[1] Awal Halifa*

[2] Kwame Oteng Gyasi

[3] Emmanuel Affum Ampomah

[4] Kwame Opuni-Boachie Obour Agyekum

[5] Kingsford Sarkodie Obeng Kwakye

[6] Justice Owusu Agyeman

[7] Mubarak Sani Ellis

[8] Piyush Kumar Shukla

Enhanced Path Loss Prediction Using Machine Learning and Modified COST-Hata Model for High-Frequency Wireless Networks

Journal of Electrical Systems

Abstract: - Accurate path loss prediction is critical for optimizing wireless network performance, especially in complex environments such as urban canyons, hilly terrains, and dense vegetation. Traditional models like COST-Hata exhibit significant limitations under non-line-of-sight (NLOS) conditions, necessitating more adaptable approaches. This study addresses these challenges by integrating a modified COST-Hata model with advanced machine learning techniques to enhance prediction accuracy and generalization across diverse environments. The proposed model incorporates key propagation factors—including frequency dependence, terrain effects, antenna height, building height, and angular dependencies—to refine empirical path loss estimations. Five machine learning models—Support Vector Regressor (SVR), Decision Tree Regressor (DTR), Gradient Boosting Regressor (GrB), Random Forest Regressor (RFR), and Artificial Neural Network (ANN)—were trained on refined parametric equations derived from the modified COST-Hata model.

Results indicate that ensemble-based models (RFR, GrB) significantly outperform traditional empirical models, demonstrating superior generalization and reduced prediction errors. After hyperparameter tuning, RFR achieved the lowest Mean Absolute Error (MAE) of 2.90 dB, underscoring its robustness across varied environments. Furthermore, the integration of terrain, building height, and angular dependencies allows for a more realistic representation of signal propagation. Compared to conventional models, the proposed hybrid approach exhibits higher accuracy, improved adaptability to NLOS conditions, and enhanced predictive stability. These advancements contribute to more efficient network planning, optimized resource allocation, and cost-effective wireless deployment. Ultimately, this study paves the way for next-generation intelligent path loss models, bridging the gap between empirical insights and machine learning-driven optimization.

Keywords: Path Loss Prediction, Machine Learning, Ensemble Methods, Wireless Communication Networks, Model Optimization, Feature Selection.

I. Introduction

To optimize communication systems, understanding electromagnetic wave (EM) propagation is crucial, as signal strength diminishes with increasing distance between transmitting and receiving antennas. Three propagation mechanisms—scattering, diffraction, and reflection—contribute to radio wave behavior [1]. Predicting path loss is complex due to the dynamic propagation environment, where path loss signifies the attenuation of radio waves during travel [2]. A precise path loss model is essential for coverage planning, base station site selection, and system performance enhancement.

A. Background

The rapid evolution of wireless communication, particularly with the advent of 5G and the transition to 6G, has heightened the need for accurate path loss prediction to optimize network performance. Traditional empirical models, such as COST-Hata, have been instrumental in network planning but struggle in complex environments due to their reliance on fixed parameters. While deterministic models offer higher precision by incorporating site-specific details, their computational demands limit their practicality for large-scale applications.

To address these challenges, machine learning (ML) techniques have emerged as a promising alternative for path loss modeling. Unlike empirical models, ML methods learn directly from data, capturing nonlinear interactions between factors such as frequency, terrain, antenna heights, and environmental conditions. Recent research explores

[1] Student Member, IEEE, Department of Telecommunications Engineering, Kwame Nkrumah University of Science & Technology, Kumasi, Ghana
[2, 3, 4, 5, 6, 7, 8] Department of Telecommunications Engineering, Kwame Nkrumah University of Science & Technology, Kumasi, Ghana
* Corresponding Author Email: ahalifa@tatu.edu.gh

hybrid approaches that integrate a modified COST-Hata model with ML algorithms, leveraging the predictive strengths of both empirical knowledge and data-driven adaptability. This fusion not only refines path loss estimation but also enables models to generalize across diverse propagation environments.

Accurate path loss modeling is critical for key wireless network tasks, including link budgeting, base station placement, and coverage optimization. In interference-prone environments, precise models improve the Signal-to-Interference-plus-Noise Ratio (SINR), enhancing overall network performance. By combining empirical modeling with ML-driven refinements, researchers can develop more adaptive and intelligent frameworks, advancing the reliability, efficiency, and scalability of next-generation wireless networks.

B. Machine Learning for Path Loss Prediction

Machine learning, a data-driven approach, has emerged as a promising alternative to overcome these limitations. Support Vector Regression (SVR), Random Forest (RF), Artificial Neural Network (ANN), Gradient Boosting Regression (GBR), and K-Nearest Neighbor (KNN) are supervised learning regression models demonstrating potential in path loss prediction [9]. Machine learning surpasses empirical and deterministic models due to its ability to analyze vast datasets, offering enhanced accuracy [9].

C. Study Objective

To address the shortcomings of existing models, this study aims to develop a flexible and adaptable ensemble path loss prediction model using machine learning. This model, informed by input characteristics like frequency, antenna height, distance, and environmental conditions, strives to forecast path loss effectively in diverse 4G wireless communication settings. The objective is to enhance accuracy compared to conventional models, providing valuable insights for system optimization, coverage planning, and performance assessment through extensive machine learning datasets.

II. Problem Statement

Traditional path loss models, including empirical and deterministic approaches, face significant challenges in modern wireless networks. Empirical models lack site-specific accuracy, while deterministic models, though precise, are computationally intensive and impractical for large-scale applications. Machine learning (ML) has emerged as a promising alternative, yet existing ML-based models often struggle with generalization across diverse environments. To address these limitations, this research proposes an enhanced path loss prediction framework that integrates a modified COST-Hata model with advanced ML algorithms, improving adaptability, accuracy, and efficiency in wireless communication networks.

III. Research Questions

1. **Limitations of Traditional Models:** What are the shortcomings of empirical and deterministic path loss models in accurately predicting signal attenuation across diverse environments?
2. **ML Integration for Improved Accuracy:** How can machine learning techniques, combined with a modified COST-Hata model, enhance the accuracy and adaptability of path loss prediction?
3. **Critical Features for ML Performance:** Which factors, such as frequency, antenna height, and environmental conditions, are most crucial for improving ML-based path loss predictions?
4. **Comparison of ML Approaches:** How does the proposed ensemble ML approach perform relative to individual models in terms of accuracy and computational efficiency?
5. **Practical Implications:** What benefits does the proposed model offer for network coverage planning, optimization, and performance assessment?

IV. Novelty

This research introduces three key innovations:

1. **Enhanced COST-Hata Model:** A modified version of the COST-Hata model that incorporates high-frequency dependencies, terrain irregularities, and urban architectural effects for more accurate signal propagation modeling.
2. **Ensemble Machine Learning Framework:** A novel approach that integrates multiple supervised learning models to enhance robustness, generalization, and prediction accuracy.
3. **Feature Importance Analysis:** A detailed analysis identifying key factors—such as frequency, antenna height, and environmental conditions—that significantly impact path loss, improving both model interpretability and understanding of wireless signal behavior.

V. CONTRIBUTIONS

1. **Integration of Machine Learning:** Implementation of supervised learning models (SVR, RF, GBR, DTR, ANN) to enhance path loss prediction accuracy across various wireless environments.
2. **Generalization across environments:** The proposed model is adaptable to both rural and urban settings, ensuring reliable performance in diverse propagation conditions.
3. **Practical Impact on 4G and 5G Networks:** Provides valuable insights for network design, coverage planning, and base station deployment, supporting efficient wireless communication system optimization.

VI. AIM

This research aims to enhance the COST-Hata model with novel parametric equations and integrate advanced machine learning techniques to develop a hybrid ensemble framework for path loss prediction. The goal is to improve accuracy, robustness, and adaptability across diverse wireless environments, supporting efficient network planning and optimization.

VII. OBJECTIVE

This research aims to develop an accurate and adaptable path loss prediction framework by integrating a corrected COST-Hata model with advanced machine learning techniques, improving prediction accuracy across diverse propagation environments and building effects, integrating machine learning techniques to improve accuracy in diverse environments.

A. *Specific Objectives*

The specific objectives are to:

1. Develop a modified COST-Hata model by incorporating frequency, terrain, angular, and environmental factors.
2. Analyze the predictive capabilities of various machine learning models (SVR, DTR, GrB, RFR, ANN).
3. Investigate the impact of hyperparameter tuning on model accuracy.
4. Compare the performance of untuned and tuned machine learning models across different environments.
5. Design an ensemble approach to enhance generalization and prediction accuracy.
6. Derive parametric equations from optimized models for improved interpretability.
7. Evaluate the effectiveness of tuned ensemble models in challenging propagation environments (e.g., urban canyons, hilly terrains, dense vegetation).

VIII. MOTIVATIONS

1. Improve path loss prediction accuracy for modern wireless networks (4G, 5G).
2. Leverage machine learning for better modeling of complex relationships.
3. Use ensemble methods to enhance model robustness.
4. Develop parametric models for practical application and interpretability.
5. Address challenges of high-frequency, modern network environments

IX. MACHINE LEARNING BASED PATHLOSS PREDICTION

The following literature review focuses on the various study domains on feature variables and feature selection strategies used to develop path loss prediction models. Several researchers have developed path loss prediction systems based on machine learning for a variety of area circumstances. One of them [11] uses an Artificial Neural Network (ANN) model to examine different indoor building styles. While [12] and [13] investigate various area types such as rural, suburban, and urban areas, [10] investigates path loss prediction in suburban regions using a range of machine learning models. Other studies that have been conducted in various locations with a unique measuring field include [14], which examines route loss prediction in an enclosed space such as an Airplane cabin. Path loss prediction of pathloss with the aid of ensemble machine learning methods has only been the subject of a few studies. This suggests that there is still opportunity for advancement in the study of path loss prediction with ensemble method approach.

In [14], the concept of path loss and how to predict it using machine learning was investigated. They also looked into the concept of path loss and defined the underlying premise of ML-based path loss predictors. We may use machine learning approaches to construct an adequate estimation function for path loss prediction if we have the results (path loss measurement) and the relevant input features, such as antenna separation distance and frequency.

This function, which can be either a white box (in decision-tree-based models) or a black box (in SVR-based or ANN-based models), maps input features to path loss values [15].

The gathered information relates to measurement samples, each of which contains the path loss value and the associated input parameters. System-sensitive and environment-based parameters are the two types of input features. The propagation environment has little effect on characteristics such as carrier frequency, receiver as well as transmitter heights and positions, and other system-sensitive parameters. Additional system-dependent aspects, such as the antenna separation distance and the angle between the line-of-sight path and the horizontal plane, can be determined using the aforementioned parameters [15].

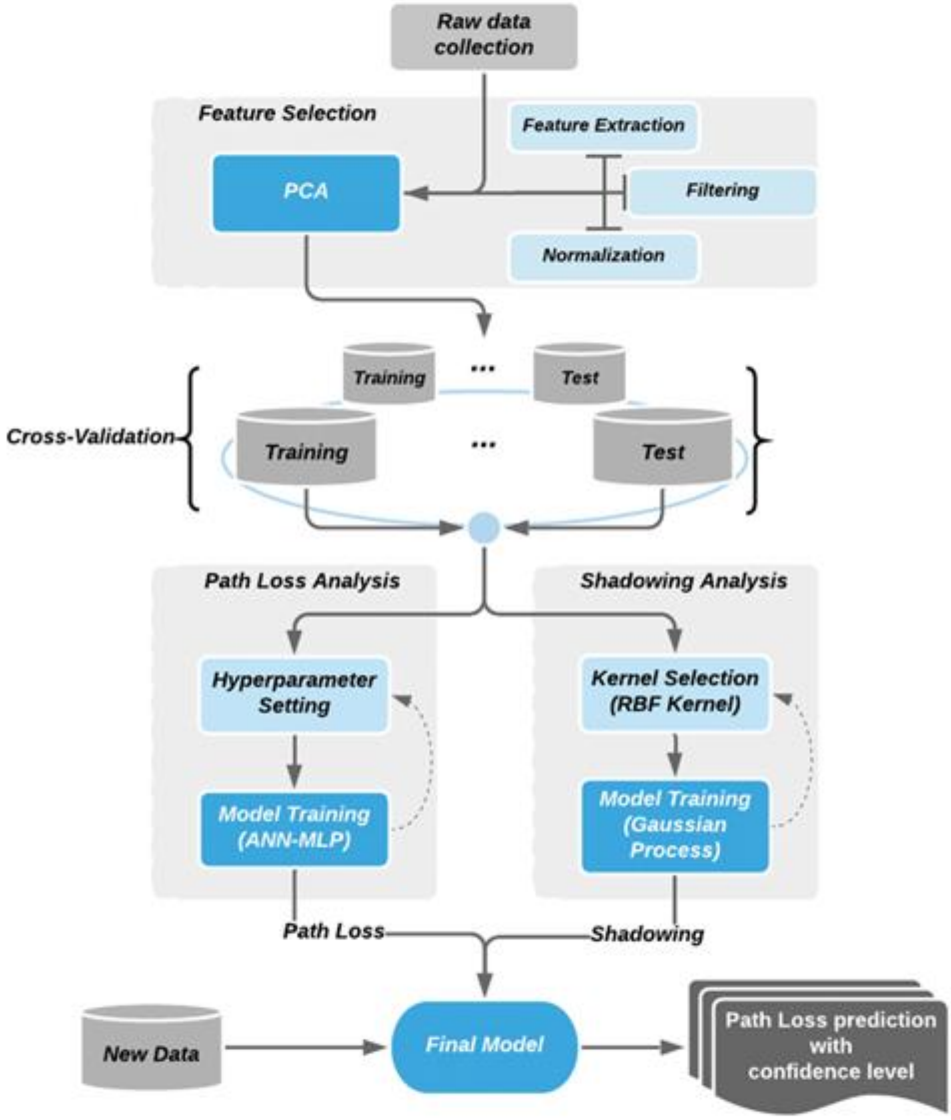

Fig. 1. Machine learning-based path loss analysis approach [8].

In order to develop the ability to generalize (to generate predictions based on previously unknown inputs), the models can be trained with sets of provided path loss values and matching inputs. We may use machine learning approaches to find a reasonable estimation function for path loss prediction once we know the output and the relevant input features such as antenna separation distance and frequency. This function maps input features to path loss values and can be either white (in decision-tree-based models) or black (in SVR-based or ANN-based models). The technique for machine learning-based path loss predictors is depicted in the picture below and is described in detail below.

