

Augmented RF Propagation Modeling

Serhat Tadik[✉], *Graduate Student Member, IEEE*, Michael A. Varner[✉], *Member, IEEE*,
Frost Mitchell, and Gregory D. Durgin[✉], *Senior Member, IEEE*

Abstract—High-speed, accurate RF propagation models are essential for interference prediction between different spectrum users. We propose a simple and effective regression-based estimation method that can be used both for correcting errors in existing propagation models, such as the Terrain Integrated Rough Earth Model (TIREM), and also as a stand-alone empirical propagation model where no prior model is required. A neural network regression model (NNRM) that is informed by real-world channel and propagation characteristics is developed to obtain accurate path loss predictions throughout a 2.9 km by 2.6 km map of the University of Utah. The proposed NNRM is trained and tested on three path loss measurement campaign data sets collected across the campus, resulting in a 58% to 87% reduction in loss difference variance when used to correct TIREM, and a 59% to 76% reduction in measured signal strength variance when used as an empirical propagation model.

Index Terms—Propagation modeling, augmented modeling, digital spectrum twin, regression, diffraction.

I. INTRODUCTION

A DIGITAL Spectrum Twin (DST) is a regularly updated digital representation of radio spectrum that observes usage within a particular geographical area. It integrates information about radio sources, regional geographical information system (GIS) data, spectrum sensing measurements, and propagation maps, to generate an estimate of spectrum activity in frequency, space, and time. With near real-time updates, DSTs are emerging tools for effective spectrum management, including guiding spectrum users and predicting future spectrum utilization [1], [2]. Additionally, DSTs increase spectrum efficiency by enabling more secondary user activity in currently unused frequency bands allocated to primary or incumbent users [1].

The creation of a DST for real-time use in wireless networks requires hybrid propagation models which have accuracies approaching deterministic modeling (e.g., ray tracing) and the speed and simplicity of empirical models (e.g., path-loss exponent). To arrive at such powerful models, we propose a combination of modeling with an established, tried-and-true broadband model (e.g., the Terrain Integrated Rough Earth

Model or TIREM) augmented with a machine learning correction factor that allows for fast, custom-tuned, error-minimized map production. These techniques lower standard deviation error and, in many instances, extend existing models into a wider range of applicable environments.

The performance of the models against radio measurements are shown for 462.7 MHz (UHF Family Radio Service) and 3.55 GHz (Citizens Broadband Radio Service). In particular, we are able to associate individual performance metrics to specific modeling elements for one of the first times in propagation modeling. For example, we find that the most important element for both 462.7 MHz FRS and 3.55 GHz CBRS bands is the shadowing angle of the mobile radio in the model augmentation mode, and distance from the stationary radio in the empirical mode.

A. Propagation Modeling Background

RF propagation modeling plays a critical role in analyzing the feasibility of communication links between terminals and potential interference with incumbent users, which are essential for automating the spectrum sharing process and maintaining the DST. Moreover, accurate propagation models can be useful in analyzing the link budgets of long-range active UHF and SHF radio frequency identification (RFID) systems [3]. There are three main categories of propagation modeling approaches: empirical, ray-tracing, and theoretical models.

Empirical propagation models use statistical methods to generate a predictive fit using data collected within a region, taking into account various factors such as frequency- and distance-dependent attenuation, antenna heights, terrain effects, and street orientation angle [4], [5], [6]. The propagation environment is usually classified as urban, semi-urban, or rural, with model parameters differing for each environment. However, the accuracy of empirical models is limited to environments similar to the ones where the measurements were taken. In dissimilar environments, changes in channel and propagation characteristics may end up with inaccurate predictions [7].

Ray-tracing models calculate path loss by considering a 3D terrain or building profile, and determining rays from a source that take different paths to a terminal. These models are used in various scenarios including monitoring of astronauts [8], assessment of lunar communication links [9], and they use time of arrival information to account for multipath and frequency-selectivity of channels, leading to accurate predictions [10]. However, their complexity makes them computationally expensive and memory intensive [11].

Manuscript received 12 April 2023; revised 17 May 2023; accepted 5 June 2023. Date of publication 12 June 2023; date of current version 30 June 2023. This work was supported by the National Science Foundation under Award 1827940. (Corresponding author: Serhat Tadik.)

Serhat Tadik, Michael A. Varner, and Gregory D. Durgin are with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: serhat.tadik@gatech.edu).

Frost Mitchell is with the Department of Computer Science, University of Utah, Salt Lake City, UT 84112 USA.

Digital Object Identifier 10.1109/JRFID.2023.3285452

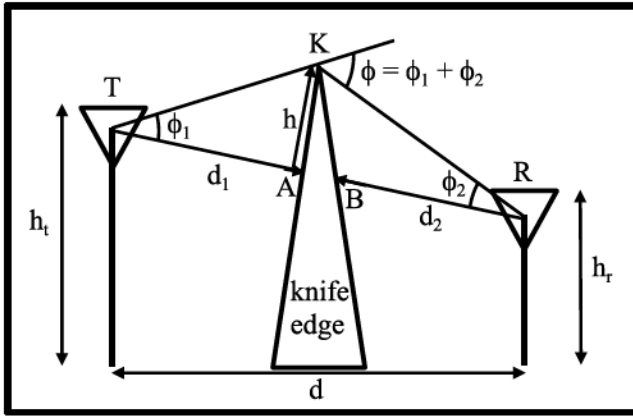


Fig. 1. Knife-edge diffraction geometry for single knife-edge.

Theoretical models of wave propagation incorporate physical phenomena such as free space path loss, reflection, and diffraction. In the case of line-of-sight (LOS) paths, both the direct wave between terminals and the ground-reflected wave are considered when estimating signal strength. For non-line-of-sight (NLOS) paths, on the other hand, diffraction and reflected multipath components are the main contributors to signal strength. The absence of discrete, defined multipath components in theoretical models sets them apart from ray-tracing models. Diffraction becomes the most important factor to consider for NLOS paths. To simplify the diffraction problem, some of the theoretical models assume that the diffracting terrain elements, such as buildings, are infinitely long and perfectly conducting half-planes, referred to as *knife-edges* (KEs), as depicted in Figure 1. Even though theoretical models are more efficient than ray-tracing models and more accurate than empirical models, their approximations and limited consideration of 3-dimensional propagation factors, such as multipath, reduce their overall accuracy.

The Irregular Terrain Model (ITM), also known as the Longley-Rice model, is a widely used theoretical propagation model optimized for irregular terrains, and is best suited for frequencies ranging from 20 MHz to 20 GHz [12]. An extended version of ITM, the Terrain Integrated Rough Earth Model (TIREM), has been developed for a variety of terrains and use cases, which is discussed thoroughly in Section II.

$$\text{Augmented} = \text{Prior} + \underbrace{G_{\text{stat}}(\theta)}_{\text{First Step}} + \underbrace{LD_{\text{pred}}}_{\text{Second Step}} \quad (1)$$

The main objective of this work is to enhance an existing propagation model (*Prior*) by taking into account the radiation pattern of the stationary radio, and addressing its limitations through the use of Equation (1), as described in Section IV. The resulting model denoted as *Augmented*, provides better performance in terms of the mean and standard deviation of the loss difference. Moreover, the augmented modeling scheme and the set of features developed for this purpose can also be used as an independent empirical model, without requiring a pre-existing theoretical model. For the testing of the proposed model, three different measurement sets collected from a college campus are compared to TIREM's predictions and a constant-power map, respectively.

II. TERRAIN INTEGRATED ROUGH EARTH MODEL

A. Overview

The Terrain Integrated Rough Earth Model (TIREM) is an RF propagation model that predicts path loss by considering parameters such as antenna characteristics, medium of propagation, and the digital terrain profile between the radios. To estimate path loss, TIREM first determines the mode of propagation between the radios, based on the frequency range. For the wide range of frequencies between 20 MHz and 1 THz, it considers terrain irregularities and distinguishes between line-of-sight (LOS) and non-line-of-sight (NLOS) channels [13].

