Machine Learning-based mmWave Path Loss Prediction for Urban/Suburban Macro Sites

Guillem Reus-Muns*†, Jinfeng Du†, Dmitry Chizhik†, Reinaldo Valenzuela†, Kaushik R. Chowdhury*
*Institute for the Wireless Internet of Things, Northeastern University, Boston, MA, {greusmuns[1], krc[5]}@ece.neu.edu
†Nokia Bell Labs, New Providence, NJ, {jinfeng.du, dmitry.chizhik, reinaldo.valenzuela}@nokia-bell-labs.com

GLOBECOM 2022 - 2022 IEEE Global Communications Conference | 978-1-6654-3540-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/GLOBECOM48099.2022.10000638

***Abstract*—Millimeter-Wave (mmWave) has great potential to provide high data dates given its large available bandwidth, but its severe path loss and high propagation sensitivity to different environmental conditions make deployment planning particularly challenging. Traditional slope-intercept models fall short in capturing large site-specific variations due to urban clutter, terrain tilt or foliage, and ray-tracing faces challenges in characterizing mmWave propagation accurately with reasonable complexity. In this work, we apply machine learning (ML) techniques to predict mmWave path loss on a link-to-link basis over an extensive set of 28 GHz field measurements collected in a major city of USA, with over 120,000 links from both urban and suburban scenarios, with over 40 dB variation for links at similar distances. Either raw environmental profile (terrain+clutter) of each link or 8 selected expert features are used to either directly predict path loss via regression-based approaches or predict the best performing option out of a pool of theoretical/empirical propagation models. Our evaluation shows that Lasso regression provides the best path loss prediction with a performance (RMSE 8.1 dB) comparable to the per-site slope-intercept fit (RMSE 8.0 dB), whereas model selection method achieves 8.6 dB RMSE, both are significantly better than the best a posteriori 3GPP model (UMa-NLOS, 10.0 dB).**

***Index Terms*—mmWave propagation, path loss estimation, machine learning.**

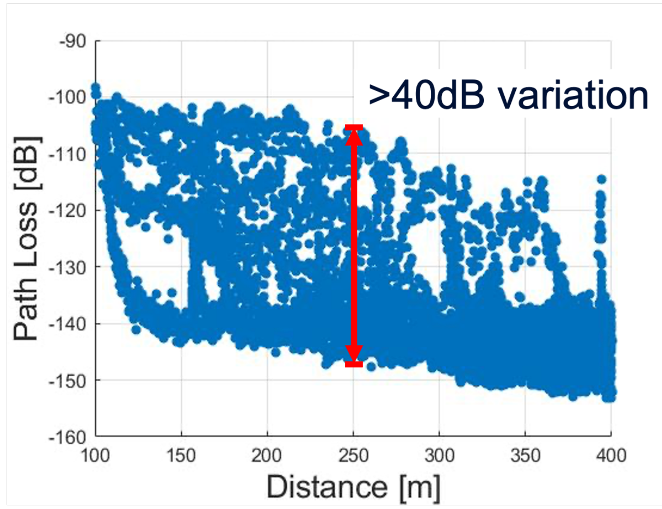

Fig. 1: Path Loss vs Distance measurements in the 28 GHz band, with over 40 dB path loss variation observed at all distances.

I. Introduction

Accurate path loss prediction is crucial in cellular deployment planning and optimization for commercial mobile networks. However, assessing the base station (BS) coverage can be particularly challenging in complex propagation environments, such as dense cities or high density foliage scenarios. Furthermore, mmWave propagation modeling is specially challenging given its sensitivity to blockage and small wavelength. Such characteristics can lead to high variations in path loss values for different links at similar distances, which forces network operators to conduct expensive on-site measurement campaigns in order to assess link coverage or interference between neighboring BSs prior to deployment. For instance, in Fig. 1 we display a set of path loss measurements in the 28 GHz band. It can be observed that over 40 dB differences in path loss are common even for BS-User Equipment (UE) links of similar distances. Given such variability, traditional slope-intercept models fall short in capturing such large variations and are not a feasible candidate for reliable mmWave path loss prediction. Alternatively, ray tracing has been proposed as a possible candidate to model complex scenarios [1], [2]. However, wavelength-scale environment information is required for proper ray tracing modeling, which makes the approach unfeasible when such fine-grained data is not in hand and often computationally intractable when detailed information is available.

