

Article

# GNSS Positioning Accuracy Enhancement Based on Robust Statistical MM Estimation Theory for Ground Vehicles in Challenging Environments

Muhammad Adeel Akram, Peilin Liu \*, Yuze Wang  and Jiuchao Qian

Brain-inspired Application Technology Center (BATC), School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; adeelakram03@gmail.com (M.A.A.); wangyuze90@sjtu.edu.cn (Y.W.); jcqian@sjtu.edu.cn (J.Q.)

\* Correspondence: liupeilin@sjtu.edu.cn

Received: 3 April 2018; Accepted: 14 May 2018; Published: 25 May 2018



**Featured Application:** A robust estimation procedure is proposed in this work targeting GNSS performance improvements. Thus this research is exclusively beneficial and applicable to a wide-ranging GNSS applications particularly for on road vehicle navigation and tracking. Despite of GNSS, the proposed scheme is appropriate to make expansions in EKF filter performance as well.

**Abstract:** Global Navigation Satellite System (GNSS) is the most reliable navigation system for location-based applications where accuracy and consistency is an essential requirement. The LSE (least squares estimator) has been used since the start of GNSS for position estimation. However; LSE is affected by outliers and errors in GNSS measurements and results in wrong user position. In this paper; we proposed a novel three-phase estimator for enhancing GNSS positioning accuracy in the presence of outliers and errors; relying upon the robust MM estimation theory. In the first phase; a subsampling process is proposed on available observations. IRWLS (iterative reweighted LS) is applied to all subsamples up to a predefined number of observations to obtain a positioning estimate and a scale factor. Secondly; IRWLS is applied up to the convergence point on a set of selected subsamples. The third phase involves the selection of optimum positioning solution having minimum scale factor. An outlier detection and exclusion process is applied on a probabilistic set of outlying observations to maintain the integrity and reliability of the position. Multiple simulated and real scenarios are tested. Results show high accuracy and reliability of the proposed algorithm in challenging environments.

**Keywords:** multi-GNSS; GNSS navigation; LSE; IRWLS; urban canyon; outliers; MM estimation

## 1. Introduction

Location-based applications are progressively getting more traction in our daily life, including applications such as transportation, communication, precise farming, and life-rescue operations. Positioning accuracy and consistency are the key requirements for the applications such as autonomous land vehicle driving, unmanned aerial vehicle (UAVs) and in life-critical applications. Global Navigation Satellite System (GNSS) is one of the most trustworthy positioning systems. The establishment of new GNSS constellations like the BeiDou and the Galileo has greatly increased the number of visible satellites around the global as well as new frequency signals and satellites are also added in existing GNSS constellations. These developments make it possible to use signals from multiple constellations to enhance the accuracy of user position. However, a key challenge for

multiconstellation GNSS positioning is to identify the trustable satellite measurements from all visible satellites for getting true position. The quality of the estimated position depends on the ability of the receiver on getting position in the presence of outlying observations and on the receiver environments. Multipath and NLoS (non-line of sight) signals can cause errors in the satellite observations for the receiver moving in challenging environments such as urban canyons, under the elevated highway or under the dense trees. Hence it is a problematic task to achieve the positioning accuracy and reliability in such environments and it is required to adopt the positioning estimation techniques to prevent the effect of possible wrong or outlier satellite observation on the user position.

A number of algorithms and methods have been proposed to improve the GNSS receiver performance in terms of the positioning accuracy. The geometric dilution of precision—(GDOP) based best satellite selection is one of the proposed methods [1–5]. The fundamental idea is to select a subset of observations having a minimum GDOP factor, though the rejection of a good observation can also lead to the minimized GDOP. Receiver autonomous integrity monitoring (RAIM) is another approach to reject outlier observations using pseudorange measurement redundancy to maintain the positioning accuracy by isolating the contaminated measurements from the good ones [6–8], and is primarily designed for aviation applications [9,10]. However, for user position estimation in urban environments, visible and available satellites are rapidly changing and it is difficult to use the RAIM for satellite/pseudorange selections. Besides the satellite/observation selection techniques, several researchers also focused to improve the estimation algorithms to enhance the positioning accuracy. Least-square estimators (LSE) has been used as the fundamental algorithm for GNSS positioning since the start [11]. LSE estimates the position of user by making a relationship between the unknown parameters and satellite to receiver geometry. However, it is assumed during LSE-based estimation, that the variance-covariance matrix of the observations, also called the dispersion matrix, is an identity matrix i.e., all the observations have same accuracy and are uncorrelated. On the other hand, real time GNSS observations have different level of accuracies depending upon the satellite elevation angle, signal strength and other radio interferences. Multiple algorithms and methods have been proposed to improve the GNSS positioning accuracy. WLSE (weighted least-squares estimation) schemes are proposed for calculating and assigning weights to the GNSS observations. In [12], authors proposed a WLSE (weighted least-squares estimator) where signal to noise ratio and observation variance are the weight model parameters. Tabatabaei [13] proposed a fuzzy weighted least squares method to enhance the GPS and the GLONASS (RUSSIAN GNSS) integrated navigation solution and 12 rules were defined based on the GDOP and the elevation angles of satellites for assigning weights to the GNSS observations. In [14], authors studied, tested and compared multiple weighing schemes for GNSS positioning relying upon the elevation angle and signal to noise ratio. In [15], two GNSS observation weighing models, sigma-elevation and CN0-elevation, are studied and compared by Sarab Tay. For the WLSE algorithms, satellite elevation angle and SNR (signal to noise ratio) are the primary focus of the researchers to compute the observation weights. However, relying on the elevation angle is not always useful because it is possible for the signal, from high elevation satellite, to be disrupted from the multipath and radio interferences especially in urban canyons. The concept of non-linear statistics (or robust statistics) has been also studied in literature by several researchers to improve the GNSS positioning accuracy. In [16], neural networks are used to improve the performance of least absolute and least-squares estimators for robust GNSS positioning. However, the primary disadvantage is the computational cost in addition to the outfitting problem. In [17], the concept of robust M-estimator is used to improve the performance for the GNSS positioning by calculating and adding the weight matrix during the innovation step. In [18] multiple robust statistical estimation techniques have been discussed including M-estimators, L1 estimators and genetic algorithm. The proposed estimation process in [18] determines a set of all possible solutions by using the minimum required number of satellites (e.g., in case of GPS, four satellites are required for positioning) using LSE. The median of remaining estimates is considered as the user position after removing possible wrong estimates and a genetic algorithm is used to minimize the estimation error. The fundamental problem for

this scheme is the number of possible solutions. The addition of new constellations increases the number of available satellites for positioning in a single epoch and the number of solutions as well with the increase in computational cost. Secondly, the positioning process starts with an unknown random initial estimate such as a least-squares estimate and the iterative reweighting scheme can ignore perfectly good measurements resulting to a wrong convergence point with a bad initial estimate. Some researchers focused on the development of techniques to minimize the effect of noise parameters, such as ionospheric, multipath etc. In [19], authors introduced the concept of the adaptive carrier smoothing to attain an optimal carrier smoothing time to minimize the cost functions related to white noise multipath and correlated multipath errors. In [20], authors studied the application of DGNSS (differential GNSS) in smartphones for enhancement of the GNSS positioning accuracy in challenged environments. However the primary drawback for DGNSS-aided navigation is that a base station must be required for transmission of useful information and without this information host system is not able to maintain the required accuracy.

In this paper, we present a novel efficient estimation process for the GNSS positioning using signal data from multiple constellations. The proposed idea is based on the nonlinear statistical estimation theory. The proposed estimator is derived from the robust MM class estimators, proposed and named by Yohai [21]. The MM estimator is a class of robust statistics for more efficient statistical data analysis with a high breakdown point. Breakdown point is the proportion measure of the outliers that can be treated efficiently without spoiling the results, and for the MM estimator, the breakdown point is 50%. For a reliable initial estimate, a weighing scheme is provided using prior knowledge of signal-to-noise ratio and satellite geometry of all visible satellites. A set of available pseudorange observations is divided into small subsets with adaptive subset size. A three-step IRWLS estimation process is applied to the individual subsample to minimize the effect of observation errors on the positioning accuracy. An outlier identification and exclusion process is proposed on the set of probabilistic outlying observations to maintain the reliability and the integrity of the positioning solution. Moreover, this paper is also presented as a concise tutorial for a better understanding of the proposed estimation algorithm in both theoretical and mathematical aspects.

The rest of the paper is organized as follows. Section 2 recalls the idea of the robust estimation theory and proposed robust estimators are reviewed. Section 3 introduces the theory and the application of the LSE model for GNSS positioning and details of the classic outlier detection exclusion process. In Section 4, mathematical and theoretical aspects of the robust MM estimation theory are presented. Details of the custom designed software receiver for performance analysis is presented in Section 5. The proposed estimation algorithm for multiconstellation positioning is explained in Section 6. Section 7 deals with the outlier detection and exclusion process in detail. In Section 8, a detailed discussion is provided related to key modifications in the original statistical estimation procedure and a very similar GNSS positioning scheme [18], based on the robust statistics, is analyzed and compared with our proposed estimation scheme. Results of the proposed estimation algorithm are analyzed and discussed against LSE and WLSE algorithms for several field and simulation tests in multiple scenarios and environments in Section 9. A conclusion is drawn in the last section with future enhancements.

## 2. Review of Robust Statistics

Robust estimation procedures have been developed and studied in statistical data analysis to replace the LSE for unknown estimation. Robust estimation algorithms perform statistical analysis of test data to detect and exclude the outlier data measurements for more accurate estimation. In normal Gaussian distribution conditions, LSE has several advantages such as simplicity of calculation. In addition, the properties of the stochastic and functional models do not change from start to end.

