

Design and Evaluation of Robust M-estimators for GNSS Positioning in Urban Environments

Omar García Crespillo^{1,2}, Alice Andreotti², Anja Grosch¹

¹*Institute of Communications and Navigation, German Aerospace Center (DLR)*

²*Swiss Federal Institute of Technology Lausanne (EPFL)*

BIOGRAPHIES

Omar García Crespillo holds a M.Sc. in Telecommunication Engineering from the University of Malaga in Spain. In 2013, he joined the Navigation department of the German Aerospace Center (DLR) where his current field of research includes multi-sensor fusion algorithms, GNSS, inertial sensors and integrity monitoring for safe ground and air transportation systems. Since 2015, he is also a PhD student at the Swiss Federal Institute of Technology (EPFL) in Lausanne.

Alice Andreotti is currently a last year Master student in Environmental Engineering at the Swiss Federal Institute of Technology in Lausanne (EPFL). She chose the specialization in environmental monitoring and modeling and she is currently working on the master thesis in the field of GNSS positioning. She received in 2017 the bachelor degree from the same university in Environmental Science and Engineering. In 2019 she stayed at the German Aerospace Center (DLR) for a 4 months internship.

Anja Grosch received the German diploma in Computer Engineering from the Ilmenau University of Technology, Germany in 2007. After her graduation, she continued working for the communications department in the area of channel coding, OFDM systems and relay networks. Since March 2008, she has joined the Institute of Communications and Navigation of the German Aerospace Center (DLR). Her main focus has been multi-sensor fusion algorithms, especially GNSS and INS integration. Furthermore, she develops and validates suitable integrity concepts optimized for different safety-of-life applications such as railway, automotive and civil aviation.

Abstract

Providing a robust position solution in urban scenarios using Global Navigation Satellites Systems (GNSS) supposes a big challenge due to the continuous presence of local threats like multipath, non-light-of-sight signals (NLOS) or interference. The impact of these threats in the pseudorange measurements is difficult to model due to the constant changing environment. Therefore they have to be dealt with at position estimator level. In this paper we leverage the knowledge from robust statistics and we design and apply Huber M-estimators to the positioning and compare its performance with respect to the classical least-squares estimator. We provide further insights on the importance of the chosen suitable pseudorange error models for the robust estimator process by processing real GNSS data in open-sky and deep urban scenarios. Under a proper design, these estimators have the potential to become a powerful option in challenging GNSS scenarios under the presence of multiple corrupted measurements or data that do not follow the assumed error models.

I. INTRODUCTION

Robust navigation in urban environments represents one of the biggest challenges in localization nowadays. With the increment of safety and liability critical applications, the need for a robust position solution against the presence of sensors faults is a mayor area of research. This demand is accentuated by the global trend of achieving higher levels of vehicle automation with more stringent requirements.

Satellite navigation is considered to be one of the key elements for positioning worldwide. However, urban scenarios are one of the most challenging environments for GNSS. Not only because of the possible limited satellite visibility due to buildings and other obstacles (that can be compensated and coasted by means of additional sensors, e.g. inertial measurement units), but mainly due to the presence of a number of local threats that can corrupt completely the position estimation and that are very difficult to model, characterize and predict. These causes are mainly the presence of high multipath and Non Line Of Sight (NLOS) signals that can appear simultaneously corrupting multiple pseudorange measurements, and interference that can cause a drop in the Carrier to Noise ratio and increment drastically the pseudorange error.

The classical and nowadays most widely used GNSS position estimator is the least squares (LS), which is known to be highly sensitive to the presence of any measurement bias or error that is not properly modelled or following a Gaussian distribution. In fact, its robustness is assumed to be zero in the sense that any unbounded measurement error would translate into an unbounded position error. For this reason, the LS estimator is typically implemented in parallel with some Fault Detection and Exclusion (FDE) function. The reliability based on the pair Estimator + FDE has been widely developed in the civil aviation community through the design of different integrity monitoring algorithms. For example, the fault detection function implemented within Receiver Autonomous Integrity Monitoring (RAIM) is based on a Chi-Square statistical test to

target the presence of single faults [1]. More recently, the Advanced RAIM is based on solution separation test statistics in order to be able to detect the presence of multiple faults or hypothesis [2]. These algorithms are accompanied of overbounded pseudorange error models, where the different error components are upper bounded in order to protect the worst-case signal in space error. In the literature, we can find several attempts to adapt the methodology from (A)RAIM to different land-based applications [3, 4]. The first problem of implementing ARAIM algorithms for urban navigation is that it is still a challenge in practice to find realistic bounds for the effect of the local threats. Another problem with the use of solution separation fault detector is that a position solution must be computed for each possible hypothesis. Whereas in aviation due to the low risk of having a satellite fault, the number of hypothesis are normally considered only for 1 or 2 simultaneous satellite faults, in urban scenarios any measurement could be in principle corrupted simultaneously. This means that a very large number of hypotheses would be required to be tested with the associated computer requirement in order to guarantee a real-time operation.

An alternative to the parallel use of estimator and fault detector can be found in the framework of robust statistics [5]. In here, the classical least-square estimator, which is a MMSE (Minimum Mean Square Error) estimator is substituted by a different cost function. Prior work on the use of robust estimators for GNSS positioning can be found in the literature, for instance in [6, 7, 8]. Early publications showed however marginal improvement with respect to the LS. This might be buried in the fact that robust estimators need a sufficient large number of measurements to performed at its best. Under the presence of contaminated data, the set of *healthy* data must be large enough. To compensate for this deficiency, the use of additional sensors or the combination of multiple time epochs has been investigated [9, 10]. Another aspect that, to the best of our knowledge has not been well addressed yet, is the impact of the pseudorange error model design on the performance of the robust estimator.

In this paper, we first consider the positioning of GNSS with multiple constellations, in particular with Galileo, GPS and GLONASS. This represents already a better conditions for the beneficial application of robust estimators in the sense that the number of available measurements is highly increased with respect to the GPS standalone case. A second aspect that is considered is the impact of the design parameters, in particular the pseudorange variance, on the estimator performance. The performance evaluation of two different error models approaches is analysed in order to provide further insight on the behavior of these estimators with respect to least squares estimator and to give more formal statements about the expected performance of them. The different analyses are supported with real measurements that were collected with our test car in different places around DLR site in Oberpfaffenhofen and in Munich. The performance analyses focus not only on the overall Root Mean Square (RMS) values achieved, but also on the maximum position error experienced over the different scenarios. The latter provides a sense of the error bounding capability of the robust estimators.

