

ReVeal: A Physics-Informed Neural Network for High-Fidelity Radio Environment Mapping

Mukaram Shahid, Kunal Das, Hadia Ushaq, Hongwei Zhang, Jiming Song,
Daji Qiao, Sarath Babu, Yong Guan, Zhengyuan Zhu, Arsalan Ahmad

Iowa State University, Ames, IA, USA 50011

Email: {mukaram, kunaldas, hadia, hongwei, jisong, daji, sarath4, guan, zhuz, aahmad}@iastate.edu

Abstract—Accurately mapping the radio environment (e.g., identifying wireless signal strength at specific frequency bands and geographic locations) is crucial for efficient spectrum sharing, enabling Secondary Users (SUs) to access underutilized spectrum bands while protecting Primary Users (PUs). However, existing models either lack generalization due to shadowing, interference, and fading, or are computationally expensive, limiting real-world applicability. To address such shortcomings, we derive a second-order Partial Differential Equation (PDE) for the Received Signal Strength Indicator (RSSI) based on an established statistical model. We then propose ReVeal (Reconstructor and Visualizer of Spectrum Landscape), a novel Physics-Informed Neural Network (PINN) that integrates the PDE residual into a neural network loss function to accurately model the radio environment using sparse Radio Frequency (RF) sensor measurements. ReVeal is validated using real-world measurement data from rural and suburban areas of the ARA testbed and benchmarked against existing methods. ReVeal outperforms traditional approaches in radio environment prediction; for example, with a Root Mean Square Error (RMSE) of only 1.95 dB, ReVeal achieves an accuracy of the order of magnitude higher than existing methods, including 3GPP and ITU-R channel models, ray tracing, and neural networks. In addition, ReVeal achieves high accuracy with low computational complexity while requiring only sparse RF sampling—for instance, just 30 training sample points across a 514-square-kilometer area. The promising results demonstrate ReVeal’s potential to advance spectrum management by enabling precise interference management between PUs and SUs.

Index Terms—Radio environment mapping, TVWS, physics informed neural network, PINN, ARA, rural regions, channel modeling.

I. INTRODUCTION

Existing spectrum sharing frameworks, such as those implemented in the TV White Space (TVWS) database and Citizens Broadband Radio Service (CBRS) Spectrum Access System (SAS), rely heavily on traditional statistical models. However, such models struggle to accurately capture the real-world spectrum occupancy and do not generalize well enough to capture shadowing and fading caused by different kinds of terrain and environmental conditions, leading to conservative approaches that over-protect the primary users (PUs) and cause discrepancies in channel availability for spectrum re-use [1]–[3]. In the meantime, deterministic models such as ray tracing require precise characterization of the complete propagation environment such as vegetation, trees, buildings, and material properties. Any errors in accurately defining these site-specific characteristics can degrade the models’ accuracy. In addition,

such deterministic models are computationally expensive to be useful for at-scale, online spectrum management in dynamic radio environments. The existing stochastic and deterministic models also typically require the transmitter’s operational parameters, such as Effective Isotropic Radiated Power (EIRP), transmitter location, and antenna characteristics, which may not be available in real-world scenarios (e.g., where strong privacy or military secrecy are desired). The aforementioned drawbacks call for new models that are generically applicable to diverse environments and that are highly accurate in capturing the impact of transmitters and environmental factors (e.g., vegetation, trees, and buildings) on receiver signal strength while not requiring comprehensive, highly accurate information about the transmitters and environment.

To address the above challenge, data-driven modeling via Spectrum Cartography (SC) offers a promising solution avenue. In SC, ground-truth wireless signal measurements from sparsely distributed RF sensors are used to accurately generate the Radio Environment Map (REM) in the geographical area of interest [1], [4]–[7]. In particular, SC treats radio environment mapping as an ill-posed inverse problem where transmitter location and RF parameters are not available, and SC uses the spatial relationship between measurements to regenerate REMs [8]–[10]. The generated REMs have a wide range of applications in wireless communications, for instance, dynamically identifying white spaces for efficient spectrum sharing, optimizing power control for interference management, and facilitating seamless handover [6], [7].

Despite their promises, existing methods for generating REMs have significant limitations. For instance, techniques such as kriging and tensor decomposition assume a uniform spatial structure, failing to capture complex variations in signal strength often observed in real-world scenarios. In addition, these models typically require dense data, leading to high computational and sensing costs [8], [11]. Similarly, while deep learning approaches are powerful for matrix or tensor completion tasks, they often lack interpretability and require vast amounts of labeled data, which are impractical to collect in real-world settings.

To fill the gap in radio environment mapping, we propose ReVeal, a novel PINN architecture for blind spectrum cartography. ReVeal uses a PDE to characterize the spatial variations in wireless signal strength, and then it incorporates the PDE as a physical constraint to a Fully Connected Neural

Network (FCNN) with random dropouts. Such an innovative approach enables ReVeal to achieve high accuracy with minimal data, capturing real-world signal variations without requiring prior knowledge of transmitter parameters or detailed environmental information (e.g., terrain). The key contributions of this paper are as follows:

- **Introduction of PINN to Spectrum Cartography:** This paper pioneers the use of PINNs in blind spectrum cartography, introducing a novel approach that integrates physical laws with data-driven learning to improve the accuracy and efficiency of radio environment mapping without requiring knowledge of transmitter operational parameters or environmental conditions.
- **PDE Form of Path Loss Model:** This paper derives a novel PDE-based on an established statistical path loss model, allowing ReVeal to capture spatial variations in signal strength caused by shadowing, thereby enabling the model to accurately account for shadowing and other environmental factors without prior knowledge of the environment. The PDE-based loss function proves superior in capturing shadowing distribution, which simple empirical path loss models fail to achieve when used with PINNs.
- **Data- and Computation-Efficient Solution:** ReVeal requires significantly fewer sample points than traditional techniques and achieves high accuracy. Its data efficiency, fast convergence, and optimized architecture, makes ReVeal well-suited for large-scale, real-time spectrum management in dynamic radio environments.
- **Real-World Outdoor Evaluation:** ReVeal has been evaluated using the ARA testbed [12] across diverse rural and suburban terrains and channel conditions. In addition, we benchmark the performance of ReVeal against statistical, deterministic, geospatial, and neural network models using real-world data collected over an area of a $19.4 \text{ km} \times 26.5 \text{ km}$ area in ARA.

The rest of the paper is structured as follows: Section II discusses related work, Section III introduces ReVeal, Section IV outlines the experiment evaluation plan, Section V presents the experimental results, and Section VI concludes the paper.