However, multiple classes of robust estimators have been introduced for data sets having multiple outlying observations and with the nonlinear behavior. A robust estimation algorithm is defined on three features: efficiency, stability and breakdown point. Efficiency is the performance of an

estimator for an assumed model. Stability refers to the size of impairment in the performance for small deviations from the model assumptions. A robust estimation process must be stable in the sense that small deviations from the model assumptions should not impair the performance. The measure of the maximal fraction of outlier observations in any data for a robust estimator, without spoiling the final estimate, is defined as the breakdown point. A robust procedure should have an optimal or near optimal efficiency at the assumed model. It should be stable in case of small deviations from the assumed model and avoid catastrophe in case of large deviations from the model. The Least Median Square Estimator (LMSE), proposed by Hampel [22] and further improved by Rousseeuw [23], is the first estimator that achieved the maximum asymptotic break down point of 1/2. Conventional LSE is based on minimizing the sum of the squared residuals. In the LMSE, the sum is replaced by the median of the squared residuals. However, the relative efficiency of this estimator is 0 with respect to LSE because of its low rate of convergence [24]. In [25], Rousseeuw and Yohai proposed the first high breakdown robust regression estimator—named S-estimator—that can achieve both high breakdown point and convergence rate. However, Hössjer proved that the regression S-estimator has maximum asymptotic efficiency 0.329 at the normal distribution [26]. Under the condition of normality, Rousseeuw introduced the concept of weighted LSE (WLSE) by designing a procedure to minimize the weight of outlying observations to reduce the effect of the outliers [27]. Later on, this concept became very popular in navigation research where weights are assigned based on the geometry (elevation and azimuth) and quality of signals [28,29]. Yohai proposed a robust estimator that can achieve nearly optimal efficiency as well as maximum breakdown point and called as MM-estimator [21]. Gervini and Yohai introduced a robust and efficient reweighted least squares estimator (REWLSE), which combines a high breakdown point with the full asymptotic efficiency for normal errors [30].

### 3. Least Squares Estimation Model for GNSS Positioning and Outlier Correction

LSE has been used for the GNSS positioning since the start and still is the key part of many GNSS positioning algorithms such as WLSE, IRLS, etc. A standalone LSE algorithm is not a trustworthy estimator especially in GNSS-challenging environments and this is proved by experimental results in Section 9. In this section, we provide a theoretical understanding of the LSE for GNSS positioning because all positioning algorithms follow the same or a very similar approach.

#### 3.1. Least Squares Position Estimation

In the classical GNSS-based position estimation process, LSE recursively calculates the offset of the receiver’s position relative to the linearization point by using measured pseudorange errors. Let  $\Delta\rho$  denotes the pseudorange error by taking the difference between the measured pseudorange  $\tilde{\rho}$  and the predicted pseudorange  $\hat{\rho}$ :

$$\Delta\rho = \tilde{\rho} - \hat{\rho} \tag{1}$$

$\hat{\rho}$  is the geometric distance between the estimated receiver position and the satellite position:

$$\hat{\rho}_l = \sqrt{(x_l^s - x^u)^2 + (y_l^s - y^u)^2 + (z_l^s - z^u)^2} + cd t_{clk} \tag{2}$$

$[x_l^s, y_l^s, z_l^s]$  are the position coordinates of the  $l$ -th satellite,  $[x^u, y^u, z^u]$  are the estimated receiver’s coordinates and  $dt_{clk}$  is the receiver clock correction.  $c$  is the speed of light. The relation between  $\Delta\rho$  and  $\Delta X^u$  (difference between the predicted  $\hat{X}^u$  and the true  $\tilde{X}^u$  receiver positions) is expressed as,

$$\Delta\rho = \mathbf{H}\Delta X^u + \zeta \tag{3}$$

$\mathbf{H}$  is the geometry matrix and  $\zeta$  is the measurement noise vector. The noise might include uncompensated atmospheric delay, multipath error, satellite clock and ephemeris errors, and the Gaussian white noise. The receiver position is recursively computed,

$$\Delta X^u = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \Delta \rho \tag{4}$$

$$\hat{X}_i^u = \hat{X}_{i-1}^u + \Delta X_i^u \tag{5}$$

$\hat{X}_i^u$  is the estimated position for  $i$ -th,  $\hat{X}_{i-1}^u$  is the estimated position in last iteration and  $\Delta X_i^u$  is the new estimated position error.

### 3.2. Receiver Autonomous Integrity Monitoring Model for Outlier Detection and Removal

RAIM is a self-monitoring process in a GNSS receiver to provide protection against the measurement faults. The classic RAIM scheme uses observation residual vector from the LSE for outlier prediction and tracking.

$$\mathbf{r}_\rho \cong \Delta \rho - \mathbf{H} \Delta X^u \tag{6}$$

$\mathbf{r}_\rho$  is the observation error residual vector. Replacing  $\Delta X^u$  from Equation (4) into Equation (6),

$$\mathbf{r}_\rho \cong (I - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T) \Delta \rho \tag{7}$$

The test statistic for fault isolation is defined as the sum of the squared residual errors,

$$v = \mathbf{r}_\rho^T \mathbf{r}_\rho \tag{8}$$

Under a fault-free positioning condition, the corresponding test statistic  $v$  follows a chi-square ( $\chi^2$ ) distribution with  $\bar{n} - \bar{v}$  degree of freedom where  $\bar{n}$  is the number of measurements and  $\bar{v}$  is the number of unknowns, and should be less than a predefined threshold  $T_{th}$ . The system false alarm probability  $P_{FA}$  and misdetection probability  $P_{md}$  are used to calculate the  $T_{th}$ .

$$P_{FA} = P\{v > T_{th} | \text{no fault}\} \tag{9}$$

$$P_{md} = P\{v < T_{th} | \text{fault in sat}\} \tag{10}$$

Mathematically  $P_{FA}$  and  $P_{md}$  are expressed as [6],

$$P_{FA} = \int_{thr}^{\infty} \left[ \frac{v^{(\frac{\bar{n}-\bar{v}}{2}-1)} e^{-\frac{v}{2\sigma^2}}}{2^{\frac{\bar{n}-\bar{v}}{2}} \Gamma(\frac{\bar{n}-\bar{v}}{2}) \sigma^{\bar{n}-\bar{v}}} \right] dv \tag{11}$$

$$P_{md} = \int_0^{thr} \left[ \frac{1}{2\sigma^2} e^{-\frac{s^2+v}{2\sigma^2}} \left(\frac{v}{s^2}\right)^{(\frac{\bar{n}-\bar{v}}{4}-\frac{1}{2})} I_{\frac{\bar{n}-\bar{v}-2}{2}}\left(\frac{s}{\sigma^2} \sqrt{v}\right) \right] dv \tag{12}$$

$\Gamma(\alpha)$  and  $I_a(t)$  are the gamma and the first order Bessel function respectively;  $\sigma^2$  is the variance of measurement error and  $s = \sqrt{E[\mathbf{r}_\rho^T] E[\mathbf{r}_\rho]}$ . If the test statistic exceeds the threshold, the presence of fault is confirmed. Referring to [7,8], the standardized outlier test is defined as,

$$q_l = \frac{|e_l^T \mathbf{r}_\rho|}{\sqrt{e_l^T Q_v e_l}} \tag{13}$$

$e_l = [0, \dots, 1, \dots, 0]^T$ , the  $l$ -th element equal to 1 while others are zeros.  $Q_v$  is the cofactor matrix of the estimated residuals and is defined as,

$$Q_v = (I - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T) \tag{14}$$

A second test is performed to isolate the faulty measurements.

$$|q_j| > |q_l| \forall l \text{ and } |q_j| > N_{\frac{\alpha_0}{2}}(0, 1) \tag{15}$$

$\alpha_0$  is the probability of false alarm.

#### 4. MM Estimation Theory for Positioning

Conventional statistical estimation schemes are based upon some assumptions about the observations. If the real observations deviate from the assumptions, estimation might deviate from the truth, such as LSE. The robust estimator tries to dynamically search a best fitting criterion to the real observations to reduce the effect of unusual data. Generally, a robust estimator minimizes the cost function  $J(\mathbf{r})$ ,

$$J(\mathbf{r}) = \sum_{l=1}^n \mu(r_l) = \sum_{l=1}^n \mu(\rho_l - \mathbf{H}_l^T X) \tag{16}$$

$\rho_l$  is the  $l$ -th observation,  $X$  is the estimated unknown (here we consider the unknown term as generic),  $\mathbf{H}_l$  is the transformation vector for the  $l$ -th observation, and  $r_l = \rho_l - \mathbf{H}_l^T X$  is the error residual for the  $l$ -th observation.  $\mu(\mathbf{r})$  is the objective function that gives the contribution of each residual to the cost function  $J(\mathbf{r})$ . For the least squares estimator  $\mu(\mathbf{r}) = \mathbf{r}^2$ , while in the robust estimation process, a scaled observation residual is minimized. The scaled estimation process minimizes the weight of out of the bound observations dynamically,

$$\min \sum_{l=1}^n \mu\left(\frac{r_l}{s}\right) = \min \sum_{l=1}^n \mu\left(\frac{\rho_l - \mathbf{H}_l^T X}{s}\right) \tag{17}$$

$s$  is the scale factor.  $\tilde{\mathbf{r}}_s$  is the scaled residual  $\mathbf{r}/s$ . Let if  $\psi$  be the derivative of the objective function for the unknowns  $X$ ,  $\psi = \mu'|_X$ . Differentiating the objective function in Equation (17) and setting the derivative to 0,

$$\begin{aligned} -\frac{1}{s} \sum_{l=1}^n \psi\left(\frac{\rho_l - \mathbf{H}_l^T X}{s}\right) \mathbf{H}_l^T &= 0 \\ \text{equivalent to } \sum_{l=1}^n \psi\left(\frac{\rho_l - \mathbf{H}_l^T X}{s}\right) \mathbf{H}_l^T &= 0 \end{aligned} \tag{18}$$

By introducing a weighting function  $w(\tilde{r}_{s,l})$  to consider the effects of each scaled residual  $\tilde{r}_{s,l}$ , Equation (18) can be modified as,

$$\sum_{l=1}^n w(\tilde{r}_{s,l}) \psi(\tilde{r}_{s,l}) \mathbf{H}_l^T = 0 \tag{19}$$

Solution of Equation (19) is a weighted least-squares problem. However, the weights depend upon the scaled residuals, the scaled residuals depend upon the estimated coefficients, and the estimated coefficients depend upon the weights. Hence, it is an iterative process, iteratively reweighted least squares (IRWLS) [31,32]. Three most popular objective functions and their corresponding weighting functions are shown in Table 1.