II. ROBUST ESTIMATION

The classical algorithm used for positioning is based on a Least Squares (LS) estimator that can be expressed as the following minimization problem:

$$\hat{\mathbf{x}}_{\text{LS}} = \arg \min_{\mathbf{x}} \sum_i \hat{r}_i^2(\mathbf{x}), \quad (1)$$

where \hat{r}_i are the normalized residuals of the i th observation over the estimated parameters \mathbf{x} . This estimator is widely used due to its powerful properties under linear models and Gaussian distributed errors. However, it presents no robustness against outliers or model mismatches. In practice, any unbounded error source can introduce an unbounded error in the position estimation. A class of estimators that generalized the minimum of the sum of residual squares as specified in Equation (1) was proposed in [11]. The so called *M-estimators* consider a more general cost function ρ that is also dependent on the residuals:

$$\hat{\mathbf{x}}_{\text{M}} = \arg \min_{\mathbf{x}} \sum_i \rho \left(\frac{r_i(\mathbf{x})}{\sigma} \right), \quad (2)$$

where $\hat{\mathbf{x}}$ are the parameters to estimate, r_i stands for the residual of the i^{th} observation and σ is considered the scale factor of the data. Considering the derivative $\psi(x) = \frac{\partial \rho(x)}{\partial x}$ of the cost function, also denoted as *influence function*, the optimization problem from Equation (2) can be solved as

$$\sum_i \psi \left(\frac{r_i}{\sigma} \right) \frac{\partial r_i}{\partial \mathbf{x}} = 0. \quad (3)$$

If the influence function is monotonically increasing, the minimization problem is convex and an unique solution exists [7]. A *weighing function* can be also defined as $w(x) = \psi(x)/x$ in order to substitute the influence function as:

$$\sum_i w(\bar{r}_i) \bar{r}_i \frac{\partial \bar{r}_i}{\partial \mathbf{x}} = 0 \quad (4)$$

which at the same time can be used to express the initial minimization problem as:

$$\hat{\mathbf{x}}_M = \arg \min_{\hat{\mathbf{x}}} \sum_i w(\bar{r}_i) \bar{r}_i^2, \quad (5)$$

where \bar{r} are the normalized residuals according to the scale factor σ . Notice that the minimization problem can be now written as a sum of weighted squared residuals. Following this structure, a suitable solver for this estimators is the Iterative Reweighted Least Squares (IRLS).

One problem with M-estimator within the context of robust regression is the need of the scale σ of the data. To overcome the possible lack of knowledge about the spread of the data, other powerful estimators like the MM-estimator or the S-estimators have been proposed that estimate jointly the scale and location [12]. Despite their better performance in robust regression, their application to GNSS positioning at pseudorange level might have some conceptual difficulties. For the robust estimation of scale, it is assumed in general that the data belongs to the same distribution. On the contrary, the error in every pseudorange from different satellites at a given epoch is normally considered as independent, and the computation of a spread (i.e., scale) among the different pseudoranges does not have an easy interpretation. Nevertheless, and in order to adapt the M-estimator to the GNSS positioning, we will consider, as it is also done with (weighted) least squares that each pseudorange will have a different scale (i.e., variance), that can here be considered as an apriori uncertainty information:

$$\hat{\mathbf{x}}_M = \arg \min_{\hat{\mathbf{x}}} \sum_i w\left(\frac{r_i}{\sigma_i}\right) \bar{r}_i^2. \quad (6)$$

For the previous reasons, in this paper we focus on M-estimators of *location* only and we will investigate the impact of the pseudorange error model into the performance of the estimator.

From the different robust cost functions that can be found in the literature, in this work, we choose two cost functions: the Huber function [11], which is a hybrid between the l1-norm and Least-Squares cost function, and the Pseudo-Huber cost function, which is very similar to the Huber but is continuously differentiable. In Table I, the reader can find the expression of the cost, influence and weight functions for the Least-Squares and the two Huber estimators.

	Cost	Influence	Weighting
Least-Squares	$\rho(r) = \frac{1}{2}r^2$	$\psi(r) = r$	$w(r) = 1$
Huber	$\rho(r) = \begin{cases} \frac{1}{2}r^2 & \text{for } r \leq k \\ k r - \frac{1}{2}k^2 & \text{for } r > k \end{cases}$	$\psi(r) = \begin{cases} r & \text{for } r \leq k \\ k & \text{for } r > k \end{cases}$	$w(r) = \begin{cases} 1 & \text{for } r \leq k \\ \frac{k}{ r } & \text{for } r > k \end{cases}$
Pseudo-Huber	$\rho(r) = \frac{1}{2} \left(\sqrt{1 + (r/k)^2} - 1 \right)$	$\psi(r) = \frac{r}{\sqrt{1 + (r/k)^2}}$	$w(r) = \frac{1}{\sqrt{1 + (r/k)^2}}$

TABLE I: Cost, Influence and Weighting functions for Least-Squares, Huber and Pseudo-Huber estimators

Huber depends on the design parameters k , which has been derived in the robust regression literature to be $k = 1.2107$ in order to obtain an asymptotic efficiency of 95% on the standard normal distribution with respect to Least-Squares [13].

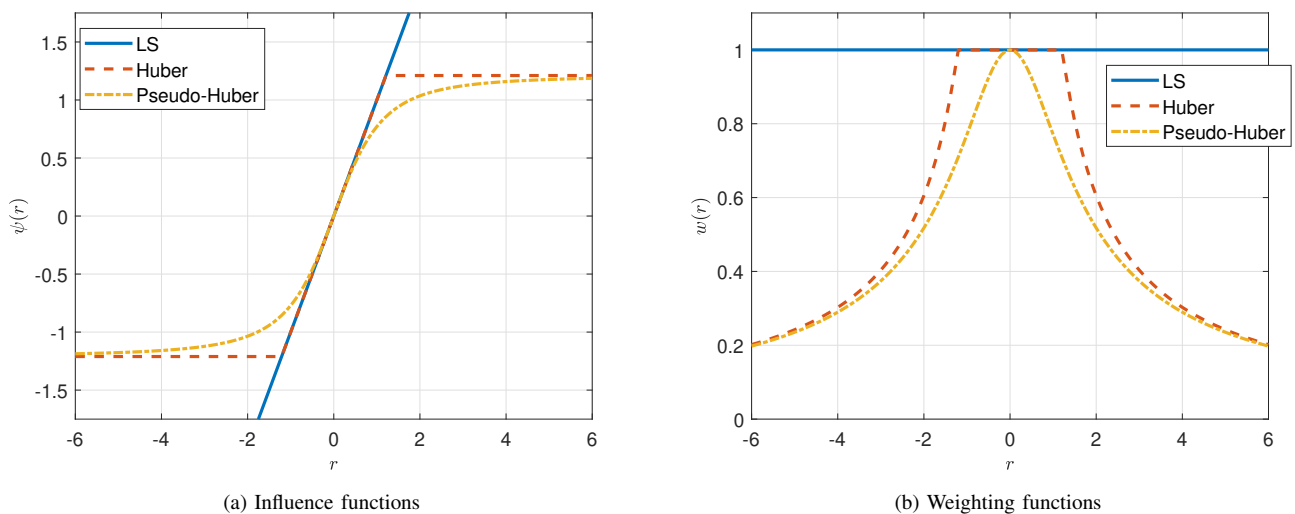


Fig. 1: Comparison of the influence and weighting functions of the LS (blue), Huber (red) and Pseudo-Huber (yellow).