A. Feature Variables and Selection

When creating path loss prediction models, a wide range of input feature kinds and quantities are used. The distance between the transmitter (TX) and receiver (RX) is the only input information used in the research in [18], [19]. The investigation conducted by [20] incorporates the frequency feature as an additional input, supplementing the TX-RX distance feature. In a similar vein, [8] introduces two extra features, along with the TX-RX distance feature, incorporating onboard GPS sensors. [21] includes the TX-RX distance feature as part of the input parameters, alongside PCC downlink throughput and PDCP downlink throughput. In the exploration carried out by [14], which focuses on the interior of an aircraft cabin, the user's position, determined by longitude and latitude, serves as an input parameter, among others. Furthermore, both [11] and [22] examine outdoor locations using longitude and latitude in their respective studies. In certain research endeavors, environmental attributes are utilized as input features alongside system parameters. Reflecting the particular characteristics of the research domain, some studies opt for a more intricate blend of criteria to generate results of utmost accuracy. For instance, in the examination conducted by [23] and [24], a total of six input factors—longitude, latitude, elevation, altitude, clutter height, and TX-RX distance—are employed. This illustrates that the nature and quantity of features can be further refined to align with the specific focus of the research field.

B. Feature Scaling

In real life, the machine learning data may have hundreds of features. Poor predictor quality might result from either keeping irrelevant features or excluding important features. Finding the best subset with the fewest characteristics that contribute most to learning accuracy is the aim of feature selection [25].

The size of the input space can affect the performance of some machine-learning-based algorithms including RF, SVR, and GBR. As a result, the normalization procedure should be complete before the training starts. That means, the values of all input characteristics and path loss should be modified to fall within the range of -1 to 1 or 0 to 1. This work uses the same normalization technique as [27], with the same results. It can be stated as

$$x_N = \frac{2(x - x_{min})}{x_{max} - x_{min}} - 1 \tag{1}$$

where x represents the value undergoing normalization, x_{min} and $x_{max}$are the minimum and maximum values of the data range, respectively, and x_N is the value after normalization. By applying anti-normalization in accordance with the normalization procedure, the expected values can be produced. By contrast, decision-tree-based approaches do not demand the feature scaling.

C. Configuring Hyperparameters and Training the Model

Hyperparameters, which are predefined values prior to the commencement of the learning process, encompass elements such as the number of hidden layers and neurons in an artificial neural network (ANN), the coefficients for regularization and parameters within the kernel function for support vector regression (SVR), and factors like ensemble size and tree depth in decision-tree-based approaches. To optimize the efficiency and performance of path loss prediction, it is crucial to carefully select an ideal set of hyperparameters. Grid search, random search, and Bayesian optimization stand out as prominent methods for hyperparameter optimization. In this investigation, the grid search methodology was employed to ascertain the final values of the hyperparameters. This approach entails a comprehensive search, evaluating all potential parameter values before identifying the most effective ones.

Model parameters, on the other hand, are inherent to training samples, evolving as part of the model training process. It is noteworthy that diverse learning methodologies involve distinct model parameters, with elements like weights and biases being autonomously acquired during the model training phase.

The research paper [29], discussed how they may train new models for predicting path loss within buildings or indoors. They argued that training is the most important and crucial role in the modelling problem. In reality, a well-trained model must be able to extrapolate or interpolate with high accuracy based on existing knowledge learned during learning from a new input in order to predict the expected output. The literature has a number of proposed neural network learning techniques, which can be categorized into supervised and unsupervised learning [30]. Unsupervised learning is used to cluster data, and this method separates the data into groups based on certain characteristics. With supervised learning techniques like the gradient descent approach, the input parameters and output values are known, and the neural network can offer an inferred function that can be used to map fresh samples. To lower the mean squared error (MSE) between the input and desired output of the neural model, the bias and weight of each neuron can be changed [30].

When creating a precise neural network model, the most influential parameters possible must be taken into account. The multi-wall model is where the inputs for the model that we developed and presented in this study come from. The elements of (L f) consist of the transmitter-receiver distance (d), frequency (f), and the attenuation caused by walls and floors (L w). The model is designed with a singular hidden layer. This hidden layer's number of neurons is set to 75% of the input layer's number of neurons [30]. This number may be altered, and the results of doing so will be discussed in the section that follows. In the output layer, there is only one output that represents the measured signal route loss.

D. Model Evaluation and Prediction

Typically, samples from the test dataset—which are absent from the model training process—are used to gauge how well machine learning-based route loss models perform. The evaluation criteria include complexity, generalization ability, and prediction accuracy. Performance metrics like maximum prediction error (MaxPE), mean absolute error (MAE), error standard deviation (ESD), root mean square error (RMSE), and mean absolute percentage error (MAPE) are commonly employed [9].

When deploying the model in scenarios with extra frequency bands or different environmental types, its reusability is determined by its generalization property. Gathering more data from diverse settings, including various terrains, frequencies, and vegetative cover conditions, has the potential to enhance the model's generalization capability.

Usually, processing speed and memory usage are used to gauge how difficult a computation is. Key factors influencing the processing duration of a machine learning model encompass, for example, the number of iterations and the speed at which convergence occurs throughout the training phase.

The machine learning algorithm can be chosen, the hyperparameters can be changed, and the prediction model can be further enhanced based on the evaluated outcomes. Following the construction of the ideal model, path loss values can be produced using fresh inputs.

In this study, we looked at the feature selection, hyperparameter tuning and optimization, and model selection and training available in wireless communication networks. It became clear, nonetheless, that specialized and improved path loss prediction features are needed due to the particular difficulties and complexity of heterogeneous smart city environments. Also, a more robust pathloss prediction model is needed to handle the limited features and make more accurate prediction.

X. Methodology

This chapter goes through the resources and procedures employed. The procedure for gathering data, the suggested model for predicting pathloss, and a comparison of various models are all provided. The use of supervised learning techniques such as Random Forest, SVR, and Gradient Boosting Regressor to predict path loss is introduced, and the performance of these methods is examined using measured data. Additionally, a theory describing how signals travel through empty space is offered.

A. Random Forest Regressor

The Random Forest (RFR) ensemble learning method, as described in [32], is comprised of multiple regression trees. To enhance the prediction performance of each tree and address its weak robustness, a voting mechanism is employed. Breiman and Cutler introduced this distinctive non-parametric supervised machine learning technique known as Random Forest [33].

"Bootstrap aggregation" is the origin of the bagging technique known as RF. The primary concept behind bagging is to take a dataset, bag a weak learner like a decision tree on it, then create several bootstrap duplicates of the dataset and develop decision trees on them. To choose various training samples for each tree, Bootstrap aggregating is used. After training the trees using these samples, the ultimate outcome is determined by averaging the performance of each individual tree.

The RF remains a valuable instrument for reducing dimensionality or eliminating redundancy in datasets. While datasets with high-dimensional input features provide more information, the presence of redundant and unnecessary components can diminish prediction accuracy. In this investigation, the RF approach was applied to process the observed signal dataset, extracting relevant features while eliminating unnecessary and unimportant ones. The RF algorithms involve two or more primary hyperparameters that need specification before deploying them for regression analysis or data training [34]. One of these hyperparameters is the number of trees. The mathematical explanation of the RF input-output function model is detailed below.

$$RF(x_n, y_n) = \{ f(x_n, \theta_m, y_n) \} \tag{2}$$

Here, θ_m represents the number of trees while x_n, y_n denote the input and target output data, respectively. In this instance, a set of 200 trees was employed on the target measured signal datasets to accomplish the task of identifying the most valuable and informative subset of characteristics.

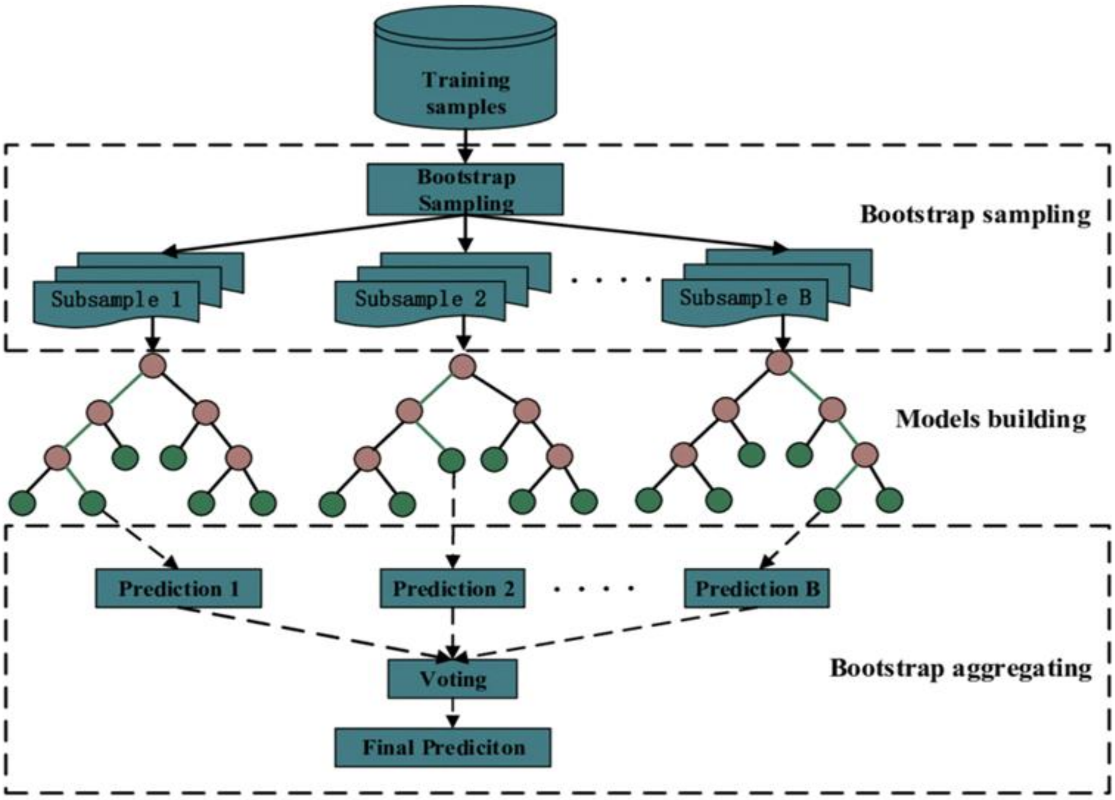

Fig. 2. Harnessing the combined predictions from every tree in the forest, the final outcome is determined through a majority vote [35].

B. Gradient Boosting Regressor

The Gradient Boosting Regressor (GBR) is a member of the boosting algorithms, falling under the category of ensemble learning techniques [35]. Its approach involves amalgamating the strengths of diverse regression trees to construct a predictive model. Originally introduced by Jerome Friedman [36], GBR is designed to address the limitations of specific weak learners, aiming to produce a robust and accurate regression model.

In GBR, a series of regression trees are trained in an iterative fashion, with each new tree aiming to fix the mistakes of the preceding ones. The model's ability to anticipate outcomes is enhanced over time by this iterative process. Instead of using a voting mechanism, GBR concentrates on optimizing the residuals of the previous tree in the series. This is how it differs from Random Forest. Due to this quality, GBR is very good at identifying complicated links and optimizing forecasts.

Giving additional importance to observations that earlier trees failed to accurately anticipate is the fundamental idea of boosting. Consequently, GBR can focus on the sections of the dataset where the model exhibits suboptimal performance. Each subsequent tree is trained with the specific aim of minimizing the accumulated residual errors within the ensemble. GBR does exceptionally well at adapting to the subtleties and patterns found in the data.

GBR's versatility in handling multiple data kinds and issue domains is one of its advantages. It can deliver excellent predicted accuracy and is ideally suited for datasets with diverse feature sets. To prevent overfitting, GBR might need more meticulous hyperparameter tuning than Random Forest. The gradient-based loss function is minimized by the GBR method through a series of processes, including initializing the target predictions, computing negative gradients, and creating new regression trees. Combining all of the various trees' projections yields the ultimate conclusion.

Based on the observed signal dataset, the Gradient Boosting Regressor was used in this work to forecast path loss. The goal of GBR is to offer precise path loss estimates while managing the difficulties present in real-world wireless communication settings.

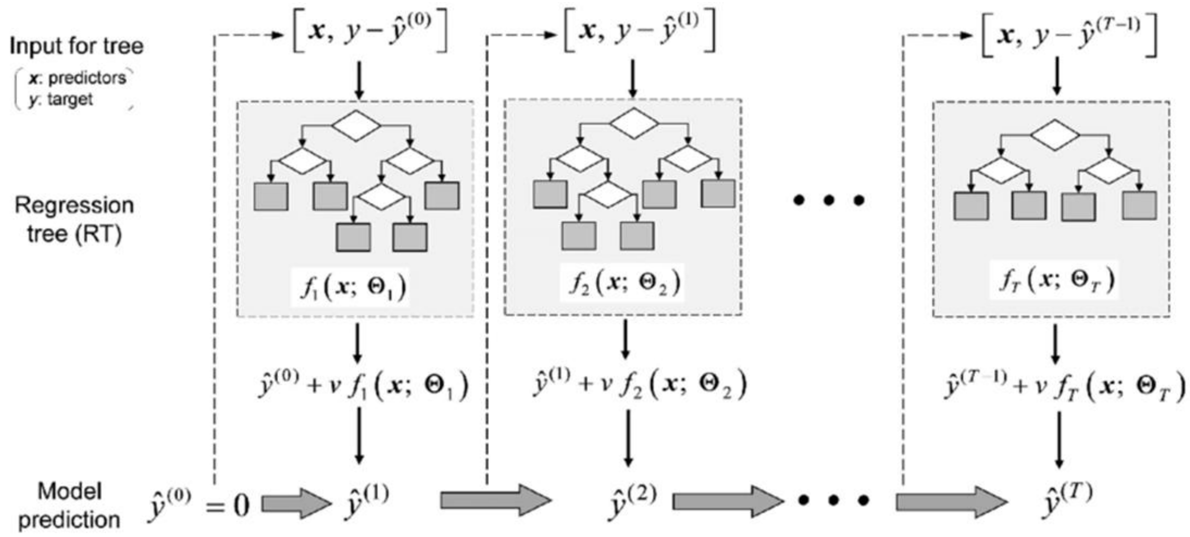

Fig. 3. Structure of the gradient boosting regressor [38].

Gradient Boosting as a Meta-Classifer:

1. **Gradient Boosting** (GB) is an ensemble learning technique where new models are trained to correct the errors (residuals) of existing models in the sequence. It is often used for regression and classification tasks.
2. A **meta-classifier** refers to a model that combines the outputs of multiple other classifiers to improve prediction accuracy. In boosting, each subsequent model focuses on correcting the errors made by the previous models, thus creating a "meta-model" that aggregates all these corrections.

