OPTIMIZING RF SIGNAL PROPAGATION MODELS USING DEEP LEARNING REGRESSOR FOR 5G NETWORKS

Gutti Pallavi[1], Yanaki Manisha[2], Yennam Srikar[3], Mrs.T.Anitha[4]
[1,2,3] UG Scholar, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Assistant Professor, Dept. of ECE, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
gutthipallavi004@gmail.com

Abstract:

The optimization of Radio Frequency (RF) signal propagation models is a critical aspect of designing and managing modern wireless communication systems, particularly for 5G networks. Traditional models, based on empirical formulas and path loss models, have served the industry for decades, but they often fail to capture the complexity of urban environments, mobility, and dynamic network conditions. In light of the rapid advancements in wireless communication, deep learning techniques offer a promising solution to enhance the accuracy and efficiency of RF signal propagation models. The Project explores the application of deep learning regressors for optimizing RF signal propagation models in 5G networks. By leveraging large-scale data from real-world network deployments and simulations, deep learning models are trained to predict signal strength, coverage, and other key performance indicators with higher accuracy and computational efficiency than traditional methods. This research highlights the limitations of conventional RF models, the need for more adaptive and accurate prediction tools, and the significance of deep learning in improving the performance of 5G networks. The proposed deep learning-based approach addresses the challenges of signal prediction in complex environments, such as urban areas with high mobility, diverse terrain, and varying environmental conditions. The results show that deep learning regressors can significantly enhance RF signal propagation modelling, providing a more robust foundation for 5G network planning, optimization, and deployment. In the early days of wireless communications, signal propagation models were based on physical principles like free- space path loss and diffraction. These models were later enhanced with empirical data to account for real-world conditions. However, the growth of mobile networks, particularly the transition from 2G to 3G, and now to 5G, has outpaced the capabilities of traditional models. With the complexity of high-frequency signals, dense urban environments, and dynamic user behaviour, traditional propagation models have become less reliable. The need for efficient RF signal propagation models has always been central to wireless communication systems. Traditional models like the Hata model, COST-231, and Okumura-Hata were developed based on empirical measurements and assumptions that work well under specific conditions but fail to adapt to the rapidly changing and complex scenarios in modern networks.

Keywords: RF propagation,5G, Deep learning, Regression, Wireless Signal prediction, Optimization, Path loss, Urban networks, AI modelling, Coverage

1. INTRODUCTION

Benchmarking different optimization algorithms is task, particularly for network-based cellular communication systems. The design and management process of these systems involves many stochastic variables and complex design parameters that demand an unbiased estimation and analysis. Though several optimization algorithms exist for different parametric modelling and tuning, an in-depth evaluation of their functional performance has not been adequately addressed, especially for cellular communication systems. The experimental data were taken from different radio signal propagation terrains around four encode cells. In order to assist the radio frequency engineer in selecting the most suitable optimization method for the parametric model tuning, three-fold benchmarking criteria comprising the Accuracy Profile Benchmark, Function Evaluation Benchmark, and Execution Speed Benchmark were employed. Cognitive radio is the enabling technology for supporting dynamic spectrum access: it addresses the spectrum scarcity problem that is faced in many countries. A cognitive radio is an intelligent radio that can be reprogrammed and reconfigured dynamically. A cognitive radio is designed to use the best available wireless channels in its surroundings. Its transceiver can automatically detect available channels in wireless spectrum and can change its transmission and reception parameters accordingly to allow more concurrent wireless communication in a given wireless band for a particular instant of time for a particular place. Such spectrum allocation is known as dynamic spectrum management (DSA).

The widespread deployment of 5G networks has raised significant challenges in optimizing the efficiency and performance of Radio Frequency (RF) signal propagation. For 5G networks, which rely on high-frequency bands and massive antenna systems, predicting RF signal strength and coverage is crucial for ensuring that end-users experience reliable and high-speed connections. Traditional RF signal prediction models often struggle to account for the increasing complexity of network environments, such as urban interference, multipath propagation, and dynamic channel conditions.

