A new method for fault identification in real-time integrity monitoring of autonomous vehicles positioning using PPP-RTK

Hassan Elsayed[1,2], Ahmed El-Mowafy[1], Kan Wang[3,4]

[1]*School of Earth and Planetary Sciences, Curtin University, Perth, Australia*

[2]*Faculty of Engineering, Suez Canal University, Ismailia, Egypt*

[3]*National Time Service Center, Chinese Academy of Sciences, Xi'an, China*

[4]*University of Chinese Academy of Sciences, Beijing, China*

Abstract

Autonomous vehicles require a real-time positioning system with in-lane accuracy. They also require an autonomous onboard integrity monitoring (IM) technique to verify the estimated positions at a pre-defined probability. This can be computationally demanding. PPP-RTK is a promising positioning technique that can serve this purpose. Since PPP-RTK is developed to process undifferenced and uncombined (UDUC) observations for both network and user sides, it provides the residuals of the individual measurements. This can be exploited to reduce the computational load consumed in the fault detection and exclusion (FDE) process, included in the IM task, without compromising the positioning availability. This research proposes filtering the faulty satellites by the network, then the hardware and location-dependent faults at the user end can be identified. This is achieved by calculating the ratio between the matching UDUC residuals of the user receiver and the nearest reference station observations. This ratio is used to rank the individual observations where the observation with the largest ratio is most likely to be the faulty one. Therefore, it is more likely to identify the faulty observation without generating and testing numerous subsets. In addition, the exclusion can be attempted per observation, which preserves observation availability, unlike the grouping techniques that perform the exclusion per satellite. The method was examined in two test cases where geodetic and commercial receivers were used. Results show that the computational load has been reduced significantly by about 85-99% compared to the solution separation and Chi-squared test methods that are commonly used for FDE.

Keywords

Integrity monitoring, fault detection and identification, autonomous vehicles, PPP-RTK, undifferenced and uncombined.

Introduction

Autonomous vehicles (AVs) require real-time positioning capabilities. This also necessitates a rigorous integrity monitoring (IM) technique to assure the reliability of the computed positions at a pre-defined probability (El‐Mowafy and Kubo 2018; Hassan et al. 2021; Wörner et al. 2016). Although many sensors are involved in operating and monitoring AVs (Li et al. 2022; Sasani et al. 2016), this work is concerned with IM of positioning based on using Global Navigation Satellite Systems (GNSS). One main step of IM is fault detection and exclusion (FDE). It is considered the most computationally demanding process of IM. This can oppose the implementation of IM for real-time applications. Many methods have been introduced to reduce the computational load. Some approaches propose selecting a limited number of satellites among the all-in-view satellites. This selection can be made based on a priori selection algorithms such as those that have a better elevation angle, weighting factors, and satellite health (Gerbeth et al. 2016; Walter et al. 2016). The reduction in the selected satellites will lead, accordingly, to a reduction in the computational load of the IM process. However, this comes at the cost of reducing IM availability. In addition, it removes satellites that may be fault-free and keeps others that may contain faults that will be removed later, hence, reducing the availability even further.

Some other approaches suggest performing what is called fault consolidation, known as Clustered Advanced receiver autonomous integrity monitoring (ARAIM). In such approaches, many satellites are combined and attempted for exclusion together to cover different fault modes in one processing. Different satellite grouping techniques were described by (Orejas and Skalicky 2016; Orejas et al. 2016). For instance, (Ge et al. 2017) attempted clustering and excluding the satellites of the same orbit plane. However, this cannot be conducted with all constellations since only the BDS constellation has different orbits. (Walter et al. 2014) clustered the satellites of the same constellation ignoring some fault modes that are less likely to happen; thus, reducing the number of the tested subsets and the computational complexity. Unlike the previous two approaches, where the number of the generated subsets changes based on the number of satellites and the fault probabilities of both the satellites and constellations, (Blanch et al. 2018) sought a further reduction in the computational load by fixing the number of the tested subsets regardless the fault probabilities of both the satellites and constellations. Although these methods managed to reduce the computational burden, this comes at the expense of compromising the availability, conservatism, and precision of the protection level (PL).

Other approaches examined the reduction of the computational load through saving in the mathematical process of estimating the solution of the tested subsets. For example, (Gunning et al. 2018) suggested that all the calculated models and corrections for the all-in-view situation can be used to estimate the solution of the generated subsets during the solution separation (SS) test, given that they are close to each other. While this included some saving in the computational load, it also included some approximation and can be only applicable, with some concerns on its impact in different situations, in some positioning techniques such as precise point positioning (PPP). Similarly, (Blanch et al. 2019) tested the replacement of the residuals covariance matrix of each subset solution, which is very computationally demanding as it requires two matrix inversions by another matrix obtained in the all-in-view case. No full matrix inversion is needed in that case. Whereas this can reduce the computational load significantly, it degrades the PL to a great extent as well. This may be accepted for some applications where up to several meters of accuracy is authorized, but this is not the case for autonomous vehicles where only in-lane accuracy is of main concern. Furthermore, (El-Mowafy and Wang 2022) proposed a method where the inverse of the covariance matrix of the state vector for any generated subset, considering single or multiple faults, is computed without inversion from the single, all-in view, normal matrix without any further inversion. It proved that this could reduce the complexity of the calculations substantially without compromising the solution availability or quality.

All of the aforementioned approaches managed to provide means to reduce the computational load. In most cases, this has adversely affected other parameters, which might be acceptable for some applications. The shared part among all of these research works is that the user still needs to test all the generated subsets to identify the faulty observation(s)/satellite(s).

We present a new process for FDE that potentially reduce the computational time compared to the current methods. The PPP-RTK approach is selected as the most suitable to provide the needed in-lane accuracy for AVs in real-time, as will be discussed in the next section. In the proposed FDE method, faults due to satellite errors will be checked by the network processing center exploiting the known ground positions of the network stations. This shall reduce the risk of having a fault due to observed satellites at the user end. While the atmospheric, location and receiver-related errors at the user end will be checked using a ratio between the residuals of the user position solution and their counterparts of the nearest reference station from the network, assumed to be fault-free. This ratio shall provide the user with an indicator of which observation(s) could be faulty and worth attempting testing for exclusion in case of a fault is detected in the overall solution. Therefore, the computational load is expected to be significantly reduced as there will be no need to form all possible subsets to identify the faulty observation(s). In addition, the availability will be maintained since the exclusion is based on the observations not satellite(s), thanks to using UDUC PPP-RTK as a positioning technique. For example, if dual frequency receiver is used where each satellite offers two code and two phase observations, the exclusion will suspect all four individual observations. This is unlike the current grouping technique where a faulty observation causes the exclusion of the satellite including all its four observations (or more in case more frequencies are observed). The method can also be combined with any of the previously stated methods to reduce the computational load even further.

The next section briefly discusses the used PPP-RTK technique and explains the rationale for its selection for the positioning of AVs. A full description of the newly proposed method, including the criteria and advantages, as well as the testing examples, is provided in the following section. In the fourth section, the results of applying the proposed new method in two different test cases are presented and discussed. The conclusion and the future work are given in the last section.

PPP-RTK for real-time positioning of autonomous vehicles

Not all GNSS-based positioning methods are suitable for AVs. For example, conventional PPP requires a long initialization time before providing a valid position (Du et al. 2021). Traditional RTK requires a dense infrastructure of base stations and a radio connection that may sometimes be interrupted (El-Mowafy and Kubo 2017). The use of networks provides redundancy and consistency of the positioning output, therefore, operating a reference network rather than multiple single reference stations would provide a more efficient solution in terms of cost against the covered area (Landau et al. 2003; Vollath et al. 2002). The methodologies of the network-RTK (NRTK) and its corrections transmitted to users differ, which include the virtual reference station (VRS) (Wanninger 2003), area broadcast (FKP), and master-auxiliary (Mac) methods (Janssen 2009; Takac and Zelzer 2008). Based on the utilized protocol, the infrastructure of the network as well as the user software can be defined. Unlike these methods, where processing is based on differencing techniques, PPP-RTK represents another method where both the network and the user process undifferenced and uncombined (UDUC) observations (Zha et al. 2021).