II. RELATED WORK

In what follows, we first review the stochastic and deterministic models typically used in today's spectrum management practice. Furthermore, we discuss geospatial models, deep learning models, and physics-informed deep learning approaches.

a) Stochastic and Deterministic Models: Channel modeling has been a fundamental component of wireless communication systems design, providing critical insights into signal propagation for interference management and network optimization [13]. Stochastic modeling techniques rely on statistical distributions or empirical equations to characterize signal distribution based on the operational parameters of the transmitter (e.g., location, height, and antenna azimuth), and line-of-sight information between the transmitter and the

receiver. The use of probabilistic techniques and statistical distributions in these models helps reduce computational complexity. However, such a simplification often comes at the expense of accuracy [14]. Since these models rely on summary statistical distributions rather than specific real-world instances, they fail to capture site-specific environmental features (e.g., vegetation, trees, and buildings). As a result, these models cannot accurately characterize wireless channel behavior at a given location, where signal propagation and shadowing effects can vary significantly across space.

Deterministic models, such as ray tracing, are site-specific and enable precise modeling of the propagation environment by considering the geographical scene, material properties, and the scattering between the transmitter and receiver [15], [16]. By incorporating the principles of physics such as reflection, diffraction, and scattering, deterministic models can accurately calculate the path loss, delay, and angle of each reflected component reaching the receiver. However, precision comes at the expense of computational complexity. Moreover, in real-world scenarios where it is difficult to precisely characterize the environment, deterministic models may still result in significant, as we will demonstrate in Section V.

Both stochastic and deterministic models require prior operational information about the transmitter, such as height, azimuth, and EIRP, which may not be realistic in Radio Dynamic Zones (RDZs) where multiple users utilize spectrum as a shared common resource pool. Moreover, such models are unable to integrate real-time spectrum usage data from RF sensors deployed in RDZs [17].

b) Geospatial and Deep Learning Models: Geospatial interpolation techniques have been the center of attention among the wireless community for generating REMs. Approaches such as kriging [18] and inverse weighted distance have been of use in modeling the spectrum occupancy based on the sparsely collected data points. However, such techniques work on the assumption of spatial stationarity and struggle to capture the nonlinear relationships, often experienced in modeling wireless channels, due to the presence of shadowing and wireless interference [18]. Furthermore, these models generally lack the ability to accommodate new terrains and varying spatial resolutions.

Deep Learning (DL) algorithms, on the other hand, are able to learn complex nonlinear spatial relationships from sparse training data [19], [20]. Various DL models such as Convolutional Neural Network [21], U-Net [22], and Generative Adversarial Network [23] have been proposed to generate spatio-temporal spectrum maps. However, DL approaches require significant amount of training data, and collecting data from real-world deployments is a time-consuming task. Furthermore, achieving a fixed dense deployment of RF sensors is often impractical due to cost and data collection overhead.

c) Physics-Informed Deep Learning: Physics-Informed Deep Learning (PIDL) has emerged as a new compelling method to solve PDEs for both forward and inverse problems. Finite Element Methods (FEMs) have been the key in solving PDEs in different engineering problems. However, while solv-

ing PDEs, FEMs are not capable of integrating real-world data without complex computationally expensive data assimilation techniques [24]. Such a limitation prevents FEMs from fully utilizing measurement data, which can cause valuable system insights to be overlooked [25]. In contrast, neural networks are naturally suited for data assimilation, as they can be trained using data of varying fidelity and modality. PINNs have been developed to bridge the gap between data-driven and physics-based methods, especially in cases where partial knowledge of the physical laws and sparse measurement data are available. By embedding physical laws directly into the neural network through residual loss terms in the objective function, PINNs can enforce the governing PDEs as soft constraints, which enable PINNs to solve forward and inverse problems using sparse and noisy data [26].

PINNs so far have been widely used in applications ranging from acoustic engineering to the modeling of flow dynamics and modeling of electromagnetic fields, however, have not been fully explored in the areas of wireless channel modeling and spectrum sharing. In this work, we explore and develop an innovative architecture that can leverage a PINN for spectrum cartography in the TVWS band utilizing the real-world data from the ARA testbed.

III. REVEAL: PINN FOR RADIO ENVIRONMENT MAPPING

Here, we focus on developing the spatial REM of a specific geographical region of interest denoted as Domain $\mathbb{D}$. The Domain $\mathbb{D}$ is discretized into $I \times J$ equally sized cells, with each cell representing a spatial location within the region of interest. We assume the presence of a transmitter X which may be located within or outside $\mathbb{D}$, as illustrated in Fig. 1. The transmitter's location and other parameters (e.g., transmission power and antenna characteristics) are unknown. A set of RF sensors Ω_n ($n = 1, 2, \dots, N$) is sparsely deployed across $\mathbb{D}$ at random to observe the Received Signal Strength Indicator (RSSI) at selected locations. Each sensor provides measurements of RSSI or power levels $P^{\text{obs}}(\Omega_n, C)$ for each channel C in a given set $\mathbb{C}$. The set of channels $\mathbb{C}$ represents the spectrum of interest, partitioned into discrete frequency

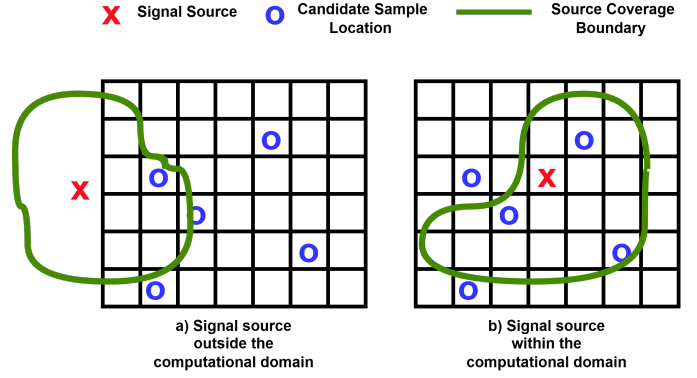


Fig. 1. Visualization of signal source placement relative to the domain/geographic-region of interest. The signal source (denoted as X) influences the candidate sample locations (denoted as O) differently depending on its position, altering the source coverage boundary (denoted in green color).

bands, with each band corresponding to a unique channel. The measurements collected by these sensors are sparse and are impacted by shadowing effects and large-scale path loss, which vary across both spatial and spectral dimensions. Our primary objective in SC is to model a function that can accurately predict the RSSI at any location within $\mathbb{D}$, enabling the generation of the REM for the domain $\mathbb{D}$. Mathematically, the objective is to minimize the error (L) between the expected observed RSSI (P^{obs}) and the expected model-predicted RSSI (P^{pred}). The notations used in this paper are summarized in TABLE I.

$$L = \frac{1}{N} \sum_{n=1}^N |P^{\text{pred}}(\Omega_n, C) - P^{\text{obs}}(\Omega_n, C)|. \quad (1)$$

A. Physics-Informed Neural Network in Reveal

Physics-informed neural networks are a recent development in scientific machine learning that leverage the ability of neural networks to learn underlying physics. The core idea behind PINNs is to incorporate the governing physical equation, typically a PDE, as a component of the neural network's loss function during training. The mean squared residual of the governing PDE, along with the data-driven loss function, is minimized to train the neural network effectively.