For LOS paths, it considers ground reflection, free space loss, and molecular absorption at higher frequencies. For NLOS paths, it makes use of a knife-edge diffraction (KED) engine to estimate path loss, taking into account diffraction losses associated with each knife edge. The KED engine is especially useful when the wave gets depolarized in its path or the polarization of the transmitter is unknown, as it provides a solution that is the geometric mean of the perpendicular- and parallel-polarized Sommerfeld diffraction solutions [14]. The overall diffraction loss value is determined based on the number of knife edges and an empirically set threshold for the average knife-edge loss [13].

The diffraction loss associated with a KE is calculated as a function of the Fresnel-Kirchhoff diffraction parameter, v , shown in Equation (2) [7].

$$v = h \sqrt{\frac{2}{\lambda} \left(\frac{1}{d_1} + \frac{1}{d_2} \right)} \approx \phi \sqrt{\frac{2}{\lambda \left(\frac{1}{d_1} + \frac{1}{d_2} \right)}} \quad (2)$$

where h , d_1 , d_2 , and ϕ (diffraction angle) are illustrated in Figure 1. The ratio of received field strength, E , to the free space field strength measured at the same location, E_0 is

$$F(v) = \frac{E}{E_0} = \frac{(1+j)}{2} \int_v^\infty \exp\left(\frac{-j\pi t^2}{2}\right) dt \quad (3)$$

These expressions are applicable only under certain conditions where $d_1, d_2 \gg h$, and $d_1, d_2 \gg \lambda$, or for relatively small diffraction angles where the path difference between LOS (T-A-B-R) and diffraction (T-K-R) paths contains many wavelengths while remaining smaller than d_1 and d_2 [15]. It has also been shown that the diffracted power obtained from Sommerfeld-polarized solutions and the KED solution diverge from one another in largely shadowed regions [14]. Moreover, the KED approximations cause a distortion in the phase of the diffracted wave specifically at largely shadowed regions.

The TIREM model does not consider the 3D geometry of the propagation environment. It rather restricts its calculations to 2D elevation profile between the two radios. Therefore, it is unable to account for multipath components or frequency selectivity of channels. It is mostly trained for long-range links [13] and underpredicts the signal strength in microcellular regions. It also assumes isotropic radiating elements whereas realistic antennas have non-uniform radiation patterns. Therefore, TIREM's predictions in short-range microcellular environments vary considerably from the reality.

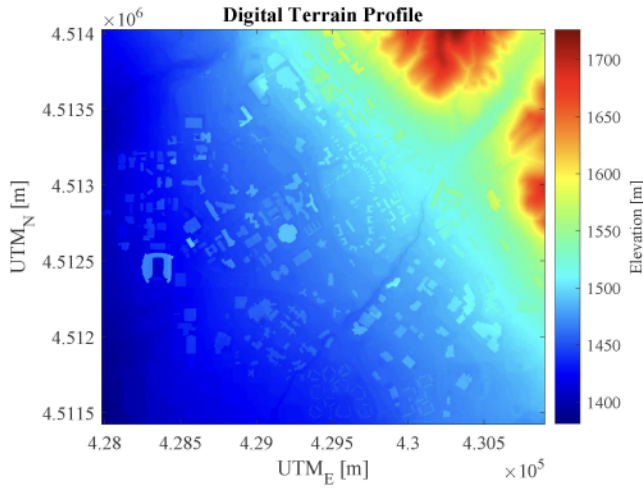


Fig. 2. Digital terrain profile showing a combination of the digital elevation map and the building layout in UTM zone 12. Map dimensions are approximately 2.6 km by 2.9 km with a 0.5 m resolution.

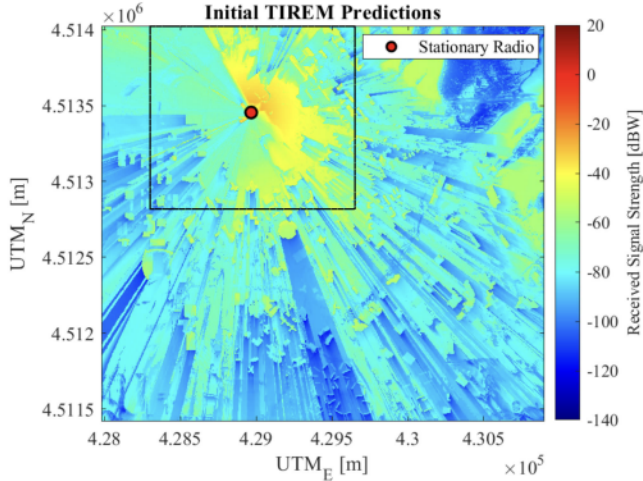


Fig. 3. TIREM's predictions for a 35 dB power adjustment (transmit power plus receiver amplifier gain for measurement system) at 462.7 MHz (FRS band) for all mobile radio locations throughout the Salt Lake City, University of Utah region. The region of radius 690 m where the measurements are taken is illustrated within the black rectangle.

B. Map Generation

To generate path loss predictions for a region, a digital terrain profile (DTP) is necessary. A DTP can be created by merging digital elevation map (DEM) data with a building layout containing building location and height information, demonstrated in [16]. In this work, the DEM of the Salt Lake Valley near Red Butte Canyon mouth is obtained from a LiDAR survey from the State of Utah [17] and combined with the university building layout to create the terrain profile shown in Figure 2.

TIREM can now perform calculations using the 2-D raster, treating each pixel as a possible pairing radio location from a stationary endpoint pixel. Initial TIREM predictions in the region are shown in Figure 3.

III. MEASUREMENTS

The proposed augmented modeling method is evaluated on three real-world RSS datasets, taken on the University of

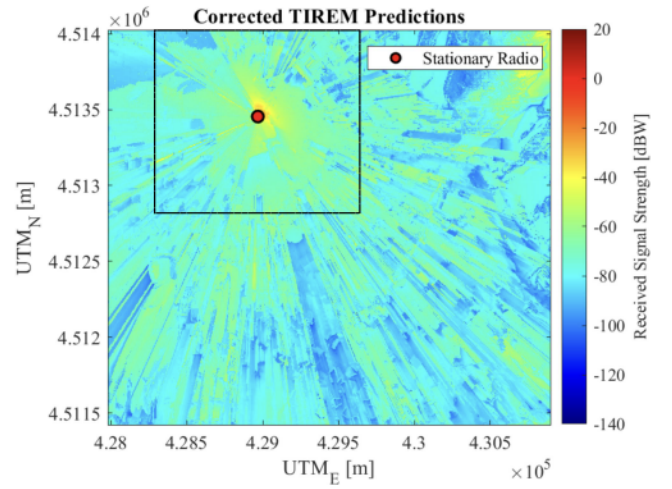


Fig. 4. TIREM 462.7 MHz (FRS band) predictions with corrections for enhanced terrain diffraction, radiation patterns, and roadway orientation. The region of radius 690 m where the measurements are taken is illustrated within the black rectangle.



Fig. 5. Rooftop (left) and Mobile (right) radio placement in the testbed.

Utah campus using the open-access POWDER platform [18]. The testbed includes Ettus USRP X310 radios mounted on buildings and B210 radios on mobile shuttles, shown in Figure 5.

The first measurement set consists of a continuous-wave transmission in the CBRS Band, $f_c = 3550.251$ MHz, from a single rooftop node, labeled USTAR in this work, with three mobile shuttles collecting power measurements. Depicted in Figure 5, the sensor antenna was mounted on the rear window of the shuttles. The shuttles followed regular routes during the collection phase, shown by the three different colors in Figure 6. Although the antenna orientation varied during the measurement campaign, samples in nearby locations also had nearly identical antenna orientations.