**Table 1.** Most popular objective and weighting functions.

Name	Objective Function	Weight Function
LSE	$\mu(\vec{r}) = \vec{r}^2$	$w(\vec{r}) = 1$
Huber	$\mu(\vec{r}) = \begin{cases} \frac{1}{2} \vec{r}^2, &  \vec{r}  \leq \alpha \\ \alpha  \vec{r}  - \frac{1}{2} \alpha^2, &  \vec{r}  > \alpha \end{cases}$	$w(\vec{r}) = \begin{cases} 1, &  \vec{r}  \leq \alpha \\ \alpha /  \vec{r} , &  \vec{r}  > \alpha \end{cases}$
Bi-square	$\mu(\vec{r}) = \begin{cases} \frac{\vec{r}^2}{2} - \frac{\vec{r}^4}{2\alpha^2} + \frac{\vec{r}^6}{6\alpha^4}, &  \vec{r}  \leq \alpha \\ \frac{\alpha^2}{6}, &  \vec{r}  > \alpha \end{cases}$	$w(\vec{r}) = \begin{cases} \left[1 - \left(\frac{\vec{r}}{\alpha}\right)^2\right]^2, &  \vec{r}  \leq \alpha \\ 0, &  \vec{r}  > \alpha \end{cases}$

$\vec{r}$  is the observation residual and it can be either true or scaled. The  $\alpha$ , for Huber and Bi-square, is a tuning constant with optimal value to attain more resistance against outliers (high breakdown point) as well as the convergence efficiency.  $\alpha$  is a tradeoff parameter in terms of robustness and efficiency. The table of  $\alpha$  is provided in Appendix A to represent the relation between efficiency, robustness and tuning constant.  $\alpha$  is selected such that to balance the performance of high breakdown point and efficiency. An outline of robust estimation procedure is defined in the following steps:

1. Find an initial estimate,  $X_0$  with IRWLS (Appendix B) or subsampling method (Appendix C). In IRWLS (simplest subsampling) procedure, all observations are used in  $X_0$  estimation. However, in subsampling procedure, subsamples (subsets) of observations are selected each with size of  $n_{sub}$ , number of observations in one subset. For each subsample, a solution is computed using IRWLS. Best candidate solution is selected as  $X_0$  and the initial residuals are estimated,  $r_l(X_0) = \rho_l - \mathbf{H}_l^T X_0$  for each observation. Then we get a residual vector  $\mathbf{r}^0 = [r_1(X_0) \dots r_n(X_0)]$
2. Multiple scale estimate functions are defined, however the following median function is employed to compute the initial scale value  $s_0$  [33],

$$s_0 = 1.4826 \times median(\mathbf{r}^0) \tag{20}$$

3. Compute initial scaled residual vector  $\check{\mathbf{r}}_s^0 = \mathbf{r}^0/s_0$ . Use the weighting function  $w(\check{\mathbf{r}}_s)$  of one method, Huber or Bi-square, shown in Table 1 to compute corresponding weights  $w_l$  for each observation.
4. Solve the estimated unknown  $X_i$  at the  $i$ -th iteration ( $i = 1, 2, 3, \dots$ ) by using the weighted least-squared method,

$$X_i = (\mathbf{H}^T \Sigma_i \mathbf{H})^{-1} \Sigma_i \mathbf{H}^T \rho \tag{21}$$

$\mathbf{H}$  is the transform matrix composed of  $\mathbf{H}_l$ , and  $\Sigma$  is the weighting matrix with the weights  $w_l (l = 1, 2, \dots, n)$  as diagonal elements.

5. Update residuals  $r_l(X_i) = \rho_l - \mathbf{H}_l^T X_i$  with the new estimate  $X_i$ , calculated in last step. Then compute scaled residual  $\check{r}_{s,l} = r_l(X_i)/s_{i-1}$ . Note that we still use last-step's scale factor to get new scaled residual because updated scale factor can only be obtained later.
6. Use the objective function  $\mu(\check{\mathbf{r}}_s)$  to compute transformed scaled residual  $\check{r}_{s,l}^H = \mu(\check{r}_{s,l})$ , and get a transformed scaled residual vector  $\check{\mathbf{r}}_i^H = [\check{r}_{s,1}^H(X_i) \dots \check{r}_{s,n}^H(X_i)]$ .
7. Update scale factor  $s_i$  with Equation (20) but with the transformed scaled residual vector  $\check{\mathbf{r}}_i^H$ . After that go back to step 4 to start a new iteration.

The first three steps are only used for initial estimate, and the last four steps are iteratively used for the later estimates.

### 5. GNSS Receiver Positioning System Architecture

The complete custom designed GNSS software receiver architecture is shown in Figure 1. The first step comprises to receive original carrier frequency GNSS signals using a wideband antenna and

signals are down-converted to IF (intermediate frequency) data for further processing in software receiver. Receiver signal tracking module processes the IF data and performs two tasks: first the satellites, originating the signal data, are acquired/identified, secondly the satellites are tracked using the stream of signals. The next step involves the transformation of signal tracking parameters to the satellite observations and navigation information. This information is used by estimation algorithm to calculate the receiver location information i.e., position, velocity, time etc.

In Table 2, specifications of the custom designed software receiver are presented. For performance comparison, LSE, WLSE and the proposed estimator are implemented under same receiver specifications. Non-coherent scalar architecture is used for tracking satellites with the early-minus-late discriminator for code phase and cross discriminator for Doppler frequency.  $C/N_0$  and other observation parameters are averaged over 1 second. To reduce the dynamic stress errors, 3rd order phase lock loop (PLL) assisted by frequency lock loop (FLL) is used. First order DLL (delay lock loop) is used for code phase tracking. For  $n + m$  visible satellites from BeiDou and GPS constellations, tracking module tracks the satellite signals using Doppler frequency and code phase ( $f, \phi$ ). Doppler frequency and Code phase information is used to estimate the satellite to receiver pseudorange rate ( $\dot{\rho}$ ) and the satellite pseudorange observations ( $\rho$ ) respectively. This information along other satellite information is transferred to position estimation module. Estimation process computes the receiver position in addition to maintain the integrity and reliability of calculated position.

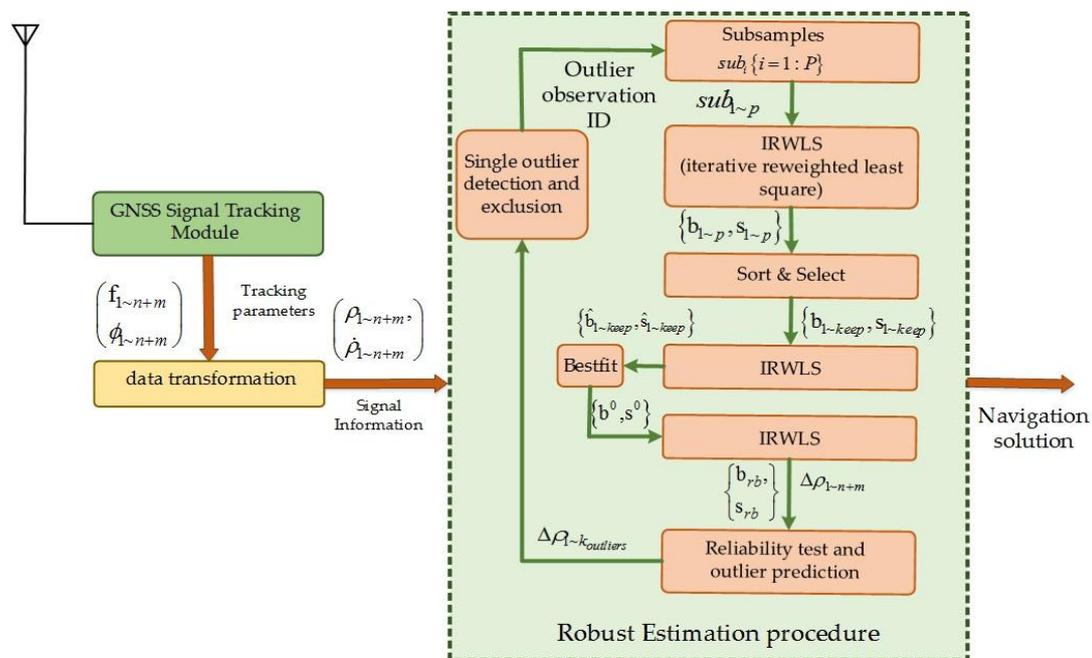


Figure 1. Implementation flowchart of proposed estimation scheme.

Table 2. Software Receiver Specifications.

BeiDou Discriminator Output (Integration Time)	GEO	2 ms
	MEO	20 ms
GPS Discriminator Output (Integration Time)	MEO	20 ms
DLL bandwidth		1 Hz
PLL bandwidth		10 Hz
Position update interval		1 s
CN0 update interval		1 s

In addition to the estimation process, an outlier identification and exclusion procedure is also introduced for verification of outlier free positioning. A reliability test is used to check whether the outlier observations is present or not. If the presence of an outlier observation is confirmed, probabilistic outlier observations are approximated. Selected observations are analyzed for the outlier selection and the corresponding observation is removed.

### 6. Proposed Multiconstellation Positioning Algorithm

The proposed positioning algorithm calculates the navigation information i.e., position, velocity etc. of the object where antenna is fixed. In this paper, BeiDou and GPS satellite observations are used in combination as a multiconstellation GNSS system. Let  $n$  be the number of observations from the first constellation and  $m$  be the number of observations from the second constellation (first can either be any GPS or BeiDou and other is the second). The first step involves the selection of subsets. In classic subsampling based estimation schemes, for a statistical data analysis, a small number of subsamples are drawn in random to process from all possible combinations and each subsample contains a minimum required number of observations. Subsamples are drawn in random because of large number of data samples and is not possible to process all combinations. In contrast, GNSS observations are finite and by using a random technique, it is possible that a good observation is thrown away. Hence it is more appropriate to consider all observations. If there are  $\bar{N}$  number of combinations from all  $n + m$  observations with  $n_{sub}$  observations in one combination, there will be  $\bar{N}$  possible estimates  $\bar{G} = \{\hat{X}_1, \hat{X}_2, \dots, \hat{X}_{\bar{N}}\}$ . In this process,  $n_{sub}$  is an essential parameter that needs to be chosen carefully because the position estimation time is directly related to number of subsamples. For example, if there are 20 satellite observations are available from two constellations and 5 unknowns (three positions, two clock corrections) are required to be estimated, minimum 5 observations are required for estimation of unknowns and there will be 15,504 number of combinations and it is not possible to process all combinations in the GNSS receiver. Using hit and trial approach, it is found that it is possible to relate the sample size  $n_{sub}$  to  $CNo$  such that the  $n_{sub} = (n + m) - k_{CNo < thr}^{obs}$ .  $k$  is the number of observations having  $CNo$  smaller than a defined threshold.

