POLITECNICO DI TORINO
Repository ISTITUZIONALE

A Post-Processing Multipath/NLoS Bias Estimation Method Based on DBSCAN

Original
A Post-Processing Multipath/NLoS Bias Estimation Method Based on DBSCAN / Guo, Y., Zocca, S., Dabove, P., Dovis, F.. - In: SENSORS. - ISSN 1424-8220. - ELETTRONICO. - 24:8(2024). [10.3390/s24082611]

Availability:
This version is available at: 11583/2987930 since: 2024-04-19T13:19:58Z

Publisher:
MDPI

Published
DOI:10.3390/s24082611

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026

sensors

MDPI

Article

A Post-Processing Multipath/NLoS Bias Estimation Method Based on DBSCAN

Yihan Guo [1,*], Simone Zocca [1], Paolo Dabove [2] and Fabio Dovis [1]

[1] Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy; simone.zocca@polito.it (S.Z.); fabio.dovis@polito.it (F.D.)
[2] Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, 10129 Turin, Italy; paolo.dabove@polito.it
* Correspondence: yihan.guo@polito.it

Abstract: Positioning based on Global Navigation Satellite Systems (GNSSs) in urban environments always suffers from multipath and Non-Line-of-Sight (NLoS) effects. In such conditions, the GNSS pseudorange measurements can be affected by biases disrupting the GNSS-based applications. Many efforts have been devoted to detecting and mitigating the effects of multipath/NLoS, but the identification and classification of such events are still challenging. This research proposes a method for the post-processing estimation of pseudorange biases resulting from multipath/NLoS effects. Providing estimated pseudorange biases due to multipath/NLoS effects serves two main purposes. Firstly, machine learning-based techniques can leverage accurately estimated pseudorange biases as training data to detect and mitigate multipath/NLoS effects. Secondly, these accurately estimated pseudorange biases can serve as a benchmark for evaluating the effectiveness of the methods proposed to detect multipath/NLoS effects. The estimation is achieved by extracting the multipath/NLoS biases from pseudoranges using a clustering algorithm named Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The performance is demonstrated using two real-world data collections in multipath/NLoS scenarios for both static and dynamic conditions. Since there is no ground truth for the pseudorange biases due to the multipath/NLoS scenarios, the proposed method is validated based on the positioning performance. Positioning solutions are computed by subtracting the estimated biases from the raw pseudoranges and comparing them to the ground truth.

Keywords: multipath; non-line-of-sight; pseudorange bias; clustering algorithm

check for updates

Citation: Guo, Y.; Zocca, S.; Dabove, P.; Dovis, F. A Post-Processing Multipath/NLoS Bias Estimation Method Based on DBSCAN. *Sensors* **2024**, *24*, 2611. https://doi.org/10.3390/s24082611

Academic Editor: Wu Chen

Received: 19 March 2024
Revised: 15 April 2024
Accepted: 16 April 2024
Published: 19 April 2024

Copyright: © 2024 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Multipath and Non-Line-of-Sight (NLoS) aspects are the most detrimental effects for Global Navigation Satellite Systems (GNSSs) in urban scenarios, affecting the large plethora of emerging applications being developed and expected to work in such an environment, such as autonomous vehicles, smart wearables, etc. [1]. Multipath interference, as the name implies, happens when a signal is received through multiple paths. This can include the direct path and one or more reflected paths, or it may involve multiple reflected paths. NLoS reception occurs when the direct path from the transmitter to the receiver is obstructed and the signals are received only through a reflected path [2].

The presence of reflected GNSS signals in the case of multipath/NLoS effects causes biases ranging from several to thousands of meters in pseudorange measurements [3]. Unlike the tropospheric and ionospheric delay contribution and relativistic effect, whose physical models allow for proper correction, mitigating the impact of multipath/NLoS effects on pseudorange measurements is challenging due to their complex physical models.

As multipath/NLoS effects pose critical threats to the widespread use of GNSS, numerous efforts have been devoted to detecting and mitigating the pseudorange biases caused by such propagation conditions. Most of the research efforts can be categorized

into three groups based on the specific stage of the user receiver at which detection and/or mitigation are applied:

The first group of solutions includes antenna-based techniques. In scenarios involving multipath/NLoS effects, the reflected GNSS signals may undergo changes in polarization and reception angles compared to the direct signals. To mitigate this issue, researchers developed advanced antenna designs, including choke rings, dual-polarized antennas, and array antennas, to minimize the reception of reflected signals [2,4,5].

The second family of multipath/NLoS detection and mitigation techniques relies on signal processing. Multipath occurs when the signals reflected from surfaces reach the receiver along with the direct signals. In GNSS baseband signal processing, the shift in auto-correlation functions caused by reflected signals can overlap with the auto-correlation function of the direct signal. As a result, the computed auto-correlation function in the Delay-Lock Loop (DLL) will be distorted due to this overlapping. To address this, various advanced signal tracking techniques have been proposed for identifying and mitigating multipath interference [6–9].

The third group of techniques operates within the positioning unit to detect and mitigate multipath/NLoS interference. Pseudorange biases induced by multipath/NLoS effects pose a threat to the consistency between the pseudorange measurements and the corresponding navigation solutions. Some approaches utilize the chi-squared test for detecting and mitigating the impact of multipath/NLoS effects [10]. Additionally, 3D city models and cameras are employed to establish the propagation model of reflected signals [11,12]. Furthermore, researchers explore the potential of machine learning algorithms to establish relationships between multipath/NLoS effects and other indicators, such as satellite elevations/azimuths and the carrier-to-noise-density ratio (C/N_0) [13–15].

In an effort to contribute to the development of methodologies mitigating multipath/NLoS effects, this research proposes a post-processing solution for estimating the pseudorange biases caused by these phenomena. In recent years, new research approaches have considered machine learning-based algorithms for detecting and mitigating multipath/NLoS effects. By extracting certain features in the positioning unit, ref. [11] proposed a classifier based on supervised machine learning to categorize GNSS pseudorange measurements into three types: clean, multipath, and NLoS. Moreover, ref. [16] adopted an unsupervised method to identify satellite signals with multipath using carrier phase, pseudorange, and carrier-to-noise ratio measurements. Additionally, ref. [17] formulated multipath modeling as a regression task, fitting the multipath errors with respect to the azimuth and elevation in the spatial domain. To implement multipath detection and mitigation, these studies utilized various features derived from GNSS raw data as well as additional information. Accurate values of pseudorange biases due to multipath/NLoS effects are pieces of information essential for these machine learning approaches to train the model, and then they are used to detect and mitigate multipath/NLoS effects. Therefore, the primary objective of this research is to provide high-precision pseudorange bias values caused by multipath/NLoS to empower machine learning-based algorithms. Furthermore, these accurately estimated pseudorange biases can serve as a standard for assessing the performance of the methods proposed for detecting and mitigating multipath/NLoS effects.

To compute the accurate pseudorange biases caused by multipath/NLoS effects, this research leverages precise corrections and models for GNSS and employs a clustering algorithm, namely Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The main novel contributions of this work can be summarized as follows:

1. Building upon the theoretical analysis of the pseudorange function, this research defines a leftover term containing the pseudorange biases caused by multipath/NLoS effects.
2. This research derives two probability distributions for the defined leftover term, which motivates the utilization of a clustering algorithm to estimate multipath/NLoS effects.
3. By utilizing a clustering algorithm, specifically DBSCAN, to isolate the other components in the defined leftover term, a procedure is proposed to estimate the values of multipath/NLoS biases if multipath/NLoS effects are present.

The paper is organized as follows: in Section 2, the mathematical foundations for the pseudorange function are established, and the definition of the leftover term is provided. Section 3 introduces the proposed method for extracting pseudorange biases caused by multipath/NLoS effects from the leftover term using a clustering algorithm. Subsequently, Section 4 outlines the details of the real-world experiments under both static and dynamic scenarios to validate the effectiveness of the proposed method. Finally, Section 5 draws the conclusions regarding this work.

2. Definition and Computation of the Leftover Term

This section recalls the derivation of the pseudorange observation equation, as outlined in [18], in order to provide the definition of the leftover term containing the pseudorange biases due to multipath/NLoS effects, along with other receiver-related error components.

The GNSS receiver generates pseudorange measurements by multiplying the signal travel time from the satellite to the receiver with the speed of light. It is noted that the superscript s is used to represent the pseudorandom noise (PRN) code of a specific GNSS satellite, and the subscript r is used to represent the GNSS receiver. The pseudorange measurement function is written as

$$\begin{aligned} p_r^s(t) = c(d_r + dt_r(t)) + \|\boldsymbol{r}^s(t) + \boldsymbol{r}_r(t)\| \\ + c\left(d^s - dt^s(t) + \delta t_{\text{stc}}^{s,\text{rel}}(t) - \delta t_{\text{clk}}^{s,\text{rel}}(t)\right) \\ + \xi_r^s(t) + I^s(t) + T^s(t) + F^s(t) + e_r^s(t) \end{aligned} \quad (1)$$

where

- $p_r^s(t)$ is the pseudorange measurement obtained from the GNSS receiver for satellite s at time t;
- c is the speed of light in vacuum;
- d_r is the signal instrumental delay of the receiver;
- $dt_r(t)$ is the clock bias of the receiver at time t;
- $\boldsymbol{r}^s(t)$ is the mass center position of the satellite under Earth Centered Earth Fixed (ECEF) frames at time t;
- $\boldsymbol{r}_r(t)$ is the antenna reference point position of the receiver under ECEF frames at time t;
- d^s is the signal instrumental delay of the satellite;
- $dt^s(t)$ is clock bias of the satellite at time t;
- $\delta t_{\text{stc}}^{s,\text{rel}}(t)$ is the delay caused by space–time curvature of the relativistic effect at time t;
- $\delta t_{\text{clk}}^{s,\text{rel}}(t)$ is satellite clock bias caused by the relativistic effect at time t;
- $\xi_r^s(t)$ is antenna phase center corrections for both transmitting and receiving antennas at time t;
- $I^s(t)$ is the error contribution of the pseudorange measurement due to the ionospheric delay, expressed in meters at time t;
- $T^s(t)$ is the error contribution of the pseudorange measurement due to the tropospheric delay, expressed in meters at time t;
- $F^s(t)$ is the error contribution of the pseudorange measurement due to the multipath/NLoS interference at time t, expressed in meters;
- $e_r^s(t)$ is the error contribution of the pseudorange measurement due to the receiver noise at time t, expressed in meters.