The corrections sent to users to deal with the observation errors are classified into the observation-state representation (OSR) and the state-space representation (SSR) (Wabbena et al. 2005). While the first provides the corrections for the combined errors, the second provides them for each error source individually, as shown in Fig. 1. Hence, differential GPS (DGPS), traditional RTK, and traditional NRTK (where a differencing technique is used) employ the OSR protocol, whereas PPP, satellite-based augmentation systems (SBAS) (Chen et al. 2022), and PPP-RTK usually use the SSR protocol. PPP-RTK has the following practical advantages compared to the rest of the methods (Zhang et al. 2019): 1) the ability to study the impact of each error source; 2) since it processes the UDUC observations, the calculated residuals are obtained for each UDUC observation, which provides the possibility of better screening of individual observations; 3) unlike most traditional NRTK methods (e.g., VRS, Mac) that require two-way communications between the network and the user(s), PPP-RTK only requires a one-way communication system with all users within the coverage area of the network; hence, reducing security and bandwidth hazards. Due to these advantages, this work proposes PPP-RTK as the most convenient method for AVs real-time positioning. The following section provides more details concerning the PPP-RTK at both the network and the user sides, based on which our method is presented.

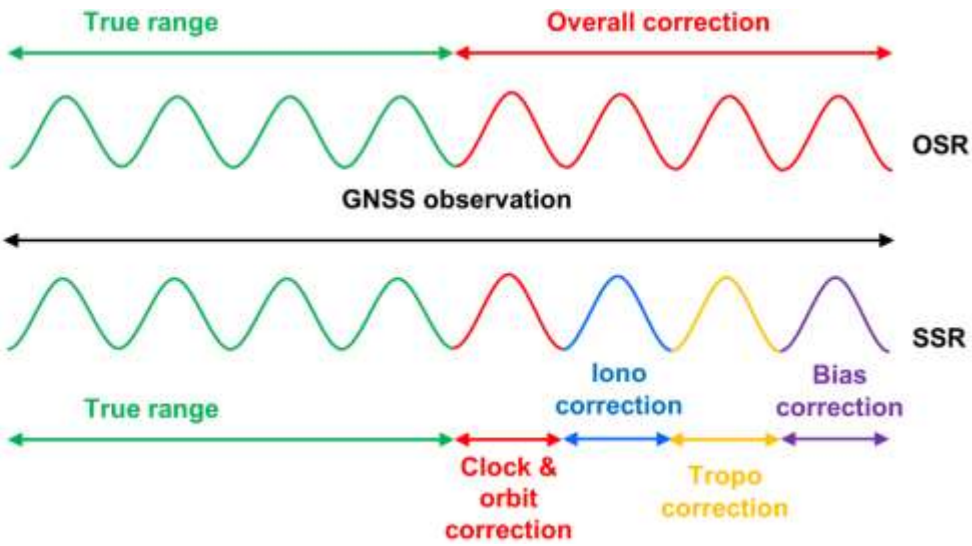

Fig. 1 Sketch of the difference between the observation-state representation (OSR) and the space-state representation (SSR)

Network Processing in the SSR mode (UDUC PPP-RTK)

PPP-RTK processing takes place on both the network and the user sides. The observation equations at the network end can be expressed as (Odijk et al. 2017):

$$E(C_R^s) = c(\bar{t}_R - \bar{t}^s) + \bar{T}_{R_W} + \mu_i \bar{I}_{R_i}^s \tag{1}$$

$$E(\varphi_R^s) = c(\bar{t}_R - \bar{t}^s) + \bar{T}_{R_W} - \mu_i \bar{I}_{R_i}^s + \frac{c}{f_i}\bar{N}_{R_i}^s + \bar{\delta}_{R_i} - \bar{\delta}_i^s \tag{2}$$

where $E(.)$ is the expected value of the observed minus computed terms; C_R^s, φ_R^s are pseudorange code and phase observations (in m); s, R refer to the observed satellite and the receiver, respectively; c represents the speed of light; t_R, t^s are the satellite and receiver clock offsets, respectively; $T_{R_W}$is the wet part of troposphere delay at the slant angle; $I_{R_i}^s$ represents the ionospheric delay on the first frequency; f_i is the frequency of the observed signal i; $N_{R_i}^s$ is the phase ambiguity; δ_{R_i} is the phase bias; $\mu_i = f_1^2 / f_i^2$ is a multiplier factor based on the frequency; $(\bar{.})$ denotes a certain representation that has been used to eliminate the rank deficiency using the S-system theory (Teunissen 1985) as shown in Table 1:

Table 1 PPP-RTK parameters representation of the network model

Parameter	Definition
$\bar{\boldsymbol{t}}^{\boldsymbol{s}}$	$t^s - t_M + [d_{IF}^s - d_{M_{IF}} - T_{M_w}]/c$
$\bar{\boldsymbol{\delta}}_{\boldsymbol{i}}^{\boldsymbol{s}}$	$\delta_i^s - d_{IF}^s + \mu_i d_{GF}^s - \delta_{M_i} + d_{M_{IF}} - \mu_i d_{M_{GF}} - \lambda_i N_{M_i}^s$
$\bar{\boldsymbol{N}}_{\boldsymbol{R_i}}^{\boldsymbol{s}}$	$\left(N_{R_i}^s - N_{M_i}^s\right) - \left(N_{R_i}^{s_P} - N_{M_i}^{s_P}\right)$
$\bar{\boldsymbol{T}}_{\boldsymbol{R_W}}$	$T_{R_W} - T_{M_W}$
$\bar{\boldsymbol{t}}_{\boldsymbol{R}}$	$t_R - t_M + (d_{R_{IF}} - d_{M_{IF}})/c$
$\bar{\boldsymbol{\delta}}_{\boldsymbol{R_i}}$	$\left(\delta_{R_i} - d_{R_{IF}} + \mu_i d_{R_{GF}} + \lambda_i N_{R_i}^{s_P}\right) - \left(\delta_{M_i} - d_{M_{IF}} + \mu_i d_{M_{GF}} + \lambda_i N_{M_i}^{s_P}\right)$
$\bar{\boldsymbol{I}}_{\boldsymbol{R_i}}^{\boldsymbol{s}}$	$I_{R_i}^s + d_{R_{GF}} - d_{GF}^s$

where M is the master station among the reference stations; $d_{M,R_{IF,GF}}$ represents the ionospheric-free (IF) and geometry-free (GF) linear combinations of the code bias for reference station (R) or master station (M); $d_{IF,GF}^s$ is the ionospheric-free (IF) and geometry-free (GF) linear combinations of the code bias for satellite; $\lambda_i = c/f_i$; s_P denotes the pivot satellite among the observed satellites.

The above equations are given for networks using dual-frequency observations, which is the case we considered. The form of the estimable parameters would be similarly modified in case more frequencies are involved.

User Processing in the SSR mode (UDUC PPP-RTK)

The user can apply the PPP-RTK processing technique by exploiting the corrections provided by the network, where at a user end:

$$E\left(C_U^s + c(\bar{t}^s)\right) = r_U^s + c\left(\bar{\bar{t}}_U\right) + \bar{T}_{U_W} + \mu_i \bar{\bar{I}}_{U_i}^s + \mu_i \bar{\bar{d}}_{U_{GF}} + e_{U_C}^s \quad (3)$$

$$E\left(\varphi_U^s + c(\bar{t}^s) + \bar{\delta}_i^s\right) = r_U^s + c\left(\bar{\bar{t}}_U\right) + \bar{T}_{U_W} - \mu_i \bar{\bar{I}}_{U_i}^s + \lambda_i \bar{N}_{U_i}^s + \bar{\bar{\delta}}_{U_i} + \epsilon_{U_\varphi}^s \quad (4)$$

where U is the user receiver; r_U^s represents the range between satellite s and the user; $e_{U_C}^s$, $\epsilon_{U_\varphi}^s$ are the code and phase observation noises that may include multipath and other location-dependent errors; $(\bar{\bar{.}})$ denotes a certain representation that is used to eliminate the rank deficiency using the S-system theory, as shown in Table 2.

