# Explainable Machine Learning for Environment-Aware Channel State Prediction in UAV-Based 6G Networks


Ladan Gholami
*dept. Information Engineering*
*University of Pisa*
Pisa, Italy
ladan.gholami@phd.unipi.it

Pietro Ducange
*dept. Information Engineering*
*University of Pisa*
Pisa, Italy
pietro.ducange@unipi.it

Arcangela Rago
*dept. Electrical and Information Engineering*
*Politecnico di Bari*
Bari, Italy
*CNIT, Consorzio Nazionale Interuniversitario per le Telecomunicazioni* Italy
arcangela.rago@poliba.it

Pietro Cassará
*Institute of Information Science and Technology*
*National Research Council (CNR)*
Pisa, Italy
pietro.cassara@isti.cnr.it

Alberto Gotta
*Institute of Information Science and Technology*
*National Research Council (CNR)*
Pisa, Italy
alberto.gotta@isti.cnr.it



***Abstract*—The emergence of 6G networks demands environment-aware communication paradigms to ensure reliable and efficient connectivity, and Channel Knowledge Maps (CKMs) offer a promising solution by mapping spatial locations to detailed channel characteristics for proactive network optimization. In this context, this paper proposes an explainable Machine Learning (ML)-based framework that uses geometrical features to predict receiver state probabilities in UAV-based mmWave communication networks. Geometrical characteristics extracted from the environment surrounding each receiver are used to train ML models, namely Decision Tree (DT), K-Nearest Neighbors (KNN), and Deep Neural Network (DNN) models, to predict three receiver states probabilities: Line-of-Sight (LOS), No-Line-of-Sight (NLOS), and Blocked. Experimental results show that the DNN model outperforms DT and KNN, achieving higher accuracy across all states, albeit with no inherent explainability. To address this, the SHapley Additive exPlanations (SHAP) method is applied to indicate feature contributions to each state prediction of the black-box DNN model. This improves the interpretability and reliability of the proposed environment-aware framework for 6G UAV-based networks.**



***Index Terms*—Channel Knowledge Map (CKM), Explainability, Geometrical Features, mmWave, Receiver State Prediction.**


## I. Introduction

The sixth-generation (6G) wireless networks are envisioned to revolutionize network connectivity to deliver ultra-high data rates, low latency, and ubiquitous coverage, enabling applications such as autonomous vehicles, tactile internet, and mixed-reality systems. To achieve these ambitious goals, environment-aware communication paradigms are emerging as key enablers, emphasizing the use of a priori knowledge about the actual surrounding environment for the design and optimization of communication networks. One approach to enable such awareness is the Channel Knowledge Map (CKM), a site-specific database that provides location-tagged channel characteristics. It allows networks to predict channel conditions proactively at any spatial point, not using solely real-time measurements. By leveraging CKM, future networks can achieve light-training communication, significantly reducing pilot overhead, especially in mmWave Unmanned Aerial Vehicle (UAV)-based deployments, where channel estimation is challenging due to mobility and blockage dynamics [1], [2].

In UAV-based mmWave communication systems, accurately determining the state of ground receivers is essential for optimizing communication quality. This arises from the inherent mobility of UAV platforms and the vulnerability of mmWave signals to attenuation and blockage by environmental obstacles, which makes receiver states highly dynamic, particularly in heterogeneous urban environments. Traditional probabilistic models, mostly estimating LOS probabilities, such as those based on distance, often fail to capture the fine-grained spatial variability in real urban environments. By integrating LOS/NLOS/Blocked probability models within the CKM framework, we develop an environment-informed probabilistic channel state model, enhancing environment awareness for improving Quality of Service (QoS), blockage avoidance, and trajectory optimization in UAV-based networks [1].