The Sagnac effect caused by Earth's rotation should also be compensated for $p_r^s(t)$ according to [18] (Chapter 19.1). By adjusting both sides of (1), we can obtain

$$\begin{aligned} F^s(t) + e_r^s(t) + c(d_r + dt_r(t)) \\ = p_r^s(t) - \|\boldsymbol{r}^s(t) - \boldsymbol{r}_r(t)\| - \zeta_r^s(t) - I^s(t) - T^s(t) \\ - c\left(d^s - dt^s(t) + \delta t_{\text{stc}}^{s,\text{rel}}(t) - \delta t_{\text{clk}}^{s,\text{rel}}(t)\right) \end{aligned} \quad (2)$$

Thanks to the advancements in the development of physical models and corrections offered by International GNSS Service (IGS), all the terms on the right-hand side of (2) can be computed in a post-processing manner, with relatively high accuracy, as outlined in Table 1. The left-hand side of (2) is designated as the leftover term $L^s(t)$ of a pseudorange measurement:

$$\begin{aligned} L^s(t) &= F^s(t) + e_r^s(t) + c(d_r + dt_r(t)) \\ &= F^s(t) + e_r^s(t) + c \cdot dt_{\text{rcv}} \end{aligned} \quad (3)$$

Table 1. Corrections and models for pseudorange measurement.

Term	Computation Method
$p_r^s(t)$	Measurements from the receiver
$\boldsymbol{r}^s(t)$	IGS final precise orbits products
$\boldsymbol{r}_r(t)$	RTK or RTK/INS positioning solutions
$\zeta_r^s(t)$	Absolute IGS phase center corrections (igs14.atx)
$I^s(t)$	Final solution of IGS combined GIMs
$T^s(t)$	Saastamoinen model
d^s	TGD provided by navigation messages
$dt^s(t)$	Clock biases of the satellites from navigation messages
$\delta t_{\text{stc}}^{s,\text{rel}}(t)$	See Appendix A
$\delta t_{\text{clk}}^{s,\text{rel}}(t)$	See Appendix B

Since d_r and $dt_r(t)$ are produced by the GNSS receiver itself, d_r remains constant under a specific epoch t for different satellites. Furthermore, the same $dt_r(t)$ is also shared for every satellite under a certain epoch t. Therefore, these two terms can be combined as a single term, which is denoted as dt_{rcv}.

3. Multipath/NLoS Bias Estimation Using a Clustering Algorithm

This section will elaborate on the method proposed by this research to estimate the pseudorange biases caused by multipath/NLoS effects. First, two probability distributions of leftover terms are generated, distinguishing between conditions with and without multipath/NLoS effects. These distinct distributions motivate the use of clustering algorithms to classify the leftover terms affected by multipath/NLoS effects from those that are not. Subsequently, we provide a detailed introduction to a density-based clustering algorithm, namely DBSCAN, and explain its suitability for this task of identifying multipath/NLoS events. Finally, this section illustrates the proposed procedure of utilizing DBSCAN for multipath/NLoS estimation.

3.1. Statistical Characterization of the Leftover Terms

Analyzing $L^s(t)$ as in (3), different components with specific features can be recognized:

1. The error contribution of the pseudorange measurement due to the multipath/NLoS interference $F^s(t)$ is usually different from one satellite to another. Furthermore, $F^s(t)$ is zero if a satellite is free from both multipath and NLoS interference.
2. The receiver noise $e_r^s(t)$ is commonly characterized by a Gaussian distribution with a zero mean and constant variance σ^2 under multipath/NLoS-free conditions [19]. However, when multipath/NLoS effects occur, the receiver noise still follows a Gaussian distribution with a zero mean and a different variance σ_F^2 [20].

3. The user clock bias term dt_{rcv} keeps the same value for every satellite for a certain epoch.

In both conditions, with and without multipath/NLoS effects, $e_r^s(t)$ is relatively small compared to $F^s(t)$. Typically, in the absence of multipath/NLoS effects, the range error due to receiver noise is in the order of ± 1 m for a geodetic-quality receiver and antenna [21]. According to [20], the variance σ_F^2 increases by approximately two times compared to σ^2 when multipath/NLoS effects occur. However, the ranging bias due to multipath can reach up to about 70 m for GPS L1 C/A signals with one-chip early-to-late spacing [22]. Moreover, NLoS conditions may induce biases in pseudorange measurements spanning several kilometers [3].

Given that dt_{rcv} and $F^s(t)$ are constant for a given epoch, the leftover term $L^s(t)$ for a specific satellite can be statistically modeled as Gaussian random variables in both conditions, whether multipath/NLoS effects occur or not.

$$L^s(t) \sim \begin{cases} \mathcal{N}\left(c \cdot dt_{\text{rcv}}, \sigma^2\right) & \text{without multipath/NLoS} \\ \mathcal{N}\left(c \cdot dt_{\text{rcv}} + F^s(t), \sigma_F^2\right) & \text{with multipath/NLoS} \end{cases} \tag{4}$$

The proposed method in this research leverages the consistency checking principle to detect and estimate the multipath/NLoS effects using leftover terms. This research assumes that at least two satellites are free from the impact of multipath/NLoS effects, which can cover most conditions according to the previous study [23]. Under this assumption, $L^s(t)$ without multipath/NLoS effects is expected to follow the distribution $\mathcal{N}\left(c \cdot dt_{\text{rcv}}, \sigma^2\right)$, forming a cluster whose size is controlled by $e_r^s(t)$. On the other hand, $L^s(t)$ with multipath/NLoS effects should follow the distribution $\mathcal{N}\left(c \cdot dt_{\text{rcv}} + F^s(t), \sigma_F^2\right)$, exhibiting a mean $F^s(t)$. Given that $F^s(t)$ is significantly larger than both σ and σ_F, the leftover terms affected by multipath/NLoS will be separated from the cluster formed by leftover terms without multipath/NLoS effects. However, it is unlikely that the leftover terms impacted by multipath/NLoS can form a consistent cluster due to the distinct values of $F^s(t)$ for each satellite resulting from different reflection paths.

Based on the previous analysis, clustering algorithms can be utilized to identify the largest cluster without multipath/NLoS, subsequently enabling the determination of leftover terms affected by multipath/NLoS.

If the assumption is that at least two satellites are free from multipath/NLoS effects, the method will be unable to form a cluster consisting of $L^s(t)$ without multipath/NLoS. Consequently, the multipath/NLoS estimation method will declare a failure rather than providing an inaccurate estimate. Additionally, the clustering method may make it difficult to identify small-value $F^s(t)$ because the quantities are close to $e_r^s(t)$. However, considering the fact that, as previously recalled, $e_r^s(t)$ is expected to range in the order of about ± 1 m, those $F^s(t)$ values close to $e_r^s(t)$ will only produce minor impact positioning errors due to their small values. As a result, the main focus of this research is on those sufficiently large values of $F^s(t)$ that will induce detrimental loss of accuracy in the navigation solutions.

3.2. DBSCAN, a Clustering Algorithm for Multipath/NLoS Estimation

After having motivated the use of a clustering algorithm as a means to distinguish the presence of multipath/NLOS effects, this section will introduce the specific clustering algorithm, DBSCAN, employed in this work, discussing the parameter selection that makes it suitable to accomplish the task.

DBSCAN is a minimum density level estimation that clusters data based on the density. This algorithm first specifies two parameters:

1. **minPts**: the minimum number of points to form a cluster.
2. $\boldsymbol{\varepsilon}$: the maximum distance between two points to consider them neighbors.

Then, every data point will be classified into three types:

1. **Core points**: the data points can find at least minPts neighbors within the radius ε.

2. **Non-core points (border points)**: within radius ε, the data points can find at least one core point but have no more than minPts neighbors.
3. **Outliers**: the data points do not satisfy either the definition of core points or the one of non-core points.

DBSCAN can be described using the flowchart in Figure 1. In the first step of DBSCAN, the RangeQuery function is employed to identify all the neighbors of a specific point. This function finds all the data points in DB whose distance to the current point p is closer than ε. Here, $dist$ is a function used to compute the distance between two data points.