Table 2 PPP-RTK parameters representation of the user model

Parameter	Definition
$\bar{\bar{\boldsymbol{t}}}_{\boldsymbol{U}}$	$[t_U + (d_{U_{IF}}/c)] - [t_M + (d_{M_{IF}}/c)]$
$\bar{\boldsymbol{I}}^{s}_{\boldsymbol{R_i}}$	$I^{s}_{U_i} + d_{M_{GF}} - d^{s}_{GF}$
$\bar{\bar{\boldsymbol{d}}}_{\boldsymbol{U_{GF}}}$	$d_{U_{GF}} - d_{M_{GF}}$
$\bar{\bar{\boldsymbol{\delta}}}_{\boldsymbol{U_i}}$	$\left(\delta_{U_i} - d_{U_{IF}} + \lambda_i N^{sp}_{U_i}\right) - (\delta_{M_i} - d_{M_{IF}} + \lambda_i N^{sp}_{M_i})$

IM of real-time positioning

ARAIM is an efficient method (Blanch et al. 2012; Blanch et al. 2011) that can be used for IM of the positioning of AVs. However, the current ARAIM methods were basically developed for aviation applications, which have many differences compared to ground applications, as discussed in (Elsayed et al. 2023). This encouraged many researchers to try to adapt ARAIM to different positioning techniques such as PPP (Du et al. 2021), RTK (Wang and El-Mowafy 2021), and PPP-RTK (Zhang et al. 2023). One main step of ARAIM is FDE. It is considered a computationally expensive step, due to the need to test a huge number of possible observation faults, which represents a hurdle for the implementation of IM in real-time applications. The SS and Chi-squared test methods are traditionally used for FDE in ARAIM (Joerger and Pervan 2014; Joerger and Pervan 2016). In the SS test, the position solution of different subsets, excluding the observation(s) that is/are checked for being faulty, one at a time, is estimated and compared against the position solution of the all-in-view observations solution at each epoch. The test statistic is expressed as:

$$\Delta\hat{x}^{k_{1\ldots n}} = \hat{x}^{k_{1\ldots n}} - \hat{x}^{0} \qquad (5)$$

where $\hat{x}^{0}$ is the all-in-view solution, while $\hat{x}^{k}$ is the solution of the subset $k_{1\ldots n}$ where n is the number of the tested subsets. This number is based on the selected fault modes that define how many suspected faulty observations should be considered for exclusion. Then, the solution difference $\Delta\hat{x}^{k_{1\ldots n}}$ of each subset is compared against a statistical threshold value as per equations (6) & (7):

$$\tau_{k_{1\ldots n}} = \frac{|\Delta\hat{x}^{k_{1\ldots n}}|}{T_{k_{1\ldots n}}} \leq 1 \;\; or \; |\hat{x}^{k_{1\ldots n}} - \hat{x}^{0}| \leq T_{k_{1\ldots n}} \tag{6}$$

$$T_{k_{1\ldots n}} = H_{fa}\, \sigma_{k_{1\ldots n}} \tag{7}$$

where assuming that the fault-free observation errors will have a normal distribution; H_{fa} is the quantile of CDF normal distribution, which is calculated based on the assigned probability for false alert (fa) and the total number of the considered fault modes, whereas $\sigma_{k_{1\ldots n}}$ is the standard deviation. The number of the required tests would be numerous when considering the possibility of concurrently multiple faulty measurements, not only single faults as mostly considered in the literature, and accordingly, the computational load would be huge.

In the case of the Chi-squared test, the residuals of the estimated position solution are scanned every epoch for faults as follows:

$$\hat{r}_U{}^T\, Q_{y_U}^{-1}\, \hat{r}_U \;\leq \chi^2_{\alpha}(df, 0) \tag{8}$$

where $\hat{r}_U$ is the residuals vector of the position solution; Q_{y_U} is the covariance matrix of the observations; α is the significance value that is decided based on the design of the IM process and the application on hand, e.g., 0.001; df is the degree of freedom, i.e., the difference between the number of observations and unknowns at the tested epoch. If the test fails at any epoch, a fault is assumed present, and exclusion must be attempted to identify that fault. The identification starts by reapplying the test on all possible subsets based on the defined fault modes. Once the test passes for a certain subset, the excluded satellite(s) is considered to be faulty. Unlike the SS test that is performed at every epoch for all possible subsets, the Chi-squared only starts the exclusion attempts when the all-in-view satellite observations solution does not pass. However, this represents a computational burden for real-time positioning when a fault is suspected. Another drawback that both the SS and Chi-squared methods share in the commonly used grouping approaches is that the subsets formation is based on testing satellites, grouping all their observations, not individual observations. Although this helps in reducing the computational load, it –unnecessarily- removes healthy observations (i.e. results in loss of information, which may be vital since real-time positioning involves a finite number of observations) and compromises the availability since all observations of the suspected satellite(s), not specifically the faulted observations, are excluded when a fault is detected.

Proposed Methodology

We classified faults into two main types according to their source. The first type is the faults due to satellite errors. This kind of fault risks both the network stations and user receivers altogether. Therefore, it is proposed that the faulty satellites due to satellite errors are checked by the network reference stations exploiting their known positions, where testing is performed in real-time in the PPP-RTK scheme. An RTCM message is proposed to be transmitted in near-real time to the network subscribers that contains a list of these faulty satellites. Users shall exclude the listed satellites before the positioning process.

The second type of fault is due to the user receiver, anomalies in the atmosphere corrections, and location-related errors such as anomalous multipath. This kind of fault emerges only on the user side and cannot be detected by the network. Therefore, a receiver autonomous integrity monitoring protocol is required by the user. In this approach, the calculated residuals of every position solution are scanned for faults using Chi-squared testing (equation (8), assuming that the squared residuals in the unbiased fault-free case would follow a Chi-squared distribution). In the case of detecting a fault, an identification process is required to identify which observations are faulty. As discussed earlier, this is a computationally expensive step that may affect the real-time positioning performance. Therefore, it is proposed, in this new method, to calculate the absolute values of the ratio between the residuals of the individual observations of the user position solution and their counterparts of the nearest reference station, assuming that the latter, after elimination of the faulty satellites, are fault-free. There is no need to use normalized values, where the elevation-angle-weighted model is typically used, and the distances between the network stations are relatively short, i.e., < 100 km, such that the standard deviations that are used to normalize the residuals would be almost the same at the rover and the reference station. The procedure can be expressed as follows:

$$\hat{r}_{U_{C,\varphi}}^{S_{1\dots m}} = y_{U_{C,\varphi}} - G_U \hat{x}_U \qquad (9)$$

$$ratio = \begin{bmatrix} \hat{r}_{U_{C_{i=1}}}^{S_{1\ldots m}} / \hat{r}_{M_{C_{i=1}}}^{S_{1\ldots m}} \\ \hat{r}_{U_{C_{i=2}}}^{S_{1\ldots m}} / r_{M_{C_{i=2}}}^{S_{1\ldots m}} \\ \hat{r}_{U_{\varphi_{i=1}}}^{S_{1\ldots m}} / \hat{r}_{M_{\varphi_{i=1}}}^{S_{1\ldots m}} \\ \hat{r}_{U_{\varphi_{i=2}}}^{S_{1\ldots m}} / \hat{r}_{M_{\varphi_{i=2}}}^{S_{1\ldots m}} \end{bmatrix} \quad (10)$$

where m is the number of the observed satellites by the user after excluding the suspected satellites by the network; $y_{U_{C,\varphi}}$ is the user code and phase observations vector; and G_U is the design matrix.

Based on the calculated ratios, the observations can be ranked in descending order for their likelihood of containing faults. The observation of the largest ratio is considered the most suspected to be faulty. Fig. 2 shows an example of the ranking criterion. It shows that in the case of ten satellites in view where two frequencies are tracked, the observations shaded in red refer to the observations with the largest ratios, while those shaded in green have the lowest ratios. The exclusion shall be attempted with the observation with the largest ratio. In addition, their highly correlated observations, if any, from the same satellite should also be considered for exclusion. The correlation coefficient is calculated as follows:

$$\rho_{A,B} = \frac{D_A^T\, Q_{\hat{r}_U}^{-1}\, D_B}{\sqrt{D_A^T\, Q_{\hat{r}_U}^{-1}\, D_A}\ \sqrt{D_B^T\, Q_{\hat{r}_U}^{-1}\, D_B}} \quad (11)$$

$$Q_{\hat{r}_U} = Q_{y_U} - G_U \left(G_U^T\, Q_{y_U}^{-1}\, G_U \right)^{-1} G_U^T \quad (12)$$

where ρ is the correlation coefficient between errors in the observations A, B; D is a zero column vector with ones at $A\ and\ B$ entries only; $Q_{\hat{r}_U}$ is the covariance matrix of the user residuals; and Q_{y_U} is the covariance matrix of the user observations.

