MULTI-FREQUENCY AIR-TO-GROUND PATH LOSS PREDICTION FOR UNMANNED AERIAL VEHICLES USING MACHINE LEARNING

DỰ ĐOÁN SUY HAO ĐƯỜNG TRUYỀN KHÔNG TRUNG - MẶT ĐẤT ĐA TẦN SỐ CHO UAV SỬ DỤNG HỌC MÁY

Duong Thi Hang *, Pham Duy Phong

Electric Power University

Ngày nhận bài: 13/6/2025, Ngày chấp nhận đăng: 11/8/2025

Abstract:

Unmanned Aerial Vehicles (UAVs) have emerged as a promising solution for modern wireless communication systems due to their flexibility, mobility, and ease of deployment. However, ensuring reliable air-to-ground (A2G) communication requires accurate channel modeling to support efficient power control and system planning. This study proposes a robust path loss prediction framework for A2G communication links using machine learning techniques, specifically the K-Nearest Neighbors (KNN) regression algorithm. The model is trained and evaluated using a publicly available dataset, with a focus on urban environments and tested across three carrier frequencies: 1 GHz, 2 GHz, and 5.8 GHz. Comparative evaluations against conventional A2G models demonstrate that the proposed approach achieves lower standard errors and narrower confidence intervals. These results highlight the model's capability to deliver accurate path loss predictions, underscoring its potential for improving the reliability and performance of UAV-based communication systems, particularly in dense urban scenarios.

Keywords:

UAV communications, air-to-ground channel, path loss prediction, KNN regression, urban environments, machine learning, wireless networks.

Tóm tắt:

Các phương tiện bay không người lái (UAV) đã nổi lên như một giải pháp tiềm năng cho các hệ thống truyền thông không dây hiện đại nhờ tính linh hoạt, khả năng di chuyển và triển khai dễ dàng. Tuy nhiên, để đảm bảo khả năng liên lạc tin cậy giữa không trung và mặt đất (A2G), cần có mô hình kênh truyền chính xác nhằm hỗ trợ kiểm soát công suất truyền và lập kế hoạch hệ thống hiệu quả. Nghiên cứu này đề xuất một khung mô hình dự đoán suy hao đường truyền đáng tin cậy cho các liên kết A2G dựa trên các kỹ thuật học máy, cụ thể là thuật toán hồi quy K-Nearest Neighbors (KNN). Mô hình được huấn luyện và đánh giá trên một bộ dữ liệu công khai, tập trung vào môi trường đô thị và được thử nghiệm với ba tần số sóng mang: 1 GHz, 2 GHz và 5.8 GHz. Kết quả so sánh với các mô hình A2G truyền thống cho thấy phương pháp đề xuất đạt sai số chuẩn thấp hơn và khoảng tin cậy hẹp hơn. Những kết quả này cho thấy khả năng dự đoán suy hao đường truyền chính xác của mô hình, làm nổi bật tiềm năng ứng dụng của nó trong việc nâng cao độ tin cậy và hiệu suất của hệ thống truyền thông dựa trên UAV, đặc biệt là trong môi trường đô thị mật độ cao.

Từ khóa:

Truyền thông UAV, kênh truyền không trung – mặt đất, dự đoán suy hao đường truyền, hồi quy KNN, môi trường đô thị, học máy, mạng không dây.

1. INTRODUCTION

The exponential growth of the Internet of Everything (IoE) has led to surging mobile data demands in 5G and beyond 5G (B5G) wireless networks. Projections estimate that global mobile traffic will reach 5016 exabytes per month by 2030 [1]. To address this, emerging technologies such as Heterogeneous Networks (HetNets), Device-to-Device (D2D) communication, Ultra-Dense Networks (UDNs), and Unmanned Aerial Vehicles (UAVs) have been identified as key enablers for future B5G systems [2].