Other common approaches to propagation study are empirical models [3], [4], which are obtained from large field measurement campaigns, and theoretical models, which are derived from electromagnetic theory derivations for modeling of the propagation environment. However, models need to be built under certain assumptions and is hard to select what model to use under unseen scenario conditions. Thus, path loss estimation techniques that can generalize to multiple scenarios while using data available at ease are necessary.

Previous works have attempted to investigate the advantages that ML can provide to path loss estimation. In particular, the works in [5]–[7] have focused on path loss interpolation, where data measurements from neighboring locations are assumed to be available during the ML model training phase. We would like to highlight that this paper does not make this assumption and all prediction algorithms are tested on unseen scenarios during training. Additionally, such works only explored sub-6 GHz bands, where the propagation characteristics are considerably less sensitive to blockage or scattering by small objects. The authors in [8] proposed a deep learning approach to estimate the path loss exponent in outdoor mmWave channels. In [9], a neural network (NN) architecture is proposed to estimate outdoors mmWave path loss given topographical information of a certain region of interest. However, both approaches are only evaluated in simulation, and cross-validation in realistic datasets is missing. To the best of our knowledge, only [10] has proposed a ML approach for mmWave path loss estimation and tested using real measurement data. However, the dataset used is for urban street canyons where both transmitter and receiver are along the same street, and the set of features are derived from LiDAR point cloud and 3D building mesh grid.

In this paper, we explore different approaches to mmWave path loss estimation for urban/suburban macro sites using machine learning. In particular, we research how to (i) conduct direct regression based on environmental features including terrain and clutter, and (ii) combine predictions from existing theoretical/empirical path loss models. All our experiments are evaluated using an extensive 28 GHz path loss measurement dataset. Main findings on best path loss predictor include:

- Lasso regression based on 8 expert-defined features, produced 8.1 dB Root Mean Square Error (RMSE), beating best of 3GPP models (UMa 10 dB RMSE), traditional global slope-intercept fit to data (9.9 dB) and Neural Network trained on a full terrain and clutter profile (9.9 dB).
- Neural network trained to select best model among 3GPP and theoretical formulas beat both 3GPP standard models and theoretical models to produce 8.6 dB RMSE.
- We analyze the importance of 8 expert-defined features by analyzing their impact to the overall performance, and identify that **blockage depth** and **parabolic fitting error** are among the most influential features in path loss prediction, in addition to distance.

The rest of the paper is organized as follows. We describe the dataset and the considered theoretical models in Sec. II-A. We describe the path loss estimation and the model selection approach in Sec. III and Sec. IV, respectively. We conclude the paper and discuss potential future work in Sec. V.

II. Background

A. Dataset

We conduct an extensive data collection campaign in a major city of the USA, with a total of 14 macro sites from 10 different rooftops (5 rooftops were urban environments and the other 5 suburban environments). In each case a 4W transmitter emitting a single tone at 28 GHz from an omnidirectional or a sector antenna was mounted on a rooftop macro site. An omnidirectional receive antenna was mounted on a vehicle roof and driven along streets, at distances reaching 500m from the transmitter. In total, mmWave path loss measurements were obtained for over 120,000 links, each of which were based on spatially averaged power from 11 samples taken within 1 meter distance to reduce the effect of small scale fading. The collected data presents an slope-intercept fit of 9.9dB RMS. The best a posterior 3GPP model, UMa-NLOS, provides 10.0 dB RMS despite the fact that neither of the sites are in rural settings.