Observed satellites										
Satellite PRN	1	2	3	4	5	6	7	8	9	10
Observations	φ1	P2	φ2	C1	φ1	φ2	φ1	C1	C1	φ2
	C1	φ2	C1	φ1	φ2	P2	C1	φ2	φ1	P2
	φ2	C1	φ1	P2	C1	φ1	P2	P2	φ2	φ1
	P2	φ1	P2	φ2	P2	C1	φ2	φ1	P2	C1

Fig. 2 Example of the ranking criterion of the observations based on the ratio of the residuals with the reference station where red-shaded observations are the most likely to be faulty

After an initial exclusion of the suspected observation(s), the position solution is re-estimated and the Chi-squared test is repeated. If the detection test continues to fail, the exclusion of the second most vulnerable observation and its highly correlated observations is attempted with and without re-inserting the previously excluded observations back into the model. These procedures are repeated until the test passes, and the faulty observation is identified. The exclusion of the faulty satellites reported by the network and the exclusion of the highly correlated observations shall significantly minimize the risk of having more than one faulty observation, and the method can be applied for the exclusion of single or two simultaneous observations. The method can be applied in the case of suspecting two simultaneous faulty observations by selecting the two observations with the largest ratio values, whether they are of the same satellite or from two different satellites. The flow chart presented in Fig. 3 shows the overall process.

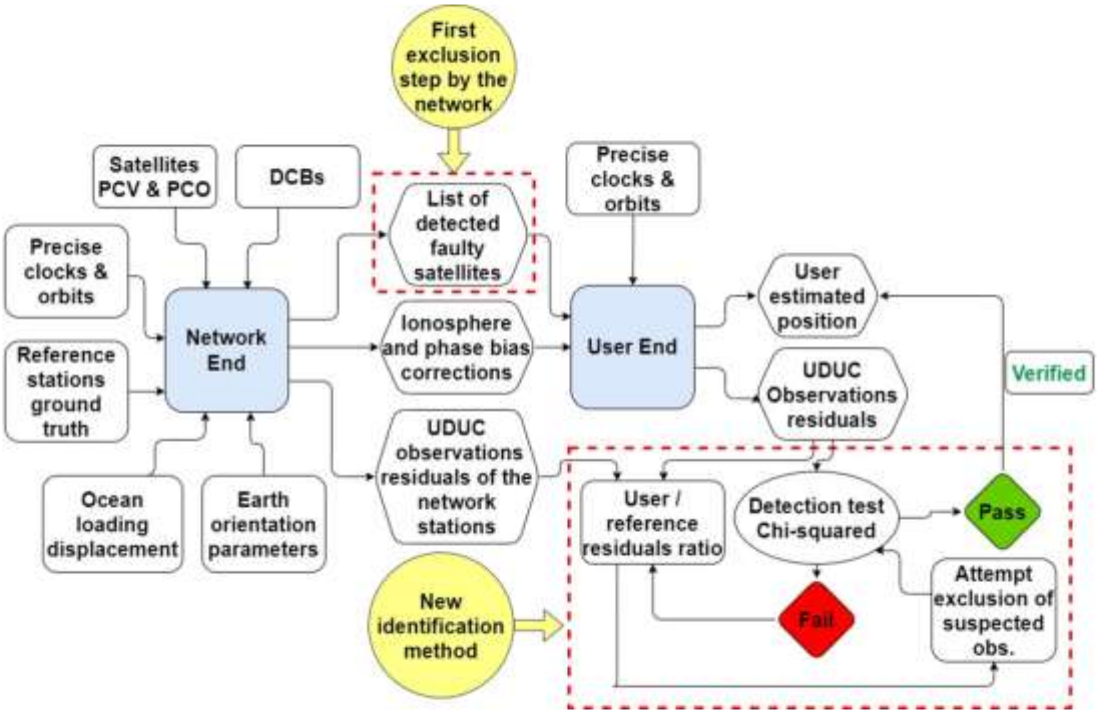

Fig. 3 Flow chart of the proposed method for FDE using PPP-RTK

Table 3 shows the main merits of the proposed method. It makes the identification process much faster due to the statistically-based selection process of a limited number of suspected observations, thereby, avoiding testing of all subsets. In addition, the identification is performed for the individual observations, maintaining the information availability, since the exclusion is performed per individual observations, not per satellites, as the case in the traditional grouping techniques. This is facilitated because the UDUC PPP-RTK offers the ability to calculate the residuals of the individual observations.

Table 3 Main advantages of the proposed method for fault identification

Advantages of PPP-RTK	Risk reduction at the user end	Improving idetification process speed	Better availability maintaining
• Precise and real time poisition method; • Ability to analysis error sources; • Reduces communication risks and the required bandwidth as one-way communications are used; • UDUC processing allows better analysis of individual observations.	• The satellites are checked by the network for any faulty satellites due to satellite faults and a list of the faulty satllites is sent to user(s) to exclude.	• The suspected observations are selected based on the ratio value without the need to generate all possible subsets; • The ratio is calculated based on the real values of the residuals	• Exclusion is attempted on observations not on satellites as it is the case in grouping techniques; • The identificaiton only takes place when the detection test fails unlike SS where it is performed every epoch and for all subsets.

Experimental test cases

The proposed method has been tested in two different situations with different parameters to verify the outcome of the proposed approach. A geodetic receiver that observes GPS legacy frequencies only is used in the first case. In the second example, a low-cost receiver was used to observe two frequencies from multi-constellation GNSS. The latter kind of receivers, due to their low cost, is anticipated to be onboard most cars, etc. The variations also extended to include the testing dates and the number of reference stations of the network, and their distances to the user. Table 4 summarises the experiments' strategy, and Fig. 4 shows a map of the receivers' distribution of the network and the user.

Table 4 Testing strategy and parameters of the two applied test cases

Parameter	**Value**	
	Test case (1)	**Test case (2)**

# of network stations	10 CORS*	7 CORS*
User receiver	Geodetic receiver (Trimble NETR9)	Commercial receiver (u-blox F9P)
Sampling interval/test period	30 sec / 10 hr	1 sec / 12 hr
Date of testing	1 July 2022	13 Oct 2022
GNSS/frequencies	GPS: C1C, L1C, C2L, L2L	GPS: C1C, L1C, C2L, L2L
		Galileo: C1C, L1C, C7Q, L7Q
		BDS: C2I, L2I, C7I, L7I
Distances to the user (km)	3 ~ 32	<1 ~ 20

*** Continuously Operating Reference Stations (CORS)**

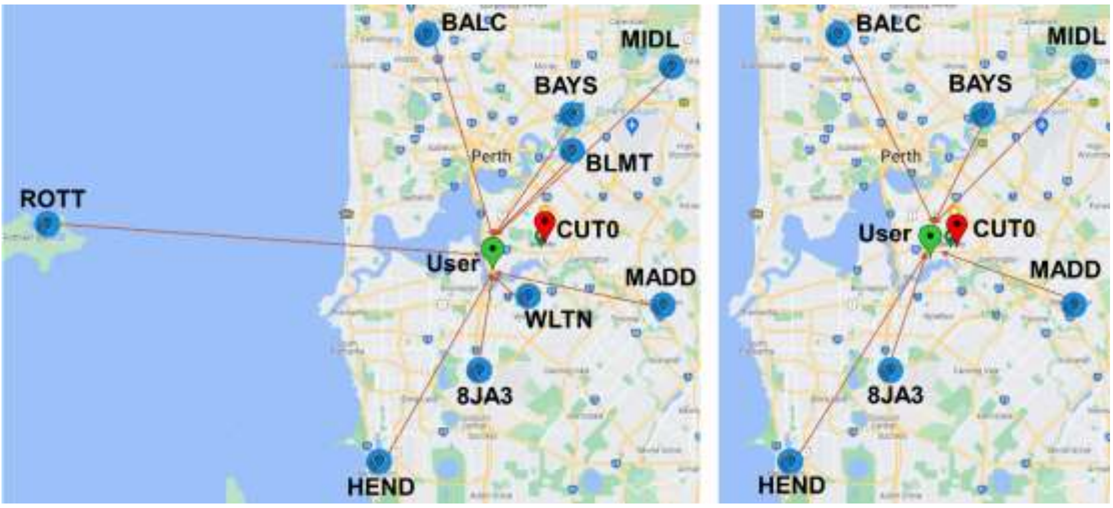

Fig. 4 Layout of the network stations (in blue), including the reference station (in red) and the user receiver (in green) for test case (1) on the left panel and test case (2) on the right panel

Results and discussion

In the two testing examples, the network processes the data and provides the error corrections as well as a list of faulty satellites to the user. The network also provides the user with the residuals of the observations of the reference stations. At the user end, the faulty satellites reported by the network are excluded, and the residuals of remaining individual observations are calculated. Fig. 5 and Fig. 6 show the residuals of the four observations of each satellite in the first test case tracked by the selected reference station (nearest to the user) and the user, respectively, during the test period. The residual behaviour was very much the same in the second test case. From the two figures, it is evident that the overall values of the user residuals are larger than their counterparts at the reference station, in particular for the code observations. This is expected since both the precautions taken in the setup of a reference station, e.g. minimizing the impact of multipath, and the processing that exploits the known position of the reference station shall help in reducing the level of the computed residuals. This is reflected in their RMS values given in Table 5.