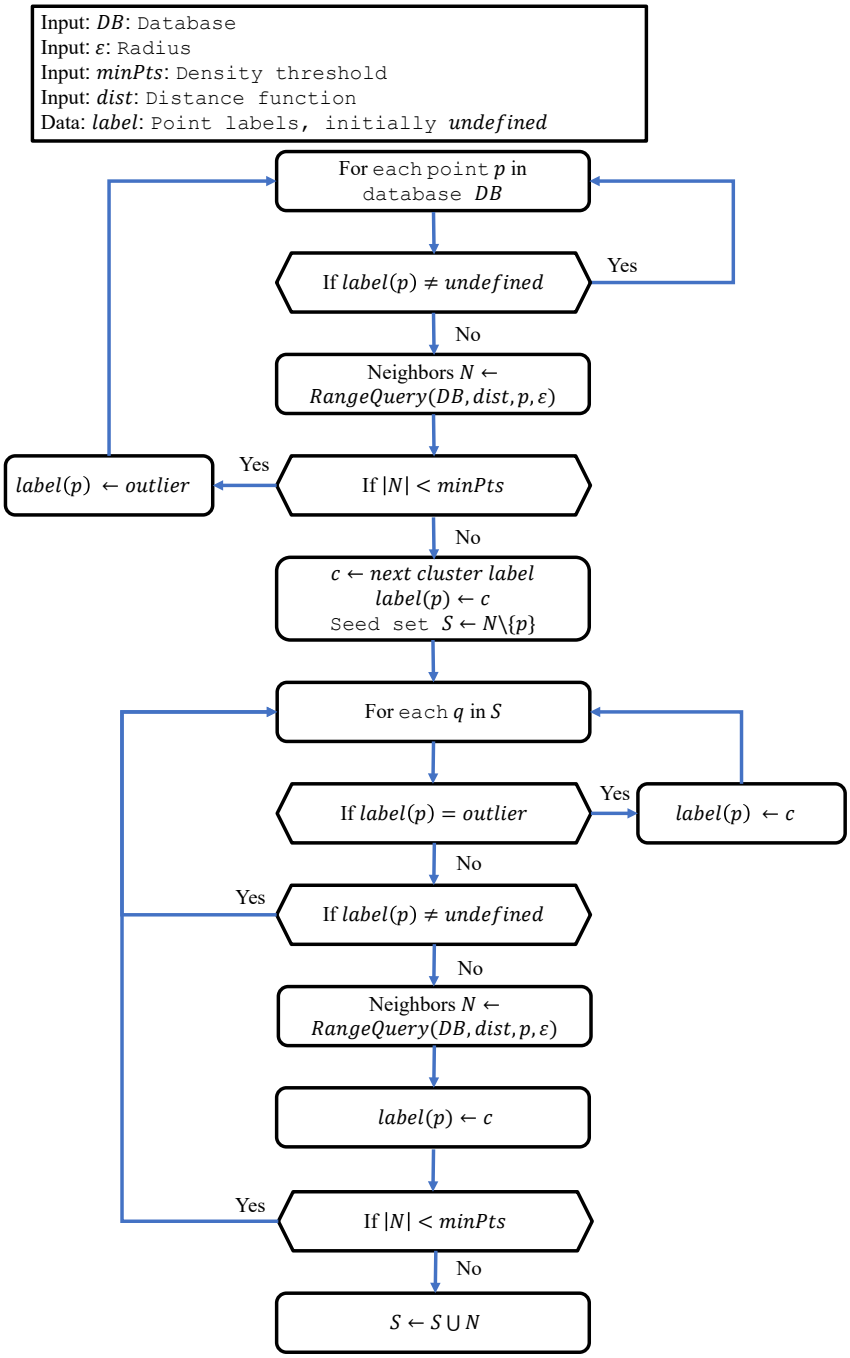


Figure 1. Flowchart of the DBSCAN algorithm.

Upon identifying a core point, all its neighbors are assigned to the same cluster as that core point. If any of these neighbors is itself a core point, the neighbors of this new

core point are classified into the same cluster. This process is iterated until all data points are clustered. Points that do not satisfy the aforementioned conditions are designated as outliers.

DBSCAN is a suitable clustering solution to address the multipath/NLoS problem since the parameter minPts establishes the minimum number of leftover terms $L^s(t)$ needed to confirm that they belong to the distribution $\mathcal{N}(c \cdot dt_{\text{rcv}}, \sigma^2)$. Meanwhile, the receiver noise $e_r^s(t)$ acts as a reference for setting the parameter ε, thus controlling the boundary of the cluster free from multipath/NLoS effects.

3.3. Implementation of Multipath/NLoS Estimation Based on DBSCAN

In total, four steps are involved in estimating the $F^s(t)$ term. The fundamental concept is to isolate $c \cdot dt_{\text{rcv}}$ from $L^s(t)$ for the pseudorange containing multipath/NLoS using the DBSCAN algorithm.

In the initial step, all M leftover terms according to (2) for a certain epoch are computed:

$$L^s(t), s = 1, \ldots, M \tag{5}$$

The second step is detecting which leftover terms are affected by multipath/NLoS. By setting the parameters ε, $minPts$, and $dist$, DBSCAN will analyze and identify clusters that meet the specified parameter configuration. Let $\tilde{L}^s(t)$ be the leftover terms belonging to the largest cluster assumed to represent the one free from multipath/NLoS effects. The leftover terms not belonging to the largest cluster are determined to be affected by multipath/NLoS effects.

Although such a condition is very rare, sometimes $L^s(t)$ values with multipath/NLoS effects from different satellites could be similar and, therefore, form a cluster. Considering the low probability of different multipath/NLoS biases being similar to each other, selecting the largest cluster further decreases the risk of incorrectly judging the multipath/NLoS-free cluster.

The third step is to estimate dt_{rcv} by computing the mean value of $\tilde{L}^s(t)$:

$$\hat{dt}_{\text{rcv}} = mean\left(\tilde{L}^s(t)\right) \tag{6}$$

The final step is to estimate the biases on pesudorange measurements due to multipath/NLoS effects. According to (3), this estimation is completed by separating $c \cdot \hat{dt}_{\text{rcv}}$ from $L^s(t)$ using the estimated $\hat{dt}_{\text{rcv}}$ obtained from the previous step.

$$\hat{F}^s(t) = L^s(t) - c \cdot \hat{dt}_{\text{rcv}}, \forall s \text{ out of the largest cluster} \tag{7}$$

4. Experiments

4.1. Static Experiment

4.1.1. Experimental Setup

A real-world experiment was carried out in a multipath scenario at the Metropolitan City of Turin (Italy), as shown in Figure 2, to demonstrate the proposed method. A Leica GS18 receiver gathered raw GNSS measurements of GPS L1 C/A signals at a rate of 10 Hz at the position depicted in Figure 3. The antenna was fixed and static for the entire collection process. Its position was georeferenced by means of an RTK solution, thus providing the position ground truth for the test. It is important to highlight that, in order to ensure the reliability of the ground truth, only the fixed RTK solutions were chosen and averaged to derive the ground truth under the conditions affected by multipath/NLoS. The GNSS data collection lasted for around 40 min. The skyplot with carrier-power-to-noise-density ratio (C/N_0) is provided in Figure 4. The receiver was deployed close to the buildings on the east side of the road but kept a short distance away from the buildings on the west side.

To obtain the $L^s(t)$ values of the pseudorange measurements, all the corrections and models were computed according to Table 1. As far as the DBSCAN is concerned, the $minPts$ and ε parameters were set as 2 and 2 m, respectively. Given the limited satellite

visibility in urban areas, *minPts* was set to 2 since at least 2 points are required to confirm membership of the same cluster. For the parameter ε, it should represent the maximum distance between two points within the same cluster. In our specific context, this maximum distance of the distribution $\mathcal{N}\left(c \cdot dt_{\text{rcv}}, \sigma^2\right)$ is dependent on the variance σ^2, which is associated with receiver noise (± 1 m), and it is then set to 2 m.

DB in this research contains all the $L^s(t)$ values, which are one-dimensional data. Hence, *dist* is set to compute the absolute value of the difference between two $L^s(t)$ values.



Figure 2. Surrounding buildings of data collection scenario.

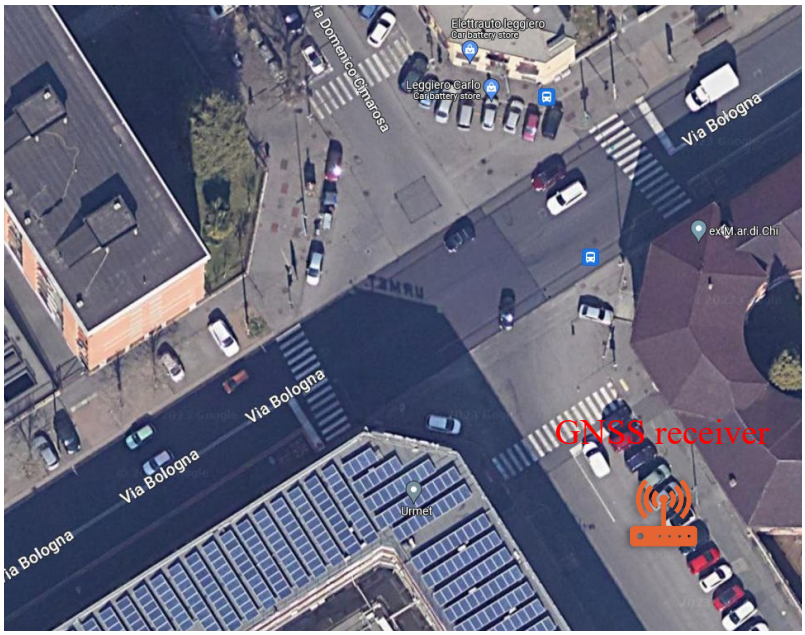


Figure 3. Location of GNSS antenna.

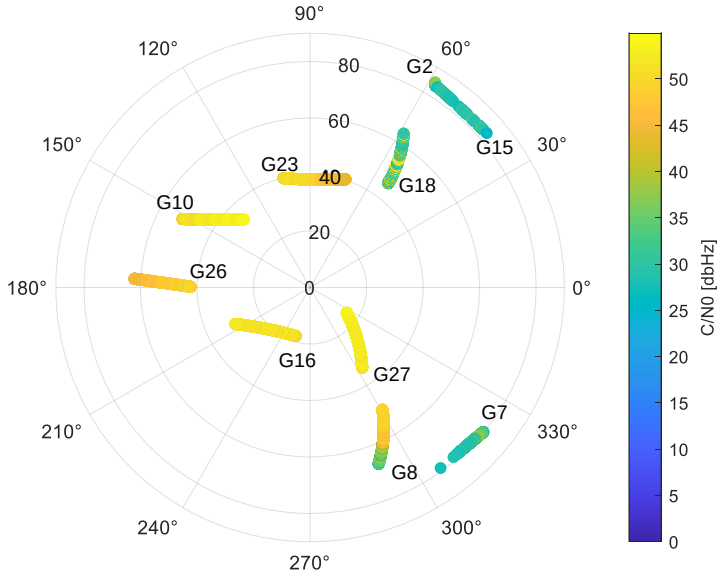


Figure 4. Skyplot with corresponding signal C/N_0.