The influence function and the weight function for the presented estimators are depicted in Figure 1. In here, we can clearly see how the Least-Squares presents an unbounded influence for an unbounded residual error. In contrast, the Huber versions limit the influence of a very large residual. A similar interpretation can be extracted from the weights plot: the Least-Squares is giving the weight of one to any measurement independently of the residual error that it poses, while the Huber and Pseudo-Huber reduces the weight for large residuals. An interesting aspect that will make easier to interpret the results of Huber estimator is the fact that for small normalized residuals it shares the same weight as Least-Squares.

III. PSEUDORANGE ERROR MODELS

The code pseudorange measurements provided by a GNSS receiver can be expressed at a certain time epoch k as:

$$P_k^{i,j} = R_k^i + b_k^j - c\delta t_k^i + I_k^i + T_k^i + M_k^i + \eta_k^{i,j}. \quad (7)$$

In here, $P_k^{i,j}$ is the code measurement for the i th satellite of constellation j , R is the true geometric distance between satellite and receiver, b_k^j is the user receiver clock bias with respect to the constellation j , $c\delta t_k^i$ is the satellite clock bias, c denotes the speed of light, I^i and T^i are the ionospheric and tropospheric propagation delays of the i th satellite, respectively. M^i is the error due to multipath and η^i is the receiver noise.

For the standard single frequency point positioning, there are available models to correct for most of the gross errors present in Equation 7.

The correction for the satellite clock is indirectly given by the broadcast navigation data. The ionosphere and troposphere propagation delays can be corrected with Klobuchar and RTCA standard models. Therefore, the linearized and single frequency corrected pseudorange can be expressed as

$$P_k^{i,j} - \tilde{P}_k^{i,j}(\mathbf{x}_k) = \mathbf{h}_k^{i,j} \Delta \mathbf{x}_k + \Delta S_k^{i,j} + \Delta I_k^{i,j} + \Delta T_k^{i,j} + mp_k^{i,j} + \eta_k^{i,j}, \quad (8)$$

where \mathbf{h} contains the line-of-sight vectors to each satellite along with the projection of clock biases to the pseudorange and $\Delta \mathbf{x}$ represents the deviation of the parameters to the linearization point. The remaining uncorrected errors are those due to imperfect ephemeris and satellite clock corrections $\Delta S_k^{i,j}$, the imperfect ionospheric $\Delta I_k^{i,j}$ and tropospheric models $\Delta T_k^{i,j}$, multipath and receiver noise.

In this paper, and in order to understand the importance of the assumed uncertainty associated with the measurements in the estimator, two error model approaches that follow different reasoning are compared. In the first one, we model all the remaining pseudorange error following a similar concept as used in (A)RAIM algorithms. That is, it is important to protect at the measurement level the possible remaining errors present in the pseudorange measurement so that the least-squares is able to provide realistic expected accuracy. Therefore, all the remaining errors ϵ can be assumed to followed a zero mean Gaussian distribution:

$$\text{Model 1: } \epsilon_k^{i,j} = \mathcal{N}\left(0, \underbrace{\sigma_{\text{URA}}^{i,j^2} + \sigma_{\text{iono}}^{i,j^2} + \sigma_{\text{tropo}}^{i,j^2}}_{\sigma_{\text{sis}}^2} + \underbrace{\sigma_{\text{mp}}^{i,j^2} + \sigma_{\eta}^{i,j^2}}_{\sigma_{\text{car}}^2}\right). \quad (9)$$

In here, the system level errors are overbounded by the σ_{URA} also provided by the navigation data. The remaining atmospheric errors can be characterised by the following variances. The residual Klobuchar ionosphere error can be overbounded by [14]:

$$\sigma_{\text{Uire},i} = F(\theta_i) \tau_{\text{vert}}(\hat{\mathbf{x}}), \quad \forall i = 1, \dots, N,$$

and the residual RTCA troposphere error can be bounded with a Gaussian distribution with zero mean and standard deviation [15]:

$$\sigma_{\text{tropo},i} = 0.12 \frac{1.001}{\sqrt{0.002001 + \sin(\theta_i)^2}}, \quad \forall i = 1, \dots, N.$$

Both remaining atmospheric errors depend on the satellite elevation angles θ_i , whereas the remaining ionospheric error also depends on the user position estimate $\hat{\mathbf{x}}$.

The multipath and noise errors are modelled for the specific antenna installation and receiver. Here we will denote it as σ_{car} . We characterize this error by an empirical estimation procedure using large set of pseudorange measurements collected in a static scenario. This estimation procedure strongly depends on the vehicle installation of the equipment, and hence, can differ strongly from vehicle to vehicle. The detailed estimation procedure can be found in Appendix A.

Notice that the remaining system and atmospheric errors after the single frequency corrections have a high level of correlation over time, so they could be also interpreted as a bias for a short time window. The stochastic modeling of them by the first model presented accounts for their possible variability over very long period.

That is why, in the second error model for the pseudorange considered in this paper we will account stochastically only for the so call σ_{car} , that is, only for the expected uncertainty of local multipath and receiver noise due to the specific installation of the antenna. Leaving the remaining system and atmospheric errors as biases that the estimator will have to deal with:

$$\text{Model 2: } \epsilon_k^{i,j} = \mathcal{N}(\Delta S_k^{i,j} + \Delta I_k^{i,j} + \Delta T_k^{i,j}, \sigma_{\text{mp}}^{i,j^2} + \sigma_{\eta}^{i,j^2}). \quad (10)$$

Please note that no explicit mention is provided to the additional multipath that the receiver may experience due to buildings or other environment. To the best of the authors knowledge there is no well established model that can account for the unpredictability of this error. Therefore in this paper, we will analyse the performance of the estimators under the presence of these additional effects without adding any particular uncertainty to the code measurement.

IV. MULTI-CONSTELLATION GNSS POSITIONING WITH ROBUST ESTIMATORS

In case that multiple GNSS constellations are used, the receiver needs to solve for the 3D position unknowns as well as for receiver clock offsets to each of the GNSS constellations. In our case, we exploit the code measurements of three different GNSS constellations: Galileo, GPS and Glonass, hence we need to solve for six unknowns:

$$\mathbf{x} = [x \quad y \quad z \quad b_u^E \quad b_u^G \quad b_u^R]^T, \quad (11)$$

where x, y, z are the position coordinates and b_u^E, b_u^G, b_u^R are the receiver clock bias with respect to the three constellations respectively.

