

A Systematic Sample Size Analysis of ML-Based Path Loss Prediction for LPWAN

Robert Bitterling^{◦•}, Christian Nettersheim^{◦•}, Jörn Hees[•] and Michael Rademacher^{◦•}

[◦]Fraunhofer FKIE
Cyber Analysis & Defense
Wachtberg, Germany

{firstname.lastname}@fkie.fraunhofer.de

[•]Hochschule Bonn-Rhein-Sieg
Department of Computer Science
Sankt Augustin, Germany

{firstname.lastname}@h-brs.de

Abstract—Low Power Wide Area Networks like LoRa are increasingly deployed for smart city applications, requiring accurate path loss prediction for effective network planning. Traditional (empirical) propagation models often exhibit limited accuracy in these scenarios. We investigate machine learning models for LoRa path loss prediction, systematically analyzing how prediction accuracy scales with training set size using real-world measurements from an urban deployment. Our approach employs a Random Forest with LiDAR-derived terrain features and k-Nearest Neighbors with coordinate data, comparing their performance against established empirical models and specialized LPWAN models. Under random pooled splits, both ML models consistently outperform the considered baseline models across the evaluated training-set sizes. At maximum training size, they achieve RMSE values below 6.5 dB compared to 9.7 dB for the best baseline, indicating accurate within-deployment interpolation. A leave-one-gateway-out check qualifies this result: RF shows placement-dependent transfer to held-out gateways, with moderate degradation for several gateways but larger errors for others, whereas coordinate-only k-NN degrades substantially when the gateway location is unseen.

Index Terms—Low Power Wide Area Network, Propagation Modeling, Machine Learning

I. INTRODUCTION

Smart city Internet of Things (IoT) deployments require wide-area, low-power connectivity. Low Power Wide Area Network (LPWAN) technologies such as Long Range (LoRa) are well suited for this setting because they provide long range, low energy consumption, and operation in unlicensed bands [1]. Establishing these networks requires careful planning to meet capacity and coverage requirements [2]. Local authorities often have less network planning expertise than Mobile Network Operators (MNOs), which makes planning difficult. A key task is to estimate signal strength across locations with propagation models. Empirical models balance computational cost, required knowledge, and accuracy, but several studies report limited accuracy in LPWAN networks [3]–[5]. This has motivated the development of refined or alternative models. For example, Rademacher *et al.* conducted a large-scale measurement campaign and derived new parameters for a Log-Distance Path Loss (LDPL) model, which outperformed traditional models [3]. Despite its improved performance, this approach is not feasible for every region, as it requires a considerable amount of work and time to

collect the necessary data. This constraint highlights the need for alternative approaches that can achieve high prediction accuracy with fewer data points and less domain-specific knowledge.

As an alternative, Artificial Intelligence (AI) and Machine Learning (ML) methods have emerged as promising candidates for predicting signal strength in wireless networks [6]. Given the cost of collecting large datasets, we ask how prediction accuracy depends on training-set size in practical smart city planning and how the models compare to traditional path-loss models. We then examine whether these gains reflect transfer beyond the measured gateways or mainly interpolation within the existing deployment. The key contributions of this work are as follows:

- We empirically demonstrate that simple ML models, Random Forest (RF) with Light Detection and Ranging (LiDAR)-derived terrain features and coordinate-based k-Nearest Neighbors (k-NN), outperform conventional propagation models under random pooled evaluation.
- We provide a systematic analysis of learning curves and training-set size effects, offering practical guidance for model selection and data collection by showing how accuracy scales with available data.
- We further add a leave-one-gateway-out (LOGO) evaluation to check how performance changes when all measurements from one gateway are held out.

To facilitate credibility in our results and enable other researchers to build upon our work, we released the complete source code of this work on our GitHub page¹.

II. BACKGROUND AND RELATED WORK

In this section, we introduce empirical propagation models, discuss their challenges for LPWANs, and then review ML for propagation loss estimation in these networks.

Classical path loss models estimate the reduction in signal strength between a transmitter and receiver due to propagation effects. An accurate estimation of this reduction is a necessity for link budget calculations in wireless networks. Classic empirical models (i.e., Okumura–Hata [7] and COST-231 Hata [8]) consider influencing factors, such as the type

¹<https://github.com/mclab-hbrs/lora-bonn-ml-pathloss>

of environment, frequency, distance, and antenna heights. However, some of these empirical models exhibit limited accuracy in LPWAN scenarios [3]–[5]. Their performance can degrade outside their original calibration environments and frequency ranges. This has motivated researchers such as Jörke *et al.* [5], Stusek *et al.* [4], and Rademacher *et al.* [3] to develop or refine empirical models specifically based on particular measurement campaigns or deployment scenarios. The process of creating or refining these models typically follows the same pattern: identify a shortcoming with current models, conduct a measurement campaign, preprocess the data, derive or refine a model, and evaluate it against the measurements.