4.1.2. Experimental Results of Multipath/NLoS Estimation

As an initial step, all the corrections and models listed in Table 1 were computed using the library provided by RTKLIB [24]. Subsequently, the pseudorange leftover terms $L^s(t)$ for all epochs were calculated applying these corrections and models. Figure 5 depicts all the $L^s(t)$ values, which are marked in different colors to distinguish the corresponding satellite pseudorandom noise. It can be seen that most $L^s(t)$ values exhibit similar trends, indicating they are influenced solely by $c \cdot dt_{\text{rcv}}$ and $e_r^s(t)$. Conversely, certain $L^s(t)$ values deviate from this trend, suggesting the presence of the large biases $F^s(t)$ caused by multipath/NLoS effects. PRN 18 shows significantly different $L^s(t)$ values compared to other PRNs, indicating a high probability of suffering from multipath/NLoS effects. Based on the skyplot in Figure 4 and the surroundings of the GNSS receiver in Figure 3, the reflected signal for PRN 18 may be attributed to the building on the west side of the road.

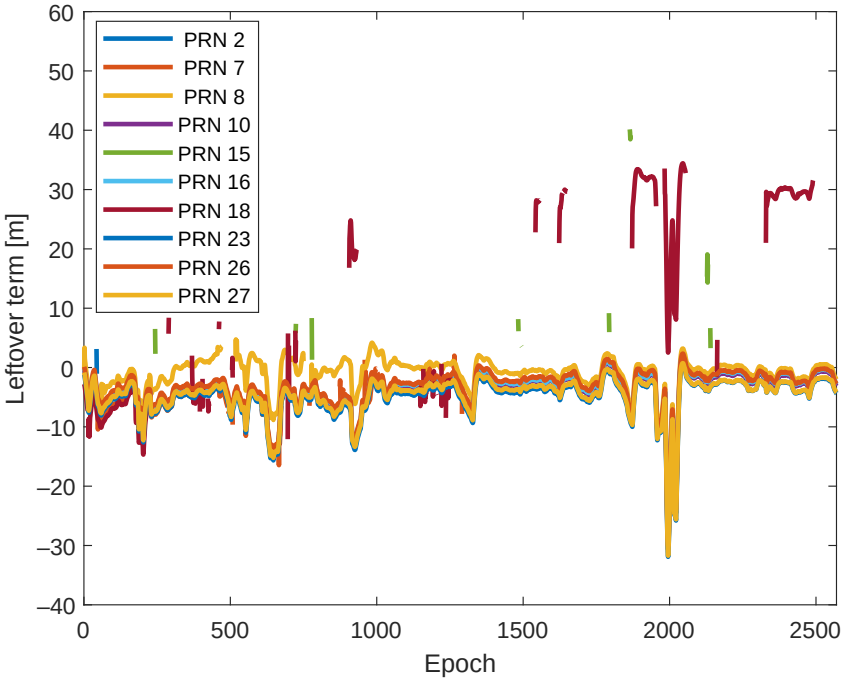


Figure 5. Computed leftover terms $L^s(t)$ for GPS satellites.

Then, Figure 6 shows all the $\tilde{L}^s(t)$ values in the largest cluster provided by DBSCAN, together with their mean value estimated as $c \cdot \hat{dt}_{\text{rcv}}$.

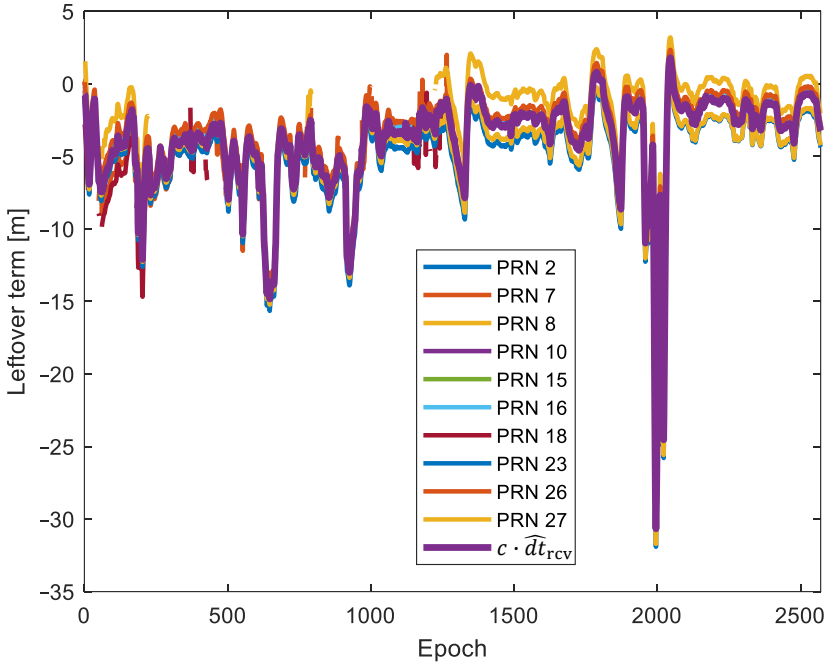


Figure 6. Leftover terms in the largest cluster determined by DBSCAN and $c \cdot \hat{dt}_{\text{rcv}}$.

Given the estimated $c \cdot \hat{dt}_{\text{rcv}}$, the $\hat{F}^s(t)$ values can be obtained for the leftover terms not belonging to the largest cluster based on (7), and they are depicted in Figure 7. It can be seen that positive values of $\hat{F}^s(t)$ are observed more frequently than negative values, consistent with the theoretical analysis in [2]. In fact, multipath can introduce both positive and negative biases to pseudoranges, while NLoS tends to produce only positive biases. Therefore, positive biases are more frequent.

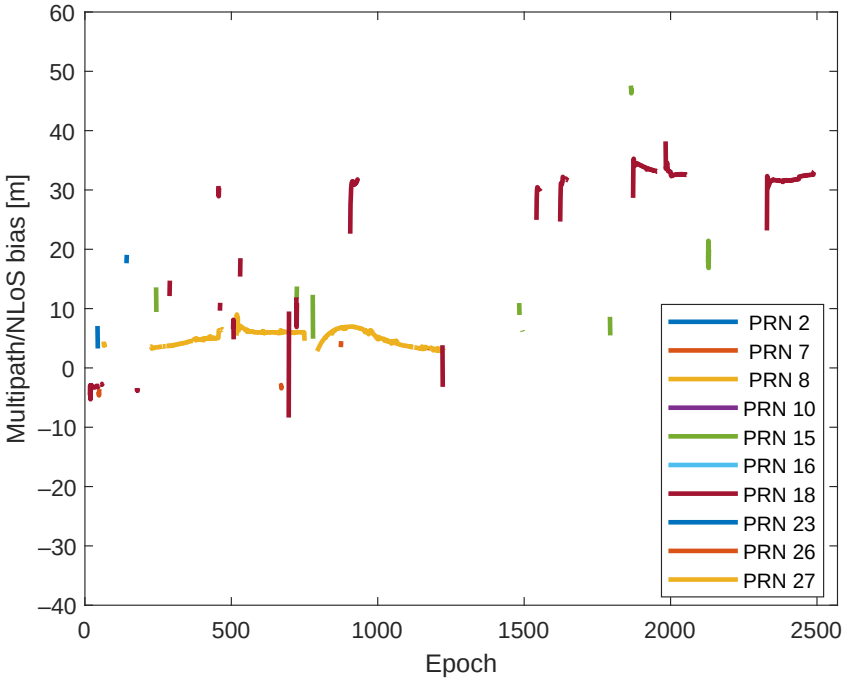


Figure 7. Estimated multipath/NLoS bias on pseudoranges.

After obtaining the $\hat{F}^s(t)$, it is necessary to prove the reliability of these estimates. Since there are no ground truth values for pseudorange measurement, in this work, we use a Generalized Least Squares (GLSs) positioning algorithm operating on pseudoranges corrected for the estimated values of $\hat{F}^s(t)$. Even if, as explained in the experimental setup, a reliable ground truth for the position can be obtained, it is not straightforward to obtain the ground truth for each pseudorange measurement since, in particular, a true value for the receiver clock $dt_r(t)$ in (1) is not known. Only the estimated value is available, and such estimates are affected by the presence of multipath/NLoS biases. For this reason, to validate the accuracy of the estimated multipath/NLoS biases, we prove the effectiveness of the method at the Position, Velocity, and Time level by comparing the accuracy of the solution obtained using pseudoranges with and without the application of the correction for the multipath/NLoS biases. Following this principle, new positioning results are computed using a new GNSS dataset, which is generated by compensating for $\hat{F}^s(t)$ in the pseudorange measurements. This compensating action is conducted by subtracting the estimated $\hat{F}^s(t)$ from the corresponding pseudorange measurements.

Figure 8 compares positioning scatters with and without multipath/NLoS compensation. Given that $\hat{F}^s(t)$ is estimated from the leftover term, which contains ground truth information, the scatter of positioning after multipath/NLoS compensation is concentrated around the true position with high accuracy. This indicates that the computed $\hat{F}^s(t)$ values closely approximate the actual pseudorange biases caused by multipath/NLoS effects. Figure 9 shows the positioning error time series in both the horizontal and vertical directions before and after the multipath/NLoS compensation. It is observed that the positioning errors are smaller in magnitude than the estimated $\hat{F}^s(t)$ for the dataset without multipath/NLoS compensation. This occurs because the locations are computed based on both the pseudorange measurements with and without multipath/NLoS effects. The pseudorange measurements without multipath can provide accurate positioning information, thereby mitigating the multipath/NLoS positioning accuracy degradation.

