically derived constraints on the DD ionospheric delays, with the constraint interval linked to baseline length. These bounds are obtained by statistical analysis and model fitting of historical GNSS data from the Hong Kong SatRef network, ensuring that the constraints are both physically meaningful and regionally adaptive. The specific procedure and rationale for determining these constraint intervals will be described in detail later. Both approaches utilize the same underlying dataset to ensure equivalent information content and fair performance comparison. The stochastic model settings for undifferenced observations utilize the exponential elevation weighting function (Shen et al. 2009) and Zenith-referenced a priori phase and code standard deviations (STDs). To mitigate ionospheric modeling errors introduced by low-elevation satellites, a cutoff elevation of 15° was applied; only satellites above 15° elevation were included in all processing steps. The LAMBDA method was employed for IAR (Teunissen 1995), verified through a ratio test (Teunissen and Verhagen 2009), while the detection, identification, and adaptation (DIA) method was used for outlier detection and elimination (Teunissen 2018).

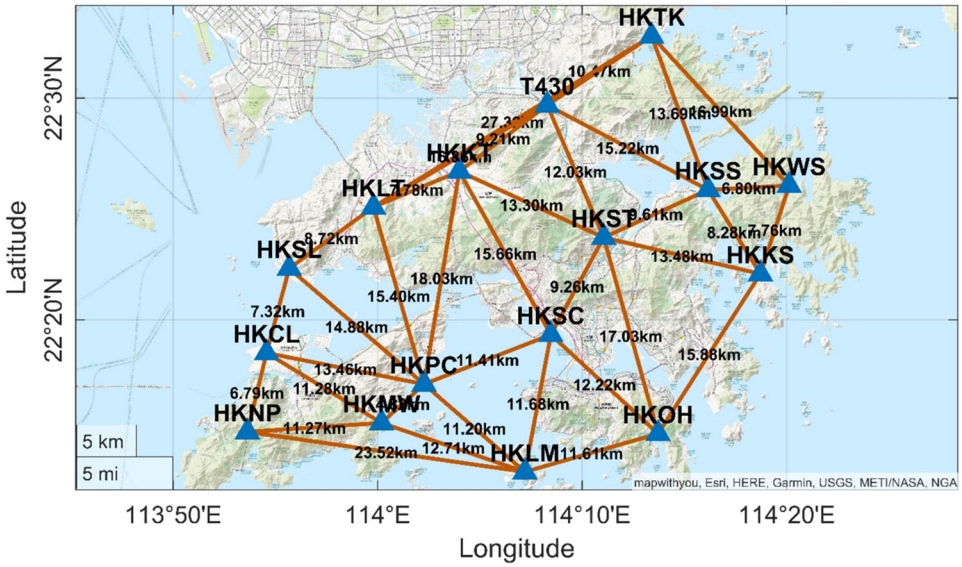

Fig. 4 Overview of the experimental baselines

Table 1 Main data processing strategies in this study for the four models

Item	Strategy details
Ionospheric constraints	IC model (Eq. (5)): inequality DD ionospheric constraints IW model (Eq. (4)): pseudo-observation constraints IFixed model (Eq. (3)): constraint on SD ionospheric delay being 0 IF model (Eq. (2)): N/A
Estimator	Kalman filter (reinitialized every 2 h)
Observations	GPS L1 + L2 with sampling interval of 30 s
Cutoff elevation angle	15°
Ionospheric delays	Estimated as white noise
Position	Estimated as white noise
Receiver clock	Estimated as white noise
Tropospheric delays	Dry part: UNB3m model; wet part: Estimated as a random walk with a process noise of $2\,\mathrm{mm}/\sqrt{\mathrm{hour}}$
Stochastic model	Elevation-dependent weighting (Shen et al. 2009); Zenith-referenced a priori phase and code STDs: 0.003 m and 0.3 m
Between-receiver DCB and phase biases	Estimated as a time constant
IAR	LAMBDA (Teunissen 1995) with a ratio test of a threshold of 3 (Teunissen and Verhagen 2009)
Outlier detection and elimination	DIA

5.2 Determination of stochastic parameters and constraint bounds

To establish both physically meaningful ionospheric delay boundaries for the ionosphere-constrained model and optimal stochastic weights for the ionosphere-weighted model, we employ a comprehensive data-driven methodology utilizing the same historical GNSS observations from the Hong Kong SatRef network, spanning from November 1 to December 31, 2023. This unified approach ensures fair comparison between the two methods by deriving their respective parameters from identical data sources.