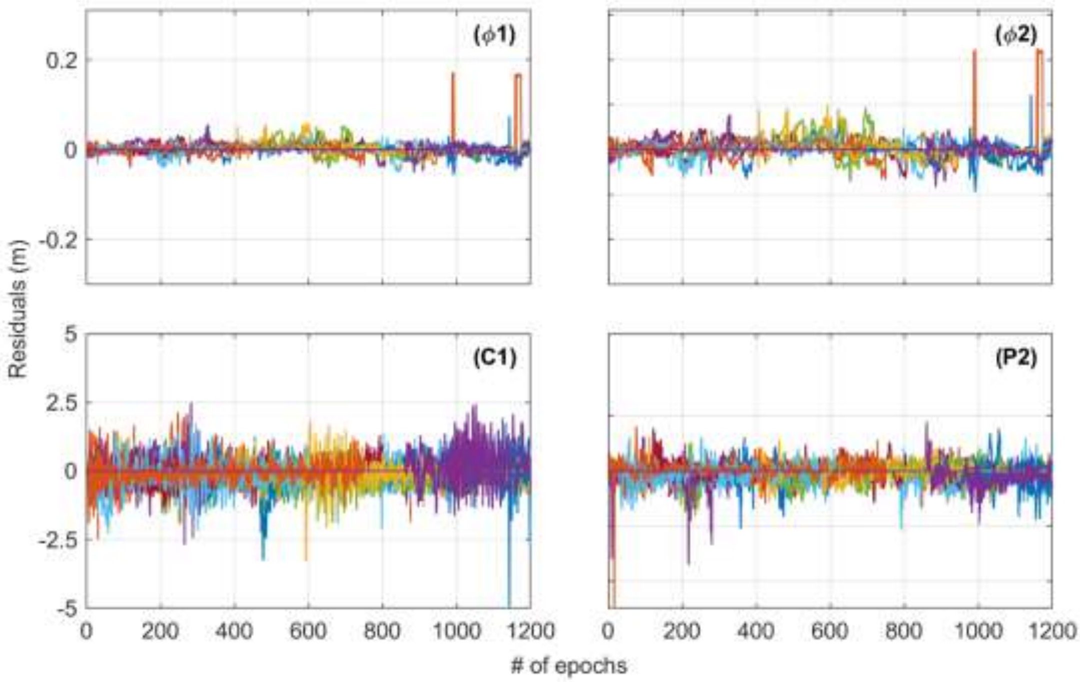

Fig. 5 Residuals of phase (top) and code (bottom) observations of L1 (left) and L2 (right) frequencies for all GPS-tracked satellites by the *reference station* during the first test where different colours represent different satellites

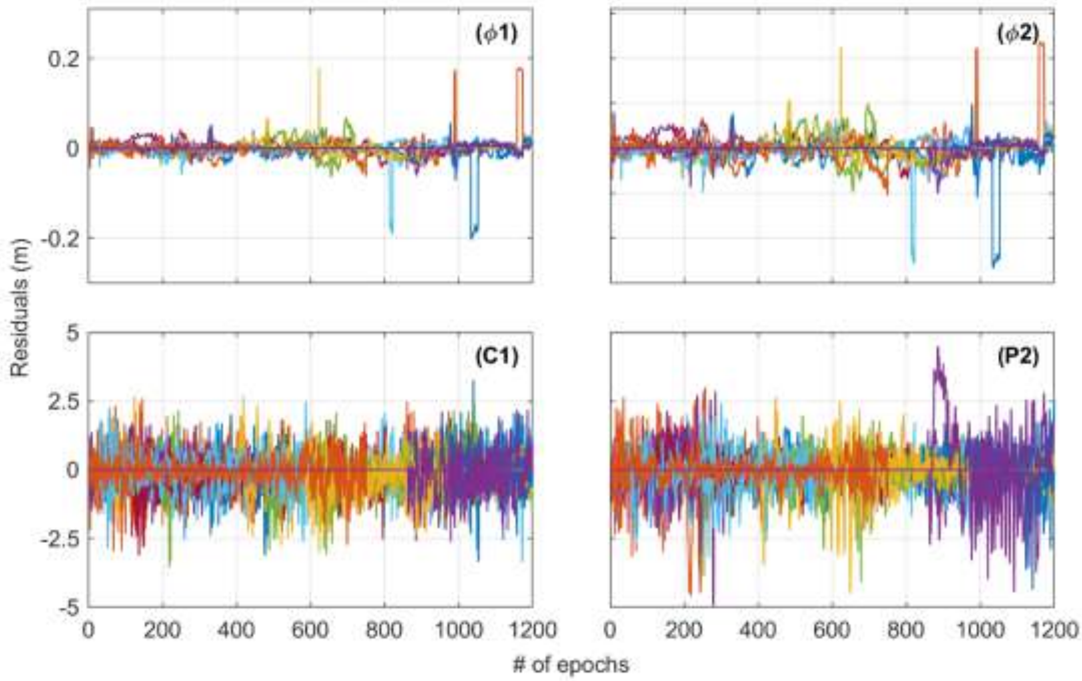

Fig. 6 Residuals of phase (top) and code (bottom) observations of L1 (left) and L2 (right) frequencies for all GPS-tracked satellites by the *user* during the first test where different colours represent different satellites

Table 5 RMS of the observation residuals of each signal type of the tracked GPS satellites for both reference and user stations over 1200 epochs period, the total length of test case (1)

	RMS (m)	
	Reference station	**User receiver**
$\boldsymbol{\varphi_1}$	0.0139	0.018
$\boldsymbol{\varphi_2}$	0.0208	0.0264
$\mathbf{C_1}$	0.4852	0.6756
$\mathbf{P_2}$	0.3498	0.736

For Fig. 7 and Fig. 8, they represent test case (1) where GPS only was used. The four plots in each figure refer to four different and independent epochs (Fig. 7 and Fig. 8 show eight different epochs) in which a fault has been detected in each of them. These epochs are representative examples among many other epochs where faults have been detected. The graph on each plot represents the absolute value of the ratio between the residuals of the observations of the user and their counterparts of the nearest reference station of the network at that epoch. For the four epochs presented in Fig. 7, the identification of the faulty observation was successful after the first iteration as per the criterion described earlier in the methodology section. In brief, in these four epochs, Chi-squared test has detected a fault. To identify which observation is the faulty one in each epoch, the ratio has been calculated and the observation with the largest ratio has been excluded. The position solution was estimated, and the residuals were computed after that exclusion. The detection test (Chi-squared test) was performed on the new residuals, and it passed, meaning that the excluded observation (the one with the largest ratio that is encircled in red in Fig. 7) was the faulty one. The four epochs presented in Fig. 8 follow the same explanation described for Fig. 7. The difference is in these four epochs the detection test did not pass after the exclusion of the observation of the largest ratio. Therefore, the observation of the second largest ratio (i.e., the observation encircled in red in Fig. 8) was excluded and the detection test passed. The eight epochs depicted in Fig. 9 and Fig. 10 are similar to those in Fig. 7 and Fig. 8, respectively, but for test case (2) where multiple constellations were observed using a commercial low-cost receiver.