The pseudorange Equation (7) is nonlinear with respect to R and a linearization step is normally necessary to solve the system of equations. The linearized measurement model can be described as:

$$\mathbf{z} = \mathbf{H}\Delta\mathbf{x} + \boldsymbol{\epsilon}, \quad (12)$$

where \mathbf{z} is the vector for all satellites in view, containing the difference between received and predicted pseudoranges based on the linearization point and expected receiver clock offset at a given time epoch according to Equation (8), \mathbf{H} is the geometry matrix that relates the measurements to the parameters to be estimated, $\Delta\mathbf{x}$ is the increments of the vector of unknowns and $\boldsymbol{\epsilon}$ contains all the errors which are not corrected. The geometric matrix can be expressed as:

$$\mathbf{H} = \begin{bmatrix} \mathbf{e}_u^{1, \text{Galileo}} & 1 & 0 & 0 \\ \mathbf{e}_u^{2, \text{Galileo}} & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{e}_u^{n, \text{Galileo}} & 1 & 0 & 0 \\ \mathbf{e}_u^{1, \text{GPS}} & 0 & 1 & 0 \\ \mathbf{e}_u^{2, \text{GPS}} & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{e}_u^{m, \text{GPS}} & 0 & 1 & 0 \\ \mathbf{e}_u^{1, \text{Glonass}} & 0 & 0 & 1 \\ \mathbf{e}_u^{2, \text{Glonass}} & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{e}_u^{k, \text{Glonass}} & 0 & 0 & 1 \end{bmatrix}, \quad (13)$$

where $\mathbf{e}_u^{i,j}$ is the unit line of sight vector pointing from the user to satellite i of constellation j . Each constellation needs special treatment in order to compute their satellites positions and parameters. More information about the necessary steps to solve for this can be found in the literature (for instance [16]).

The linearized measurement equation can be solved via least-squares as:

$$\Delta\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T \cdot \mathbf{z}. \quad (14)$$

The vector of corrected measurements follows a non standard Gaussian distribution with a diagonal covariance Σ described by the pseudorange error model uncertainties. We can use this apriori knowledge to weight the different measurements in a weighted least-squares estimator (WLS):

$$\Delta\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W} \cdot \mathbf{z}, \quad (15)$$

where typically the WLS estimator is designed such that $\mathbf{W} = \Sigma^{-1}$.

Now, as we have mentioned in the previous section, in order to apply the robust estimators over the same measurement equation, one solution consist on the implementation of an Iterative Reweighting Least Squares (IRLS). Basically, we build a weighted LS estimator that adjust at the same time the specific weights given to each measurement iteratively according to the robust functions given in Table I:

$$\Delta\hat{\mathbf{x}}^i = (\mathbf{H}^T\mathbf{W}^i\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}^i \cdot \mathbf{z}, \quad (16)$$

the index i is used to indicate here the algorithm iteration. The diagonal matrix \mathbf{W}^i is computed at each iteration based on the normalized residuals:

$$\mathbf{W}^i = f(\Sigma^{-1/2}(\mathbf{z} - \mathbf{H}\Delta\hat{\mathbf{x}}^{i-1})). \quad (17)$$

The different steps for the implementation of the robust estimator is given in Algorithm 1.

Algorithm 1 Robust Estimator

```
1: Initial point  $\mathbf{x}_0$ 
2: while !convergence do
3:   Compute Satellite Positions and Clock
4:   Compute Atmospheric corrections
5:   Correct Pseudoranges
6:   Initialize  $\mathbf{W} = \mathbf{I}$ 
7:   while !convergence do
8:     Linearize and Prepare  $\mathbf{z}, \mathbf{H}$ 
9:     Solve WLS
10:    Compute residuals  $\mathbf{r} = \mathbf{z} - \mathbf{H}\Delta\hat{\mathbf{x}}$ 
11:    Update  $\mathbf{W}$  with  $w_i = f_{\text{robust}}\left(\frac{r_i}{\sigma_i}\right)$ 
12:   end while
13:   Update  $\mathbf{x}_0$ 
14: end while
```

V. MEASUREMENT CAMPAIGN

In this section, we describe the measurement setup and scenarios that were used to analyse and validate the presented algorithm with real GNSS data. In total, we have collected the 19 hours of static data obtained in an open-sky scenario and several dynamic drives in urban scenario in Munich, Germany.

A. Experimental Setup

The measurement campaign was conducted between the 24th and the 25th of July 2019 by means of the DLR's *Safety-of-life* (*Sol*) testing vehicle. The antenna was installed on the roof of the car as shown in Figure 2. In particular, the front antenna is a Multi-frequency-multi-constellation (MFMC) *navXperience 3G+C* antenna connected to a *JAVAD Delta* receiver. The reference position was computed with the Post-Processed Kinematic service CSRS-PPP provided by the Canadian Geodetic Survey of natural Resources. The solution provides centimeter level accuracy in open sky and showed good availability and decimeter level accuracy in the urban environment.



Fig. 2: Antenna setup on car's roof.



Fig. 3: Car placement in the open-sky static scenario.

B. Open sky scenario

The DLR car was parked at a location where we were able to ensure good open-sky conditions during the whole data collection. This is also depicted in Figure 3. The collected data was used for two different purposes: the whole data-set (18 hours, 5Hz) were processed to model the on-board local multipath and noise component of the equipment setup (i.e., σ_{car}). Afterwards, one hour of data was selected and processed to evaluate the performance of the estimators in the open sky scenario.

C. Urban scenario

In order to test the algorithm in challenging GNSS scenarios, we recorded data during a car ride in Munich, Germany. The specific scenario shown in Figure 4 consisted of different challenging environments for GNSS, including a residential area,

an urban canyon and the presence of very tall buildings in some parts. The trajectory, which was driven twice, is represented in yellow in Figure 4. The starting point is represented by a blue star, whereas the end is depicted by an orange star. The different relevant areas are highlighted by red rectangles.

