



Improving millimetre-wave path loss estimation using automated hyperparameter-tuned stacking ensemble regression machine learning

Johnson O. Afape^a, Alexander A. Willoughby^a, Modupe E. Sanyaolu^a, Obiseye O. Obiyemi^b, Katileho Moloi^c, Janet O. Jooda^d, Oluropo F. Dairo^{a,e,*}

^a Department of Physical Sciences, Faculty of Natural Sciences, Redeemer's University, Ede, Osun State, Nigeria

^b Space Science Centre, Department of Electrical Power Engineering, Durban University of Technology, Durban, South Africa

^c Department of Electrical Power Engineering, Durban University of Technology, Durban, South Africa

^d Department of Computer Engineering, Faculty of Engineering, Redeemer's University, Ede, Osun State, Nigeria

^e Department of Electrical and Electronic Engineering, Faculty of Engineering, Redeemer's University, Ede, Osun State, Nigeria

ARTICLE INFO

Index Terms:

Stacking ensemble-regression machine learning model

Automated hyperparameter tuning

Composite raytracing-image-method

Millimetre-wave

ABSTRACT

Path loss prediction is a crucial aspect of designing and operating wireless communication systems, especially in the millimetre-waves (mmWaves) frequency bands. However, these bands are associated with climate-related challenges: rain attenuation, and free space path loss. To address these challenges, an advanced stacking ensemble-regression machine learning (SEML) model with automated hyperparameter tuning (AHT) was proposed. The AHT-SEML model leverages multiple base regressors integrated with a meta-regressor. The model's performance was optimised using the AHT tuning technique. The AHT-SEML model's efficiency was tested using simulated path loss data from a Composite 3D Raytracing-Image-Method propagation model across four sub-Saharan cities, at mmWaves frequencies. The AHT-SEML model's performance was compared to three empirical path loss models, namely Close-In (CI), Floating Intercept (FI), and Alpha-Beta-Gamma (ABG), using evaluation metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). AHT-SEML outperformed other models in the four cities across all frequencies and scenarios with the highest Index of Agreement and lowest Bayesian information criterion. Model confidence set (MCS) analysis with CI benchmark indicates that all the models except AHT-SEML performed below the critical t-value of 2.3530 at 95% confidence level with a degree of freedom of 3, implying no significant differences in their MAEs compared to the CI. However, AHT-SEML's t-statistic values exceed this critical t-value, indicating statistically significant differences and better performance than the CI benchmark models. Similarly, F-statistics of 29.45 and 26.54 correspond to p-values of 1.91×10^{-14} and 2.50×10^{-13} for MAE and RMSE, respectively, corroborating significant differences in the AHT-SEML's performance.

1. Introduction

The next generation of wireless communication systems is expected to rely heavily on millimetre-wave (mmWave) technology [1,2] due to its ability to deliver high data rates and low latency. However, accurate prediction of path loss is essential for the effective use of mmWave technology. The signal propagation characteristics at higher frequencies, including the effects of atmospheric absorption, reflection, and diffraction, make the path loss models for mmWave technology complex [3].

It is essential to have an accurate path loss model for planning, designing, and constructing wireless networks, especially when dealing with millimetre waves, which have sensitive channel characteristics. This model helps operators make detailed estimations before deployment. There are three main methods for path loss modelling: empirical methods [4,5], deterministic methods [6], and machine learning-based methods [7].

The empirical methods naturally use measurement results gathered in a standard path loss scenario to create a model based on the best curve-fitting mechanisms. In these methods, the transmitter and receiver

* Corresponding author.

E-mail addresses: afape7885@run.edu.ng (J.O. Afape), willoughbya@run.edu.ng (A.A. Willoughby), sanyaolum@run.edu.ng (M.E. Sanyaolu), obiseyeo@dut.ac.za (O.O. Obiyemi), katilehom@dut.ac.za (K. Moloi), joodaj@run.edu.ng (J.O. Jooda), dairof@run.edu.ng (O.F. Dairo).

distances (Tx and Rx), the carrier frequency (f), the path loss exponent (PLE), and the shadow fading standard deviation (σ) are the usually considered propagation parameters [8,9].

The most popular deterministic method for mmWaves path loss modelling is 3D raytracing. It is built on geometric optics that mimic the behaviour of physical propagation, including attenuation, reflection, and scattering [10]. Although deterministic methods require much computational power, they are more accurate than empirical ones [11].

Machine learning algorithms offer several advantages over traditional 3D raytracing methods, including speed, efficiency, generalisation, flexibility, scalability, data-driven approach, real-time applications, and reduced computational cost [10,12]. Machine learning models can provide rapid predictions or generate outputs without extensive computational simulations, generalise patterns and relationships, adapt to diverse inputs and tasks, and efficiently handle large datasets and complex scenes [13]. They can also be trained on various types of data, enabling tasks like denoising, inpainting, reconstruction, and scene understanding. Machine learning techniques can be highly scalable, enabling efficient handling of large datasets and complex scenes [14].

Machine learning algorithms can better capture the richness and complexity of real-world signal propagation phenomena such as multi-path propagation, diffraction, and scattering than simplistic ray tracing models. Machine learning models may develop realistic channel impulse responses and spatial channel models that accurately represent the behaviour of millimetre-wave signals in varied propagation environments by learning from large-scale measurement datasets and high-fidelity simulations [12,15].

Masood et al. [15] proposed a Machine Learning (ML) approach for radio wave propagation modelling and RSS estimation. It estimates RSS using smart predictors, including transmitter parameters and propagation environment characteristics. Despite sparse training data, Deep Neural Networks (DNN) outperforms other ML techniques, providing a 25% increase in prediction accuracy and a 12x decrease in prediction time compared to ray tracing. According to Ahmadien et al. [12], ray tracing (RT) or ray launching (RL) simulations require 3D models but may not be readily available or practical, and high computational costs prevent real-time applications compared to machine learning models, even when 3D models are available.

Machine learning-based path loss modelling is considered a regression problem based on empirical data. In this path loss modelling, the characteristics determined from data in the propagation environment and measurement location are used as input, and the measured path loss values are employed during training. The transmitter and receiver distances, the carrier frequency (f), the transmitter and receiver's antenna properties, sharp-edge diffraction, obstacle refraction and penetration, and reflection are the commonest characteristics that machine learning-based models for path loss in urban scenarios take into account [16].

Also, machine learning is assumed to be a replacement technique for building path loss models with incredibly accurate prediction capabilities based on training data, even when in-depth knowledge of a particular propagation environment is unavailable [17]. The flexible model architecture of the machine learning algorithms allows them to learn from data and make predictions. Machine learning algorithms can be used in millimetre-wave channel parameters' predictions to train the models using historical data and discover the relationships between the channel parameters and the propagation environments. After which, the trained models can be used to predict any propagation channel parameter. On the other hand, choosing the appropriate algorithm and hyperparameters for path loss prediction is a complex process because different algorithms and hyperparameters perform differently for different datasets. The predictions of base models are combined as effectively as possible using ensemble learning techniques. To put it briefly, they combine the individual models that perform best to create a better meta-model. Machine learning models' prediction accuracy and computational efficiency are higher in many scenarios than empirical

and deterministic models [18].

In recent years, machine learning (ML) has been employed in several studies to predict path loss in various radio propagation environments, especially in urban and suburban scenarios. Support vector regression (SVR) was suggested by Timoteo et al. [7] as a technique for forecasting path loss in an urban outdoor environment. According to simulation findings, SVR performs better than the two empirical propagation models (Okumura-Hata and Ericsson 9999 Models). Popoola et al. [37] used an Extreme Learning Machine (ELM) algorithm to develop an optimal path loss prediction model for outdoor propagation scenarios. The results of the Artificial Neural Network – Back Propagation (ANN-BP), COST-231, Okumura-Hata, and ELM models, when in comparison with the training dataset's target variable, have RMSE values of 2.449, 6.116, 7.456, and 2.896 dB, respectively, and regression coefficient (R) values of 0.973, 0.935, 0.935, and 0.959, respectively, on prediction accuracy. In addition, Popoola et al. [19] also developed an optimal model for path loss predictions at 1.8 GHz, utilising the algorithm for the feed-forward neural network (FFNN). The generated ANN model outperformed the Hata, COST 231, ECC-33, and Egli models regarding prediction accuracy and generalisability.

Furthermore, different models, including artificial neural networks, support vector regression, and random forests, were trained and evaluated based on measured data by Zhang et al. [20]. The machine-learning-based algorithms were demonstrated to perform better than the log-distance model. Cheng et al. [21] proposed a new Enhanced local area multi-scanning (E-LAMS) algorithm that provides a convolutional neural network (CNN) structure with four subnetworks and feature-sharing layers to improve path loss prediction accuracy by achieving RMSE of 8.59 dB in the test scenarios. Nguyen and Cheema [22] used a feed-forward deep neural network (DNN) model to predict path loss of 13 different frequencies from 0.8 GHz to 70 GHz simultaneously in an urban and suburban environment in a non-line-of-sight (NLOS) scenario. The outcomes demonstrated that the proposed DNN-based path loss model outperformed the multi-frequency alpha-beta-gamma (ABG) path loss model in terms of prediction accuracy R^2 and mean square error (MSE), which was improved by roughly 6 dB.

Support vector regression (SVR) and radial basis function (RBF) models were introduced by Ojo et al. [23] to perform path loss predictions utilising measurement data produced by the 3G network in Lefkiosa and Kyrenia, Northern Cyprus. The findings demonstrated that the two machine learning models predicted path loss more accurately than empirical models such the Cost-231, SUI, Egli, Free space, and Cost-231 W-I. Conversely, the SVR model performed the best across all indices, with RMSE values in rural, suburban, and urban environments of 1.378 dB, 1.4523 dB, and 2.1568 dB, respectively. Ojo et al. [23] developed machine learning-based bagging and blending-ensemble path loss prediction models using data from Nigeria's 4G LTE wireless networks. According to the results, the created bagging-ensemble path loss prediction model provided the datasets with the lowest errors (MSE = 0.0011 dB, Sum of Squared Errors (SSE) = 0.6069 dB, Mean Absolute Error (MAE) = 0.0245 dB, and R-squared (R^2) = 0.7484). Sotirovdis et al. [24] used ensemble learning and oversampling techniques to provide a satisfactory path loss model for flying base stations (FBSs) using the datasets obtained through ray tracing simulations.

Path loss data collected in several rural regions across Greece were used to train the models as Moraits et al. [25] examined stacking, voting, bagging, and gradient-boosted tree ensemble methods. The findings demonstrated that all ensemble models improved prediction accuracy and superseded all single-model-based techniques. The best performance of the newly proposed ensemble techniques is demonstrated by the stacking ensemble method, which uses five base learners and a bespoke DNN as a meta-learner.

However, most of these studies only examined the performance of individual algorithms rather than combining several algorithms to improve prediction accuracy. Multiple algorithms are combined through ensemble learning, which has been demonstrated to increase prediction

accuracy and decrease overfitting [25]. Stacking is a typical ensemble strategy to improve prediction accuracy, combining numerous base models and a meta-model [24]. Stacking ensemble models combine multiple base models, enhancing prediction accuracy by leveraging the strengths of each model while mitigating their weaknesses, reducing overfitting risk, and capturing complex patterns and relationships in the data that may be missed by simpler models [26,27]. Stacking ensemble models offer flexibility in model selection and combination, allowing them to perform well across different datasets and prediction tasks. Additionally, stacking can perform automatic feature engineering by learning to combine base models, improving accuracy by capturing higher-order interactions [28].