The majority of existing studies focus solely on estimating LOS probability, which can be classified into empirical [3]–[6], deterministic methods such as Ray Tracing (RT), and geometry-based models [7]–[9]. For instance, Zhu et al. [8] proposed a geometry-based model that incorporates variations in transmitter and receiver height alongside urban parameters such as building height distributions, spacings, and widths, while considering the Fresnel zone. In another study, Saboor et al. [9] developed a model that captures the influence of User Equipment (UE) positions along streets, integrating elevation and azimuth angles, as well as building and street dimensions, within an artificially constructed urban grid using ITU-defined parameters. Pang et al. [5] employed a hybrid approach combining K-Nearest Neighbors (KNN) and a neural network, both trained on RT simulation datasets; however, their reliance on synthetically generated cities limits their applicability. Similarly, the work [6] utilized a Graph Neural Network to predict LOS probability while accounting for UAV mobility, but their model also suffers from constraints imposed by synthetic data. In particular, the work [10] was the first to introduce probabilistic models covering LOS, NLOS, and Blocked states, designed primarily for terrestrial networks; however, their formulations do not extend to UAV-based scenarios as they overlook urban geometry and transmitter height effects. Overall, most UAV communication models simplify environmental representation and rely primarily on distance-based features, neglecting the real, intricate geometrical characteristics of urban environments. As a result, the primary aim of this study is to estimate the probabilities of LOS, NLOS, and Blocked states for UEs at various locations within a large-scale UAV cell using geometrical characteristics describing the surrounding areas. Unlike prior models, the proposed approach divides the environment into equal-sized spatial sections, each characterized by distinct geometrical or morphological features, depicting city districts, enabling a more comprehensive estimation of channel state probabilities across the deployment area.

In this study, after presenting the adopted methodology (Section II), we explore the use of Machine Learning (ML) techniques, namely Decision Tree (DT), Deep Neural Networks (DNN), and KNN, to map environmental features to probabilistic channel states in Air-to-Ground (A2G) communication and then compare their predictive performance (Section III). Furthermore, we perform an explainability analysis through SHapley Additive exPlanations (SHAP) to gain insight into the decision-making processes of the models, utilizing data from Florence city to evaluate their generalization ability (Section IV). Finally, we conclude the paper, summarizing the main contributions and outlining directions for future work.

## II. Methodology

This study uses urban 3D maps from ten different cities, each covering a $1\text{km}^2$ area. The selected cities include Dubai, Florence, London, Madrid, Munich, New York, Ottawa, Paris, Tokyo, and Vienna, offering various urban layouts ranging from dense metropolitan, high-rise deployment to typical Manhattan-like districts. This diversity ensures that trained models are exposed to a wide variety of structural layouts and environmental conditions, enhancing their generalization capability across different deployment scenarios.

The 3D city models are acquired from OpenStreetMap [11] and converted into DXF files using Blender software. The DXF format is chosen for 3D representation of urban sections due to its compatibility with Wireless InSite, as well as its ability to preserve detailed geometrical representations with accurate XYZ vertex coordinates. Unlike other common 3D formats, DXF files retain complex building geometries without simplification, ensuring high fidelity in urban structure representation. The 3D city maps maintain their real-world geographic orientation throughout the preprocessing pipeline, without any rotations, preserving true spatial alignment.

For propagation modeling, we employ *Wireless InSite*, a high-precision RT simulator by REMCOM [12], renowned for its accuracy for 3D propagation simulations. The UAV transmitters are placed in the center at 150 meters height above terrain, operating at a 28 GHz mmWave carrier frequency with 30 dBm output power, employing an omnidirectional antenna oriented downward to provide full coverage of the lower hemisphere. The building walls are modeled as solid concrete. The receivers are located with 20-meter spacing from each other, forming a grid, at an altitude of 2 meters from the ground surface, resulting in 2500 receivers per each city cell, ensuring complete spatial coverage. The noise figure of each receiver is 3 dB. Each UAV city cell is partitioned into equal square sections of $100\text{m}^2$, as this area can provide an optimal trade-off between feature consistency and resolution granularity, and 25 receivers are arranged in a 5x5 grid inside each section.

