A bagging tree-based pseudorange correction algorithm for global navigation satellite system positioning in foliage canyons

Fan Qin¹, Linxia Fu², Yuanqing Wang¹ and Yi Mao²

Abstract
Global navigation satellite system is indispensable to provide positioning, navigation, and timing information for pedestrians and vehicles in location-based services. However, tree canopies, although considered as valuable city infrastructures in urban areas, adversely degrade the accuracy of global navigation satellite system positioning as they attenuate the satellite signals. This article proposes a bagging tree-based global navigation satellite system pseudorange error prediction algorithm, by considering two variables, including carrier to noise $C/N_0$ and elevation angle $\theta$, to improve the global navigation satellite system positioning accuracy in the foliage area. The positioning accuracy improvement is then obtained by applying the predicted pseudorange error corrections. The experimental results shows that as the stationary character of the geostationary orbit satellites, the improvement of the prediction accuracy of the BeiDou navigation satellite system solution (85.42% in light foliage and 83.99% in heavy foliage) is much higher than that of the global positioning system solution (70.77% in light foliage and 73.61% in heavy foliage). The positioning error values in east, north, and up coordinates are improved by the proposed algorithm, especially a significant decrease in up direction. Moreover, the improvement rate of the three-dimensional root mean square error of positioning accuracy in light foliage area test is 86% for BeiDou navigation satellite system/global positioning system combination solutions, while the corresponding improvement rate is 82% for the heavy foliage area test.

Keywords
Foliage canyon, BDS, GPS, machine learning

Date received: 20 January 2021; accepted: 12 April 2021
Handling Editor: Lyudmila Mihaylova

Introduction
Nowadays, with the rapid development of intelligent and modernized urbanization, severe environmental issues like overcrowding, air pollution, climate change, and ecocide are emerging in the social and economic development of cities.¹ To promote the green and sustainable urbanization, the tree canopy which is full of layers of leaves and branches that cover the ground or on the roadsides are regarded as a functional management of environmental benefits and mental health. In addition, global navigation satellite system (GNSS) is also widespread as an underpinning technology for the intelligent transportation system in the modern and smart cities.² However, the multipath interference and
non-line-of-sight (NLOS) receptions, which could degrade the GNSS positioning performance in city areas, are caused by signal blockage, reflection, and diffraction of the tree canopy.

Many studies have shown the performance of the GNSS in the foliage area. Positioning accuracy and precision are examined in typical peripheral canopies on Irish forest roads. Seasonal parameters of deciduous tree foliage are studied in affecting GNSS signal attenuation. So far, various research studies have been conducted to find robust approaches to detect, classify, and eliminate the NLOS reception and multipath interference. The main solutions of NLOS reception/multipath interference elimination methods could be classified as antenna design, signal processing, and measurement-based modeling methods.

It has been shown that the antenna design method such as dual-polarization antenna, choke-ring antenna, and antenna array technology are significant in NLOS and multipath detection and mitigation at low elevation angles. However, high-cost and bulky-size make it inappropriate for widespread application in either highly built area or dense foliage streets.

The signal processing method is about mitigating the multipath interference by minimizing the space between the receiver code correlators. A more comprehensive review of signal processing-based multipath mitigation methods is described. Furthermore, Newton method-based fast iterative maximum-likelihood algorithm (FIMLA), advanced receiver-architecture vector tracking-based method, and probabilistic neural network (PNN)-based methods are functional methods to mitigate multipath. However, signal processing-based method is invalid for NLOS signal elimination.

Measurement-based method is considered as integrating other source of information such as sensors, spatial, and geometric information with the GNSS systems to compensate the signal blockage to enhance the performance of positioning. For the GNSS with sensor information measurement-based method, the integrated GNSS with inertial navigation system (INS) was proposed which provide timely position, velocity, and attitude information by dead reckoning system for positioning improvement of a moving object. High accuracy-level INS system is always expensive for ground-use which is functional but not recommended. Utilizing dual-frequency observations, code-minus-carrier measurements (CMC) can be obtained to estimate code multipath error and evaluate multipath conditions when carrier-phase and code-phase measurements are available.

