Proceedings of the 2021 AIAA/IEEE 40th Digital Avionics Systems Conference (DASC), 3-7 October, San Antonio, USA
DOI:10.1109/DASC52595.2021.9594434

A Machine Learning Based GNSS Performance Prediction for Urban Air Mobility Using Environment Recognition

Oguz Kagan Isik
Cranfield University
Bedfordshire, UK
o.isik@cranfield.ac.uk

Ivan Petrunin
Cranfield University
Bedfordshire, UK
i.petrunin@cranfield.ac.uk

Gokhan Inalhan
Cranfield University
Bedfordshire, UK
inalhan@cranfield.ac.uk

Antonios Tsourdos
Cranfield University
Bedfordshire, UK
a.tsourdos@cranfield.ac.uk

Ricardo Verdeguer Moreno
Spirent Communications PLC
Crawley, UK
ricardo.verdeguermoreno@spirent.com

Raphael Grech
Spirent Communications PLC
Crawley, UK
raphael.grech@spirent.com

***Abstract* — As the primary navigation source, GNSS performance monitoring and prediction have critical importance for the success of mission-critical urban air mobility and cargo applications. In this paper, a novel machine learning based performance prediction algorithm is suggested considering environment recognition. Valid environmental parameters that support recognition and prediction stages are introduced, and K-Nearest Neighbour, Support Vector Regression and Random Forest algorithms are tested based on their prediction performance with using these environmental parameters. Performance prediction results and parameter importances are analyzed based on three types of urban environments (suburban, urban and urban-canyon) with the synthetic data generated by a high quality GNSS simulator.**

Keywords* — *GNSS, machine learning, performance prediction, environment recognition, environment classification, integrity, urban air mobility.

I. Introduction

Global Navigation Satellite System (GNSS) is the primary navigation source when available, thanks to its high accuracy, global coverage and low cost. However, its navigation performance can considerably vary based on the particular operational environment due to GNSS is vulnerable to Line-of-Sight (LoS) blockage, bad satellite geometry and multipath reflections [1].

With the growing usage of Unmanned Aerial Systems (UASs) for mission-critical applications, reliable UAS positioning and timing in GNSS challenging or denied environments becomes increasingly important. Performance requirements to guarantee reliability has already been defined for aircraft by aviation authorities with Performance Based Navigation (PBN) concept [2], but the requirements and regulations for UASs are still under investigation and mission dependent. As a most recent example, operational requirements of UASs [3] were published by Civil Aviation Authority (CAA) for supporting medical Urban Air Mobility and Cargo (UAM/C) missions during the COVID-19 pandemic.

Especially in urban areas, meeting safety requirements is very challenging. In UAM/C use cases, UASs are envisioned to operate in a mixture of urban and suburban environments, where positioning, navigation and timing (PNT) accuracy, availability and continuity from GNSS could vary significantly. Thus, it is anticipated that either certain flight zones would need to be avoided, considering GNSS performance, or different means of navigation aiding would need to be active in such zones to ensure desired navigation performance.

In the literature, integrity monitoring has been researched without considering environment types or urbanization levels [4][5]. These works aim to specify general integrity criteria with respect to satellite geometry, satellite visibility, signal strength etc., in order to ensure reliable GNSS performance. Considering GNSS performance differentiates based on environment, some other studies examine how integrity parameters such as Dilution of Precision (DOP), number of visible satellites, cut-off elevation, Signal-to-Noise Ratio (SNR) change in different environments [6][7][8]. Results show that while the impact of LoS obscuration on GNSS accuracy and integrity can be specified with DOP, number of visible satellites and cut-off elevation, SNR is useful to detect multipath effects.

Based on the assumption that these parameters change according to the environmental conditions and represents the effect of different error sources, the idea of environment recognition has emerged. An environment recognition algorithm is suggested in [9] to distinguish the five most common urban environment types based on GNSS blockage and signal strength. Another algorithm for determining LoS blockage characteristics around a railway is presented in [10] with clustering GNSS elevation and azimuth. And in [11], GNSS SNR data is classified for different environment types based on elevation angles.

This study aims to investigate and develop a novel machine learning based GNSS performance prediction method using environment recognition for UAM vehicles operating in urban environments. The GNSS performance prediction is a key enabling technology for the design of flight trajectories and the development of geo-fenced or segregated airspaces considering predicted operational PNT accuracy and integrity.