To construct the feature set for inputs of the ML-based channel state model, we extracted a set of geometrical descriptors characterizing each square section:

- *3D Distance:* The direct Euclidean distance from the UAV transmitter to the center of each section, incorporating both horizontal displacement and vertical height differences. This feature also encapsulates UAV elevation and terrain elevation.
- *Number of Buildings:* Represents the total count of distinct building objects within each section, related to region density.
- *Building Height Statistics:* Includes *minimum, maximum, mean, standard deviation, variance, and median* of building heights. These statistics describe the vertical distribution profile, revealing patterns such as whether the area

contains mostly tall structures, sparse short buildings, or a mixture of low-rise buildings with occasional towers. We also extract the *number of buildings above average height*, which quantifies how many buildings exceed the mean height, showing dominant tall structures that may block or reflect signals.

- *Building Density Ratio:* The fraction of the total section area occupied by building footprints, capturing urban compactness and man-made footprint. Moreover, the feature *unoccupied area* has a complementary role for this feature.
- *Number of 3D Faces and Vertices:* Measures the geometrical complexity of building facades within each section. Higher counts reflect more intricate architectural designs, which influence scattering, diffraction, and non-specular reflections.

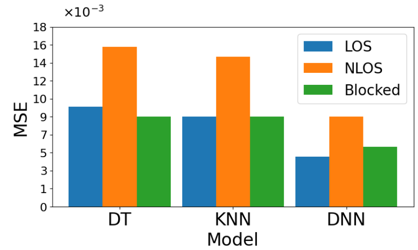
Each feature is selected to comprehensively characterize the spatial, structural, and morphological attributes of the environment, enabling the ML models to learn complex relationships between environmental geometry and channel state probabilities.

The Wireless InSite parameters include up to six reflections and one diffraction, capturing dominant propagation paths in urban environments. No transmission through buildings is allowed, reflecting the high penetration losses of mmWave frequencies.

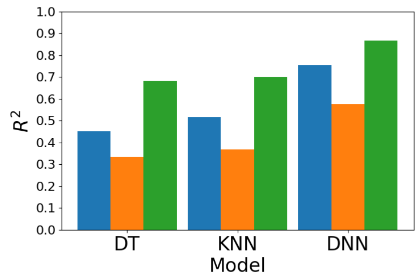
ML techniques have demonstrated strong capabilities in solving both classification and regression problems across numerous scientific fields, owing to their capability in identifying hidden patterns within the data [13], and have also shown potential in improving UAV systems for tasks such as detection, security and surveillance [14]. Instead of proposing analytical closed-form formulas, this study employs ML approaches to model the complex relationship between geometrical features and receiver state probabilities as a multi-output regression problem. Among the approaches tested, DT, KNN, and DNN [15] are selected for evaluation. DT is a non-parametric algorithm that divides data into hierarchical segments based on optimal feature splits, offering intuitive interpretability at the cost of overfitting. KNN operates as an instance-based learner, predicting outputs by averaging the outcomes of the most similar training examples, effectively capturing local data patterns without requiring explicit parameterized training, although it is sensitive to outliers. DNN, consisting of multiple layers with nonlinear activation functions, excels at learning hierarchical and abstract representations, making it well-suited for modeling the complex and nonlinear relationship between environmental and channel state data.

## III. Performance Evaluation

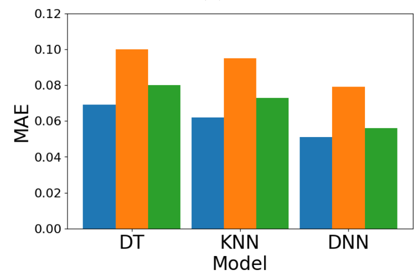
The effectiveness of the proposed ML models was evaluated using three standard regression metrics: coefficient of determination ($R^2$), Mean Squared Error (MSE), and Mean




(a)




(b)




(c)

Fig. 1: Performance evaluation of the selected models on the test set in terms of (a) MSE, (b) $R^2$, and (c) MAE for LOS (blue), NLOS (orange), and Blocked (green) states.