The evolution of wireless communication networks, particularly the rollout of 5G, has brought forward the need for highly accurate and scalable methods to predict RF signal strength. The motivation behind this research is rooted in the complexity and dynamic nature of modern communication systems, where traditional RF modelling techniques often fall short in dealing with the intricacies of 5G networks.

The use of machine learning models specifically deep learning models like CNN has shown great promise in various fields, from image recognition to natural language processing. This research seeks to harness the power of deep learning to analyse and predict RF signal strength more effectively than traditional methods. The motivation extends to exploring how these models can generalize to unseen data and adapt to real-time changes in network conditions, which is crucial for the dynamic and evolving nature of 5G network.

2. LITERATURE SURVEY

In recent years, the system design, deployment, and management of wireless radio frequency (RF) networks have become more tasking and complicated [1]. The intricacies and complications may be attributed to many dynamic factors. The advancement and constant evolution of different cellular network technologies, accompanied by different deployment procedures and management costs, can be a prominent factor [2]. In addition, frequent changes in localized environmental features such as houses, buildings, and trees, plus the varying weather condition around these networks, can be another significant factor [3].

Constant increasing traffic of mobile subscribers in the networks with different multimedia service quality demands could also be a key factor [4]. Remarkably, cell site acquisition is becoming more problematic due to the limited availability of suitable sites in a built-up area and neighbouring residents that generally work against such site installations, probably because of frequently rumoured electromagnetic radiations that emanate from them [5].

In order to cope with or overcome the aforementioned key challenges, the RF engineers must also be ready to explore techniques and efforts at the network design/deployment phase or optimization/management phase when in operation [6]. The propagation loss model is a key tool that the RF employs to estimate the cell radius and signal attenuation losses during and after cellular system network design/deployment [7]. These signal propagation models usually contain some unknown parameters that must be accurately determined in correspondence with experimental data from the terrain of interest. Inaccuracies in RF propagation modelling and their parameter estimation can compromise effective network planning, management, optimization, and operational activities [8]. The impact can be enormous regarding poor service quality, resource input wastage, and time costs. This key problem is often called the propagation model parameters identification problem in the telecommunication network engineering domain.

Optimization algorithms play a crucial role in enhancing the accuracy and efficiency of predictive analytics by finding the optimal values of model parameters. Accurate modelling and estimating parametric values for cellular network-based propagation models is a dynamic optimization problem due to the different nonlinearities involved [9]. The interaction of the transmitted waves with different propagation mediums and terrain features around the receiver causes its strength to attenuate and degrade, thus resulting in what is known as signal propagation loss. During the cellular radio frequency (RF) network design or optimization phase of an existing one, the RF engineer uses signal propagation models to estimate the characteristics of signal attenuation losses that occur between the transmitting stations and the receiver stations [10]. These signal propagation models usually contain some unknown parameters that must be accurately determined in correspondence with experimental data from the terrain of interest. Different benchmarking and comparative works exist on numerical and global optimization performance impacts for real-time applications but not within the domain of intricate RF propagation modelling and parameter tuning problems. In [11], deterministic local and stochastic global optimization methods were investigated and compared to identify and estimate unknown kinetic model parameters systematically.

This paper identifies and provides an overview of the common existing numerical and global optimization methods. The second focus is to benchmark the precision performance of the identified numerical and global optimization methods with practical case studies from different radio signal propagation terrains. In [12], both stochastic and deterministic global optimization algorithms were studied for nonlinear biological modelling and parameter estimation. The stochastic methods provided lower processing time from their results but with poor convergence to a global minimum under a limited iteration number. On the contrary, the deterministic methods yielded preferred solutions regarding convergence quality but huge computational weights.

Several global optimization algorithms are benchmarked with standard functions for practical applications, presented in [13]. The authors discovered that the Hybrid Differential Evolution and Adaptation Evolution Strategy Algorithm was better for complex objective functions than the Hook–Jeeves and particle swam optimization, which attained better global minimum convergence for less complex objective functions.