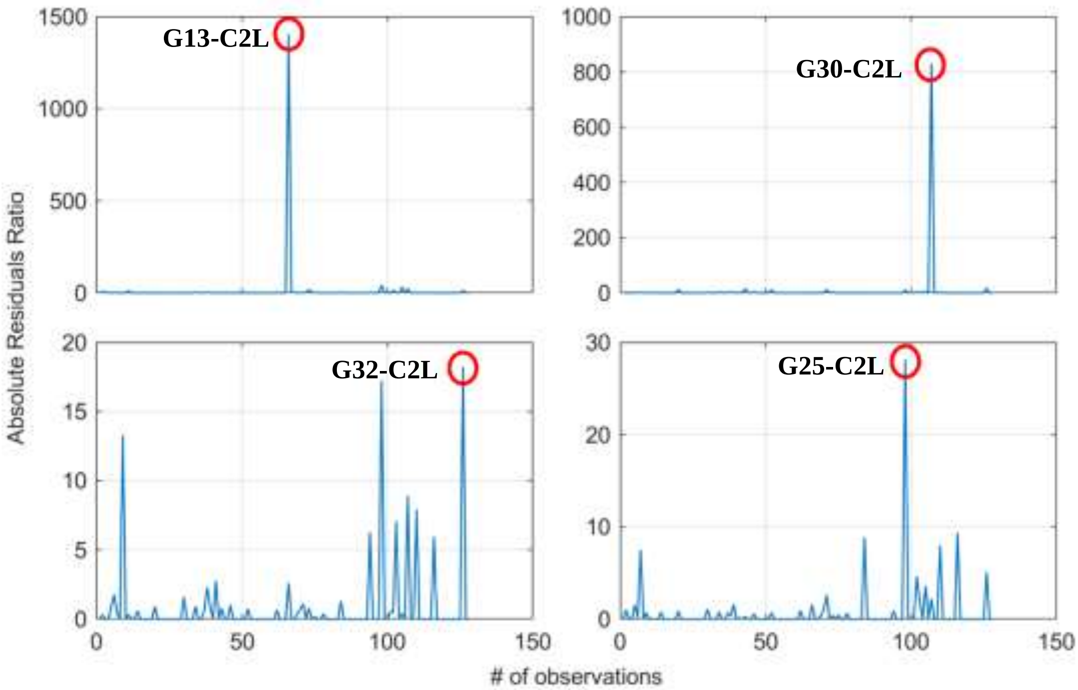

Fig. 7 Absolute ratios between the residuals of the user receiver and reference station observations at four different epochs as representative examples in *the first test* where a fault has been detected among GPS observations only. The encircled observations, with *the largest ratio***,** are the faulty observations. The satellite PRN and the observation type of each faulty observation are mentioned in the text within each respective plot

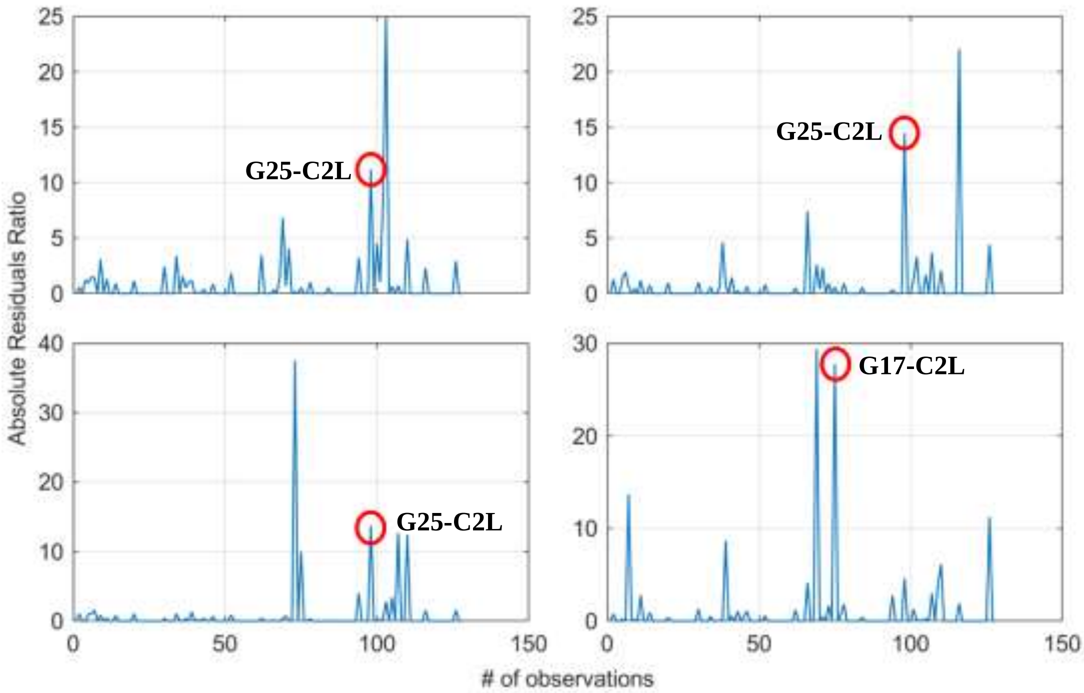

Fig. 8 Absolute ratios between the residuals of the user receiver and reference station observations at four different epochs as representative examples in *the first test* where a fault has been detected among GPS observations only. The encircled observations with *the second largest ratio***,** are the faulty observations. The satellite PRN and the observation type of each faulty observation are mentioned in the text within each respective plot

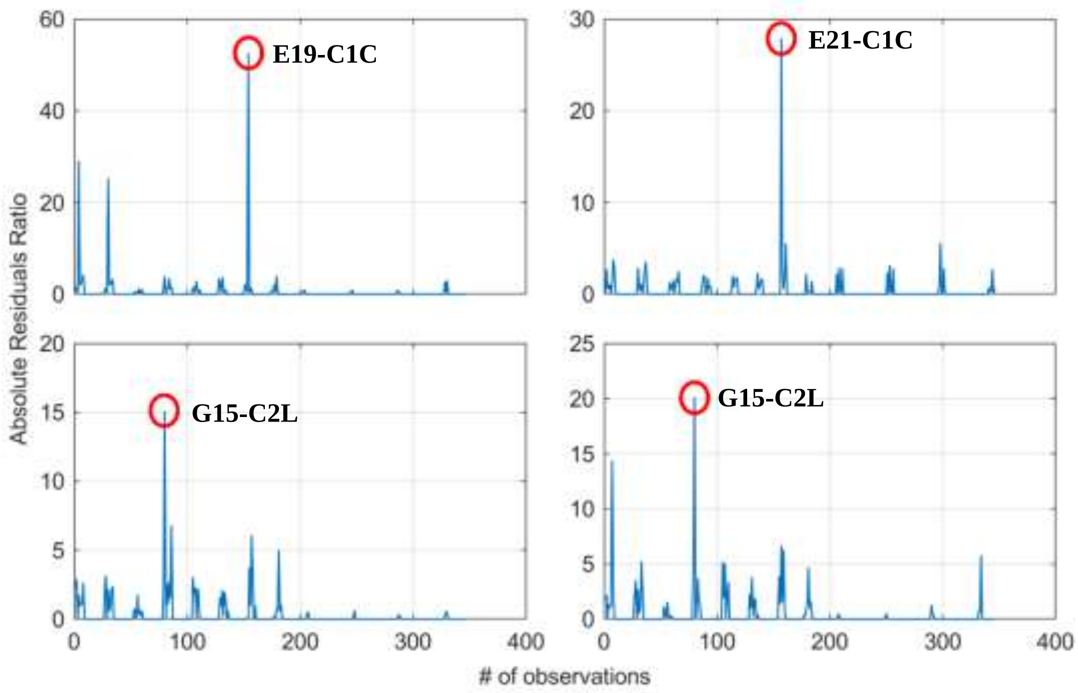

Fig. 9 Absolute ratios between the residuals of the user receiver and reference station observations at four different epochs as representative examples in *the second test* where a fault has been detected among GPS, Galileo and BDS observations. The encircled observations with *the largest ratio***,** are the faulty observations. The satellite PRN and the observation type of each faulty observation are mentioned in the text within each respective plot