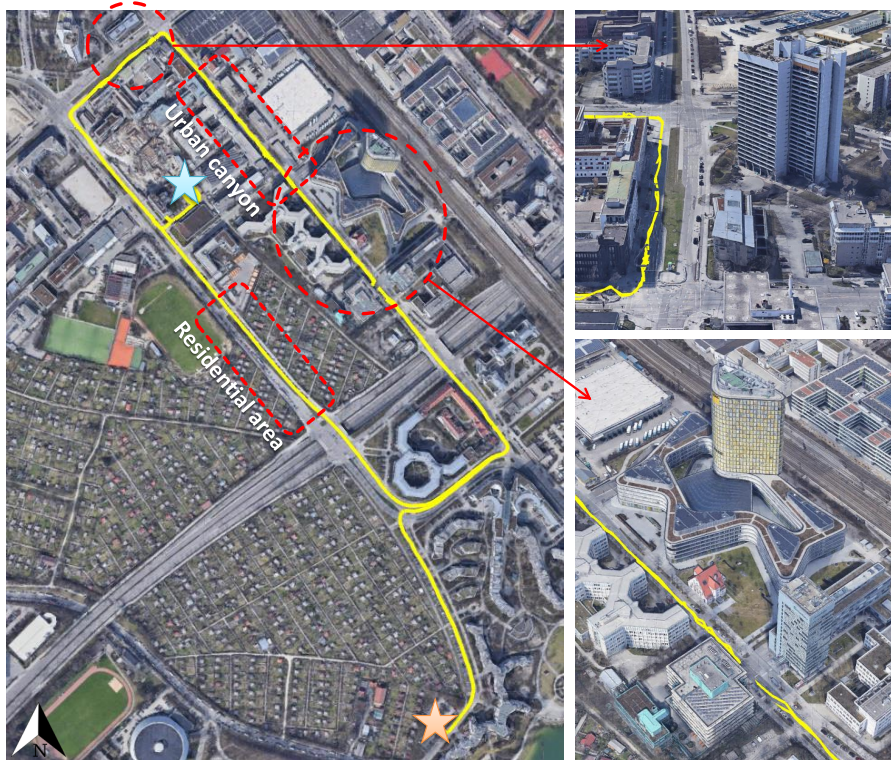


Fig. 4: Challenging urban scenario in Munich, Germany: reference trajectory is represented by a yellow line, starting point by a blue star, end point by an orange star, areas of particular interest are highlighted by red circles and rectangles (extracted from *GoogleEarthPro*).

VI. RESULTS

A. Multipath and Noise Error Modeling

In total, we collected MFMC GNSS code and carrier-phase measurements with $f_{\text{data}} = 5$ Hz data rate over 18 h. The nominal multipath due to the vehicle and the receiver noise is modeled based on the satellite data sequences that are tracked at least 60 minutes in order to guarantee a proper removal of the ambiguity component as explained in Section III. The presence of cycle slips was also monitored. After clustering our estimates with respect to the satellite elevation angle, we analyzed the empirical distribution of the multipath and noise error estimate for each cluster.

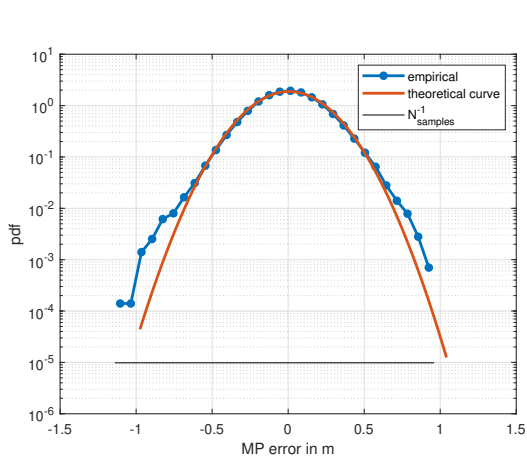
In Figure 5a, you can see exemplarily the empirically two sided probability density function in a logarithmic scale for elevations cluster 50 – 60 deg. Parallel in Figure 5b, the associated Q-Q (quantile-quantile) plot with respect to the normal distribution having the observed mean and variance for the same elevation bin is presented.

We can see how for certain elevation angles the distribution does not perfectly follow a Gaussian distribution and present some heavy tails. The non-Gaussianity of the error models is one of the weakness of Least-Squares estimators that can be better handled by robust regression.

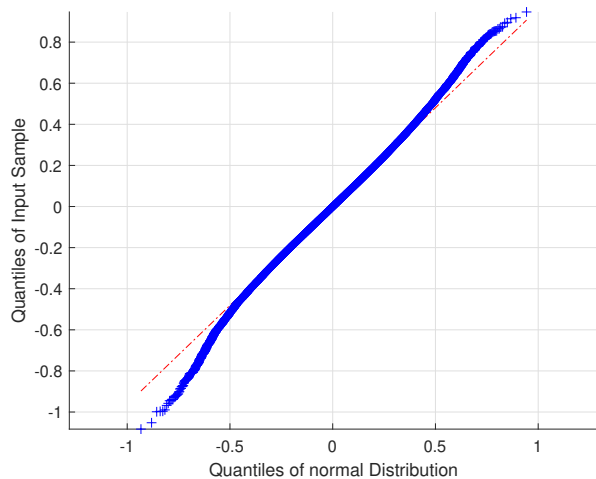
The obtained standard deviation with respect to the satellite elevation angle is presented in Figure 6. Please note that we used an unequal elevation bin sizes, i.e., for low elevations $5 \text{ deg} \leq \theta_i \leq 30 \text{ deg}$ we used a bin size of one degree, for $30 \text{ deg} < \theta_i \leq 50 \text{ deg}$ a bin size of two degree was applied and for higher elevations $\theta_i > 50 \text{ deg}$ the bin size was increased to five degrees. As we can see the extracted standard deviation is declining with increasing satellite elevation angle. GPS and Galileo present a similar level, whereas Glonass shows the same trend but higher values for all elevation angles.

B. Performance in Static-Open Sky Scenario

As a first evaluation, we assess the performance of the robust estimators with Huber and Pseudo-Huber functions with respect to the Least-Squares in open-sky scenario. In Figure 7a, we can see the results of the error in horizontal for one



(a) Two-sided probability density function of the observed vehicle multipath and noise estimates and given satellite elevation el angles plotted in logarithmic scale.



(b) Quantile-Quantile-plot of the empirical distribution versus the theoretical Normal distribution having the empirical statistics $N(\mu_{car}(\theta), \sigma_{car}^2(\theta))$.

Fig. 5: Assessment of the statistical characteristics of the local vehicle multipath and noise for satellite elevation angles between 55 and 60 deg seen by our land-based receiver setup.