The described empirical modeling process closely resembles supervised ML workflows. In these workflows, models are trained on labeled data to predict quantities such as path loss [6]. The key difference is that, while empirical models typically use a fixed functional form and adjust only certain parameters, ML algorithms can learn more flexible relationships from data without a predefined structure. Recent surveys, such as the one by Seretis and Sarris [6], have shown that ML models, including both Artificial Neural Network (ANN)-based and non-ANN models like RF and k-NN, can achieve higher accuracy than traditional empirical models for modeling radio wave propagation. Random Forests are frequently highlighted for their strong performance.

Although there are numerous studies applying AI techniques in wireless communications, a review of the literature reveals that only a limited number have focused specifically on ML to model radio wave propagation in LoRa scenarios. Some of the identified studies focus on ML and LoRa [9, 10], while others use ML only as a support tool rather than as their primary research focus [11, 12].

Alobaidy *et al.* [11] developed a hybrid ML model for water quality monitoring in a tropical environment. The approach combined free space path loss with linear regression and a bagged trees ensemble, using 17 radio and environmental features. The data were preprocessed with standard ML methods and augmented with ray tracing and a 3D building map. Their model outperformed conventional models.

Gonzalez-Palacio *et al.* [12] aimed to save energy in LoRa networks by proposing an Adaptive Data Rate (ADR) algorithm that incorporates environmental factors. They evaluated multiple ML models, including multiple linear regression, ANNs, support vector regression, and RFs, to predict path loss statistics. RFs performed best.

Ballestrin *et al.* [9] developed and benchmarked ML models for LoRa path loss prediction using ANNs, RFs, Support Vector Regression (SVR), Linear Regression (LR), XGBoost, and decision trees. The models used a dataset enriched with elevation data and were compared to Okumura and log-distance baselines. All ML models except SVR with a sigmoid kernel outperformed the traditional models, with RFs, XGBoost, and decision trees performing best. The study reported limited details on preprocessing and feature enrichment.

Liu *et al.* [10] addressed the limits of dominant land cover features by developing a deep neural network based on Bidirectional Long Short-Term Memory (Bi-LSTM). They collected 30,000 urban data points using two gateways and six end nodes. The model used sequences of classified land cover types, distance, and height difference as inputs. Land cover was automatically classified from remote sensing images with Support Vector Machine (SVM) using a Radial Basis Function (RBF) kernel, distinguishing Line of Sight (LOS) and Non-Line-of-Sight (NLOS) types such as trees, buildings, water, and roads. Training on these features allowed the model to approximate propagation behavior and generalize to new environments. It halved the error of conventional methods.

In contrast to previous studies, our work focuses on the amount of training data required for accurate ML-based LoRa propagation-loss prediction. Rather than constructing complex or hybrid ML architectures or employing deep learning models, we emphasize simple, easy-to-use ML models. Furthermore, our study does not concentrate solely on maximizing prediction performance or benchmarking different ML approaches. Instead, we investigate these models as an extension to Rademacher *et al.*'s [3] work. Specifically, we examine how training-set size affects model performance.

III. METHODOLOGY

To examine how training-set size affects LoRa path-loss prediction, we designed an experimental setup using real-world measurements and terrain data. Our approach emphasizes a comparative evaluation of traditional and ML-based models. We detail dataset preparation, features, model selection, and experimental setup in the following. An overview is shown in Figure 1, and we refer to the numbered steps.

We use the final post-processed dataset from the study by Rademacher *et al.* [3] rather than collecting new measurements ❶. The data were collected in Bonn, Germany, an urban area of approximately 330,000 inhabitants. Uplink Received Packet Power (RPP) was recorded from outdoor end devices mounted on garbage trucks over 230 days, using nine gateways citywide at 868 MHz, 125 kHz bandwidth, and Spreading Factor (SF) 12. After the original cleaning, 131,057 samples remained. For this work, we removed samples and gateways outside the city limits, reducing the set to 114,465 samples and eight gateways ❷. From measured RPP and known system parameters (e.g., antenna gains), we computed path loss. To reduce redundancy, we removed duplicates and applied Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [13] to sender and receiver coordinates, grouping spatially proximate samples and consolidating overlaps, yielding 64,294 samples ❸.