To solve the SC problem, we propose a PINN architecture as illustrated in Fig. 2. The network input consists of a set of spatial locations of the measurement points Ω_n representing the geographical locations of the RF sensors within the Domain $\mathbb{D}$. For a given channel C , the output of the PINN P^{pred} represents the expected RSSI values at the specified locations. During training, P^{pred} is compared with the observed values at each training sample to compute the data-driven loss. Apart from the data-driven loss, the residual of the governing-physics PDE is incorporated to further minimize the overall error of the PINN. The calculated loss is fed into the optimizer and is used to update the network's weights and biases accordingly. To prevent overfitting, multiple dropout layers are included in the neural network architecture.

TABLE I
NOTATION SUMMARY

Notation	Description
$\mathbb{D}$	Domain representing the specific geographical region of interest
$I \times J$	Grid dimensions of the domain $\mathbb{D}$
Ω_n	Set of RF sensors sparsely deployed across the region $\mathbb{D}$; $n = 1, 2, \dots, N$
$P^{\text{obs}}(\Omega_n, C)$	Observed power measurements (RSSI) at sensor locations Ω_n for channel C
$\mathbb{C}$	Set of channels or spectrum of interest, partitioned into discrete frequency bands
P^{pred}	Predicted power measurements (RSSI) by the model for sensor locations
L	Objective function representing the error between observed and predicted values

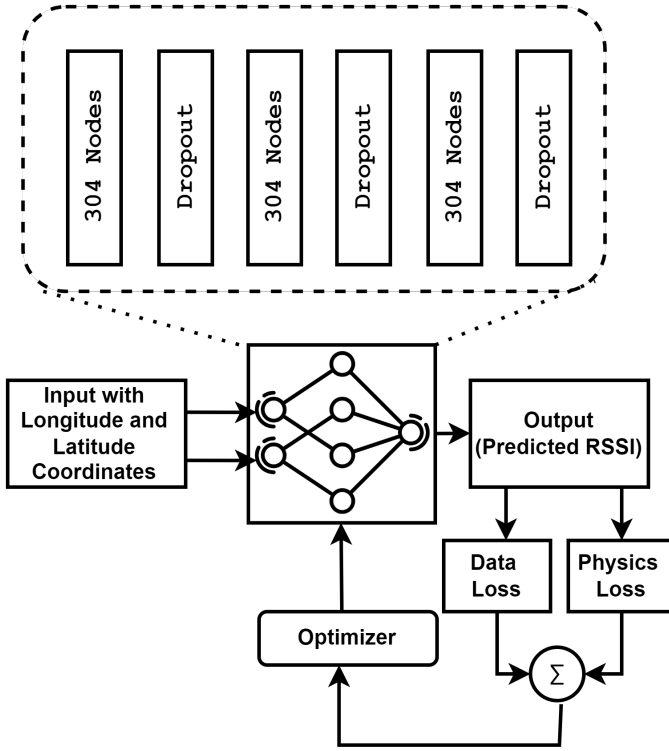


Fig. 2. PINN architecture of ReVeal

The overall loss function of the PINN ensures that the model not only minimizes the Mean Absolute Error (MAE) based on the measured data, but also adheres to the underlying physics, helping the model to generalize better during the training process. The overall loss function for the PINN optimization can be defined as

$$L_{\text{total}} = (1 - \lambda)L_d + \lambda L_p, \quad (2)$$

where

$$L_d = \frac{1}{N} \sum_{n=1}^N |P^{\text{pred}}(\Omega_n, C) - P^{\text{obs}}(\Omega_n, C)|, \quad (3)$$

and

$$L_p = \frac{1}{N} \sum_{n=1}^N |\nabla^2 P^{\text{pred}}(\Omega_n, C) - \nabla^2 P^{\text{obs}}(\Omega_n, C)|. \quad (4)$$

Here, L_d refers to the loss calculated from the labeled data points collected from the sensors. The physics-driven loss L_p is defined by Eqn. (4), where ∇^2 represents the Laplacian of both the predicted RSSI and observed RSSI at the sample location (a detailed analogy of introducing L_p is discussed in Eqn. (9)). The parameter λ in Eqn. (2) is a variable that controls the weightage given to the data-driven or physics-driven loss during the training of the neural network.

B. Governing-Physics PDE in ReVeal

A key design decision in ReVeal is selecting the physics model that governs the spatial dynamics of RSSI. To achieve

this, we derive a second-order PDE based on well-established wireless signal path loss models. Specifically, a well-known path loss model is given by

$$P_r(x, y) = P_T - \left[10\eta \log_{10} \left(\frac{\sqrt{(x - x_T)^2 + (y - y_T)^2}}{d_0} \right) \right] + Z_\sigma. \quad (5)$$

Eqn. (5) describes the received signal power $P_r(x, y)$ at any given instance in space (x, y) , as a function of the transmit power P_T , distance from the transmitter $\sqrt{(x - x_T)^2 + (y - y_T)^2}$, path loss exponent η , reference distance d_0 , and an added shadowing factor Z_σ denoting a specific realization of the underlying random spatial variations¹. Taking the expectation of both sides of Eqn. (5), we get

$$E[P_r(x, y)] = P_T - \left[10\eta \log_{10} \left(\frac{\sqrt{(x - x_T)^2 + (y - y_T)^2}}{d_0} \right) \right] + E[Z_\sigma]. \quad (6)$$

The second-order partial derivative of Eqn. (6) with respect to x results in

$$\frac{\partial^2 E[P_r(x, y)]}{\partial x^2} = \frac{10\eta}{\ln(10)} \cdot \frac{(y - y_T)^2 - (x - x_T)^2}{[(x - x_T)^2 + (y - y_T)^2]^2} + \frac{\partial^2 E[Z_\sigma]}{\partial x^2}. \quad (7)$$

Similarly, the second order partial derivative of Eqn. (6) with respect to y results in

$$\frac{\partial^2 E[P_r(x, y)]}{\partial y^2} = \frac{10\eta}{\ln(10)} \cdot \frac{(x - x_T)^2 - (y - y_T)^2}{[(x - x_T)^2 + (y - y_T)^2]^2} + \frac{\partial^2 E[Z_\sigma]}{\partial y^2}. \quad (8)$$

Adding Eqns. (7) and (8), we get

$$\frac{\partial^2 E[P_r(x, y)]}{\partial x^2} + \frac{\partial^2 E[P_r(x, y)]}{\partial y^2} = \frac{\partial^2 E[Z_\sigma]}{\partial x^2} + \frac{\partial^2 E[Z_\sigma]}{\partial y^2}. \quad (9)$$

The left hand side of Eqn. (9) represents the Laplacian operator as defined in Eqn. (4) while the right-hand side captures variations in signal strength due to shadowing caused by heterogeneous terrain and buildings within in the domain. The PDE constraint in Eqn. (9) explicitly encodes the spatial behavior of shadowing without requiring the knowledge of transmitter's location, ensuring that ReVeal does not merely fit the training data but also generalizes effectively in unseen regions.