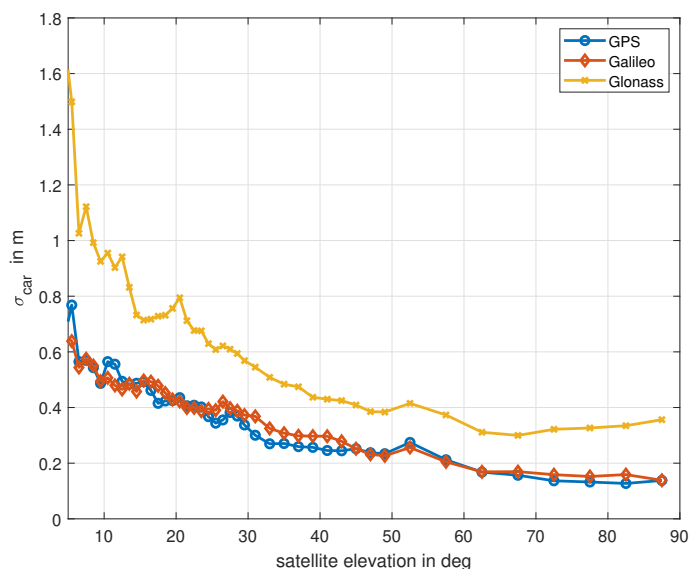
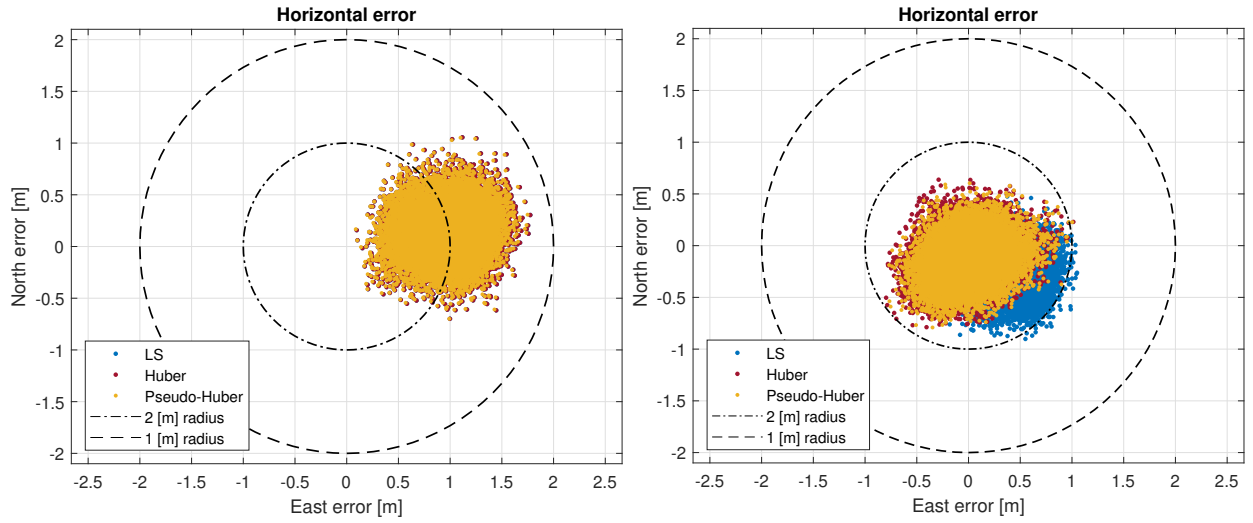


Fig. 6: Vehicle multipath and receiver noise standard deviation versus satellite elevation angle for GPS L1, Galileo E1 and Glonass L1.

hour of data and using the model 1 that overbounds the system and atmospheric errors. In Figure 7b, the horizontal error is depicted for the same data with with the Model 2 that considers only multipath and noise for the variance.

Under a large uncertainty model (Model 1), we can see how all the estimators present almost the same behaviour. In fact, LS and Huber have exactly the same error values while pseudo Huber differs only slightly. Using a large *a priori* uncertainty for the pseudoranges translates into a large acceptance of error until the robust estimators starts deweighting measurements. In the case of a more restricted model (Model 2), we can see in Figure 7b that the robust estimators are less biased as compared to LS. This improvement in the error may come from the remaining ionospheric, system level and tropospheric errors that are not accounted for in the correction models, but that the robust estimator is able to adjust by reducing the weight given to measurements that present a larger residual with respect to the overall consensus. The robust estimator is only able to become sensitive (to deweight) to these errors if the *a priori* uncertainty is close to the same order of magnitude.

The value for the 2D RMS error averaged over one hour and the horizontal maximum error is reported in Table II for the second model.



(a) Applying Error Model 1 ($\sigma_{\text{SIS}}^2 + \sigma_{\text{car}}^2$).

(b) Applying Error Model 2 (only σ_{car}^2).

Fig. 7: Scatter plot of the horizontal error for different estimators over one hour.

Estimator	Least Squares	Huber	Pseudo Huber
2D RMS error [m]	0.57	0.34	0.33
Maximum horizontal error [m]	1.16	1.04	1.03

TABLE II: Position performance comparison in the static scenario between three different estimators (Model 2).

C. Performance in Urban Scenario

We provide first a qualitative assessment of the performance of the different algorithms in the urban scenario selected. In Figure 8, we can see the overview of the vehicle trajectory as computed by the different algorithms.

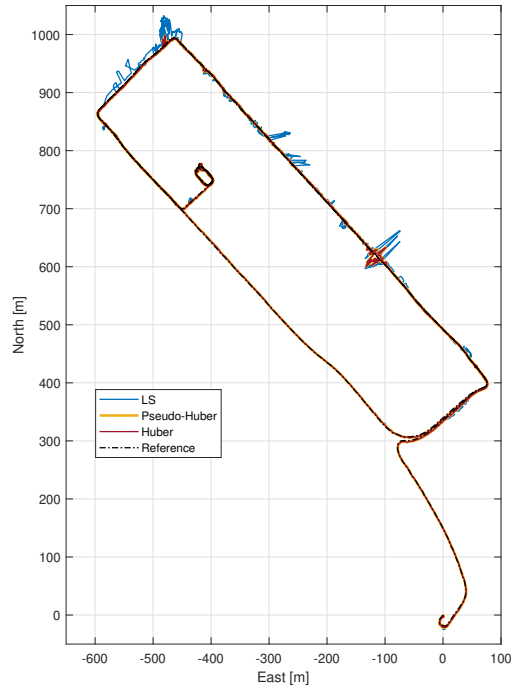


Fig. 8: Estimated trajectories in urban scenario.

For the evaluations of this section, we use the second error model that uses only the multipath and noise uncertainty because this has been found to provide the best performance. The numerical results with both models will be also presented and compared at the end.

Looking a bit closer to two of the areas where larger errors have been observed, we can see in Figure 9a how it seems that the discrepancies between the reference and the solution computed with the LS are significantly larger as compared to the robust estimators. Remarkably in Figure 9b we see that the robust estimators are able to provide better solution along an urban canyon, while the LS present continuous jumps, due probably to the presence of high multipath or NLOS signals.

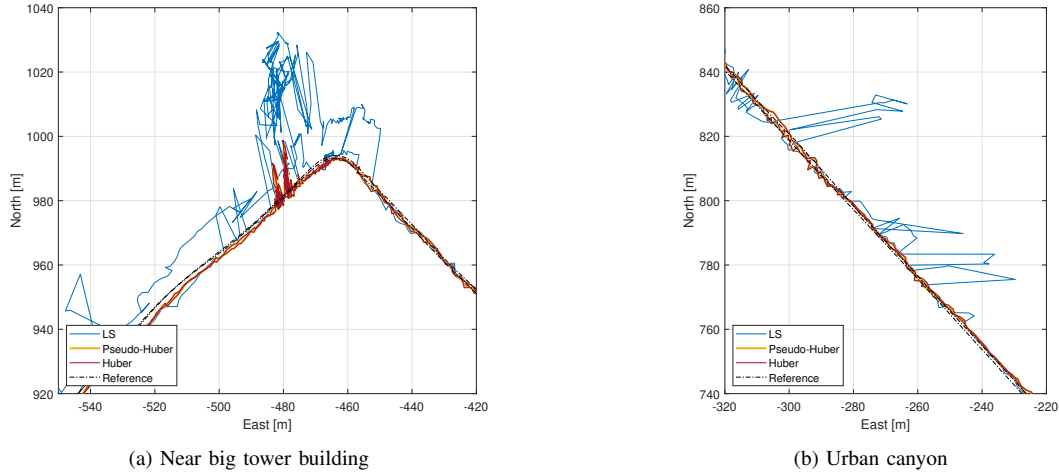


Fig. 9: Example of trajectories provided by different estimators

A more quantitative assessment is provided in Figure 10a where the horizontal error of the different estimators is shown over time. We clearly see how the robust estimators are more resilient to very large errors in certain moments as compared to the LS.