In [14], five different global optimization algorithms were investigated for benchmarking to reconstruct and optimize nano-optical shape parameters. From the investigation, the Bayesian optimization method was reported to outperform other algorithms, such as differential evolution and particle swam, in terms of run times. A similar approach involving different optimizers is presented in [15] for the panel data model. It found that the computational success rate of the optimizers varies proportionally with the nature of the problem being handled by them. In [16], the cumulative density function is explored as an indicator to benchmark the performance of stochastic global optimization algorithms on test data sets' analysis. The result reveals that the algorithms with the pure random search performed preferably better. In [17], numerical-based optimization techniques focusing on Levenberg–Marguardt (LM) and Gauss–Newton (GN) algorithms were investigated to compare their performance on the propagation model parameter optimization and prediction analysis. With the application focus on loss data taken from built-up areas, the results showed that the LM outperformed the GN in terms of precision accuracies. In [18], Particle Swarm Optimization (PSO) and random forest (RF) were applied comparatively to tune and identify the parameters of the signals' attenuation models. The authors found that the PS method attained the most preferred precision performance by 22–25% across the study locations, using maximum absolute error as the indicator.

In [19], neural networks, support vector machine, and random forest were benchmarked with traditional path loss models like the COST 231-Walfisch Ikegami model and COST 231-Hata model. The authors disclosed that random forest yielded the best precision performance in path loss prediction.

Through the propagation modelling and benchmarking process, it was found in [20] that the proposed Light GBM model, which is a machine learning-based developed modelling algorithm, outperforms the empirical models by 65% in terms of prediction accuracy and decreased by 13 x in prediction time when matched with ray-tracing. The notable performance was achieved even with thin training data sets. Also, via detailed benchmarking processes in [20], the authors developed hybrid particle swarm–random forest and vector statistics–neural network models for propagation loss modelling and observed that their proposed models attained preferred prediction accuracies compared to traditional approaches.

In [21], the predictive modelling performance of four popular machine learning methods consisting of support vector regression, neural networks, gradient tree boosting, and random forest was compared with empirical path loss models after incorporating crossed walls' number into them. From among the four learning-based methods engaged, the gradient tree boosting displayed the best generalization and prediction capacities.

3. PROPOSED METHODOLOGY

The project is focused on optimizing the prediction of Radio Frequency (RF) signal propagation models, specifically for 5G networks. The goal is to predict signal strength based on various environmental and technical features that influence the propagation of signals in a wireless network. The project utilizes machine learning and deep learning techniques to achieve this optimization, with an emphasis on applying regression models for more accurate predictions.

Steps Involved:

The primary objective of this project is to predict **Signal Strength**, which serves as the dependent (target) variable in the dataset. Accurate signal strength prediction is crucial for **5G network planning**, ensuring **optimal coverage, efficient resource utilization, and minimal interference**. To achieve this, the project focuses on **optimizing existing signal propagation models** using advanced **machine learning and deep learning techniques**. By carefully selecting **relevant features** and applying **optimized regression techniques**, the goal is to enhance prediction accuracy and minimize errors.

A **hybrid approach** is employed, integrating both **traditional machine learning models** such as **KNN and AdaBoost**, alongside **deep learning models like CNN**. This approach allows for a **comparative analysis** of different techniques, facilitating the selection of the most effective method for **signal strength prediction**. By leveraging the strengths of both **machine learning and deep learning**, the project aims to develop a robust predictive model that enhances the efficiency of network planning and optimization in **5G environments**.