In homogeneous environmental conditions or simple free-space settings, the signal variation due to shadowing is zero, thus the right-hand side of Eqn. (9) remains zero and does not vary across space. However, in real-world environments, where shadowing effects are present, the right-hand side of Eqn. (9) becomes non-zero. Thus, a good model must precisely capture the impact of shadowing and ensure that the second-order derivatives of the predicted signal strength in Domain $\mathbb{D}$ closely match the second-order derivatives of the observed signal strength in $\mathbb{D}$. To achieve this, we define the physics-driven loss term L_p as shown in Eqn. (4).

¹Temporal variation caused by fading is beyond the scope of this study and is left as a part of the future work.

C. Learning Criteria and Algorithm of ReVeal

During training, the network optimizes a composite loss function that integrates both data-driven and physics-driven losses. After the forward pass, the model predicts the expected RSSI at the training sample points. The model then computes the error between the predicted and observed values using the data-driven loss defined in Eqn. (3). In addition, leveraging the network's automatic differentiation capability, the model calculates the second-order spatial derivatives of the predicted RSSI values, which are used to evaluate the residuals of the governing-physics PDE defined in Eqn. (4). The physics-driven loss term minimizes these residuals, ensuring consistency with the governing-physics PDE relative to the observed RSSI. During back-propagation, the optimizer adjusts the network's weights and biases to minimize the composite loss, effectively balancing both data-driven and the physics-driven loss components.

The trade-off between the physics-driven loss and the data-driven loss is controlled by the parameter λ , which determines their relative importance. Without the data-driven loss term, the model lacks a starting point for optimization, as the physics-driven loss term, based on a PDE, does not provide sufficient guidance to align predictions with real-world observations. Therefore, an appropriate choice of λ is essential to achieve an optimal balance between empirical accuracy and physical consistency during training, as further elaborated in Section V.

D. Optimizing the PINN Architecture

Hyper-parameter tuning is a crucial yet tedious task in designing machine learning algorithms, including PINNs. The optimal selection of parameters—such as the number of hidden layers, number of neurons per layer, activation function, and learning rate—significantly impacts both the performance and convergence of the model. In literature, advanced hyper-parameter tuning algorithms utilizing techniques such as grid search and random search have been employed for parameter selection, including Autotune [27] and SMAC [28]. However, most of such libraries require a predefined search space from the user to identify the best parameters that minimize an objective function. On the other hand, libraries such as Optuna [29] provide users with the flexibility to define a dynamic search space and employ advanced optimization techniques such as Tree-Structured Parzen Estimator (TPE) for dynamic and efficient hyper-parameter tuning and optimal parameter selection. Therefore, in this study, we use Optuna [29] as the hyper-parameter optimization library to select parameters based on the spatially sampled data points. The resulting optimized hyper-parameters are presented in TABLE II. Bringing together all the aforementioned design choices, Algorithm 1 summarizes the ReVeal algorithm.

IV. EXPERIMENTAL SETUP

We evaluate ReVeal using real-world data collected from the ARA testbed [12]. ARA is a first-of-its-kind wireless living lab located around Iowa State University, spanning an area over

TABLE II
HYPER-PARAMETERS OF ReVEAL

Hyper-parameter	Value
Number of input features	2
Number of hidden layers	3
Number of neurons per layer	304
Activation function	ReLU
Dropout rate	0.2
Number of output features	1
Learning rate	0.00369

Algorithm 1 ReVeal for Generating REM

- 1: Define the domain $\mathbb{D}$ of interest as a grid of $I \times J$ cells.
- 2: Define the locations Ω_n of the RF sensors within $\mathbb{D}$.
- 3: Load the observed RSSI values P^{obs} from the sensors at locations Ω_n for each channel C .
- [Initialize the PINN model]**
- 4: $model = \text{PINN}(\text{hyper-parameters})$
- [Define the Loss function]**
- 5: $L_d = \text{MAE}(P^{\text{pred}}(\Omega_n, C), P^{\text{obs}}(\Omega_n, C))$
- 6: $L_p = \text{MAE}(\nabla^2 P^{\text{pred}}(\Omega_n, C), \nabla^2 P^{\text{obs}}(\Omega_n, C))$
- 7: $L_{\text{total}} = (1 - \lambda)L_d + \lambda L_p$
- [Train the PINN model]**
- 8: **for** $epoch = 1$ **to** num_epochs **do**
- 9: **for** $batch$ **in** $batches(\Omega_n, P^{\text{obs}})$ **do**
- 10: Predict the power measurements P^{pred} for the current $batch$: $P^{\text{pred}} = model(batch)$
- 11: Calculate the total loss: $loss = L_{\text{total}}$
- 12: Update the model parameters using the optimizer: $optimizer.step(loss)$
- 13: **end for**
- 14: **end for**
- [Obtain the predicted REM]**
- 15: $REM = model(\mathbb{D})$

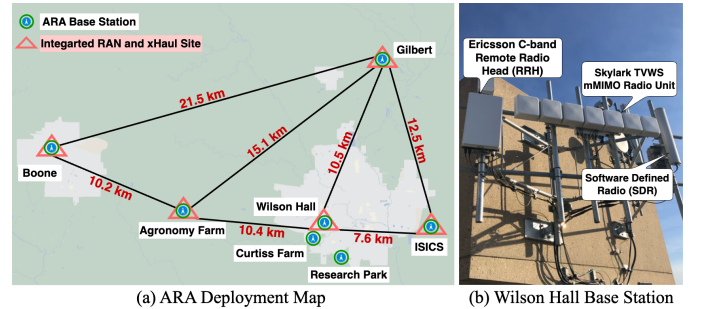


Fig. 3. ARA Deployment

500 square kilometers and covering research and producer farms along the rural communities of Central Iowa. Fig. 3 illustrates the deployment of seven ARA Base Station (BS) sites. Among these, Wilson Hall, Boone, ISICS, and Gilbert are equipped with the SkyLark massive MIMO (mMIMO) platform operating in the TVWS band. Meanwhile, Wilson Hall, Agronomy Farm, Research Park, and Curtis Farm sites include a C-Band Commercial-Off-the-Shelf (COTS) mMIMO

system from Ericsson operating at 3450–3550 MHz band. In addition, all BS sites feature Software Defined Radios (SDRs) operating at 3400–3600 MHz, supporting fully programmable, end-to-end, whole-stack 5G experiments using open-source systems such as OpenAirInterface and srsRAN.

a) Data Collection Site: For REM modeling, we collected real-world data using the Skylark mMIMO systems deployed at the Wilson Hall BS. The sample points were selected around the Wilson Hall covering an area spanning 514 square kilometers. At the time of sampling, all other BSes were turned off. Before initiating the actual sampling process, we ensured that no other transmitters were operating at the same frequency as the TVWS band used by the Skylark BS. The coverage range of the Skylark BS around Wilson Hall encompasses diverse terrains, ranging from rural communities to suburban areas of Downtown Ames, as illustrated in Fig. 4. These different terrains exhibit varying distributions of shadowing and fading effects. As shown in Fig. 5, the measurement sampling instances indicate different levels of shadowing, even when the distance from the BS remains the same.