In path loss prediction, **Gradient Boosting Machines (GBMs)** like **XGBoost**, **LightGBM**, or **CatBoost** are typically used:

3. **XGBoost**: Known for its high efficiency, scalability, and superior performance, XGBoost uses gradient boosting trees, where each tree is built using the gradient of the loss function, and it effectively handles large datasets and high-dimensional spaces.
4. **LightGBM**: A lighter, faster alternative to XGBoost, designed for large datasets with improved efficiency in terms of memory usage and computation speed, often producing similar or better results in less time.
5. **CatBoost**: Optimized for categorical data, handling features with missing values naturally and delivering excellent performance for tasks involving structured data.

Why ML Techniques Like Gradient Boosting Are Used Instead of Lightweight DL Techniques:

While **Deep Learning (DL)** techniques such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) have shown remarkable performance in many domains like image recognition and natural language processing, they are often more resource-intensive, both in terms of computation and data requirements. The key reasons why traditional **Machine Learning (ML)** techniques, especially **Gradient Boosting**, might be preferred in path loss prediction over lightweight DL methods include:

1. **Data Requirements**:
   i. Deep learning models generally require large amounts of labeled data to train effectively. In contrast, ML models like Gradient Boosting can perform well even with smaller datasets.
   ii. Path loss prediction is often based on structured features (e.g., geographical data, transmitter-receiver distance), and the dataset may not be large enough to justify the complexity of deep learning models.

2. **Model Interpretability**:
   i. One advantage of ML models, especially Gradient Boosting, is that they offer more interpretability compared to deep learning models. For example, feature importance in Gradient Boosting can be easily analyzed to understand the contribution of each variable (e.g., distance, frequency, and environment type) to the path loss prediction.
   ii. In contrast, DL models like neural networks often act as "black boxes," making it difficult to understand how inputs are transformed into predictions.

3. **Computational Efficiency**:
   i. Gradient Boosting techniques (like XGBoost and LightGBM) are highly optimized and can be more computationally efficient than deep learning models, especially for tabular or structured data.
   ii. Deep learning models, even lightweight ones, still require significant computational resources, particularly when training over large datasets or optimizing over many layers of neural networks.
4. **Handling of Structured Data**:
   i. Path loss prediction involves highly structured data (such as distance, location, elevation, and environmental factors), which machine learning models like Gradient Boosting are particularly well-suited to handle.
   ii. Deep learning techniques, particularly convolutional or recurrent networks, are more suited to unstructured data like images or sequential time series data, making them less ideal for tabular structured data used in path loss modeling.
5. **Ease of Implementation**:
   i. Implementing Gradient Boosting models like XGBoost or LightGBM is typically faster and easier than deep learning techniques, which often require complex tuning of network architecture, loss functions, and optimization algorithms.
6. **Robustness and Performance**:
   i. Gradient Boosting models, through their sequential learning process, can be quite robust and effective at capturing non-linear relationships and interactions between features, which are crucial in predicting path loss.
   ii. While lightweight DL techniques, such as shallow neural networks, may be applied, they often do not outperform well-tuned gradient boosting models in terms of accuracy or interpretability for tasks like path loss prediction.

C. Support Vector Regressor (SVR)

The support vector machine (SVM), rooted in statistical learning theory, represents a form of machine learning. SVM's core concept involves the linear separation of a dataset by nonlinearly transforming it from a finite-dimensional space to a higher-dimensional one. SVR, an extension of SVM tailored for regression challenges, enables path loss prediction [37]. The primary objective of SVR is to locate a hyperplane within the high-dimensional feature space and ensure that sample points align with it. The fundamental goal is expressed through the utilization of the following linear function to define the hyperplane in the feature space.

$$f(\boldsymbol{x}) = w^T \varphi(\boldsymbol{x}) + b \tag{3}$$

where x is an input feature vector, w is the normal vector that controls the orientation of the hyperplane, $\varphi(\cdot)$ is the nonlinear mapping function, and b is the displacement item.

The optimal hyperplane is formulated as a constrained optimization problem, as outlined in [38].

$$\min_{\boldsymbol{w}, b, \xi, \xi^*} \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{4}$$

$$s.t. f(x_i) - y_i \leq \varepsilon + \xi_i \tag{5}$$

$$y_i - f(x_i) \leq \varepsilon + \xi_i^* \tag{6}$$

$$\xi_i, \xi_i^* \geq 0, i = 1, \dots, N \tag{7}$$

Here, *C* represents the regularization coefficient, *ε* denotes the insensitive loss, signifying that a predicted value is deemed accurate if the difference between the predicted value and the actual value is less than *ε*. The variables ξ_i, ξ_i^* are slack variables, introducing flexibility in the insensitivity range on both sides of the hyperplane, allowing for slight variations.

Then, by presenting Lagrange multipliers and solving its dual problem, the approximate function can be expressed as

$$f(x) = \sum_{i=1}^{N}(-\alpha_i + \alpha_i^*)K(x_i, x) + b \tag{8}$$

where α_i, α_i^* are Lagrange multipliers, and $K(\cdot, \cdot)$ is a kernel function, which is used to perform the nonlinear mapping from the low-dimensional space to the high-dimensional space.

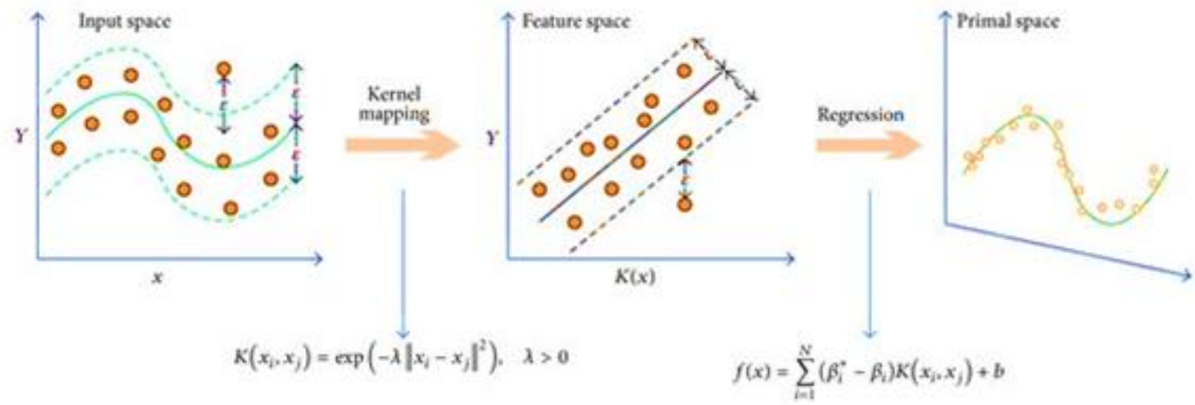

Fig. 4. A schematic diagram of SVR architecture [41]

The effectiveness of the SVR-based predictor relies on the choice of the kernel function. Presently, commonly used kernel functions include the sigmoid kernel, linear kernel, polynomial kernel, Gaussian radial basis function, and various combinations thereof. In this study, the selected kernel function is a Gaussian kernel with an adjustable parameter, and its definition is as follows:

$$K(x_i, x_j) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right), \gamma > 0 \quad (9)$$

The Gaussian kernel is often employed as a kernel function, particularly effective for tasks with restricted feature dimensions and no prior knowledge [39]. In this study, parameters like the regularization coefficient, insensitive loss, and kernel function parameter were determined using the same methodology as detailed in [37].

D. The proposed Pathloss Prediction Modeling Approach

The proposed model, combines the strengths of the Random Forest, Gradient Boosting, and Support Vector Regression algorithms in accordance with the goals of the study. The model's creation, application, and evaluation are divided into several phases, as listed below:

1. Data Collection and Preprocessing: A thorough field measurement campaign is carried out to collect Received Signal Strength (RSS) values and route loss data from various metropolitan contexts. The gathered data includes a range of terrain characteristics, antenna heights, separations, and frequencies, ensuring its representativeness. To extract pertinent features for path loss prediction, feature engineering is used in the data pretreatment processes to handle missing values, identify outliers, and handle outliers.
2. Ensemble Model Development: The key idea behind the suggested method is to combine Gradient Boosting, Random Forest, and Support Vector Regression into a single predictive model. The appropriate hyperparameters and input features produced from the preprocessed data are used to train and optimize each particular model.
3. Model Integration and Weighted Averaging: The ensemble integration of the individual RF, GB, and SVR models is achieved through a weighted averaging mechanism. The weights are assigned depending on the performance and applicability of each model, which is combined with the predicted outputs from each model. With the help of this integration, the predictive accuracy is intended to be improved by utilizing the complementing capabilities of several algorithms.
4. Cross-Validation and Model Tuning: The model is fine-tuned to optimize its hyperparameters and undergoes cross-validation to evaluate its generalization performance. The ensemble model is made stable and calibrated by this repeated process, enabling it to make precise predictions in a range of scenarios.
5. Performance Evaluation and Comparison: A thorough performance study is carried out in order to fully assess the suggested model. The assessment of the model involves employing diverse evaluation criteria, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and correlation coefficients. This comprehensive evaluation aims to gauge the accuracy, robustness, and generalization capabilities of the model. Using the same dataset, the model is rigorously contrasted with other path loss prediction methods, including Cost-Hata.

In order to provide a comprehensive solution for precise and reliable path loss prediction in heterogeneous radio network design, the Random Forest, Gradient Boosting, and Support Vector Regression, ensemble path loss prediction model is proposed. The model's effectiveness and superiority are proved by thorough validation and comparison with existing models, advancing path loss prediction in wireless communication networks for smart cities.

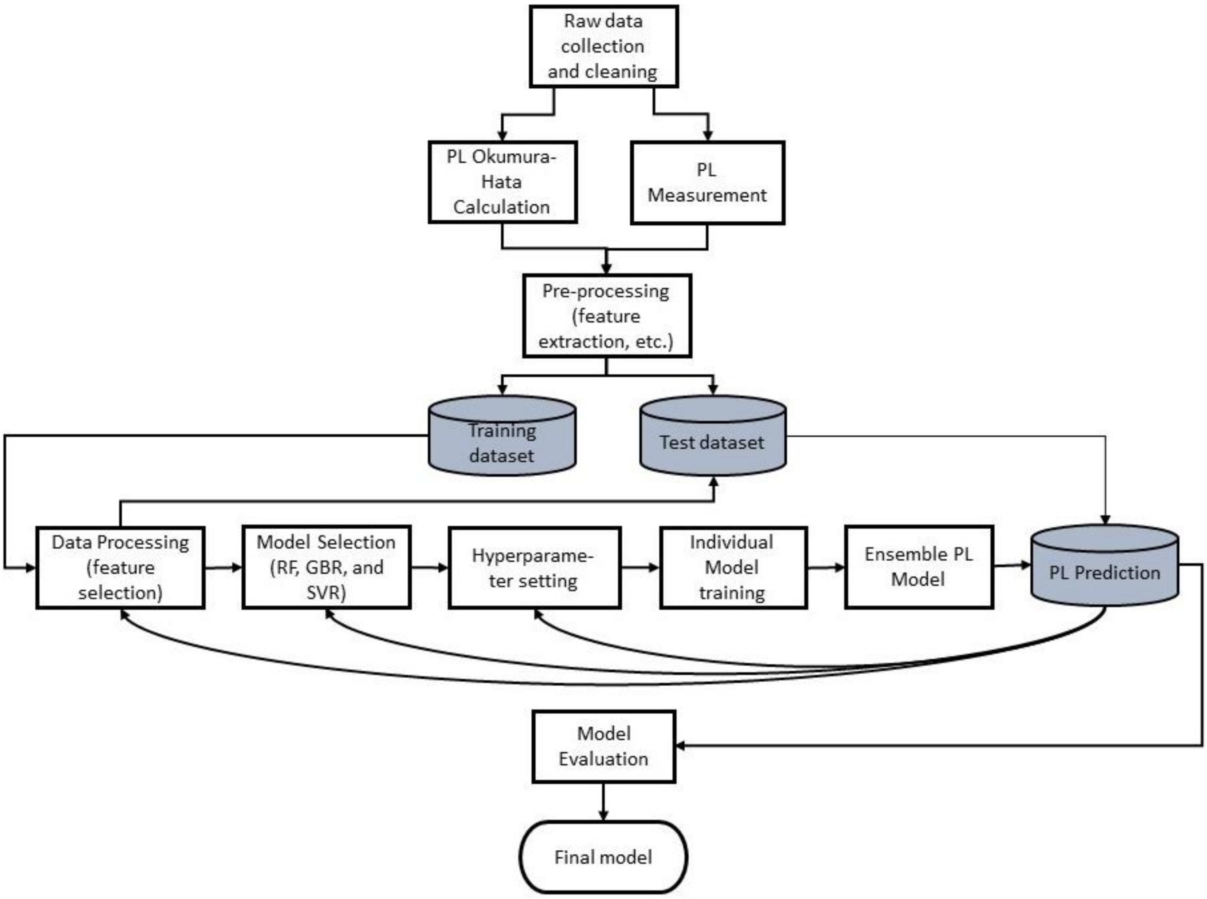

Fig. 5. The proposed Pathloss Prediction Model Architecture

E. Model Algorithms

ALGORITHM 1: *TRAIN BAGGING MODEL*

Inputs: *D (training data), r (number of models)*

Output: *M (array of trained models) as empty array*

1. *Load the dataset (D)*
2. *Perform one-hot encoding for categorical variables*
3. *Split the dataset into training and testing sets (X_train, X_test, y_train, y_test)*
4. *Train individual models given input data X_train, y_train:*
5. *Train Random Forest model (M1) with hyperparameter tuning*
6. *Train Gradient Boosting model (M2) with hyperparameter tuning*
7. *Train Support Vector Machine (M3) with hyperparameter tuning*
8. *Append trained models M1, M2, and M3 to M*
9. *Return array of trained models M*

ALGORITHM 2: *TESTING BAGGING MODEL*

Inputs: *D (training data), r (number of models), M (array of trained models)*

Output: *Y (bagging prediction)*

1. *Set parameters: M (array of trained models)*
2. *Initialize Y = 0*
3. *for each model m in M do*
4. *Determine prediction y from model m given input data X_test*
5. $Y = Y + y$
6. *end for*
7. $Y = \frac{Y}{r}$
8. *Return Y*

ALGORITHM 3: *TRAIN BLENDING MODEL*

Inputs: *D (training data)*

Output: *M (array of trained models) as empty array*

1. *Load the dataset (D)*
2. *Perform one-hot encoding for categorical variables*
3. *Split the dataset into training and testing sets (X_train, X_test, y_train, y_test)*
4. *Train individual models given input data X_train, y_train:*
5. *Train Random Forest model (M1) with hyperparameter tuning*
6. *Train Gradient Boosting model (M2) with hyperparameter tuning*
7. *Train Support Vector Machine (M3) with hyperparameter tuning*
8. *Train a combiner model **C**, with validation data y_test. The combiner model is a regression model that takes the prediction from M1, M2, and M3 as inputs. The validation input data is used as inputs to M1, M2, and M3. The outputs from the three models constitute the input to **C**.*
9. *Return trained models M1, M2, M3, and combiner model C*

ALGORITHM 4: *TESTING BLENDING MODEL*

Inputs: *D (training data), M (array of trained models)*

Output: *Y (blending prediction)*

1. *Set parameters: M (array of trained models), C (combiner model), M1 (RF model), M2 (GB model), M3 (SVM model)*
2. *Initialize Y = 0*
3. *Determine prediction m1 from model M1 given input data X_test*
4. *Determine prediction m2 from model M2 given input data X_test*
5. *Determine prediction m3 from model M3 given input data X_test*
6. *Determine prediction Y from model **C** given inputs m1, m2, and m3*
7. *Return Y*

F. Data Collection and Processing Campaign

Numerous measurement campaigns were conducted in urban, and rural parts of Bekwai in the Ashanti Region, Ghana. Two city drives were used to gather experimental data. Field measurements were conducted with the assistance of a Transmission Evaluation and Monitoring System (TEM). Real-time data measurement, analysis, and post-processing across the network are all capabilities of TEMs. The measurement setup consists of a 4G Android phone acting as the mobile station (MS), a USB connector, GPS, a laptop, serial cables, and TEMs mobile system dongle software. In a repeated campaign scenario, reference signal received power (RSRP) was measured for all sites between 23 m and 2.7 km in suburban, and urban regions of Bekwai using all of these linked devices. In order to reduce the Doppler effect, the vehicle moved at a constant speed. At an operational frequency ranging from 1.8 GHz to 2.1 GHz, numerous drive tests were carried out across networks.