The Cumulative Distribution Functions (CDFs) of these positioning errors are provided in Figure 10. The CDF illustrates the distribution of positioning errors, serving a crucial role in assessing the enhanced positioning accuracy after multipath/NLoS compensation. The 3D Root Mean Squared Error (RMSE) can be improved from 13.92 m to 3.01 m (reduced by 78%) using the new dataset that compensates for multipath/NLoS biases.

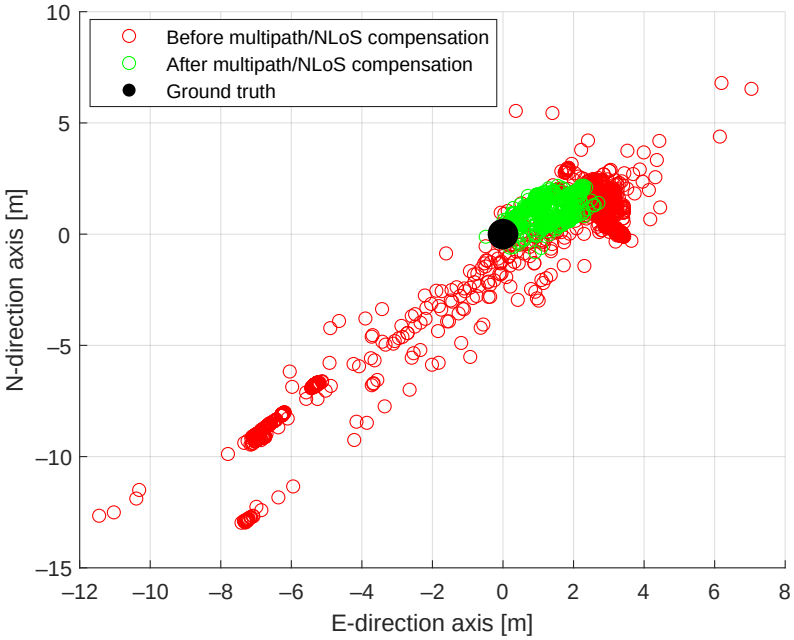


Figure 8. Positioning scatters before and after pseudorange multipath/NLoS compensation in the horizontal plane.

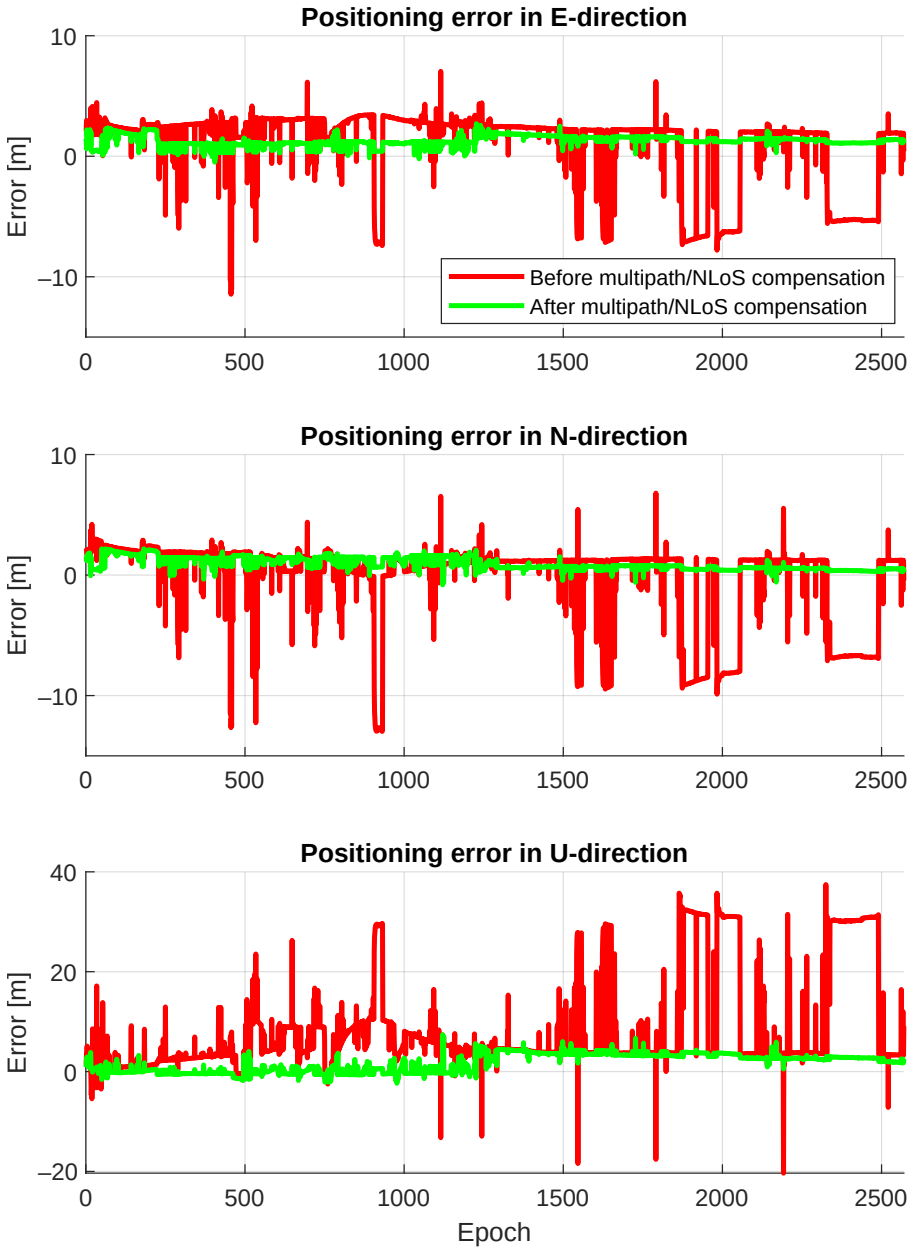


Figure 9. Positioning error time series before and after pseudorange multipath/NLoS compensation in East, North, and Up directions.

Table 2 presents the horizontal and vertical errors (at the 95th percentile of the CDFs) derived from the GNSS datasets both before and after the multipath/NLoS compensation. The positioning solutions after the multipath/NLoS compensation in Figure 8 still contain slight biases. However, these positioning results meet the standard accuracy performance criteria for GPS, as outlined in [25].

Table 2. Horizontal and vertical positioning errors at 95th percentile of the CDFs.

	Horizontal [m]	**Vertical [m]**
Before Compensation	10.94	31.04
After Compensation	2.54	4.22

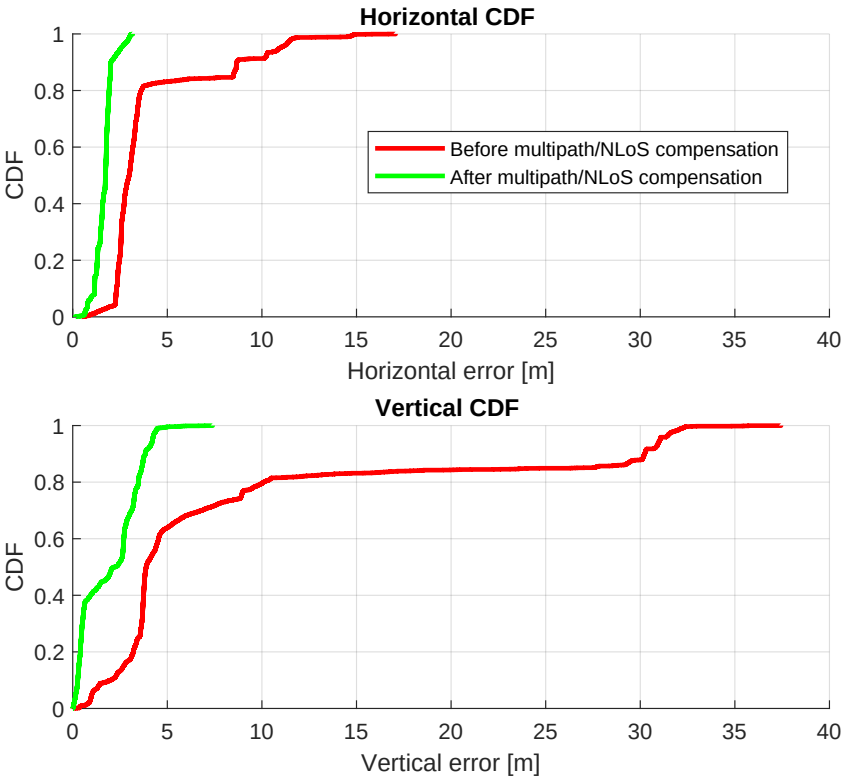


Figure 10. CDFs of positioning errors before and after pseudorange multipath/NLoS compensation in horizontal and vertical directions.

4.2. Dynamic Experiment

4.2.1. Experimental Setup

The dynamic experiment utilizes the UrbanNav dataset provided by Hong Kong Polytechnic University [26]. The specific dataset used, named Odaiba, was collected within the challenging urban canyons of Tokyo. This experiment focuses on single-frequency GPS L1 C/A signals. The GNSS data were acquired at a rate of 10 Hz using Trimble NetR9. The ground truth information was provided by the integration of RTK and INS, offering an RMSE of 5 cm and a frequency of 10 Hz under the multipath/NLoS conditions. The trajectory for the entire experiment is illustrated in Figure 11. It is important to note that the experiment was carried out based on a segment of the entire trajectory featuring strong multipath/NLoS interference.

The settings for DBSCAN totally inherit the parameters provided in Section 4.1.1.



Figure 11. Vehicle trajectory for the Odaiba dataset [27].