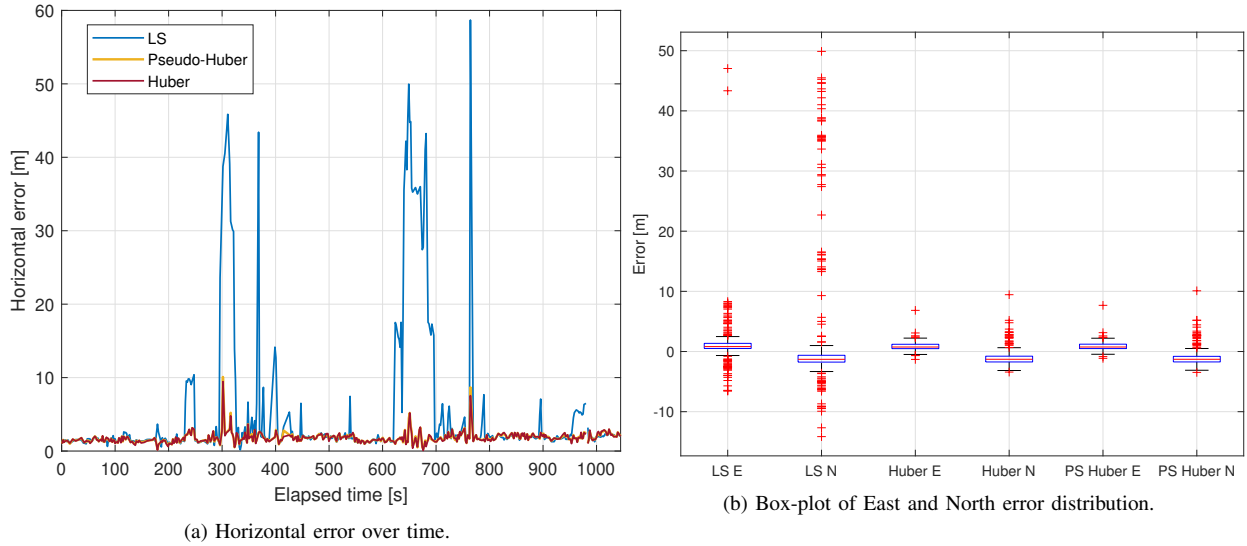


Fig. 10: Horizontal position error in urban scenario for different estimators.

Another way of seeing the dispersion of the east and north error is provided in the boxplot of Figure 10b. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. We can see how LS present very high errors whereas the robust one keeps the error bounded at the position level.

Table III resumes the horizontal Root-Mean-Square error (2D RMS) and the maximum horizontal error for the three estimators using the second model. In accord with the previous observation the higher values are observed for the LS approach. Indeed, the maximum error goes up to 58.69 m in the LS case compared to the 9.45 m of the Huber estimator and 10.09 m for the Pseudo-Huber. This important difference shows the improvement that is given by the robust estimators in computing the positioning solution for the analysed configuration. Huber and Pseudo-Huber present similar performance.

Estimator	Least Squares	Huber	Pseudo-Huber
2D RMS error [m]	11.32	1.95	1.98
Maximum horizontal error [m]	58.69	9.45	10.09

TABLE III: Position Performance comparison in the urban scenario between three different estimators

D. Summary of results

Table IV summarizes the main results of the different evaluations for the Horizontal RMS and Maximum Error. In parallel to the different scenarios, we provide the information when using the two different philosophies or models for the pseudorange stochastic errors.

Scenario	Pseudorange Error Model used	Least Squares	Huber	Pseudo Huber
Horizontal RMS Error [m]				
Open sky	Model 1 ($\sigma^2 = \sigma_{\text{sys}}^2 + \sigma_{\text{car}}^2$)	1.07	1.07	1.06
Open sky	Model 2 ($\sigma^2 = \sigma_{\text{car}}^2$)	0.57	0.34	0.33
Urban	Model 1 ($\sigma^2 = \sigma_{\text{sys}}^2 + \sigma_{\text{car}}^2$)	13.16	7.35	8.17
Urban	Model 2 ($\sigma^2 = \sigma_{\text{car}}^2$)	11.32	1.95	1.98
Maximum Horizontal Error [m]				
Open sky	Model 1 ($\sigma^2 = \sigma_{\text{sys}}^2 + \sigma_{\text{car}}^2$)	1.80	1.80	1.78
Open sky	Model 2 ($\sigma^2 = \sigma_{\text{car}}^2$)	1.16	1.04	1.03
Urban	Model 1 ($\sigma^2 = \sigma_{\text{sys}}^2 + \sigma_{\text{car}}^2$)	64.02	58.99	55.32
Urban	Model 2 ($\sigma^2 = \sigma_{\text{car}}^2$)	58.69	9.45	10.09

TABLE IV: Overview of the results obtained within the different measurement campaigns and error models

VII. DISCUSSIONS

In this work, we have evaluated the performance of least-squares and Huber M-estimators for the GNSS positioning in an open-sky scenario and in urban automotive scenarios. In particular, we focused the evaluation on the impact of the apriori uncertainty or scale that was used for each pseudorange measurement. From the GNSS perspective, each one of the two models provides a different interpretation of the error sources (stochastic vs biases) that can be justified by looking at the change of the error over different time periods.

With respect to the open sky scenario and by using the model with a larger variance, we have seen that the robust estimators provide practically the same performance as LS. When not accounting for the system and atmospheric errors in the apriori variance (Model 2), the robust estimators are able to deweight those measurement with a higher remaining error and therefore some improvement is achieved.

Please note that M-estimators are non-optimal estimators and under Gaussian and well modelled errors, they present theoretically a loss of efficiency with respect to LS (which is optimal). However, in real GNSS positioning, due to the different error sources, it is arguable when a perfect model match would occur. We have seen that even in open-sky scenarios, it appears that the remaining, not corrected errors or the presence of longer tail error distributions (as seen in Figure 5) may have larger impact than the possible loss of efficiency of the estimator. This is suggested by the results of Table IV in Open sky for both Model 1 and 2.

With respect to the urban scenario, we have also seen that when using the larger variance model, the robust estimators were only able to reduce the maximum error from 64.02 meters (LS) to 58.99 (Huber) or 55.32 (pseudo-Huber) meters. Whereas the smaller variance model reduced the maximum error from 58.69 meters (LS) to 9.45 (Huber) or 10.09 meters (pseudo-Huber). From this, it is clear that if the apriori variances are too large, the robust estimator may accept a large multipath or NLOS error as nominal behaviour and is less likely to deweight that measurement. In other words, if we account for the overbounded and large possible system and atmospheric errors within the apriori variance, the robust estimator is not able to identify and adjust the presence of multipath.