For the ionosphere-weighted model parameter estimation, we apply least squares variance component estimation (Teunissen and Amiri-Simkooei 2008) to determine ionospheric pseudo-observation variances. The procedure assumes a zero ionospheric delay and then applies the ionosphere-float model to estimate the variances of ionospheric delays by treating them as unknown parameters in the adjustment process (Mi et al.2019b). The ionospheric variance is subsequently modeled using empirical relationships derived from the extracted statistics, following established approaches for characterizing the spatial and temporal correlation structure of atmospheric delays. It is important to highlight that previous research utilizing ionosphere-weighted models has demonstrated that the magnitude of ionospheric delays is influenced by both the baseline length and the satellite elevation angles (Zhou and Wang 2013; Mi et al. 2019b). However, in our analysis of the Hong Kong SatRef network with a 15° elevation cutoff, no significant correlation was found between DD ionospheric delays and satellite elevation angles. Consequently, we chose not to incorporate satellite elevation angles into our modeling approach. In this study, we adopt a baseline-length-dependent linear variance model

of 1.08 mm/km, where the standard deviation scales linearly with baseline distance d in kilometers. This empirical relationship is derived from the Hong Kong SatRef network and provides a more suitable model for the test region compared to conventional approaches, e.g., 0.96 mm/km (Odolinski et al. 2015).

For the ionosphere-constrained model boundary determination, we utilize the same Hong Kong SatRef dataset to extract DD ionospheric delay statistics. Utilizing Eq. (2), we fix the coordinates of each station and apply forward–backward processing combined with LAMBDA-based IAR to maximize the number of successfully fixed ambiguities. For each baseline, ambiguity-fixed solutions are derived to ensure reliable separation of SD ionospheric delay effects from ambiguity errors. These fixed solutions enable the extraction of DD ionospheric delays on an epoch-by-epoch basis, using the satellite with the highest elevation angle as the reference. This process results in time series datasets that represent the variability of ionospheric delays for each baseline length.

The extracted DD ionospheric delay boundaries are assumed to vary approximately linearly with baseline length. This assumption is physically reasonable, as ionospheric delay decorrelation typically increases roughly linearly with the distance between stations. To simplify the model and enforce physical symmetry, we apply a symmetric treatment: the absolute value of the lower bound is averaged with the upper bound to form symmetric boundaries around zero. This approach reduces model complexity while maintaining adherence to physical expectations.

We employ ordinary least squares (OLS) regression (Pohlmann and Leitner 2003) to fit a linear model to the symmetric DD ionospheric delay boundary data. The fitting minimizes the sum of squared errors (SSE) between the observed boundary data points and the fitted linear curve, yielding the best linear parameters under the assumptions of independent, zero-mean, homoscedastic errors. The fitted linear function thus provides a baseline-length-dependent boundary:

$$\mathrm{DDI}(L) = a \times L + b \tag{21}$$

where $\mathrm{DDI}(L)$ is the boundary at baseline length L and a, b are the fitted slope and intercept coefficients.

We evaluate the fitting quality using the coefficient of determination R^2, defined as

$$R^2 = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}} \tag{22}$$

where SSE is the residual sum of squares and SST is the total sum of squares of the observed data. An R^2 value close to 1 indicates that the linear model explains most of the variance in the boundary data, confirming the reasonableness of the linear assumption. Values near zero would suggest poor fit and the need for more complex models.

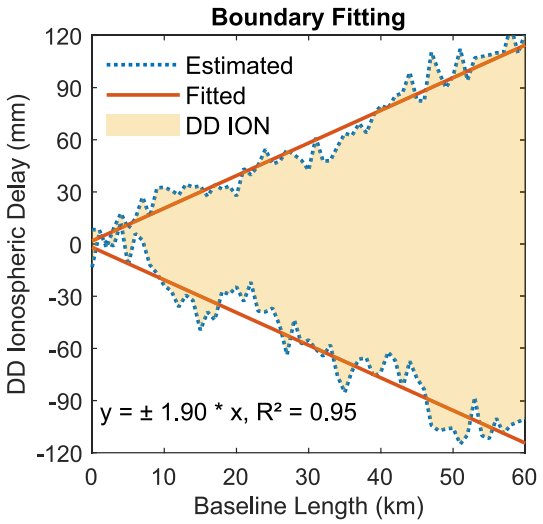

Fig. 5 Linear fit of DD ionospheric delay boundaries with baseline length

Figure 5 illustrates the fitted linear boundaries alongside the empirical DD ionospheric delay bounds extracted from the Hong Kong SatRef data. In our analysis, the DD ionospheric delay boundaries were modeled as symmetric linear functions of baseline length. The fitting process yields $a = \pm 1.9$ mm/km and $b = 0$, indicating that the upper and lower bounds of the DD ionospheric delay envelope can be represented as straight lines with slopes of + 1.9 mm/km and − 1.9 mm/km, respectively, and zero intercept. This means that the DD ionospheric delay is expected to increase or decrease linearly with baseline length at a rate of 1.9 mm per kilometer. The OLS linear fit achieved an excellent coefficient of determination, $R^2 = 0.95$, indicating that the linear model explains 95% of the variance in the observed ionospheric delay boundaries. This strong linear relationship confirms the physical expectation that ionospheric delay variability increases approximately linearly with baseline length in the Hong Kong region. In the subsequent development of the ionosphere-constrained model, these linear bounds incorporating the ± 1.9 mm/km slope will be directly applied as inequality constraints on the DD ionospheric delays.