Because the presence of contaminated observations, the final position estimate  $\hat{X}_{rb}$  must meet the following condition,

$$\sum_{l=1}^{n+m} \mu \left( \frac{r_l(\hat{X}_{rb})}{s_{rb}} \right) = \min \tag{22}$$

$s_{rb}$  is the corresponding scale factor to the robust estimate  $\hat{X}_{rb}$ .

The proposed estimation procedure attains the required accuracy and consistency in two stages. The first stage is aimed to compute a reliable initial estimate. The second stage tries to attain the required accuracy. The two-stage algorithm is described in the following.

#### 6.1. $X_0$ Estimation (Stage 1)

1. Select  $P$  subsamples of size  $n_{sub}$  from all  $n + m$  observations. The number of the subsamples is,

$$\bar{N} = \binom{n + m}{n_{sub}}, \quad \#(P) = n_{sub} \tag{23}$$

$n_{sub}$  represents the number of observations in one subsample. The value  $n_{sub}$  should always meet the criteria of  $g_n < n_{sub} < n + m$ .  $g_n$  is the minimal number of observations required for positioning.  $n_{sub}$  is selected so that at least one combination contains all observations with high  $CNo$ . However, Equation (23) is not directly applicable to a multiconstellation positioning system. For a multiconstellation positioning system, the number of unknowns are  $3 + \mathfrak{R}$ , here  $\mathfrak{R}$  is the number of constellations used in positioning (clock correction parameters). In order to attain the benefits of the multiconstellation positioning, subsamples must contain at least one

observations from each constellation, for example a two constellations based positioning system with five observations independently, maximum four observations can be selected from either constellation to fulfill the minimum requirement as minimum five observations are required for positioning.

- Let  $p_k$  is a subsample belonging to  $P$  where  $P = \{p_1, p_2, \dots, p_{\bar{N}}\}$ . We compute the initial estimate  $X_{p_k}^0$  with the WLSE method,

$$X_{p_k}^0 = \left( \mathbf{H}_{p_k}^T \Sigma_{p_k}^0 \mathbf{H}_{p_k} \right)^{-1} \mathbf{H}_{p_k}^T \Sigma_{p_k}^0 \rho_{p_k} \tag{24}$$

Here  $\mathbf{H}_{p_k}$ ,  $\rho_{p_k}$  and  $\Sigma_{p_k}^0$  are the geometry matrix, the observation vector and the weight matrix for the subset  $p_k$ .  $\Sigma_{p_k}^0$  is a diagonal observation weights matrix  $diag(1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_{n_{sub}})$  and initial weights are calculated by using following relation [34].

$$\sigma_l^2 = \frac{\lambda^2}{2(T \cdot 10^{(CNo)/10})^2} + \frac{\lambda^2}{4T \cdot 10^{(CNo)/10}} \tag{25}$$

$T$  is the predetection integration time and  $\lambda = c/f_{code}$  is the code wavelength.  $CNo$  is the carrier-to-noise ratio for corresponding satellite signal.

- Compute the residual vector  $\mathbf{r}_{p_k}^0$ , scale factor  $s_{p_k}^0$  and standardized residuals,

$$\begin{aligned} \mathbf{r}_{p_k}^0 &= \rho_{p_k} - \mathbf{H}_{p_k} X_{p_k}^0 \\ s_{p_k}^0 &= 1.4826 \times \underset{\sim}{median}(\mathbf{r}_{p_k}^0) \\ \mathbf{r}_{s,p_k}^0 &= \frac{\mathbf{r}_{p_k}^0}{s_{p_k}^0} \end{aligned} \tag{26}$$

- Tukey bi-weight function (Table 1) is adopted to update the weighting matrix  $\Sigma_{p_k}$ .

$$\begin{aligned} w_l &= 1 \quad \text{if } \left| \underset{\sim}{r}_{s,p_k,l} \right| < \alpha \\ w_l &= 0 \quad \text{if } \left| \underset{\sim}{r}_{s,p_k,l} \right| \geq \alpha \end{aligned} \tag{27}$$

Here  $\alpha$  is the tuning factor to balance the accuracy and the breakdown robustness and can be selected from the look up table in Appendix A.  $\alpha = 4.658$  is selected to maintain 95% efficiency. Then updated weighting matrix will be  $\Sigma_{p_k} = diag(\{w_l : l = 1, 2, \dots, n_{sub}\})$ .

- Use Equation (24) again with updated weight matrix to compute the unknowns, then we get the estimate  $X_{p_k}$  for the subsample set  $p_k$ .

### 6.2. $X_{rb}$ Estimation (Stage 2)

- Apply steps 1~5 to all subsets and we get a set of subsample estimates and their corresponding scale factors,

$$\begin{aligned} \beta &= \{X_{p_1}, \dots, X_{p_{\bar{N}}}\} \\ \tilde{s} &= \{s_{p_1}, \dots, s_{p_{\bar{N}}}\} \end{aligned} \tag{28}$$

Next sort up the scale vector  $\tilde{s}$  in ascending order as  $\tilde{s}_{asc}$ . Take the 5 minimal scale factors from  $\tilde{s}_{asc}$  and their corresponding estimates (the number 5 is determined empirically by the authors, it can be changed). Up to now we obtained the most robust 5 initial estimates  $\tilde{\beta} = \{X_1, \dots, X_5\}_{asc}$ , their relative scale factors  $\tilde{s} = \{s_1, \dots, s_5\}_{asc}$  and corresponding subsamples  $\tilde{P} = \{p_{X_1}, \dots, p_{X_5}\} \cdot p_{X_4}$

is the subsample used to compute  $X_q$  where  $X_q \in \widetilde{\beta}$ . We apply the above steps 1~5 iteratively to the selected subsets  $\widetilde{P}$ . The iteration stops only when the convergence point is achieved,

$$\frac{\|X_i^q - X_{i-1}^q\|}{\|X_i^q\|} < k_{cc} : q = 1, 2, \dots, 5 \tag{29}$$

$k_{cc}$  is set as 0.001. The only difference in this step is the weighting function that we adopt Tukey bi-weight function type 2,

$$w_{l,i}^q = \begin{cases} \left\{ \left[ 1 - \left( \frac{\widetilde{r}_{l,i}^q}{\alpha} \right)^2 \right]^2 & \left| \widetilde{r}_{l,i}^q \right| \leq \alpha \\ 0 & \left| \widetilde{r}_{l,i}^q \right| > \alpha \end{cases} \tag{30}$$

Once all the selected 5 subset estimates have got converged, the one with minimal scale factor will be taken as the most robust initial estimate  $X_{rb}^0$  for the final robust estimation.

- Based on the robust initial estimate  $X_{rb}^0$  and its scale factor  $s_{rb}^0$ , we use all  $n + m$  observations at this step to get the final robust estimate. The computation procedure is the same as step 1~5, or equivalently as the IRWLS algorithm in Appendix B. This iteration will not stop until the convergence point, Equation (29) is met, and the final estimate  $X_{rb}$  is obtained.

### 7. Outlier Identification and Exclusion

The RAIM is a commonly used technique in the GNSS receivers for assurance of integrity and reliability of the estimated position. The fundamental task of the RAIM is the identification and exclusion of outlying observations for error free positioning. In the classic RAIM approach, the least-squares residual is used as the test statistic for outlier identification. As a rule of thumb, if the sum of squared residuals follows the chi-squared distribution, presence of outlier is rejected, otherwise observations are verified for outlier and the observation with the highest residual error is considered as the outlier. The primary drawback of this approach is that it is also possible that a wrong estimated position or exclusion of a wrong observation may lead to successful chi-square test. Conversely, in this paper, our proposed estimation technique itself maintains the integrity of estimated position by selection of the best fit observations. If  $\bar{\epsilon}$  proportion of the observations are outliers, the probability of the outlier free observation is  $\bar{\mu} = (1 - \bar{\epsilon})^{n_{sub}}$ ,  $n_{sub}$  is the number of observations in one subsample and the probability of at least one good subsample is  $1 - (1 - \bar{\mu})^{\bar{N}}$ ,  $\bar{N}$  is the total number of subsamples.

However, to enhance the positioning accuracy and reliability, we also proposed a two-step outlier detection and exclusion process. The proposed outlier detection and exclusion scheme is explained in the following,

- A predetection test is performed to predict the possible outlier observations prior to the FDE (fault detection and exclusion) process. A predefined threshold level is used to predict the outlier observations based on the confidence level and on the confidence interval. A confidence interval is the range of values those act as the good estimates and the confidence level is the proportion of the confidence interval that contains the true value of their corresponding parameter. The confidence interval can be calculated from data [35,36]. However, the desired confidence level is defined by the user. Let  $cl$  is the confidence level in predetection test, a threshold  $th_{cl}$  can be computed using normal inverse cumulative distribution function (also called quantile function).

$$t_{cl} = F^{-1}(\kappa | \widetilde{\mu}, \widetilde{\sigma}) = \{t_{cl} : F(t_{cl} | \widetilde{\mu}, \widetilde{\sigma}) = \kappa\} \tag{31}$$

Here  $\kappa = (cl + 1)/2$  is the probability for corresponding confidence level with mean  $\tilde{\mu}$  and standard deviation  $\tilde{\sigma}$  and  $F(t_{cl} | \tilde{\mu}, \tilde{\sigma}) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \int_{-\infty}^{t_{cl}} e^{-\frac{(t-\tilde{\mu})^2}{2\tilde{\sigma}^2}} dt$  is the cumulative distribution function (CDF). The scaled observation residual is computed as,

$$\bar{u}_{sc} = \bar{S}Y \tag{32}$$

Here  $Y = (\tilde{\rho} - \hat{\rho})/s_{sc}$  and  $\bar{S} = I - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$ . As soon as  $\gamma = \sum_{l=1}^{n+m} \bar{u}_{sc,l}^2$  follows  $\chi^2$  distribution with  $\bar{n} - \bar{v}$  degree of freedom, the estimated position is considered as reliable, where  $\bar{n}$  is the number of observations used for integrity monitoring and  $\bar{v}$  is the number of unknown to be estimated.