Huber and pseudo-Huber provided in general similar performances. Although pseudo-Huber presents a more aggressive cost function, it has shown slightly larger error values as compared to Huber. Although the difference appears to be not so significant, this could be further investigated in the future.

The errors obtained in this paper are computed based on a PPP-like reference position. This position showed up to 60 cm error accuracy along the urban scenario. Although a better reference position will be obtained in the future based on a precise GNSS+INS solution, the accuracy of the current one does not hide the main message and conclusions obtained in this paper.

When designing the multipath and noise model for Glonass, in this work we modeled the variance per elevation as a average model for all FDMA frequencies due to the limited number of available measurements. With the availability of larger static dataset, a better tuning of the model should be possible for each satellite (i.e., specific frequency).

Robust estimators present therefore very promising performance in diverse situations, in particular when multiple corrupted measurements are received. Further work is however still required to be able to assess rigorously their bounding capability and estimated variance at the position domain.

VIII. CONCLUSIONS

One of the main challenges to bring GNSS reliability standards from aviation to ground transportation is due to the presence of local threats that are scenario specific and therefore very difficult to model or predict. Therefore the classical approaches used in aviation that highly rely on models are difficult to apply in urban scenarios. The use of robust estimators are here presented as a promising alternative to classical least-squares estimators. These estimators may have more power to handle the presence of non Gaussian errors, mismodel of the errors or the highly corrupted GNSS measurements due to multipath, NLOS or interference. The design of these estimators, in particular the specific assumed pseudorange error model that is used to normalize the adjustment have a critical impact in the final error performance and robustness as it has been here presented. One alternative for the design that provided improvements in the maximum observed error in urban scenarios of more than 80% and kept the RMS in below two meters with only code measurements, has been presented. The results obtained suggest that this estimator have the potential to become a powerful option for challenging GNSS scenarios.

APPENDIX A

VEHICLE NOISE AND MULTIPATH MODELING

The estimation of the multipath and noise component due to the antenna installation within the vehicle consist of the following steps:

- 1) Collection of a large set of consecutive MFMC GNSS measurements
- 2) Calculation of the corresponding satellite elevation angles
- 3) Computing the multipath estimate for each observed satellite
- 4) Clustering of the estimates according to the satellite elevation angle
- 5) Computation of the statistical characteristics for each cluster

The vehicle depending multipath estimate for each satellite i and a certain GNSS frequency (e.g., L1) at a given epoch k can be isolated with the following technique:

$$M_k^i = \text{CMC}_k^i - 2\hat{I}_k^i - \hat{b}^i, \quad (18)$$

where $\text{CMC}_k^i = P_k^i - \Phi_k^i$ is the well-known Code-minus-Carrier (CMC) observable, the second term denotes the ionospheric delay estimate, that can be computed using the carrier phase measurements in two frequencies. For example, for the ionospheric estimation on L1 with the use of L2, we would have $\hat{I}_{L1,k}^i = \frac{f_{L2}^2}{f_{L1}^2 - f_{L2}^2} (\Phi_{L1,k}^i - \Phi_{L2,k}^i)$. The last term \hat{b}^i represents mainly the linear combination of the integer ambiguity of the carrier-phase measurements. It can be computed by taking the average of consecutive ionosphere corrected CMC measurements,

$$\hat{b}^i = \frac{1}{N} \sum_{k=1}^N (\text{CMC}_k^i - 2\hat{I}_k^i).$$

Please note that in order to have a good estimate of the average value, a large number of consecutive data without cycle slips is necessary.

REFERENCES

- [1] M. Joerger, F.-C. Chan, and B. Pervan, "Solution Separation versus Residual-based RAIM," *Navigation*, vol. 61, no. 4, pp. 273–291, 2014.
- [2] Working Group C - ARAIM Subgroup, "Milestone 3 report," EU/US Cooperation on Satellite Navigation, Tech. Rep., 2016.
- [3] A. Grosch, O. García Crespillo, I. Martini, and C. Günther, "Snapshot Residual and Kalman Filter based Fault Detection and Exclusion Schemes for Robust Railway Navigation," in *2017 European Navigation Conference (ENC)*, May 2017, pp. 36–47.
- [4] N. Zhu, J. Marais, D. Bétaille, and M. Berbineau, "Gnss position integrity in urban environments: A review of literature," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 9, pp. 2762–2778, Sep. 2018.
- [5] P. Huber, *Robust Statistics*. Wiley, N. Y., 1981.
- [6] N. L. Knight and J. Wang, "A Comparison of Outlier Detection Procedures and Robust Estimation Methods in GPS Positioning," *Journal of Navigation*, vol. 62, no. 4, pp. 699–709, Oct. 2009.
- [7] H. Pesonen, "Robust estimation techniques for GNSS positioning," in *Proceedings of NAV07-The Navigation Conference and Exhibition, 31.10.-1.11. 2007, London, England, 2007*.

- [8] S. Gaglione, A. Innac, S. P. Carbone, S. Troisi, and A. Angrisano, "Robust estimation methods applied to GPS in harsh environments," in *2017 European Navigation Conference (ENC)*, May 2017, pp. 14–25.
- [9] O. Garcia Crespillo, D. Medina, J. Skaloud, and M. Meurer, "Tightly Coupled GNSS/INS Integration based on Robust M-estimators," in *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. Monterey, CA: IEEE, Apr. 2018, pp. 1554–1561.
- [10] R. M. Watson and J. N. Gross, "Robust Navigation In GNSS Degraded Environment Using Graph Optimization," *arXiv:1806.08899 [cs]*, Jun. 2018, arXiv: 1806.08899.
- [11] P. J. Huber, "Robust regression: asymptotics, conjectures and monte carlo," *The Annals of Statistics*, pp. 799–821, 1973.
- [12] J. Fox, "Robust regression," *An R and S-Plus companion to applied regression*, 2002.
- [13] R. Maronna, R. D. Martin, and V. Yohai, *Robust statistics*. John Wiley & Sons, Chichester. ISBN, 2006, vol. 1.
- [14] D. Salós, C. Macabiau, A. Martineau, B. Bonhoure, and D. Kubrak, "Nominal GNSS pseudorange measurement model for vehicular urban applications," in *IEEE/ION Position, Location and Navigation Symposium*, May 2010, pp. 806–815.
- [15] RTCA/SC-159, "RTCA/DO-229C: Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment," RTCA, Tech. Rep., 2001.
- [16] J. S. Subirana, J. M. J. Zornoza, and M. Hernández-Pajares, *GNSS DATA PROCESSING*. ESA Communications, May 2013, vol. Volume 1: Fundamentals and Algorithms.