To eliminate the multipath effects effectively, the correctness and robustness of distinguishing and classifying the direct, NLOS, multipath signals are essential. The strength of the signal which is defined as the carrier to noise ($C/N_o$) value is a traditional signal identification indicator. The signal with higher $C/N_o$ value is defined as line-of-sight (LOS), while those with lower $C/N_o$ value is defined as NLOS and multipath. In addition, the elevation angle ($\theta_e$) is another important factor which is combined with $C/N_o$ to examine the signal types. While it is indicated that the signals from the lower elevation satellites are generally received as NLOS in the urban environment. Besides, the machine learning-based signal reception classification has been widely studies in recent years. The support vector machine (SVM)-based method treats elevation and azimuth angle as the key features to eliminate the multipath effects. Yozevitch et al. selected signal strength, elevation, and other variables as the training indicators for decision trees-based method to identify the types of receiving signals. Sun et al. defined nine indicators which are derived from the raw global positioning system (GPS) measurements for the adaptive neuro fuzzy inference system (ANFIS)-based algorithm to classify the signal types. A gradient boosting decision tree (GBDT)-based signal classification method was proposed to improve the signal classification accuracy.

To mitigate NLOS and multipath caused by the built-up constructions, a three-dimensional (3D) digital model which provides additional geometric information such as the height of the building, tilt, orientation, and area of the roof is treated as an available measure. Groves indicates that the main limitation of the existing 3D digital models is that they assume that the roof of the buildings are flat, whereas the most roofs of the buildings are rough and even have different shapes in reality.

What’s more, it is further difficult for the 3D digital models to present other detailed roadside objects information, such as the moving vehicles, growing trees foliage, and changing irregular shapes of leaves. Above all, it is concluded that 3D digital models cannot be effective to solve the NLOS reception and multipath problem in foliage canyons, while GNSS and INS integration-based methods are effective with high equipment cost in solving foliage canyons.

Thus, to improve the positioning performance in foliage canyons, we have developed a bagging tree-based GNSS positioning algorithm with a combined strategy of pseudorange error prediction and correction. The contributions are summarized as follows:

1. A bagging tree-based pseudorange error prediction and correction algorithm is proposed by using the signal strength and elevation angle as the inputs.

2. The GNSS positioning accuracy in foliage canyon could be improved by the pseudorange correction combined strategy. The field test has validated that the proposed algorithm could
effectively enhance the GNSS positioning performance in both light and heavy foliage areas. The positioning accuracy of the single or multiple GNSS, that is, BeiDou navigation satellite system (BDS) only and BDS/GPS, is analyzed and compared.

Algorithm

Algorithm framework

The proposed bagging tree-based pseudorange determination algorithm is presented in Figure 1. The main processes could be divided into the training and testing stages.

In the training stage, first, an offline data set is created which are GNSS raw pseudorange measurements including LOS, NLOS, and multipath signals from a known point in the foliage area. Second, calculating the pseudorange error which is equal to the difference between the raw pseudorange measurement and the corresponding geometric ranges from the known station coordinate and satellite ephemeris. Then, as the input features for machine learning, the signal strength ($C/N_0$) and elevation angle ($\theta_e$) at each epoch will be labeled with the pseudorange error to obtain the training data set. Finally, utilizing this training data set, a bagging tree-based algorithm is used to develop the rules about the relationship between the input features and the corresponding pseudorange errors. More details about each step will be discussed in the following sections.

During the online testing stage, the extracted rules from the offline stage are being used to predict the pseudorange errors according to the data of newly received variables (i.e. signal strength and elevation angle). Then, the position solutions are calculated with the pseudorange error proposed by the pseudorange correction algorithm based on the predicted pseudorange errors which are treated as the corrections to the newly raw pseudorange measurements.

Feature selection

Signal strength and elevation angle are selected as the inputs:

1. Signal strength ($C/N_0$)

The strength of the signal is represented by the $C/N_0$ value. $C/N_0$ is defined as the ratio of the received signal power to the weighted noise power spectral density. It is known that high-elevation satellites should have a high $C/N_0$, while NLOS satellites should have a lower received $C/N_0$, although NLOS with high $C/N_0$ can occur. Thus, this indicator is not fully sufficient to identify the received signal.