2. Otherwise, a set of observations is selected as the potential outliers based on the following relation,

$$\{\rho_l \in \{outliers\} | \bar{u}_{sc,l} > t_{cl} \sqrt{\bar{u}_{sc}} \wedge \rho_l \leftrightarrow \bar{u}_{sc}\} \tag{33}$$

Let  $\varphi = \arg_{\bar{u}_{sc,l} > t_{cl}} (\rho_{outlier,k_1}, \rho_{outlier,k_2}, \dots, \rho_{outlier,k_n})$  is the set of predicted outlier observations.  $k_l$  is corresponding observation number. The FDE on the predicted set of outliers is applied as follows,

$$\varepsilon_l = \frac{|h_{k_l}^T \bar{u}_{sc,k_l}|}{\sqrt{\bar{S}_{ll,k_l}}} \tag{34}$$

Any faulted observation will satisfy the  $|\varepsilon_j| \geq |\varepsilon_l| \forall l$ .

### 8. Discussion

Here we study some key modifications and properties of the proposed algorithm. As described in Section 6, the fundamental algorithm is a two-stage process; the first stage involves the initialization of a best fit estimate in the presence of outliers while the second stage involves the convergence of the initial estimate. Subsampling process is the key point to find an initial robust estimate. Subsampling is also the most critical task as the number of subsamples determines the computational cost. Thus the number of observations per subsample is the most significant factor in determining the computational load. In our proposed scheme,  $n_{sub}$  is a selectable factor, hence it is possible to tune the computational cost of algorithm.

The LSE estimator has low complexity level in terms of computational cost. IRWLS contains more complex structure than simple LSE as it requires additional arithmetic operations related to residual evaluation and updating of the observation weights and these operations are performed in iterative manner. In statistics, a huge amount of observation data, either qualitative or quantitative, is needed to be analyzed before making final decision. There are numerous statistical estimation techniques, such as robust and multivariate regression procedures [31,32], those provide general interpretation of data based on random samples drawn from original dataset as it is not possible to process all datasets, and additionally enough processing power is available with no real time requirements. On the other hand, the dataset such as GNSS has very small number of observations with certain characteristics, such as GDOP, and random sampling technique can lead to a wrong estimate. A multiconstellation system can provide better reliability and integrity, thus each subsample should contain the observations from all available constellations. And the  $n_{sub}$  factor is managed in such a way that at least one subsample must contain all best fit satellites from both constellations.

Several authors proposed similar approaches for GPS positioning based on robust statistics. The idea in [18], is discussed in details to compare with our proposed method as there are some similarities among two algorithms. The key idea, in [18], is to find all possible solutions based on

the minimum number of required observations per solution, and a genetic algorithm is applied for outlier detection and exclusion. CNo of GPS and BeiDou visible satellites is shown in Figure 2 for opensky data.

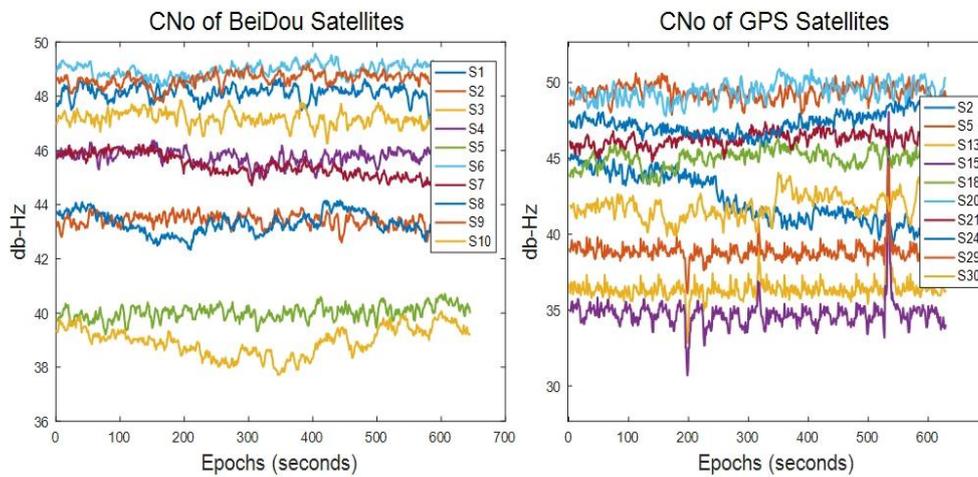


Figure 2. CNo of BeiDou and GPS-visible satellites for open sky data.

Signals from all visible satellites have well CNo. Thus it is hard for any observation to be an outlier and applying subsampling theorem in such a scenario is an impractical approach. And multi-GNSS positioning further increases the complexity in using minimum number of required number of observations. On the minimum requirement criteria basis, there will be multiple subgroups those will contain either GPS only or BeiDou only observations, thus leaving the advantages of multiple constellation positioning.

### 9. Performance Validation and Analysis

Multiple scenarios and environmental circumstances are assessed to validate the performance of the proposed estimator. Considered scenarios comprise static as well as dynamic environments. Two static locations are rooftop (for line of sight signal propagation model) and a location surrounded by skyscrapers (for multipath in urban environment) in the downtown of Shanghai. In Figure 3a,b, the two locations are shown. In open sky scenario (roof top) all signals are line of sight with no multipath or other interferences. In contrast, there are multipath as well as other radio interferences in urban canyon scenario.



Figure 3. GNSS antenna position (a) open sky (rooftop) (b) urban canyon (downtown Shanghai).

For stationary scenarios, accuracy is examined in terms of 90% position error radius (RCEP = radial circular error position) and position error mean distance (PMD) from the center. Results are analyzed against LSE and WLSE (weighted least squares estimation) algorithms. In this paper, we applied the SIGMA – ε weight model [37] for observation weight estimation.

$$\sigma_l^2(i) = a. \exp(-CNo/10) \tag{35}$$

$\sigma_l(i)$  is the variance of  $l$ th observation at  $i$ th epoch.  $a$  is the scaling constant and depends upon the noise characteristics of the tracking loop or carrier tracking loop noise bandwidth. And the WLSE is defined as;

$$\Delta x = \left( \mathbf{H}^T \Sigma^{-1} \mathbf{H} \right)^{-1} \mathbf{H}^T \Sigma^{-1} \Delta \rho \tag{36}$$

$\Sigma^{-1}$  is the weight matrix where  $\Sigma$  is the diagonal observation variance-covariance matrix.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} \tag{37}$$

### 9.1. Open Sky Scenario

In the open sky scenario, the only errors are Gaussian white noise from tracking loops, atmospheric and receiver/satellite clock corrections. In Table 3, GNSS signal information for open sky scenario is represented.

Horizontal radial error for open sky environment is shown in the Figure 4 in east and north error terms. There is a very little difference between RCEP of three models. However, PMD of the proposed scheme is larger than the other two schemes (except the difference is substantively small). The primary cause of better performance of LSE/WLSE is the presence of Gaussian white noise. As RCEP provides the 90% position error radius, results indicate that WLSE and the proposed schemes are more robust although the mean position error distance of LSE is smaller. Two additional simulations are also created by modifying the noise characteristics of observations.

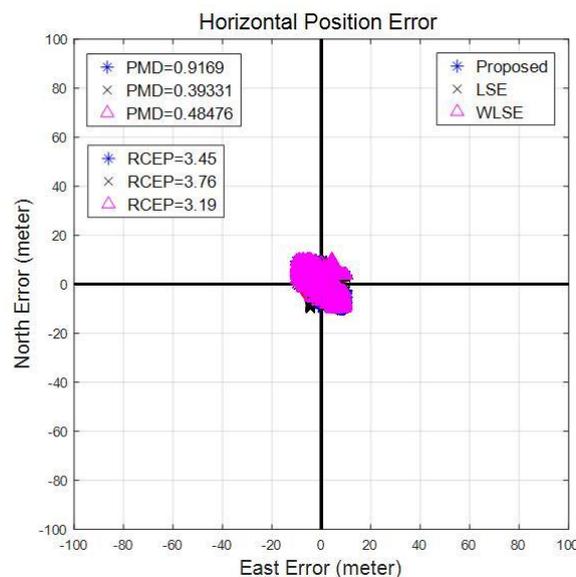


Figure 4. Horizontal radial error of open sky static scenario.

**Table 3.** GNSS signal information for open sky scenario.

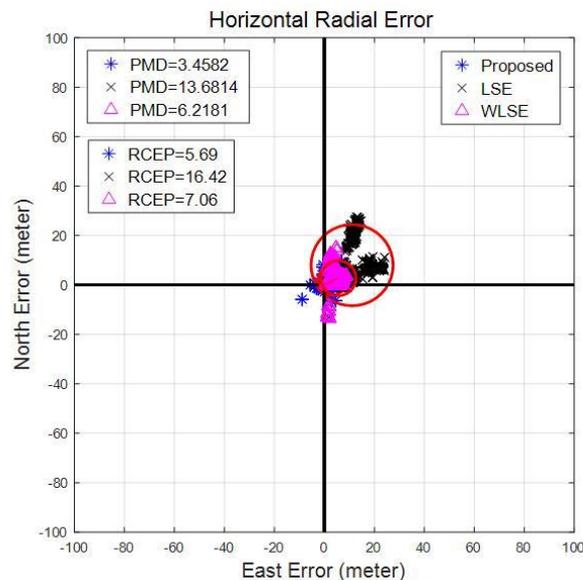
GNSS	Signal Type	Intermediate Frequency	No. of Satellites	Satellite IDs
BeiDou	B1I	−6.902 (MHz)	10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
GPS	L1-CA	7.42 (MHz)	10	2, 5, 13, 15, 18, 20, 21, 24, 29, 30

9.1.1. Emulated Scenario 1

The statistics of the first emulated scenario are presented in the following Table 4. This scenario is designed to evaluate the positioning accuracy in the presence of fixed outliers. Measurement errors stay fixed for a satellite over a defined period of time. The measurement error is kept small because a large error is easy to detect by a simple algorithm however it is more difficult to detect if the error is not very large and show significant effect. Results, in Figure 5, showed that the performance of our proposed estimation scheme is better than LSE and WLSE in both RCEP and PMD.

**Table 4.** Statistics of Emulated Scenario 1.

Time (s)	Outliers	Bias (m)	PRN with Bias	
			BeiDou	GPS
1–100	0	0	0	0
101–200	1	500	3	0
201–300	2	[500, 900]	4, 9	0



**Figure 5.** Horizontal radial error for emulated scenario 1.

9.1.2. Emulated Scenario 2

The objective of this scenario is to evaluate the positioning reliability and accuracy for unstable observation data. Figure 6 represents the dynamic environment simulation scenario.