Among these, UAVs stand out for their flexibility, rapid deployment, and strong line-of-sight (LoS) capabilities. They can support various applications, including emergency recovery, public safety, ultra-reliable low-latency communication (URLLC), and enhanced mobile broadband (eMBB), making them ideal for augmenting terrestrial 5G networks [3].

Accurate modeling of wireless channels is essential for the design of UAV-assisted systems, with **path loss** being a critical parameter that influences received signal strength and network performance [4]. Conventional approaches for path loss modeling include empirical, deterministic, and machine learning (ML)-based methods [5], each offering distinct advantages. Recent studies have proposed path loss models for air-to-ground (A2G) communication using parameters such as carrier frequency, distance, UAV altitude, and ray-tracing-based simulations [6]–[8]. ML-based techniques, including KNN, Regression Trees, and Neural Networks, have shown high potential for modeling complex propagation behavior [9]. Unlike existing studies that focus on single-frequency or theoretical datasets, our study combines physics-based features with KNN regression across multiple frequencies and validates the model using a dense urban dataset. This hybrid and practical approach enhance prediction robustness in real-world UAV deployment scenarios.

This study proposes a KNN-based ML model to predict A2G path loss in dense urban environments. The main contributions are as follows :

(i) An extended feature set is constructed using the free-space path loss model;

(ii) A simple yet effective KNN regression model is applied to predict path loss using a publicly available dataset.

The remainder of this paper is structured as follows : Section 2 presents the problem formulation ; Section 3 discusses the data description and modeling approach ; Section 4 presents results and discussion and Section 5 concludes the paper.

2. PROBLEM FORMULATION

The Close-In (CI) reference distance path loss model is a widely adopted empirical model used for estimating wireless signal attenuation. It is mathematically expressed as Equation (1):

$$PL_{CI}(f,d) = PL(f,d_0) + 10nlog_{10}\left(\frac{d}{d_0}\right) + X_\sigma \quad (1)$$

where $PL_{CI}(f,d)$ is the path loss (in dB) at frequency f (in Hz) and distance d (in meters), $PL(f,d_0)$ is the free-space path loss at the reference distance d_0, n is the path loss exponent (PLE) that depends on the propagation environment, X_σ is a zero-mean Gaussian random variable representing shadow fading, with standard deviation σ, d_0 is the reference distance, typically set to 1 meter. The free-space path loss at the reference distance d_0 is computed by Equation (2):

$$PL(f,d_0) = 20log_{10}\left(\frac{4\pi d_0 f}{c}\right) \quad (2)$$

where c is the speed of light in free space ($c \approx 3 \times 10^8 m/s$). In this model, the shadow fading term X_σ captures the random variations in path loss caused by obstacles and environmental factors. Its stochastic nature makes accurate path loss prediction challenging, especially in dynamic environments like urban air-to-ground (A2G) communication scenarios.

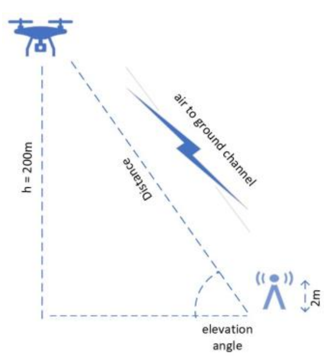

Figure 1. System model of UAV-to-ground communication

The system setup considered in this study is illustrated in Figure 1, where a communication link is established between an Unmanned Aerial Vehicle (UAV) and a ground-based receiver. In this configuration:

- The UAV operates at a fixed altitude of 200 meters above ground level.
- The receiver is positioned at 2 meters above ground level.
- The dataset includes features such as the horizontal distance between the UAV and the receiver, and the elevation angle of the communication link.

The choice of 200 meters for UAV altitude is based on common deployment practices in urban environments, which provide a balance between coverage and LoS probability. However, future work should examine the impact of varying UAV altitudes, as different heights may lead to distinct path loss behaviors due to changes in reflection and diffraction patterns.