A key idea in this work is to use terrain-based features as input to ML-based LPWAN modeling. To obtain terrain-based features, we employ ❹ “SignalServer”², a multithreaded RF-propagation software based on “SPLAT!”³. “SignalServer”

²<https://github.com/lmux/Signal-Server/tree/85589ed6>

³<https://www.qsl.net/kd2bd/splat.html>

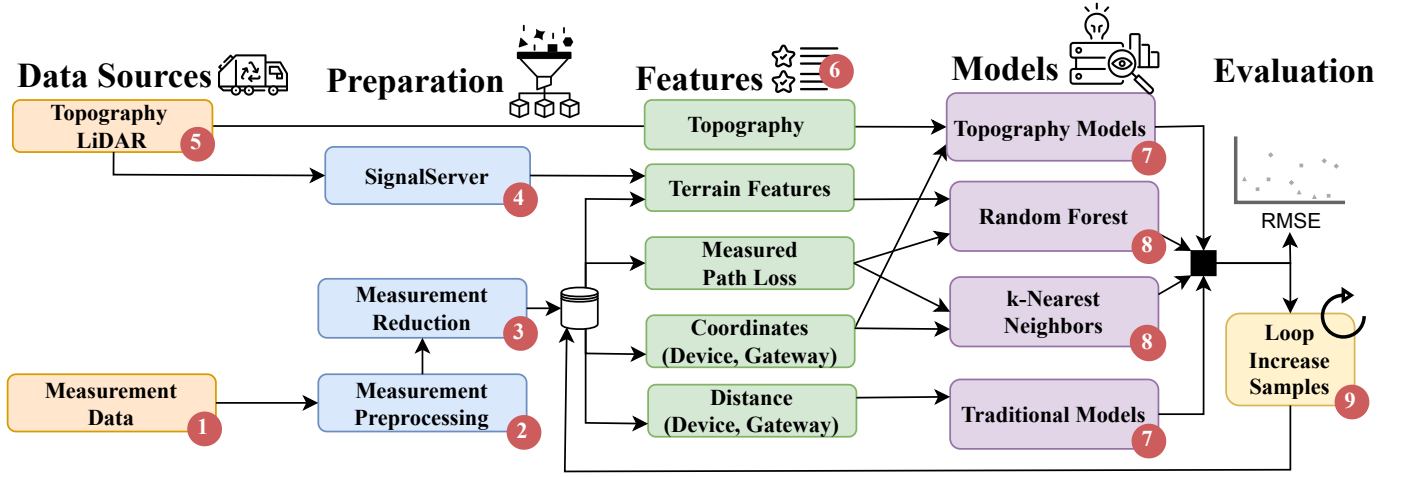


Fig. 1. Overview of the experimental pipeline. LiDAR and measurement data are prepared. From these sources, we derive features (converted LiDAR/topographic inputs, SignalServer outputs, measured path loss, coordinates, and distances) to train and evaluate models (deterministic topography-based, traditional empirical, RF, and k-NN). Performance is assessed by RMSE. The experiment is repeated across multiple training set sizes.

can produce 360° polar coverage maps as well as 2-D point-to-point path profiles. LiDAR topographic data (provided by the state of NRW in Germany) were converted from LAZ (EPSG:25832) to ASCII grid (EPSG:4326) for “SignalServer” ⑤. From the path profiles for all data points, we derive features including antenna height Above Mean Sea Level (AMSL) at transmitter and receiver, receiver antenna height required to clear the first Fresnel zone, receiver antenna height required to clear all detected obstructions, receiver antenna height required to clear 60% of the first Fresnel zone, Free-Space Path Loss (FSPL), downtilt angles to transmitter and receiver, azimuths to transmitter and receiver, and distance to receiver (Figure 2). After preprocessing and feature engineering ⑥, the available inputs are topographic data, terrain-derived features, coordinates, and distance, with path loss as the target variable.

Traditional empirical and terrain-based models serve as baseline approaches in our experiments ⑦ and include both widely used empirical models such as Okumura-Hata, ECC33 [14], ITM [15], and ITWOM [16], and specialized models such as Dortmund [5], Ghent [17], Oulu [18], and Bonn [3]. Following the protocol of Rademacher *et al.* [3], we selected the same set of models to ensure consistency and comparability between studies. This selection strategy reflects two common scenarios encountered in radio propagation research: the use of established, widely recognized models, and the development or adoption of site-specific models based on dedicated measurement campaigns. Models such as ECC33, COST-Hata, Ericsson, Egli, ITM, and ITWOM were computed using “SignalServer”, while Okumura-Hata and specialized models (Dortmund, Ghent, Oulu, and Bonn) were implemented directly by us.

As ML models, we selected k-NN and RF for methodological and practical reasons ⑧. Prior work suggests that, outside deep learning, common regressors yield similar path-loss accuracy. We chose k-NN for simplicity and interpretability

and RF for strong prior performance [9, 12] and out-of-the-box usability. Both have low computational overhead.