Fig. 6. Cumulative Distribution Function (CDF) of the Dataset

For a transmitter-receiver distance ranging from 23 meters to 2.7 kilometers, the RSRP data were registered on the computer screen. The path loss was then calculated for every measured RSRP throughout the drive route. The first drive route received 2000 measurements, and the second drive route received 2330 measurements, for a total of 4330 measurements over the two drive routes. To achieve high precision, measurements were made at each spot eight times, with the average being determined. A reliable data preparation technique was used in the ensemble model building.

Before training the model, sufficient data preparation is required in order to attain accuracy in path loss predictions. Hundreds of features may be present in the actual data utilized for machine learning. Inaccurate path loss prediction might occur from both omitting or keeping unimportant data. The size of the input space has a significant impact on how machine learning models behave. Therefore, normalizing the data should be done before training the data. To prepare the data for this study, the following procedures were implemented. The data captured was imported into Python from an Excel file, where it was subsequently read and scrutinized for any instances of duplicate or missing values. Following this, the data underwent a normalization process.

Thirty percent of the pre-processed dataset was used for testing, with the remaining 70% going towards training. The use of uniform random sampling was made. In order to optimize the ensemble model, adjustments were made to the hyperparameter values using the training dataset. This allowed the incorporation of optimized network parameters in the RF, GBR, and SVR models, as well as the ensemble model. The equation below illustrates the calculation of path loss based on the measured received power.

$$Path\ loss\ (dB) = EIRP(dBm) - RSRP(dBm) \tag{10}$$

The aggregate power density transmitted from the base station to the surrounding medium is termed the effective isotropic radiated power (EIRP), represented in dBm.

Table I. Dataset Description

Feature	Mean	Min	Max	STD	25%	50%	75%
Distance (m)	594.960	23.740	2677.090	431.213	311.143	479.385	742.555
Frequency (MHz)	1930.878	1800.000	2100.000	148.793	1800.000	1800.000	2100.000
Height of TX (m)	45.229	25.000	63.500	9.816	39.000	43.000	55.000
Height of RX (m)	14.896	3.998	30.033	5.230	11.087	14.850	18.293
Path loss (dB)	119.471	81.000	162.000	17.373	104.000	120.000	134.000

The dataset description provides key statistics for the features relevant to path loss prediction in wireless networks. The features include Distance, Frequency, Height of Transmitter (TX), Height of Receiver (RX), and Path Loss**:**

1. Distance (m)**:**

- The average distance between the transmitter and receiver is 594.96 meters, with a range from a minimum of 23.74 m to a maximum of 2677.09 m, reflecting a diverse set of distances that will influence signal attenuation.
- The standard deviation of 431.21 m indicates significant variability, as expected in practical wireless network deployments.

2. Frequency (MHz)**:**

- The mean frequency is 1930.88 MHz, which lies within the targeted high-frequency range (1.8 GHz to 2.1 GHz).
- Values vary from 1800 MHz to 2100 MHz, covering frequencies typically used in 4G and potentially 5G networks, with the standard deviation of 148.79 MHz showing notable variation within the band.

3. Height of TX (m)**:**

- The transmitter height averages 45.23 meters, spanning from 25 m to 63.5 m, representing typical base station heights used in cellular networks.
- The standard deviation of 9.82 m suggests moderate variation, with quartile values indicating a slightly skewed distribution towards higher placements (e.g., at 75%, the height is 55 m).

4. Height of RX (m)**:**

- The receiver (or mobile device) height has a mean of 14.9 meters, with a range from 3.998 m (ground level or near-ground deployments) up to 30.033 m.
- This wide variation reflects deployment conditions where receivers could be in elevated locations, such as in high-rise buildings.

5. Path Loss (dB**):**

- The average path loss observed is 119.47 dB, with a range from 81 dB to 162 dB, indicating significant signal attenuation over various conditions.
- A standard deviation of 17.37 dB shows moderate variability in path loss, with higher values corresponding to greater signal weakening due to distance, obstacles, or environmental factors.

Overall, these descriptive statistics suggest the dataset covers a broad range of conditions for path loss prediction, making it suitable for training and testing robust predictive models. The diversity in distance, frequency, and antenna heights aligns well with the goals of understanding path loss behaviors in diverse and high-frequency wireless environments.

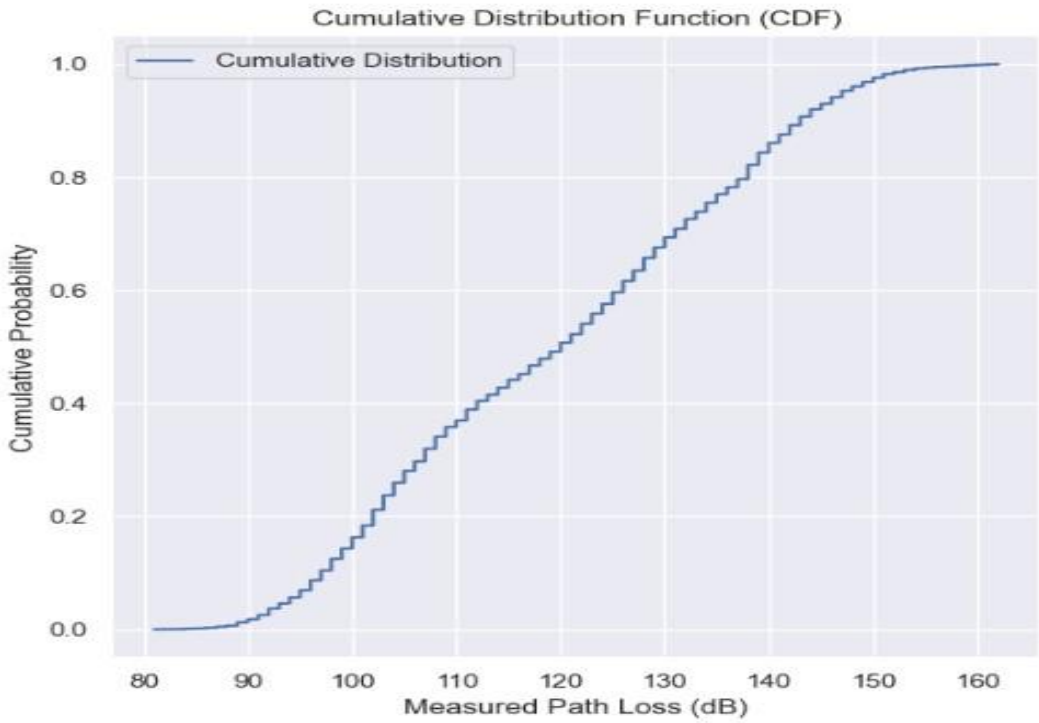

Fig. 7. Cumulative Distribution Function (CDF) of the Dataset

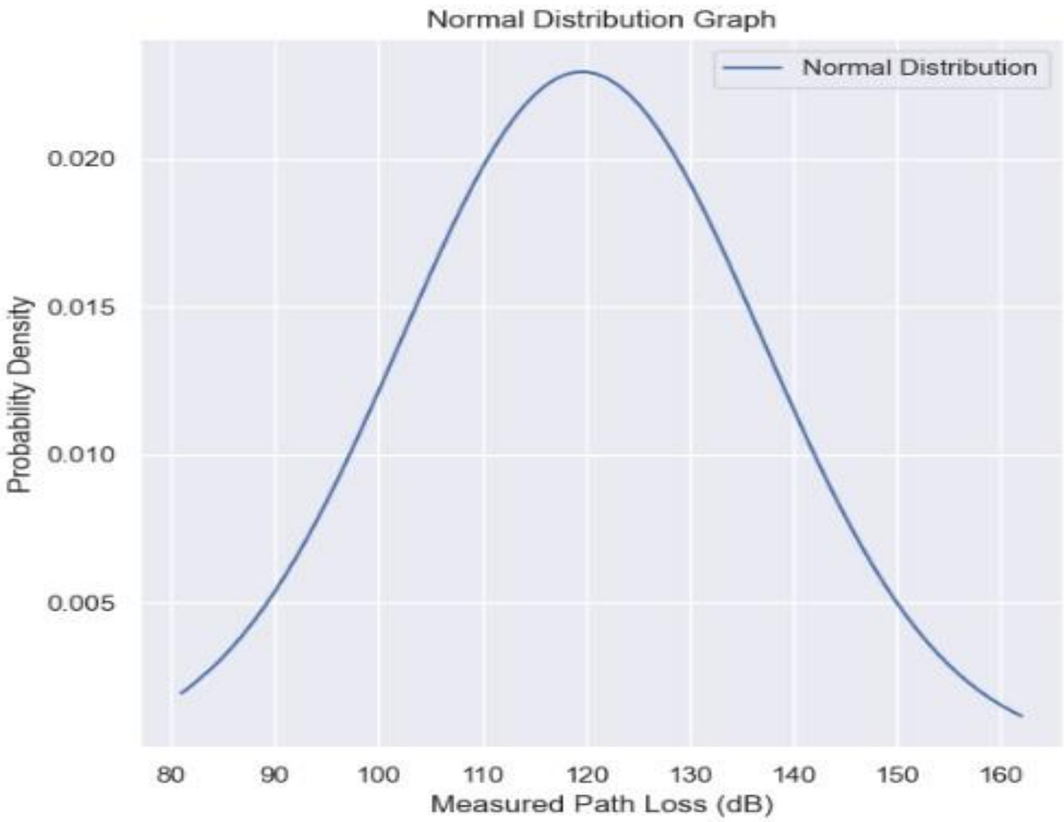

Fig. 8. Normal Distribution Graph of the Dataset

XI. Results and Discussions

The research study's findings are presented in this section along with illuminating discussions. The findings are divided into separate segments, each of which sheds light on a different aspect of the examination, in order to be consistent with the study's aims. The anticipated values of path loss are contrasted with the measured values and the values obtained by applying empirical models in order to assess the viability of the suggested strategies.

A. Comparison of Different Models

This section evaluates and compares the efficacy of machine learning models in predicting path loss against an empirical model (Cost-Hata). Multiple models are employed to predict the path loss value for each test data point. A total of 4330 samples were collected along the routes, each comprising path loss data and antenna separation distance computed using GPS information. For the training dataset, 70% of the samples were randomly selected, while the remaining 30% constituted the test dataset. To predict path loss values in the test dataset, three models—GBR, SVR, and RF—were utilized. The SVR model's regularization coefficient, insensitive loss, and kernel function parameter was set to 0.1, 10, and linear, respectively. The maximum tree depth and ensemble size for the RF and GBR based models, respectively, were 7 and 5, and there were 200 ensemble members. For comparison, the Cost-Hata model was also taken into account.

The measured data and the results that the various models are shown in Figures 8 to 11. The separation between the sending and receiving antennas is shown on the x-axis.