This study applies machine learning to predict path loss in air-to-ground channels across three carrier frequencies: 1 GHz, 2 GHz, and 5.8 GHz. Due to the limitations of traditional models in capturing urban propagation complexity, data-driven regression offers a promising alternative for UAV-based communication scenarios.

3. DATA DESCRIPTION AND MODELING APPROACH

3.1. Dataset

The dataset used in this study is publicly available and was generated through high-fidelity simulations using Wireless InSite software. The simulation was conducted in a realistic urban environment modeled from CADMAPPER, covering 0.748 km² with 1,419 buildings, approximately 98% of which are high-rise structures [8]. The dataset includes three input features: UAV-receiver distance, elevation angle, and carrier frequency. After preprocessing, 36,753 records were retained and randomly split into training (80%) and testing (20%) sets.

3.2. Feature Selection

Table 1. Feature set description

No	Name	Type	Description
1	Frequency	Numeric	The carrier frequency of the system (GHz)
2	Distance	Numeric	The distance between the UAV and the receiver (m)
3	Angle	Numeric	The elevation angle created by the transmitter/receiver line straight to the earth (degree)
4	FSPL	Numeric	The path loss is calculated from the log-distance model (dB)

The free-space path loss (FSPL) model is widely used for its simplicity but often falls short in urban environments due to multipath and shadowing effects. To improve prediction accuracy, we adopt a K-Nearest Neighbors (KNN) regression algorithm, which captures nonlinear relationships between features and path loss. In addition to the basic inputs—distance, elevation angle, and frequency—we include the log-distance path loss as an extended feature. This hybrid approach combines physical modeling with data-driven learning. The complete feature set is listed in Table 1.

3.3. Performance Evaluation

To assess model performance, we use two standard metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) [10], defined in Equations (3) and (4). These quantify the deviation between predicted and actual path loss values on the test set:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y_i}| \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y_i})^2} \quad (4)$$

Where y_i and $\hat{y_i}$ represent the actual and predicted path loss values for the i^{th} test sample, respectively, and N denotes the total number of test samples. These metrics allow for consistent accuracy comparisons across models and frequencies.

3.4. The proposed algorithm

The flowchart of the proposed A2G path loss prediction algorithm is shown in Figure 2. The process begins with data preprocessing to remove invalid entries and handle missing values, followed by splitting the dataset into training and testing sets. An additional feature—path loss calculated using Equation (1)—is added to improve prediction accuracy.

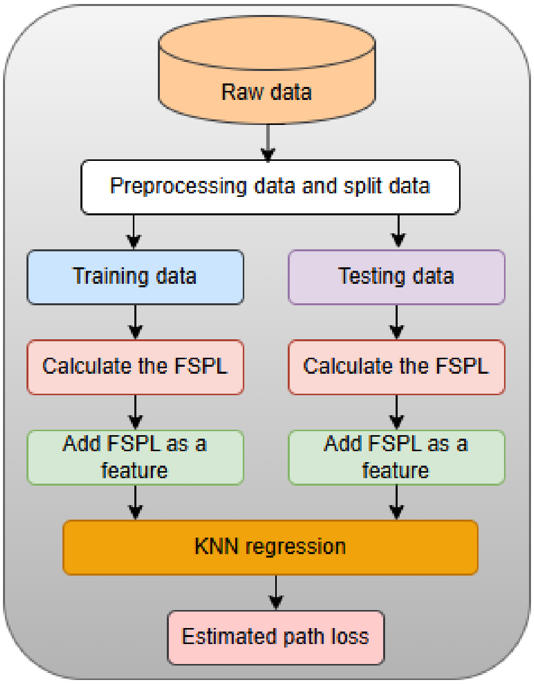

Figure 2. Flowchart of the proposed algorithm

The final input is fed into a KNN regression model with k=35 and p=2, and performance is evaluated on the test set. The parameter k=35 was selected based on cross-validation experiments to balance bias and variance, while p=2 corresponds to the Euclidean distance metric, which is suitable for the spatial characteristics of the input features.