B. Models

For comparison purposes with the proposed solutions and to support the model selection approach described in Sec. IV, in this subsection we describe the models that were considered in this paper:

- **RMa-NLOS**: 3GPP Rural Macro Cellular NLOS empirical model. Path loss is described with the following equation:

$$\begin{aligned} PL_{\text{RMa-NLOS}} &= 161.04 - 7.1\log_{10}(W) + 7.5\log_{10}(h) \\ &-(24.37 - 3.7(h/h_{BS})^2)\log_{10}(h_{BS}) + (43.42 \\ &-3.1\log_{10}(h_{BS}))(\log_{10}(d) - 3) + 20\log_{10}(f_c) - \\ &(3.2(\log_{10}(11.75h_{UT}))^2 - 4.97) \end{aligned} \tag{1}$$

  where W (default 20 m) is the average street width, h is the average clutter height (default 15 m), h_{BS} is the base station (BS) height, d is the distance between the BS and the receiver, f_c is the center frequency and h_{UT} is the receiver height.
- **UMa-NLOS**: 3GPP Urban Macro Cellular NLOS empirical model. Path loss is described with the following equation:

$$\begin{aligned} PL_{\text{UMa-NLOS}} &= 13.54 + 39.08\log_{10}(d) + \\ &20\log_{10}(f_c) - 0.6(h_{UT} - 1.5) \end{aligned} \tag{2}$$

- **Parabolic**: Analytical model in [11], which approximates the terrain between the transmitter and receiver as a parabola. The fitted parabola is then used to model the propagation of a certain link. The solution is expressed as a sum of modal contributions, expressed in terms of Airy functions, parametrized by distance, fitted terrain curvature and antenna heights [11].
- **Slab+Att**: Analytical model in [12], which models the propagation in a macro cellular scenario with an outdoor terminal on the street level as the path loss in free space plus the attenuation through the street clutter. In particular, we consider the expression [12] below:

$$PL_{\text{slab+ATT}} = \frac{\lambda^2 h^2}{8\pi^2 d^4}(1 + |\Gamma_{\text{g}}|^2) + \frac{\lambda^2 e^{(-\kappa_{\text{v}} r_{\text{v}})}}{16\pi^2 d^2} \tag{3}$$

  where Γ_{g} is the ground reflection coefficient, κ_{v} is the vegetation absorption coefficient (1.6dB/m), and r_{v} is the through-vegetation distance.

III. Path Loss Estimation

In this section, we describe the different ML approaches considered for path loss estimation. We explore two kinds of input data: (i) terrain and clutter vertical cut between transmitter and receiver and (ii) a set of expert-defined features that are considered to be relevant to mmWave propagation.

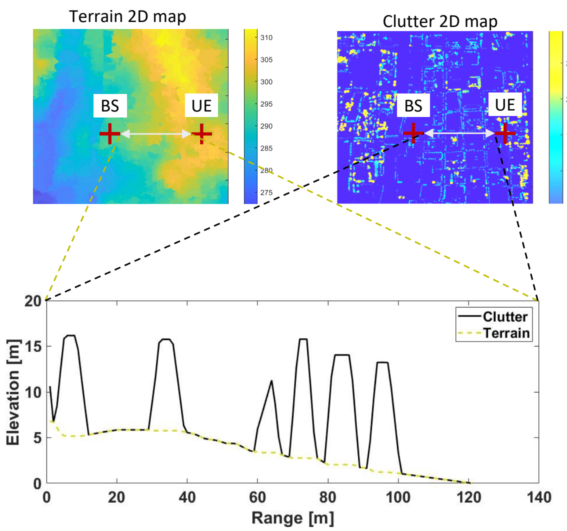

Fig. 2: Terrain and clutter profiles between the Tx-Rx (bottom) are generated from detailed maps of the environment (top).