To the best of our knowledge, research is yet to be done on employing an automated hyperparameter tuning stacking ensemble regressor for predicting millimetre-waves path loss at these frequencies and locations.

This paper proposes a stacking-ensemble regression machine learning model with base regressors, including Linear Regression, Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, SV Regressor, and Gradient Boosting Regressor, with a meta-regressor: Extreme Gradient Boosting Regressor (XGBR) for mmWave path loss prediction. An automated hyperparameter tuning technique was used to choose the ideal hyperparameters for the ensemble regressor. Our suggested approach seeks to increase the accuracy of path loss prediction and reduce the selection bias of individual algorithms. The proposed model used data from a 3D ray tracing simulation of mmWave signal propagation in a real-world setting. The stacking ensemble model was seen to outperform the popular empirical path loss models like the Close-In (CI) Free-Space Reference Distance Large-scale, the Floating Intercept (FI), and the Alpha-Beta-Gamma (ABG) in terms of prediction accuracy.

The article is structured as follows: Section II covers the methodology, Section III presents the result analysis, Section IV offers the conclusion, and Section V discusses the study's limitations and recommendations.

2. Methodology

The proposed stacking ensemble regressor consists of six base regressors, namely Linear Regression, Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, Support Vector Regressor, and Gradient Boosting Regressor, while the Xtreme Gradient Boosting Regressor is used as the meta-regressor. Nine (9) different machine learning algorithms namely Linear Regression, Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, Gradient Boosting Regressor, Support Vector Regressor, Xtreme Gradient Boosting Regressor, Stacking Ensemble Model1 and the Stacking Ensemble Model2, were compared with the proposed stacking ensemble path loss model. Seven of the ML models were single-based models while the remaining were stacking ensemble models. The Stacking Ensemble Model1 makes use of five base regressors, which are Linear Regression, Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, and Gradient Boosting Regressor, with the Xtreme Gradient Boosting Regressor as the meta-regressor. Similarly, the Stacking Ensemble Model2 is designed with four base regressors, which are the Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, and Gradient Boosting Regressor, with the Xtreme Gradient Boosting Regressor as the meta-regressor. The empirical models used in this study are Alpha-Beta-Gamma (ABG), Close-In (CI) Free-Space Reference Distance, and Floating Intercept (FI) large-scale path loss models.

2.1. Stacking ensemble model

To create the stacking ensemble, the predictions of the base regressors were merged and then fed into the meta-regressor. The training

set was divided into two parts: the first was used to train the base regressors, while the second was used to create predictions for the meta-regressor. The predictions of the base regressors from the second part were then fed into the meta-regressor to obtain the final predictions [29].

2.1.1. Base regressors

The six base regressors considered are Linear Regression, Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, Support Vector Regressor, and Gradient Boosting Regressor. They are all machine learning techniques used for regression analysis. Regression analysis is a statistical approach for estimating the interaction between dependent and independent variables. The regression analysis is used primarily to develop a model that predicts the value of the dependent variable based on the value of the independent variables.

(i). Linear Regressor

Linear regression algorithm is a method that fits a line to the observed data to explain a linear connection between the input and output variables. It is a parametric model that predicts the outputs by making certain assumptions about the data. The predicted output, y_{LR} is given as Equation (1):

$$y_{LR} = w_1 * X_1 + w_2 * X_2 + w_3 * X_3 + \dots + w_n * X_n + b \quad (1)$$

where $X_1, X_2, X_3, \dots, X_n$ are the input features, and $w_1, w_2, w_3, \dots, w_n$ are the corresponding weights or coefficients assigned to each feature. These weights represent the strength of the relationship between each feature and the output, while b is the bias term or intercept. The bias term represents the value of y_{LR} when all input features are zero. In order to make accurate predictions, the model learns the values of the weights and bias term from the training data. This is typically done by minimising a cost function, such as mean squared error (MSE), which measures the difference between the predicted values and the actual target values in the training data. The outcome is a linear equation that best fits the observed data points when the weights and bias term are set to optimal values.

(ii). Histogram Gradient Boosting Regressor

The HGBR algorithm is an ensemble method that uses decision trees to produce several weak learners. These learners are then combined to form a strong learner using a gradient-boosting approach. By creating many decision trees and merging their predictions, HGBR produces a final forecast. The algorithm also uses histogram to display the data distribution in each tree's leaf nodes, which improves training efficiency and memory usage. HGBR is known for its ability to handle datasets with many features and complex relationships between independent and dependent variables. The predicted output, y_{HGBR} is represented as Equation (2):

$$y_{HGBR} = \sum_{i=1}^m (\gamma_i * h_i(X)) \quad (2)$$

where i is an index that denotes a specific decision tree, m is the total number of the decision trees, $h_i(X)$ represents the prediction of the i -th decision tree for the input X , and γ_i is the corresponding weight of the i -th decision tree. These weights are determined through the gradient-boosting process and indicate the contribution of each tree's prediction to the final ensemble prediction.

(iii). K Neighbours Regressor

The K Neighbours Regressor is a non-parametric regression method. This technique determines the value of the dependent variable based on

Table 1

Auto-generated optimal hyperparameters.

REGRESSOR	HYPERPARAMETERS									
	learning_rate	max_depth	max_iter	neighbours	weights	p	n_estimators	C	gamma	kernel
Liner Regression	–	–	–	–	–	–	–	–	–	–
HGBR	0.1	7	500	–	–	–	–	–	–	–
KNR	–	–	–	20	uniform	1	–	–	–	–
RFR	–	20	–	–	–	–	1000	–	–	–
SVR	–	–	–	–	–	–	–	10	scale	rbf
GBR	0.5	5	–	–	–	–	–	–	–	–
AHT-SEML	0.1	10	–	–	–	–	1000	–	–	–

the k-nearest neighbours of the newly added data point. The method calculates the average (mean) of the target values of the k-nearest neighbours, y_{KNNR} , which is represented by Equation (3):

$$y_{KNNR} = \left(\frac{1}{k} \right) * \sum_{i=1}^k (y_i) \quad (3)$$

where i denotes the individual nearest neighbor, k is the number of nearest neighbours, and y_i is the target value of the i -th neighbor.

(iv). Random Forest Regressor

The Random Forest Regressor is a tree-based ensemble regressor algorithm that utilises decision trees to build a forest of weak learners, which are then combined to form a strong learner by using bagging methods. The algorithm combines all of the decision trees' predictions to provide a final prediction. This algorithm is capable of providing accurate predictions on datasets with many features and is robustness to overfitting. Additionally, it is helpful in identifying which input features are important for making predictions through its feature importance estimates. The predicted output, y_{RFR} , is given as Equation (4):

$$y_{RFR} = \sum_{i=1}^m (\gamma_i * h_i(X)) \quad (4)$$

where i is an index that denotes the individual decision tree, m is the total number of the decision trees, $h_i(X)$ represents the prediction of the i -th decision tree for the input X , and γ_i is the corresponding weight of the i -th decision tree. Unlike some other ensemble methods, such as Gradient Boosting and Histogram Gradient Boosting Regressors, Random Forest does not explicitly assign weights to trees. Instead, the ensemble's predictions are aggregated without weighting, and each tree contributes equally to the final prediction.

(v). Support Vector Regressor

The Support Vector Regressor is a kernel-based algorithm that uses support vector machines to transform data to a higher-dimensional plane that best fits the training data while minimising the margin violations, that is, data points that fall outside the margin. Unlike other algorithms, it does not aim to split data into distinct classes but instead focuses on predicting a continuous target variable. It works well on datasets with a small number of features. The predicted output, y_{SVR} is represented as Equation (5):

$$y_{SVR} = w_1 * X_1 + w_2 * X_2 + w_3 * X_3 + \dots + w_n * X_n + b \quad (5)$$

where $X_1, X_2, X_3, \dots, X_n$ are the input features, $w_1, w_2, w_3, \dots, w_n$ are the corresponding weights or coefficients assigned to each feature. These weights represent the strength of the relationship between each feature and the output. The bias term or intercept is represented by b . In contrast to linear regression, SVR can use various kernel functions such as linear, polynomial, and radial basis function to map data into a higher-dimensional plane. This allows it to capture complex non-linear re-

lationships between the input features and the target variable.

(vi). Gradient Boosting Regressor

The Gradient Boosting Regressor is a tree-based ensemble algorithm. It produces a final prediction by constructing numerous decision trees sequentially, with each tree aiming to correct the errors or residuals made by the previous trees and integrating their predictions. The technique works iteratively and uses a loss function to evaluate the model's performance and then adjusts the model parameters to minimise the loss function during each iteration. This process repeats for a specified number of iterations or until a convergence criterion is met. The final prediction, y_{GBR} , which is the sum of the predictions of all the individual trees, is obtained from Equation (6).

$$y_{GBR} = \sum_{i=1}^m (\gamma_i * h_i(X)) \quad (6)$$

where i is an index that denotes the individual decision tree, m is the total number of the decision trees, $h_i(X)$ represents the prediction of the i -th decision tree for the input X , and γ_i is the corresponding weight of the i -th decision tree. The weights are determined through the gradient-boosting process and indicate the contribution of each tree's prediction to the final ensemble prediction.

2.1.2. Meta-regressor

2.1.2.1. Xtreme Gradient Boosting Regressor. The Xtreme Gradient Boosting Regressor was utilised as the meta-regressor to integrate the predictions of the base regressors. The XGBR Regressor is an improved version of the Gradient Boosting Regressor, which employs a tree-based ensemble technique that is fast, scalable, and accurate.

Let the output of each base regressor be denoted as y_i , where i ranges from 1 to 6. Then, the meta-regressor (XGBR) will combine these base regressors as follows:

$$y_{XGBR} = \sum_{i=1}^{m=6} (\alpha_i * y_i) \quad (7)$$

where α_i represents the weight (importance) of the i -th base regressor in the ensemble.

Finally, the overall prediction of the Stacking Ensemble-Regression Machine Learning model with Automated Hyperparameter Tuning (AHT-SEML) for omnidirectional path loss prediction is the output of the meta-regressor:

$$y_{final_prediction} = y_{XGBR} \quad (8)$$

The model will learn the optimal values of the weights $w_1, w_2, w_3, \dots, w_n$, γ_i , and α_i during the training process using the data and automated hyperparameter tuning techniques.

2.1.3. Hyperparameter tuning

This technique performs a lookup over a set of hyperparameters for each of the base regressors and the meta-regressor.

RandomizedSearchCV function from the sci-kit-learn module [29] was used to evaluate the performance of each set of hyperparameters and select the ideal set of hyperparameters with the lowest MSE on the testing set as shown in Table 1. The choice of this function was mainly based on the limited available computational resources (cost). However, many rounds of hyperparameter tuning and cross-validation were undertaken to ensure the robustness of the chosen hyperparameters as well as the model's generalisation performance. The automatic hyperparameter tuning was confined to the training period to offer a comprehensive evaluation and reduce overfitting risk using the RandomizedSearchCV optimisation algorithm. Also, this process was adopted to ensure that the model was based solely on historical data and did not incorporate future knowledge to maintain the integrity of the forecasting process and reduce the risk of bias.