Furthermore, our results show that varying the reference satellite does not significantly affect the statistical determination of the DD ionospheric delay boundaries. This invariance supports the robustness of our boundary estimation method and confirms that the constraints incorporated into the proposed model are physically meaningful and consistent regardless of the differencing scheme.

When the QP solution repeatedly lands exactly on the prescribed DD delay limits, it serves as an automatic anomaly

alert indicating that the ionospheric estimate may be untrustworthy. Rather than relying on fixed weights or manual resets, we employ a statistically driven expand-and-retry strategy that scales directly with the baseline length. As soon as a satellite's DD delay estimate binds the upper or lower bound, we temporarily widen that satellite's interval by applying a 30% expansion to the baseline-specific boundary DDI(L), resulting in expanded limits of $\pm 1.3\text{DDI}(L)$. This multiplicative scaling automatically adapts to baseline-dependent ionospheric variability while maintaining statistical rigor, as DDI(L) represents the 100% envelope of observed DD delays at baseline length L. The 30% expansion captures the vast majority of genuine ionospheric excursions without admitting extreme outliers, encompassing over 95% of observed DD delays across all baselines in the Hong Kong SatRef network while preserving the interpretability of hard bounds. If the expanded bounds still bind after resolving QP, we fall back to the unconstrained ionosphere-float solution for that epoch, ensuring that over-constraint cannot introduce bias. Compared to the ionosphere-weighted method requiring real-time variance estimation and subjective weight tuning, our boundary-adjustment mechanism avoids per-epoch variance computation while delivering equivalent or superior robustness by explicitly containing only truly anomalous estimates through this baseline-adaptive expansion strategy.

5.3 Experimental results

To systematically explore the differences in IAR performance among the four models, we categorized the baselines of the Hong Kong SatRef network based on their lengths. Specifically, we grouped the baselines into intervals of five kilometers, rounding each baseline length to the nearest node for statistical analysis. For instance, a baseline measuring 11 km was classified under the 10-km node, while a 14-km baseline was grouped with the 15-km node, and so on. This classification approach allowed for a structured comparison of model performance across varying baseline lengths, providing insights into how each model adapts to changes in baseline distance.

Figure 6 illustrates the IAR success rates for GPS-only data across the four models, where one week of data from the Hong Kong SatRef network is used and the strategies in Table 1 are adopted. The Kalman filter is restarted every two hours to provide statistically significant conclusions. The ionosphere-float model, due to its lack of constraints on ionospheric delays, exhibits a relatively low model strength, resulting in an IAR success rate of approximately 80%. This rate shows no significant correlation with baseline length, highlighting the model's limitations in handling ionospheric variability effectively. In contrast, the ionosphere-fixed model exhibits a marked correlation between IAR success rates and baseline length, with performance deteriorating as baseline length increases. Specifically, when the baseline exceeds 10 km, the ionosphere-fixed model performs worse than the ionosphere-float model. This decline can be attributed to the unrealistic assumption that the SD ionospheric delay is zero, which becomes increasingly untenable at longer baselines. Given that the ionosphere-fixed model records the lowest IAR success rates for baselines exceeding 10 km in Fig. 6, we will in the following analyses exclude the model from further performance comparisons.

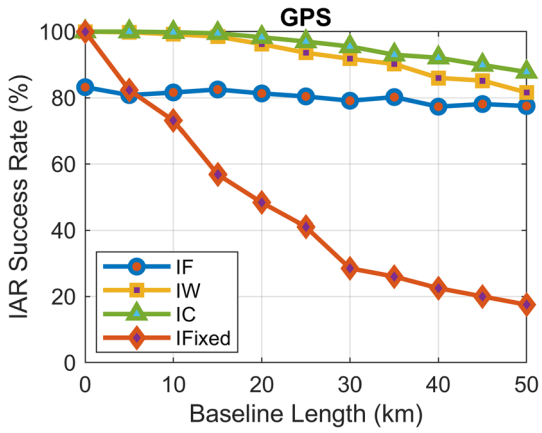

Fig. 6 IAR success rate for baselines of different lengths in Hong Kong SatRef over one week with GPS-only

The ionosphere-weighted model demonstrates excellent performance within 15 km, achieving an IAR success rate close to 100%. This performance is attributed to the model's assumption of the SD ionospheric delay being zero, which, when combined with the empirically derived baseline-length-dependent stochastic model (1.08 mm/km), allows for assigning appropriate uncertainty to ionospheric pseudo-observations based on the regional characteristics of the Hong Kong SatRef network. However, as baseline length increases, the assumption becomes less valid, leading to a significant decline in success rate. This decline is particularly pronounced in low-latitude regions like Hong Kong, where ionospheric activity is high.