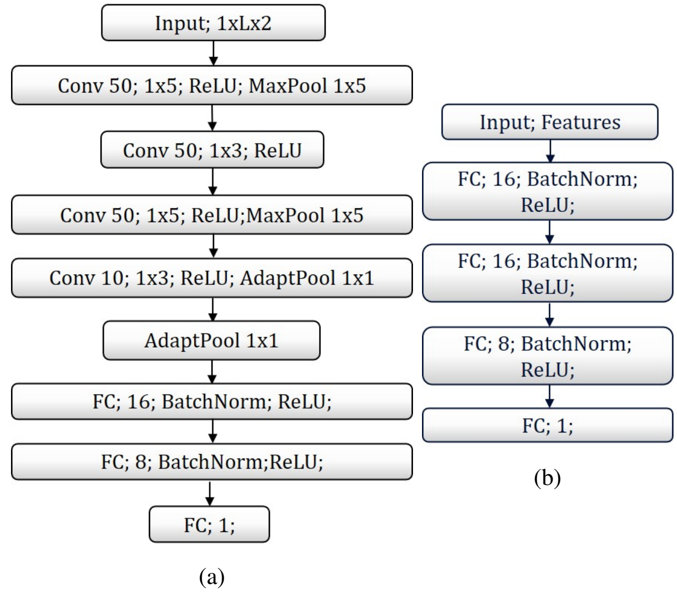

Fig. 3: NN regression architectures using (a) profiles and (b) expert features.

A. Profile-based Path Loss prediction

The 2D vertical cut between the transmitter and the receiver, referred as *profile*, is hypothesised to capture most of the relevant information for propagation modeling, which is necessary for path loss estimation. In particular, we use terrain and clutter (trees and buildings) information, as displayed in Fig. 2. The two different profiles are stored as vectors of size $1 \times L$ and are concatenated along a third dimension, forming a matrix of size $1 \times L \times 2$. We train use the neural network showed in Fig. 3a to perform path loss estimation. Notice that the input of the NN will vary given that the distance between the BS and the UE (L) will not be constant. In order to cope with non-constant input size, we use an adaptive pooling layer with output size 1×1 such that the FC layers will always get an input with constant size. We train the network with a learning rate of $3e^{-4}$, ADAM optimizer with weight decay of 0.1. To reduce overfitting, we define an early stop rate when the training accuracy reaches 6.5 dB RMSE.

B. Expert Feature-based Path Loss Prediction

We define a set of expert features obtained from environment information. We list and define such features below:

- **Log distance**: Logarithm of the 3D distance between the transmitter and receiver.
- **Foliage/Building Blockage Depth**: Given the direct path between transmitter and receiver, we measure the how much depth is blocked by trees or buildings.
- **Parabolic curve fit**: Scalar that defines the curvature of a parabolic fit on the terrain between the transmitter and the receiver.
- **Parabolic curve fit error**: Goodness of the parabolic fit expressed as the RMSE between the parabolic fit and the terrain.
- **Clutter height at terminal** (h_{tc}): Clutter vertical height at the receiver neighboring location.
- **Relative BS height above clutter** (h_{bc}): BS height relative to the neighboring clutter
- **Ratio** α $(\frac{h_{BS}-h_{tc}}{d})$: Ratio between the difference of the BS height and the clutter around the terminal height and the distance.
- **Ratio** β $(\frac{h_{bc}}{d})$: Ratio between the clutter height at the base station and the distance between BS and the terminal.

The set of 8 features is used as an input to different ML models to predict the mmWave path loss. We list and describe such ML models below:

1) *Random Forest*: Technique that combines predictions from other models to output a new prediction (ensemble learning). Here, a number of decision trees are combined to predict a single path loss value. We use a total of 100 trees with no depth constraint.
2) *Fully-Connected Neural Network*: We show the used architecture in Fig. 3b, which comprises a set of FC layers with batch normalization and ReLU activation functions. We train the model for 50 epochs with a learning rate of $3e^{-4}$, weight decay of $1e^{-2}$ and additive Gaussian noise-based data augmentation on the input features with a standard deviation of 0.1.
3) *Linear Regression (Lasso)*: Optimizes the regression weights while imposing a l_1 norm penalty on the regression coefficients. The regression weights are optimized to minimize the MSE error.