For k-NN, we use only the latitude and longitude of transmitter and receiver as features, with path loss as the target. This defines a simple, low-input ML baseline. The coordinates implicitly encode link distance and local spatial context, while avoiding the terrain-derived features used by RF. This allows us to compare a coordinate-based learner with both traditional low-input propagation models and the terrain-feature-based RF model. The value $k = 31$ with distance weighting was selected via grid search over $k \in \{1, 3, 5, 7, 9, 11, 15, 21, 31, 41, 51, 61, 71, 81, 91, 101, 151, 201, 251, 301, 351\}$,

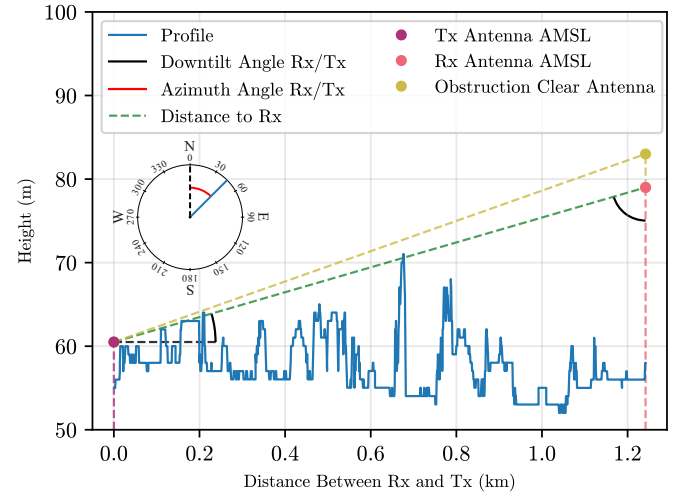


Fig. 2. Illustration of selected terrain features used as input variables for the RF models. Displayed features are: “Tx antenna height AMSL”, “Rx antenna height AMSL”, “Downtilt angle to Tx”, “Downtilt angle to Rx”, “Distance to Rx”, “Rx antenna height to clear all obstructions detected”, “Azimuth to Tx”, and “Azimuth to Rx”. The “Profile” curve is included for illustration only and is not used as an input feature. All features were generated using “SignalServer”. Additional model features not displayed include: “Rx antenna height to clear the first Fresnel zone”, “Rx antenna height to clear 60% of the first Fresnel zone”, and “FSPL”.

TABLE I
RMSE (dB) COMPARISON OF MODELS ACROSS TRAINING-SET SIZES.

Sample Size	This Work			Baseline Models			
	RF	k-7	k-31	Bonn	C. Ha.	ITM	ITW.
100	8.36	9.56	10.02				
200	8.07	8.83	9.37				
500	7.74	8.11	8.47				
4,500	7.02	6.95	6.95	9.68	12.39	21.75	22.58
15,000	6.65	6.58	6.47				
48,220	6.34	6.32	6.15				

Note: The test-set size was held constant for all experiments. Baseline models (Bonn, COST-Hata, ITM, ITWOM) were not trained on the dataset. Therefore, their RMSE values remain approximately constant across training-set sizes. Minor differences arise from random sampling, not systematic changes in model effectiveness.

testing uniform and distance weighting. Distance weighting with $k = 31$ yielded the lowest validation RMSE. However, to address potential oversmoothing and poor performance at low sample sizes, a variant with $k = 7$ was also evaluated in iterations with limited training data. Although incorporating the explicit distance between transmitter and receiver as a feature in k-NN was considered, we opted not to do so, as the coordinate-based approach essentially provides a spatial lookup and maintains simplicity.

To go beyond simple region-bound coordinate features and to provide inputs comparable to terrain-based models such as ITM and ITWOM, we used the terrain information available through SignalServer. ITM and ITWOM require topography data as input, which SignalServer processes directly to compute their propagation predictions. The same SignalServer path-profile analysis that enables these traditional models also provides the terrain-based features described above, including elevation data, Fresnel zone clearances, and azimuth information. To leverage these engineered features in a machine learning context, we employed RF (using YDF⁴), which learns patterns from these terrain-derived features rather than processing the raw topographic data.

Our primary evaluation measures how accuracy scales with training-set size. Following [9], we used a 75/25 random pooled split: a test set with a fixed size of 16,074 samples was held out, and training subsets of increasing size were drawn at random from the remaining data [9], with each configuration repeated by resampling the training subset. This pooled split cannot indicate how far the models transfer beyond the measured gateways. Rather than relying on a second, independent dataset, we use a LOGO check at the largest training size to gain initial insight into this: all samples from one gateway are held out for testing and the model is trained on the remaining gateways. We report each LOGO metric as the unweighted mean over the eight held-out gateways, so that every gateway contributes equally rather than in proportion to its number of measurements.