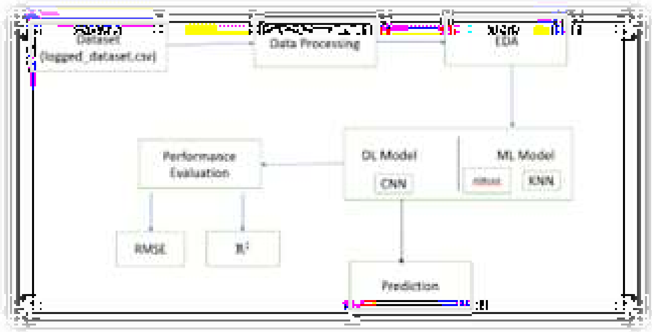

FIGURE 1. Proposed system Block Diagram

The workflow begins with Data Collection and Preprocessing, where the dataset (*logged_data.csv*) is imported, missing values are handled, duplicate rows are removed, and categorical variables are encoded using Label Encoding. Additionally, numerical features are normalized using Standard Scaler to improve model performance. Next, in the Exploratory Data Analysis (EDA) phase, correlation analysis is performed using heatmaps to understand feature relationships with Signal Strength, followed by feature selection to retain only the most relevant features. The dataset is then split into training (80%) and testing (20%) sets using train_test_split from scikit-learn.

For Model Building and Training, both traditional machine learning models (KNN Regressor, AdaBoost Regressor) and a deep learning model (CNN) are trained. The CNN model consists of convolution layers, max-pooling layers, dropout layers, and dense layers, and is trained using the prepared dataset. The trained models are evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 Score, and their predictions are visualized using scatter plots comparing actual vs predicted values.

Once trained, the models are saved for deployment, with KNN and AdaBoost stored as .pkl files using joblib, and the CNN model saved as an .h5 file. If pre-trained models exist, they are loaded from disk to optimize computational resources. The trained models are then used for predictions on new data (*testdata.csv*), ensuring that the test data undergoes the same preprocessing steps. Predictions are stored alongside actual values for further analysis. Finally, Result Analysis and Reporting compares the performance of KNN, AdaBoost, and CNN, visualizes the results using scatter plots, and summarizes their performance using evaluation metrics, providing insights into the effectiveness of different modelling approaches for Signal Strength Prediction.

The Improved CNN Model is an advanced version of a Convolutional Neural Network (CNN) designed specifically for predicting RF signal strength in 5G networks. Unlike traditional CNNs used for image classification, this model has been adapted for regression tasks, allowing it to predict continuous values such as signal strength based on various environmental and network factors. CNNs are well-suited for this task because they can automatically learn spatial hierarchies and patterns from raw input data, eliminating the need for extensive manual feature engineering.

The architecture of the model includes multiple 1D convolutional layers, followed by max-pooling layers to reduce dimensionality, dropout layers to prevent overfitting, and fully connected (dense) layers for the final prediction. This structure enables the model to efficiently capture temporal and spatial dependencies in the RF signal data, making it more robust in generalizing across different network conditions. The model is trained on features such as geographical data, network topology, and environmental factors, making it a powerful tool for optimizing 5G network planning and improving coverage accuracy.

4. EXPERIMENTAL ANALYSIS

The implementation of the deep learning-based RF signal propagation modeling system for 5G networks follows a structured approach to achieve accurate signal strength predictions and optimize network planning. The process begins with setting up the necessary libraries and dependencies. Warnings are suppressed to avoid unnecessary logs, and essential data science and machine learning libraries such as pandas, numpy, scipy, seaborn, and matplotlib are imported for data manipulation, analysis, and visualization. The scikit-learn library is utilized for tasks like data splitting, scaling, and implementing regression models, while TensorFlow/Keras is used for building the deep learning-based Convolutional Neural Network (CNN) model.

The next step involves data loading and preprocessing. The dataset (logged_data.csv) is loaded into a pandas DataFrame, and its structure is analyzed by checking for missing values, duplicated records, and overall data distribution. Categorical features are encoded into numerical values using LabelEncoder, ensuring the data is machine-learning compatible. A heatmap is generated to visualize correlations between features and the target variable (Signal Strength), providing insights into feature relationships. The dataset is then split into input features (X) and the target variable (y), followed by further splitting into training and test sets, with 80% allocated for training and 20% for testing. Standardization is applied using StandardScaler to normalize the feature values, improving the model's efficiency. Additionally, synthetic data is generated using make_regression to augment the dataset and enhance model generalization.