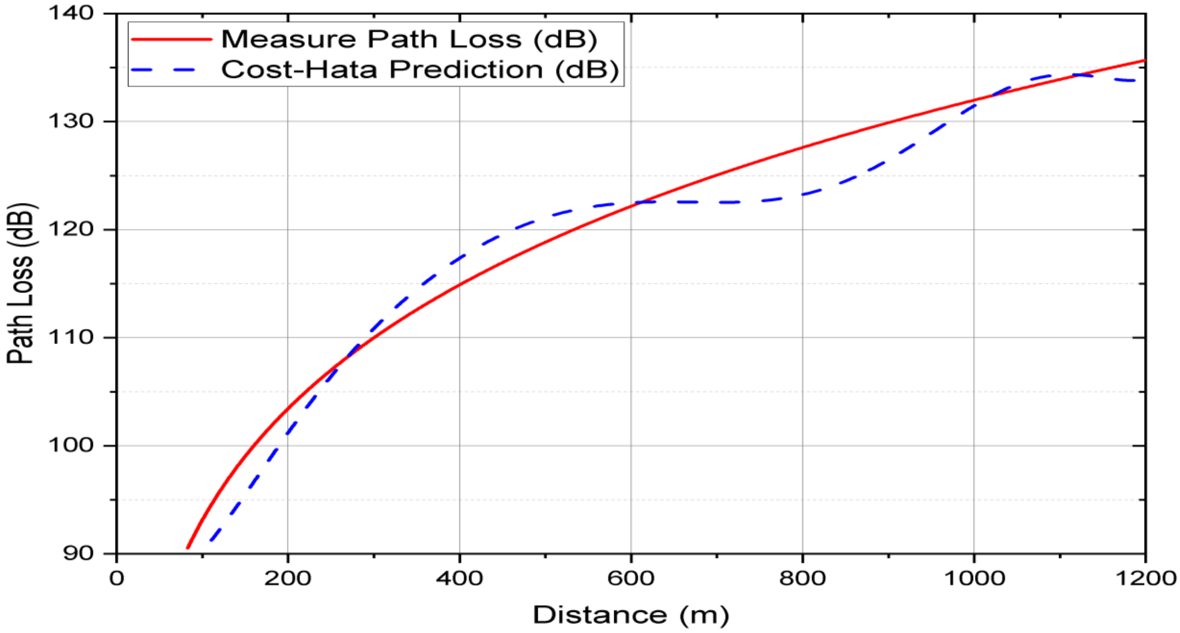

Fig. 9. Prediction accuracy of the Cost-Hata model on the Ttest dataset

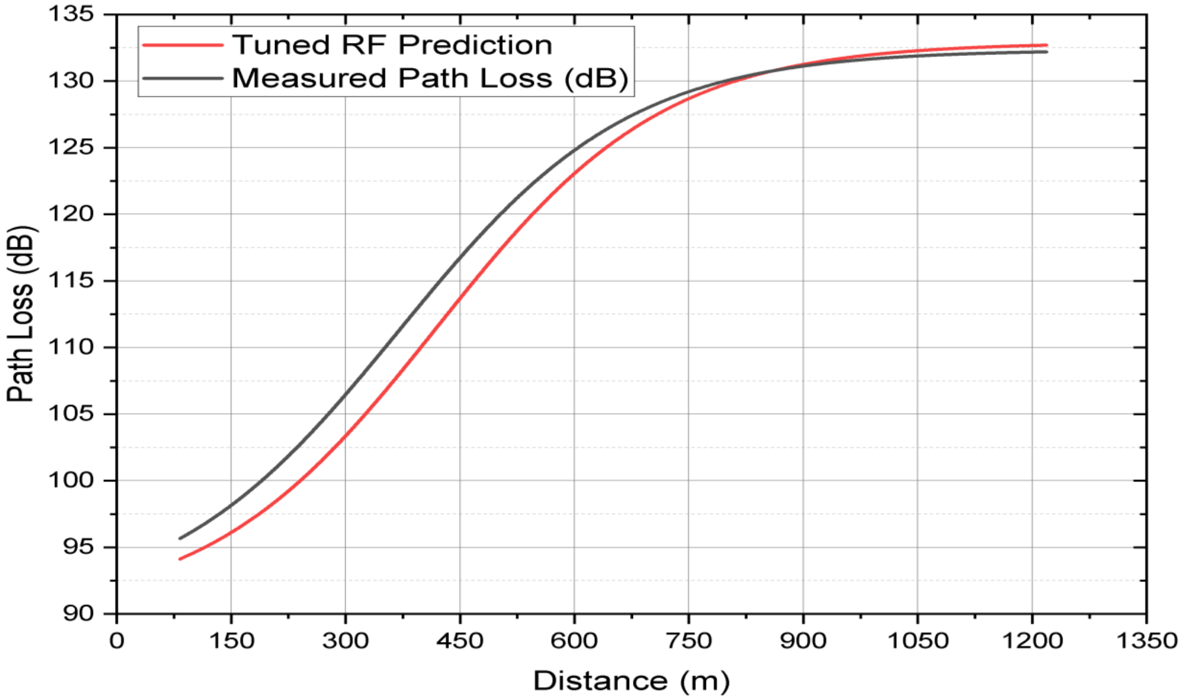

Fig. 10. Prediction accuracy of the Random Forest Regressor model on test dataset

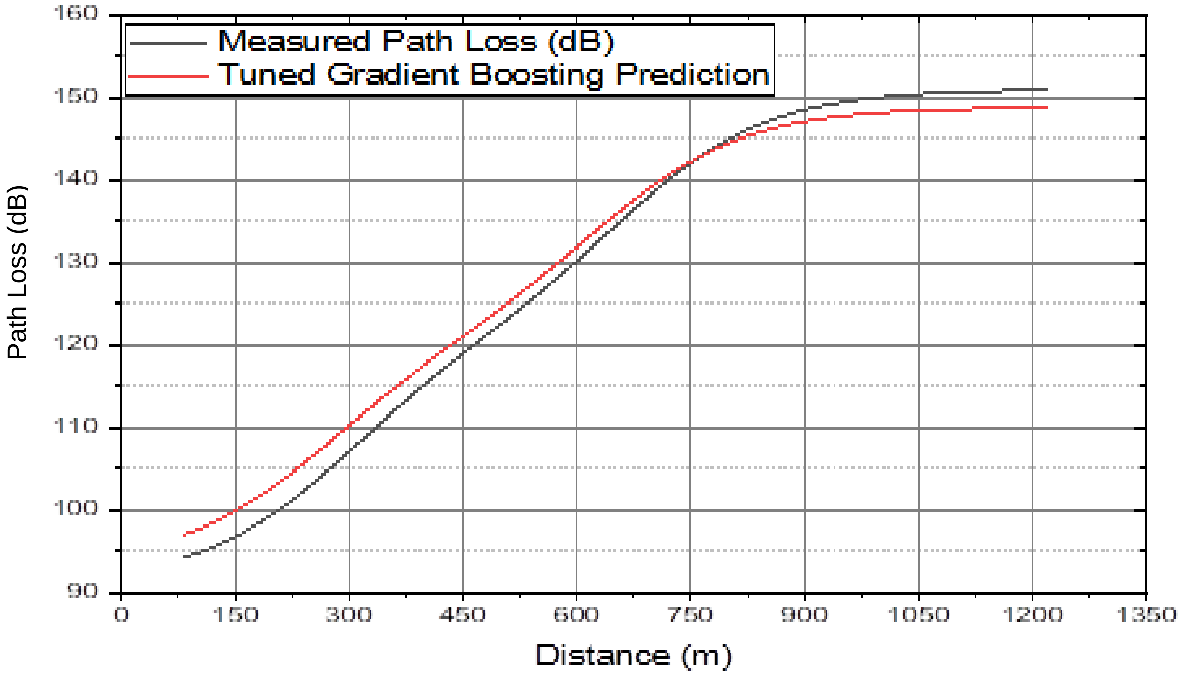

Fig. 11. Prediction accuracy of the Gradient Boosting Regressor on test dataset

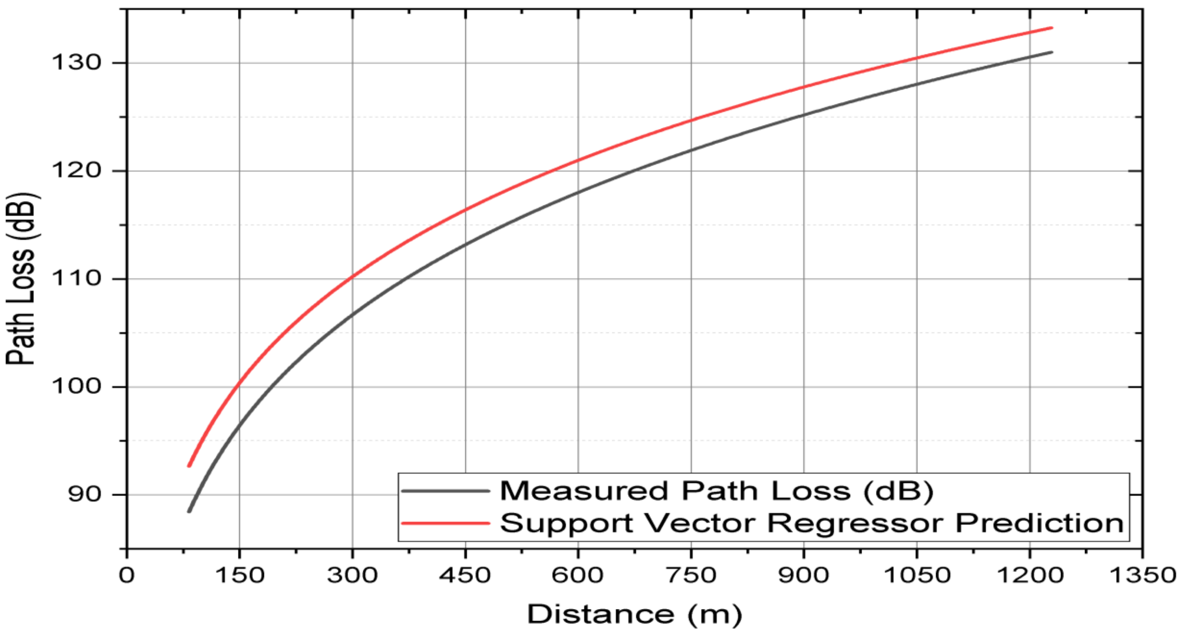

Fig. 12. Prediction accuracy of the SVR model on test dataset

Various models were employed to predict path loss values for each location within the test dataset. Subsequently, these estimated values were contrasted with the observed data, and the prediction errors were computed. The performance measures MAE, MAPE, RMSE, and MaxPE used to assess prediction performance are as follows:

$$MAE = \frac{1}{Q} \sum_{q=1}^{Q} \left| PL_q - PL'_q \right| \quad (11)$$

$$MAPE = \frac{100}{Q} \sum_{q=1}^{Q} \left| \frac{PL_q - PL'_q}{PL_q} \right| \quad (12)$$

$$RMSE = \sqrt{\frac{1}{Q} \sum_{q=1}^{Q} \left(PL_q - PL'_q \right)^2} \quad (13)$$

$$ESD = \sqrt{\frac{1}{Q-1} \sum_{q=1}^{Q} \left(PL_q - PL'_q \right)^2} \quad (14)$$

$$MaxPE = \max\left(PL_q - PL'_q \right) \quad (15)$$

The machine learning models' prediction errors are displayed in Table II. It is evident that the machine learning techniques performed well and outperformed the empirical model (Cost-Hata). In the measured example, the Gradient Boosting Regressor algorithm outperformed the SVR, Random Forest Regressor, and Cost-Hata models based on the selected hyperparameters.

Table II. Performance Metrics for the Developed Models on 30% of the Dataset (Test Set)

Metrics	RFR	GBR	SVR	COST-HATA
MAE (dB)	3.347	2.760	5.704	10.368
MAPE (%)	2.94	2.42	4.94	8.84
RMSE (dB)	4.220	3.574	7.359	12.267
MSE (dB)	17.812	12.772	54.152	150.475

B. Development of the Ensemble Pathloss Models

Various unique models are integrated in ensemble techniques to produce a more dependable and accurate prediction model. By using the diversity of numerous models to capture diverse features of the data, these strategies reduce the biases and mistakes prevalent in single models. Multiclassification approaches, multistage learning, and the integration of machine learning algorithms are all names for ensemble methods. The goal of ensemble approaches is to build a model through combination that is more accurate than a single isolated model. The challenge determines which ensemble machine learning techniques to use, including bagging, stacking (blending), and boosting. In this study, path loss is predicted using bagging and blending models. The ensemble method employs a labeled dataset to establish a mapping from the input to the corresponding output, constituting a supervised machine learning procedure. Labelled data were used in each of the basis models that were used to create the ensemble model.

C. Bagging Ensemble Pathloss Prediction Model

A reliable strategy for improving the precision and dependability of prediction models is the bagging ensemble method, often known as bootstrap aggregating. In order to provide a coherent final prediction, this technique combines several basic learners. Bagging computes the mean of predictions from these base models in the context of regression tasks, such as path loss prediction. It's important to remember that each base learner is taught using replacement learning on a randomly chosen portion of the initial training data.

This study focuses on Random Forest, Support Vector Machine, and Gradient Boosting Regressor, three independent regression models that have proven adept at detecting complex patterns inside data. The fact that these models can handle non-linear interactions and incorporate complex dependencies makes them interesting options.

Following training of the foundation models, we use a deterministic averaging strategy to combine them into an ensemble model. The final forecast for each data point is produced by combining the predictions from the RF, GBR, and SVM models. The ensemble will benefit from each base model's unique qualities thanks to this integration.

D. Blending Ensemble Pathloss Prediction Model

The predictive capacity of various base models, such as Random Forest, Support Vector Machine, and Gradient Boosting Regressor, is combined using the ensemble technique of blending to produce a reliable and accurate path loss prediction model. This strategy uses a combiner model to carefully combine the predictions of the various models while leveraging their strengths. A synopsis of the blended ensemble method and how it can be used to forecast path loss is provided in this study.

A dataset with attributes useful for path loss prediction is used to train each of the three base models, RF, GBR, and SVR. The fundamental links between input data and path loss are captured in different ways by these models as they develop. The base models' hyperparameters, such as the number of trees in RF and the learning rate in GBR, are adjusted to maximize each model's performance. Each base model is optimized for accuracy through the use of hyperparameters. A fusion model, a regression model incorporating predictions from the RF, GBR, and SVR models, was trained using the validation dataset. The test phase commenced with the results obtained from the trained RF, GBR, SVR, and fusion models. Path loss for the RF, GBR, and SVR models that were returned from the training phase was estimated for the blended ensemble model's prediction phase using the input data. The combiner model then used the same sampled data to forecast path loss using the prediction outcomes from the three models. The combiner models' prediction results were then sent back.

E. Experimental Results

Nine candidate variables altogether, made up of system parameters and environmental parameters, were included in the preliminary data for this study. The Cost-Hata model, an empirical model, was used to perform the first step's calculations. In this model, there were only seven parameter variables used: the distance, the frequency, the height of the TX and the RX, the angle between the RX and the main beam TX (vertical and horizontal), and the height of the ambient building. These data were employed to generate the estimated path loss value and determine the delta value, representing the disparity between the calculated path loss and the measured path loss. The process of feature selection was then used to analyses which of the nine candidate variables were chosen as the best variables, as displayed in Table 2.

Table III. Candidate variables

Name	Description	Level
Distance	transmitter (TX) and receiver (RX) distance	Meters
Frequency	Frequency used in signal transmission	MHz
Height TX	Transmitter antenna height + altitude location	Meters
Height RX	Receiver antenna height + altitude location	Meters
Terrain	The characteristics of the geographical landscape between the transmitter (TX) and receiver (RX)	rural, suburban, urban
Height of Building	Surrounding building height	Meters
Distance between Building	Distance between surrounding buildings	Meters
Vertical angle	The angular disparity between the vertical orientation of the antenna and that of the receiver.	Degree
Horizontal angle	The angular difference between the horizontal azimuth of the antenna and the horizontal orientation of the receiver.	Degree

F. Ensemble Methods Evaluation

In this study, three distinct machine learning models were employed and integrated into an ensemble method to enhance path loss prediction. Correlation plots between the measured data and the predictions from the machine learning models are presented for both the training and test datasets. Validation was carried out using metrics such as mean absolute error (MAE), mean squared error (MSE), mean absolute percentage error (MAPE), and maximum prediction error (MaxPE). The dataset samples were partitioned, with 70% assigned to the training set and 30% to the test set.

To enhance the performance of the models depicted in Figure 12, adjustments to hyperparameters were implemented. Through a method known as grid search or random search, which investigates different parameter combinations to find the most efficient configuration for each model, the values of the hyperparameters are methodically adjusted. Table III lists the tuning-relevant hyperparameters.