The remainder of the paper is organized as follows: in Section II, selected urban environment specifications are introduced, and methods for environment recognition are discussed. Section III presents the methodology of the performance prediction algorithm and simulation setup. Results are given and discussed in Section IV. And finally, Section V summarize the results, and a brief conclusion is given.

II. Environmental Awareness

GNSS performance can be predicted by monitoring GNSS data collected by the receiver. Additionally, some environmental parameters can enhance the prediction performance if applicable. However, the error model of a GNSS receiver is not the same in different environments. Thus, environment awareness is crucial in order to produce better error models. For this purpose, in this section, the environment types that cause differentiation in error models will be defined first. Then environment recognition approach

© 20XX IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

and the parameters that can be used for GNSS performance prediction will be specified.

A. Environment Types

Different environment types can be defined based on the density of an urban area. The most common environment types used in the literature are open-sky, suburban, semi-urban, urban canyon and forest, considering LoS blockage and multipath reflections [9]. As we are interested in predicting GNSS performance in urban areas, three urbanization levels are selected and defined considering cut-off elevation and multipath effects:

1) Suburban: Defined as a settlement area consisting mostly of 2 or 3-storey houses and cut-off elevation angle is maximum of 10 degrees. The multipath effect is too low.

2) Urban: Defined as a settlement area covered with apartments and cut-off elevation angle is 10 to 30 degrees. Multipath reflections have more impact than suburban areas.

3) Urban-Canyon: An area with lots of very tall buildings sheathed with highly reflective materials. Cut-off elevation is 20 to 60 degrees with dense multipath impact.

Environment types are considered as defined above while generating environment-specific data, and GNSS performance prediction tests are conducted with this data set.

B. Environment Recognition

Based on the definitions of environment types, it can be seen that the key separators between the environments are satellite visibility and multipath reflections. While DOP, number of visible satellites, cut-off elevation angle gives information about satellite visibility and geometry, variation in SNR is beneficial to detect multipath reflections. These observables have been used in the literature in two ways. The first one is the direct usage of observables recorded by the receiver [10]. In this case, inputs of classification or clustering algorithms are:

- DOP
- Number of visible satellites
- Elevation angles of satellites
- Azimuth angles of satellites
- SNR of satellites

As all these data are directly readable in navigation messages by a receiver, this approach is applicable for in-flight recognition.

The second way is to use relative values of observables. In this approach, values of observables under optimum conditions (open sky, no multipath) should be known to calculate differences or rates based on the values from these optimum conditions. Suggested environment classification inputs are [9]:

- DOP ratio
- Blockage coefficient
- Signal strength attenuation
- Signal strength fluctuation coefficient

As it is investigated in [9], environment recognition results are significantly better with using relative values as input of classification algorithm. However, this approach is not suitable for in-flight recognition unless a calibration flight is held in advance.

In this study, the data is generated under predefined environment conditions with a GNSS simulator. Hence, we ensure that environment recognition errors do not affect GNSS performance prediction results. The training data generation stage is explained in Methodology, and the characteristics of training data are analyzed in Results.

III. Methodology

A. Simulation Setup

Environment types and impact of error sources are unpredictable in a real test environment, and producing a realistic error model requires a high amount of data. Even the required amount of data can be collected from clearly differentiated environments, calculating positioning errors is not possible due to ground truths are unknown. Therefore, a GNSS Simulator was preferred to simulate environments and generate GNSS signals. With Spirent GSS7000 GNSS simulator, signals from each satellite can be generated considering defined environment types, atmospheric effects, obscuration levels and multipath reflections. Suburban, urban and urban-canyon environment scenarios are defined in the simulator as explained in Section II, and L1 frequency signals are generated for only GPS constellation for 12 hours (one orbital cycle of GPS satellites). Simulated environments are illustrated in Fig. 1-3.