The proposed ionosphere-constrained model maintains a nearly 100% IAR success rate within 15 km, showcasing the benefits of imposing constraints on DD ionospheric delays. Although the success rate decreases with longer baselines, the model consistently outperforms the ionosphere-weighted model. For instance, at a baseline length of 40 km, the success rates for the ionosphere-constrained and ionosphere-weighted models are 92.1% and 85.8%, respectively, representing an improvement of 7.3%. This improvement underscores the effectiveness of the ionosphere-constrained model in mitigating the adverse effects of ionospheric variability.

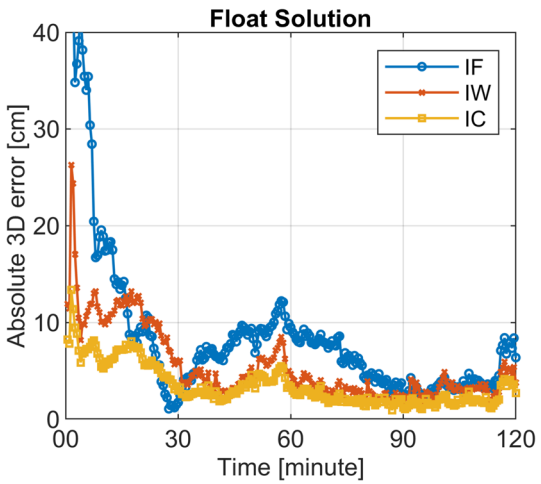

Fig. 7 Absolute 3D positioning errors of ambiguity-float solutions for the three models in a 2-h session for baseline HKKT-HKNP with a length of 28 km

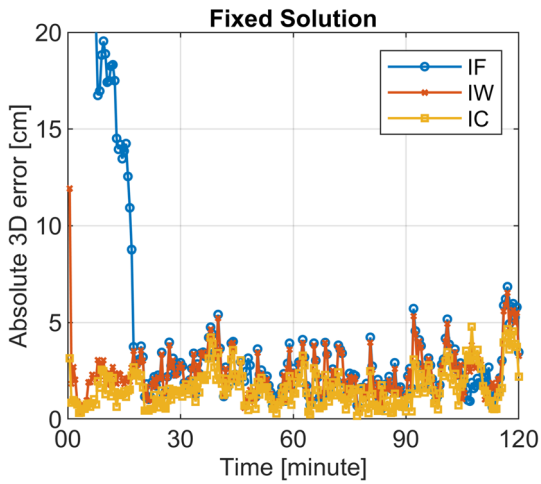

Fig. 8 Absolute 3D positioning errors of ambiguity-fixed solutions for the three models in a 2-h session for baseline HKKT-HKNP with a length of 28 km

The ionosphere-constrained model's ability to reduce errors in ionospheric parameters by constraining DD ionospheric delays is a key feature that enhances the estimation accuracy of other parameters, including the positioning parameters of interest. Figure 7 presents the absolute 3D positioning errors for the HKKT-HKNP baseline under ambiguity-float solutions, comparing the ionosphere-constrained model with the ionosphere-float and ionosphere-weighted models. The HKKT-HKNP baseline was selected, at 28 km, since it provides a challenging scenario for evaluating model performance.

The ionosphere-float model, which estimates ionospheric delays for each satellite, requires a longer convergence time and exhibits instability in the ambiguity-float solution. Using a 10 cm threshold, the ionosphere-float model takes approximately 15 min to converge. Additionally, its results exhibit larger fluctuations compared to the other two models due to the model strength. The ionosphere-weighted model, while also showing a good positioning convergence, achieves overall smaller errors compared to the ionosphere-float model. The initial convergence behavior largely depends on the stochastic model of the ionospheric pseudo-observations, which is a disadvantage of this model. More sophisticated modeling could address this, but it involves complex calculations and frequent updates, making it impractical, especially in regions with active ionospheric conditions like Hong Kong.

The proposed ionosphere-constrained model significantly outperforms the other two models in terms of accuracy and convergence time for the ambiguity-float solution, thanks to the benefits of ionospheric constraints. The ionosphere-constrained model converges to a 10 cm accuracy within just four minutes, much faster than the other models. Furthermore, its positioning error is substantially smaller, demonstrating the model's robustness, and reliability in challenging ionospheric conditions.

Figure 8 provides results for the fixed solutions, reinforcing the necessity of preprocessing ionospheric delay, either in an ionosphere-weighted or ionosphere-constrained form. The ionosphere-float model takes 16 min to successfully IAR, whereas the ionosphere-weighted and ionosphere-constrained models require approximately 3.5 and 2 min, respectively. The superior performance of the ionosphere-constrained model is expected, as the accuracy of the float solution is crucial for successful IAR, and the proposed model has an advantage over the ionosphere-weighted model in this regard.