C. Results

We evaluate the performance of the profile-based approach and the expert features-based approach for the previously described techniques. In Fig. 4, we show the performance of each approach as a bar plot. The results are obtained applying k-fold cross-validation [13] using the data from each site as

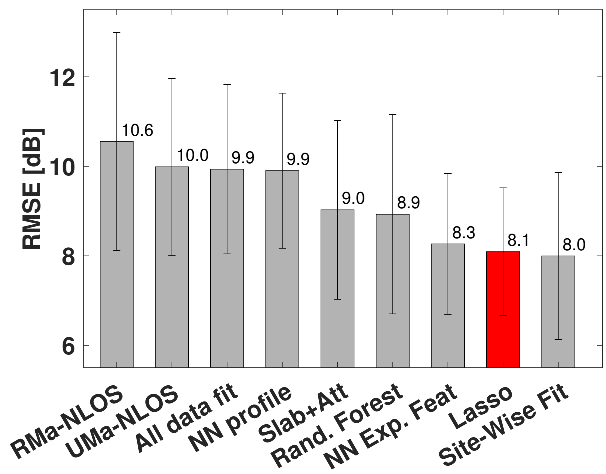

Fig. 4: Path loss estimation results. The error bars indicate ± one standard deviation of the testing RMSE across 14 different sites.

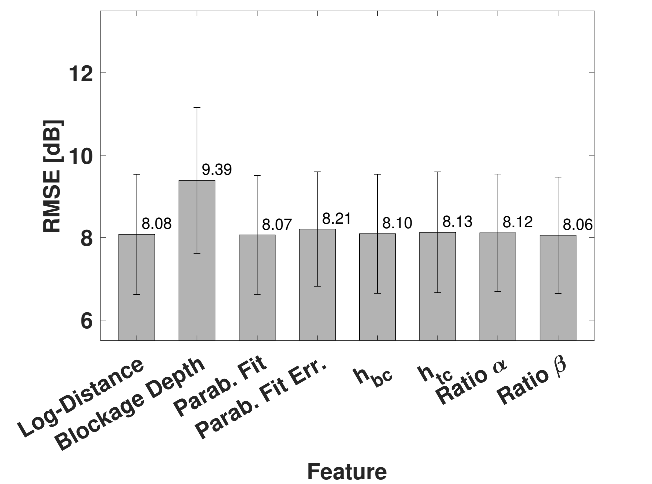

Fig. 5: RMSE obtained after removing each feature from the input set. The presented results are obtained using Lasso regression. The RMSE using all 8 features is 8.1 dB. The error bars indicate ± one standard deviation of the testing RMSE across 14 different sites.

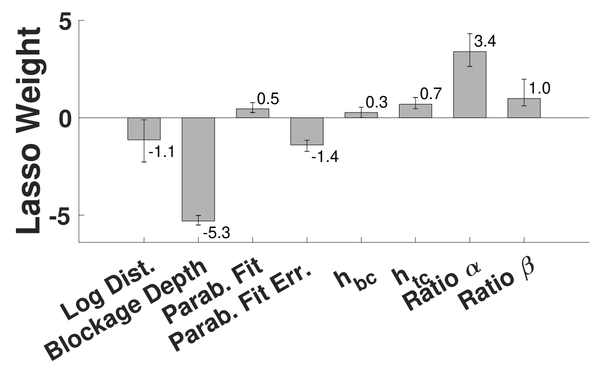

Fig. 6: Lasso regression weights for the selected expert features. The error bars indicate max/min observed deviation from the mean RMSE.

a dataset partition (k=14). Thus, we train each model using data from all sites except one, which is used to evaluate the test performance. This process is repeated for all sites. It can be observed how the Lasso regression with expert features provides the best performance, followed closely by the NN with expert features. The Lasso regression model achieves an RMSE of 8.1 dB, comparable to the slope intercept fit on a per-site basis (8.0 dB). Interestingly, while the *profile* input shares the same information as the selected features, the model overfits to an RMS of 9.9 dB.

D. Feature importance

Here, we analyze the impact of the described set of features in Sec. II-A toward the regression prediction. In particular, we study the impact on the RMSE after removing one feature and keeping the remaining 7.

In Fig. 5, we show the obtained RMSE while discarding one feature a time. This analysis was done selecting the best performing algorithm (Lasso regression). The regression model is fit using all the features except the discarded one. Next, we evaluate each model on all sites following the k-fold cross-validation process described in the previous subsection. The obtained performance after a certain feature is excluded in comparison to the case where all features are used, it can be interpreted as a measure of importance of that feature. It can be observed that the most relevant feature is *blockage depth*, whose removal increases the average RMSE to 9.39 dB, in contrast to 8.1 dB when all 8 features are used. Additionally, in Fig. 6 we plot the Lasso regression weights for every feature. It can be observed how the highest weight corresponds to *blockage depth*. This results matches the previous analysis, where this feature was identified to heavily impact the estimation error.