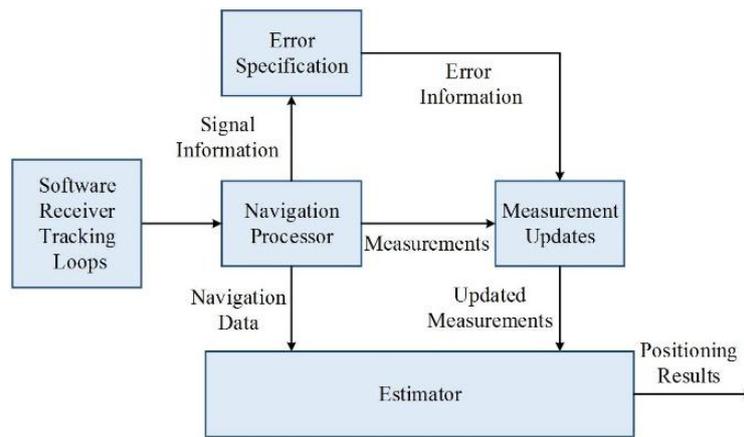


Figure 6. Flowchart of observation error simulation in true navigation data.

Specifications include the details about PRN, measurement error, number of outliers etc. Navigation module receives the signal data from tracking loops and errors are introduced in the real measurements dynamically based on the error specifications. PRN/Sat IDs are selected at random and are changed every 10 s. Results are presented in Figure 7. Position mean distance error of our proposed estimator is significantly smaller as well as RCEP. On the other hand, WLSE is also performed well against LSE. One important point to consider is that if the observation is effected by un-modeled noise sources such as radio interferences, the performance of WLSE is degraded though observations with high CNo because the weight estimation scheme for WLSE is based on the observation CNo.

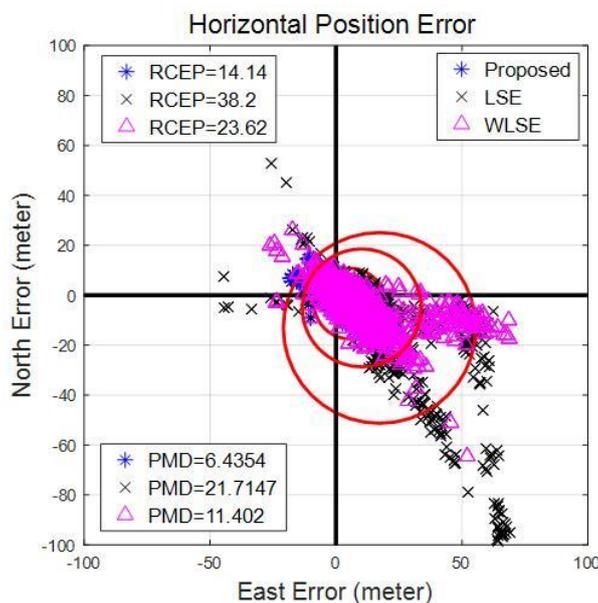


Figure 7. Error plot for emulated scenario 2.

### 9.2. Multipath Scenario in Urban Canyons

This scenario is designed to analyze the performance of the proposed algorithm in severe urban environments. Multipath and line of sight problem are the primary source of errors in urban environments. Experiment location is surrounded by high skyscrapers as shown in the Figure 1b. In Figure 8 CNo of visible satellite is shown from GPS and BeiDou satellites.

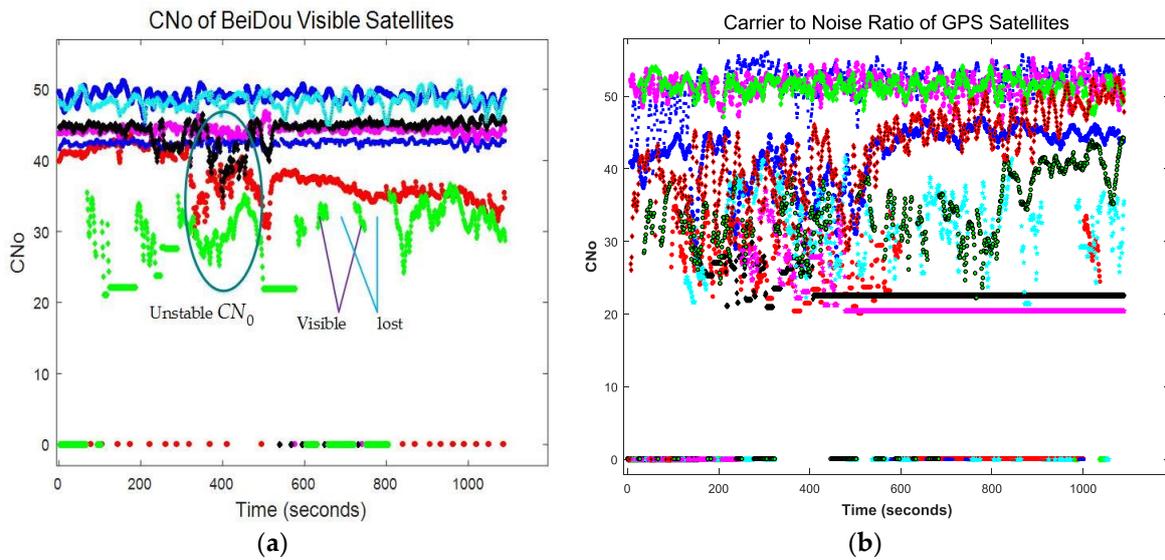


Figure 8. Carrier-to-noise ratio of visible satellites (a) BeiDou (b) GPS.

Figure 8 demonstrates that CNo of both BeiDou and GPS is unstable. And the disruption in line represents the loss of satellite visibility. Moreover the multipath signal interferences introduce additional errors in observations.

Positioning results are provided in the Figure 9 in terms of RCEP and PMD. As RCEP represents the accuracy of estimated position during most of time, it can be found in the figure that the proposed estimation technique significantly improved the positioning accuracy.

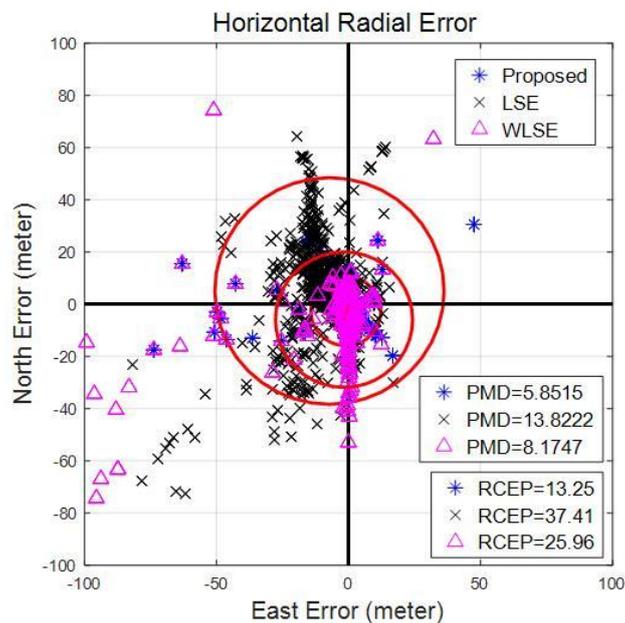


Figure 9. HRE of proposed estimator against LSE and WLSE.

### 9.3. High Dynamic Moving Scenario

The most critical requirement for every GNSS receiver is the performance under high dynamics when moving in the severe surroundings. Examples of harsh surroundings are moving under trees, under elevated roads and in urban canyons with partial visible sky. High dynamic characterizes the condition when the Doppler shift may vary in the range of  $\pm 60$  kHz. Although this is difficult to

achieve in normal urban canyons, still at carrier-to-noise ratio of the order of 30–40 dB-Hz and sudden loss of tracking satellites make it challenging to maintain the navigation accuracy. The following scenario is designed to evaluate the dynamic performance of the proposed estimation procedure against classic LSE and WLSE algorithms. GNSS signal data is recorded in remote rural area near shanghai during driving on the road with antenna fixed on the roof of van. Ephemeris information and other navigation data is produced by offline processing of signal data.

Experiment trajectory of vehicle is shown in Figure 10. The experiment route is designed so that there must be multiple full and partial sky visibility regions. Figure 11 illustrates the  $CN_0$  of GPS satellites. As the  $CN_0$  of satellites represent the signal characteristics, observations from the satellites, with unstable  $CN_0$ , can lead the estimator to wrong approximation.

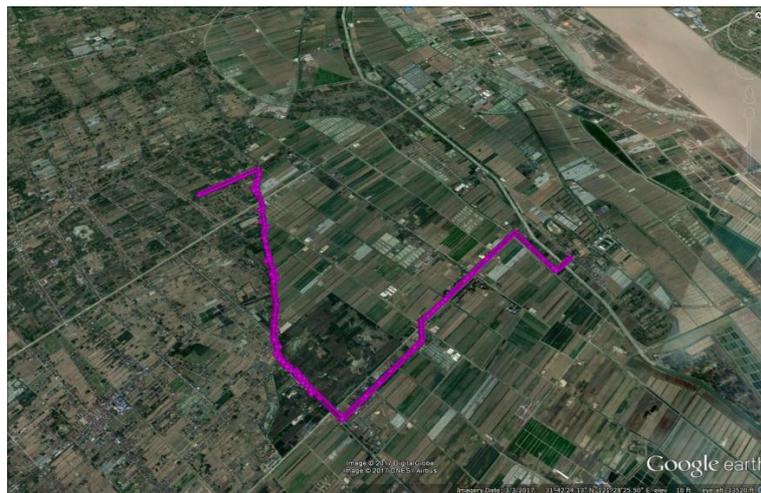


Figure 10. Vehicle route for data logging (courtesy of Google Earth).