The model training and evaluation phase consists of implementing multiple predictive models. A K-Nearest Neighbors (KNN) regressor is trained on the scaled dataset, and if a pre-trained model (knn_regressor.pkl) is available, it is loaded instead of retraining. The performance of the KNN model is evaluated using regression metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score. Similarly, an AdaBoost regressor is trained and evaluated using the same approach. The deep learning-based CNN model is then implemented using Keras, comprising multiple convolutional layers, max-pooling layers, dropout layers for regularization, and fully connected layers for regression output. The model is compiled with the Adam optimizer and Mean Squared Error loss function. If a pre-trained CNN model (improved_cnn_model.h5) exists, it is loaded for predictions; otherwise, it is trained, saved, and evaluated.

Evaluation metrics and visualization play a crucial role in assessing model performance. A custom function (calculateRegressionMetrics) calculates key regression metrics, including MAE, MSE, RMSE, and R^2 score. A scatter plot is generated to compare actual vs. predicted values, providing a visual representation of the model's accuracy. For the CNN model, another function (calculate_metrics) evaluates performance using the same metrics and visualizations. After evaluation, the model is applied to test data (testdata.csv). The test dataset is reshaped for CNN compatibility, and predictions are generated, with results stored in a DataFrame for further analysis. The trained KNN and AdaBoost models are saved using joblib, while the CNN model is stored using Keras' .save() function, allowing future reuse without retraining.

The dataset used in this project is crucial for modeling RF signal propagation in 5G networks. It includes real-world logged RF signal data, along with synthetic data for additional robustness. The dataset, stored in CSV format, consists of thousands of rows with various features categorized into network parameters, environmental parameters, user equipment data, and weather data. Key network parameters include frequency (Hz), transmission power (dBm), and bandwidth (MHz), which directly influence signal strength. Environmental parameters such as building density, vegetation index, and terrain type affect signal propagation, while user equipment data (latitude, longitude, and device specifications) provide additional context. Weather conditions, including temperature, humidity, and rainfall, contribute to signal attenuation. The target variable, signal strength (RSS), ranges between -120 dBm (weak) to -30 dBm (strong), serving as the primary prediction objective.

Data preprocessing is essential to prepare the dataset for machine learning models. This includes handling missing values, encoding categorical data like terrain type into numerical values, and scaling continuous variables for consistency. Synthetic data is generated using make_regression from Scikit-Learn to create additional samples, ensuring the model encounters diverse scenarios for improved robustness. By following this systematic approach, the project successfully implements a deep learning-based RF signal propagation modeling system, capable of making accurate signal strength predictions to optimize 5G network planning.

FIGURE 2. Sample Dataset to check the Signal Strength.