IV. EVALUATION RESULTS

We evaluated a range of models, including RF and k-NN ($k = 7$ and $k = 31$), as described in Section III. To maintain clarity in the results section, we report results for a representative subset of the evaluated models.

Table I presents the Root Mean Squared Error (RMSE) of various models, comparing traditional models with Machine Learning models for training-set sizes ranging from 100 to 48,220. The models discussed include the COST 231 Hata model, the LDPL Bonn model, and the terrain-based ITM and ITWOM models, together with the ML models k-NN ($k = 7$ and $k = 31$) and the RF model. At small sample sizes, the LDPL Bonn model achieves an RMSE comparable to the k-NN models, while the RF model outperforms all traditional models. As the sample size increases, the ML models, particularly RF and k-NN with $k = 31$, show substantial improvements in RMSE and reliably outperform the baseline models. The baseline models' RMSE values remain nearly constant, with Bonn around 9.68 dB, COST 231 Hata at 12.39 dB, ITM at 21.75 dB, and ITWOM at 22.58 dB. In contrast, the RF and k-NN models reduce their RMSE values to below 6.5 dB at the largest sample sizes. Overall, under the random pooled evaluation, the ML models consistently outperform the considered traditional models across the evaluated training-set sizes.

Figure 3 provides a more detailed analysis of machine learning models. For training-set sizes between 100 and 500, the RF model consistently outperforms both k-NN models. The performance gap between the RF and k-NN models narrows as the sample size increases. Both k-NN models begin with similar RMSE values, but the model with $k = 7$ improves faster than $k = 31$. This is likely due to $k = 7$ being more appropriate for smaller sample sizes. The improvement is substantial, with the k-NN models achieving a reduction in RMSE of approximately 1.5 dB and the RF model improving by about 0.6 dB as the sample size increases to 500. Although the central tendency of the RMSE curves is favorable, the spread is non-negligible, especially for k-NN: the k-NN bands overlap substantially at small sample sizes, whereas the RF band intersects those of the k-NN models only rarely and primarily at larger sample sizes. Overall, the bands narrow with increasing sample size, indicating reduced variability.

When shifting from a detailed view to coarser increments, the trend of decreasing mean RMSE persists. Up to roughly 3,250 samples the RF maintains the lowest RMSE. At this point both k-NN models overtake the RF. The first vertical reference line marks this intersection. At around 4,750 samples the k-NN with $k = 31$ overtakes the other k-NN, marked by the second vertical line. From this stage onward, the k-NN with $k = 7$ and the RF model exhibit closely matched RMSE values, reaching 6.32 dB and 6.34 dB, respectively. At the maximum training-set size, the k-NN model with $k = 31$ achieves the lowest RMSE at 6.15 dB. Overall, a clear improvement trend emerges with diminishing returns: for RF,

⁴<https://ydf.readthedocs.io/en/stable/>

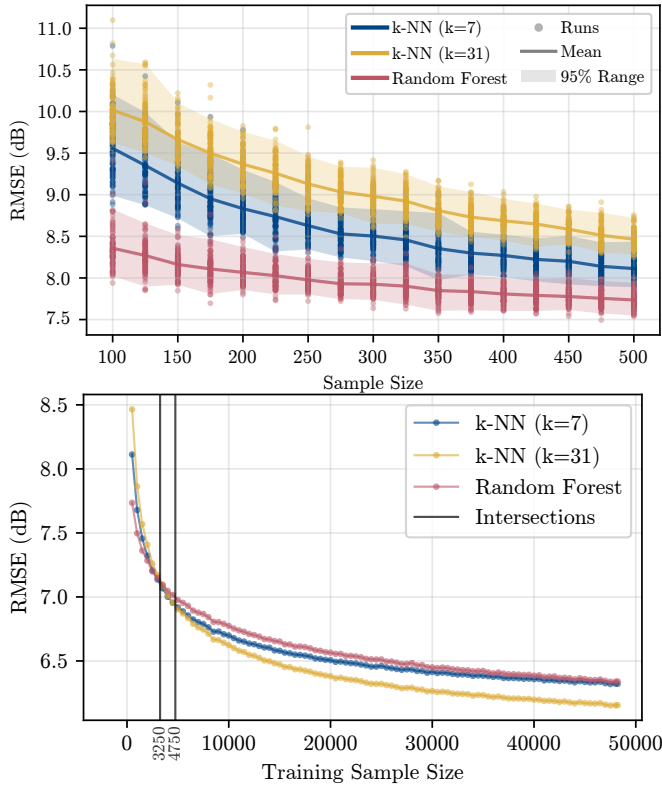


Fig. 3. Mean RMSE (dB) versus training-set size. Top: sizes 100–500 over 100 independent runs, with mean lines and 95% envelopes; variability shrinks as the sample size grows and RF leads across this range. Bottom: sizes 500–48,220, where accuracy improves for all models with diminishing returns; both k-NN variants overtake RF at approximately 3,250 samples, and $k = 31$ overtakes $k = 7$ at approximately 4,750 samples.