Absolute Error (MAE) [16]. The $R^2$ score measures how well the predicted values approximate the actual data by indicating the proportion of variance explained by the model, and is computed as (1):

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{1}$$

where $\hat{y}_i$ is the predicted value and $\bar{y}$ is the mean of the actual values as: $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$, with $N$ is the total number of data. The MSE calculates the average of squared differences between actual and predicted values, penalizing larger errors more heavily:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2. \tag{2}$$

Additionally, MAE quantifies the average magnitude of prediction errors, offering an interpretable measure of typical

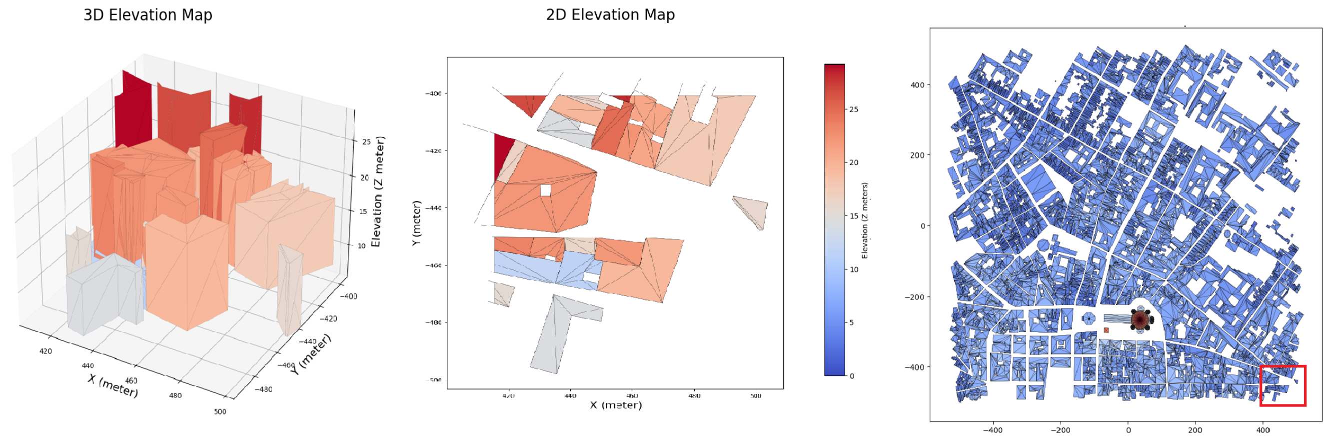



Fig. 2: 3D and 2D Elevation map of the chosen area, with its location on the Florence cell.

prediction error:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i| \,. \tag{3}$$

Fig. 1 summarizes the performance of the models for LOS, NLOS, and Blocked state probabilities using these three evaluation metrics. All models were evaluated using 10-fold cross-validation, and the resulting performance metrics for LOS, NLOS, and blockage probabilities exhibited low variability, with standard deviations of less than 0.1, indicating robust performance across different data partitions.

The results in Fig. 1 indicate that KNN performs moderately well, outperforming DT in all metrics, highlighting its strength in using local patterns. However, its performance remains inferior to that of DNN due to limitations in handling high-dimensional feature interactions. DT, on the other hand, shows the highest prediction errors, reflecting its limited capacity to generalize for high nonlinearity. In general, the DNN model emerges as the most effective approach for the environment-aware channel state prediction model in this study, despite its inherent interpretability challenges.

## IV. Explainability Analysis

In ML-based models, explainability plays a crucial role in understanding how different input features influence the final decisions of the model. Explainability techniques provide insights into feature contributions, enabling verification of whether the behavior of the ML model aligns with physical expectations and ensuring confidence in its deployment for wireless communication applications, particularly for UAV-based networks [13], [17]. One of the most promising eXplainable Artificial Intelligence (XAI) techniques that provide understanding and interpretability for various black-box ML models is the SHAP method [13], [18].