2.2. The empirical models

The mathematical derivations of the average values of PLE (n), α^{FI} , α^{ABG} , β^{ABG} , and γ^{ABG} parameters in the CI, FI, and ABG omnidirectional path loss models were computed with closed-form solutions that minimise the SF standard deviations (σ^{CI} , σ^{FI} , σ^{ABG}) through the MMSE fit on the path loss data from the simulation distances at 28, 38, 60, and 73 GHz, respectively, as defined in Ref. [5,30,31].

(i) . The Close-In (CI) Free-Space Reference Distance Large-scale Path Loss Model

The Close-In (CI) free-space reference distance large-scale path loss model [5] is a single slope model with physical relevance tied to the free-space loss at a specific reference distance, typically 1 m for millimetre-waves applications.

Close-In free space reference distance large-scale path loss, PL^{CI} is given as:

$$PL^{CI}(f, d)[dB] = PL_0 + 10n \log_{10}\left(\frac{d}{d_0}\right) + \chi_{\sigma}^{CI} \quad (9)$$

where the free space path loss, PL_0 , is [8]

$$PL_0 = 10 \log_{10}\left(\frac{4\pi d_0}{\lambda}\right)^2 \quad (10)$$

d denotes the distance between the transmit antenna and the receive antenna, d_0 is the close-in free space reference distance of 1 m, λ represents the transmit frequency's wavelength, n is the single model parameter, the path loss exponent (PLE) for a particular frequency band or environment [8], χ_{σ}^{CI} is the large-scale shadow fading with a standard deviation, σ^{CI} .

(ii) . The Floating Intercept (FI) Path Loss Model

Floating Intercept large-scale path loss, PL^{FI} is given as [30,31]:

$$PL^{FI}(d)[dB] = \alpha^{FI} + 10\beta^{FI} \log_{10}(d) + \chi_{\sigma}^{FI} \quad (11)$$

where α^{FI} denotes the floating intercept in dB, β^{FI} represents a coefficient known as line slope (path loss exponent) that characterises the distance dependency of the path loss, χ_{σ}^{FI} represents a zero mean Gaussian random variable that stands for shadow factor, and σ^{FI} denotes the shadow factor standard deviation in decibels.

(iii) . The Alpha-Beta-Gamma (ABG) Path Loss Model

The ABG reference distance and frequency of large-scale path loss, PL^{ABG} , is given as (Akdeniz et al., 2014; [10,32]):

$$\begin{aligned} PL^{ABG}(f, d)[dB] &= 10 \alpha^{ABG} \log_{10}\left(\frac{d}{1m}\right) + \beta^{ABG} + 10\gamma^{ABG} \log_{10}\left(\frac{f}{1GHz}\right) \\ &+ \chi_{\sigma}^{ABG} \end{aligned} \quad (12)$$

where α^{ABG} and γ^{ABG} represent the coefficients showing how path loss varies with frequency and distance, respectively, β^{ABG} stands for the optimised offset value for path loss in dB, f denotes the carrier frequency in GHz, d is the 3D transmitter-receiver (T-R) separation distance in metres, and χ_{σ}^{ABG} is the shadow fading with standard deviation, σ^{ABG} describes large-scale signal fluctuations about the mean path loss over distance.

2.3. Data processing

The dataset used in this study was generated from simulated path loss data using a Composite 3D Raytracing-Image-Method (CRTI) propagation model developed in MATLAB R2023a [35,36]. CRTI takes into account site-specific atmospheric conditions, topography and building materials, building constructions, building heights, and city geographical layouts during simulations. The data contain the parameters for the outdoor propagation of mmWaves at 28, 38, 60, and 73 GHz in four Nigerian cities: Abuja, Ibadan, Lagos, and Port Harcourt. The geographical longitudes and latitudes of the cities are as follows: 9.07° N, 7.49° E; 6.46° N, 3.39° E; 7.45° N, 3.89° E; 4.78° N, 7.01° E. The dataset consists of 2800, 3100, 3050, and 2800 samples for the cities, respectively.

To quicken convergence and boost the algorithms' generalisation ability, data transformation procedures were employed. To achieve this objective, all input and output parameters underwent normalisation using a standard scaler from Scikit Learn [29] before the training phase. The standard scaler normalised the data and brought all features to a similar scale to increase the accuracy and efficacy of the machine learning models. The only categorical variables, LOS and NLOS scenarios, were encoded into 0 and 1, respectively.

Data Splitting was done by dividing the dataset into ratio 0.7:0.2:0.1 for training, validation, and testing sets. The automatic hyperparameter tuning was performed using the entire dataset, including both the training and validation sets. The testing set was used to evaluate how the model would perform on the unseen data. Additionally, each dataset has ten input features, such as distance between the transmitter and receiving antennas, scenario types, frequency, free-space path loss, longitude and latitude, average temperature, average air pressure, average water vapour density, average rainrate, and an output value that represents the omnidirectional path loss.

2.4. Performance metrics

The accuracy of the path loss model was evaluated using performance matrices such as mean absolute error (MAE), root mean square error (RMSE), Index of agreement (IoA), and Bayesian information criterion (BIC).

$$MAE = \frac{1}{N} \sum_{k=0}^N (PL_S - PL_P) \quad (13)$$

$$MSE = \frac{1}{N} \sum_{k=0}^N (PL_S + PL_P)^2 \quad (14)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=0}^N (PL_S + PL_P)^2} \quad (15)$$

Table 2

Performance metrics of different machine learning models.

City	Freq (GHz)	Scenario	Linear Regression		Histogram Gradient Boosting Regression		KN Regression		Random Forest Regression		Support Vector Regression		Gradient Boosting Regression		XGBOOST Regression		Stacking Ensemble Model1		Stacking Ensemble Model2		AHT-SEML Path Loss Model	
			MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Abuja	28	LOS	18.465	25.830	10.650	16.337	17.181	26.456	10.078	16.990	14.397	26.305	13.789	20.374	9.183	14.369	8.615	13.632	10.253	16.003	6.404	10.809
		NLOS	22.343	31.022	12.887	19.621	20.789	31.774	12.194	20.405	9.876	15.971	16.684	24.470	11.111	17.257	10.424	16.372	12.406	19.220	6.396	10.778
	38	LOS	18.096	24.280	10.437	15.357	16.838	24.869	9.876	15.971	14.109	30.628	13.513	19.152	8.999	13.507	8.442	12.814	10.048	15.043	6.182	10.622
		NLOS	21.896	26.346	12.629	16.664	20.374	26.985	11.950	17.330	14.266	29.467	16.351	20.782	10.889	14.656	10.215	13.904	12.158	16.323	6.822	10.888
	60	LOS	20.087	28.519	11.585	18.038	18.690	29.211	10.962	18.759	15.661	31.897	14.999	22.495	9.989	15.865	9.371	15.051	11.153	17.669	7.961	11.032
		NLOS	24.305	29.118	14.018	18.417	22.615	29.825	13.264	19.153	18.949	32.568	18.149	22.968	12.087	16.198	11.339	15.368	13.495	18.041	7.518	10.905
	73	LOS	18.694	28.296	10.782	17.897	17.394	28.982	10.202	18.612	20.589	35.084	13.959	22.319	9.297	15.741	8.721	14.933	10.380	17.531	6.893	10.88
		NLOS	22.620	28.890	13.046	18.273	21.047	29.591	12.345	19.003	18.759	32.987	16.891	22.788	11.249	16.071	10.553	15.247	12.560	17.899	7.032	10.762
Lagos Island	28	LOS	17.020	26.598	9.817	16.823	15.837	27.243	9.289	17.495	18.745	32.979	12.709	20.980	8.464	14.796	7.940	14.037	9.450	16.479	6.772	9.215
		NLOS	22.168	33.638	12.786	21.276	20.626	34.454	12.098	22.126	13.283	28.180	16.553	26.533	11.024	18.712	10.342	17.753	12.308	20.841	7.015	9.031
	38	LOS	22.414	32.363	12.928	20.469	20.855	33.148	12.232	21.288	17.263	36.952	16.737	25.528	11.147	18.003	10.457	17.080	12.445	20.051	7.559	9.074
		NLOS	24.606	35.032	14.192	22.157	22.895	35.882	13.429	23.043	16.196	29.487	18.374	27.633	12.237	19.488	11.479	18.489	13.662	21.705	6.118	9.046
	60	LOS	29.773	35.768	17.172	22.623	27.703	36.636	16.249	23.528	19.598	30.107	22.232	28.214	14.806	19.898	13.890	18.877	16.532	22.161	6.772	9.299
		NLOS	32.349	38.532	18.658	24.371	30.100	39.467	17.655	25.346	26.827	32.434	24.156	30.394	16.087	21.435	15.092	20.336	17.962	23.873	6.703	9.143
	73	LOS	29.474	36.229	17.000	22.914	27.424	37.107	16.085	23.830	24.443	30.495	22.009	28.577	14.657	20.154	13.750	19.120	16.365	22.446	6.488	9.281
		NLOS	29.453	36.221	16.987	22.909	27.405	37.099	16.074	23.825	24.425	30.488	21.993	28.570	14.647	20.149	13.740	19.116	16.354	22.441	6.902	9.209
Ibadan	28	LOS	20.870	30.950	12.037	19.576	19.419	31.701	11.390	20.358	17.308	26.052	15.584	24.413	10.379	17.217	9.736	16.334	11.588	19.176	6.453	10.936
		NLOS	27.124	40.583	15.644	25.669	25.238	41.568	14.803	26.695	22.494	34.160	20.254	32.012	13.489	22.576	12.654	21.418	15.061	25.144	6.83	10.167
	38	LOS	25.448	32.385	14.678	20.484	23.678	33.171	13.888	21.302	21.104	27.260	19.003	25.545	12.655	18.016	11.872	17.092	14.130	20.065	7.282	10.277
		NLOS	27.178	47.123	15.675	29.805	25.288	48.266	14.832	30.996	22.538	39.665	20.294	37.170	13.515	26.214	12.679	24.870	15.090	29.196	6.87	10.253
	60	LOS	29.895	39.772	17.243	25.155	27.816	40.736	16.315	26.161	24.792	33.477	22.324	31.372	14.867	22.125	13.947	20.990	16.599	24.641	7.448	10.875
		NLOS	32.912	43.052	18.983	27.230	30.623	44.096	17.962	28.318	27.294	36.238	24.576	33.959	16.367	23.949	15.354	22.721	18.274	26.673	7.016	10.484
	73	LOS	26.334	37.033	15.188	23.423	24.502	37.932	14.372	24.360	21.839	31.172	19.664	29.212	13.096	20.601	12.285	19.545	14.622	22.945	7.292	10.327
		NLOS	30.019	42.715	17.314	27.017	27.932	43.750	16.383	28.097	24.895	35.954	22.416	33.693	14.929	23.762	14.005	22.543	16.668	26.464	7.981	10.241
Port Harcourt	28	LOS	26.718	34.514	15.410	21.830	24.860	35.351	14.581	22.703	22.157	29.052	19.951	27.225	13.287	19.200	12.465	18.215	14.835	21.384	6.728	11.172
		NLOS	29.419	35.000	16.968	22.137	27.373	35.849	16.055	23.022	24.397	41.862	21.968	27.608	14.630	19.470	13.725	18.472	16.335	21.685	7.646	11.35
	38	LOS	29.481	36.961	17.004	23.378	27.431	37.857	16.089	24.312	24.669	40.276	22.014	29.155	14.661	20.561	13.754	19.507	16.369	22.900	7.408	11.462
		NLOS	32.655	40.009	18.834	25.305	30.384	40.980	17.822	26.317	27.081	43.597	24.384	31.559	16.239	22.257	15.234	21.115	18.132	24.788	7.732	11.47
	60	LOS	33.341	40.850	19.230	25.837	31.022	41.841	18.196	26.870	32.768	44.514	24.897	32.222	16.580	22.725	15.554	21.559	18.512	25.309	7.602	11.238
		NLOS	36.061	44.730	20.799	28.291	33.553	45.815	19.680	29.422	35.603	47.953	26.927	35.283	17.933	24.883	16.823	23.607	20.023	27.713	7.109	11.353
	73	LOS	31.045	40.014	17.905	25.308	28.886	40.984	16.943	26.320	32.438	45.086	23.182	31.562	15.438	22.259	14.483	21.118	17.237	24.791	7.829	11.629
		NLOS	36.100	40.487	20.821	25.608	33.590	41.469	19.702	26.631	35.642	45.620	26.957	31.936	17.953	22.523	16.842	21.368	20.045	25.084	7.542	11.529

Table 3

Performance metrics of the empirical and proposed models.