1) Distance as a feature: Generally, distance is an indispensable feature for modeling path loss/attenuation. However, in Fig. 5 it can be observed that removing *distance* does not seem to impact the overall performance. Here, we clarify that although explicit impact of distance measurement is seemingly small as shown in Fig. 5 and 6, the distance information is implicitly embedded in the ratio α. We show this in Fig. 7, where the it can be observed how the ratio is strongly correlated with the link distance. In particular, we observe the data follows a fit of the form $a * d^b$, with $a = 315.8$ and $b = -1.05$. Notice that this results is expected, given that the distance is in the denominator of the previously defined ratio. Thus, a dependency of the type $\approx 1/d$ is expected.

IV. Model Selection

In this section, we describe an approach that leverages from existing mmWave theoretical/empirical path loss prediction models to obtain a single prediction.

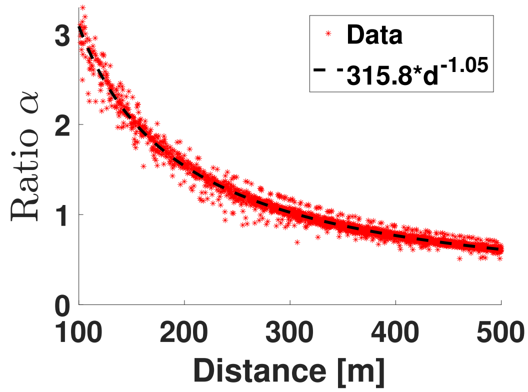

Fig. 7: Ratio α vs distance. Given their strong correlation, radio α and distance can be used interchangeably to model path loss.

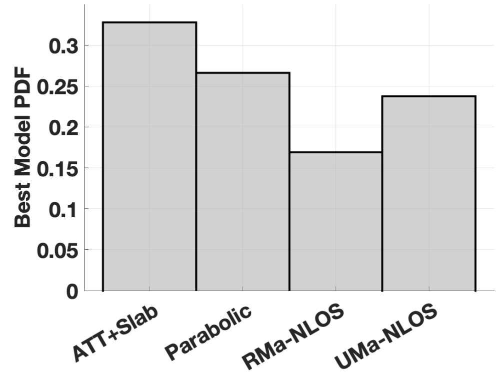

Fig. 8: PDF of Genie-aided model selection, RMSE 5.1 dB

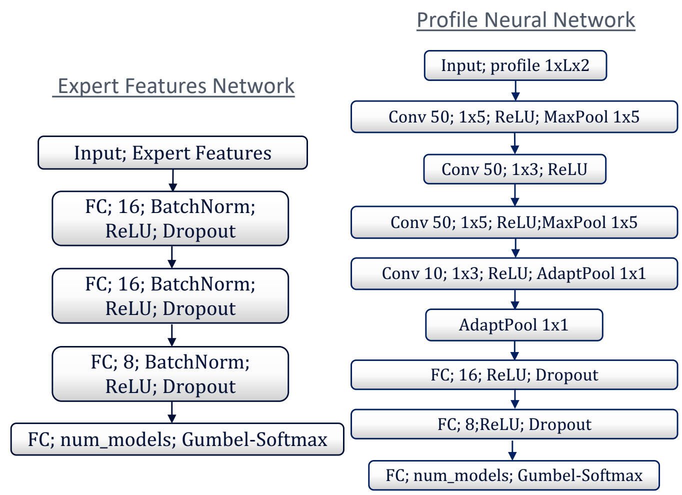

Fig. 9: Model Selection NN Architectures

A. Motivation of the approach

Existing propagation models try to capture specific environment characteristics (i.e. terrain curvature, foliage attenuation) by formulating each one of those phenomenons through electromagnetic theory or trying to fit complex formulas into large measurement data. Each individual model might perform well at certain sites and it is common to see different models complementing each other. For instance, the Slab+Att model achieves an RMSE of 5.8 dB at the site *A*, while the parabolic model falls short with an RMSE of 11.3 dB. On the other hand, the Slab+Att and the parabolic models achieve an RMSE of 11.7 and 8.2 dB, respectively, at a different site (i.e. Site *B*). Thus, the complex mmWave propagation in urban/suburban scenarios makes it particularly challenging to develop a *one-fits-all* propagation model. Here, we propose to intelligently combine a set of pre-existing models based on each particular link terrain and clutter information.