This consistent antenna orientation is not the case in our other evaluation datasets, taken from [19], where a portable two-way radio (BaoFeng BF-F8HP) transmitted in the FRS band (462.7 MHz) while two rooftop radios (USTAR and Honors) were used to collect power measurements, shown in Figure 7.

In the CBRS band, measurements were taken with mobile receivers and a fixed transmitter, but in the FRS band, measurements were taken with a mobile handheld transmitter and fixed receivers. Consequently, an extra variable of error is embedded in this set because the antenna is handheld without a gyroscope to keep the radiation pattern consistent across transmit

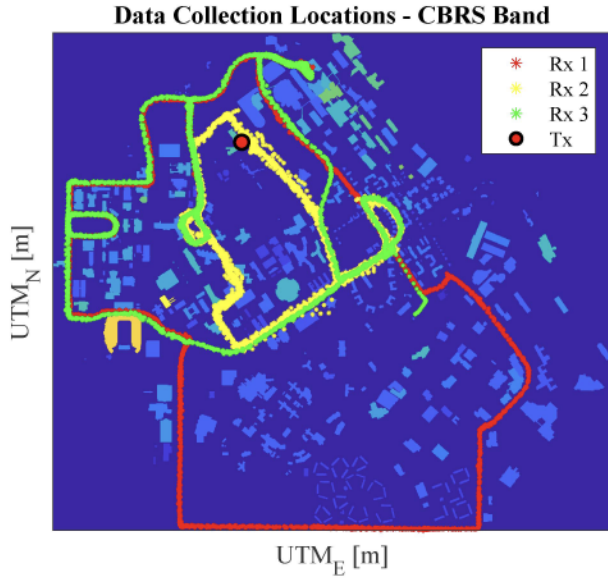


Fig. 6. Data collection locations for the first dataset with 3550.25 MHz (CBRS band) frequency. The single red dot shows the stationary radio's location and different shuttle routes are shown in different colors.

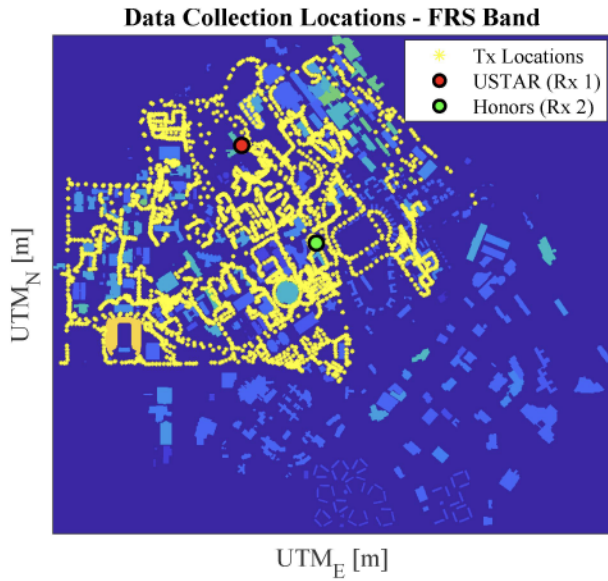


Fig. 7. Data collection locations for the second and third datasets with 462.7 MHz (FRS band) frequency. The red and green dots show the two stationary radio locations and the yellow dots are the mobile radio locations.

locations. Although the device has a whip antenna, considered omni-directional in the horizontal plane, test data with a rotating antenna showed a power delta (maximum - minimum) of 11.5 dB, compared to only 2.8 dB for a stationary antenna.

Some FRS data was taken with the transmitter on a bicycle, potentially introducing more drastic changes in radiation patterns. With these unknown variables of antenna orientation and transmit power, it is assumed that the two FRS datasets pose a greater challenge for consistent propagation modeling.

The maximum ranges of samples from the stationary radio that have received power values above the noise floor are listed in Table I along with the number of samples considered.

As the distance between the terminals increases, the received power values gradually decrease and ultimately fall

TABLE I
PARAMETERS RELATED TO THE MEASUREMENT SETS

Set Name	Frequency (MHz)	Maximum Range (m)	# Samples
CBRS USTAR	3550.251	350	1221
FRS USTAR	462.7	690	1614
FRS Honors	462.7	400	1050

to the noise floor. The maximum ranges up to which the signal strength is above the noise floor are listed for each experiment in Table I along with the number of samples considered.

IV. THEORY

This section presents our augmented modeling scheme for improved path loss predictions, along with an explanation of the working principles of the neural network regression model (NNRM) and the features that are fed into it.

A. Augmented Modeling Scheme Overview

The overall augmentation approach includes generating predictions per pixel using TIREM, as shown in Figure 3, and producing corrections at the pixel level to be reintroduced and improve the original predictions.

Previous efforts to generate corrections of this nature have produced modest improvements, such as using simple minimum mean squared error (MMSE) techniques to compensate for pessimistic shadowing losses, or the integration of clutter models to further categorize corrections of an environmental nature, [16], [20]. However, these approaches only consider a limited number of the factors which impact propagation, where plenty other existing ones are left unaccounted for.

A two-step procedure is adopted to correct the prior propagation model's predictions, expressed in Equation (1). In the first step of this procedure, the radiation pattern of radios is accounted for, which is neglected by most of the existing theoretical models including TIREM. Accordingly, the gain pattern, $G_{stat}(\theta)$, of the stationary radio is generated. Then the sum of the prior model's prediction and the stationary radio's gain is computed at each pixel's elevation angle θ , and this process is repeated for every pixel in the map. If the signal is received from or transmitted to a non-line-of-sight mobile radio, the corresponding observation angle is used instead of the elevation angle.

The second step aims to compensate for other factors that contribute to the path loss difference between the measurements and the prior model's predictions. These factors are either imperfectly handled by the prior model or not considered at all. Features associated with these factors are designed and fed into a fully connected neural network regression model (NNRM) and pixel-level predictions of the loss difference are obtained, denoted as LD_{pred} . Specifically, the difference between the measurements and the prior model's signal strength, denoted as $LD = Measurements - Prior$, is calculated and used as the target to be predicted in the NNRM. The augmented predictions offering more accurate predictions are then obtained by adding LD_{pred} to the values obtained at the end of the first step.

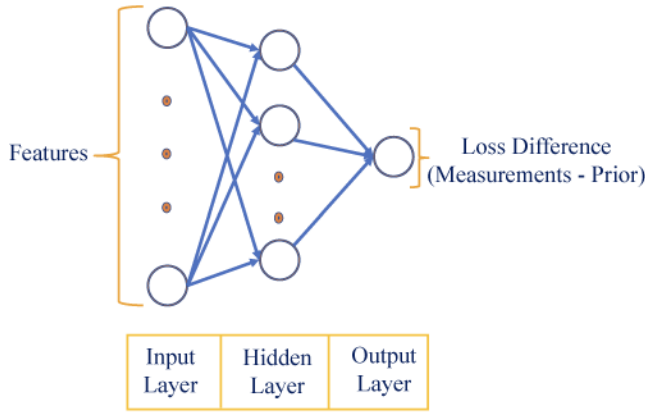


Fig. 8. The neural network regression model.

B. Neural Network Regression Model

A fully connected neural network regression model as shown in Figure 8 is used to predict the LD . This model is chosen because it utilizes an activation function to identify non-linear relationships between features (predictors) and the loss difference (target), thereby incorporating a non-linear component, unlike linear models. Moreover, it doesn't require prior knowledge of the distribution of the data, unlike some other nonlinear models such as the polynomial regression model. In other words, it is able to learn the underlying patterns and relationships in the data without making assumptions about its distribution.