4. RESULTS AND DISCUSSION

The proposed model was implemented on a Lenovo ThinkPad laptop equipped with 8 GB RAM, using Python 3.9 as the development environment. The total execution time for processing the entire dataset was 0.95 seconds.

To further analyze the multi-frequency performance, we computed MAE and RMSE separately for each frequency. The KNN model maintained consistently low errors across all frequencies, with slightly higher accuracy at 1 GHz and 2 GHz due to lower attenuation. These results confirm the model's generalization capability across different frequency bands used in UAV communications.

Figure 3 presents the relationship between elevation angle and path loss as observed in the raw input data, the free-space model, and the proposed KNN-based model. The results show that the KNN regression approach aligns more closely with the actual measurement data compared to the free-space model, particularly in complex urban propagation scenarios.

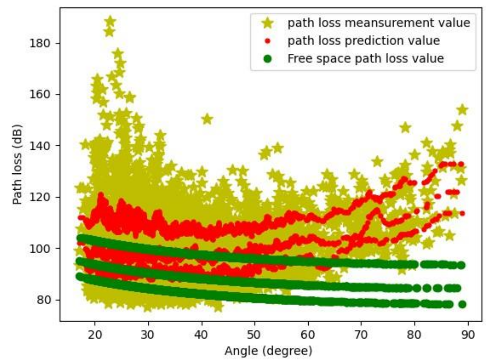

Figure 3. The relationship between elevation angle and path loss in the A2G channel

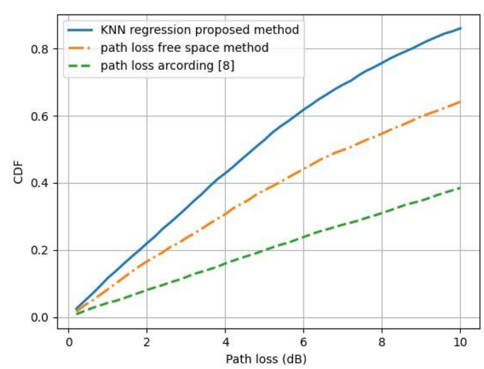

Figure 4. The CDF path loss

The cumulative distribution function (CDF) of prediction errors is depicted in

Figure 4, clearly demonstrating the improved accuracy of the machine learning model over the traditional models. Specifically, the probability of the path loss error being less than 5 dB reaches 50% with the KNN model, compared to only 38% with the free-space model and 20% with the model proposed in [8]. Similarly, for errors below 10 dB, the cumulative probability is 84% using KNN, while the free-space and [8] models achieve only 63% and 39%, respectively.

Table 2. Statistical comparison of models

Model	MAE (dB)	RMSE (dB)
Free-space	9.295	12.634
Model from [8]	16.173	20.459
KNN Regression	**5.821**	**7.956**

As shown in Table 2, the proposed KNN regression model delivers superior performance, achieving the lowest MAE and RMSE values. This improvement can be attributed to the fact that the dataset is derived from a dense urban environment with a high density of tall buildings, where signal propagation is influenced by not only line-of-sight (LOS) components but also significant contributions from reflection, diffraction, and non-line-of-sight (NLOS) effects.

While the free-space model fails to account for these complex interactions, the KNN-based machine learning model is able to implicitly learn such propagation behaviors from data, making it a robust solution for predicting path loss in UAV-enabled A2G communication systems. These results highlight the strong potential of data-driven techniques to supplement or replace traditional models, especially in challenging urban environments.

5. CONCLUSION

This study proposed a KNN-based machine learning model to predict air-to-ground path loss in dense urban environments. By incorporating both physical and data-driven features, the model achieved significantly lower MAE and RMSE compared to the free-space model and an existing method in [8]. The results confirm the effectiveness of machine learning in capturing complex propagation characteristics. Future work may explore deep learning models and apply the method to dynamic UAV scenarios.