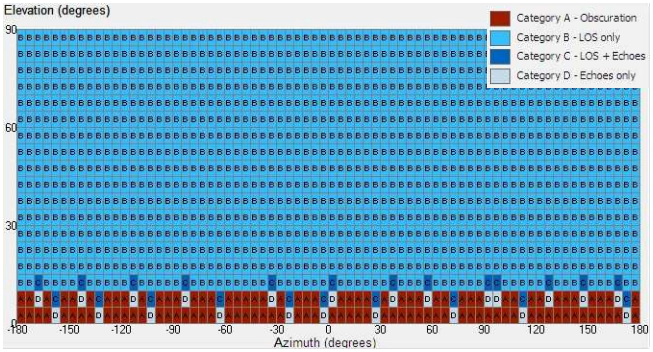


Fig. 1. Simulated Suburban Environment

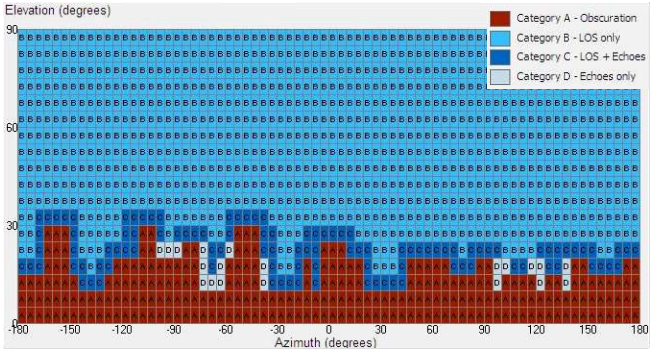


Fig. 2. Simulated Urban Environment

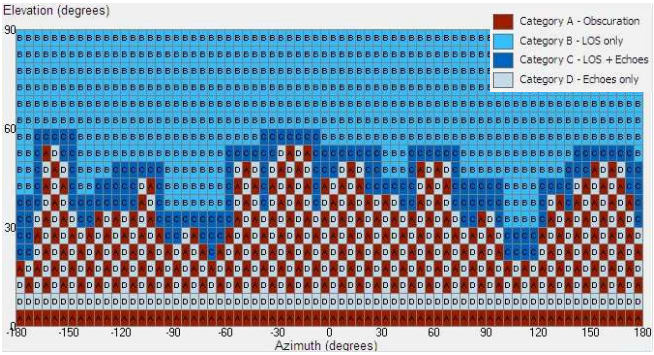


Fig. 3. Simulated Urban-canyon Environment

Generated signals are given to a low-cost GNSS receiver, u-blox NEO-M8T, to observe positioning errors for each environment. Position errors are calculated from the difference between the receiver's position estimation and ground truth of the simulator. HDOP, VDOP, satellite IDs, elevations and signal strengths are obtained via extracting NMEA outputs from the receiver. Then, extracted data are processed in Python and MATLAB in order to train GNSS performance predictors with supervised learning algorithms.

B. Data Processing

After NMEA data extraction stage, the raw data need to be pre-processed before being used in training. Initially, recorded data samples where the number of visible satellites is less than four are removed because the position cannot be estimated with less than four satellites by the receiver. The remaining data are used for generating training inputs. According to the literature discussed in Section II, the following parameters are selected as training inputs:

- HDOP/VDOP
- Number of visible satellites
- Minimum elevation angle of visible satellites (deg)
- Mean signal strength of visible satellites (dB)

These parameters are easy to collect and calculate during the flight as well as beneficial to build environment-related error models.

After unavailable parts are removed and selected training parameters are calculated, the parameters are normalized with respect to their mean (μ_x) and standard deviation (σ_x), in order to enhance training performance, as below:

$$X_{norm} = (X - \mu_x) / \sigma_x \quad (1)$$

Three supervised learning algorithms, K-Nearest Neighbours (KNN), Support Vector Regressor (SVR) and Random Forest, have been trained with normalized parameters for each environment type. Additionally, in order to ensure each learning algorithm are optimized, hyperparameters of each learning model are tuned with grid search.

A graphical representation of the prediction architecture is shown in Fig 4.