Once a value is predicted in the output layer as a result of the forward pass, it is compared to the ground truth, i.e., LD , through a cost function such as the Ridge (L2) regularized mean squared error (MSE) function shown in Equation (4), where N is the number of samples fed into the model, LD_s is the actual path loss difference between the measurement and TIREM's prediction for the sample s , $g_{\hat{\beta}}(X_s)$ (which can also be denoted as $LD_{pred,s}$) is the path loss difference predicted by the model which is a function of model parameters, $\{b, w\} \in \hat{\beta}$, ρ is the regularization parameter, and M is the total number of parameters constituting the model.

$$C(\hat{\beta}, X, LD) = \underbrace{\frac{1}{N} \sum_{s=1}^N (LD_s - g_{\hat{\beta}}(X_s))^2}_{\text{MSE}} + \underbrace{\rho \sum_{p=1}^M \hat{\beta}_p^2}_{\text{Regularization}} \quad (4)$$

The regularization term in the cost function penalizes large model parameters such that the negative effect of multicollinearity between different independent variables is minimized. The defined cost function is then minimized with a backpropagation algorithm that adjusts model parameters iteratively. The limited-memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm (LBFGS) is used in this work as a cost function minimization technique [21]. This is a technique requiring only a limited amount of memory whose main driver is the gradient of the loss function with respect to model parameters.

The comparison of several configurations of the model is carried out with a cross-validation procedure. The dataset is

divided into training, validation, and testing sets. The training and validation sets were then used to determine the optimal number of hidden layers (1 to 3), the number of nodes in each hidden layer (from a range between 4 and 32), the activation function (from either rectified linear unit (ReLU) or sigmoid), and the regularization parameter, ρ (from a range between 10^{-4} and 10^{-1}). Once the optimal model configuration is determined, the former two sets are combined into a single training set and the trained model is tested on the testing set. The resulting optimal hyperparameters and the corresponding statistics are shared in Section V. The optimal NNRM model is then used to make predictions about the LD throughout the map including the pixels that have no measurements. The corrected map is illustrated in Figure 4.

C. Features

The performance of the NNRM is highly dependent on the quality of the features it's fed. To address the factors that are not considered by the *Prior* model, features that are correlated with LD have been designed. This approach leads to a reduction in the mean and standard deviation of LD , improving the overall predictions.

TIREM is designed mainly for long-range, macrocellular links. Therefore, it tends to underpredict signal strength for short-range microcellular links. Additionally, ground reflections in LOS paths only affect the propagation after a critical distance, which changes the functional dependency of path loss on distance [7]. Furthermore, the non-uniform radiation pattern of mobile radios causes a distance-dependent loss difference, whereas TIREM assumes isotropic radiation. The logarithmic distance to the stationary radio is introduced as a feature to eliminate variation caused by those factors.

Even though the radiation pattern of the stationary radio is taken into account in the first step in the overall correction scheme, the radiation pattern of the mobile radio is still incorrectly represented in TIREM as isotropic. One way to correct that is to calculate elevation angles for the mobile radio and use those angles as features in NNRM.

While the radiation pattern depends on both the elevation and azimuth angles, it is important to note that the azimuthal orientation of the mobile radio is highly variable while the elevation angle remains fairly stationary across the pixels. In other words, the radio is likely to rotate in a 2D horizontal plane during measurements because of the random orientation of roads with respect to the stationary radio. However, it is expected to be stationary vertically most of the time. Hence, the azimuth angle of the radio with respect to the stationary radio does not provide a steady feature that can be used to correct for the radiation while using the elevation angles can serve as a reliable indicator of variations in signal strength due to the radiation pattern. The elevation angles for all pixels in the map with respect to a stationary endpoint are shown in Figure 9.

TIREM's predictions are based only on the 2D vertical cross-section between the radios, disregarding the 3D nature of signal propagation in real-world environments. As a result, TIREM's predictions may be significantly inaccurate when

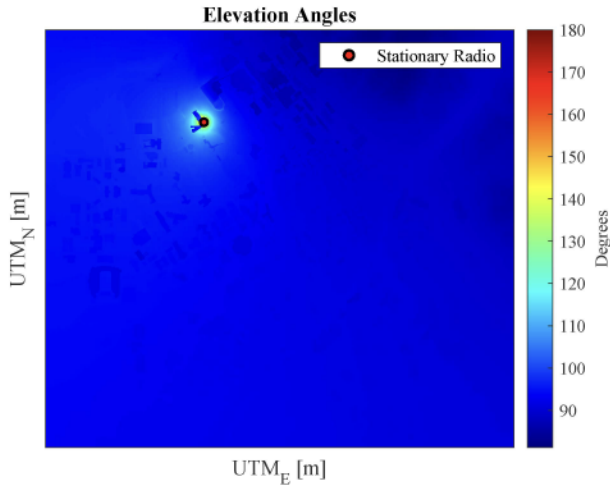


Fig. 9. Elevation angles for all pixels in the map with respect to a stationary endpoint. The angles are calculated taking the positive z-axis as 0 degrees.

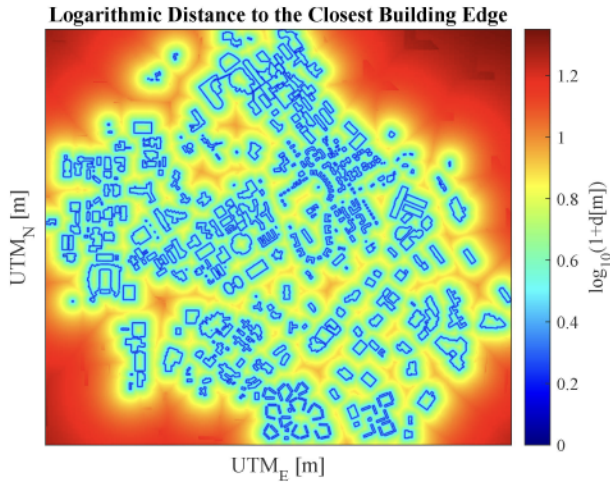


Fig. 10. Logarithmic distance of a mobile radio to the closest building edge around it.

signals propagate via alternative 3D paths that contribute significantly to the overall signal strength. To solve this problem, one approach is to provide NNRM with a measure of distance to the closest building edge, as shown in Figure 10. This information can be useful in identifying less-varying, clear channels in LOS scenarios and clutter density affecting the expected number of multipath components in NLOS scenarios. Despite its contribution to the prediction performance, it should be noted that the generation of this feature is computationally challenging. This is one of the main reasons why a theoretical model, like TIREM, is preferred over a ray-tracing model.

TIREM predicts signal strength separately for the LOS and NLOS paths, as discussed in Section II. The two cases have different factors affecting path loss. The calculation of LOS path loss is relatively simple and less prone to errors compared to NLOS path loss, which requires additional approximations. For example, obstacles are represented as semi-infinite half-planes, and the expressions for field strength are only valid for relatively small diffraction angles. These approximations

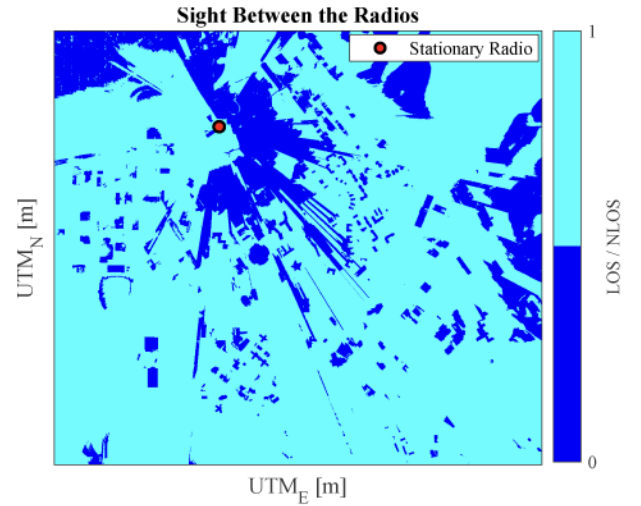


Fig. 11. Binary sight information between the stationary radio and all possible pairing radio locations in the map. The darker blue pixels show line-of-sight, and lighter blue pixels show non-line-of-sight links.