To illustrate the behavior of the model, a SHAP explainability analysis was performed using a region with a relatively dense urban layout of Florence city, i.e., Section 91. Fig. 2 depicts both the 3D and 2D elevation views of this section, highlighting the distribution of buildings and the section position relative to the UAV transmitter.

Fig. 3 shows the SHAP force plots for the specific section under analysis, with the features with positive contributions on the left and the features with negative contributions on the right, respectively, by also indicating the magnitude and direction of the contributions. In particular, the plots reveal the contributions of various geometrical features to the predicted probabilities of LOS, NLOS, and Blocked states. Note that the base value is the model's average prediction across all data, and the bold number shows the final prediction for a specific instance. Red and blue arrows represent features that increase and decrease the prediction, respectively. Moreover, the size of each arrow shows how strong that feature's influence is, and the related value is its SHAP value. For the LOS prediction, a larger unoccupied area (large size of the relative red arrow and high SHAP value) combined with a moderate density ratio increases the likelihood of LOS, while a greater 3D distance (the first blue arrow from the left) to the UAV transmitter significantly decreases this probability. In the case of NLOS prediction, the 3D distance, the number of buildings, and the number of buildings above the average height, i.e., the first three red arrows starting from the right, contribute positively, pushing the output towards higher NLOS prediction, whereas the density ratio (the first blue arrow from the left) exerts a negative influence, helping the Blocked state. Finally, for the Blocked probability, a higher density ratio (the first red arrow starting from the right) increases the likelihood of blockage, but a greater 3D distance and a higher number of buildings (in blue) reduce the predicted probability of the Blocked state. This plot jointly shows the model's understanding of propagation conditions in this urban environment.


Fig. 3: SHAP force plots for the section under analysis.

## V. Conclusions

This paper presents an ML-based framework for environment-aware receiver state prediction for A2G UAV-assisted mmWave communication, leveraging detailed geometrical features extracted from the urban region where the receiver is located. By incorporating morphological attributes into the prediction process, the proposed approach contributes to the broader vision of CKM for proactive and intelligent network optimization. We train and evaluate three state-of-the-art ML models, namely DT, KNN, and DNN, using the extracted features to estimate the probabilities of LOS, NLOS, and Blocked states. The results indicate that the DNN model demonstrated superior predictive performance for all states across all evaluation metrics, highlighting its ability to capture complex relationships between the urban environment and the channel state. However, due to the black-box nature of the DNN model, interpretability remains a challenge. To address this, we employ SHAP to analyze feature contributions and provide transparency into the decision-making process of the model. The explainability analysis reveals meaningful insights into how specific environmental factors affect each of the predicted receiver states. As for the future, we will explore additional explainable artificial intelligence techniques to further enhance model interpretability and support more trustworthy and robust decision-making in complex urban communication scenarios. Moreover, a promising direction involves leveraging high-resolution LiDAR data to enhance the accuracy and adaptability. Integrating LiDAR-based 3D environmental sensing into the proposed framework for UAV systems could enable real-time CKM refinement, thereby improving channel state prediction in dynamic urban environments.

## Acknowledgement

This work is supported by the European Union under the Italian National Recovery and Resilience Plan (PNRR) PE00000001 - program "RESTART", Mission 4 Issue 2 Investment 1.4, and by the HORIZON-CL4-2021- SPACE-01 project "5G+ evoluTion to mutioRbitAl multibaNd neTwORks" (TRANTOR) No. 101081983. Moreover, we acknowledge Tarun Chawla and Remcom Inc. for the Wireless InSite X3D ray tracer used for generating the numerical analysis. Additionally, the contribution of Pietro Ducange to this research has been supported by the Italian Ministry of University and Research (MUR) in the framework of the *FoReLab* and *CrossLab* projects (*Departments of Excellence*) and by the PNRR - M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - *FAIR - Future Artificial Intelligence*

*Research* - Spoke 1 *Human-centered AI*. The contribution of Arcangela Rago has also been supported by the PRIN project HORUS (grant no. 2022P44KA8), funded by the Italian MUR.