Figure 9 presents the statistical results of the convergence times for all baselines in Hong Kong SatRef over a one-week period, utilizing a 10 cm threshold in the 3D direction, with the Kalman filter being restarted every two hours. These results further corroborate our previous conclusions, which can be summarized as follows. First, the convergence times for both the float and fixed solutions derived from the ionosphere-constrained model and the ionosphere-weighted model are significantly faster than those obtained from the ionosphere-float model. This improvement can be attributed to the effective constraints imposed by the ionosphere. Second, the ionosphere-constrained model applies bounds on the DD ionospheric delay, allowing for the optimal solution to be achieved through the WLS QP. Consequently, both the float and fixed solutions show marked enhancements compared to the ionosphere-weighted model, which relies on ionospheric pseudo-observations with predetermined weights. This finding underscores the meaningful role of WLS QP

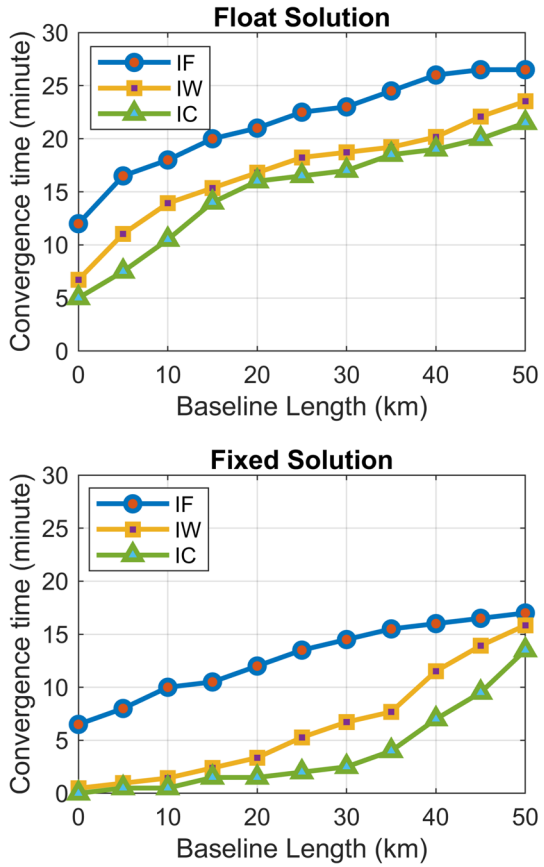

Fig. 9 Convergence time analysis for the 3D direction of float and fixed solutions to achieve 10 cm with GPS-only and a 30-s sampling interval across various baseline lengths over one week in Hong Kong SatRef

with additional inequality constraints in enhancing GNSS RTK positioning.

To intuitively compare the RTK positioning performance of the three models, we report the results for the baseline on Day of Year (DOY) 001 of 2024 as a representative example. Figure 10 shows the positioning results based on correctly fixed solutions for the three ionosphere processing strategies, depicting horizontal position scatter and vertical position time series. These results were obtained by comparing the estimated positions to precise benchmark coordinates derived from long-term observations.

From the comparison, it is evident that the positioning accuracy of the ionosphere-weighted model is improved compared to the ionosphere-float model due to the increased number of redundant observations with additional ionospheric pseudo-observations. Notably, compared to the ionosphere-float model, the accuracy of the ionosphere-weighted model in the east (E), north (N), and up (U) directions is improved by 57.1%, 50.0%, and 38.6%, respectively. The ionosphere-constrained model exhibits even higher positioning accuracy than the other two models, as demonstrated by the horizontal position scatter and vertical position time series. The positioning accuracy of the ionosphere-constrained model in the three directions is improved by 66.7%, 65.4%, and 59.6% compared to the ionosphere-float model, and by 22.2%, 30.8%, and 34.3% compared to the ionosphere-weighted model. These results validate the effectiveness of the proposed ionosphere-constrained model.

The ionosphere-constrained model delivers a clear advantage around UTC 07:00, when severe ionospheric disturbances disrupted positioning for all approaches. During this interval, the ionosphere-float method failed to resolve ambiguities and suffered large position errors due to the extreme ionospheric disturbances. The ionosphere-weighted method also underperformed as its fixed pseudo-observation weights could not accommodate the unexpected disturbance, leading to degraded positioning accuracy. By contrast, the ionosphere-constrained approach uses empirically derived bounds on double-differenced ionospheric delays to automatically intercept abnormal increments. When the unconstrained solution exceeds these bounds, the quadratic programming solver automatically truncates the estimate to the nearest limit and triggers either a boundary expansion procedure or a fallback to the unconstrained solution if necessary. This "safety-valve" mechanism prevents physically implausible ionospheric increments from polluting the ambiguity estimation process, and it activates only when needed, leaving the solution unchanged under normal conditions. As a result, the ionosphere-constrained model rapidly recovers float solutions and achieves high ambiguity-resolution success rates even during intense ionospheric activity, ensuring more reliable RTK positioning.