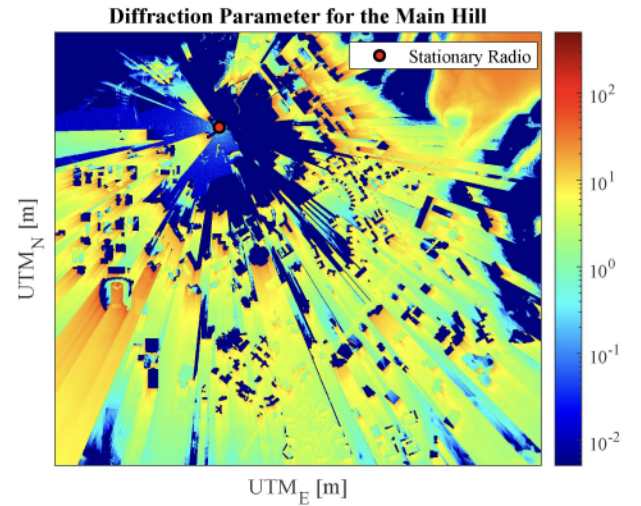


Fig. 12. Diffraction parameter for the main knife edge having the largest Fresnel zone path clearance ratio calculated for 462.7 MHz (FRS band) frequency.

of the KED engine can result in the overprediction of path loss for NLOS paths. To resolve this issue, the sight between the radios is introduced as a feature to correct for LOS and NLOS paths separately, as shown in Figure 11.

Another layer of correction can be applied for NLOS paths by determining how deeply a radio is shadowed by a building or terrain element. The fact that KED relative field strength expression in Equation (3) is only valid for small diffraction angles, or similarly for small Fresnel-Kirchhoff diffraction parameters, ν , means that TIREM's predictions for deeply shadowed regions are likely to be more erroneous than less-shadowed regions. To solve this issue, the diffraction parameter, ν , can be calculated for the main obstacle (having the largest Fresnel zone path clearance ratio) blocking the path between the radios, using Equation (2). The resulting parameters are illustrated in Figure 12.

With a similar motivation, the shadowing angles (whose geometry is shown in Figure 13) for the first and last KEs are

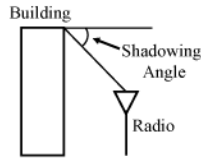


Fig. 13. The geometry of shadowing angle. The horizontal line extending out of the building is parallel to the ground.

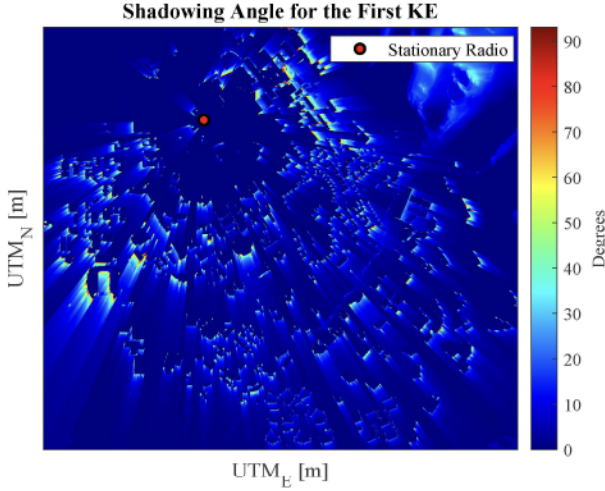


Fig. 14. Shadowing angle for the first knife-edge along the path. The lighter colors indicate larger shadowing.

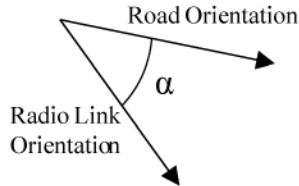


Fig. 15. Road & radio link orientations.

calculated for all pixels in the map and the calculated shadowing angles at the vicinity of the mobile radio are illustrated in Figure 14.

As the mobile antenna is significantly lower than most buildings, signals tend to propagate along the streets, leading to clearer channels and stronger received signals for radio links aligned with the primary road orientation at a given location [6], [22], [23]. However, TIREM does not account for this guiding effect, resulting in potential discrepancies between predictions and measurements.

This problem is approached by using GIS and OpenStreetMap (OSM) software to obtain a road map of the University of Utah region. Principal component analysis is then carried out in the neighborhood of each road pixel to determine its primary road orientation. Finally, the radio link orientation between a stationary radio and a radio located at the road pixel is calculated and compared to the road orientation, as shown in Figure 15.

The alignment between the road orientation and the radio link direction is calculated by computing the absolute value of the cosine of the angle between them, i.e., $\cos(\alpha)$. The resulting alignment map is presented in Figure 16, where roads

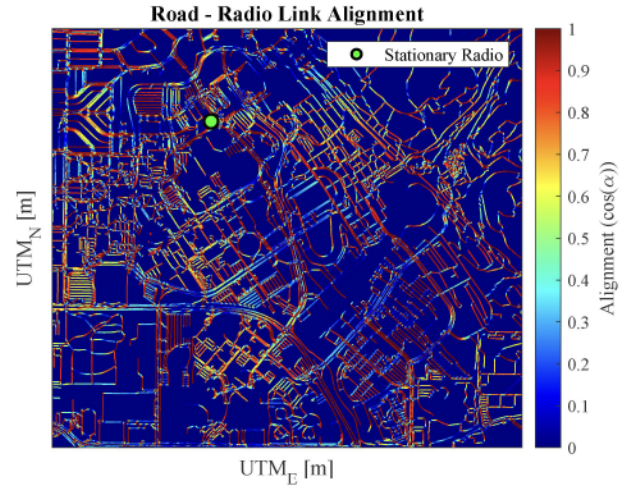


Fig. 16. Alignment between radio link and road orientations. The redder roads are more aligned with the radio link.

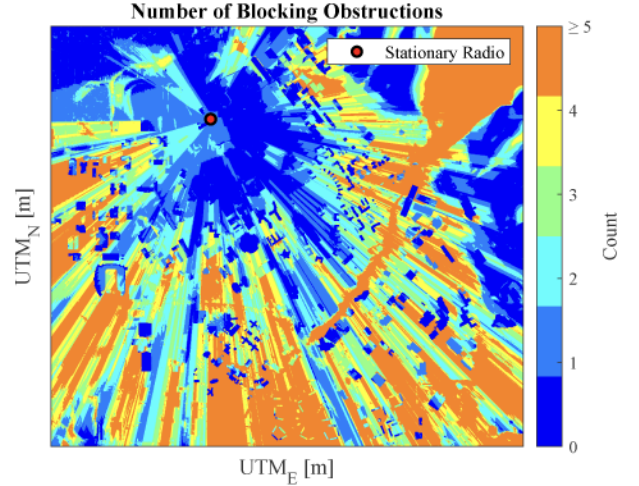


Fig. 17. Number of blocking obstructions between the two radios. Different colors represent different numbers of buildings obstructing the path.

that are more closely aligned with the radio link show larger values.

TIREM's diffraction loss calculation method includes empirical thresholding based on the number of knife edges and the mean loss due to these knife edges [13]. As the number of blocking obstructions in the signal path increases, multipath components might contribute more to the overall signal strength, leading to greater discrepancies. The number of blocking obstructions in the path between the two radios is calculated, and the resulting map is shown in Figure 17.