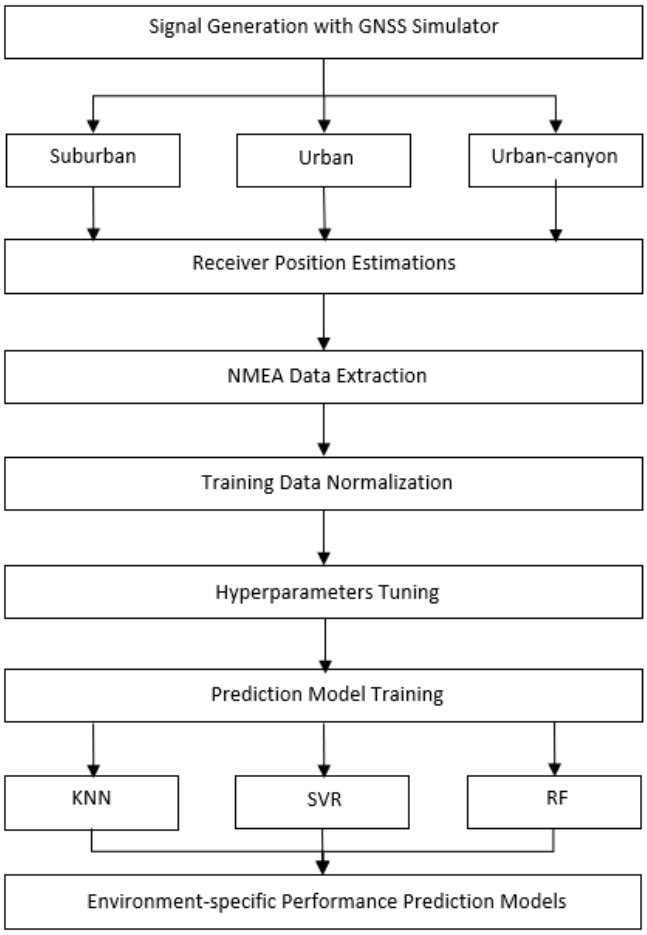


Fig. 4. GNSS Performance Prediction Architecture

IV. Results

A. Training Data Characteristics

Before pursuing GNSS performance prediction results, initially, training data characteristics were investigated. Thanks to the simulated data, we are able to show exact differences in horizontal and vertical error characteristics based on different environments. The variation of each training parameter can be seen as well. Horizontal and vertical error characteristics of training data are statistically analyzed and summarized in Table I for each environment.

TABLE I. Training Data Error Characteristics

Environment Type	Horizontal Absolute Error (m)		Vertical Absolute Error (m)	
	μ	σ	μ	σ
Suburban	15.48	36.08	20.11	1.40
Urban	82.64	88.76	23.16	4.09
Urban-Canyon	1076.47	750.52	65.09	22.40

According to the results, it is hard to build a general error model for GNSS navigation that suits all the environment types efficiently as error characteristics vary dramatically with respect to environmental conditions. This is the idea behind using environment recognition for supporting GNSS error prediction. Especially in denser urban areas like urban-

canyon, either horizontal and vertical characteristics are much worse than suburban or urban areas.

Now we would like to investigate the environmental training parameters to see whether their variations are similar to positioning error variation. Training parameter characteristics are statistically analyzed and summarized in Table II for each environment.

TABLE II. TRAINING PARAMETER CHARACTERISTICS

Parameters	Suburban		Urban		Urban-Canyon	
	μ	σ	μ	σ	μ	σ
HDOP	0.87	0.14	1.17	0.28	1.47	1.39
VDOP	1.51	0.17	2.21	0.82	2.00	1.87
Number of Visible Satellites	9.98	1.35	8.11	1.44	7.65	1.38
Minimum Elevation Angle (º)	11.62	2.77	16.34	4.83	9.26	3.97
Mean Signal Strength (dB)	45.59	0.66	45.86	0.99	45.16	2.18

Based on the characteristics of training parameters, a logical variation can be observed between suburban and urban areas as a result of the difference in LoS blockage levels. In urban environment, DOPs and minimum elevation angle is higher than suburban, while number of visible satellites is lower. Nevertheless, in the urban-canyon area, although there is a higher building obscuration, minimum elevation angles are less because reflected signals still can be detected and mislead the receiver. Besides, mean signal strength does not vary much with respect to environments, but the standard deviation of mean signal strength becomes higher in denser urban areas.

As an overall discussion based on these characteristics, error variation between suburban and urban areas is mainly caused by the effect of LoS blockage, and impact of multipath reflection is low. However, in urban-canyon environment, because of the high impact of multipath reflections in addition to blockage, the GNSS error model differs significantly, and it is expected that an overall error predictor would not work properly under this kind of environment with high multipath reflections. Thus, environment recognition can enhance the error prediction performance in urban environments as the environment type can be considered before building an error prediction model with a learning algorithm.