6 Conclusions

This paper introduces an ionosphere-constrained RTK model that formulates DD ionospheric delay estimation as a convex quadratic programming problem with simple baseline-length-dependent bounds derived from historical data. Unlike the ionosphere-weighted model, which demands complex variance component estimation and iterative weight tuning sensitive to local ionospheric dynamics, our method applies transparent and physically interpretable bounds. Its core innovation is a conditional activation mechanism: When the unconstrained float solution exceeds empirical limits, the optimizer either adjusts the estimate to a feasible interior value within the prescribed bounds or directly projects the result to the corresponding boundary, which provides a built-in anomaly warning. If the unconstrained solution naturally satisfies all constraints, the model simply returns the original float solution with negligible overhead. This constraint-driven approach improves float precision and accelerates IAR, resulting in faster and more accurate RTK positioning without introducing additional stochastic complexity.

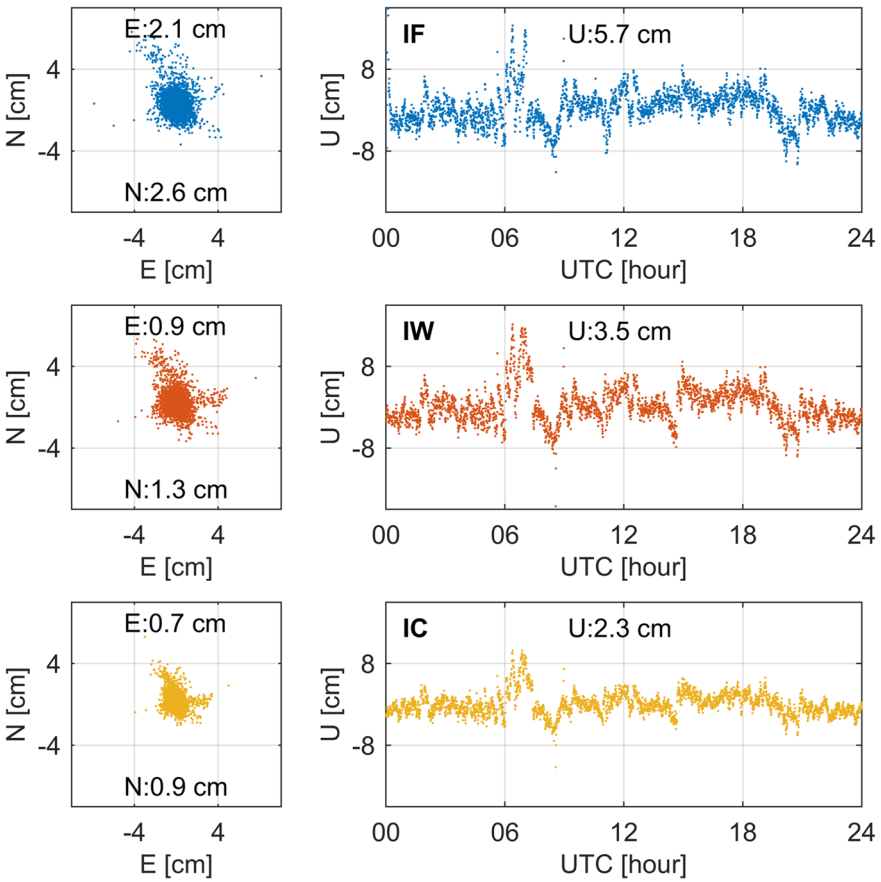

Fig. 10 Horizontal position scatter and vertical time series (IF top row, IW middle row, and IC bottom row) with ambiguity-fixed solutions for the three models for baseline HKKT-HKNP with a length of 28 km

Utilizing controlled experimental configurations, we demonstrate the critical importance of appropriately chosen ionospheric constraints in enhancing estimation accuracy through comprehensive sensitivity analysis. While overly lenient constraints fail to impact the solution, appropriately calibrated constraints guide DD ionospheric estimates toward physically reasonable values, thereby improving the accuracy of essential parameters such as positioning and float ambiguities. The analysis of error distributions further underscores the efficacy of well-calibrated constraints, as they improve the precision and reliability of parameter estimates. The systematic evaluation demonstrates that properly configured constraints significantly enhance the IAR success rate, with performance improvements directly correlating with constraint accuracy levels.