RMSE improves by 0.62 dB from 100 to 500 samples, by 1.08 dB from 500 to 15,000 samples, and by only 0.31 dB from 15,000 to 48,220 samples.

RMSE is an aggregate measure and does not reveal the distribution of individual errors, which Figure 4 addresses. In general, apart from the direction of underestimation or overestimation, the prediction error behavior of the models aligns with the trends observed in the performance line plots. Among baseline models, most tend to underestimate path loss, except ITM, whose distribution is centered near zero but has a wider spread, especially on the overestimation side. ITWOM is less dispersed but shifted toward positive prediction errors, indicating consistent underestimation. Non-terrain-based baselines show smaller overall spreads than terrain-based models. For ML models (RF, k-NN with $k = 7$ and $k = 31$), we show training sizes of 500 and 48,220. Their error distributions are more compact, with narrower and darker central bins, than the baselines, although at 500 samples the errors of all model distributions are a bit more spread out. At the maximum sample size, the maximum spread of the prediction errors remains relatively unchanged, but the central error distribution narrows. Large, infrequent prediction errors persist, indicating either inherent variability in the measurements or remaining unmodeled factors. The Bonn model, a domain-specific model derived from a variant

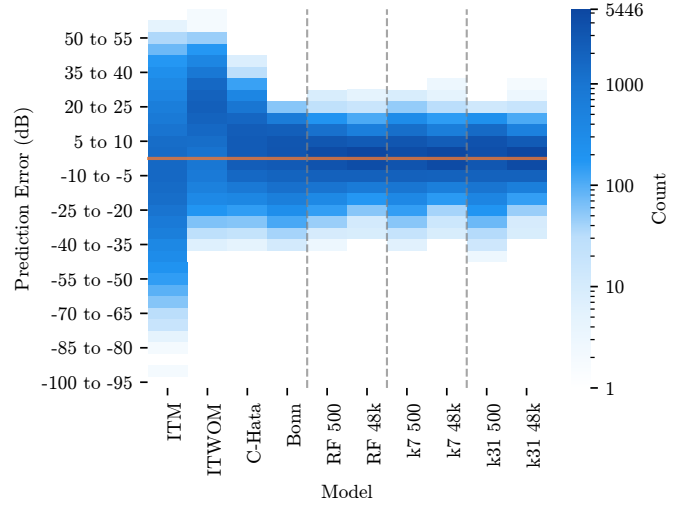


Fig. 4. Distribution heatmap of path-loss prediction errors (prediction error, dB) from one iteration across prediction error bins (5 dB per bin) for multiple models. Logarithmic color scale (white = zero occurrences; darker blue = higher frequency). Abbreviations: ITM = Irregular Terrain Model; ITWOM = Irregular Terrain with Obstructions Model; C-Hata = COST-Hata; RF = Random Forest; k7/k31 = k-NN ($k = 7$ and $k = 31$). Parentheses indicate training set sizes (500 vs. 48k).

TABLE II
RMSE (dB) FOR THE RANDOM POOLED SPLIT VERSUS
LEAVE-ONE-GATEWAY-OUT (LOGO) AT THE LARGEST TRAINING SIZE.

Model	Random	LOGO	Δ
RF	6.34	8.92	+2.58
k-NN ($k = 7$)	6.32	13.73	+7.41
k-NN ($k = 31$)	6.15	13.28	+7.13

Note: LOGO trains on 7 of 8 gateways and tests on the held-out one, averaged with equal weight across all 8 gateways. The random baseline is the largest-size (48,220-sample) result from Table I.

of the dataset [3], stands out among the baseline models, as its prediction error distribution closely matches that of the ML models trained with 500 samples in both spread and central tendency.

To gain a first insight into how far the models transfer beyond the measured gateways, we perform a LOGO evaluation. Table II compares the random split and LOGO at the largest training size. Both k-NN variants degrade sharply, roughly doubling their RMSE from about 6.3 dB to about 13.5 dB. This is expected, as the coordinate-only k-NN acts as a spatial lookup: no training samples share the held-out gateway coordinate, leaving k-NN without nearby receiver-side neighbors for interpolation. This supports the interpretation that coordinate-only k-NN mainly captures spatial interpolation within the measured deployment rather than transferable propagation behavior. The RF model degrades far less, from 6.34 dB to 8.92 dB (an increase of 2.58 dB), indicating that its terrain-derived features transfer partially to unseen gateways.