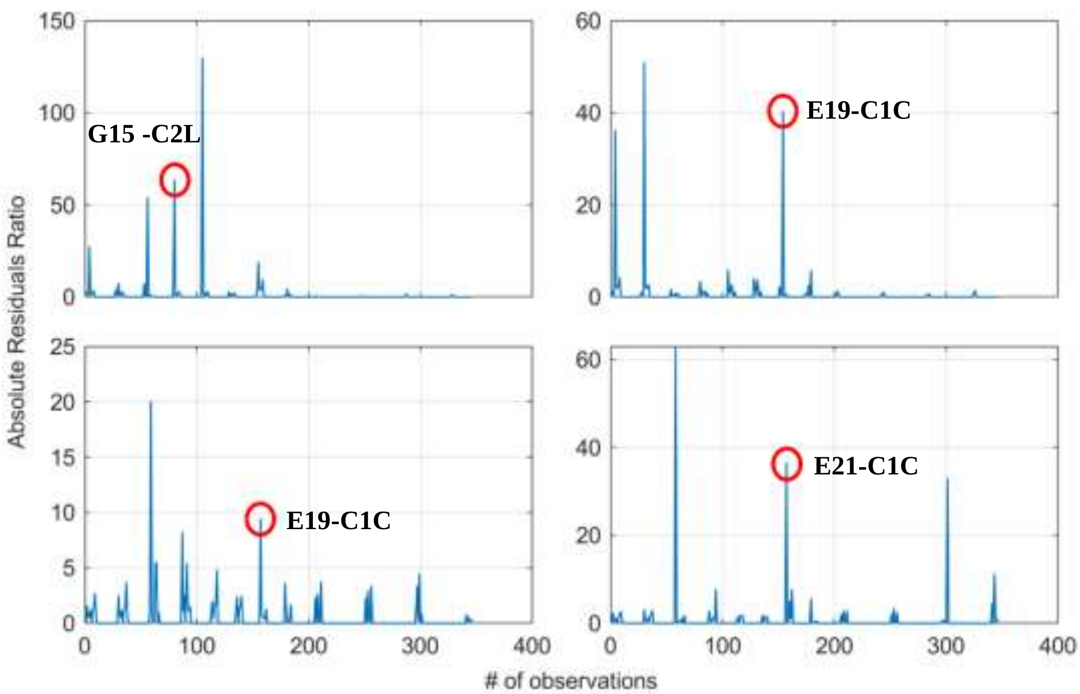

Fig. 10 Absolute ratios between the residuals of the user receiver and reference station observations at four different epochs as representative examples in *the second test* where a fault has been detected among GPS, Galileo and BDS observations. The encircled observations with *the second largest ratio***,** are the faulty observations. The satellite PRN and the observation type of each faulty observation are mentioned in the text within each respective plot

It is noted that the new approach using the user/reference residuals ratio was not *always* able to identify the faulty observation(s) from the first exclusion attempt and more testing cycles were needed. This can be explained as follows: from equation (10), it can be shown that the more the reference station residuals are accurate, the more the ratio will be sensitive to identify a fault from the first attempt. However, in some cases, the best fit of the observations with the final solution performed at the reference station can produce large residuals for some observations due to their specific errors such as multipath, imperfection of the bias model, etc., which are reflected in their residuals. This can cause an increase in the value of some residuals of the reference stations, and as a result, the ratio related to these observations may not be the highest in case their counterparts are the faulty ones at the user end. However, the case, where the largest ratio is unrepresentative of the faulty observation after a few exclusion attempts, was infrequent during the test cases.

Table 6 shows the overall statistics of the two testing cases in terms of the identification potentials of the new approach. 112 and 110 faulty epochs were found in testing examples one and two, respectively. The ratio method identified the faulty observation(s) at the first exclusion attempt in about 16 -19% of the number of faulty epochs in the two test cases increased to 76 – 81%, after six exclusion attempts.

Table 6 Percentage of identification success in the two testing examples

# of exclusion attempts	Test case (1)%	Test case (2)%
First exclusion attempt	19.64	16.36
First and second exclusion attempts	39.29	30
First to third exclusion attempts	50.89	46.36

First to fourth exclusion attempts	60.71	61.81
First to fifth exclusion attempts	71.43	69.1
First to sixth exclusion attempts	81.25	76.36

Table 7 presents a comparison between the proposed identification method and the SS as well as Chi-squared methods in terms of their ability to reduce the computational load as a factor in the number of the required observation subsets in the two tests. The new method saved around 85% and 98% of the computational load of the FDE process in the first test case compared to Chi-squared and SS, respectively, while it saved about 94% and 99.999% in the second test compared to the two methods. The percentage of reducing the computational load is proportional to parameters such as observation period, sampling interval, number of observations at each epoch, and the number of detected epochs with faulty observations. This is because the number of the required subsets for testing increases significantly with the increase of these parameters in the case of using SS and Chi-squared methods. Whereas the number of the generated and tested subsets when using the new ratio method is only dependent on the number of performed iterations needed to identify the faulty observations. This example shows how the new ratio method is effective in significantly reducing the computational load of the FDE process especially when a high sampling rate is required.

Table 7 Comparison between the proposed identification method and the solution separation and Chi-squared identification methods concerning the ability to save in the computational load

Testing example	**Test case (1)**			**Test case (2)**		
Observation period (hr)	10			12		
Sampling interval (sec)	30			1		
Average No. of observations	32			76		
No. of faulty epochs	112			110		
Identification method	SS	Chi-2	Ratio	SS	Chi-2	Ratio

No. of tested subsets in the one-satellite-out fault mode	38400	3584	507	~ 3 million	8360	493
% of computational load saving the FDE process when using the ratio method	~ 98	~ 85		~99.99	~ 94	

Conclusion

Autonomous navigation of vehicles, drones, and others requires real-time precise positioning with efficient integrity monitoring capability. PPP-RTK positioning method can cover a wide area and provide precise corrections with fast solution initialization and has the additional advantage of IM, i.e., providing the residuals for the individual observations by processing the UDUC observations. Current FDE methods represent a major challenge for real-time applications as it encompasses the generation and testing of numerous observation subsets to identify faulty observations when detected, especially when multiple faults take place concurrently. To reduce the computational load, some suggested methods, such as the grouping technique, result in the loss of valuable information from the observations of the removed satellite.

We propose a new approach that can reduce the computational load of the FDE process without affecting the observation availability. We suggest excluding faulty satellites at the network station exploiting the known positions of the stations, and sending this information to users. For errors due to the user environment or due to imperfect error treatment, when a fault has been detected at a certain epoch, the ratio between the observation residuals of the user receiver and the closest reference station, assuming that the latter is fault-free, is to be calculated. The highest ratio can indicate the faulty observation(s) so that their exclusion is attempted to avoid checking for solutions from all possible numerous observation subsets to identify the fault as done by the traditional methods. Moreover, the exclusion will be based on screening the individual observations, not the whole satellite, which maintains the observation availability due to processing UDUC observations.

Two representative tests were performed to demonstrate the performance of the proposed method. The first included a geodetic receiver that tracked GPS observations only, and the second test comprised a low-cost receiver that is most likely to be used in

AVs observing multi-GNSS constellations measurements. In the two tests, the new ratio method provided consistent performance where the faulty observations have been identified from the first exclusion attempt in 16 – 19% of the epochs where a fault has been detected, while it took up to six exclusion attempts to identify the faulty observation in around 76 – 81% of the faulty epochs. When compared to the commonly used FDE methods, such as the SS test and conventional Chi-squared test, it takes only <1% and 15%, respectively, of the time required for detection and identification. This is based on the observing period and interval, the number of the faulty epochs, i.e., 112 and 110, and the average number of observations at each epoch, i.e., 32 and 76, in the two test examples, respectively. The future work plans include testing in a kinematic mode where the receiver is mounted on top of a moving vehicle. Also, it includes involving testing more frequencies, and for more challenging environments.

Acknowledgment

The first author acknowledges the Egyptian Ministry of Higher Education and Scientific Research for providing him with a scholarship to complete his Ph.D. studies at Curtin University, Australia. The School of Earth and Planetary Sciences, Curtin, University for partial support of this study. This research is partially funded by the Australian Research Council Discovery Project (Grant No. DP 190102444) and the National Time Service Center, Chinese Academy of Sciences (CAS) (No. E167SC14). The authors acknowledge HxGN SmartNet for providing access to Australian GNSS network data used in this study.