4.2.2. Experimental Results of Multipath/NLoS Estimation

Figure 12 illustrates all the $L^s(t)$ values for the pseudorange measurements computed using the corrections and models outlined in Table 1. It is worth noting that, due to the Trimble receiver's relatively large clock bias and drift, the dominating factor in the $L^s(t)$ values is the term $c \cdot dt_{\text{rcv}}$. Consequently, the visual impact of $L^s(t)$ values influenced by multipath/NLoS effects may not be immediately apparent in Figure 12. However, the DBSCAN algorithm is still able to recognize and segregate $L^s(t)$ values affected by multipath/NLoS effects. This is achieved by identifying deviations from the general trend formed by $L^s(t)$ values that are free from multipath/NLoS effects. This can be observed in the magnified section of Figure 12.

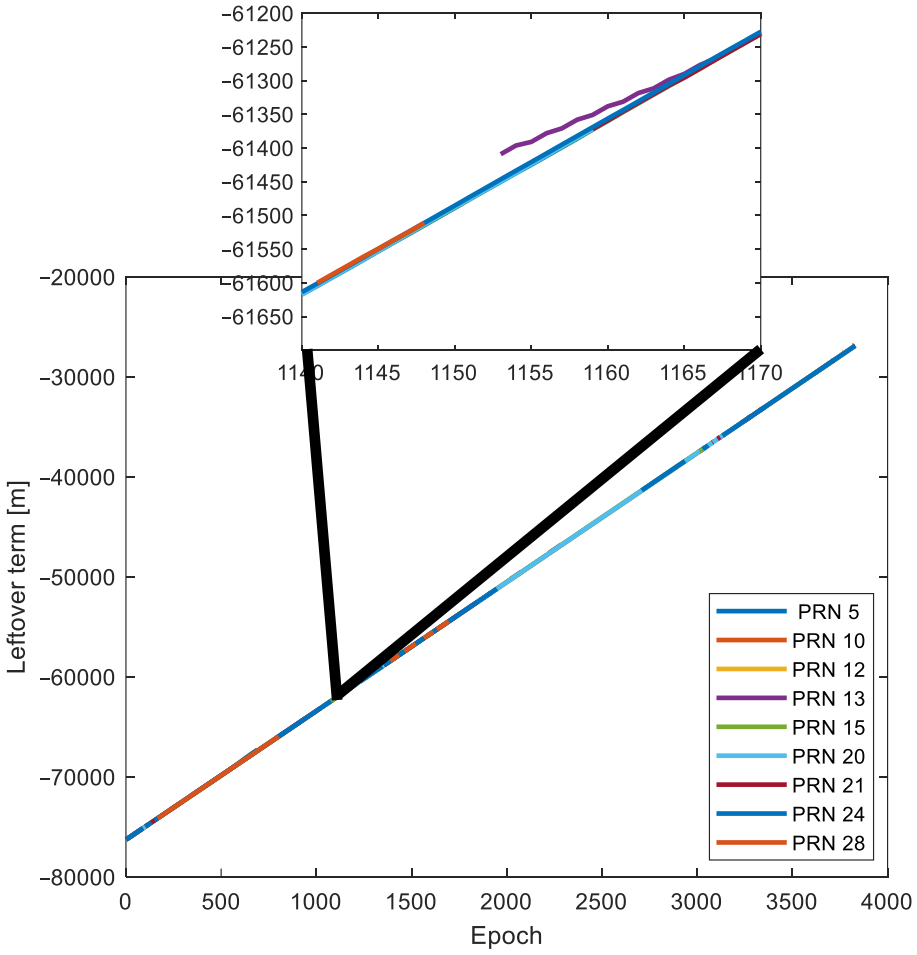


Figure 12. Computed leftover terms $L^s(t)$ for GPS satellites.

To mitigate the impact of clock bias and drift on visualization, the DBSCAN algorithm is initially employed to identify $L^s(t)$ values unaffected by multipath/NLoS effects within the largest cluster. Subsequently, the time series of the estimated $c \cdot \hat{dt}_{\text{rcv}}$ is computed by averaging all the $L^s(t)$ values within this cluster for each epoch. Figure 13 illustrates the time series of $L^s(t) - c \cdot \hat{dt}_{\text{rcv}}$ within the largest cluster. As per (3), $L^s(t) - c \cdot \hat{dt}_{\text{rcv}}$ should only be contributed by receiver noises for the cluster free from multipath/NLoS effects. The majority of the values in Figure 13 fall within the ± 2 m range, which has a similar level to the receiver noise. This observation validates the analysis in Section 3.1, which further shows the effectiveness of the DBSCAN algorithm.

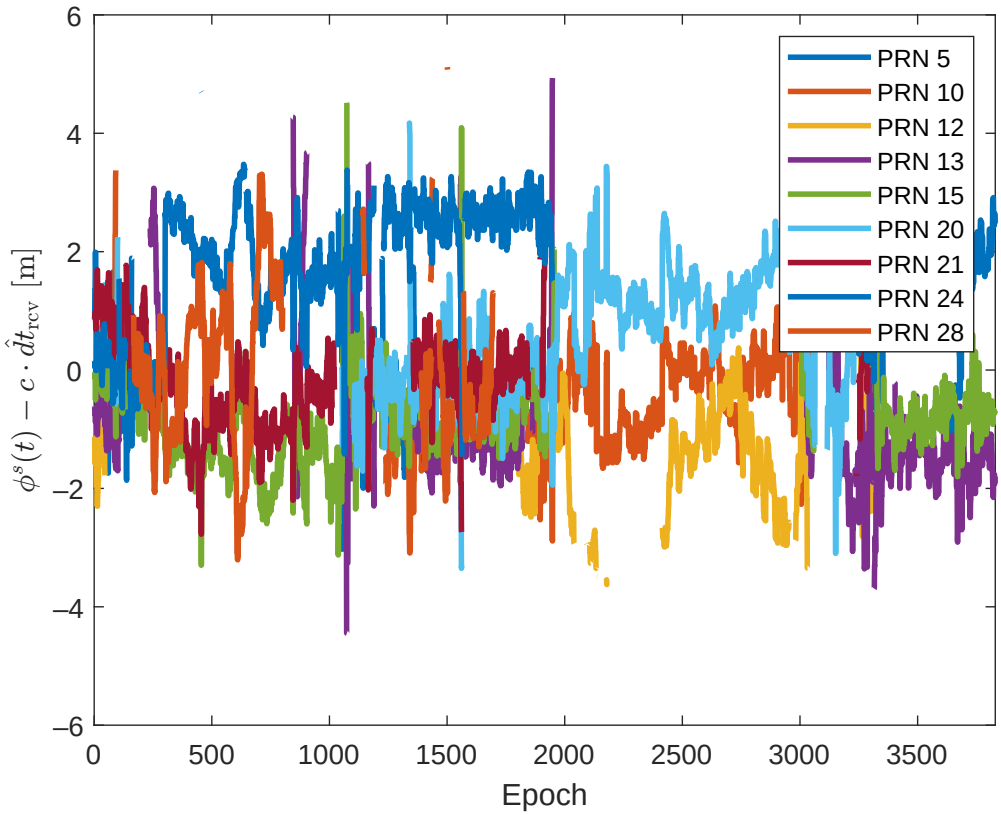


Figure 13. $L^s(t) - c \cdot \hat{dt}_{\mathrm{rcv}}$ in the largest cluster determined by DBSCAN.

Then, the $\hat{F}^s(t)$ values can be separated from the leftover terms out of the largest cluster based on (7). Figure 14 presents $\hat{F}^s(t)$ for each satellite. It is observed that PRN 5 produced significant multipath/NLoS biases, reaching up to 120 m. This is more likely caused by NLoS rather than multipath, considering that the maximum multipath bias for GPS L1 C/A signals is estimated to be around 70 m, as discussed in the theoretical analysis in [18].

Like the static dataset, a new GNSS dataset is generated by applying these $\hat{F}^s(t)$ to pseudorange measurements to evaluate the effectiveness of the estimated $\hat{F}^s(t)$ values. Subsequently, new positioning results are obtained using these adjusted pseudoranges.

Figure 15 shows a comparison of positioning trajectories with and without multipath/NLoS compensation. The positioning accuracy after the multipath/NLoS compensation proves that these computed $\hat{F}^s(t)$ values are close to the real pseudorange biases caused by multipath/NLoS effects. Figure 16 illustrates the time series of positioning errors in both the horizontal and vertical directions before and after the multipath/NLoS compensation. The CDFs of these positioning errors are presented in Figure 17. The 3D RMSE shows a significant improvement, decreasing from 26.0103 m to 5.5094 m (reduced by 79%), using the new dataset that compensates for multipath/NLoS biases.

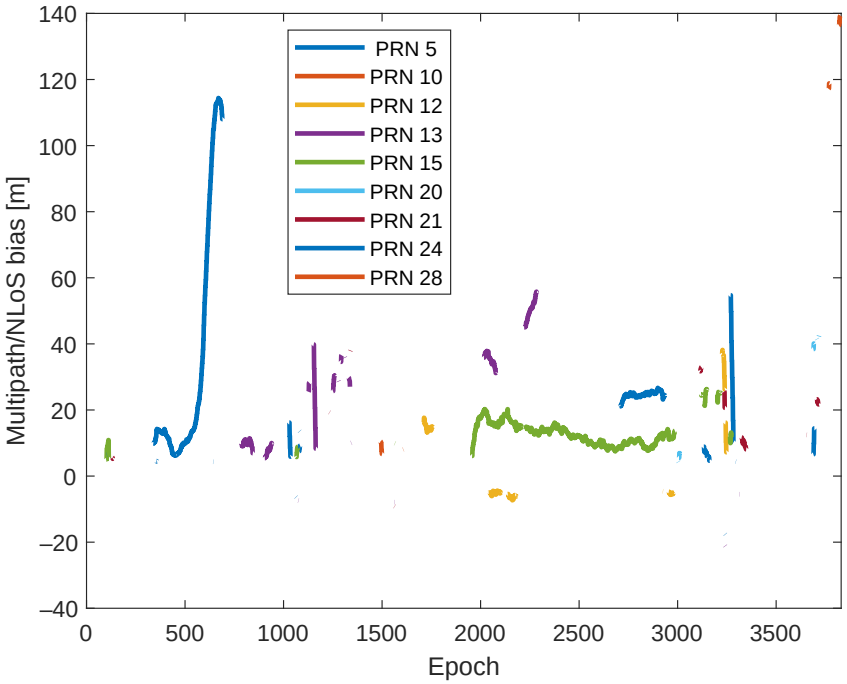


Figure 14. Estimated multipath/NLoS bias on pseudorange.