This aggregate, however, hides substantial gateway-to-gateway variation. Figure 5 shows the RF error for each held-out gateway. For five of the eight gateways, the RMSE remains between 6.6 dB and 7.6 dB, close to the random

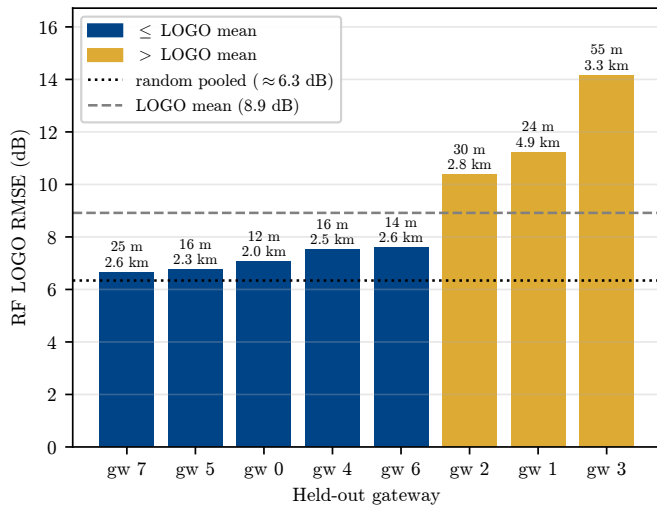


Fig. 5. RF path-loss RMSE for each held-out gateway under LOGO (trained on the other seven gateways), sorted by error and annotated with the gateway’s antenna height and mean link distance. The dotted line marks the random pooled error (about 6.3 dB), the dashed line the eight-gateway LOGO mean. Five held-out gateways remain close to the random pooled error, whereas three gateways with larger antenna height or mean link distance drive the average up.

pooled error of about 6.3 dB. This suggests that held-out gateways with characteristics similar to the training gateways can be predicted with only moderate degradation relative to the random pooled result. The remaining three gateways raise the average considerably, reaching 10.4 dB, 11.2 dB and 14.2 dB. These three gateways stand out by their site characteristics: an unusually large antenna height of 55 m, a mean link distance of 4.9 km, and an antenna height of 30 m.

V. DISCUSSION AND FUTURE WORK

Across random pooled splits ML models outperform the considered empirical and terrain-based baselines. RF is strongest at very small sample sizes and remains favorable up to roughly 3,250 samples. With more data, coordinate-based k-NN becomes competitive and achieves the lowest RMSE at the largest training size, albeit by a small margin. This behavior is expected because k-NN acts mainly as a spatial interpolation method: as sampling density increases, the nearest measured links tend to lie closer to each prediction location. The gain per additional sample diminishes, however, indicating a practical limit to the accuracy attainable from more data alone.

The pooled-split results, however, can overestimate performance when the target gateway was not represented during training, since links from the same gateways can appear in both training and test data. Our LOGO check, which isolates this transfer setting, shows a more nuanced result. RF remains competitive on average and transfers with moderate degradation to held-out gateways that are similar to those in the training data. For some gateways, errors increase significantly. In our data, these high-error cases coincide with unusual antenna heights or different link-distance distributions. Since we did not perform a dedicated causal analysis or pre-

deployment outlier-detection study, we treat these characteristics as descriptive indicators rather than confirmed causes.

For practitioners, model choice should reflect the available data, the target accuracy, and the effort of feature extraction. With only a few hundred labeled links, RF yields the best accuracy, at the cost of deriving terrain features. With a few thousand samples or more, coordinate-only k-NN becomes equally competitive for within-deployment interpolation while requiring only link coordinates. Beyond roughly 3,250 samples the two lie within about 0.2 dB, a gap that should be interpreted with caution given the limited hyperparameter tuning and the differing feature sets. When data or compute is highly constrained, a traditional empirical model such as COST-231 Hata, which requires no local measurements, may still suffice.

These findings come with limitations. They are based on one urban region, environment, technology, and frequency, so whether terrain-feature models transfer across cities, technologies, or frequencies remains open. We also did not perform extensive hyperparameter tuning, which could shift absolute errors and the small differences between ML models.

Two directions follow for future work. First, for within-deployment interpolation, the central question is how to design a small measurement campaign that still achieves the maximum tolerable RMSE, selecting sampling locations that best represent the environment while reducing the variability seen at low sample sizes. Second, replication on independent deployments in other cities, frequencies, and technologies would test the external validity left open above.