2. Elevation angle ($\theta_e$).

Generally speaking, the elevation angles of satellites influence the satellite visibility. It is certified that a signal of higher elevation angle is less possible to be blocked by the surroundings. Thus, the satellite elevation is served as a probabilistic LOS indicator.
Pseudorange error labeling process

Pseudorange error calculation. Theoretically, the raw pseudorange measurement can be derived by multiplying the light speed $c$ with signal propagation time which is the difference value between the departure time from the satellite $t_s$ and the arrival time from the receiver $t_r$, expressed as

$$\rho = (t_r - t_s) \times c$$  \hspace{1cm} (1)

However, the pseudorange measurement is always affected by several errors, such as ionospheric delay, tropospheric delay, receiving clock and satellite ephemeris error, receiving dynamic error, and multipath. The correction of the pseudorange measurement can be improved as

$$\rho = r + \Delta \rho = r + c \times (\tau' - \tau^*) + I + T + \varepsilon$$  \hspace{1cm} (2)

where the geometric range is denoted by $r$; $\Delta \rho$ is the pseudorange error; $\tau'$ and $\tau^*$ are the receiver clock error and the satellite clock error, respectively; $I$ and $T$ are the ionospheric delay and the tropospheric delay, respectively; $\varepsilon$ is the errors from other sources such as multipath error, receiver noise, and antenna delay.

The geometric range $r$ in equation (2) can also be expressed by the following equation

$$r = \sqrt{(X_R - X_S)^2 + (Y_R - Y_S)^2 + (Z_R - Z_S)^2}$$  \hspace{1cm} (3)

where $(X, Y, Z)_R$ is the coordinate of the GNSS receiver $R$, and $(X, Y, Z)_S$ is the coordinate of the satellites.

As the receiver clock error, satellite clock error, ionospheric delay, and tropospheric delay could be moderated by proper models, and their corresponding residuals are small enough to ignore, and the main errors are considered as multipath errors.

Labeling process. Then, label the pseudorange error $\Delta \rho$ to the corresponding values of input variables. From the above data, we can obtain the set that consists of GNSS measurements, containing $C/N_0$, $\theta_\alpha$, and pseudorange error at each epoch in the offline labeling phase. Then, the labeled data set is created to train.

Bagging tree-based pseudorange correction algorithm

Ensemble learning, as a kind of machine learning, aims at increasing the accuracy of the results by combining multiple models instead of using a single model. Compared with the ordinary machine learning methods, the ensemble methods try to create a series of hypotheses and combine them to use.\(^{29}\) Bagging and boosting are the two main methods of ensemble machine learning. A gradient-boosting decision tree algorithm was designed as a signal classifier.\(^{19}\) Compared with the boosting technique, bagging is advantageous of minimizing the decision tree training error, reducing the variance of a predictor, and avoiding overfitting.\(^{31,32}\) A bagging procedure is classified as bootstrapping and aggregation. Bootstrapping is a sampling technique where samples are selected from the origin data set with replacement randomly. Aggregation is a procedure of combing all the possible outcomes of the prediction for the final prediction. The bagging method is shown in Figure 2.

In this study, the bagging method is proposed to form the stronger and robust prediction rules. Step 1 in bagging is training the multiple learners from different bootstrap samples. A bootstrap sample is obtained by subsampling the training data set with replacement. And the sample size is as the same as that of the training data set.\(^3\) Step 2 in bagging is aggregating the multiple learners by averaging for regression.

The process of the bagging regression tree-based pseudorange error prediction method is as follows:

1. Define an original training data set of $n$ examples:

$$L = \{(x_1, \Delta \rho_1), (x_2, \Delta \rho_{12}), (x_3, \Delta \rho_3), \ldots (x_n, \Delta \rho_n)\}$$  \hspace{1cm} (4)

2. Where $x_i = (C/N_0, \theta_\alpha), i = 1, 2, 3, \ldots n$. Here, $n$ is the total number of the samples, and $\Delta \rho_n$ is the labeled pseudorange error of $x_i$.

3. To form a sequence number of the weaker predictor $w(x, L_s)$ by the decision tree algorithm on the sequence of the training data set. Let $x_i$ be the input, and $L_s$ is predicted by $w(x, L_s)$.

4. A bagged predictor $W_{\text{aver}}$ is obtained as a final predictor by the average of the $w(x, L_s)$ over $S$

$$W_{\text{aver}} = \frac{\sum_{i=1}^{s} w(x_i, L_s)}{s}$$  \hspace{1cm} (5)

5. Once the final predictor $W_{\text{aver}}$ is obtained, the corresponding pseudorange errors could be predicted.

6. Positioning solutions are calculated. The newly testing pseudorange measurement can be corrected by subtracting the predicted pseudorange error obtained above.