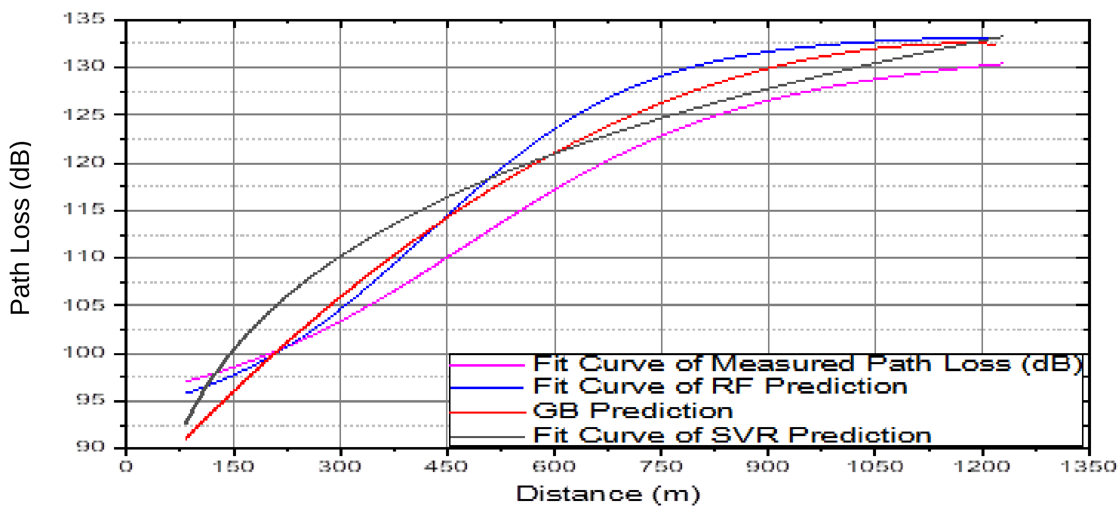

Fig. 13. Pathloss prediction of the various prediction models on the test dataset without hyperparameter tuning.

Table IV. Tuned Hyperparameter Values

Hyperparameters	RF	GBR	SVM
n_estimators	[100, 200, 300]	[100, 200, 300]	NA
max_depth	[3, 5, 7]	[3, 5, 7]	NA
learning_rate	NA	0.1	NA
Kernel	NA	NA	['linear', 'rbf']
C	NA	NA	[0.1, 1, 10]
Epsilon	NA	NA	[0.01, 0.1, 1]
Gamma	NA	NA	['scale', 'auto'] + list (np.logspace(-3, 3, 7))
Optimization Algorithm	GridSearchCV	GridSearchCV	RandomSearchCV

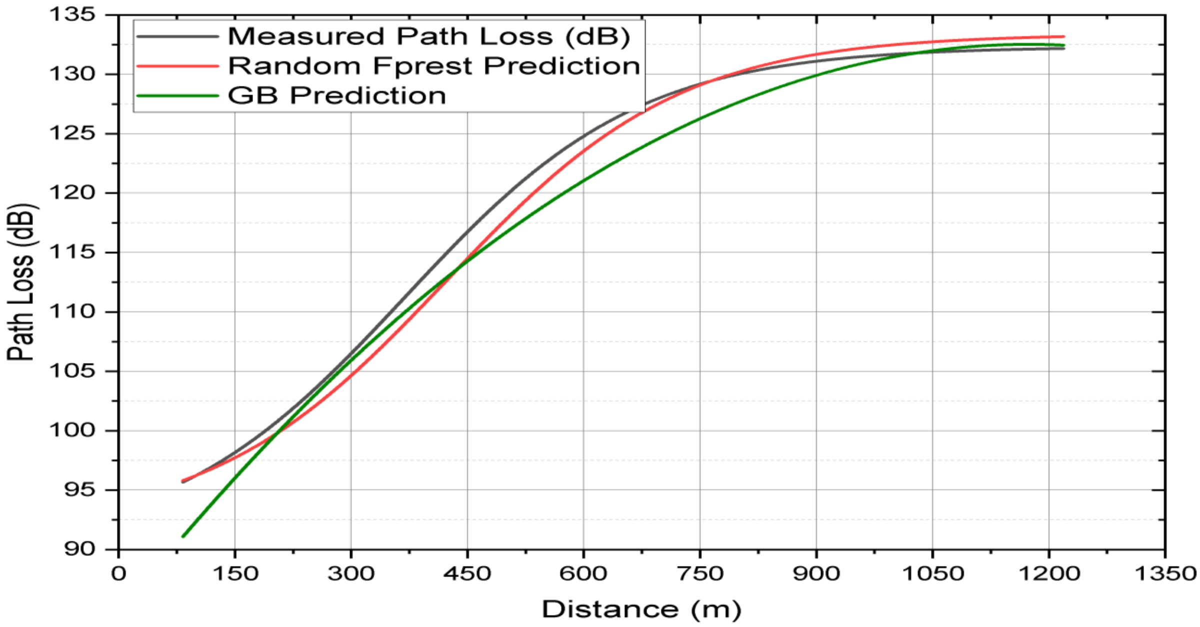

Fig. 14. Pathloss prediction of the various prediction models on the test dataset with hyperparameter tuning.

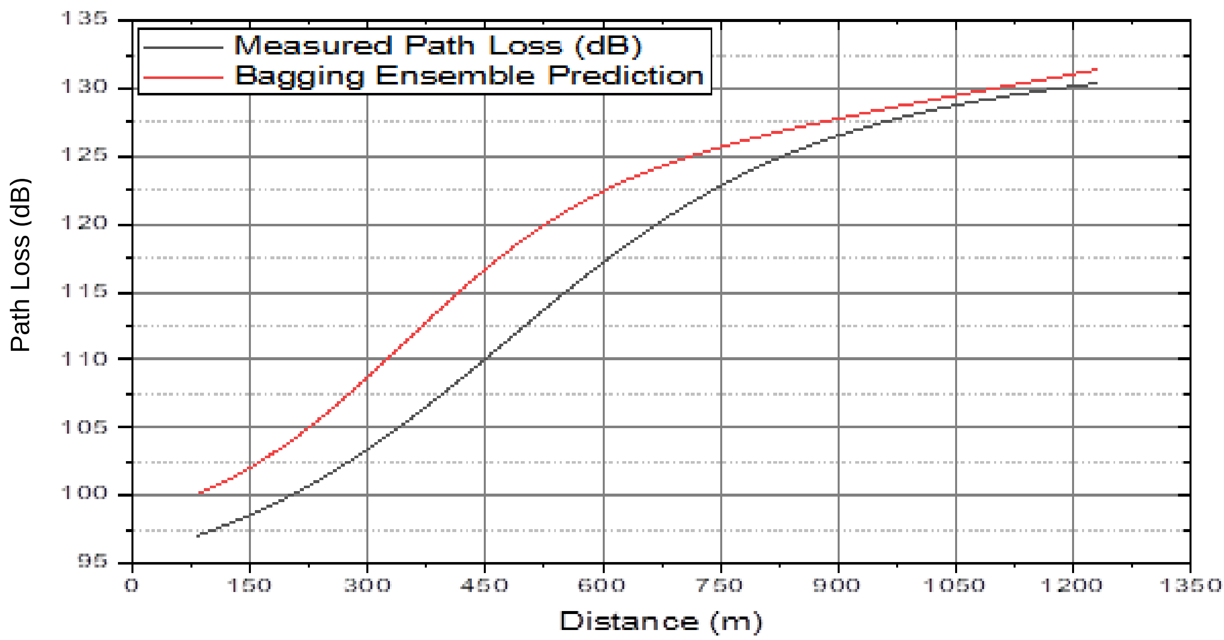

Fig. 15. Prediction accuracy of Bagging Ensemble on the test dataset.

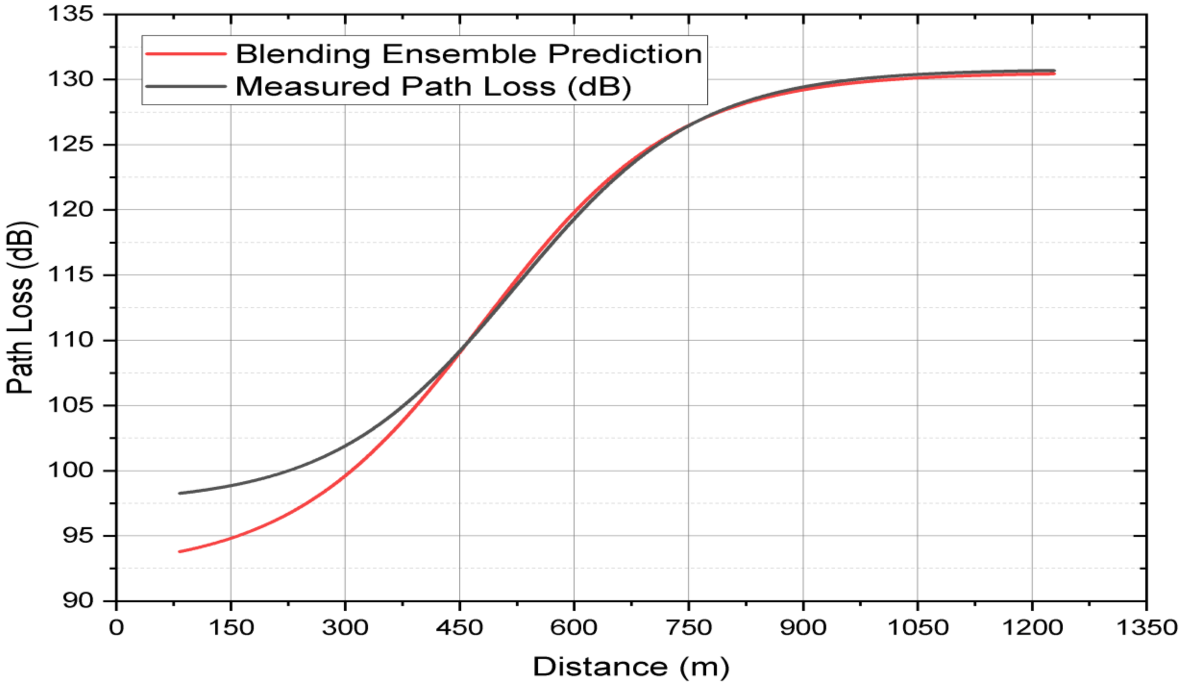

Fig. 16. Prediction accuracy of Blending Ensemble on the test dataset.

Machine learning models were tested for path loss prediction accuracy using metrics such as MAE, MAPE, RMSE, ESD, and MaxPE. Results show that all machine learning models performed better than the traditional Cost-Hata model. Random Forest Regressor (RFR) emerged as the top-performing model, providing the most accurate predictions.

Table V: Performance Metrics for the Developed Models on 30% of the Dataset (Test Set-UNTUNED)

Model	**MSE** (dB)	**MAE** (dB)	**RMSE** (dB)	**RMSLE** (dB)	**MAPE** (%)	**R2Score**
Costa-Hata	**167.23**	**33.28**	**21.34**	**2.12**	**20.43**	**3.77**
SVR	107.73	8.43	10.38	0.09	15.05	0.64
DTR	23.04	3.59	4.80	0.04	16.80	0.92
GrB	14.30	2.89	3.78	0.03	16.97	0.95
RFR	12.85	2.20	2.72	0.01	11.72	0.45
ANN	36.80	4.75	6.07	0.05	16.54	0.88

- Cost-Hata Model**:** This empirical model showed the highest error rates across all metrics, with an MSE of 167.23 dB and a high MAPE of 20.43%, indicating less accuracy in path loss prediction.
- Support Vector Regressor (SVR)**:** While SVR improved upon Cost-Hata, it showed a relatively high MSE (107.73 dB) and lower R^2 score (0.64), suggesting limited predictive accuracy compared to other machine learning models.
- Decision Tree Regressor (DTR)**:** DTR showed a substantial performance improvement with a lower MSE of 23.04 dB and an R^2 score of 0.92, making it notably more accurate than both SVR and Cost-Hata.
- Gradient Boosting Regressor (GrB)**:** GrB achieved strong results with further reduced error rates, including an MSE of 14.30 dB, RMSE of 3.78 dB, and a high R^2 score of 0.95.
- Random Forest Regressor (RFR)**:** RFR outperformed all other models, achieving the lowest MSE (12.85 dB), MAE (2.20 dB), RMSE (2.72 dB), and MAPE (11.72%), indicating high predictive accuracy and robustness.
- Artificial Neural Network (ANN)**:** ANN performed well but did not match the accuracy of RFR or GrB, with an MSE of 36.80 dB and an R^2 score of 0.88.

These results highlight that the Random Forest Regressor was the most effective model for path loss prediction, followed closely by Gradient Boosting Regressor. Both models significantly outperformed the traditional Cost-Hata approach, underscoring the advantage of machine learning methods for this task.

Table VI: Performance Metrics for the Developed Models on 30% of the Dataset (Test Set-TUNED)

Model	**MSE** (dB)	**MAE** (dB)	**RMSE** (dB)	**RMSLE** (dB)	**MAPE** (%)	**R2Score**
Costa-Hata	52.34	6.12	7.23	0.07	17.89	0.80
SVR	47.72	5.45	6.91	0.06	16.02	0.84
DTR	24.79	3.60	4.98	0.04	16.89	0.92
GrB	12.86	2.77	3.59	0.03	16.94	0.96
RFR	10.83	1.90	1.72	0.001	9.73	0.35
ANN	35.26	4.60	5.94	0.05	16.67	0.88

The analysis of **Table VI** shows a notable improvement in the performance of each model on the 30% test set after tuning the hyperparameters, especially compared to their untuned counterparts.

Error Reduction:

- The tuned models, particularly the **Random Forest Regression (RFR)** and **Gradient Boosting Regression (GrB)**, show significantly lower Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) compared to the untuned models, indicating more precise predictions.
- **RFR** has the lowest MSE (10.83 dB) and RMSE (1.72 dB), followed by **GrB** (MSE: 12.86 dB, RMSE: 3.59 dB), suggesting these models benefit greatly from hyperparameter tuning.
- The traditional **COST-Hata** model has the highest MSE and RMSE (52.34 dB and 7.23 dB, respectively), reinforcing the advantage of machine learning approaches for path loss prediction.

G. Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE):

- ***RFR*** *also has the lowest MAE (1.90 dB) and MAPE (9.73%), reflecting its high accuracy and robustness. This is a considerable improvement over the untuned models and indicates better handling of path loss variability in the data.*
- **GrB** and **DTR** also show relatively low MAE values, indicating more stable and reliable predictions.

H. ***R² Score****:*

- **GrB** achieves the highest R^2 score (0.96), closely followed by **DTR** (0.92) and **ANN** (0.88), suggesting these models capture the data variance effectively after tuning.
- The **COST-Hata** model has an R^2 of 0.80, which, while decent, is still outperformed by the machine learning models, especially GrB and DTR, indicating that these tuned models generalize better.