V. RESULTS & ANALYSIS

Path loss values obtained from collected measurements are used for two different modeling approaches:

- *Model Correction*: to correct initial TIREM predictions throughout the map (*Prior* in Equation (1) is TIREM's predictions).
- *Empirical*: to make predictions when no prior model is available (*Prior* in Equation (1) is some constant power

TABLE II

PEARSON CORRELATION COEFFICIENTS BETWEEN FEATURES, AND LOSS DIFFERENCE (LD), LINE-OF-SIGHT LOSS DIFFERENCE (LOS LD), AND NON-LINE-OF-SIGHT LOSS DIFFERENCE (NLOS LD). THE BLUE VALUES ON THE LEFT-HAND SIDE REPRESENT TIREM LD CORRECTION RESULTS WHEREAS THE RED VALUES ON THE RIGHT-HAND SIDE REPRESENT EMPIRICAL RESULTS, WHICH ARE OBTAINED DIRECTLY FROM OBSERVATIONS WITHOUT RELYING ON A PRIOR MODEL

Feature Name	3.55 GHz (CBRS) USTAR						462.7 MHz (FRS) USTAR						462.7 MHz (FRS) Honors					
	LD	LD	LOS LD	LOS LD	NLOS LD	NLOS LD	LD	LD	LOS LD	LOS LD	NLOS LD	NLOS LD	LD	LD	LOS LD	LOS LD	NLOS LD	NLOS LD
LOS / NLOS	0.75	-0.31	N/A	N/A	N/A	N/A	0.62	-0.43	N/A	N/A	N/A	N/A	0.63	-0.43	N/A	N/A	N/A	N/A
Logarithmic Distance to the Closest Edge	-0.13	0.06	-0.09	-0.03	-0.31	0.15	-0.01	0.20	0.17	0.21	0.02	0.14	-0.15	0.20	0.21	0.21	-0.12	0.14
Logarithmic Distance to the Stationary Radio	-0.01	-0.82	-0.65	-0.83	-0.42	-0.76	-0.15	-0.67	-0.22	-0.66	-0.29	-0.73	-0.23	-0.67	-0.39	-0.66	-0.44	-0.73
Elevation Angle for the Mobile Radio	-0.40	-0.66	-0.65	-0.82	-0.49	-0.55	-0.43	-0.41	-0.54	-0.54	-0.37	-0.56	-0.36	-0.41	-0.47	-0.54	-0.34	-0.56
Road-Radio Link Alignment	-0.04	0.02	0.01	0.03	-0.04	0.00	-0.08	-0.09	-0.26	-0.07	-0.04	-0.07	-0.16	-0.09	-0.14	-0.07	-0.11	-0.07
Number of Blocking Obstructions	0.72	-0.37	N/A	N/A	0.27	-0.31	0.50	-0.52	N/A	N/A	0.13	-0.42	0.67	-0.52	N/A	N/A	0.41	-0.42
Logarithmic Shadowing Angle for the First KE	0.54	-0.28	N/A	N/A	0.34	-0.23	0.50	-0.41	N/A	N/A	0.14	-0.19	0.68	-0.41	N/A	N/A	0.42	-0.19
Logarithmic Shadowing Angle for the Last KE	0.45	-0.24	N/A	N/A	0.34	-0.20	0.58	-0.27	N/A	N/A	0.30	0.13	0.76	-0.27	N/A	N/A	0.66	0.13
Logarithmic Fresnel-Kirchhoff Diffraction Parameter for the Main KE	0.76	-0.34	N/A	N/A	0.26	-0.25	0.62	-0.45	N/A	N/A	0.12	-0.28	0.65	-0.45	N/A	N/A	0.43	-0.28

TABLE III

INDIVIDUAL FEATURE CONTRIBUTIONS TO THE MODEL PERFORMANCE CALCULATED FROM THE INITIAL STANDARD DEVIATION OF LD, AND THE REDUCED STANDARD DEVIATION WHEN THE RELEVANT FEATURE ALONE IS FED INTO THE NNRM. AN AVERAGE OF 20 RANDOM SEEDS. LARGER VALUES INDICATE GREATER CONTRIBUTION. THE BLUE VALUES ON THE UPPER SIDE OF THE SLASH REPRESENT TIREM LD CORRECTION RESULTS WHEREAS THE RED VALUES ON THE LOWER SIDE OF THE SLASH REPRESENT EMPIRICAL RESULTS, WHICH ARE OBTAINED DIRECTLY FROM OBSERVATIONS WITHOUT RELYING ON A PRIOR MODEL

Feature Name	3.55 GHz (CBRS) USTAR	462.7 MHz (FRS) USTAR	462.7 MHz (FRS) Honors
	σ_{init}	σ_{init}	σ_{init}
	13.56 11.08	12.65 10.27	14.18 12.47
Individual Feature Contributions ($\sqrt{\sigma_{init}^2 - \sigma_{init+feature}^2}$)			
LOS / NLOS	10.67 2.77	7.71 5.02	8.21 4.58
Logarithmic Distance to the Closest Edge	2.33 2.07	0.00 3.33	1.44 2.80
Logarithmic Distance to the Stationary Radio	4.74 9.60	2.47 6.03	1.90 9.26
Elevation Angle for the Mobile Radio	9.55 8.80	7.31 4.11	5.10 8.09
Road-Radio Link Alignment	0.38 0.00	1.49 0.98	0.44 0.00
Number of Blocking Obstructions	10.76 3.66	7.78 5.51	9.32 4.66
Logarithmic Shadowing Angle for the First KE	10.85 3.60	7.71 4.77	9.03 4.31
Logarithmic Shadowing Angle for the Last KE	10.20 3.80	8.05 4.99	10.48 7.37
Logarithmic Fresnel-Kirchhoff Diffraction Parameter for the Main KE	10.80 3.68	7.76 5.32	9.19 4.72

corresponding to transmit power plus receiver amplifier gain for measurement system).

The blue values in Tables II, III, IV, and V correspond to the results obtained from the first modeling approach, i.e., the model correction mode, while the red values denote the results obtained from the second approach, i.e., the empirical mode.

The Pearson correlation coefficients between the features and LD as well as the LOS/NLOS components of the LD for

the three measurement sets are presented in Table II. Most of the features are either correlated with the LOS component of LD, NLOS component of LD, or both. Note that some features such as the diffraction parameter and shadowing angle or LOS/NLOS and the number of blocking obstructions are correlated with each other to a certain degree. However, the regularization term in the cost function reduces the negative effects of multicollinearity between features. Features having

TABLE IV

MMSE AND NNRM LOSS DIFFERENCE STATISTICS, AN AVERAGE OF 20 RANDOM SEEDS. THE BLUE VALUES ON THE UPPER SIDE OF THE SLASH REPRESENT TIREM LD CORRECTION RESULTS WHEREAS THE RED VALUES ON THE LOWER SIDE OF THE SLASH REPRESENT EMPIRICAL RESULTS, WHICH ARE OBTAINED DIRECTLY FROM OBSERVATIONS WITHOUT RELYING ON A PRIOR MODEL

Set Name	Loss Difference Statistics					
	$m_d - m_{NNRM}$	σ_d	σ_{MMSE}	σ_{NNRM}	$\sigma_d - \sigma_{NNRM}$	$\frac{(\sigma_d^2 - \sigma_{NNRM}^2)}{\sigma_d^2}$
3.55 GHz (CBRS) USTAR	$\frac{-0.06}{-0.23}$	$\frac{13.56}{11.08}$	$\frac{6.98}{6.13}$	$\frac{4.94}{5.46}$	$\frac{8.62}{5.62}$	$\frac{86.73\%}{75.72\%}$
462.7 MHz (FRS) USTAR	$\frac{-0.35}{-0.41}$	$\frac{12.65}{10.27}$	$\frac{9.19}{6.94}$	$\frac{8.23}{6.61}$	$\frac{4.42}{3.66}$	$\frac{57.67\%}{58.58\%}$
462.7 MHz (FRS) Honors	$\frac{-0.03}{0.66}$	$\frac{14.18}{12.47}$	$\frac{9.08}{7.07}$	$\frac{8.18}{7.01}$	$\frac{6.00}{5.46}$	$\frac{66.72\%}{68.40\%}$