B. Error Prediction Performance

In order to show the impact of environments on both horizontal and vertical error prediction performance, three different error models were trained by K-Nearest Neighbour (KNN), Support Vector Regression (SVR) and Random Forest (RF) learning algorithms, with simulated environment specific data. Prediction results were analyzed with respect to Mean Absolute Error (MAE) and Coefficient of Determination (R^2). While MAE represents the prediction error of the trained model, R^2 is a parameter to show how well the model fits the real data, and its value is between 0 and 1. Table III summarize the prediction performance results of horizontal and vertical error for each environment type and each learning algorithm.

TABLE III. PREDICTION PERFORMANCE RESULTS

Environment Type	Learning Algorithm	Horizontal		Vertical	
		MAE (m)	R^2	*MAE (m)*	R^2
Suburban	***KNN***	6.35	0.67	0.16	0.87
	SVR	10.57	0.07	0.37	0.43
	RF	6.22	0.68	0.15	0.89
Urban	***KNN***	19.22	0.79	0.55	0.90
	SVR	48.58	0.33	1.85	0.39
	RF	18.63	0.84	0.41	0.95
Urban-Canyon	***KNN***	181.39	0.79	5.10	0.79
	SVR	496.01	0.16	15.44	0.17
	RF	144.45	0.85	3.67	0.88

When the results are compared based on learning algorithms, RF algorithm gives the best performance for both horizontal and vertical error prediction. However, prediction errors getting higher with increasing urbanization regardless of algorithm. It means that the same performance cannot be reached with the same training parameters for different environments. Therefore, looking at the importance of parameters during the training of RF algorithm would be a good idea to see which one has more impact on prediction performance for each environment type. Table IV and V shows the parameter importance for horizontal and vertical error prediction, respectively.

TABLE IV. PARAMETER IMPORTANCES FOR HORIZONTAL ERROR PREDICTION

Environment Type	Relative Importance			
	HDOP	*Number of Visible Satellites*	*Minimum Elevation Angle*	*Mean Signal Strength*
Suburban	0.54	0.03	0.34	0.09
Urban	0.41	0.15	0.37	0.07
Urban-Canyon	0.56	0.11	0.22	0.11

TABLE V. PARAMETER IMPORTANCES FOR VERTICAL ERROR PREDICTION

Environment Type	Relative Importance			
	VDOP	*Number of Visible Satellites*	*Minimum Elevation Angle*	*Mean Signal Strength*
Suburban	0.64	0.07	0.25	0.04
Urban	0.53	0.11	0.33	0.03
Urban-Canyon	0.46	0.17	0.23	0.14

The parameter importance analysis shows that the most important parameters are DOP coefficients and minimum

elevation angle. And then, it shows the importance of satellite geometry on prediction performance. These two parameters are much more effective than the other two parameters, especially under low LoS blockage and multipath effect. In order to discuss the impact of LoS blockage independent from multipath reflections, it is better to compare suburban and urban environments. Based on this comparison, while importances of DOP coefficients are decreasing under higher LoS blockage, importances of number of visible satellites and minimum elevation angle increase. This situation can be explained by that the position error mostly depends on their sight angle and geometry if more than enough satellites are visible. Besides, mean signal strength has a very limited effect and does not chance with blockage level if multipath effects are not intense.

On the other hand, if we compare urban and urban-canyon environments to see the impact of multipath on GNSS error prediction, it can be seen that especially the impact of mean signal strength parameter getting higher in urban-canyon environment with increasing intensity of multipath effects. Because signal strengths fluctuate as a result of reflected GNSS signals and even the signals from blocked satellites can arrive at the receiver weaker than normal strength. Moreover, the information collected from these signals misleads the training algorithm in urban-canyon environments as non-visible satellites are counted as visible by the receiver. It also causes miscalculations for minimum elevation angle and DOP values and degrades the error prediction performance in urban-canyon environment.

V. Conclusion

In this study, the benefits of environment recognition on GPS performance prediction were investigated with synthetic data generated by a GNSS simulator. In this context, firstly, potential challenging environment types were defined concerning LoS blockage and multipath intensity. Suitable environment recognition parameters were specified as inputs of learning algorithms. Then, the simulated environment-specific data were analyzed statistically, and error characteristics were summarized with the characteristics of training parameters. And lastly, GNSS error predictors were trained with KNN, SVR and RF learning algorithms, and their performance and importance of training parameters were assessed with respect to the environments.