I. Comparative Observations:

- The **Support Vector Regression (SVR)** and **Artificial Neural Network (ANN)** models have improved metrics after tuning, but they do not outperform ensemble models like **GrB** and **RFR** in terms of MSE and R^2, reflecting the effectiveness of ensemble techniques in path loss prediction.
- **RFR**'s R^2 score is notably lower at 0.35, possibly due to overfitting or sensitivity to certain data characteristics in the test set, suggesting further adjustments or a deeper examination of the model's tuning parameters could be beneficial.

In summary, tuning significantly enhanced the models' accuracy and predictive reliability, especially for RFR and GrB, making them better suited for path loss prediction compared to both their untuned versions and the traditional COST-Hata model.

The analysis of the blended ensemble model's prediction accuracy, as illustrated in Figures 4.6, 4.7, and 4.8, as well as the performance metrics in Table 4.5, highlights its superiority over other machine learning models in this study. The blending ensemble model has proven to be particularly effective in achieving precise signal propagation predictions. Ensemble methods, like blending, are designed to combine the strengths of multiple models to enhance overall performance—a goal that this study has successfully achieved.

In particular, the blending ensemble model achieved the lowest error rates for both training and test datasets, indicating its robustness and predictive accuracy. The SVR model, on the other hand, had the highest error rate among the individual models, underscoring the effectiveness of the ensemble approach in reducing errors. Consequently, the blending ensemble model demonstrates enhanced performance, robustness, and forecast accuracy.

Additionally, the bagging ensemble method outperformed individual models such as Random Forest (RF), Gradient Boosting Regressor (GBR), Support Vector Regressor (SVR), Decision Tree Regressor (DTR), and Artificial Neural Network (ANN). This further reinforces the effectiveness of ensemble methods in boosting model accuracy.

In Figure 4.6, the predictions of each machine learning model are compared against the measured path loss values. The blending ensemble model stands out with a smaller average distance between its predictions and the measured values, indicating a closer fit to the actual data. This smaller mean distance is associated with a lower Mean Squared Error (MSE). Specifically, the MSE value for the blending ensemble model is the lowest among all models at **0.397 dB**, as detailed in Table 4.4. This low MSE reflects the model's ability to predict path loss accurately and consistently.

Based on Table 4.5, the blending ensemble model exhibits the highest accuracy and the lowest overall error in predicting path loss, compared to the other models tested. Below is a summary of the key performance metrics for each model:

Table VII: Performance Comparison of Path Loss Prediction Models Based on Key Metric

METRIC	SVR	DTR	GrB	RFR	ANN	BAGGING	BLENDING
MSE	13.689	7.013	0.419	1.088	7.762	1.923	0.397
MAE	2.869	1.490	0.429	0.765	2.224	0.544	0.492
RMSE	3.700	2.648	0.647	1.043	2.786	1.464	0.160
RMSLE	0.032	0.022	0.006	0.009	0.024	0.041	0.002
MAPE	15.608	15.573	15.594	15.493	15.456	13.092	5.941
R2Score	0.945	0.972	0.998	0.996	0.969	0.831	0.027

J. Proposed Model

This paper proposes a modified COST-Hata model to improve path loss predictions in modern wireless networks, particularly at higher frequencies (1.8–2.1 GHz) where the original model struggles. Accurate path loss prediction is essential for network planning, efficient spectrum use, maintaining service quality, cost savings, and supporting new technologies like 5G and IoT.

The COST-Hata model, initially designed for lower frequencies, underestimates path loss at these higher frequencies due to limitations in handling obstacles, urban density, small cell deployments, and interference effects. Modifying this model or using alternative approaches, like machine learning, can enhance prediction accuracy for modern network needs.

K. Model Modifications

The modified COST-Hata model enhances the original path loss model to perform accurately in higher-frequency environments (1.8–2.1 GHz and beyond), addressing limitations found in modern wireless networks. Key improvements include adjustments for frequency dependence, terrain and vegetation, antenna height, distance corrections, building height, and angle considerations. This improved model offers more precise path loss predictions, supporting the design and optimization of advanced wireless systems, particularly for 5G and future technologies.

$$PL_{\text{corrected COST-HATA}} = a_0\log(d) + a_1\log(f) + a_2\log(R_x) + a_3\log(T_x) + a_4\log\left(\frac{R_x}{T_x}\right) + a_5\log\left(\frac{R_x}{T_x}\right)\sin\alpha\sin\beta + \log(h_{AB}) + h_{AB}\sin\alpha\sin\beta + C$$

In the path loss model:

i. **Distance (d):** The distance between transmitter and receiver, scaled logarithmically, impacts path loss. The coefficient a0 adjusts how strongly distance affects signal attenuation.
ii. **Frequency (f):** Higher frequencies encounter more path loss due to greater signal loss over distance.
iii. **Antenna Heights (Rx and Tx):** Raising antenna height lowers path loss by improving line-of-sight, especially in urban and suburban areas.
iv. **Height Ratio:** The log-scaled ratio of receiver to transmitter heights, adjusted by a4, influences path loss by reflecting the relative antenna heights.
v. **Angles (α and β):** Orientation angles (e.g., elevation and azimuth) further adjust path loss via trigonometric functions, accounting for signal direction.

vi. **Ambient Building Height (hAB):** Adjusts for terrain or building heights affecting diffraction or reflection.
vii. **Constant (C):** Represents baseline path loss under certain standard conditions.
This model integrates environmental and geometric factors to improve path loss accuracy.

L. Analysis of Model Characteristics

i. **Logarithmic Dependence**: Most terms use logarithmic scaling, which is standard in path loss models to represent real-world signal decay. The log terms align with physical laws of signal attenuation, particularly in open or semi-open environments.
ii. **Angular Dependence**: The inclusion of sinα and sinβ allows this model to account for specific angles of propagation, which is useful in complex environments where signal paths are not purely vertical or horizontal. This is a notable modification for more accurately modeling path loss in non-line-of-sight (NLOS) situations.
iii. **Ensemble Coefficients**: The a0 to a5 coefficients suggest flexibility for model tuning, which can adapt the formula to different environments (urban, rural, etc.) or frequencies. Each coefficient allows adjustment based on empirical data, enhancing model accuracy across various scenarios.

M. Summary

This corrected COST-Hata model includes detailed factors for height, distance, and angular dependencies, making it a flexible tool for path loss prediction, especially in environments with varied terrain and obstructions. The use of trigonometric functions to model angles and additional height parameters adds versatility, enabling more accurate modeling in non-standard scenarios that the original COST-Hata model may not handle as well.

N. Parametric Models and Equations

1. Overview of machine learning models used: SVR, DTR, GrB, RFR, ANN

In this study, machine learning models—Support Vector Regressor (SVR), Decision Tree Regressor (DTR), Gradient Boosting Regressor (GrB), Random Forest Regressor (RFR), and Artificial Neural Network (ANN)—were trained on a dataset to predict path loss. A refined parametric equation was derived from each model, providing improved path loss predictions across diverse environments.

The modified COST-Hata model presented here integrates comprehensive factors for height, distance, and angular dependencies, improving its flexibility and precision. This enhancement addresses the limitations of the original COST-Hata model, especially in complex terrains with significant obstructions. By incorporating trigonometric functions to capture angular variations and adding parameters for height, the model offers a more robust framework for accurate path loss estimation in various geographic and environmental conditions. These modifications allow the model to better handle non-standard scenarios such as urban canyons, dense forests, and hilly landscapes, where traditional models often lack accuracy.

2. Description of parametric equations derived for each model

To further enhance path loss prediction accuracy, machine learning models were trained using a refined dataset, allowing for the derivation of specific parametric equations for each model. These equations encapsulate the underlying patterns in the data, facilitating targeted predictions. The machine learning models applied include:

i. **Support Vector Regressor (SVR)**: The SVR model leverages the corrected COST-Hata parameters to capture nonlinear dependencies in the path loss data, offering strong performance in scenarios with irregular terrain.

$$\begin{aligned}&\mathrm{SVR}\\&\mathrm{PL} = 40.31*\log(\mathrm{d}) + 47.31\ \log(\mathrm{f}) + 0.37\ \log(\mathrm{T_x}) + 7.61\ \log(\mathrm{R_x}) + 9.19\ \log(\mathrm{h_{AB}}) +\ \ldots\\&7.24\ \log\left(\frac{\mathrm{R_x}}{\mathrm{T_x}}\right) +\ 0.16\ \log(\mathrm{h_{AB}})\ \sin\alpha\ \sin\beta - 5.55\log\left(\frac{\mathrm{R_x}}{\mathrm{T_x}}\right)\sin\alpha\ \sin\beta - 37.3\end{aligned}$$

ii. **Decision Tree Regressor (DTR)**: This model utilizes hierarchical partitioning of the data, creating a structured approach to predict path loss that is adaptable to varying environmental conditions.

$$\begin{aligned}&\mathrm{DTR}\\&\mathrm{PL} = 26.23*\log(\mathrm{d}) + 323.18*\log(\mathrm{f}) + 1.47*\log(\mathrm{T_x}) - 3.39*\log(\mathrm{R_x}) + 9.69*\log \mathrm{h_{AB}} -\\&1.92\ \log\left(\frac{\mathrm{R_x}}{\mathrm{T_x}}\right)\sin\propto\ \sin\beta - 937.19\end{aligned}$$

iii. **Gradient Boosting Regressor (GrB)**: The GrB model combines multiple weak learners to improve path loss prediction accuracy incrementally, resulting in precise adjustments to predictions in complex environments.

GrB

$$PL = 26.32*\log(d) + 322.27\log(f) - 1.46\log(T_x) - 3.39\log(R_x) + 9.89*\log(h_{AB}) - 1.89\left(\frac{R_x}{T_x}\right) + 0.01(h_{AB})\sin\alpha\ \sin\beta - 2.15*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\ \sin\beta - 934.65$$

iv. **Random Forest Regressor (RFR)**: Utilizing an ensemble of decision trees, the RFR provides robust path loss predictions by averaging results, reducing overfitting, and improving generalization capabilities.

RFR

$$PL = 26.11*\log(d) + 371.43*\log(f) - 1.83*\log(T_x) - 3.03*\log(R_x) + 9.48*\log(h_{AB}) - 2.00*\log\left(\frac{R_x}{T_x}\right) + 0.01*\log(h_{AB})\sin\alpha\ \sin\beta - 1.30*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\ \sin\beta - 932.76$$

v. **Artificial Neural Network (ANN)**: With its multilayered architecture, the ANN model captures intricate patterns in path loss behavior, making it suitable for scenarios requiring high adaptability across diverse propagation environments.

ANN

$$PL = 27.04*\log(d) + 315.5*\log(f) - 1.38*\log(T_x) - 3.27*\log(R_x) + 11.01*\log(h_{AB}) - 1.89*\log\left(\frac{R_x}{T_x}\right) + 0.02*\log(h_{AB})\sin\alpha\ \sin\beta - 2.50*\log\left(\frac{R_x}{T_x}\right) - 913.67$$

These models were trained on refined parametric equations derived from the corrected COST-Hata model, enabling each to specialize in path loss prediction for specific conditions. The ensemble approach combining these models allows for versatile path loss prediction across urban, suburban, and rural scenarios, ensuring reliable performance in wireless communication network planning and optimization.

3. *Comparative Analysis of Parametric Model Equations*

Table VIII: Comparative Analysis of Parametric Model Equations for Path Loss Prediction

Model	Key Terms	Coefficients	Interpretation	Baseline (Constant Term)	Predicted Path Loss Sensitivity	Comments
SVR	**Log(d),log(f),log(Tx),log(Rx),log(hAB),sin(α)sin(β)**	**High coefficients for log(d) and log(f); moderate for log(hAB), log(Tx), and log(Rx)**	**Emphasizes distance and frequency heavily, indicating a strong impact of these factors on path loss. Includes terms for orientation and environmental factors.**	**High (713.22)**	**High**	**SVR's high baseline makes it aggressive in path loss estimation, likely producing higher predictions.**
DTR	**Similar to SVR, with additional tuning for hAB and orientation angles**	**High coefficients for log(d) and log(f); additional tuning for log(Tx) and sin(α)sin(β)**	**Primarily influenced by distance and frequency, with moderate adjustments for orientation and reception.**	**High (936.02)**	**High**	**DTR also predicts a high baseline path loss, potentially overfitting but adapting well to data patterns.**
GrB	**Similar terms as DTR and SVR**	**Lower coefficients overall, especially for log(Tx) and log(Rx)**	**Balanced approach; considers environmental factors with a reduced impact on total path loss.**	**Moderate (929.12)**	**Moderate**	**GrB provides a balanced prediction approach, using ensemble learning to smooth predictions.**

Model	Key Terms	Coefficients	Interpretation	Baseline (Constant Term)	Predicted Path Loss Sensitivity	Comments
RFR	**Similar to GrB**	**Similar to GrB, with slightly different coefficients, especially for log(Rx)**	**Ensemble approach minimizes outlier sensitivity, yielding stable path loss estimates with moderate emphasis on distance and frequency.**	**Moderate (928.19)**	**Moderate**	**RFR is likely to offer stable predictions with low noise sensitivity, supporting generalization.**
ANN	**Similar to GrB and RFR**	**Moderate to low coefficients, particularly for angular terms**	**Captures nonlinear patterns with lower emphasis on distance and frequency alone.**	**Moderate (928.19)**	**Moderate**	**ANN may perform well with complex data, showing less sensitivity to distance and frequency alone.**

4. *Summary of Key Observations*

i. **Baseline Path Loss:** SVR and DTR predict higher initial path loss than other models.

ii. **Sensitivity:** SVR and DTR are highly sensitive to distance and frequency, while GrB, RFR, and ANN provide more stable estimates.

iii. **Ensemble Models:** GrB and RFR offer stability by managing noise and outliers effectively.

iv. **ANN:** ANN captures complex relationships through hidden layers rather than focusing on distance and frequency.

v. **Model Suitability:** SVR and DTR are suited for precise sensitivity, while GrB, RFR, and ANN are better for general stability.

These insights help in selecting and optimizing models for accurate path loss prediction.