B. Feasibility of the approach

In order to quantify the potential of the approach, we define a *genie-aided* model selector that is able to choose the *best* model on a per-link basis. Here, the *best* model is the one that provides the minimum RMSE. We consider the set of models defined in Sec. II-B. In Fig. 8, we show the best selected model PDF, evaluated in the dataset described in Sec. II-A. We observe that by selecting the best model out of a small set of 4 options, an overall RMSE of 5.1 is achieved, more than 4 dB lower than the slope-intercept fit to all data (9.9 dB). Additionally, the PDF shows how different models complement each other, given that all models are selected a relevant amount of times.

C. Approach

Similarly to the path loss approaches described in Sec. III, we define two model selection techniques, which use either 2D terrain/clutter profile information or a set of expert features. In Fig. 9, we show the two different NN architectures for each of the different approaches. They share the same fundamental structure with the architectures in Fig. 3. Instead, the last layer has as many neurons as the number of models considered (4 in our case) and we use the gumbel-softmax function [14] with hard-decisions right at the output, which returns one-hot vectors of the best predicted model. The gumbel-softmax is an approach that uses the reparameterization trick to approximate the sampling of discrete data while keeping the continuous derivative properties, necessary for back-propagation. Next, the one-hot vector is multiplied with a vector containing all model predictions. Thus, the output of the network will be the path loss prediction of the model which is estimated to better predict the path loss for a certain link. The network is trained by minimizing the MSE between the predictions and the measured path loss for every link.

D. Results

We apply *k*-fold cross-validation to evaluate the performance of each model at every different site, as described in Sec. III-C. We show the obtained results in Fig. 9. It can be observed that, similarly to the path loss prediction presented in the previous section, model selection also performs better using the set of manually selected expert features in

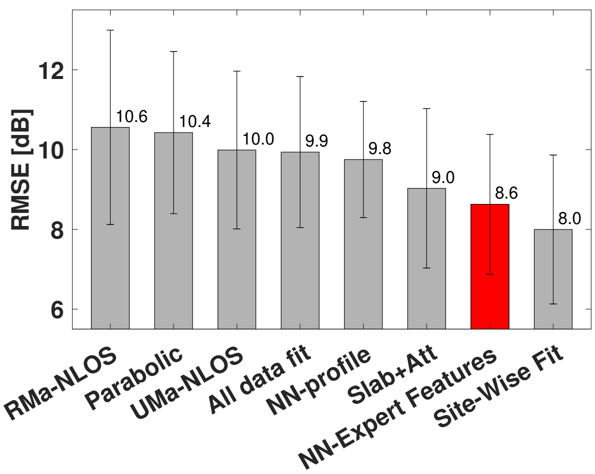

Fig. 10: Model Selection RMSE results. The error bars indicate $\pm$ one standard deviation of the testing RMSE across different sites.

comparison to using profile information. Here, we achieve an RMSE of 8.6 dB using expert features and 9.8 dB using 2D profile information. Notice that the expert features approach achieves a lower RMSE than the best performing theoretical model (Slab+Att). However, the obtained performance is still far from the lower bound of 5.1 dB, as shown in Fig. 8. This is believed to be due to the lack of generalization (i.e. overfitting) of our models, which present a gap of > 2 dB between the training and testing RMSE and > 3 dB with the *genie-aided* model selection.