According to the results, RF has the best error prediction performance with suggested parameters for all environments. However, when we check the training efficiency based on environment types, the performance of RF algorithm degraded, especially in urban-canyon environment. Parameter importance analysis shows that while the selected parameters work well in the training algorithm to predict errors sourced by LoS blockage, performance prediction model cannot be trained appropriately where multipath errors are dominant. This issue can be solved with additional training parameters which show similar variation characteristics with multipath errors.

Overall, this research shows that GNSS performance can be predicted better by considering environment type thanks to environment recognition as the GNSS error models varies based on environment types. Different training parameters can be selected according to dominant error sources in different environments. Additionally, with environment recognition approach, error range in a certain environment can be specified. This would be useful for mission-critical applications to decide whether the flight environment is risky based on mission requirements.

Acknowledgment

Authors would like to thank Spirent Communication PLC for their support in technical aspects.
Oguz Kagan Isik presents his special thanks to Republic of Turkey, Ministry of National Education as a funded PhD student.

References

[1] J. Zidan, E. I. Adegoke, E. Kampert, S. A. Birrell, C. R. Ford, and M. D. Higgins, “GNSS Vulnerabilities and Existing Solutions: A Review of the Literature,” IEEE Access, early access, vol. 4, 2020.

[2] ICAO, “Performance-based Navigation (PBN) Manual,” 2013.

[3] Civil Aviation Authority, “CAP1915, Unmanned Aircraft Systems BVLOS Operations in Support of the COVID-19 Response – Requirements , Guidance & Policy,” 2020.

[4] R. Sabatini, T. Moore, and C. Hill, “A new avionics-based GNSS integrity augmentation system: Part 1 - Fundamentals,” J. Navig., vol. 66, no. 3, pp. 363–384, 2013.

[5] N. Zhu, J. Marais, D. Betaille, and M. Berbineau, “GNSS Position Integrity in Urban Environments: A Review of Literature,” IEEE Trans. Intell. Transp. Syst., vol. 19, no. 9, pp. 2762–2778, 2018.

[6] M. Tahsin, S. Sultana, T. Reza, and M. Hossam-E-Haider, “Analysis of DOP and its preciseness in GNSS position estimation,” in 2nd International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT, 2015.

[7] J. Januszewski, “Sources of Error in Satellite Navigation Positioning,” TransNav, Int. J. Mar. Navig. Saf. Sea Transp., vol. 11, no. 3, pp. 419–423, 2017.

[8] O. K. Isik, J. Hong, I. Petrunin, and A. Tsourdos, “Integrity analysis for GPS-based navigation of UAVs in urban environment,” Robotics, vol. 9, no. 3, pp. 1–20, 2020.

[9] Y. Wang et al., “Urban environment recognition based on the GNSS signal characteristics,” Navig. J. Inst. Navig., vol. 66, no. 1, pp. 211–225, 2019.

[10] J. Tan, J. Wang, and D. Lu, “GNSS data driven clustering method for railway environment scenarios classification,” Proc. 14th IEEE Conf. Ind. Electron. Appl. ICIEA 2019, pp. 2026–2031, 2019.

[11] R. U. R. Lighari, M. Berg, E. T. Salonen, and A. Parssinen, “Classification of GNSS SNR data for different environments and satellite orbital information,” in 11th European Conference on Antennas and Propagation, EUCAP 2017, 2017, pp. 2088–2092.



https://dspace.org/

A machine learning based GNSS performance prediction for urban air mobility using environment recognition

Isik, Oguz Kagan; Petrunin, Ivan; Inalhan, Gokhan; Tsourdos, Antonios; Moreno, Ricardo Verdeguer ...

2021-11-15
IEEE
https://doi.org/10.1109/DASC52595.2021.9594434
Isik OK, Petrunin I, Inalhan G, et al., (2021) A machine learning based GNSS performance prediction for urban air mobility using environment recognition. In: 2021 AIAA/IEEE 40th Digital Avionics Systems Conference (DASC), 3-7 October 2021, San Antonio, USA

Downloaded from a DSpace repository. https://dspace.org/