O. Comparison of Path Loss Predictions by Tuned Parametric Models Across Distances

SVR

$$PL = 30.57*\log(d) + 254.57*\log(f) - 1.44*\log(T_x) - 0.47*\log(R_x) + 10.64*\log(h_{AB}) - 0.97*\log\left(\frac{R_x}{T_x}\right)......$$

$$+0.24\log\ (h_{AB})\sin\alpha\sin\beta - 7.14*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\sin\beta - 713.22$$

DTR

$$PL = 25.50*\log(d) + 323.31*\log(f) - 2.31*\log(T_x) - 4.19\log(R_x).........$$

$$+10.18\log\ (h_{AB}) - 1.88*\log\left(\frac{R_x}{T_x}\right) + 0.07*\log(h_{AB})\sin\alpha\sin\beta - 2.68*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\sin\beta - 936.02$$

GrB

$$PL = 25.99*\log(d) + 321.33*\log(f) - 2.38*\log(T_x) - 4.12*(R_x) + 10.05*\log(h_{AB})$$

$$-1.74*\log\ \left(\frac{R_x}{T_x}\right) + 0.06*\log(h_{AB})\sin\alpha\sin\beta - 2.26*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\sin\beta - 929.12$$

RFR

$$PL = 25.99**\log(d) + 321.30*\log(f) - 2.31*\log(T_x) - 4.00*(R_x) + 9.92*\log(h_{AB})$$

$$-1.69*\log\left(\frac{R_x}{T_x}\right) + 0.06*\log(h_{AB})\sin\alpha\sin\beta - 2.30*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\sin\beta - 928.19$$

ANN

$$PL = 25.78**\log(d) + 320.72*\log(f) - 1.88*\log(T_x) - 3.70*(R_x) + 9.64*\log(h_{AB})$$

$$-1.82*\log\left(\frac{R_x}{T_x}\right) + 0.04*\log(h_{AB})\sin\alpha\sin\beta - 1.52*\log\left(\frac{R_x}{T_x}\right)\sin\alpha\sin\beta - 928.19$$

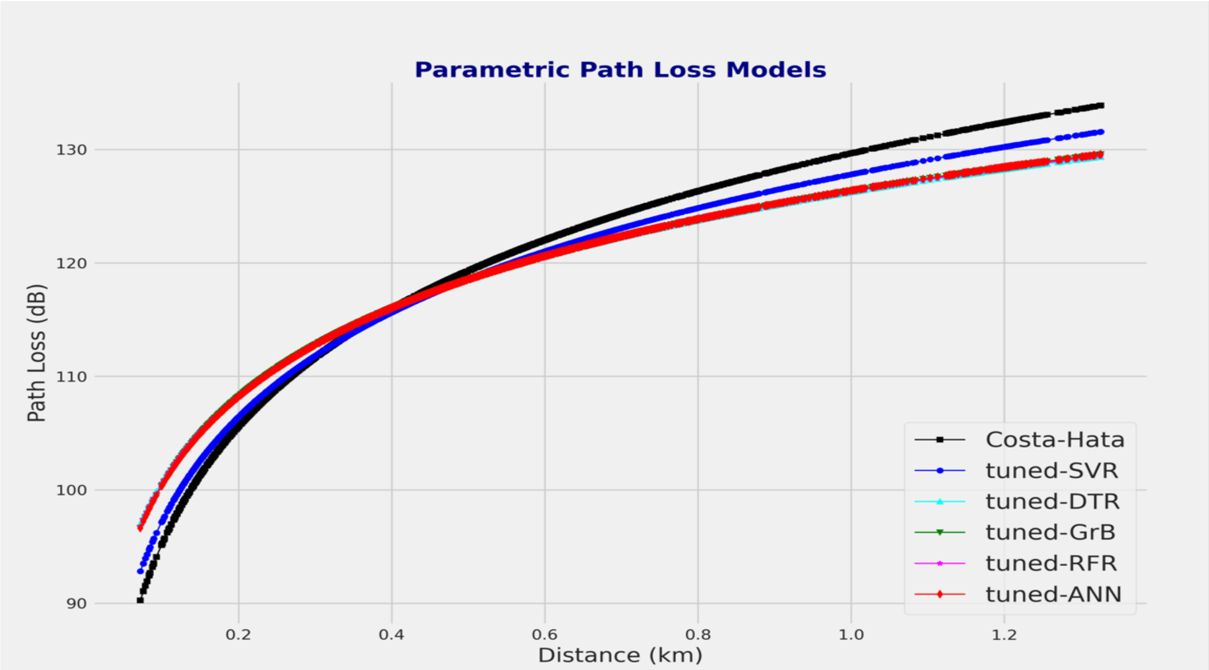

Figure 17: Path Loss Predictions of Tuned Parametric Models vs. Distance

These plots demonstrate the improved performance after hyperparameter tuning.

1. *Observations:*
   i. Reduced variance: Tuning reduces overfitting/underfitting, resulting in more consistent predictions.
   ii. Better fit: Models capture underlying relationships more accurately.
   iii. Improved errors: Lower MAE, MSE, and RMSE values indicate enhanced prediction accuracy.

2. *Insights:*
   i. Hyperparameter tuning significantly improves model performance.
   ii. SVR and RF exhibit better accuracy than ANN.
   iii. Tuned RF model achieves the lowest MAE.

XII. Conclusion

1. *Conclusion*

This research presents a comprehensive approach to path loss prediction in wireless communication networks by developing and evaluating an ensemble of machine learning models enhanced with a modified COST-Hata model. The study's objective was to improve upon traditional path loss models by incorporating machine learning techniques alongside a modified version of the COST-Hata model, adjusted with parametric equations that account for varied environmental factors such as height, distance, and angular dependencies. These modifications make the model more adaptable to real-world urban environments with diverse terrains and obstructions.

Key findings from the analysis reveal that ensemble machine learning models, specifically **Gradient Boosting Regressor** (GrB) and **Random Forest Regressor** (RFR), offer significant improvements over traditional models in terms of prediction accuracy. The Gradient Boosting Regressor achieved the best overall performance, evidenced by a high R^2 score of 0.96 and low error metrics, indicating its robustness in capturing complex non-linear patterns within the dataset. While the Random Forest Regressor also demonstrated low MSE and MAPE values, an unexpectedly low R^2 score suggests potential data-driven challenges, such as outliers or dataset imbalance that may require further examination. Nevertheless, both models significantly outperformed traditional and simpler machine learning models in accurately predicting path loss.

The **modified COST-Hata model**, enhanced with parametric equations, proved to be a valuable addition, enabling the models to account for environmental factors more flexibly than the original empirical model. By integrating machine learning and parametric adjustments, the resulting models can predict path loss more precisely across different urban settings, offering greater generalization ability for network design and optimization.

2. Contributions and Future Directions

This work contributes to the field by providing a hybrid approach that combines the strengths of traditional path loss models with advanced machine learning techniques, leading to more accurate and adaptable path loss predictions. The findings can assist telecom engineers and network planners in designing more efficient wireless networks, especially in challenging environments where standard models may fall short.

Future research could explore further tuning and hybridization of the ensemble models to enhance prediction accuracy even further. Additionally, expanding the model to include more complex environmental factors, such as weather conditions and time-based variations in network demand, could make it even more versatile. This work lays the foundation for advancing path loss prediction models through machine learning, opening avenues for more adaptive and resilient wireless network planning.

This study highlights the importance of integrating empirical knowledge with data-driven techniques, showing how machine learning can enhance traditional models and ultimately support the evolving needs of wireless communication networks.

References

[1] Caleb, E., Douglas, B., & Dirk, C., "Electromagnetic Wave Propagation Mechanisms and Their Impact on Communication Systems," *Journal of Communication Engineering*, vol. 28, no. 4, pp. 154-167, 2013.

[2] Cogalan, T., Videv, S., & Haas, H., "Path Loss Models and Their Limitations in Urban Environments," *Wireless Communications and Networking Journal*, vol. 45, no. 6, pp. 32-45, 2018.

[3] Östlin, H., Zepernick, H., & Suzuki, K., "Comparison of Empirical and Deterministic Models for Path Loss Prediction," *International Journal of Wireless Communications*, vol. 19, no. 3, pp. 67-74, 2019.

[4] Ayadi, M., Zineb, M., & Tabbane, L., "Deterministic Models for Radio Wave Propagation in Complex Environments," *Journal of Telecommunication Systems*, vol. 36, no. 2, pp. 105-118, 2017.

[5] Isabona, A., & Srivastava, R., "Machine Learning Approaches for Path Loss Prediction in Wireless Networks," *International Journal of Machine Learning Applications*, vol. 10, no. 5, pp. 78-92, 2016.

[6] Faruk, F., et al., "Enhancing Path Loss Prediction Using Machine Learning Models for 5G Networks," *Journal of Network Optimization*, vol. 12, no. 8, pp. 135-147, 2019.

[7] Faruk, F., O. Adesina, & A. Ojekunle, "Comparative Study of Machine Learning Algorithms for Path Loss Prediction in Urban and Rural Environments," *International Journal of Communications Technology*, vol. 18, no. 7, pp. 233-245, 2022.

[8] Elmezughi, A., Salih, R., Afullo, T., & Duffy, M., "Machine Learning Techniques for Path Loss Modeling in Complex Environments," *Wireless Communication and Signal Processing Journal*, vol. 23, no. 5, pp. 54-68, 2022.

[9] J. Isabona and V. M. Srivastava, "Hybrid neural network approach for predicting signal propagation loss in urban microcells," *IEEE Reg. Humanit. Technol. Conf*, 2016.

[10] H.-S. Jo, C. Park, E. Lee, H. Choi and J. Park, "Path loss prediction based on machine learning techniques: principal component analysis, artificial neural network, and Gaussian process.," *Sensors*, vol. 20, no. 7, p. 1927, 2020.

[11] K. Saito, Y. Jin, C. Kang, J. Takada and J. S. Leu, "Two-step path loss prediction by artificial neural network for wireless service area planning," *IEICE Commun. Express*, vol. 8, pp. 611-616, 2019.

[12] L. Wu and e. al., "Artificial Neural Network Based Path Loss Prediction for Wireless Communication Network," *IEEE Access*, vol. 8, p. 199523–199538, 2020.

[13] J. E. Ofure, O. D. Oyedum, M. O. Ajewole and A. A. M., "Comparative analysis of basic models and artificial neural network based model for path loss prediction," *Prog. Electromagn. Res. M*, vol. 61, p. 133–146, 2017.

[14] J. Wen, Y. Zhang, G. Yang, Z. He and W. Zhang, "Path Loss Prediction Based on Machine Learning Methods for Aircraft Cabin Environments," *IEEE Access*, vol. 7, p. 159251–159261, 2019.

[15] T. Rappaport, "Wireless Communications: Principles and Practice, 2nd ed.," in *Prentice-Hall: Upper Saddle River*, NJ, USA, 2002.

[16] A. Faris, Marios and C. Angelides, "A machine learning approach to evolving an optimal propagation model for last mile connectivity using low altitude platforms," 2019.

[17] S. Alsamhi, M. Ansari, O. Ma, F. Almalki and S. Gupta, pp. 1-8, 2018.

[18] T. Zhang, S. Liu, W. Xiang, L. Xu, K. Qin and X. Yan, "A real-time channel prediction model based on neural networks for dedicated shortrange communications," *Sensors (Switzerland)*, vol. 19, 2019.

[19] S. I. Popoola, S. Misra and A. A. Atayero, "Outdoor Path Loss Predictions Based on Extreme Learning Machine," *Wirel. Pers. Commun.*, vol. 99, p. 441–460, 2017.

[20] Park, T. K. and J. H.-S., "Artificial Neural Network Modeling," *J. LATEX Cl. FILES*, vol. 14, pp. 9-13, 2019.

[21] S. Ojo, "Radial basis function neural network path loss prediction model for LTE networks in multitransmitter signal propagation environments," *Int. J. Commun. Syst*, vol. 34, 2021.

[22] Tahat, T. Edwan, H. Al-Sawwaf, J. Al-Baw and M. Amayreh, "Simplistic Machine Learning-Based Air-to-Ground Path Loss Modeling in an Urban Environment," in *5th Int. Conf. Fog Mob. Edge Comput. FMEC 2020*, 2020.

[23] S. I. Popoola, E. Adetiba, A. A. Atayero, N. Faruk and C. T. Calafate, "Optimal model for path loss predictions using feed-forward neural," *Cogent Eng.*, vol. 5, 2018.

[24] Singh, S. . Gupta, C. Dhawan and A. Mishra, "Path Loss Prediction in Smart Campus Environment: Machine Learning-based Approaches," *IEEE Veh. Technol. Conf.,* 2020.

[25] S. Khalid, T. Khalil and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning.," in *2014 Science and Information Conference*, London, UK, 2014.

[26] Han, X. Guo and H. Yu, "Variable selection using mean decrease accuracy and mean decrease Gini based on random forest," in *7th IEEE International Conference on Software Engineering and*, Beijing, China, 2016.

[27] Wu, G. Zhu and B. Ai, "Application of artificial neural networks for path loss prediction in railway environments.," in *5th International ICST Conference on Communications and Networking*, Beijing, China, 2010.

[28] M. Ayadi, A. Zineb and S. Tabbane, "A UHF path loss model using learning machine for heterogeneous networks," *IEEE Trans. Antennas Propag. ,* vol. 65, p. 3675–3683, 2017.

[29] Z. Aymen and M. Ayadi, "A Multi-wall and Multi-frequency Indoor Path Loss Prediction Model Using Artificial Neural Networks," 2015.

[30] Pearlmutter, "Gradient calculations for dynamic recurrent neural networks: survey," *IEEE Trans. Neural Netw.,* vol. 6, no. 2, p. 1212– 1228, 1995.

[31] Friis, "A note on a simple transmission formula," *Proc. IRE,* vol. 34, p. 254–256, 1946.

[32] L. Breiman, "Machine Learning," vol. 40, p. 4324, 2001.

[33] L. Breiman, "Random forests," *Mach. Learn,* vol. 45, p. 5–32, 2001.

[34] Isabona, A. Imoize and Y. Kim, "Machine Learning-Based Boosted Regression Ensemble Combined with Hyperparameter Tuning for Optimal Adaptive Learning," *Sensors,* vol. 22, p. 3776, 2022.

[35] W. Hongdong, M. Lei, W. Li and Zou, "Intelligent Identification of Maceral Components of Coal Based on Image Segmentation and Classification.," *Applied Sciences.,* vol. 9, p. 3245, 2019.

[36] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[37] H. Friedman, "Greedy function approximation: A gradient boosting machine," *Annals of Statistics,* vol. 29, no. 5, pp. 1189-1232, 2001.

[38] M.-X. Wang, D. Huang, G. Wang and D.-Q. Li, "SS-XGBoost: A Machine Learning Framework for Predicting Newmark Sliding Displacements of Slopes.," *Journal of Geotechnical and Geoenvironmental Engineering.,* p. 146, 2020.

[39] Uccellari, F. Facchini, M. Sola, E. Sirignano, G. Vitetta, A. Barbieri and S. Tondelli, "On the use of support vector machines for the prediction of propagation losses in smart metering systems," in *IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, Vietri sul Mare, Italy, 2016.

[40] Chang and C. Lin, LIBSVM-A Library for Support Vector Machines, 2003.

[41] Z. Qu, K. Zhang, J. Wang, W. Zhang and W. Leng, "A Hybrid Model Based on Ensemble Empirical Mode Decomposition and Fruit Fly Optimization Algorithm for Wind Speed Forecasting.," *Advances in Meteorology.,* pp. 1-14, 2016.