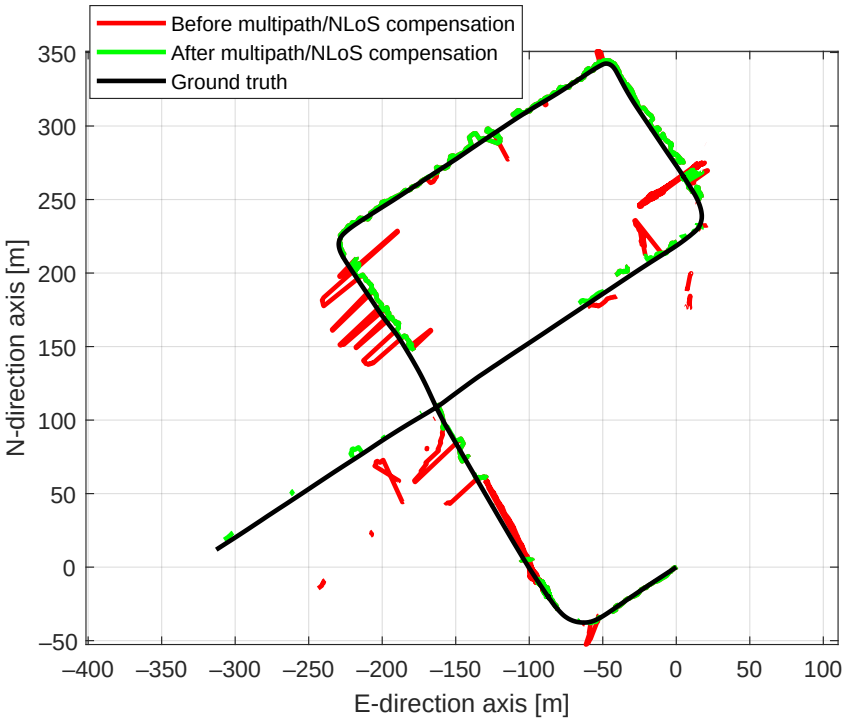


Figure 15. Positioning trajectory before and after pseudorange multipath/NLoS compensation in the horizontal plane.

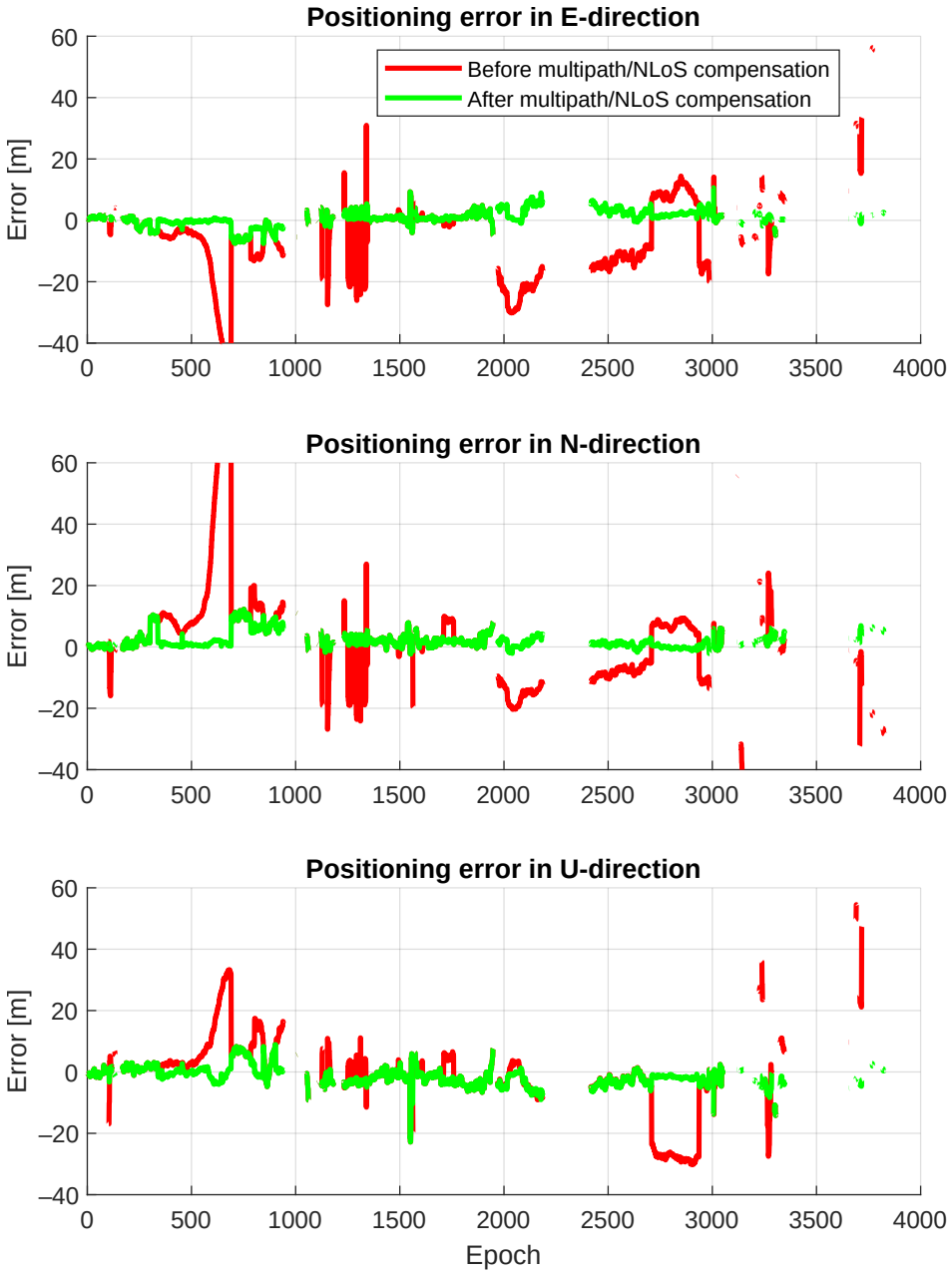


Figure 16. Positioning error time series before and after pseudorange multipath/NLoS compensation in East, North, and Up directions.

Table 3 provides the horizontal and vertical errors, specifically the 95th percentile of the CDFs, obtained from the GNSS datasets before and after the compensation for multipath/NLoS effects.

Table 3. Horizontal and vertical positioning errors at 95th percentile of the CDFs.

	Horizontal [m]	**Vertical [m]**
Before Compensation	38.75	28.68
After Compensation	9.98	6.43

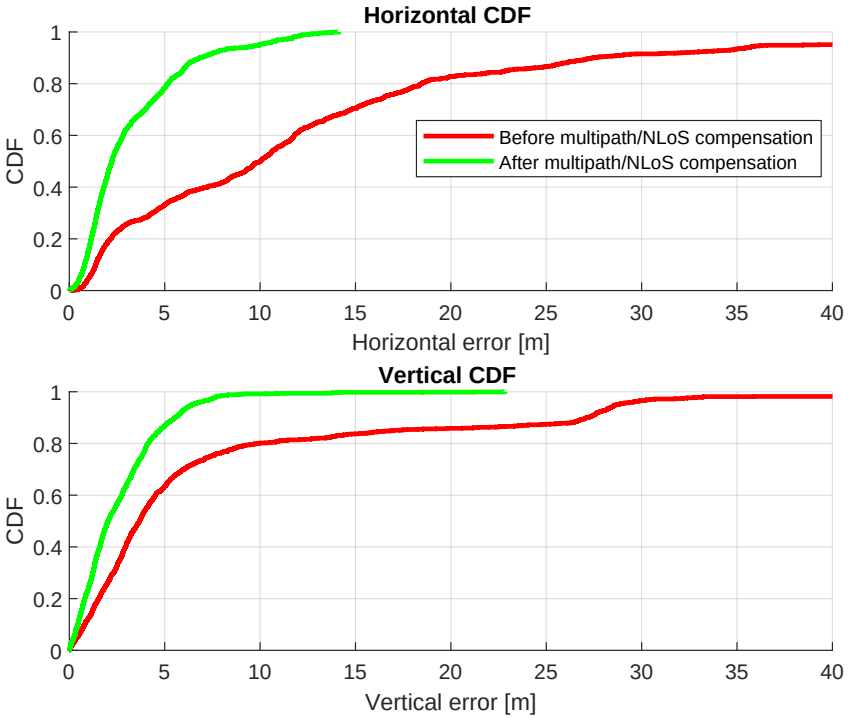


Figure 17. CDFs of positioning errors before and after pseudorange multipath/NLoS compensation in horizontal and vertical directions.

5. Conclusions

This study proposed a post-processing method to calculate the pseudorange biases resulting from multipath/NLoS effects. Leveraging the physical models and corrections from IGS, pseudorange leftover terms were computed. The application of the DBSCAN algorithm facilitated the separation of the receiver clock parameters.

The novelty of this paper can be summarized in three key aspects. Firstly, the paper defined a novel leftover term derived from insights into the pseudorange equation. Secondly, through theoretical analysis, two probability models for this leftover term were established under conditions with and without multipath/NLoS effects. The analysis demonstrates that leftover terms without multipath/NLoS effects can form a cluster. Finally, the paper proposed a post-processing procedure based on a clustering algorithm for estimating multipath/NLoS biases.