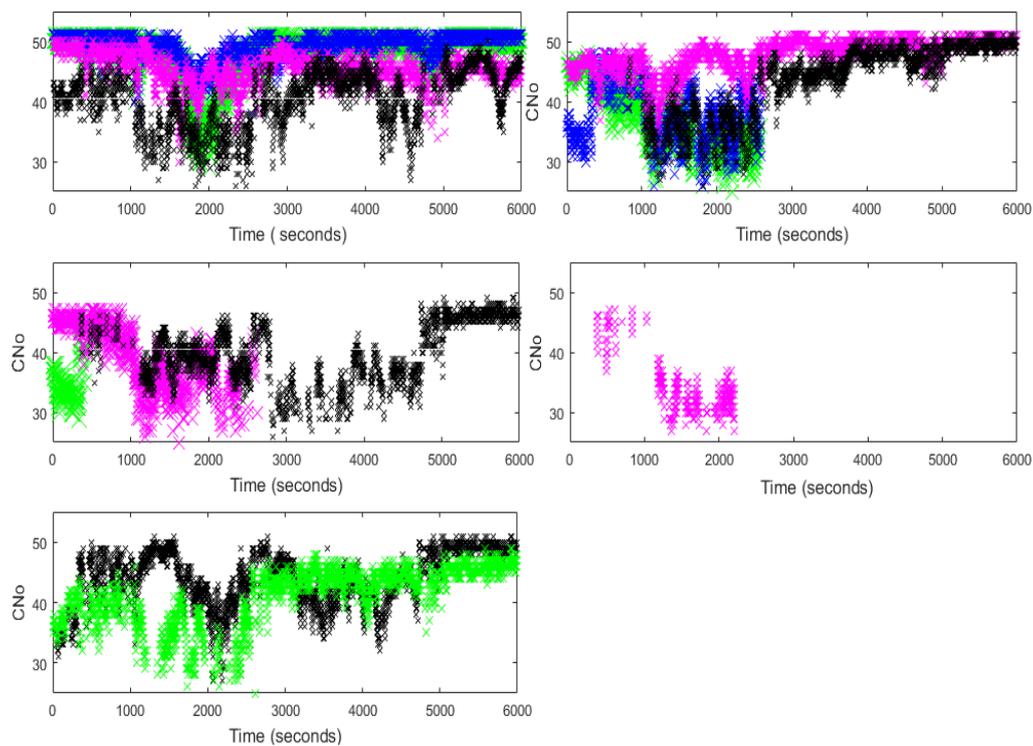
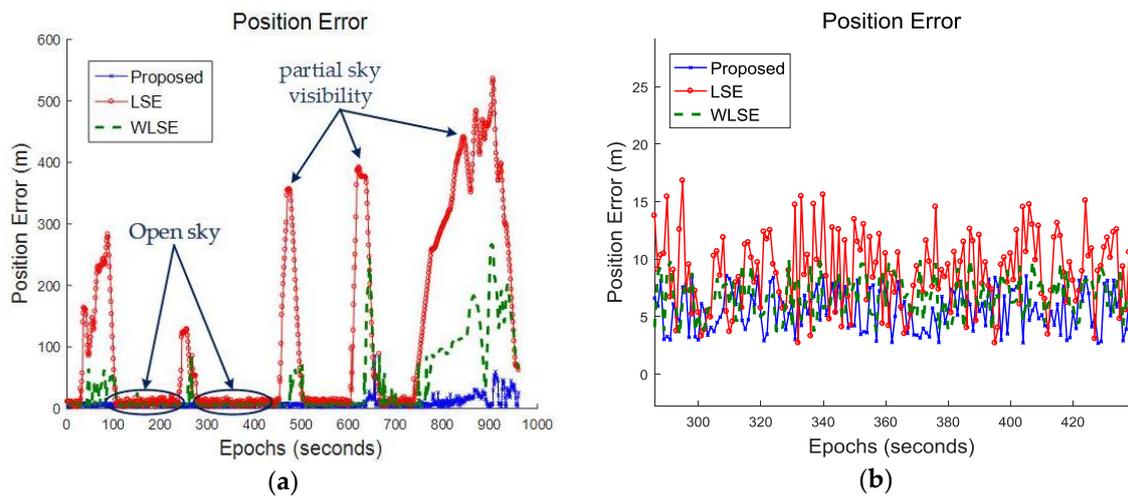
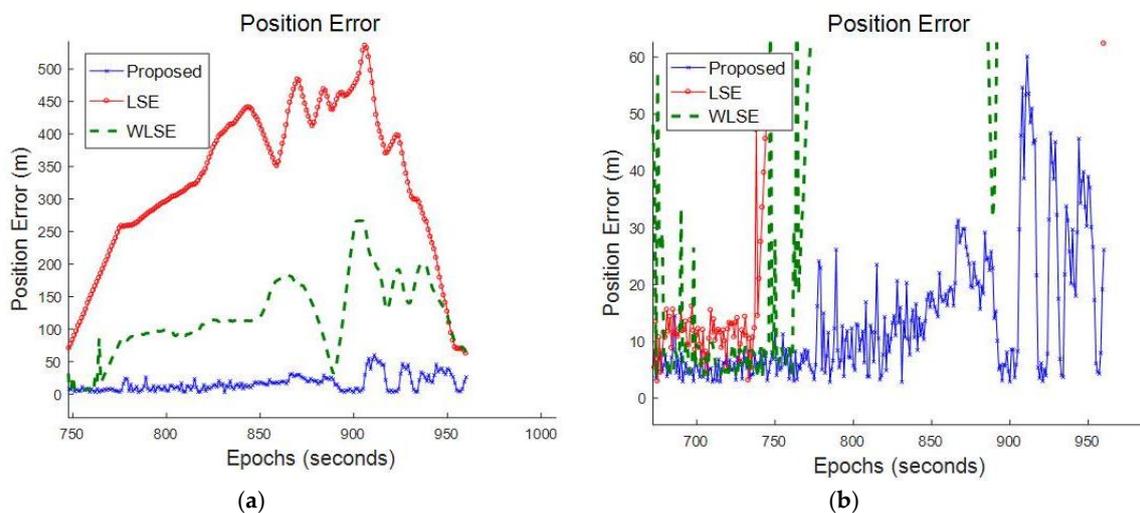


Figure 11.  $CN_0$  of visible GPS satellites.

Figure 12a shows the errors in the estimated position from LSE, WLSE and our proposed algorithms. As described before, there are several open sky and partial visible sky regions, mentioned with arrows. In Figure 12b, position errors for open sky are zoomed. Estimated positions errors, for the duration of driving under trees, are showed in Figure 13. From Figure 11b, it is found that the performance of three algorithms is similar under (LoS) line-of-sight signal propagation conditions such as open sky. In LoS conditions, signals are biased with Gaussian-only white noise, hence LSE showed high positioning accuracy.



**Figure 12.** (a) Estimated position error from three algorithms; (b) estimated position error during open sky region (zoomed in).



**Figure 13.** Estimated position errors when during driving under trees (zoomed in).

However, the circumstances are quite different for NLoS (non-line of sight) conditions i.e., driving under trees. During NLoS conditions, other than Gaussian errors are dominant such as multipath, signal diffraction or reflection etc. We cannot say that the standalone positioning accuracy of the proposed scheme is very high (Figure 13b), however there is a significant improvement in positioning accuracy in comparison to LSE and WLSE.

Positioning error for moving scenario from three algorithms are analyzed in Table 5. Maximum and RMS (root mean square) position errors from the proposed estimation algorithm are significantly lower than LSE and WLSE. The maximum error are resulted when vehicle was moving under dense trees and majority of the received observations were corrupted.

**Table 5.** Analysis of position errors from three algorithm for moving scenario.

Error \ Algorithm	LSE	WLSE	Proposed
min	1.4615	3.3358	2.6673
max	535.8151	266.4498	94.1103
mean	111.6460	39.3734	10.7022
std	144.4107	54.1297	11.8657
RMS	182.5304	66.9329	15.9787

### 10. Conclusions and Future Perspectives

A novel, real-time GNSS positioning estimator is presented in this paper for land vehicles in challenging environments. The proposed estimation procedure is based on robust MM estimation theory. The presence of outlier observations in GNSS measurements significantly affect the positioning results. This method is capable of approximating the user position with accuracy and reliability in the presence of outlying observations by using the most trusted observations for position estimation and excluding outliers. The proposed algorithm is compared against LSE and WLSE algorithms in terms of 90% positioning accuracy. Extensive field and simulation tests are carried out to analyze the performance of the proposed scheme in diverse environments. Results demonstrate that the proposed scheme could be a promising scheme for GNSS-based positioning with high accuracy under diverse conditions, especially in harsh and challenging environments, because LSE and WLSE are also capable to provide similar accuracy in normal outlying free environments, however they are not able to maintain the accuracy in challenging environments. For future extensions, we plan to explore the applications of the algorithm in more complex GNSS-based positioning architectures such as vector-based GNSS tracking receivers and fusion of GNSS measurements with other sensors such as IMU (Inertial Measurement Unit) and vision sensors to further improve the positioning accuracy. We also want to explore the application of the proposed estimation scheme (especially the robust initialization stage) for integration with the EKF (Extended Kalman Filter) in order to enhance the filter performance.

**Author Contributions:** P.L. help out in resolving the research direction; M.A.A. and P.L. designed the experiments; P.L. facilitated the analysis and experimental tools; Y.W. and J.Q. helped to perform and log the experimental data; M.A.A. analyzed the test data; M.A.A. wrote the paper.

**Acknowledgments:** The authors wish to express their gratitude for the support of Prof. Peilin Liu and BATC (Brain-inspired Application Technology Center) in providing facilities and equipment to complete this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

Relative efficiency for different breakdown points using Tukey bisquare function

**Table A1.** Relation between Efficiency and Break Down Point (BDP).

BDP	Efficiency	ff	K
50%	28.7%	1.547	0.1995
45%	37.0%	1.756	0.2312
40%	46.2%	1.988	0.2634
35%	56.0%	2.251	0.2957
30%	66.1%	2.560	0.3278
25%	75.9%	2.973	0.3593
20%	84.7%	3.420	0.3899
15%	91.7%	4.096	0.4194
0.10%	96.6%	5.182	0.4475

### Appendix B

#### Iterative Reweighted Least Squares Algorithm (IRWLS)

1. Defined convergence constant  $k_{cc}$  and calculate initial estimate using weighted least squares using  $X_0 = (\mathbf{H}^T \Sigma \mathbf{H})^{-1} \Sigma \mathbf{H}^T \rho$
2. Compute initial residual vector using  $\mathbf{r}_0 = \rho - \mathbf{H}X_0$
3. Calculate initial standardized residuals;  $\widetilde{\mathbf{r}}_{s,0} = \frac{\mathbf{r}_0}{s_0}$
4. Tukey bi-weight function type1 with tuning constant  $\alpha$  is applied.

a.  $w_{l,0} = 1$  if  $|\widetilde{r}_{s,l}| < \alpha$

b.  $w_{l,0} = 0$  if  $|\widetilde{r}_{s,l}| \geq \alpha$

5.  $\Sigma_0 = \text{diag}(\{w_{l,0} : l = 1, 2, \dots, n\})$

6. while

a.  $X_i = (\mathbf{H}^T \Sigma_{i-1} \mathbf{H})^{-1} \mathbf{H}^T \Sigma_{i-1} \rho$

b. If  $\frac{\|X_i - X_{i-1}\|}{\|X_i\|} < k_{cc}$ ; break

c. If  $\frac{\|X_i - X_{i-1}\|}{\|X_i\|} \geq k_{cc}$ ; continue

d.  $\mathbf{r}_i = \rho - \mathbf{H}X_i$

e.  $s_i = 1.4826 \times \text{median}|\mathbf{r}_i|$

f.  $\widetilde{\mathbf{r}}_{s,i} = \frac{\mathbf{r}_i}{s_i}$

g. Calculate  $\Sigma_i = \{w_{1,i}, \dots, w_{n,i}\}$  using Tukey bi-weight function type 2.  $n$  is the number of observations.

$$w_{l,i} = \begin{cases} \left\{ \left[ 1 - \left( \frac{|\widetilde{r}_{s,i}^l|}{\alpha} \right)^2 \right]^2 & |\widetilde{r}_{s,i}^l| \leq \alpha \\ 0 & |\widetilde{r}_{s,i}^l| > \alpha \end{cases}$$

Here  $l = \{1, \dots, n\}$

h. go to step a.