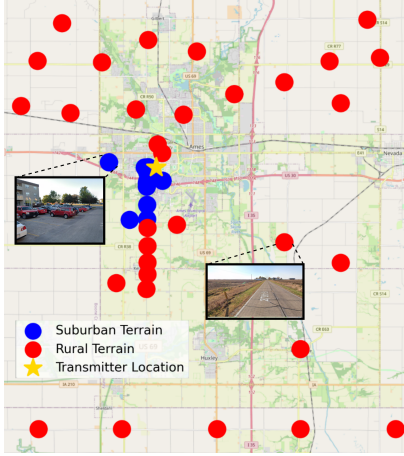


Fig. 4. Sampling Locations and Corresponding Terrain Conditions

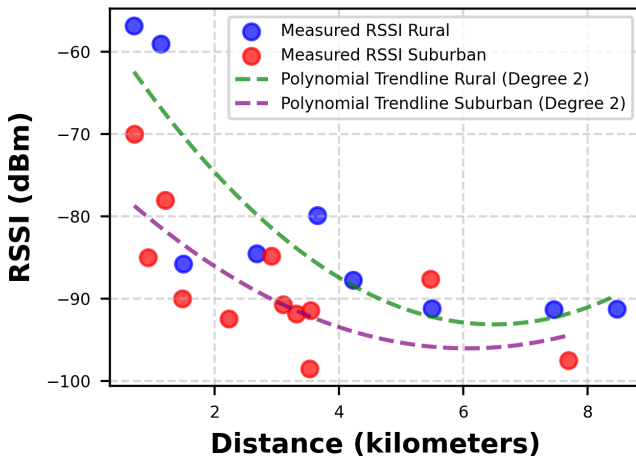


Fig. 5. Variation in Channel Conditions across Rural and Suburban Terrains

b) Spectrum Sensing Equipment: The ARA BS sites are equipped with spectrum sensing equipment, specifically the Keysight N6841-A RF sensors connected to Keysight N6850A omni-directional antennas, which monitor spectrum activities across the bands of interest. Apart from the fixed spectrum sensors, the Keysight N9952A FieldFox, equipped with an N6850A omni-directional antenna, and the Keysight NEMO Handy handheld measurement solution are used to capture RSSI values at various spatial points around the BS site.

c) Spatial Sampling Strategy: Selecting appropriate spatial samples is crucial to reducing sampling overhead while maintaining modeling accuracy. A spatially-balanced sampling strategy is essential in spatial modeling to minimize prediction errors across the entire domain. The primary objective of this approach is to identify the most informative sample points from all the candidate locations in the dataset [30]. Since spatial points in close proximity often exhibit similar or identical RSSI values, it is preferable to select spatially distant points that are representative of the overall population. To this end, we employ Local Pivotal Method (LPM) [31] to determine the spatial sampling locations in our study, as shown in Fig. 4.

V. EXPERIMENTAL RESULTS

The computational experiments for evaluating ReVeal were conducted on a workstation equipped with an Intel® Xeon® processor operating at 3.40 GHz with 32 GB DDR4 RAM. ReVeal is compared against stochastic models, including 3GPP TR 38.901 [32] and ITU-R IMT-2020 [33]. Deterministic models such as ray tracing using Sionna [34] is also implemented for the specific ARA regions where real-world measurement data is collected. Furthermore, ReVeal is benchmarked against classical machine learning models commonly used for generating REMs, including the kriging and standard FCNNs. To assess the importance of using the PDE-based path loss model in ReVeal, we also compare it to that incorporate alternate physics models, specifically 3GPP TR 38.901 and ITU-R IMT-2020. These variants are denoted as PINN with 3GPP TR 38.901 Model and PINN with ITU-R IMT-2020 Model, respectively. Unless stated otherwise, all PINN-based models (including ReVeal) use $\lambda = 0.999$.

TABLE III presents the performance of various benchmark algorithms used for channel modeling. The table includes several key metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared, and computation time for each model. ReVeal achieves an RMSE of the order of magnitude smaller than other methods while maintaining a relatively low computation time. Although the statistical models of 3GPP TR 38.901 and ITU-R IMT-2020 have low computation times of 3.5 s and 3.4 s, respectively, they exhibit high modeling errors, with RMSE values of 17.25 dB and 11.13 dB, respectively. The low computation complexity but high prediction errors in these statistical models result from their reliance on predefined formulas that capture statistical signal propagation behavior without incorporating site-specific environmental features (e.g., vegetation, trees, and buildings)

TABLE III
PERFORMANCE OF DIFFERENT PATH LOSS MODELS ACROSS THE WHOLE DOMAIN

Model	RMSE (dB)	MAE (dB)	R-Squared	Computation Time (seconds)
3GPP TR 38.901 Model	17.25	15.93	-0.81	3.5
ITU-R IMT-2020 Model	11.13	10.51	0.25	3.4
Ray-Tracing with Sionna	26.96	25.74	-3.42	> 690
Kriging	12.02	13.65	0.15	39
FCNN	10.59	10.23	0.24	39.4
PINN with 3GPP TR 38.901 Model	23.80	17.69	-3.02	9.5 (With early stopping)
PINN with ITU-R IMT-2020 Model	12.35	13.35	0.21	9.9 (With early stopping)
ReVeal	1.95	2.15	0.95	8.9 (With early stopping)

present in the ARA testbed during evaluation. The computation time for data-driven models such as kriging and FCNN is higher than that of statistical models, however, lower than the deterministic ray tracing model. The ray tracing model is the most computationally intensive, as it requires detailed environment modeling and the ability to trace individual propagation path between the transmitter and the receiver.