The performance evaluation using the Hong Kong SatRef further validates the significant advantages of the proposed ionosphere-constrained model in RTK positioning. The results highlight the superior performance of the proposed model in IAR and positioning accuracy. The ionosphere-constrained model consistently achieves nearly 100% IAR success rates within 15 km and maintains a distinct advantage over the ionosphere-weighted model at longer baselines, such as 40 km, where it improves success rates by 7.3%. This model's ability to impose effective DD ionospheric constraints enhances the estimation accuracy of critical parameters, including position and float ambiguities, and significantly reduces convergence time. The findings confirm that the ionosphere-constrained model not only accelerates convergence but also achieves higher positioning accuracy, improving by over 20% compared to the ionosphere-weighted model.

This research offers a novel solution to the challenges of weak model strength and the difficulty of achieving fast IAR in the traditional ionosphere-float model. By integrating conditional constraint optimization within the weighted least squares framework, this study demonstrates the potential of constraint-based approaches for fast and precise GNSS positioning. Future work will focus on three key areas: proving DD constraint invariance to reference satellite selection, integrating robust estimators into the QP framework, and developing adaptive boundary scaling for extreme ionospheric conditions. These investigations will refine theoretical foundations and broaden operational applicability of constraint-based RTK positioning.

Acknowledgements This work was supported in part by the University Grants Committee of Hong Kong under the General Research Fund (GRF) (Grant No. 15212525 and 15229622), the National Natural Science Foundation of China (Grant No. 42404052), and the Research

Centre for Low-Altitude Economy (RCLAE), The Hong Kong Polytechnic University (Grant No. P0058167).

Author contributions XM and WC designed this method; XM and XC conducted the experiments and analyzed the data. XM drafted this manuscript; RO, YW, and YY revised this manuscript.

Data availability The RINEX data with sampling rate at the 30 s of the Hong Kong Satellite Positioning Reference Station Network are available for download at: https://rinex.geodetic.gov.hk/rinex3/

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Altman A, Gondzio J (1999) Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. Optim Methods Softw 11(1–4):275–302

Amiri-Simkooei A, Teunissen P, Tiberius C (2009) Application of least-squares variance component estimation to GPS observables. J Survey Eng 135(4):149–160

Best M, Chakravarti N (1990) Active set algorithms for isotonic regression - a unifying framework. Math Program 47(1):425–439

Best M, Kale J (2000) Quadratic programming for large-scale portfolio optimization. In: Financial Services Information Systems. Auerbach Publications, pp 531–548

Boggs P, Tolle J (1995) Sequential quadratic programming. Acta Numer 4:1–51

Boland N (1997) A dual-active-set algorithm for positive semi-definite quadratic programming. Math Program 78(1):1–27

Bretthauer K, Shetty B (1997) Quadratic resource allocation with generalized upper bounds. Oper Res Lett 20(2):51–57

Cai Y, Cheng P, Meng X, Tang W, Shi C (2011) Using network RTK corrections and low-cost GPS receiver for precise mass market positioning and navigation applications. In: 2011 IEEE Intelligent Vehicles Symposium (IV), pp 345–349

Cimini G, Bemporad A (2017) Exact complexity certification of active-set methods for quadratic programming. IEEE Trans Automat Contr 62(12):6094–6109

Coleman T, Hulbert L (1989) A direct active set algorithm for large sparse quadratic programs with simple bounds. Math Program 45(1):373–406

Fan J, Zhang L (1998) Real-time economic dispatch with line flow and emission constraints using quadratic programming. IEEE Trans Power Syst 13(2):320–325

Fesanghary M, Mahdavi M, Minary-Jolandan M, Alizadeh Y (2008) Hybridizing harmony search algorithm with sequential quadratic programming for engineering optimization problems. Comput Methods Appl Mech Eng 197(33–40):3080–3091

Frank M, Wolfe P (1956) An algorithm for quadratic programming. Nav Res Logist Q 3(1–2):95–110

Grejner-Brzezinska D, Kashani I, Wielgosz P (2005) On accuracy and reliability of instantaneous network RTK as a function of network geometry, station separation, and data processing strategy. GPS Solut 9:212–225

Hadas T, Teferle F, Kazmierski K, Hordyniec P, Bosy J (2017) Optimum stochastic modeling for GNSS tropospheric delay estimation in real-time. GPS Solut 21:1069–1081

Jiang H, Ralph D (2000) Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. SIAM J Optim 10(3):779–808

Landau H, Chen X, Kipka A, Vollath U (2007) Latest developments in Network RTK modeling to support GNSS modernization. In: Proceedings of the 2007 national technical meeting of the institute of navigation pp 851–859.