Field test and results

Experiment setup

In order to validate the proposed pseudorange prediction and determination algorithm in solving the foliage canyon problems, the field test was carried out at foliage area in Jiangjun Road Campus of Nanjing...
University of Aeronautics and Astronautics, China. This test consists of two stages, one stage is to obtain the ground truth and the other stage is to collect the testing data. To obtain the ground truth, two locations in the open sky area near the foliage area were selected to collect reference data for more than 1 h, and then the network real-time kinematic (NRTK) services from Qianxun SI Ltd was used to obtain the centimeter-level positioning solutions for these locations. The accurate coordinates (centimeter-level accuracy) of the reference points in the foliage areas are then obtained by the total station Leica TS02 based on these two control points. The testing data were collected at 10 Hz data rate by a NovAtel OEM7 receiver in two points. One is in a light foliage area and the other one is in a heavy foliage area, as shown in Figures 3 and 4.

The summary of the testing data sets is shown in Table 1. In Test 1, BDS-only and BDS/GPS combination observations are defined as Scenario 1 and Scenario 2, respectively. In Test 2, BDS-only and GPS/BDS combination observations are defined as Scenario 3 and Scenario 4, respectively. For each testing scenario, 4000 samples are selected for training and 9000 samples are selected for testing.

Results analysis

To assess the positioning performance, the satellite visibility and the dilution of precision (DOP) value are two indicators which are measured and evaluated. The satellite visibility of BDS-only and BDS/GPS combination observations in partial epochs is shown in Figures 5 and 6. As shown below, the number of visible satellites of the multiple constellation strategy is superior to the single constellation strategy. In addition, as the severe signal blockage in heavy foliage, the number of visible
satellites fluctuates more dramatically in both BDS-only and BDS/GPS scenarios than that shown in Figure 5.

The position dilution of precision (PDOP) and horizontal dilution of precision (HDOP) of BDS-only and BDS/GPS combination observations are shown in Figures 7 and 8. It can be seen from the results that both the PDOP and HDOP values of BDS + GPS are lower than that of BDS-only. Besides, compared with the HDOP values in all scenarios, the higher value of PDOP demonstrate that the positioning performance in three dimensions is inferior to that of the horizontal ones. The above situation is more significant as the signal blockage and multipath inference are caused by the heavily surrounding trees in the up direction.

The testing results are shown in Figures 9 and 10, respectively. It can be seen that both the BDS-only and BDS/GPS positioning solutions by the proposed method are closer to the ground truth than that of the conventional results.

In order to verify the performance of the proposed algorithm, the pseudorange error prediction accuracy are evaluated and shown in Figures 11–13. If the absolute value of the difference between the actual measured pseudorange error and the predicted pseudorange error is less than the actual value, it indicated that the proposed algorithm is valid, and invalid when it does not. In BDS-only scenario, the prediction accuracy by the proposed algorithm is 82.45% in light foliage area and 81.48% in heavy foliage area. In the BDS/GPS scenario, the prediction accuracy of the BDS solution and GPS solution by the proposed algorithm is 85.42% and 70.77% in light foliage area, respectively, while the prediction accuracy of the BDS solution and GPS solution

| Field test ID | Test 1: light foliage | Test 2: heavy foliage |
|---------------|-----------------------|-----------------------|
| Scenario 1    | Scenario 2            | Scenario 3            | Scenario 4            |
| Constellations| BDS-only              | BDS/GPS               | BDS-only              | BDS/GPS               |
| Training/Testing sample | 4000/9000 | 4000/9000               | 4000/9000             | 4000/9000             |

BDS: BeiDou navigation satellite system; GPS: global positioning system.
Figure 7. DOP value at light foliage area.

Figure 8. DOP value at heavy foliage area.

Figure 9. Positioning results in light foliage: (a) BDS-only scenario; (b) BDS/GPS scenario.