The experimental results, derived from two real-world datasets in challenging GNSS scenarios for both static and dynamic conditions, demonstrate the effectiveness of the proposed method. Firstly, the successful estimation of receiver clock parameters $\hat{d}t_{\mathrm{rcv}}$ was achieved using the DBSCAN method. Secondly, the pseudorange biases due to multipath/NLoS effects were isolated from the leftover terms. Finally, compensating for the estimated pseudorange biases in the measurement results in a significant improvement in 3D positioning accuracy. The 3D positioning RMSE was reduced by 78% (from 13.9217 m to 3.0121 m) and by 79% (from 26.0103 m to 5.5094 m) for the static dataset and dynamic dataset, respectively. The experimental results highlight the effectiveness of the proposed method.

While the proposed method has demonstrated promising results, as outlined earlier, its application is subject to certain limitations. For instance, in scenarios where the interference leading to multipath/NLoS effects is particularly strong, biases may affect all the measurements. Under such extreme conditions, the clustering algorithm may fail as there may be insufficient measurements unaffected by the influence of multipath/NLoS effects.

In the future, several aspects will be considered for extending the proposed multipath/NLoS estimation method, including its application to multi-constellation GNSS

systems. Additionally, it would be valuable to explore the factors that influence the accuracy of the clustering algorithms employed in this context.

Author Contributions: Conceptualization, Y.G.; methodology, Y.G.; software, Y.G.; validation, Y.G., S.Z., P.D. and F.D.; formal analysis, S.Z. and F.D.; resources, P.D.; data curation, P.D.; writing—original draft preparation, Y.G.; writing—review and editing, S.Z. and F.D.; supervision, F.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Chinese Scholarship Council (grant code 202008310182).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The static GNSS data used in this research are available from Yihan Guo (yihan.guo@polito.it) upon reasonable request; The dynamic GNSS data can be downloaded from https://github.com/weisongwen/UrbanNavDataset (accessed on 25 October 2023).

Acknowledgments: Yihan Guo acknowledges the help from Shuyang Li (Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy) for the discussion regarding the clustering algorithms.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Shapiro Effect Correction for $\delta t_{\mathbf{stc}}^{s,\mathbf{rel}}(t)$

Earth's gravitational field causes the Shapiro effect, which results in a time delay for GNSS signal propagation. The correction for the Shapiro effect can be compensated by

$$\delta t_{\mathrm{stc}}^{s,\mathrm{rel}} = \frac{2\mu}{c^3} \ln\left(\frac{\|\boldsymbol{r}^s\| + \|\boldsymbol{r}_r\| + \rho_r^s}{\|\boldsymbol{r}^s\| + \|\boldsymbol{r}_r\| - \rho_r^s}\right) \tag{A1}$$

where

- μ is the gravitational constant of Earth.

Appendix B. Relativistic Clock Correction for $\delta t_{\mathbf{clk}}^{s,\mathbf{rel}}(t)$

The relativity affects the satellite clock through both the satellite motion and the unsteady gravitation field. Although the oscillator frequency of the satellite clock has been intentionally offset to compensate for the relativistic effect, the elliptical satellite orbits cause deviations to the set offset. The compensation due to the orbit eccentricity can be calculated from

$$\delta t_{\mathrm{clk}}^{s,\mathrm{rel}} = -\frac{2}{c^2}\sqrt{a\mu}e\sin E \tag{A2}$$

- a is the orbit semimajor axis;
- e is the orbit eccentricity;
- E is the eccentric anomaly of the satellite.

References

1. Xie, P.; Petovello, M.G. Measuring GNSS multipath distributions in urban canyon environments. *IEEE Trans. Instrum. Meas.* **2014**, *64*, 366–377.
2. McGraw, G.A.; Groves, P.D.; Ashman, B.W. Robust positioning in the presence of multipath and NLOS GNSS signals. In *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2020; Volume 1, pp. 551–589.
3. Strode, P.R.; Groves, P.D. GNSS multipath detection using three-frequency signal-to-noise measurements. *GPS Solut.* **2016**, *20*, 399–412.
4. Groves, P.D.; Jiang, Z.; Skelton, B.; Cross, P.A.; Lau, L.; Adane, Y.; Kale, I. Novel multipath mitigation methods using a dual-polarization antenna. In Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2010), Portland, OR, USA, 21–24 September 2010; pp. 140–151.
5. Vagle, N.; Broumandan, A.; Jafarnia-Jahromi, A.; Lachapelle, G. Performance analysis of GNSS multipath mitigation using antenna arrays. *J. Glob. Position. Syst.* **2016**, *14*, 4.

6. McGraw, G.A.; Braasch, M.S. GNSS multipath mitigation using gated and high resolution correlator concepts. In Proceedings of the 1999 National Technical Meeting of the Institute of Navigation, San Diego, CA, USA, 25–27 January 1999; pp. 333–342.
7. Garin, L.; Rousseau, J.M. Enhanced strobe correlator multipath rejection for code & carrier. In Proceedings of the 10th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1997), Kansas City, MO, USA, 16–19 September 1997; pp. 559–568.
8. Van Nee, R.D.; Siereveld, J.; Fenton, P.C.; Townsend, B.R. The multipath estimating delay lock loop: Approaching theoretical accuracy limits. In Proceedings of the 1994 IEEE Position, Location and Navigation Symposium-PLANS'94, Las Vegas, NV, USA, 11–15 April 1994; IEEE: Piscataway, NJ, USA, 1994; pp. 246–251.
9. Chen, X.; Dovis, F.; Pini, M.; Mulassano, P. Turbo architecture for multipath mitigation in global navigation satellite system receivers. *IET Radar Sonar Navig.* **2011**, *5*, 517–527.
10. Zhu, N.; Betaille, D.; Marais, J.; Berbineau, M. Extended Kalman filter (EKF) innovation-based integrity monitoring scheme with C/N 0 weighting. In Proceedings of the 2018 IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI), Palermo, Italy, 10–13 September 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
11. Hsu, L.T. GNSS multipath detection using a machine learning approach. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
12. Suzuki, T.; Kitamura, M.; Amano, Y.; Hashizume, T. High-accuracy GPS and GLONASS positioning by multipath mitigation using omnidirectional infrared camera. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 311–316.
13. Lee, Y.; Park, B. Nonlinear regression-based GNSS multipath modelling in deep urban area. *Mathematics* **2022**, *10*, 412.
14. Suzuki, T.; Nakano, Y.; Amano, Y. NLOS multipath detection by using machine learning in urban environments. In Proceedings of the 30th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2017), Portland, OR, USA, 25–29 September 2017; pp. 3958–3967.
15. Guermah, B.; EL Ghazi, H.; Sadiki, T.; Guermah, H. A robust GNSS LOS/multipath signal classifier based on the fusion of information and machine learning for intelligent transportation systems. In Proceedings of the 2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Marrakech, Morocco, 21–23 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 94–100.
16. Savas, C.; Dovis, F. Multipath Detection based on K-means Clustering. In Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019), Miami, FL, USA, 16–20 September 2019; pp. 3801–3811.
17. Pan, Y.; Möller, G.; Soja, B. Machine learning-based multipath modeling in spatial domain applied to GNSS short baseline processing. *GPS Solut.* **2024**, *28*, 9.
18. Teunissen, P.J.; Montenbruck, O. *Springer Handbook of Global Navigation Satellite Systems*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 10, pp. 561–563.
19. Blanch, J.; Walter, T.; Enge, P.; Lee, Y.; Pervan, B.; Rippl, M.; Spletter, A. Advanced RAIM user algorithm description: Integrity support message processing, fault detection, exclusion, and protection level calculation. In Proceedings of the 25th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2012), Nashville, TN, USA, 17–21 September 2012; pp. 2828–2849.
20. Khanafseh, S.; Kujur, B.; Joerger, M.; Walter, T.; Pullen, S.; Blanch, J.; Doherty, K.; Norman, L.; de Groot, L.; Pervan, B. GNSS multipath error modeling for automotive applications. In Proceedings of the 31st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2018), Miami, FL, USA, 24–28 September 2018; pp. 1573–1589.
21. Zhang, X.; Tao, X.; Zhu, F.; Shi, X.; Wang, F. Quality assessment of GNSS observations from an Android N smartphone and positioning performance analysis using time-differenced filtering approach. *GPS Solut.* **2018**, *22*, 1–11.
22. Braasch, M.S. GPS multipath model validation. In Proceedings of the Position, Location and Navigation Symposium-PLANS'96, Atlanta, GA, USA, 22–25 April 1996; IEEE: Piscataway, NJ, USA, 1996; pp. 672–678.
23. Sokhandan, N.; Curran, J.T.; Broumandan, A.; Lachapelle, G. An advanced GNSS code multipath detection and estimation algorithm. *GPS Solut.* **2016**, *20*, 627–640.
24. Takasu, T. RTKLIB: An Open Source Program Package for GNSS Positioning. 2020. Available online: https://www.rtklib.com/rtklib.htm (accessed on 1 January 2021).
25. National Coordination Office for Space-Based Positioning, Navigation, and Timing. GPS Standard Positioning Service (SPS) Performance Standard; Technical Report. 2020. Available online: https://www.gps.gov/technical/ps/ (accessed on 18 March 2024).
26. Hsu, L.T.; Kubo, N.; Wen, W.; Chen, W.; Liu, Z.; Suzuki, T.; Meguro, J. UrbanNav: An open-sourced multisensory dataset for benchmarking positioning algorithms designed for urban areas. In Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021), Online, 20–24 September 2021; pp. 226–256.
27. *UrbanNavDataset*. Available online: https://github.com/IPNL-POLYU/UrbanNavDataset (accessed on 5 October 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.