## References

[1] Y. Zeng, J. Chen, J. Xu, D. Wu, X. Xu, S. Jin, X. Gao, D. Gesbert, S. Cui, and R. Zhang, "A Tutorial on Environment-Aware Communications via Channel Knowledge Map for 6G," *IEEE communications surveys & tutorials*, vol. 26, no. 3, pp. 1478–1519, 2024.

[2] X. Wang, Y. Shi, T. Wang, Y. Huang, Z. Hu, and Z. Jiang, "Can Channel Knowledge Map Help to Predict Instantaneous MIMO Channel State Information?" in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2024, pp. 1–6.

[3] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 16)," 3GPP, Tech. Rep. TR 138.901, 2020. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/16.01.00_60/tr_138901v160100p.pdf

[4] *WINNER II Channel Models*, Inf. Soc. Technol, 2008. [Online]. Available: https://www.researchgate.net/publication/259900906_IST-4-027756_WINNER_II_D112_v12_WINNER_II_channel_models

[5] M. Pang, Q. Zhu, Z. Lin, F. Bai, Y. Tian, Z. Li, and X. Chen, "Machine Learning based altitude-dependent empirical LOS probability model for Air-to-Ground communications," *Frontiers of Information Technology & Electronic Engineering*, vol. 23, no. 9, pp. 1378–1389, 2022.

[6] F. Ali, Y. Pan, Q. Zhu, N. Ahmad, K. Mao, and H. Ullah, "Height-dependent LOS probability model for A2G channels incorporating airframe shadowing under built-up scenario," *Electronics Letters*, vol. 60, no. 18, p. e70032, 2024.

[7] ITU, "Propagation data and prediction methods required for the design of terrestrial broadband radio access systems operating in a frequency range from 3 to 60 GHz," Tech. Rep. ITU-R P.1410-5, 2023. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1410-6-202308-I!!PDF-E.pdf

[8] Q. Zhu, F. Bai, M. Pang, J. Li, W. Zhong, X. Chen, and K. Mao, "Geometry-based Stochastic Line-of-Sight Probability Model for A2G Channels under Urban Scenarios," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 7, pp. 5784–5794, 2022.

[9] A. Saboor, E. Vinogradov, Z. Cui, A. Al-Hourani, and S. Pollin, "A Geometry-Based Modelling Approach for the Line-of-Sight Probability in UAV Communications," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 364–378, 2023.

[10] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[11] OpenStreetMap contributors, "Planet dump retrieved from https://planet.osm.org ," https://www.openstreetmap.org , 2017.

[12] Remcom, "Wireless InSite 3D Wireless Propagation Software,," 2025, [Online]. Available: https://www.remcom.com/wireless-insite-propagation-software.

[13] H. Sun, Y. Liu, A. Al-Tahmeesschi, A. Nag, M. Soleimanpour, B. Canberk, H. Arslan, and H. Ahmadi, "Advancing 6g: Survey for explainable ai on communications and network slicing," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 1372–1412, 2025.

[14] M. Mrabet, M. Sliti, and L. B. Ammar, "Machine learning algorithms applied for drone detection and classification: benefits and challenges," *Frontiers in Communications and Networks*, vol. 5, p. 1440727, 2024.

[15] C. Zheng, X. Hong, D. Ding, S. Vargaftik, Y. Ben-Itzhak, and N. Zilberman, "In-network machine learning using programmable network devices: A survey," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 2, pp. 1171–1200, 2024.

[16] M. Rahman, H. Sarwar, M. A. Kader, T. Gonçalves, and T. T. Tin, "Review and empirical analysis of machine learning-based software effort estimation," *IEEE Access*, vol. 12, pp. 85 661–85 680, 2024.

[17] T. Senevirathna, V. H. La, S. Marcha, B. Siniarski, M. Liyanage, and S. Wang, "A survey on xai for 5g and beyond security: Technical aspects, challenges and research directions," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 2, pp. 941–973, 2025.

[18] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in neural information processing systems*, vol. 30, 2017.