V. Conclusions and Future Work

In this paper we proposed two different ML approaches to estimate path loss in the 28 GHz band, by either directly predicting the link path loss from input parameters or by formulating the problem as a model selection problem, where we aim to select the best fitting model out of a pool of pre-existing theoretical approaches. Both techniques were evaluated using either 2D profile information or a set of selected expert features. We conclude that the expert features based approach using Lasso linear regression achieves the best performance, even outperforming models that use detailed 2D profiles data. In particular, we achieve an RMSE of 8.1 dB, which is comparable to the 8.0 dB RMSE of the per-site slope-intercept fit, only available after data is collected at all sites. Similarly, expert features also provide superior results to 2D profile information for the model selection approach, which highlights the importance of careful data pre-processing. We believe that some open questions will need to be tackled in future work. For instance, while profiles are assumed to be important to model signal propagation, extra information such as clutter around the receiver should be investigated as well. On the model selection approach, further understanding on the overfitting cause is needed, given its high potential if proper generalization was achieved.

Acknowledgment

We would like to thank Guillaume Roland and Pierre Sabatier for their effort in defining the measurement equipment and process and carrying out the high quality path loss measurements. This material is based on work supported by the U.S. National Science Foundation under award no. CNS 1923789.

References

[1] T. A. Thomas, H. C. Nguyen, G. R. MacCartney, and T. S. Rappaport, “3d mmwave channel model proposal,” in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*. IEEE, 2014, pp. 1–6.

[2] K. Rizk, J.-F. Wagen, and F. Gardiol, “Two-dimensional ray-tracing modeling for propagation prediction in microcellular environments,” *IEEE Transactions on Vehicular Technology*, vol. 46, no. 2, pp. 508–518, 1997.

[3] 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 Transactions on antennas and propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.

[4] Y. d. J. Bultitude and T. Rautiainen, “Ist-4-027756 winner ii d1. 1.2 v1. 2 winner ii channel models,” *EBITG, TUI, UOULU, CU/CRC, NOKIA, Tech. Rep*, 2007.

[5] Y. Egi and C. E. Otero, “Machine-learning and 3d point-cloud based signal power path loss model for the deployment of wireless communication systems,” *IEEE Access*, vol. 7, pp. 42 507–42 517, 2019.

[6] S. P. Sotiroudis, S. K. Goudos, and K. Siakavara, “Neural networks and random forests: A comparison regarding prediction of propagation path loss for nb-iot networks,” in *2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST)*. IEEE, 2019, pp. 1–4.

[7] M. E. Morocho-Cayamcela, M. Maier, and W. Lim, “Breaking wireless propagation environmental uncertainty with deep learning,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5075–5087, 2020.

[8] J.-Y. Lee, M. Y. Kang, and S.-C. Kim, “Path loss exponent prediction for outdoor millimeter wave channels through deep learning,” in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–5.

[9] V. V. Ratnam, H. Chen, S. Pawar, B. Zhang, C. J. Zhang, Y.-J. Kim, S. Lee, M. Cho, and S.-R. Yoon, “Fadenet: Deep learning-based mm-wave large-scale channel fading prediction and its applications,” *IEEE Access*, vol. 9, pp. 3278–3290, 2020.

[10] A. Gupta, J. Du, D. Chizhik, R. A. Valenzuela, and M. Sellathurai, “Machine learning-based urban canyon path loss prediction using 28 ghz manhattan measurements,” *IEEE Transactions on Antennas and Propagation*, 2022. Online available at arXiv:2202.05107.

[11] D. Chizhik, J. Moilanen, S. Klein, L. Maestro, and R. A. Valenzuela, “Analytic propagation approximation over variable terrain and comparison to data,” in *2020 14th European Conference on Antennas and Propagation (EuCAP)*, 2020, pp. 1–3.

[12] D. Chizhik, J. Du, and R. A. Valenzuela, “Universal path gain laws for common wireless communication environments,” *IEEE Transactions on Antennas and Propagation*, 2021.

[13] M. Stone, “Cross-validatory choice and assessment of statistical predictions,” *Journal of the royal statistical society: Series B (Methodological)*, vol. 36, no. 2, pp. 111–133, 1974.

[14] E. Jang, S. Gu, and B. Poole, “Categorical reparameterization with gumbel-softmax,” *arXiv preprint arXiv:1611.01144*, 2016.