City	Freq (GHz)	Scenario	Close-In (CI) Path Loss Model				Floating Intercept (FI) Path Loss Model				ABG Path Loss Model				AHT-SEML Path Loss Model			
			MAE	RMSE	IoA	BIC	MAE	RMSE	IoA	BIC	MAE	RMSE	IoA	BIC	MAE	RMSE	IoA	BIC
Abuja	28	LOS	19.044	22.846	0.673	8356.846	18.695	22.042	0.647	8200.444	17.723	21.067	0.735	8161.272	6.404	10.809	0.987	4309.370
		NLOS	14.832	17.283	0.625	7939.004	14.746	17.068	0.602	7790.422	14.823	17.139	0.684	7753.209	6.396	10.778	0.918	4093.901
	38	LOS	20.696	25.2	0.699	8691.120	19.25	23.693	0.673	8528.462	19.646	24.067	0.764	8487.723	6.182	10.622	1.027	4481.745
		NLOS	15.812	18.331	0.686	8523.983	15.929	18.344	0.660	8364.453	14.43	16.971	0.750	8324.498	6.822	10.888	1.007	4395.557
	60	LOS	23.985	27.714	0.632	7855.435	21.141	24.545	0.608	7708.417	20.768	24.227	0.691	7671.596	7.961	11.032	0.928	4050.808
		NLOS	24.491	28.059	0.726	9276.099	24.691	28.152	0.699	9102.493	23.987	27.559	0.794	9059.012	7.518	10.905	1.066	4783.401
	73	LOS	15.74	17.525	0.676	8494.734	13.659	14.942	0.650	8335.751	9.404	11.571	0.738	8295.933	6.893	10.88	0.992	4380.475
		NLOS	16.171	17.816	0.682	8749.576	16.079	17.866	0.656	8585.824	17.631	19.459	0.746	8544.811	7.032	10.762	1.002	4511.889
Lagos Island	28	LOS	27.099	29.946	0.647	8811.926	24.13	22.876	0.725	8580.569	19.884	22.585	0.738	8552.332	6.772	9.215	0.999	1732.293
		NLOS	15.688	18.093	0.602	8371.330	15.88	18.157	0.675	8151.540	15.839	18.116	0.686	8124.716	7.015	9.031	0.929	1645.678
	38	LOS	30.141	33.14	0.673	9164.404	26.355	29.124	0.754	8923.791	26.357	29.127	0.767	8894.425	7.559	9.074	1.039	1563.394
		NLOS	17.918	20.481	0.660	8988.165	17.951	20.499	0.740	8752.180	20.141	22.729	0.753	8723.379	6.118	9.046	1.019	1766.939
	60	LOS	41.475	45.448	0.608	8283.211	36.266	40.209	0.682	8065.734	36.435	40.38	0.694	8039.192	6.772	9.299	0.939	1628.355
		NLOS	26.03	29.184	0.699	9781.238	25.515	28.878	0.783	9524.431	25.64	28.999	0.797	9493.089	6.703	9.143	1.079	1922.845
	73	LOS	34.254	38.583	0.650	8957.323	30.208	34.062	0.729	8722.148	27.925	31.776	0.741	8693.446	6.488	9.281	1.003	1760.876
		NLOS	19.872	22.599	0.656	9226.043	19.64	22.442	0.736	8983.812	23.376	26.379	0.748	8954.249	6.902	9.209	1.013	1813.702
Ibadan	28	LOS	15.921	16.997	0.750	13674.826	15.243	16.172	0.764	13446.086	16.075	16.882	0.761	13529.501	6.453	10.936	0.999	1888.856
		NLOS	18.3	21.095	0.697	12991.084	17.373	20.009	0.711	12773.782	17.829	20.472	0.708	12853.026	6.83	10.167	0.929	1794.414
	38	LOS	20.847	21.659	0.780	14221.819	20.377	21.112	0.795	13983.930	20.961	21.663	0.791	14070.681	7.282	10.277	1.039	2058.853
		NLOS	20.096	23.616	0.765	13948.322	19.32	22.238	0.780	13715.008	18.867	21.775	0.776	13800.091	6.87	10.253	1.019	1964.411
	60	LOS	28.244	29.672	0.705	12854.336	26.887	28.082	0.719	12639.321	26.592	27.821	0.715	12717.731	7.448	10.875	0.939	1775.525
		NLOS	33.075	37.748	0.809	15179.056	32.768	37.098	0.826	14925.156	32.486	36.82	0.822	15017.747	7.016	10.484	1.078	2096.631
	73	LOS	25.269	26.435	0.753	13900.460	24.666	25.659	0.768	13667.946	24.663	25.656	0.764	13752.738	7.292	10.327	1.003	1920.023
		NLOS	25.05	29.088	0.760	14317.474	24.366	28.016	0.776	14077.985	25.175	28.813	0.772	14165.320	7.981	10.241	1.013	1977.623
Port Harcourt	28	LOS	52.554	54.007	0.533	15716.755	47.349	49.68	0.605	15417.857	46.262	48.713	0.608	15367.376	6.728	11.172	0.999	2196.482
		NLOS	52.218	54.481	0.496	14930.918	50.23	52.58	0.563	14646.964	47.998	50.501	0.565	14599.007	7.646	11.35	0.929	2086.658
	38	LOS	51.867	53.513	0.554	16345.426	46.06	48.501	0.629	16034.571	46.296	48.719	0.632	15982.071	7.408	11.462	1.039	2284.341
		NLOS	51.842	54.047	0.544	16031.090	49.206	51.789	0.617	15726.214	48.6	51.214	0.620	15674.723	7.732	11.47	1.019	2240.412
	60	LOS	49.265	51.635	0.501	14773.750	43.118	46.376	0.569	14492.786	42.783	46.06	0.571	14445.333	7.602	11.238	0.939	2064.693
		NLOS	59.409	63.769	0.575	17445.598	55.457	60.432	0.653	17113.821	55.096	59.946	0.657	17057.787	7.109	11.353	1.079	2438.095
	73	LOS	48.243	50.227	0.535	15976.082	43.122	45.756	0.608	15672.252	42.882	45.535	0.611	15620.937	7.829	11.629	1.004	2232.724
		NLOS	48.347	50.996	0.541	16455.364	44.853	47.729	0.614	16142.419	45.955	48.779	0.617	16089.565	7.542	11.529	1.014	2299.706

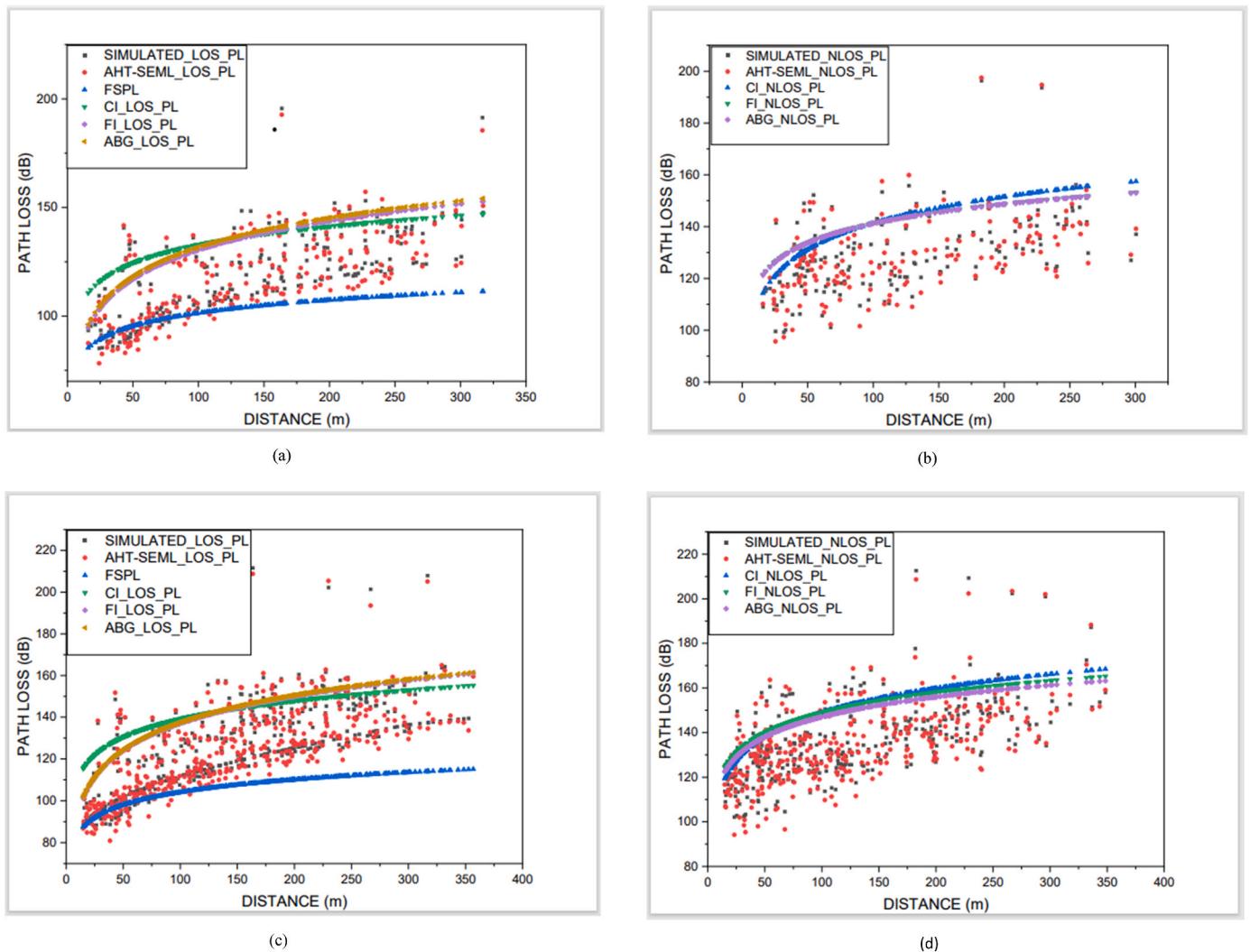


Fig. 1. Omnidirectional Path Loss for (a) the LOS condition at 28 GHz, (b) the NLOS condition at 28 GHz, (c) the LOS condition at 38 GHz, (d) the NLOS condition at 38 GHz, (e) the LOS condition at 60 GHz, (f) the NLOS condition at 60 GHz, (g) the LOS condition at 73 GHz, and (h) the NLOS condition at 73 GHz in Abuja City, respectively.

$$IoA = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (16)$$

$$BIC = N \times \log_{10} \left(\frac{\sum (PL_S - PL_P)^2}{N} \right) + k \times \log_{10}(N) \quad (17)$$

where N denotes the total number of samples, k represents the number of parameters in the model, PL_s is the simulated or empirical path loss in dB as the case may be, PL_p is the predicted path loss in dB, O_i is the observed value, P_i is the predicted value, \bar{O} is the mean of the observed values, and n is the number of data points.