For locations at varying distances from the Skylark BS, Fig. 6 compares the actual RSSI values and those predicted by

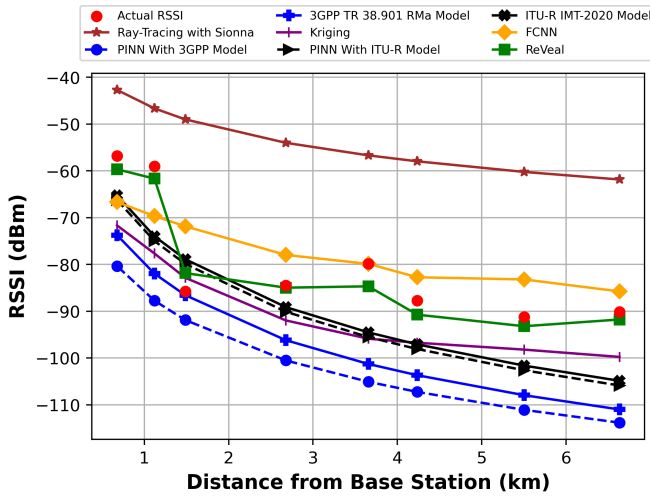


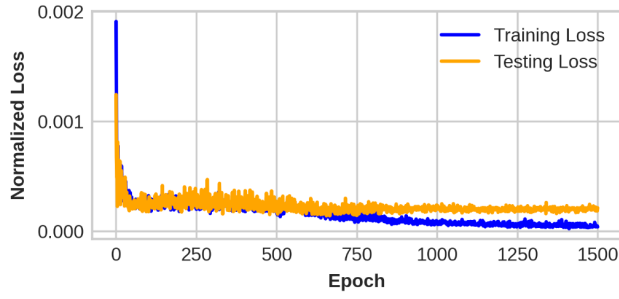
Fig. 6. Comparison of Predicted and Actual RSSI Values at Different Distances from the BS

different methods. ReVeal demonstrates strong performance across nearly all distances, effectively capturing RSSI variations caused by shadowing. In particular, ReVeal closely follows the changes in RSSI influenced by terrain conditions, showcasing its ability to model real-world signal behavior more accurately than existing methods. Traditional statistical models tend to underestimate signal strength by tens of dBs. The kriging model exhibits higher errors in certain scenarios due to its method of estimating values at unmeasured locations by calculating a weighted average of surrounding spatial points. Such an interpolation approach assumes that the underlying spatial field follows a specific correlation structure, and any deviation—such as in complex environments with

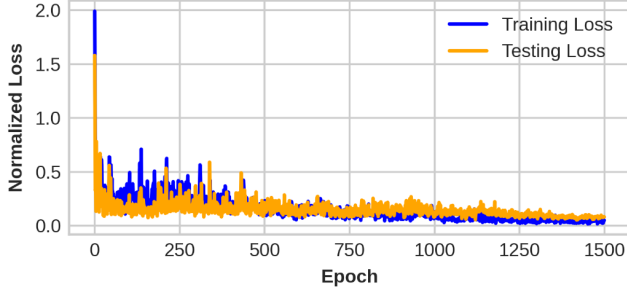
significant signal strength variations—can lead to larger prediction errors. Similarly, ray tracing relies heavily on precise characterization of propagation conditions, including factors such as vegetation and building materials. However, obtaining such detailed environmental data for large, complex outdoor settings is practically infeasible. As a result, ray tracing performs poorly due to the lack of accurate environmental data. In addition, PINNs with the 3GPP TR 38.901 Model and ITU-R IMT-2020 Model attempt to align with the behavior of the underlying statistical models, and the inherent limitations of the underlying models restrict the accuracy of the PINNs in such cases.

To further highlight the importance of incorporating physical-domain knowledge in radio environment mapping, Fig. 7 compares training and testing losses of ReVeal and FCNN. In Fig. 7(a), ReVeal exhibits rapid convergence, with its training loss decreasing quickly and stabilizing around 300 epochs. Both training and testing losses in ReVeal start at 0.0020, whereas in FCNN, the loss exceeds 2. This demonstrates that the incorporation of physics-driven loss helps the neural network learn more effectively, ultimately improving ReVeal's overall performance. Fig. 7(b) further shows that the FCNN model experiences greater fluctuations and higher errors, with both the training and testing losses decreasing more slowly and erratically, suggesting that FCNN struggles to converge and generalize as effectively as ReVeal.

To evaluate ReVeal's capability of generating accurate REMs with sparse training samples, Fig. 8 presents the performance of different data-driven models with varying sample sizes, where PINN models (including ReVeal) use $\lambda = 0.999$. We observe that the RMSE decreases as the number of training samples increases. However, PINN models that utilize statistical models as their underlying physics framework hardly show any improvement, as they tend to replicate the behavior of the statistical model itself. With 16 spatial samples, the average RMSE for ReVeal was slightly over 5 dB, whereas it decreased to 1.95 dB with 30 training samples. As expected, computation complexity increases with the number of training samples. By employing early stopping, where training halts once the desired accuracy is reached, ReVeal could be trained and used to visualize REM in approximately 8.9 seconds. In contrast, other methods require significantly more computation



a) Training vs Testing Loss of ReVeal



b) Training vs Testing Loss of FCNN

Fig. 7. Comparison of Training and Test Loss for ReVeal and FCNN Model

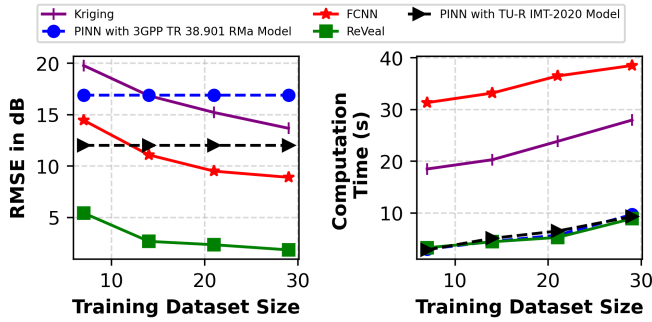


Fig. 8. Performance Analysis with other Spectrum Cartography Techniques

time, for instance, with kriging and FCNN taking around 28 seconds and 39 seconds, respectively.