Li B, Shen Y, Feng Y, Gao W, Yang L (2014) GNSS ambiguity resolution with controllable failure rate for long baseline network RTK. J Geod 88:99–112

Li P, Cui B, Hu J, Liu X, Zhang X, Ge M, Schuh H (2022) PPP-RTK considering the ionosphere uncertainty with cross-validation. Satell Navig 3(1):10

Luo Z, Ma W, So A, Ye Y, Zhang S (2010) Semidefinite relaxation of quadratic optimization problems. IEEE Signal Process Mag 27(3):20–34

Mi X, Zhang B, Yuan Y (2019a) Multi-GNSS inter-system biases: estimability analysis and impact on RTK positioning. GPS Solut 23:1–13

Mi X, Zhang B, El-Mowafy A, Wang K, Yuan Y (2023) Undifferenced and uncombined GNSS time and frequency transfer with integer ambiguity resolution. J Geod 97(2):13

Mi X, Zhang B, Yuan Y (2019b) Stochastic modeling of between-receiver single-differenced ionospheric delays and its application to medium baseline RTK positioning. Meas Sci Technol 30(9):095008

Momoh J, Guo S, Ogbuobiri E, Adapa R (1994) The quadratic interior-point method solving power-system optimization problems. IEEE Trans Power Syst 9(3):1327–1336

Moore T, Hill C, Norris A, Hide C, Park D, Ward N (2008) The potential impact of GNSS/INS integration on maritime navigation. J Navig 61(2):221–237

Odijk D (2000a) Improving ambiguity resolution by applying ionosphere corrections from a permanent GPS array. Earth Planets Space 52:675–680

Odijk D (2000b) Stochastic modelling of the ionosphere for fast GPS ambiguity resolution. In: Geodesy beyond 2000: the challenges of the first decade IAG general assembly birmingham, pp. 387–392

Odijk D (2000c) Weighting ionospheric corrections to improve fast GPS positioning over medium distances. In: Proceedings of the 13th international technical meeting of the satellite division of the institute of navigation (ION GPS 2000) pp 1113–1123

Odijk D, Khodabandeh A, Nadarajah N, Choudhury M, Zhang B, Li W, Teunissen P (2017) PPP-RTK by means of S-system theory: Australian network and user demonstration. J Spatial Sci 62(1):3–27

Odijk D, Van Der Marel H, Song I (2000) Precise GPS positioning by applying ionospheric corrections from an active control network. GPS Solut 3:49–57

Odolinski R, Teunissen P, Odijk D (2015) Combined GPS plus BDS for short to long baseline RTK positioning. Meas Sci Technol 26(4):045801

Parkins A (2011) Increasing GNSS RTK availability with a new single-epoch batch partial ambiguity resolution algorithm. GPS Solut 15:391–402

Pohlmann J, Leitner D (2003) A comparison of ordinary least squares and logistic regression (1). Ohio J Sci 103(5):118–126

Psychas D, Verhagen S, Liu X, Memarzadeh Y, Visser H (2019) Assessment of ionospheric corrections for PPP-RTK using regional ionosphere modelling. Meas Sci Technol 30(1):014001

Rizos C (2007) Alternatives to current GPS-RTK services and some implications for CORS infrastructure and operations. GPS Solut 11:151–158

Serafini T, Zanghirati G, Zanni L (2005) Gradient projection methods for quadratic programs and applications in training support vector machines. Optim Methods Softw 20(2–3):353–378

Shen Y, Li B, Xu G (2009) Simplified equivalent multiple baseline solutions with elevation-dependent weights. GPS Solut 13:165–171
Teunissen P (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geod 70(1):65–82
Teunissen P (2018) Distributional theory for the DIA method. J Geod 92(1):59–80
Teunissen P, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. Surv Rev 41(312):138–151
Teunissen P (1997) The geometry-free GPS ambiguity search space with a weighted ionosphere. J Geod 71:370–383
Teunissen P (1998) The ionosphere-weighted GPS baseline precision in canonical form. J Geod 72:107–117
Teunissen P (2004) Penalized GNSS ambiguity resolution. J Geod 78(4):235–244
Teunissen P, Amiri-Simkooei A (2008) Least-squares variance component estimation. J Geod 82:65–82
Teunissen P, Khodabandeh A (2015) Review and principles of PPP-RTK methods. J Geod 89(3):217–240
Teunissen P, Khodabandeh A (2021) A mean-squared-error condition for weighting ionospheric delays in GNSS baselines. J Geod 95(11):118
Xu H (2012) Application of GPS-RTK technology in the land change survey. Proc Eng 29:3454–3459
Yang T, Sun D (2020a) Global navigation satellite system fault detection protection level using parametrized quadratic programming. J Aerosp Inf Syst 17(11):615–623
Yang T, Sun D (2020b) Global navigation satellite systems fault detection and exclusion: a parameterized quadratic programming approach. IEEE Trans Aerosp Electron Syst 56(4):2862–2871
Zhou P, Wang J (2013) Stochastic ionosphere models for precise GNSS positioning: sensitivity analysis. J Glob Position Syst 12(1):53–60