Also, the models were compared using the Model confidence set (MCS) with CI path loss model as the benchmark to calculate the average mean absolute error (AVG MAE), standard deviation (STD) and the t-statistics across all cities for each model, frequency, and scenario. The absolute t-statistic value (t-statistic) of each model, frequency, and scenario was compared to the critical t-value with the chosen confidence level of 95% and degrees of freedom of 3 to determine if the difference is statistically significant. Models with t-statistics greater than the critical value are considered significantly different.

Generally, the choice of the CI model is recognised as the benchmark for calculating outdoor millimetre-wave path loss in 5G urban networks due to its simplicity, accuracy, stability, and widespread use [5]. It offers comparable goodness of fit in LOS and NLOS scenarios with just one parameter, ensuring stable behaviour across different frequencies and distances. Thereby, enhancing simulation accuracy and repeatability [5]. Research consistently finds it to be the best path loss model for outdoor millimetre-wave propagation, making it essential for reliable 5G system design and network planning [33,34].

$$STD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\text{AVG_MAE}_i - \text{AVG_MAE}_{CI})^2} \quad (18)$$

$$t = \frac{\text{AVG_MAE}_i - \text{AVG_MAE}_{Cl}}{\sqrt{\frac{(\text{STD}_i)^2}{n} + \frac{(\text{STD}_{Cl})^2}{n}}} \quad (19)$$

$$\text{Degree of freedom} = n - 1 \quad (20)$$

where n denotes total number of cities (4), i represents the model, STD_i , STD_{CI} , AVG_MAE_i and AVG_MAE_{CI} are the standard deviation and average mean absolute error of each model and CI model, respectively.

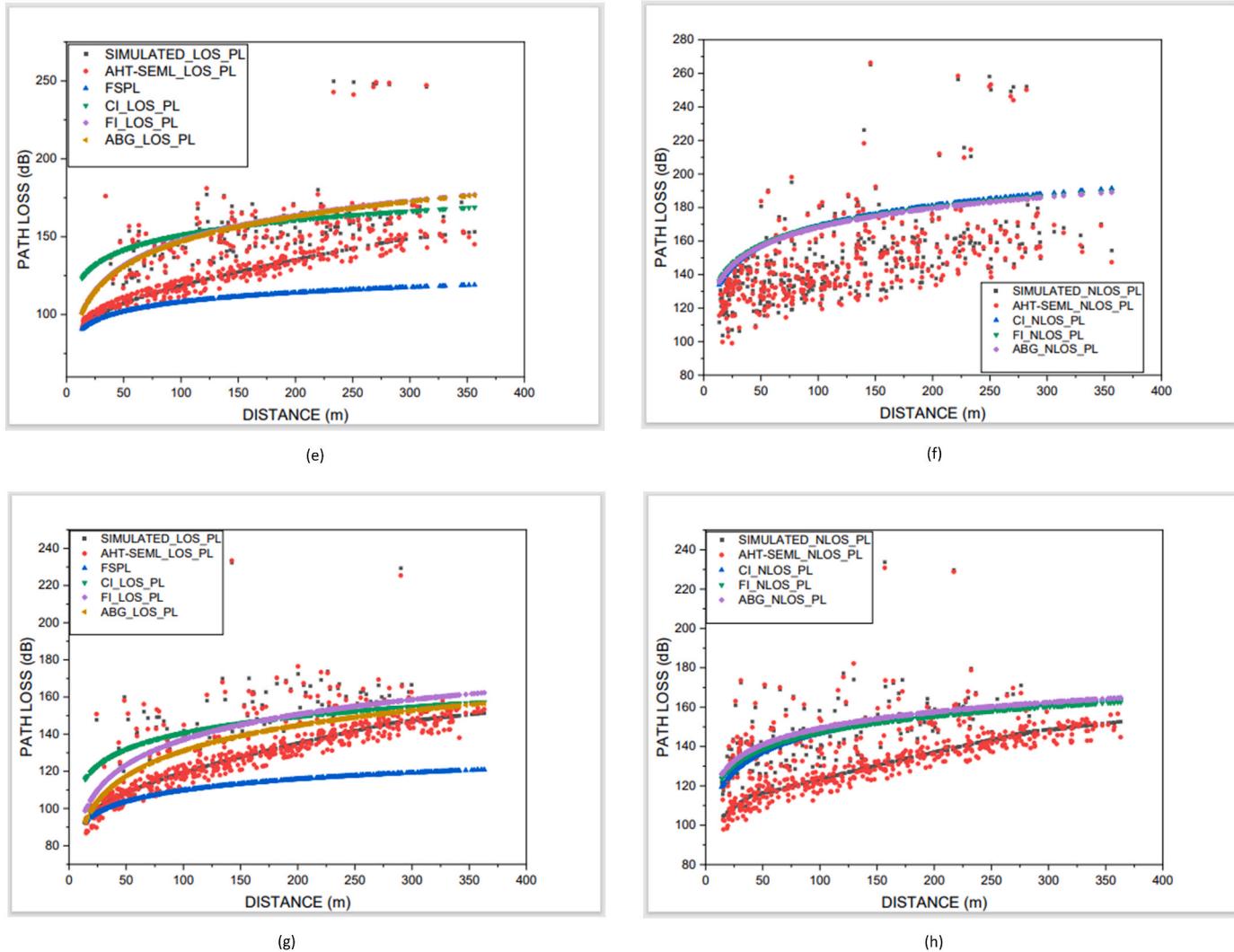


Fig. 1. (continued).

3. Result analysis

To come up with a robust machine learning model that will predict the path loss data with the best possible accuracy and consistency, ten different machine learning algorithms were tested as depicted in Table 2. The results showed significant variations in performance across frequencies and scenarios. The three ensemble models outperformed other single-based models. The Stacking Ensemble Model1 consists of five base regressors namely Linear Regression, Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, and Gradient Boosting Regressor, with the Xtreme Gradient Boosting Regressor as the meta-regressor. The Stacking Ensemble Model2 is made up of four base regressors, which are the Histogram Gradient Boosting Regressor, K Neighbours Regressor, Random Forest Regressor, and Gradient Boosting Regressor, with the Xtreme Gradient Boosting Regressor as the meta-regressor. The proposed AHT-SEML path loss model performed the best with the lowest MAE and RMSE. Among the single-base regression models, Xtreme Gradient Boosting Regressor performance is the best and closely followed by the Random Forest model. On the other hand, the linear regression and KN regression models showed the least performance.

Table 3 shows the performance metrics of four different path loss models: Close-In (CI), Floating Intercept (FI), Alpha-Beta-Gamma (ABG), and the Automated Hyperparameter-Tuning Stacking Ensemble Machine Learning (AHT-SEML) models. The models were tested across

four Nigerian cities: Abuja, Lagos Island, Ibadan, and Port Harcourt – at four different frequencies: 28, 38, 60, and 73 GHz. The results highlighted two different scenarios - LOS and NLOS - for each frequency.

3.1. Models comparison

The performance of the different models was evaluated based on four metrics: MAE, RMSE, IoA, and BIC. These metrics measure the deviation of the predicted values from the actual values, with a smaller value indicating better performance of the model.

The comparison of ten (10) Machine Learning models in Table 2 shows that only the AHT-SEML model has a combination of MAE and RMSE with the lowest values of 6.182 dB and 9.031 dB under LOS and NLOS scenarios in Abuja and Lagos Island at 38 GHz and 28 GHz, respectively.

Table 3 shows that for the CI path loss model, the lowest MAE (15.740 dB) and RMSE (17.525 dB) were recorded for the LOS condition at 73 GHz in Abuja City. The same model also yielded its lowest MAE (14.832 dB) and RMSE (17.283 dB) for the NLOS condition at 28 GHz in Abuja City.

The FI path loss model recorded the lowest MAE (13.659 dB) and RMSE (14.942 dB) for the LOS condition at 73 GHz in Abuja City. The same model yielded the lowest MAE (14.746 dB) and RMSE (17.068 dB) for the NLOS condition at 28 GHz.

For the ABG path loss model, the lowest MAE (9.404 dB) and RMSE

(11.571 dB) were recorded for the LOS condition at 73 GHz in Abuja City. The same model yielded the lowest MAE (14.430 dB) and RMSE (16.971 dB) for the NLOS condition at 38 GHz in Abuja City.

Table 3 results show that the Automated Hyperparameter-Tuning Stacking Ensemble Machine Learning (AHT-SEML) path loss model has the lowest MAE of 6.182 dB in Abuja City and RMSE of 9.074 dB in Lagos Island at 38 GHz for the LOS scenario. Similarly, the AHT-SEML recorded the lowest MAE of 6.118 dB at 38 GHz and RMSE of 9.031 dB at 28 GHz in Lagos Island for the NLOS scenario.

Moreover, **Table 3** also indicates that for both LOS and NLOS conditions in the four cities at 28, 38, 60, and 73 GHz, the MAE and RMSE values of the AHT-SEML in dB range from 6.118 to 7.981 and 9.031 to 11.629, respectively. In contrast, for the CI, FI and ABG path loss models, the MAE values in dB range from 14.832 to 59.409, 13.659 to 55.457, and 9.404 to 55.096, respectively. Additionally, their RMSE values in dB range from 17.283 to 63.769, 14.942 to 60.432, and 11.571 to 59.946, respectively.

The IoA and BIC of the AHT-SEML models were the highest and lowest, respectively, in comparison to the other empirical models.