Figure 10. Positioning results in heavy foliage: (a) BDS-only scenario; (b) BDS/GPS scenario.
by the proposed algorithm is 83.99% and 73.61% in heavy foliage area, respectively. The testing results show that the proposed algorithm displays better performance in BDS solutions compared to the GPS solutions. This performance advantage of BDS is mainly because that BDS consists of not only the medium earth orbit (MEO) satellites but also the geostationary orbit (GEO) satellites and inclined geosynchronous orbit (IGSO) satellites. The modeling in the proposed algorithm is more effective for the GEO satellites as the GEO satellites are almost stationary, while it is less effective for the MEO satellites and IGSO satellites, which are with large motions.

In addition, the positioning error values in the east, north, and up direction by both the conventional single point positioning method and the bagging tree-based pseudorange correction-based positioning method are computed and depicted in Figures 14 and 15. As the conventional results in these two tests show, the East and North coordinate error values are less than 5 m, while the up coordinate error values are more than 30 m. Compared with the conventional results, the proposed methods could decrease the error value significantly, especially in the up direction.

The comparison of the root mean square error (RMSE) for the horizontal 2D position and 3D position about both the conventional single point positioning method and the bagging tree-based pseudorange correction-based positioning method is shown in Tables 2 and 3, respectively.

In light foliage area, the 2D and 3D position RMSE for BDS-only observation is 2.97 m and 22.21 m, respectively. After combing the BDS and GPS observation, the 2D RMSE is improved to 2.82 m while the 3D RMSE is degraded to 22.86 m. In heavy foliage area, the conventional positioning method results also show that the 3D RMSE of the combination of BDS and GPS is 23.51 m which is 22.61 m in BDS-only scenario. It is indicated that the combined BDS/GPS positioning accuracy is degraded compared with BDS-only solutions as the GPS satellites are not contributing to the positioning accuracy, even introducing more multipath and NLOS signals when there is enough high-quality BDS satellites.

The performance of the proposed bagging tree-based pseudorange correction method in improving the positioning accuracy in various scenarios is shown above. In light foliage area, the proposed bagging tree-based method proposed a 14.39% (2D) and 77.61% (3D)
Figure 14. Positioning error at light foliage area.

Figure 15. Positioning error at heavy foliage area.

Table 2. GNSS positioning accuracy in Test 1: light foliage area.

| RMSE (m)          | Scenario 1 | Scenario 2 |
|-------------------|------------|------------|
|                   | BDS-only   | BDS/GPS    |
|                   | 2D         | 3D         | 2D         | 3D         |
| Conventional      | 2.97       | 22.21      | 2.82       | 22.86      |
| improvement       | 14.39%     | 77.61%     | 26.89%     | 85.79%     |
| The proposed      | 2.54       | 4.97       | 2.06       | 3.25       |
| improvement       | 14.39%     | 77.61%     | 26.89%     | 85.79%     |

BDS: BeiDou navigation satellite system; GPS: global positioning system.
improvement in the RMSE value of the BDS-only positioning accuracy than that of the conventional positioning method. And 26.89% (2D) and 85.79% (3D) improvement in the RMSE value of the BDS/GPS positioning accuracy than that of the conventional positioning method. In heavy foliage area, the proposed bagging tree-based method provide a 13.82% (2D) and 61.69% (3D) improvement in the RMSE value of the BDS-only positioning accuracy than that of the conventional positioning method. And 28.17% (2D) and 82.17% (3D) improvement in the RMSE value of the BDS/GPS positioning accuracy than that of the conventional positioning method. It can be concluded that the proposed algorithm can achieve a significant improvement in the 2D and 3D positioning accuracy, especially in 3D solutions. In addition, various scenarios indicated that the percentage improvement of the 3D RMSE value is significantly higher than that of the 2D RMSE value, which indicated that the up direction is influenced severely by the foliage canyon environment.