TABLE V

NNRM OPTIMAL HYPERPARAMETERS OBTAINED AS A RESULT OF CROSS-VALIDATION. THE BLUE VALUES ON THE UPPER SIDE OF THE SLASH REPRESENT TIREM LD CORRECTION RESULTS WHEREAS THE RED VALUES ON THE LOWER SIDE OF THE SLASH REPRESENT EMPIRICAL RESULTS, WHICH ARE OBTAINED DIRECTLY FROM OBSERVATIONS WITHOUT RELYING ON A PRIOR MODEL

Set Name	Optimal Hyperparameters			
	# Hidden Layers	# Nodes	Regularization Parameter, ρ	Activation Function
3.55 GHz (CBRS) USTAR	$\frac{1}{1}$	$\frac{8}{8}$	$\frac{1 \times 10^{-2}}{1 \times 10^{-1}}$	$\frac{\text{ReLU}}{\text{ReLU}}$
462.7 MHz (FRS) USTAR	$\frac{1}{1}$	$\frac{8}{8}$	$\frac{1 \times 10^{-1}}{1 \times 10^{-1}}$	$\frac{\text{ReLU}}{\text{ReLU}}$
462.7 MHz (FRS) Honors	$\frac{1}{1}$	$\frac{8}{8}$	$\frac{1 \times 10^{-2}}{1 \times 10^{-2}}$	$\frac{\text{ReLU}}{\text{ReLU}}$

large correlations with LOS LD such as elevation angles, and distance to the stationary radio are the main factors of correction in LOS paths. Features having large correlations with NLOS LD such as shadowing angles, and the diffraction parameters are the main factors of correction in NLOS paths. For each measurement set, the features with the largest correlations are in bold. The results show that the elevation angle, distance to the stationary radio, shadowing angle, diffraction parameter for the main hill, and sight between the radios are among the most correlated features.

The individual contributions of features to the correction performance are presented in Table III. As expected, features with the largest correlations tend to contribute the most. Specifically, in model correction mode, the diffraction-related features dominate performance, while in empirical mode, the distance and elevation angle features are the main contributors. The effect of distance to the closest edge feature on performance is moderate, while the contribution of road-radio link alignment is less clear due to relatively low correlations.

However, as the maximum range of the experiment increases (e.g., FRS USTAR set), the contribution of the road-radio link alignment feature becomes more visible. This is because the effect of road and radio link alignment on signal strength can only be observed when there are enough roads with varied alignments in the measurement area and there are enough measurements taken along those roads. At appreciable distances, roads and their clearings can act loosely as waveguides, allowing more signal to reach further regions easier than TIREM might predict due to immediate, direct path obstacles.

The two correction techniques, namely MMSE and NNRM, are subsequently compared. Table IV lists the initial and corrected LD statistics for each technique. The testing results show that NNRM outperforms MMSE by up to 2.04 dB in terms of corrected LD standard deviation as it can consider nonlinear feature-target relations and interactions between different features. It is numerically more stable compared to MMSE since no matrix inversion is required. Notably, the newly suggested augmented modeling scheme and techniques can reduce LD variance by up to 86.73% in the correction mode and by up to 75.72% in the empirical mode.

The reason that CBRS USTAR measurement showed the best correction performance is that the mobile radio used in the campaign was a higher-grade device and relatively stable in the elevation plane. Additionally, the relatively short maximum range considered during the CBRS USTAR measurement is the region where the features such as elevation angles and shadowing angles are correlated the most with the LD. The highly nonisotropic transmitter antenna used in the FRS measurement campaign in the azimuth plane as well as the arbitrary orientation of the antenna during measurements cause unrecoverable variations in the signal strength further contributing to the variance of the predictions in FRS sets.

Hyperparameters leading to the reported LD statistics are listed in Table V. Best performances are obtained with 1 hidden layer and 8 nodes within the hidden layer for all three sets but with varying regularization parameters.

The corrected TIREM predictions throughout the map for the FRS USTAR experiment are illustrated in Figure 4. The region where measurements are taken is bounded by the black rectangle. In the corrected map, the average overall signal

strength is increased by an amount approximately equal to the original average of LD. Other immediately noticeable changes in the corrected map are that in LOS paths signal strength predictions become smaller whereas in NLOS paths they become larger, eliminating the abrupt change in signal strength around obstacles. It is important to note that the confidence in the corrected values is greater within the black rectangle compared to the outside of the measurement region. This is mainly because the samples taken in a region reflect the propagation and channel characteristics of that particular region. There is no guarantee that the distribution of the LD is going to be similar for a given feature inside and outside of the measurement region. However, the extrapolated LD predictions provide an initial correction for the signal strength outside of the region if common features are determinable.

The resulting statistics illustrate the change in standard deviation when the transmit power and the receiver amplifier gain are known and taken into account. However, the same correction scheme can be applied even when these parameters are unknown. In other words, the model can predict the LD with an unknown power bias, likely with degraded performance compared to the case where the parameters are known.

VI. CONCLUSION & FUTURE WORK

The work presented here considers the channel and propagation characteristics of radio waves and attempts to correct for enhanced terrain diffraction, radiation patterns, sight, distance, shadowing, roadway orientation, and multipath effects. It is based on a fully connected neural network regression model that outperforms the traditionally used MMSE technique as it is able to consider nonlinear feature-target relations, and interactions between different features. Moreover, it is numerically more stable compared to MMSE. Although the corrections are designed for and applied to a specific propagation model, namely TIREM, in the model correction mode, this approach can be adapted to any model facing similar limitations.

The suggested augmented modeling approach reduces test LD variance by up to 86.73% in the correction mode and by up to 75.72% in the empirical mode. Two correction techniques, namely MMSE and NNRM, are compared in terms of standard deviation reduction performance. The results show that NNRM outperforms MMSE by up to 2 dB in terms of corrected LD standard deviation.

Once the desired features and initial predictions are generated for a given transmitter in the region of interest, the augmentation process takes no longer than a few minutes as the NNRM model is a simple 2-layer architecture requiring a limited amount of computation. The time required in the initial process of generating features and *Prior* predictions depends heavily on the map size and resolution, and the *Prior* model. Generating the initial TIREM predictions and features except for the distance to the closest edge feature throughout a map of size 2.9 km by 2.6 km with 0.5 m resolution takes approximately 7 hours in MATLAB on a 12th Gen Intel(R) Core(TM) i7-12700H computer. However, since the initial predictions and most of the features at a pixel are independent of those at

other pixels, the computation is parallelizable among multiple machines.

While the current work has made significant progress in augmenting single-transmitter propagation models, there is a potential for testing the designed features at larger distances with a larger and more diverse set of samples observing the generalizability of the proposed method and the correction performance. Another path that is worthwhile to pursue is to design new features and generalize the tuning process for the case where multiple transmitters are operating simultaneously across the region, a process called *macro-tuning*.

An area for potential future research is to use alternative models that can achieve satisfactory performance, even when the number of available samples is insufficient. In that case, a semi-supervised regression model may be used to learn a function that maps predictors to the target [24], [25]. The idea behind semi-supervised learning is to use the labeled (measured) data to train a supervised regression model, then use the learned function to make predictions for the unlabeled data. The predicted values for the unlabeled data are then used to improve the accuracy of the regression model. This iterative process that leverages the unlabeled data along with the labeled ones may lead to improved performance even when the amount of data is insufficient.