3.2. Scenario analysis

(1) LOS scenario

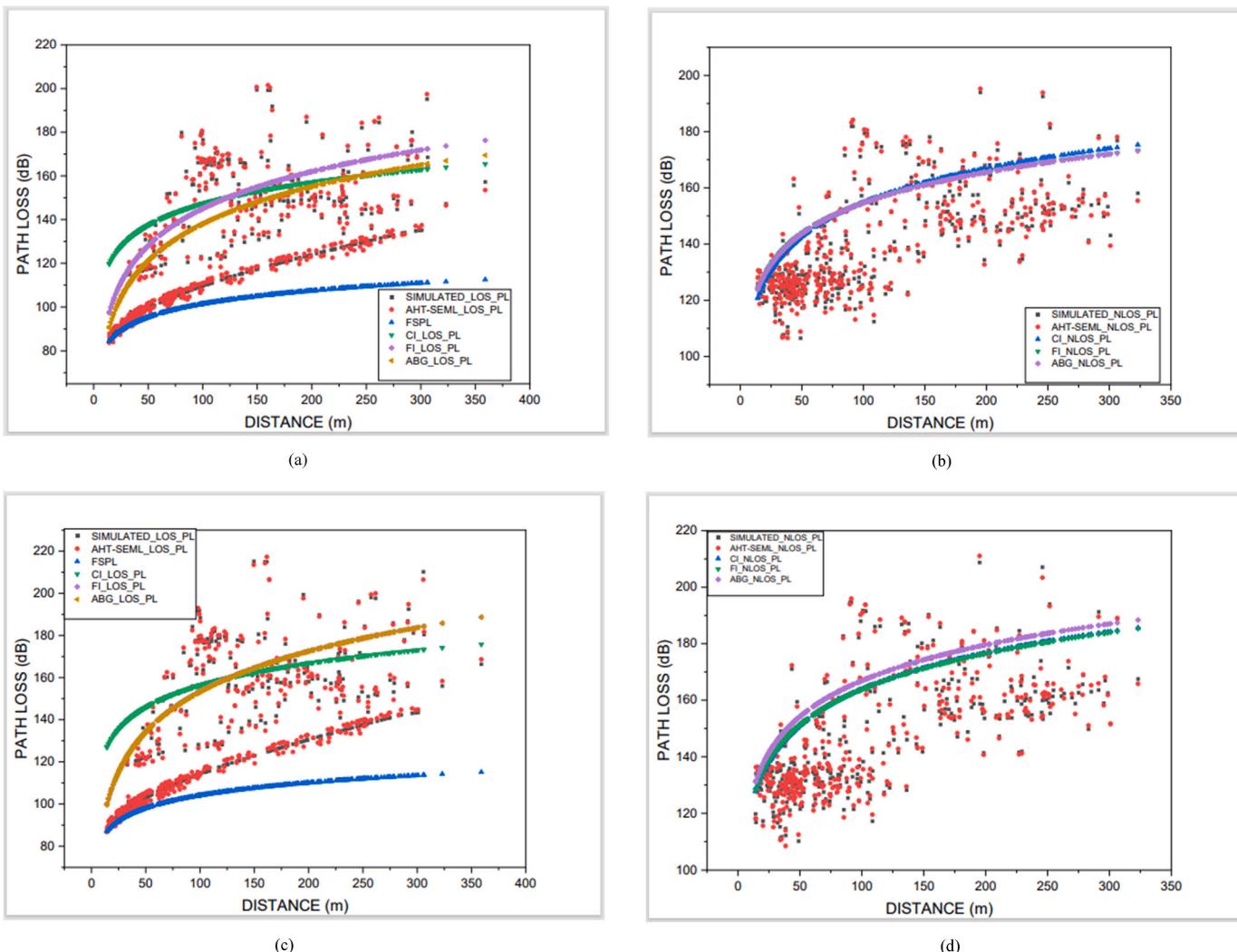


Fig. 2. Omnidirectional Path Loss for (a) the LOS condition at 28 GHz, (b) the NLOS condition at 28 GHz, (c) the LOS condition at 38 GHz, (d) the NLOS condition at 38 GHz, (e) the LOS condition at 60 GHz, (f) the NLOS condition at 60 GHz, (g) the LOS condition at 73 GHz, and (h) the NLOS condition at 73 GHz in Lagos Island, respectively

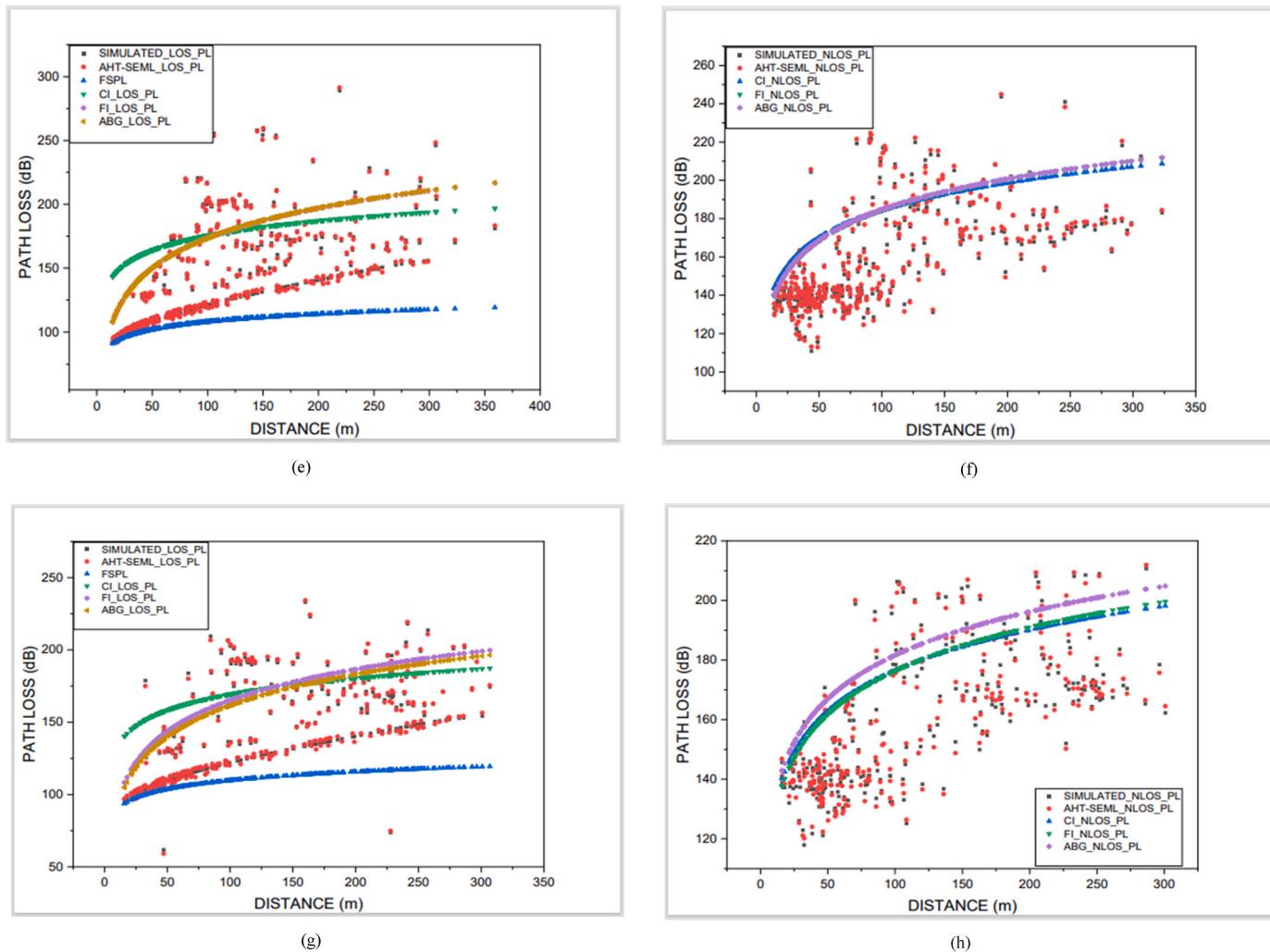


Fig. 2. (continued).

45.535 dB.

The AHT-SEML path loss model demonstrated consistent performance across various frequencies. At 28 GHz, the MAE values ranged from 6.404 to 6.778 dB, and RMSE values from 9.215 to 11.172 dB. At 38 GHz, the MAE values ranged from 6.182 to 7.559 dB, and RMSE values from 9.074 to 11.462 dB. At 60 GHz, the MAE values ranged from 6.772 to 7.961 dB, and RMSE values from 9.299 to 11.238 dB. Finally, at 73 GHz, the MAE values ranged from 6.488 to 7.829 dB, and RMSE values from 9.281 to 11.629 dB. This indicates the accuracy and stability of the AHT-SEML path loss model across different frequency bands.

(2) NLOS scenario

In the NLOS scenario, the performance of various path loss models, including the CI, FI, ABG, and AHT-SEML models, was evaluated across different frequencies using MAE and RMSE metrics.

The CI path loss model showed variable performance at different frequencies. For instance, at 28 GHz, the MAE values ranged from 14.832 to 52.218 dB, and the corresponding RMSE values ranged from 17.283 to 54.481 dB. At higher frequencies, the model's performance demonstrated similar variability. These results suggest that the accuracy and variability of the CI path loss model can differ significantly across frequencies.

The FI path loss model also exhibited varying performance across frequencies. At 28 GHz, the MAE values ranged from 14.746 to 50.230 dB, and the associated RMSE values ranged from 17.068 to 52.580 dB.

The model's performance showed similar patterns at higher frequencies. These findings highlight the fluctuating accuracy and variability of the FI path loss model at different frequencies.

The ABG path loss model's performance showed similar variability across frequencies. At 28 GHz, the MAE values ranged from 14.823 to 47.998 dB, and the RMSE values ranged from 17.139 to 50.501 dB. The model's performance showed similar variability at higher frequencies, highlighting the potential for fluctuating accuracy and variability at different frequencies.

In contrast, the AHT-SEML path loss model demonstrated consistent performance across frequencies. At 28 GHz, the MAE values ranged from 6.396 to 7.646 dB, and the RMSE values ranged from 9.031 to 11.350 dB. The model's performance showed similar consistency at higher frequencies, indicating the accuracy and stability of the AHT-SEML path loss model across different frequency bands.

Overall, the AHT-SEML path loss model showed the lowest deviation from simulated path loss values compared to the other models for the four frequencies (28, 38, 60, and 73 GHz) in both scenarios. This suggests that the machine learning model can accurately capture the effects of various parameters affecting path loss, such as distance, frequency, scenarios, and environmental conditions, and can provide more accurate signal strength predictions than the empirical models.

As shown in Figs. 1–4 for the AHT-SEML, CI, FI, and ABG path loss models, the omnidirectional path loss values are mostly higher for NLOS scenarios than LOS scenarios. This indicates that obstacles between the transmitter and receiver have a significant impact on signal propagation