### Conclusion

This article proposed a bagging tree-based method to improve the GNSS positioning accuracy in solving foliage canyons by predicting the pseudorange errors and correcting the pseudorange measurement. Based on the bagging tree method, the pseudorange error predicted from the input features of signal strength, and elevation angle, the positioning solutions are calculated with the corrected pseudorange measurement. Field tests are carried out in light and heavy foliage canyons, and the results indicate that the performance of the satellite visibility and DOP indicators of the BDS/GPS scenario are superior to the BDS-only scenario. Besides, like the severe signal blockage and multipath inference in heavy foliage area, the number of visible satellites fluctuates more dramatically and the PDOP value is higher. Moreover, the test results show that the proposed bagging tree-based pseudorange correction method could achieve a significant improvement in positioning accuracy compared to the conventional positioning calculation. The pseudorange error prediction accuracy is evaluated which demonstrated that the proposed algorithm tends to perform better in BDS satellites for the stationary character of the GEO satellites. The BDS/GPS error value in the east, north, and up direction by both the conventional method and the bagging tree-based pseudorange correction-based method illustrated that the up coordinate error values could be decreased significantly from more than 30 to 10 m. Furthermore, as the signal blockage and multipath by the tall foliage in the up direction, the improve rate of the 3D RMSE value in much higher than the 2D RMSE value. With the proposed algorithm, the BDS-only solution in light foliage area provides a 14.39% and 77.61% improvement in 2D and 3D RMSE, and the BeiDou/GPS combination solution in light foliage area shows a 26.89% and 85.79% improvement in 2D and 3D RMSE. With the proposed algorithm, the BDS-only solution in heavy foliage area provides a 13.82% and 61.69% improvement in 2D and 3D RMSE, and the BeiDou/GPS combination solution in heavy foliage area shows a 28.17% and 82.17% improvement in 2D and 3D RMSE. In future studies, it is advised to improve the weighting strategies in the multi-GNSS constellation system for the positioning improvements.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research work was supported by The National Key Technology R&D Program of China (2019YFB1600703) and the Higher Education Discipline Innovation Project111 (grant number: B20035).

### ORCID iD

Yi Mao [https://orcid.org/0000-0002-8624-772X](https://orcid.org/0000-0002-8624-772X)

### References

1. Ulmer JM, Wolf KL, Backman DR, et al. Multiple health benefits of urban tree canopy: the mounting evidence for a green prescription. *Health Place* 2016; 42: 54–62.
2. Meng Q and Hsu LT. Integrity for autonomous vehicles and towards a novel alert limit determination method. *Proc IMechE, Part D: Journal of Automobile Engineering* 2021; 235: 969–1006.

3. Martin AA, Holden NM, Owende PM, et al. The effects of peripheral canopy on DGPS performance on forest roads. *Int J Forest Eng* 2001; 12: 71–79.

4. Pirti A. The seasonal effects of deciduous tree foliage in CORS-GNSS measurements (VRS/FKP). *Tehn Vjesn* 2016; 23: 679–774.

5. Sun R, Wang G, Fan Z, et al. An integrated urban positioning algorithm using matching, particle swarm optimized adaptive neuro fuzzy inference system and a spatial city model. *IEEE Trans Veh Technol* 2020; 69: 4842–4854.

6. Sun R, Wang G, Cheng Q, et al. Improving GPS code phase positioning accuracy in urban environments using machine learning. *IEEE Internet Things J* 2020, https://www.researchgate.net/publication/345976151_IOT_improving_gps_code_phase_positioning_accuracy_in_urban_environments_using_machine_learning

7. Jiang Z and Groves PD. NLOS GPS signal detection using a dual-polarisation antenna. *GPS Sol* 2014; 18: 15–26.

8. Groves PD, Jiang Z, Rudi M, et al. A portfolio approach to NLOS and multipath mitigation in dense urban areas. In: *Proceedings of the 26th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2013)*, Manassas, VA, 30 September 2013, pp.3231–3247. Manassas, VA: The Institute of Navigation.

9. Van Dierendonck A, Fenton P and Ford T. Theory and performance of narrow correlator spacing in a GPS receiver. *J Navig* 1992; 39: 265–283.

10. Braasch MS. GPS multipath model validation. In: *Proceedings of the position, location and navigation symposium (PLANS’96)*, Atlanta, GA, 22–25 April 1996, pp.672–678. New York: IEEE.

11. Sahmoudi M, Amin MG and Landry R. Acquisition of weak GNSS signals using a new block averaging pre-processing. In: *Proceedings of the 2008 IEEE/ION position, location and navigation symposium*, Monterey, CA, 5–8 May 2008, pp.1362–1372. New York: IEEE.

12. Hsu LT, Gu Y and Kamijo S. NLOS correction/exclusion for GNSS measurement using RAIM and city building models. *Sensors* 2015; 15: 17329–17349.