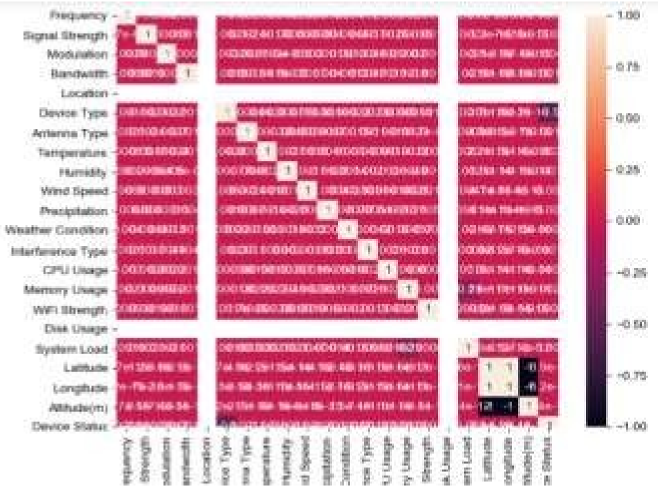

FIGURE 3. Heat map for column importance

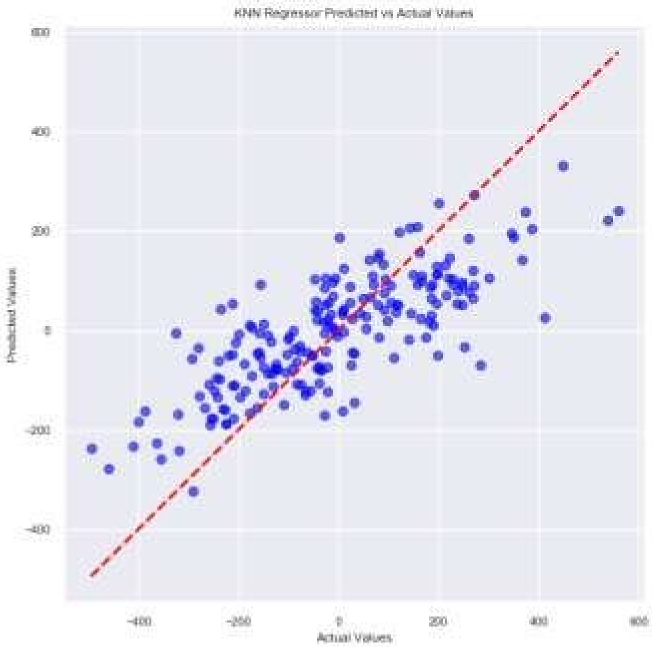

FIGURE 4. Illustration of confusion matrix obtained using KNN model.

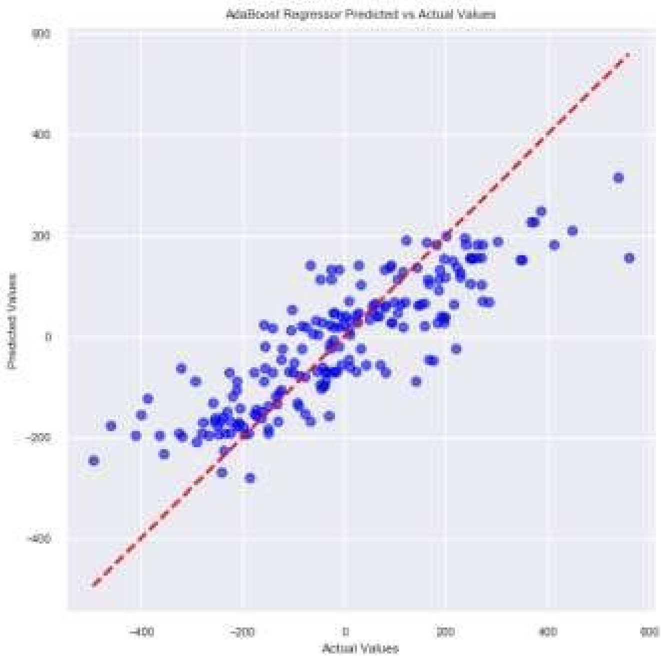

FIGURE 5. Illustration of confusion matrix obtained using AdaBoost model.

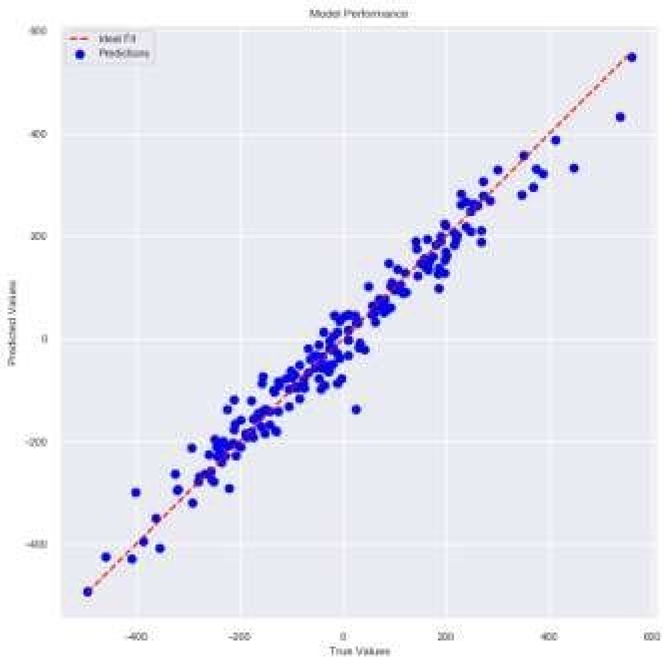

FIGURE 6. Illustration of confusion matrix obtained using CNN model.