REFERENCES

- [1] G. D. Durgin, M. A. Varner, N. Patwari, S. K. Kasera, and J. Van der Merwe, "Digital spectrum twinning for next-generation spectrum management and metering," in *Proc. IEEE 2nd Int. Conf. Digit. Twins Parallel Intell. (DTPI)*, 2022, pp. 1–6.
- [2] G. D. Durgin et al., "Digital spectrum twinning and the role of RFID and backscatter communications in spectral sensing," in *Proc. IEEE Int. Conf. RFID Technol. Appl. (RFID-TA)*, 2021, pp. 89–92.
- [3] I. B. Shirokov, "Increasing of operation range of system of RFID and positioning," *IEEE J. Radio Freq. Identification*, vol. 4, no. 4, pp. 444–451, Dec. 2020.
- [4] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.
- [5] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Electr. Commun. Lab.*, vol. 16, pp. 825–873, Jan. 1968.
- [6] J. Walfisch and H. Bertoni, "A theoretical model of UHF propagation in urban environments," *IEEE Trans. Antennas Propag.*, vol. 36, no. 12, pp. 1788–1796, Dec. 1988.
- [7] T. Rappaport, *Wireless Communications: Principles and Practice*. Hoboken, NJ, USA: Prentice Hall, 1996.
- [8] N. Panunzio, C. Occhiuzzi, and G. Marrocco, "Propagation modeling inside the international space station for the automatic monitoring of astronauts by means of epidermal UHF-RFID sensors," *IEEE J. Radio Freq. Identification*, vol. 5, no. 2, pp. 174–181, Jun. 2021.
- [9] R. C. Toonen, S. L. Booth, B. W. Welch, and M. J. Zemba, "Optimizing lunar map partitioning for multipath fade loss analyses," *IEEE J. Radio Freq. Identification*, vol. 6, pp. 284–291, Mar. 2022.
- [10] K. Schaubach, N. Davis, and T. Rappaport, "A ray tracing method for predicting path loss and delay spread in microcellular environments," in *Proc. Veh. Technol. Soc. 42nd VTS Conf. Front. Technol.*, vol. 2, 1992, pp. 932–935.
- [11] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015.
- [12] G. A. Hufford, A. G. Longley, and W. A. Kissick, "A guide to the use of the ITS irregular terrain model in the area prediction mode," U.S. Dept. Commer., Nat. Telecommun. Inf. Admin., Washington, DC, USA, NTIA Rep. 82-100, Apr. 1982.
- [13] D. Eppink and W. Kuebler, *TIREM/SEM Handbook*. Chicago, IL, USA: IIT Res. Inst., Sep. 1986.

- [14] G. D. Durgin, "The practical behavior of various edge-diffraction formulas," *IEEE Antennas Propag. Mag.*, vol. 51, no. 3, pp. 24–35, Jun. 2009.
- [15] G. Millington, R. Hewitt, and F. Immirzi, "Double knife-edge diffraction in field-strength predictions," *Proc. IEE-Part C Monographs*, vol. 109, pp. 419–429, Sep. 1962.
- [16] M. A. Varner, F. Mitchell, J. Wang, K. Webb, and G. D. Durgin, "Enhanced RF modeling accuracy using simple minimum mean-squared error correction factors," in *Proc. IEEE 2nd Int. Conf. Digit. Twins Parallel Intell. (DTPPI)*, 2022, pp. 1–5.
- [17] "State of utah acquired lidar data - wasatch front," in *Utah Automated Geographic Reference Center and Utah Geological Survey, Distributed by OpenTopography*, New Delhi, India, UAGRC and UGS: Jan. 2015.
- [18] J. Breen et al., "Powder: Platform for open wireless data-driven experimental research," in *Proc. 14th Int. Workshop Wireless Netw. Testbeds Exp. Eval. AMP Characterization*, New York, NY, USA, 2020, pp. 17–24. [Online]. Available: <https://doi.org/10.1145/3411276.3412204>
- [19] F. Mitchell, A. Baset, S. K. Kasera, and A. Bhaskara, "A dataset of outdoor RSS measurements for localization," Oct. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.7259895>
- [20] C. R. Anderson, "An integrated terrain and clutter propagation model for 1.7 GHz and 3.5 GHz spectrum sharing," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5804–5818, Jul. 2022.
- [21] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.
- [22] G. D. Durgin, "Electro-magnetic propagation modeling," U.S. Patent 7 433 652 B2, Oct. 7, 2008.
- [23] F. Ikegami, T. Takeuchi, and S. Yoshida, "Theoretical prediction of mean field strength for urban mobile radio," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 299–302, Mar. 1991.
- [24] C. Cortes and M. Mohri, "On transductive regression," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 19, 2006, pp. 305–312.
- [25] M. Timilsina, A. Figueroa, M. d'Aquin, and H. Yang, "Semi-supervised regression using diffusion on graphs," *Appl. Soft Comput.*, vol. 104, Jun. 2021, Art. no. 107188.



Serhat Tadik (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering and physics from Boğaziçi University in 2022. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the Georgia Institute of Technology. His research interests include the application of machine learning techniques in propagation modeling, spectrum allocation, and data science specifically focusing on graph analysis.



Michael A. Varner (Member, IEEE) received the B.S.E.E. degree from the Rose-Hulman Institute of Technology in 2015 and the M.S.E.E. degree from Georgia Tech in 2017, where he is currently pursuing the Ph.D. degree with the Propagation Group under the supervision of Prof. G. D. Durgin. He joined the School of Electrical and Computer Engineering, Georgia Institute of Technology in 2015. His previous research areas include the development of next-gen electromagnetic compatibility methods and devices, novel antenna and microwave structure design, and digital spectrum twins and augmented propagation models. His current primary research is in ambient scatter communications, the focus of his doctoral dissertation, for the creation of spectrum agile and RF power recycling networks. He is a previous recipient of the Georgia Tech Presential Fellowship.



Frost Mitchell received the B.S. degree in computer engineering from Utah State University in 2018. He is currently pursuing the Ph.D. degree in computer science from the University of Utah. His research interests include high-performance machine learning and ML applications for localization, propagation, and spectrum allocation.



Prof. Gregory D. Durgin (Senior Member, IEEE) received the B.S.E.E., M.S.E.E., and Ph.D. degrees from Virginia Polytechnic Institute and State University in 1996, 1998, and 2000, respectively. He joined the faculty of the School of Electrical and Computer Engineering, Georgia Tech in Fall 2003, where he serves as a Professor. In 2001, he was awarded the Japanese Society for the Promotion of Science Postdoctoral Fellowship and spent one year as a Visiting Researcher with the Morinaga Laboratory, Osaka University. He has authored *Space-Time Wireless Channels* in 2002, the first textbook in the field of space-time channel modeling which has influenced multiple generations of commercial cellular technologies. He founded the Propagation Group (<http://www.propagation.gatech.edu>) with Georgia Tech, a research group that studies radiolocation, radio measurement, RFID-related technology, and applied electromagnetics. He has received best paper awards for articles coauthored in the IEEE TRANSACTIONS ON COMMUNICATIONS (1998 Stephen O. Rice Prize), *IEEE Microwave Magazine* in 2014, and IEEE RFID Conference in 2016, 2018, and 2019 as well as the 3rd Place 2020 Nokia Bell Labs Prize for "Hyper-RFID: A Revolution for the Future of RFID." He is a winner of the NSF CAREER Award as well as numerous teaching awards, including the Class of 1940 Howard Ector Outstanding Classroom Teacher Award at Georgia Tech in 2007. He has served on the editorial staff for IEEE RFID VIRTUAL JOURNAL, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE JOURNAL ON RFID. He also serves as the President Elect for the IEEE Council of RFID (CRFID). He served as an IEEE CRFID Distinguished Lecturer from 2015 to 2018 and IEEE CRFID VP of Conferences from 2020 to 2021, and as the general/executive chair of many IEEE conferences. His educational channel #profdurgin on YouTube instructs viewers on engineering electromagnetics and RFID-related topics, having drawn over 13 000 subscribers and over 1 million views. He is a frequent consultant to industry, advising numerous multinational corporations on wireless technology.