Data Availability

Data used, generated and analysed in this study will be made available upon reasonable request from the corresponding author.

Authors contributions

Hassan Elsayed: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft; **Ahmed El-Mowafy:** Conceptualization, Methodology, Resources, Writing - Review & Editing, Supervision; **Kan Wang:** Writing - Review & Editing, Supervision.

References

Blanch J, Gunning K, Walter T, De Groot L, Norman L. Reducing computational load in solution separation for Kalman filters and an application to PPP integrity. Proc. ION ITM 2019, Institute of Navigation, Reston, Virginia, USA, January 28-31, 720-729

Blanch J, Walter T, Enge P. Fixed subset selection to reduce advanced RAIM complexity Proc. ION ITM 2018, Institute of Navigation, Reston, Virginia, USA, January 29-1, 88-98

Blanch J, Walter T, Enge P, Lee Y, Pervan B, Rippl M, Spletter A. Advanced RAIM user algorithm description: Integrity support message processing, fault detection, exclusion, and protection level calculation. Proc. ION GNSS 2012, Institute of Navigation, Nashville, Tennessee, USA, September 17-21, 2828-2849

Blanch J et al. A proposal for multi-constellation advanced RAIM for vertical guidance. Proc. ION GNSS 2011, Institute of Navigation, Portland, Oregon, USA, September 20-23, 2665-2680

Chen J, Zhang Y, Yu C, Wang A, Song Z, Zhou J (2022) Models and performance of SBAS and PPP of BDS. Satellite Navigation 3:4

Du Y, Wang J, Rizos C, El-Mowafy A (2021) Vulnerabilities and integrity of precise point positioning for intelligent transport systems: Overview and analysis. Satellite Navigation 2:1-22

El-Mowafy A, Kubo N (2017) Integrity monitoring of vehicle positioning in urban environment using RTK-GNSS, IMU and speedometer. Measurement Science and Technology 28:055102

El-Mowafy A, Wang K (2022) Integrity monitoring for kinematic precise point positioning in open-sky environments with improved computational performance. Measurement Science and Technology 33:085004

El-Mowafy A, Kubo N (2018) Integrity monitoring for positioning of intelligent transport systems using integrated RTK-GNSS, IMU and vehicle odometer. IET Intelligent Transport Systems 12:901-908

Elsayed H, El-Mowafy A, Wang K (2023) Bounding of correlated double-differenced GNSS observation errors using NRTK for precise positioning of autonomous vehicles. Measurement 206:112303

Ge Y, Wang Z, Zhu Y (2017) Reduced ARAIM monitoring subset method based on satellites in different orbital planes. GPS Solutions 21:1443-1456

Gerbeth D, Martini I, Rippl M, Felux M. Satellite selection methodology for horizontal navigation and integrity algorithms. Proc. ION GNSS 2016, Institute of Navigation, Portland, Oregon, USA, September 12-16, 2789-2798

Gunning K, Blanch J, Walter T, de Groot L, Norman L. Design and evaluation of integrity algorithms for PPP in kinematic applications. Proc. ION GNSS 2018, Institute of Navigation, Miami, Florida, USA, September 24-28, 1910-1939

Hassan T, El-Mowafy A, Wang K (2021) A review of system integration and current integrity monitoring methods for positioning in intelligent transport systems. IET Intelligent Transport Systems 15:43-60

Janssen V (2009) A comparison of the VRS and MAC principles for network RTK. Proc. IGNSS 2009, Queensland, Australia, December 1-3, 13pp

Joerger M, Pervan B. Solution separation and Chi-Squared ARAIM for fault detection and exclusion. Proc. IEEE/ION PLANS 2014, Institute of Navigation, Monterey, California, USA, May 5-8, 294-307

Joerger M, Pervan B (2016) Fault detection and exclusion using solution separation and chi-squared ARAIM. IEEE Transactions on Aerospace and electronic systems 52:726-742

Landau H, Vollath U, Chen X. Virtual reference stations versus broadcast solutions in network RTK–advantages and limitations. Proc. GNSS 2003, Graz, Austria, April 22-24, 22-25

Li X, Li X, Li S, Zhou Y, Sun M, Xu Q, Xu Z (2022) Centimeter-accurate vehicle navigation in urban environments with a tightly integrated PPP-RTK/MEMS/vision system. GPS Solutions 26:124

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

Orejas M, Skalicky J. Clustered ARAIM. Proc. ION ITM 2016, Institute of Navigation, Monterey, California, USA, January 25-28, 224-230

Orejas M, Skalicky J, Ziegler U. Implementation and testing of clustered ARAIM in a GPS/Galileo receiver. Proc. ION GNSS 2016, Institute of Navigation, Portland, Oregon, USA, September 12-16, 1360-1367

Sasani S, Asgari J, Amiri-Simkooei A (2016) Improving MEMS-IMU/GPS integrated systems for land vehicle navigation applications. GPS solutions 20:89-100

Takac F, Zelzer O. The relationship between network RTK solutions MAC, VRS, PRS, FKP and i-MAX. Proc. ION GNSS 2008, Institute of Navigation, Savannah, Georgia, USA, September 16-19, 348-355

Teunissen P (1985) Zero order design: generalized inverses, adjustment, the datum problem and S-transformations. In: Grafarend E W, Sansò F (eds) Optimization and design of geodetic networks, Springer, Berlin, Heidelberg, pp 11-55

Vollath U, Landau H, Chen X, Doucet K, Pagels C. Network RTK versus single base RTK-understanding the error characteristics. Proc. ION GPS 2002, Institute of Navigation, Portland, Oregon, USA, September 24-27, 2774-2781

Wabbena G, Schmitz M, Bagge A. PPP-RTK: precise point positioning using state-space representation in RTK networks. Proc. ION GNSS 2005, Institute of Navigation, Long Beach, California, USA, September 13-16, 2584-2594

Walter T, Blanch J, Enge P. Reduced subset analysis for multi-constellation ARAIM. Proc. ION ITM 2014, Institute of Navigation, San Diego, California, USA, January 27-29, 89-98

Walter T, Blanch J, Kropp V (2016) Satellite selection for aviation users of multi-constellation SBAS. InsideGNSS: Red Bank, NJ, USA:50-58

Wang K, El-Mowafy A (2021) Effect of biases in integrity monitoring for RTK positioning. Advances in Space Research 67:4025-4042

Wanninger L (2003) Virtual reference stations (VRS). GPS Solutions 7:143-144

Wörner M, Schuster F, Dölitzscher F, Keller CG, Haueis M, Dietmayer K. Integrity for autonomous driving: A survey. Proc. IEEE/ION PLANS 2016, Institute of Navigation, Savannah, Georgia, USA, April 11-14, 666-671

Zha J, Zhang B, Liu T, Hou P (2021) Ionosphere-weighted undifferenced and uncombined PPP-RTK: theoretical models and experimental results. GPS Solutions 25:1-12

Zhang B, Chen Y, Yuan Y (2019) PPP-RTK based on undifferenced and uncombined observations: theoretical and practical aspects. Journal of Geodesy 93:1011-1024

Zhang W, Wang J, El-Mowafy A, Rizos C (2023) Integrity monitoring scheme for undifferenced and uncombined multi-frequency multi-constellation PPP-RTK. GPS Solutions 27:68

Authors Autobiography

Hassan Elsayed is a Ph.D. candidate at School of Earth and Planetary Sciences, Curtin University, Australia. He obtained his M.Sc. degree in evaluating the GNSS performance in highly dynamic environments. His research interests include GNSS precise positioning, integrity monitoring of real-time applications, errors overbounding and health monitoring.

Ahmed El-Mowafy is a Professor and Director of Graduate Research, School of Earth and Planetary Sciences, Curtin University, Australia. He obtained his Ph.D. from the University of Calgary, Canada, in 1995 and has more than 200 publications in precise positioning and navigation using GNSS, quality control, POD, integrity monitoring and estimation theory.

Kan Wang is a Professor at the National Time Service Center, Chinese Academy of Sciences. She received her Ph.D. in GNSS advanced modeling from ETH Zurich in 2016 and worked at Curtin University until 2021. Her research interests include high-precision GNSS positioning, LEO precise orbit determination, SBAS and integrity monitoring.