Model name	RMSE	R^2-score
KNN	126.33	0.58
AdaBoost Regressor	109.03	0.69
CNN	1422.1124	96.26

TABLE 1. Comparison of all models.

5. CONCLUSION

The integration of deep learning techniques, particularly the improved Convolutional Neural Network (CNN) model, into RF signal propagation modelling for 5G networks presents a significant advancement over traditional methods. While traditional RF signal models such as path loss, ray tracing, and empirical measurements have been integral to network planning and optimization, they often face limitations in accuracy, computational efficiency, and adaptability to complex, dynamic environments. These traditional models also struggle to handle the intricacies of modern 5G networks, which involve highly dense deployments, high-frequency signals, and rapidly changing environmental factors.

The proposed deep learning-based approach leverages CNNs to analyse large datasets and learn complex patterns in RF signal propagation, improving the accuracy and efficiency of signal strength predictions. By training the model on diverse data, the improved CNN model can adapt to different environments, making it a promising tool for real-time network optimization and planning. In addition, the model's ability to continuously improve with more data and its potential for real-time predictions make it suitable for dynamic 5G environments.

REFERENCES

[1] Tataria, H.; Haneda, K.; Molisch, A.F.; Shafi, M.; Tufvesson, F. Standardization of Propagation Models: 800 MHz to 100 GHz—A Historical Perspective. 2020. Available online: http://arxiv.org/abs/2006.08491.

[2] Viswanathan, H.; Mogensen, P.E. Communications in the 6G Era. *IEEE Access* 2020, *8*, 57063–57074.

[3] Shen, S.; Zhang, W.; Zhang, H.; Ren, Q.; Zhang, X.; Li, Y. An Accurate Maritime Radio Propagation Loss Prediction Approach Employing Neural Networks. *Remote Sens.* 2022, *14*, 4753.

[4] Oueis, J.; Strinati, E.C. Uplink traffic in future mobile networks: Pulling the alarm. In *Cognitive Radio Oriented Wireless Networks, Proceedings of the 11th International Conference, CROWNCOM 2016, Grenoble, France, 30 May–1 June, 2016*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 583–593.

[5] Fehske, A.; Fettweis, G.; Malmodin, J.; Biczok, G. The global footprint of mobile communications: The ecological and economic perspective. *IEEE Commun. Mag.* 2011, *49*, 55–62.

[6] Imoize, A.L.; Udeji, F.; Isabona, J.; Lee, C.-C. Optimizing the Quality of Service of Mobile Broadband Networks for a Dense Urban Environment. *Future Internet* 2023, *15*, 181.

[7] Mohammadjafari, S.; Roginsky, S.; Kavurmacioglu, E.; Cevik, M.; Ethier, J.; Bener, A.B. Machine learning-based radio coverage prediction in urban environments. *IEEE Trans. Netw. Serv. Manag.* 2020, *17*, 2117–2130.

[8] Lim, S.Y.; Yun, Z.; Iskander, M.F. Propagation measurement and modeling for indoor stairwells at 2.4 and 5.8 GHz. *IEEE Trans. Antennas Propag.* 2014, *62*, 4754–4761.

[9] Morita, Y.; Rezaeiravesh, S.; Tabatabaei, N.; Vinuesa, R.; Fukagata, K.; Schlatter, P. Applying Bayesian optimization with Gaussian process regression to computational fluid dynamics problems. *J. Comput. Phys.* 2022, *449*, 110788.

[10] Zakaria, Y.A.; Hamad, E.K.I.; Elhamid, A.S.A.; El-Khatib, K.M. Developed channel propagation models and path loss measurements for wireless communication systems using regression analysis techniques. *Bull. Natl. Res. Cent.* 2021, *45*, 54.