To gain insight into Reveal's modeling accuracy, Fig. 9(a) presents the reconstructed RSSI map for $\lambda = 0.999$. The spatial variation in signal strength is evident, with a distinct high-signal region surrounded by weaker areas. Such a spatial variability demonstrates ReVeal's ability to effectively capture signal propagation patterns. Fig. 9(b) shows the absolute error for both training points (circles) and testing points (squares), where the color intensity represents the magnitude of error. Overall, the error remains relatively low across most locations, indicating strong generalization performance. To further analyze the probability distribution of the absolute modeling error in ReVeal, Fig. 10 presents its Empirical

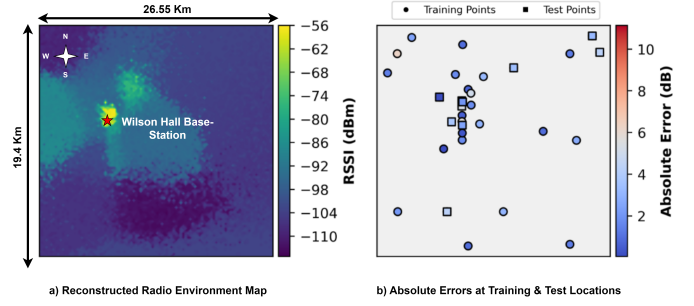


Fig. 9. Comparison of Test and Training Error over Regenerated Field

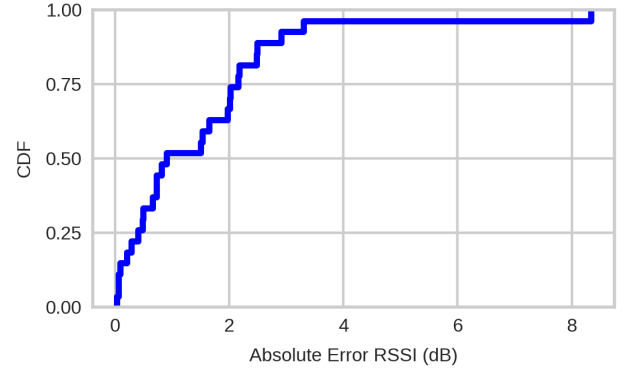


Fig. 10. ECDF of Absolute Errors

Cumulative Distribution Function (ECDF). The 25th percentile, median (50th percentile), and 75th percentile absolute errors are 1.02 dB, 1.31 dB, and 2.39 dB, respectively.

Fig. 11 illustrates the impact of varying λ in Eqn. (2) on the overall RMSE. The parameter λ controls the trade-off between the data-driven loss and the physics-driven loss in the overall loss function. As λ increases, greater emphasis is placed on the physics-driven loss, leading to a reduction in overall error. However, when $\lambda = 1$, the data-driven loss (L_d) is completely disregarded, causing the RMSE to increase

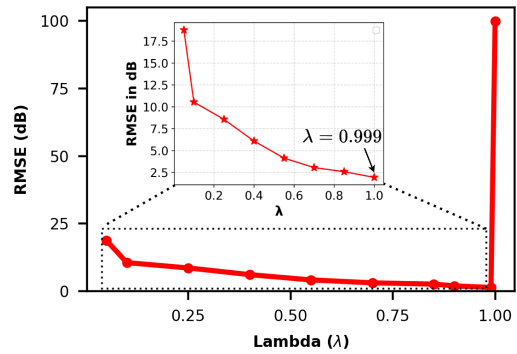


Fig. 11. Impact of Varying λ on ReVeal Performance

significantly, highlighting the critical role of L_d in guiding the model toward a high-performance solution. Specifically, the data-driven loss ensures that the model adheres to the real-world measurements while maintaining consistency with the governing-physics PDE. The aforementioned behavior underscores the importance of fine-tuning λ to effectively balance the contributions of data-driven and physics-driven loss terms in PINNs for accurate prediction.

VI. CONCLUDING REMARKS

Radio environment mapping from a sparse set of spatial measurements plays crucial role in dynamic spectrum sharing. In this paper, we derived a PDE-based formulation of wireless path loss model and used it to develop ReVeal, a physics-informed neural network for accurate radio environment mapping in real-world settings. By integrating domain-specific physical principles into the neural network architecture, ReVeal effectively addressed the challenge of sparse measurements and achieved high accuracy and efficiency. ReVeal is evaluated using the first-of-its-kind ARA wireless testbed, demonstrating significant improvements over existing methods and showcasing its potential for real-world applications. ReVeal also opens avenues for future research, such as its application in designing dynamic spectrum sharing mechanisms and incorporating temporal dynamics for predicting spectrum usage trends.

ACKNOWLEDGMENT

This work is supported in part by the NSF awards 2130889, 2112606, 2212573, 2229654, and 2232461, NIFA award 202167021-33775, and PAWR Industry Consortium. The authors would also like to thank the Microsoft Research team, particularly Ranveer Chandra, Tusher Chakraborty, and Suraj Jog, for intriguing discussions on related topics.