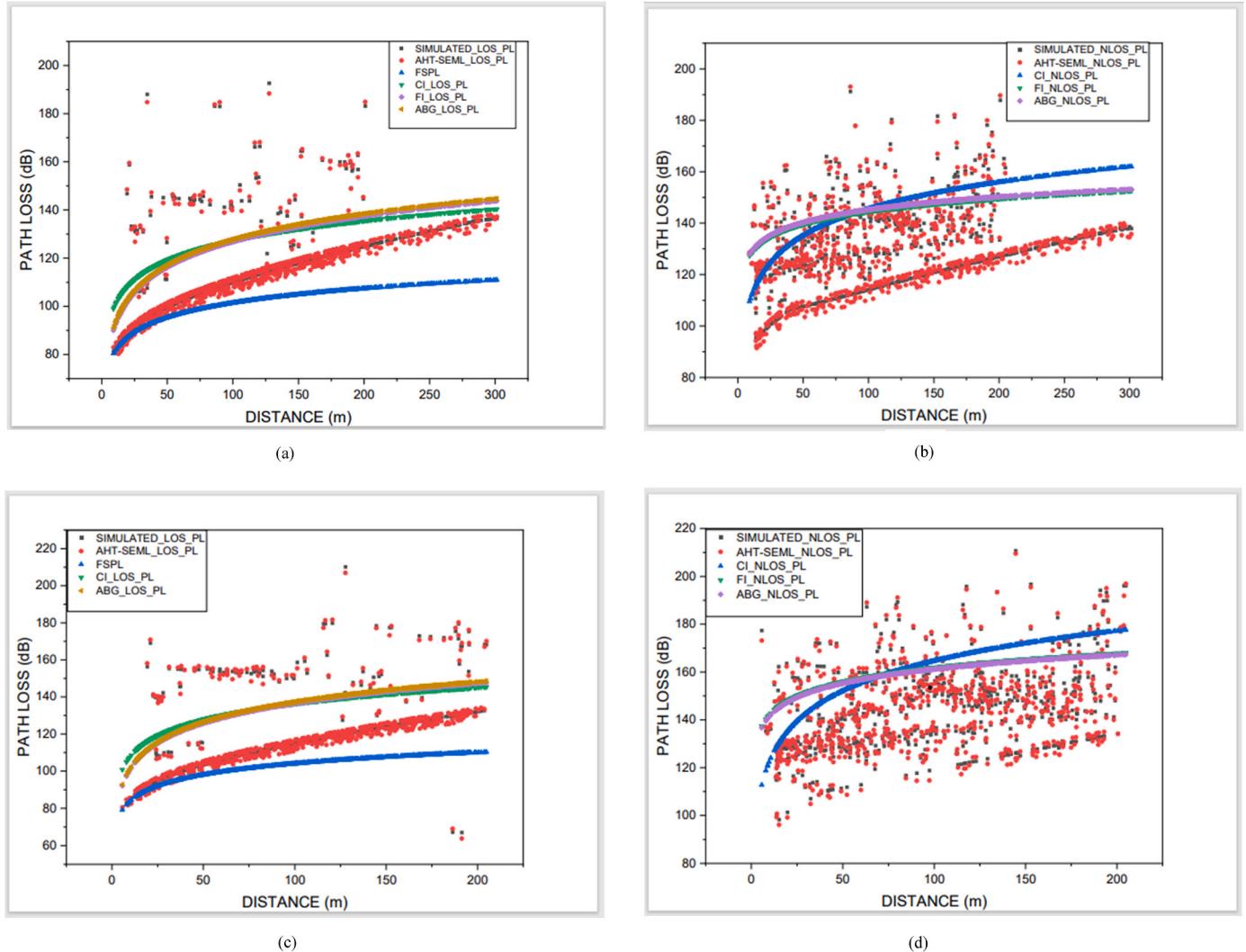


Fig. 3. Omnidirectional Path Loss for (a) the LOS condition at 28 GHz, (b) the NLOS condition at 28 GHz, (c) the LOS condition at 38 GHz, (d) the NLOS condition at 38 GHz, (e) the LOS condition at 60 GHz, (f) the NLOS condition at 60 GHz, (g) the LOS condition at 73 GHz, and (h) the NLOS condition at 73 GHz in Ibadan City, respectively.

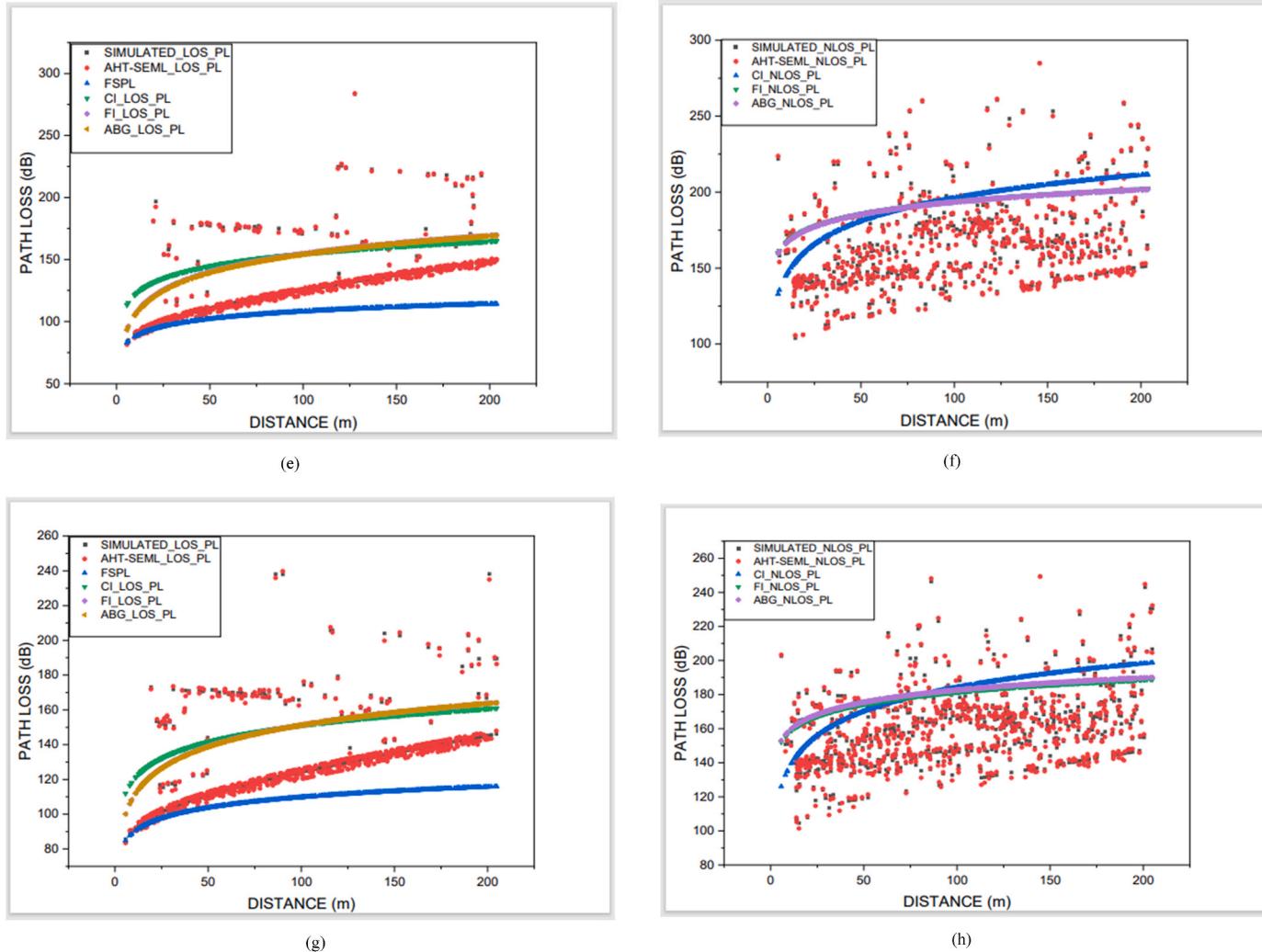


Fig. 3. (continued).

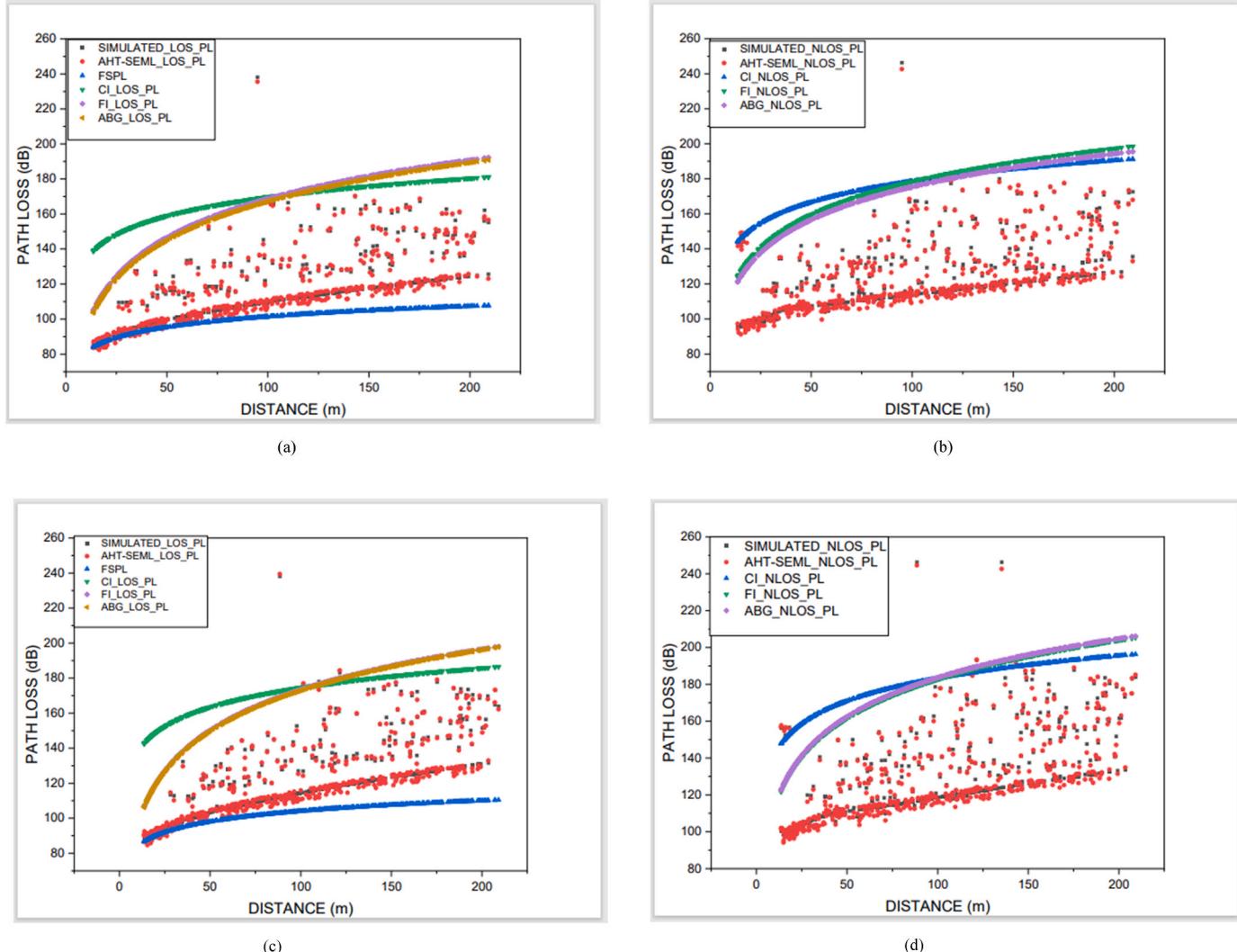
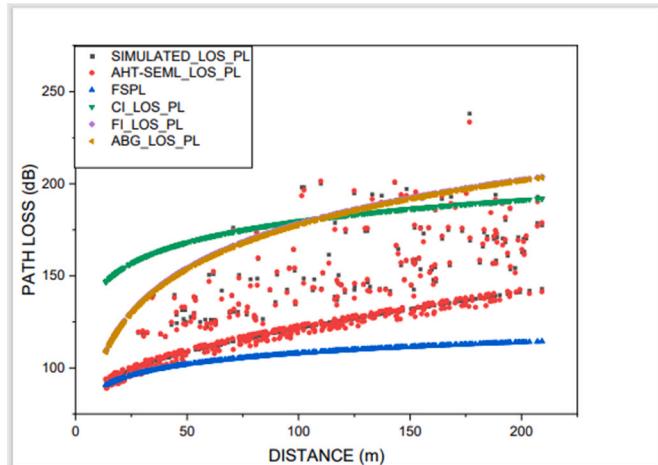
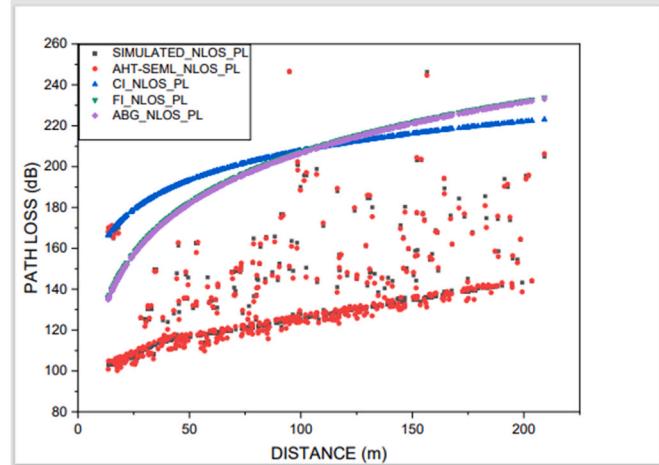