REFERENCES

[1] Chowdhury, M.Z.; Shahjalal, M.; Ahmed, S.; Jang, Y.M. 6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions. IEEE Open J. Commun. Soc. 2020, 1, 957–975. [Google Scholar] [CrossRef]

[2] Alzubaidi, O.T.H.; Hindia, M.N.; Dimyati, K.; Noordin, K.A.; Wahab, A.N.A.; Qamar, F.; Hassan, R. Interference Challenges and Management in B5G Network Design: A Comprehensive Review. Electronics 2022, 11, 2842. https://doi.org/10.3390/electronics11182842.

[3] Bin Li, Zesong Fei, Senior Member IEEE, and Yan Zhang, Senior Member IEEE. UAV Communications for 5G and Beyond: Recent Advances and Future Trends. IEEE Internet of thing Journal. arXiv:1901.06637v1 [cs.NI] 20 Jan 2019.

[4] M. Polese, M. Giordani, T. Zugno, A. Roy, S. Goyal, D. Castor, and M. Zorzi: . In Tegrated Access and Backhaul in 5G mmWave Networks: Potential and Challenges. IEEE Commun. pp. 62–68. Mar (2020)

[5] Ki Joung Jang, Sejun Park, Junseok Kim, Youngkeun Yoon, Chung-Sup Kim, Young-Jun Chong, And Ganguk Hwang: Path Loss Model Based on Machine Learning Using Multi-Dimensional Gaussian Process Regression. IEEE Open Access Journal. Vol 10 (2022)

[6] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang: Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks-With a Focus on Propagation Models. IEEE Trans. Antennas Propag. pp. 6213–6230. Dec (2017)

[7] Qiuming ZHU, Mengtian YAO, Fei BAI, Xiaomin CHEN, Weizhi ZHONG, Boyu HUA, Xijuan YE "A general altitude-dependent path loss model for UAV-to-ground millimeter-wave communications", Frontiers of Information Technology & Electronic Engineering, 2021 22(6):767-776

[8] Nour El-Din Safwat, Fatma Newagy, Ismail M. Hafez, "Air-to-ground channel model for UAVs in dense urban environments", IET Communications 2020, Vol. 14 Iss. 6, pp. 1016-1021

[9] Ashraf Tahat , Talal Edwan , Hamza Al-Sawwaf , Jumana Al-Baw , Mohammad Amayreh, "Simplistic Machine Learning-Based Air-to-Ground Path Loss Modeling in an Urban Environment" 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC)

[10] Y Isabona and V. M. Srivastava, "Hybrid neural network approach for predicting signal propagation loss in urban microcells," in Proceedings of the 2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), pp. 1–5, Agra, India, December 2016.

Authors Biography

Duong Thi Hang, was born in Bac Giang, Vietnam. She received her Bachelor's degree in 2000, her Master's degree in 2006, and completed her Ph.D. in 2025, all from the University of Engineering and Technology, Vietnam National University, Hanoi. She is currently a lecturer at Electric Power University (EPU), Hanoi, Vietnam. Her main research interests include indoor localization using machine learning, optimization algorithms, and UAV-based communication systems. Her work focuses on developing efficient, interpretable, and scalable solutions for real-time positioning challenges, with an emphasis on combining machine learning and wireless technologies to enable smarter environments and autonomous systems.

Pham Duy Phong is the Dean of the Faculty of Electronics and Telecommunications at the Electric Power University, Hanoi, Vietnam. He received the B.E degree in Telecommunications Engineering from University of Communications and Transport, Hanoi, in 2000 and the Master degree from Hanoi University of Science and Technology, Hanoi, Vietnam in 2007. He received the Ph.D degree in theTelecommunications Engineering at Vietnam Research Institute of Electronics, Informatics and Automation, Hanoi, Vietnam in 2013. He was a researcher in Research Institute of Posts and Telecommunications (2000-2005). His main research interests include wireless communications, antenna design for wireless communications, underwater acoustic communications, electromagnetic interference on telecommunication systems due to power systems, earthing and lightning protection.