7. end

### Appendix C

Let  $nm = n + m$  are the total number of observations and  $n_{sub}$  is the number of observations in one subsample and  $\overline{N} = \binom{nm}{n_{sub}}$  are the total number of subsamples. For  $n_{sub} \leq nm$ , subsampling is described with the following steps,

1. Select  $P = \{p_1, p_2, \dots, p_{\overline{N}}\}$  subsamples of size  $n_{sub}$  and  $\overline{N} = \binom{nm}{n_{sub}}$  number of subsamples.

$$\{(\mathbf{H}_k, \rho_k) : k \in P\}, P \subset \{1, \dots, nm\}, \#(P) = n_{sub}$$

2. Compute initial estimate for each subsample using weighted least squares estimation (WLSE)

$$X_{p_k} = (\mathbf{H}_{p_k}^T \Sigma_{p_k} \mathbf{H}_{p_k})^{-1} \Sigma_{p_k} \mathbf{H}_{p_k}^T \rho_{p_k}$$

Here  $\{p_k : k = 1, \dots, \overline{N}\}$  are the subsamples.

## References

1. Moka, E.; Paul, A. Cross A Fast Satellite Selection Algorithm for Combined GPS and GLONASS Receivers. *J. Navig.* **1994**, *47*, 383–389.
2. Zhang, M.; Zhang, J. A Fast Satellite Selection Algorithm: Beyond Four Satellites. *IEEE J. Sel. Top. Signal Process.* **2009**, *3*, 740–747. [[CrossRef](#)]
3. Zhu, J.J. Calculation of geometric dilution of precision. *IEEE Trans. Aerosp. Electron. Syst.* **1992**, *28*, 893–895.
4. Wei, M.; Wang, J.; Li, J. A new satellite selection algorithm for real-time application. In Proceedings of the International Conference on Systems and Informatics (ICSAI), Yantai, China, 19–20 May 2012; pp. 2567–2570.
5. Kong, J.; Mao, X.; Li, S. BDS/GPS satellite selection algorithm based on polyhedron volumetric method. In Proceedings of the IEEE/SICE International Symposium on System Integration (SII), Tokyo, Japan, 13–15 December 2014; pp. 340–345.
6. Zhang, X.Y.; Huang, Z.G.; Li, R. RAIM analysis in the position domain. In Proceedings of the IEEE/ION Position Location and Navigation Symposium (PLANS), Indian Wells, CA, USA, 4–6 May 2010; pp. 53–59.
7. Hewitson, S.; Lee, H.K.; Wang, J. Localizability Analysis for GPS/Galileo Receiver Autonomus Integrity Monitoring. *J. Navig.* **2004**, *57*, 245–259. [[CrossRef](#)]
8. Baarda, W. *A Testing Procedure for Use in Geodetic Networks*. Netherlands Geodetic Commission; Publications on Geodesy, New Series 2, No. 5; Rijkscommissie voor Geodesie: Delft, The Netherlands, 1968.
9. Blanch, J.; Ene, A.; Walter, T.; Enge, P. An Optimized Multiple Hypothesis RAIM Algorithm for Vertical Guidance. In Proceedings of the 20th ITM, The Institute of Navigation, Fort Worth, TX, USA, 25–28 September 2007; pp. 2924–2933.
10. Martineau, A.; Macabiau, C.; Mabillean, M. GNSS RAIM assumptions for vertically guided approaches. In Proceedings of the 22nd International Technical Meeting of The Satellite Division of the Institute of Navigation, Savannah, GA, USA, 22–25 September 2009; The Institute of Navigation: Savannah, GA, USA, 2009; pp. 2791–2803.
11. Kaplan, E.D.; Hegarty, C.J. *Understanding GPS, Principles and Applications*, 2nd ed.; Artech House: Norwood, MA, USA, 2005.
12. Li, J.; Wu, M. The Improvement of Positioning Accuracy with Weighted Least Square Based on SNR. In Proceedings of the 2009 5th International Conference on Wireless Communications, Networking and Mobile Computing, Beijing, China, 24–26 September 2009; pp. 1–4.
13. Tabatabaei, A.; Mosavi, M.R.; Khavari, A. Reliable Urban Canyon Navigation Solution in GPS and GLONASS Integrated Receiver Using Improved Fuzzy Weighted Least-Square Method. *Wirel. Pers. Commun.* **2017**, *94*, 3181–3196. [[CrossRef](#)]
14. Rahemi, N. Accurate Solution of Navigation Equations in GPS Receivers for Very High Velocities Using Pseudorange Measurements. *Adv. Aerosp. Eng.* **2014**, *2014*, 435891. [[CrossRef](#)]
15. Tay, S.; Marais, J. Weighting models for GPS Pseudorange observations for land transportation in urban canyons. In Proceedings of the 6th European Workshop on GNSS Signals and Signal Processing, Munich, Germany, 5–6 December 2013.
16. Jwo, D.J.; Chang, S.J. Outlier Resistance Estimator for GPS Positioning. In Proceedings of the 2006 IEEE International Conference on Systems, Man and Cybernetics, Taipei, Taiwan, 8–11 October 2006; pp. 657–662.
17. Rao, K.D.; Swamy, M.N.S.; Plotkin, E.I. GPS navigation with increased immunity to modeling errors. *IEEE Trans. Aerosp. Electron. Syst.* **2004**, *40*, 2–11.
18. Fallahi, K.; Cheng, C.T.; Fattouche, M. Robust Positioning Systems in the Presence of Outliers Under Weak GPS Signal Conditions. *IEEE Syst. J.* **2012**, *6*, 401–413. [[CrossRef](#)]
19. Kim, E.; Walter, T.; Powell, J.D. Adaptive carrier smoothing using code and carrier divergence. In Proceedings of the ION NTM, San Diego, CA, USA, 22–24 January 2007; pp. 114–152.
20. Yoon, D.; Kee, C.; Seo, J.; Park, B. Position Accuracy Improvement by Implementing the DGNSS-CP Algorithm in Smartphones. *Sensors* **2016**, *16*, 910. [[CrossRef](#)] [[PubMed](#)]
21. Yohai, V.J. High breakdown point and high efficiency robust estimates for regression. *Ann. Stat.* **1987**, *15*, 642–656. [[CrossRef](#)]
22. Hampel, F.R. Beyond location parameters: Robust concepts and methods. *Bull. Int. Stat. Inst.* **1975**, *46*, 375–391.
23. Rousseeuw, P.J. Least median of squares regression. *J. Am. Stat. Assoc.* **1984**, *79*, 871–880. [[CrossRef](#)]

24. Davies, L. The asymptotics of S-estimators in the linear regression model. *Ann. Stat.* **1990**, *18*, 1651–1675. [[CrossRef](#)]
25. Rousseeuw, P.J.; Yohai, V. *Robust Regression by Means of S-estimates*. *Robust and Nonlinear Time Series Analysis*; Lecture Notes in Stati; Springer: New York, NY, USA, 1984; Volume 26, pp. 256–272.
26. Hössjer, O. On the optimality of S-estimators. *Stat. Probab. Lett.* **1992**, *14*, 413–419. [[CrossRef](#)]
27. Rousseeuw, P.J.; Leroy, A. *Robust Regression and Outlier Detection*; Wiley: New York, NY, USA, 1987.
28. Bryan, C.G. Weighted Least-Squares Orbit Estimation using GPS SPS Navigation Solution Data. Master's Thesis, San Jose State University, San Jose, CA, USA, 1995.
29. Delgado, N.B.; Nunes, F. Satellite selection based on WDOP concept and convex geometry. In Proceedings of the 5th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC), Noordwijk, The Netherlands, 8–10 December 2010; pp. 1–8.
30. Gervini, D.; Yohai, V.J. A class of robust and fully efficient regression estimators. *Ann. Stat.* **2002**, *30*, 583–616. [[CrossRef](#)]
31. Rey, W.J. *Introduction to Robust and Quasi-Robust Statistical Methods*; Springer: Berlin/Heidelberg, Germany, 1983.
32. Huber, P.J. *Robust Statistics*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009.
33. Draper, N.R.; Smith, K. *Applied Regression Analysis*, 3rd ed.; Wiley: New York, NY, USA, 1998.
34. Parkinson, B.; Spilker, J. *The Global Positioning System: Theory and Applications*, ser. *Progress in Astronautics and Aeronautics*; American Institute of Aeronautics and Astronautics, Inc., SW: Washington, DC, USA, 1996; Volume 1.
35. Javanmard, A.; Montanari, A. Confidence Intervals and Hypothesis Testing for High-Dimensional Statistical Models. In Proceedings of the Advances in Neural Information Processing Systems 26 NIPS, Lake Tahoe, CA, USA, 5–10 December 2013; pp. 1187–1195.
36. Cetin, M.; Aktas, S. Confidence Intervals Based on Robust Estimators. *J. Mod. Appl. Stat. Methods* **2008**, *7*, 21. [[CrossRef](#)]
37. Brunner, F.; Hartinger, H.; Troyer, L. GPS signal diffraction modelling: The stochastic SIGMA-D model. *J. Geodesy* **1999**, *73*, 259. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).