VI. CONCLUSION

We evaluated simple machine learning models for LP-WAN path loss prediction using real-world LoRa measurements and LiDAR-derived terrain features in an urban setting. Under random pooled evaluation, both models consistently outperformed the traditional empirical and terrain-based baselines across training-set sizes, reaching RMSE below 6.5 dB versus about 9.7 dB for the best baseline, with clear diminishing returns as data grew.

These results show that data-driven models deliver accurate path-loss estimates for within-deployment interpolation with modest data requirements. The LOGO check qualifies this finding: RF shows partial, placement-dependent transfer to newly added gateways, whereas coordinate-only k-NN largely reflects interpolation within the measured deployment.

REFERENCES

- [1] J. Bodenhausen *et al.*, “Securing wireless communication in critical infrastructure: Challenges and opportunities,” in *Mobile and Ubiquitous Systems: Computing, Networking and Services*, A. Zaslavsky *et al.*, Eds. Cham: Springer Nature Switzerland, 2024, pp. 333–352.
- [2] T. Horstmann, M. Rademacher, M. Roobi, and S. Weckmann, “Evaluation of lora in a real-world smart city: Selected insights and findings,” in *Mobile Communication - Technologies and Applications; 27th ITG-Symposium*, 2023, pp. 91–96.
- [3] M. Rademacher, H. Linka, T. Horstmann, and M. Henze, “Path Loss in Urban LoRa Networks: A Large-Scale Measurement Study,” *IEEE Vehicular Technology Conference*, vol. 2021-September, 2021.

- [4] M. Stusek *et al.*, “Accuracy Assessment and Cross-Validation of LP-WAN Propagation Models in Urban Scenarios,” *IEEE Access*, vol. 8, pp. 154 625–154 636, 2020.
- [5] P. Jorke, S. Bocker, F. Liedmann, and C. Wietfeld, “Urban channel models for smart city IoT-networks based on empirical measurements of LoRa-links at 433 and 868 MHz,” in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, vol. 2017-October. IEEE, 2017, pp. 1–6.
- [6] A. Seretis and C. D. Sarris, “An Overview of Machine Learning Techniques for Radiowave Propagation Modeling,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 6, pp. 3970–3985, 1 2021.
- [7] M. Hata, “Empirical Formula for Propagation Loss in Land Mobile Radio Services,” *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980.
- [8] European Commission: Directorate-General for the Information Society and Media, *COST Action 231 – Digital mobile radio towards future generation systems: Final Report*. Publications Office, 1999.
- [9] R. Ballestrin, J. F. Feijó, M. Feldman, and I. Müller, “Exploring Machine Learning Techniques for Path Loss Prediction in LoRa Networks,” *2024 19th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 1–6, 7 2024.
- [10] L. Liu, Y. Yao, Z. Cao, and M. Zhang, “DeepLoRa: Learning accurate path loss model for long distance links in LPWAN,” *Proceedings - IEEE INFOCOM*, vol. 2021-May, 5 2021.
- [11] H. A. Alobaidy *et al.*, “Low-Altitude-Platform-Based Airborne IoT Network (LAP-AIN) for Water Quality Monitoring in Harsh Tropical Environment,” *IEEE Internet of Things Journal*, vol. 9, no. 20, pp. 20 034–20 054, 10 2022.
- [12] M. Gonzalez-Palacio *et al.*, “Machine-Learning-Based Combined Path Loss and Shadowing Model in LoRaWAN for Energy Efficiency Enhancement,” *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10 725–10 739, 6 2023.
- [13] M. Ester *et al.*, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *kdd*, vol. 96, no. 34, 1996, pp. 226–231.
- [14] I. T. Union, “Prediction methods for the terrestrial land mobile service in the VHF and UHF bands,” ITU, ITU-R Recommendation P.529-3, Oct. 1999, withdrawn (suppressed 22 Oct. 2001).
- [15] G. Hufford, W. Kissick, and A. Longley, “A guide to the use of the its irregular terrain model in the area prediction mode,” Institute for Telecommunication Sciences, Tech. Rep., Apr. 1982.
- [16] S. E. Shumate, “Longley-rice and itu-p.1546 combined: A new international terrain-specific propagation model,” in *2010 IEEE 72nd Vehicular Technology Conference - Fall*, 2010, pp. 1–5.
- [17] G. Callebaut and L. Van der Perre, “Characterization of lora point-to-point path loss: Measurement campaigns and modeling considering censored data,” *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1910–1918, 2020.
- [18] J. Petäjäjärvi *et al.*, “On the coverage of LPWANs: Range evaluation and channel attenuation model for LoRa technology,” *2015 14th International Conference on ITS Telecommunications, ITST 2015*, pp. 55–59, 1 2016.