REFERENCES

- [1] M. Shahid, K. Das, T. U. Islam, C. Somiah, D. Qiao, A. Ahmad, J. Song, Z. Zhu, S. Babu, Y. Guan, T. Chakraborty, S. Jog, R. Chandra, and H. Zhang, "Wireless spectrum in rural farmlands: Status, challenges and opportunities," Jul. 2024. [Online]. Available: <https://arxiv.org/abs/2407.04561>
- [2] R. Ramjee, S. Roy, and K. Chintalapudi, "A critique of FCC's TV white space regulations," vol. 20, no. 1, Aug. 2016. [Online]. Available: <https://doi.org/10.1145/2972413.2972421>
- [3] M. Ante, J. Molina, E. Trinidad, and L. Materum, "A survey and comparison of TV white space implementations in Japan, the Philippines, Singapore, the united kingdom, and the United States," *International Journal of Advanced Technology and Engineering Exploration*, vol. 8, pp. 780–796, Jul. 2021.
- [4] S. Bhattarai, "Spectrum efficiency and security in dynamic spectrum sharing," Ph.D. dissertation, Virginia Tech, Apr. 2018, doctoral Dissertation. [Online]. Available: <http://hdl.handle.net/10919/82872>
- [5] Federal Communications Commission, "Facilitating shared use in the 3.1–3.55 GHz band," Jan. 2020, federal Register, Vol. 85, No. 14, Pages 3888–3903. [Online]. Available: <https://www.federalregister.gov/documents/2020/01/22/2020-00535/facilitating-shared-use-in-the-31-355-ghz-band>
- [6] Y. S. Reddy, A. Kumar, O. J. Pandey, and L. R. Cenkeramaddi, "Spectrum cartography techniques, challenges, opportunities, and applications: A survey," *Pervasive Mob. Comput.*, vol. 79, no. C, Jan. 2022. [Online]. Available: <https://doi.org/10.1016/j.pmcj.2021.101511>
- [7] Y. Teganya, D. Romero, L. M. L. Ramos, and B. Beferull-Lozano, "Location-free spectrum cartography," *Trans. Sig. Proc.*, vol. 67, no. 15, p. 4013–4026, Aug. 2019. [Online]. Available: <https://doi.org/10.1109/TSP.2019.2923151>
- [8] S. Timilsina, S. Shrestha, and X. Fu, "Quantized radio map estimation using tensor and deep generative models," *Trans. Sig. Proc.*, vol. 72, pp. 173–189, Nov. 2023. [Online]. Available: <https://doi.org/10.1109/TSP.2023.3336179>
- [9] D. Romero and S.-J. Kim, "Radio map estimation: A data-driven approach to spectrum cartography," *IEEE Signal Processing Magazine*, vol. 39, no. 6, pp. 53–72, Nov. 2022.
- [10] C. R. Dietlein, "Wide-area spectrum cartography," National Telecommunications and Information Administration, Institute for Telecommunication Sciences, Boulder, CO, USA, Tech. Rep., accessed: 2024-11-21. [Online]. Available: <https://its.ntia.gov/>
- [11] H. Sun and J. Chen, "Integrated interpolation and block-term tensor decomposition for spectrum map construction," *IEEE Transactions on Signal Processing*, vol. 72, pp. 3896–3911, Aug. 2024.
- [12] T. U. Islam, J. O. Boateng, M. Nadim, G. Zu, M. Shahid, X. Li, T. Zhang, S. Reddy, W. Xu, A. Atalar, V. Lee, Y.-F. Chen, E. Gosling, E. Permatasari, C. Somiah, Z. Meng, S. Babu, M. Soliman, A. Hussain, D. Qiao, M. Zheng, O. Boyraz, Y. Guan, A. Arora, M. Selim, A. Ahmad, M. B. Cohen, M. Luby, R. Chandra, J. Gross, and H. Zhang, "Design and implementation of ARA wireless living lab for rural broadband and applications," Aug. 2024. [Online]. Available: <https://arxiv.org/abs/2408.00913>
- [13] J. Zhang, J. Lin, P. Tang, Y. Zhang, H. Xu, T. Gao, H. Miao, Z. Chai, Z. Zhou, Y. Li, H. Gong, Y. Liu, Z. Yuan, X. Liu, L. Tian, S. Yang, L. Xia, G. Liu, and P. Zhang, "Channel measurement, modeling, and simulation for 6G: A survey and tutorial," *arXiv preprint arXiv:2305.16616*, Oct. 2023. [Online]. Available: <https://arxiv.org/abs/2305.16616>
- [14] H. Tataria, K. Haneda, A. F. Molisch, M. Shafi, and F. Tufvesson, "Standardization of propagation models for terrestrial cellular systems: A historical perspective," *International Journal of Wireless Information Networks*, vol. 27, pp. 340–364, Mar. 2020. [Online]. Available: <https://link.springer.com/article/10.1007/s10776-020-00500-9>
- [15] J.-H. Lee and A. F. Molisch, "A scalable and generalizable pathloss map prediction," *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 17793–17806, Nov. 2024.
- [16] T. Orekondy, P. Kumar, S. Kadambi, H. Ye, J. Soriaga, and A. Behboodi, "WiNeRT: Towards neural ray tracing for wireless channel modelling and differentiable simulations," in *Proceedings of the 11th International Conference on Learning Representations (ICLR 2023)*, Jul. 2023. [Online]. Available: <https://openreview.net/forum?id=tPKKXeW33YU>
- [17] M. Zheleva, C. R. Anderson, M. Aksoy, J. T. Johnson, H. Affinnih, and C. G. DePree, "Radio dynamic zones: Motivations, challenges, and opportunities to catalyze spectrum coexistence," *IEEE Communications Magazine*, vol. 61, no. 6, pp. 156–162, Jan. 2023.
- [18] P. Maiti and D. Mitra, "Complexity reduction of ordinary Kriging algorithm for 3D REM design," *Physical Communication*, vol. 55, p. 101912, Oct. 2022.
- [19] A. Chaves-Villota and C. A. Viteri-Mera, "DeepREM: Deep-learning-based radio environment map estimation from sparse measurements," *IEEE Access*, vol. 11, pp. 48 697–48 714, May. 2023.
- [20] S. Sarkar, D. Guo, and D. Cabric, "ProSpire: Proactive spatial prediction of radio environment using deep learning," in *2023 20th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Dec. 2023, pp. 177–185.
- [21] H. Cheng, H. Lee, and S. Ma, "CNN-based indoor path loss modeling with reconstruction of input images," in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, Nov. 2018, pp. 605–610.
- [22] M. Mallik, S. Kharbech, T. Mazloun, S. Wang, J. Wiart, D. P. Gaillot, and L. Clavier, "EME-Net: A U-net-based indoor EMF exposure map reconstruction method," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, Jun. 2022, pp. 1–5.
- [23] X. Han, L. Xue, F. Shao, and Y. Xu, "A power spectrum maps estimation algorithm based on generative adversarial networks for underlay cognitive radio networks," *Sensors*, vol. 20, no. 1, Aug. 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/1/311>
- [24] J. D. Toscano, V. Oommen, A. J. Varghese, Z. Zou, N. A. Daryakenari, C. Wu, and G. E. Karniadakis, "From PINNs to PIKANs: Recent advances in physics-informed machine learning,"

arXiv preprint arXiv:2410.13228, Oct. 2024. [Online]. Available: <https://arxiv.org/abs/2410.13228>

- [25] D. E. Johnson, "Numerical solutions to Poisson equations using the finite-difference method," *IEEE Antennas and Propagation Magazine*, vol. 56, no. 6, pp. 158–162, Jan. 2014. [Online]. Available: <https://ieeexplore.ieee.org/document/6931698>
- [26] M. Raissi, P. Perdikaris, and G. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0021999118307125>
- [27] P. Koch, O. Golovidov, S. Gardner, B. Wujek, J. Griffin, and Y. Xu, "Autotune: A derivative-free optimization framework for hyperparameter tuning," Apr. 2018.
- [28] F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Sequential model-based optimization for general algorithm configuration," in *Proceedings of the 5th International Conference on Learning and Intelligent Optimization*, ser. LION'05. Berlin, Heidelberg: Springer-Verlag, Jan. 2011, p. 507–523. [Online]. Available: https://doi.org/10.1007/978-3-642-25566-3_40
- [29] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Mar. 2019, pp. 2623–2631.
- [30] A. Ivanov, K. Tonchev, V. Poulkov, A. Manolova, and A. Vlahov, "Limited sampling spatial interpolation evaluation for 3D radio environment mapping," *Sensors*, vol. 23, no. 22, Sep. 2023. [Online]. Available: <https://www.mdpi.com/1424-8220/23/22/9110>
- [31] A. Grafstr  , N. L. Lundstr  , L. Schelin *et al.*, "Spatially balanced sampling through the pivotal method," *Biometrics*, vol. 68, Jun 2012.
- [32] 3GPP, "3GPP TS 38.901—Study on channel model for frequencies from 0.5 to 100 GHz," Mar. 2021, accessed: 2024-12-13. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3173>
- [33] ETSI, "ETSI TR 138 901 V14.0.0— technical report: Study on channel model for frequencies from 0.5 to 100 GHz," Apr. 2020, accessed: 2024-12-13. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/14.00.00_60/tr_138901v140000p.pdf
- [34] J. Hoydis, S. Cammerer, F. Ait Aoudia, A. Vem, N. Binder, G. Marcus, and A. Keller, "Sionna: An open-source library for next-generation physical layer research," *arXiv preprint*, Mar. 2023.