13. Ziedan NI. Multipath channel estimation and pattern recognition for environment-based adaptive tracking. In: *Proceedings of the ION GNSS 2012*, Nashville, TN, 17–21 September 2012, pp.394–407, https://www.researchgate.net/publication/288046169_Multipath_Channel_Estimation_and_Pattern_Recognition_for_Environment-Based_Adaptive_Tracking

14. Zhong M, Guo J and Cao Q. On designing PMI Kalman filter for INS/GPS integrated systems with unknown sensor errors. *IEEE Sens J* 2014; 15: 535–544.

15. Chiang KW, Duong TT and Liao JK. The performance analysis of a real-time integrated INS/GPS vehicle navigation system with abnormal GPS measurement elimination. *Sensors* 2013; 13: 10599–10622.

16. Gu Y, Hsu LT and Kamijo S. GNSS/onboard inertial sensor integration with the aid of 3-D building map for lane-level vehicle self-localization in urban canyon. *IEEE Trans Veh Technol* 2015; 65: 4274–4287.

17. Blanco-Delgado N and de Haag MU. Multipath analysis using code-minus-carrier for dynamic testing of GNSS receivers. In: *Proceedings of the 2011 international conference on localization and GNSS (ICL-GNSS)*, Tampere, 29–30 June 2011, pp.25–30. New York: IEEE.

18. Smolyakov I, Rezaee M and Langley R. Resilient multipath prediction and detection architecture for low-cost navigation in challenging urban areas. *J Navig* 2020; 67: 397–409.

19. Sun R, Wang G, Zhang W, et al. A gradient boosting decision tree based GPS signal reception classification algorithm. *Appl Soft Comput* 2020; 86: 105942.

20. Wang L, Groves PD and Ziebart MK. Smartphone shadow matching for better cross-street GNSS positioning in urban environments. *J Navig* 2015; 68: 411–433.

21. Ahmad KAB, Sahmoudi M, Macabiau C, et al. Reliable GNSS positioning in mixed LOS/NLOS environments using a 3D model. In: *Proceedings of the European navigation conference (ENC)*, Vienna, 23–25 April 2013, pp.1–9. Bangalore, India: HAL.

22. Phan QH, Tan SL, McLoughlin I, et al. A unified framework for GPS code and carrier-phase multipath mitigation using support vector regression. *Adv Artif Neural Syst* 2013; 2013: 240564.

23. Phan QH, Tan SL and McLoughlin I. GPS multipath mitigation: a nonlinear regression approach. *GPS Sol* 2013; 17: 371–380.

24. Yozevitch R, Moshe BB and Weissman A. A robust GNSS LOS/NLOS signal classifier. *J Navig* 2016; 63: 429–442.

25. Sun R, Hsu LT, Xue D, et al. GPS signal reception classification using adaptive neuro-fuzzy inference system. *J Navig* 2019; 72: 685–701.

26. Peyraud S, Bétaille D, Renault S, et al. About non-line-of-sight satellite detection and exclusion in a 3D map-aided localization algorithm. *Sensors* 2013; 13: 829–847.

27. Groves PD. Shadow matching: a new GNSS positioning technique for urban canyons. *J Navig* 2011; 64: 417–430.

28. Sánchez JS, Gerhmann A, Thevenon P, et al. Use of a FishEye camera for GNSS NLOS exclusion and characterization in urban environments. In: *Proceedings of the 2016 international technical meeting of the institute of navigation*, Monterey, CA, 25 January 2016. Manassas, VA: The Institute of Navigation.

29. Xu H, Angrisano A, Gaglione S, et al. Machine learning based LOS/NLOS classifier and robust estimator for GNSS shadow matching. *Satell Navig* 2020; 1: 15.

30. Zhou ZH. Ensemble learning. In: Li SZ and Jain A (eds) *Encyclopedia of biometrics*. Boston, MA: Springer, 2009, pp.270–273.

31. Breiman L. Bagging predictors. *Mach Learn* 1996; 24: 123–140.

32. Bühlmann P and Yu B. Analyzing bagging. *Ann Stat* 2002; 30: 927–961.

33. Meng Q, Liu JY, Zeng QH, et al. Neumann-Hoffman code evasion and stripping method for BeiDou software-defined receiver. *J Navig* 2017; 70: 101–119.