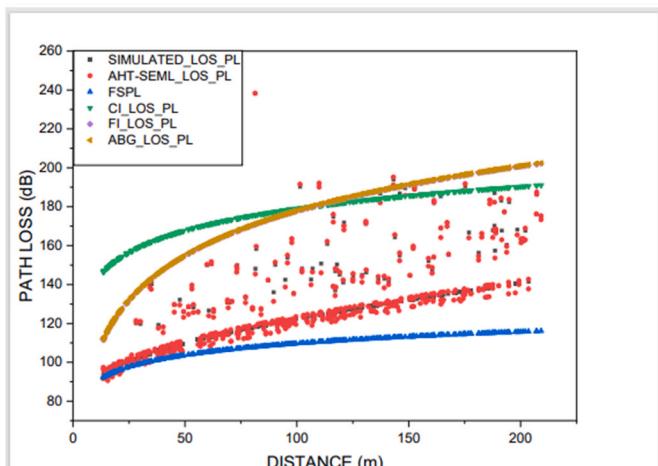
Fig. 4. Omnidirectional Path Loss for (a) the LOS condition at 28 GHz, (b) the NLOS condition at 28 GHz, (c) the LOS condition at 38 GHz, (d) the NLOS condition at 38 GHz, (e) the LOS condition at 60 GHz, (f) the NLOS condition at 60 GHz, (g) the LOS condition at 73 GHz, and (h) the NLOS condition at 73 GHz in Port Harcourt City, respectively.



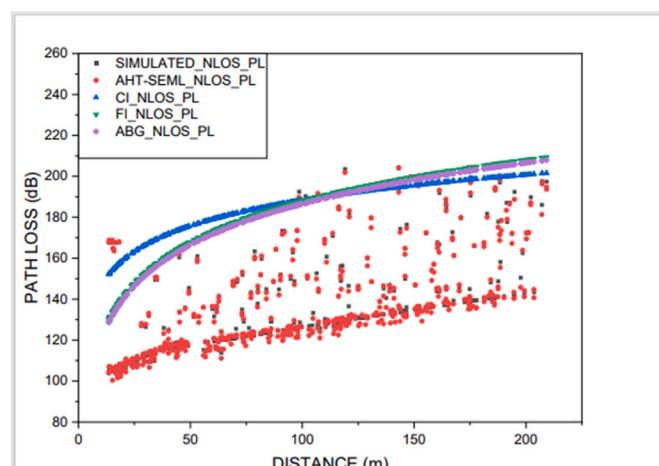
(e)



(f)



(g)



(h)

Fig. 4. (continued).

Table 4
Model confidence set analysis.

City	Freq (GHz)	Scenario	Close-In (CI) Path Loss Model		Floating Intercept (FI) Path Loss Model				ABG Path Loss Model	
			CI_AVG_MAE	CI_STD	FI_AVG_MAE	FI_STD	t - statistic	t - statistic (t_st)		
The Combined Cities	28	LOS	28.6545	16.6143	26.3543	14.4666	-0.2088	0.2088	t_st < t_crv	24.9860
		NLOS	25.2595	18.0328	24.5573	17.1489	-0.0564	0.0564	t_st < t_crv	24.1223
	38	LOS	30.8878	14.6671	28.0105	12.4304	-0.2993	0.2993	t_st < t_crv	28.3150
		NLOS	26.4170	17.0400	25.6015	15.7979	-0.0702	0.0702	t_st < t_crv	25.5095
	60	LOS	35.7423	11.6931	31.8530	9.7602	-0.5107	0.5107	t_st < t_crv	31.6445
		NLOS	35.7513	16.2085	34.6078	14.3654	-0.1056	0.1056	t_st < t_crv	34.3023
	73	LOS	30.8765	13.8270	27.9138	12.2515	-0.3207	0.3207	t_st < t_crv	26.2185
		NLOS	27.3600	14.4575	26.2345	12.8681	-0.1163	0.1163	t_st < t_crv	28.0343

across all four frequencies. Furthermore, the AHT-SEML model was found to be more accurate than the three empirical models at predicting path loss for all four frequencies and both LOS and NLOS conditions in the four cities considered. In contrast, the three empirical models tended to over-predict the path loss at all the frequencies and in both scenarios.

3.3. Complex analysis

This report presents a further comparative analysis of the path loss models using the IoA and BIC evaluation metrics to assess the accuracy and complexity of the models as also observed in [Table 3](#). The IoA assesses the agreement between predicted and observed values, while BIC accounts for the model's complexity and penalises overfitting.

The results show significant differences in model performance across scenarios and frequencies. In LOS scenarios, models generally exhibited higher IoA values and lower BIC values, indicating better agreement with observed data and model fit. However, in NLOS scenarios, the performance of the models varied, with some models demonstrating reduced accuracy and increased complexity.

The AHT-SEML model consistently outperformed the CI, FI and ABG models across all the scenarios, exhibiting the highest IoA values and lowest BIC values. This model demonstrated robustness in both LOS and NLOS conditions, highlighting its suitability for diverse propagation environments.

3.4. Statistical analysis

The findings from [Table 4](#) revealed that, except for the AHT-SEML model, all models exhibited absolute t-statistic values below the critical t-value of 2.3530. This suggests that the differences in MAE compared to the CI benchmark model are not statistically significant at the 95% confidence level. However, the AHT-SEML model stands out, with its absolute t-statistic values surpassing the critical value. This indicates a statistically significant difference, confirming that AHT-SEML performs significantly better than the CI benchmark model.

Importantly, while the FI, ABG, and AHT-SEML models all demonstrate lower average MAEs compared to the CI benchmark model, only the AHT-SEML model shows a statistically significant difference. This deduction is supported by the MCS analysis, which is conducted at a 95% confidence level with a degree of freedom of 3.

In addition, the results from one-way ANOVA statistical analysis conducted to compare the accuracies of the four path loss models using two metrics with a chosen significance level of 0.05: MAE and RMSE are as follows.

H0. there is no significant difference in the mean accuracies of the path loss models.

H1. there is at least one path loss model with a different mean

accuracy.

For the MAE, an F-statistic of 29.45 with a corresponding p-value of 1.91×10^{-14} was obtained. The low p-value suggests strong evidence against the null hypothesis, indicating that there is a significant difference in the mean accuracies among the path loss models based on MAE. This result provides initial confirmation that the performance of the models varies significantly.

Similarly, for the ANOVA results on RMSE, an F-statistic of 26.54 with a p-value of 2.50×10^{-13} was obtained. Again, the low p-value indicates a significant difference in the mean accuracies among the models based on RMSE. This finding corroborates the previous result obtained from MAE, further supporting the conclusion that there are significant differences in model performance.

4. Conclusion

This study presented a stacking ensemble-regression machine learning model with automated hyperparameter tuning (AHT-SEML) for accurate path loss prediction in mmWave systems. The proposed model combines various base regressors and a meta-regressor to increase prediction accuracy. Simulated path loss data from four Nigerian cities at different frequencies were used for model training and evaluation.

Three empirical path loss models and the AHT-SEML model's performance were compared. The results showed that the AHT-SEML model achieved the lowest MAE and RMSE values in both LOS and NLOS scenarios across the four cities (Abuja, Lagos, Ibadan, and Port Harcourt) and frequencies (28, 38, 60, and 73 GHz) considered. The AHT-SEML's MAE and RSME values in dB range from 6.118 to 7.981 and 9.031 to 11.629, respectively. For CI, FI and ABG path loss models, the MAE values, in dB, vary from 14.832 to 59.409, 13.659 to 55.457, and 9.404 to 55.096, respectively. Similarly, their RSME values in dB range from 17.283 to 63.769, 14.942 to 60.432, and 11.571 to 59.946, respectively. These indicate that the AHT-SEML model can accurately predict path loss in diverse conditions and frequencies.

Also, at the 95% confidence level with degree of freedom of 3, the MCS analysis supports the conclusion that while the FI, ABG, and AHT-SEML models all demonstrate lower average MAEs compared to the CI benchmark model, only the AHT-SEML model shows a statistically significant difference.

Moreover, the one-way ANOVA statistical analysis compared the accuracies of the four path loss models using their MAEs and RMSE. An F-statistic of 29.45 with a corresponding low p-value of 1.91×10^{-14} suggests strong evidence against the null hypothesis, indicating that there is a significant difference in the mean accuracies among the path loss models based on MAE.

Likewise, the ANOVA results from RMSE, having an F-statistic of 26.54 with a low p-value of 2.50×10^{-13} , indicate a significant

ABG Path Loss Model				AHT-SEML Path Loss Model					
ABG_STD	t - statistic	t - statistic (t_st)	t-statistic comparison with the critical t-value (t_crv = 0.23530)	AHT-SEML_AVG_MAE	AHT-SEML_STD	t - statistic	t - statistic (t_st)	Critical t-value (t_crv)	t-statistic comparison with the critical t-value
14.2695	-0.3350	0.3350	t_st < t_crv	6.5893	0.1876	-2.6560	2.6560	2.3530	t_st > t_crv
15.9661	-0.0944	0.0944	t_st < t_crv	6.9718	0.5190	-2.6274	2.6274	2.3530	t_st > t_crv
12.3340	-0.2685	0.2685	t_st < t_crv	7.1078	0.6275	-3.2397	3.2397	2.3530	t_st > t_crv
15.5871	-0.0786	0.0786	t_st < t_crv	6.8855	0.6608	-2.3907	2.3907	2.3530	t_st > t_crv
9.8462	-0.5361	0.5361	t_st < t_crv	7.4458	0.4979	-4.8355	4.8355	2.3530	t_st > t_crv
14.3425	-0.1339	0.1339	t_st < t_crv	7.0865	0.3360	-3.5362	3.5362	2.3530	t_st > t_crv
13.7323	-0.4780	0.4780	t_st < t_crv	7.1255	0.5724	-3.4325	3.4325	2.3530	t_st > t_crv
12.3728	-0.2191	0.2191	t_st < t_crv	7.3643	0.4953	-2.7645	2.7645	2.3530	t_st > t_crv

difference in the mean accuracies among the models based on RMSE. These findings corroborate the previous results obtained from MAE. Hence these results further support the deduction that there are statistically significant differences in the AHT-SEML model's performance.

These models' complexity and formulation impacted their ability to capture the underlying relationships between input variables and path loss. Also, the effectiveness of hyperparameter tuning influenced the AHT-SEML model's accuracy.

The findings underscore the importance of considering scenario-specific characteristics when evaluating path loss models. LOS scenarios typically offer better signal propagation conditions, leading to higher model accuracy and simplicity. In contrast, NLOS scenarios introduce additional challenges, such as signal blockage and multipath fading, which may require more sophisticated modelling approaches.

The comparative analysis highlights the influence of propagation scenarios on path loss model performance. Network planners can make informed decisions regarding model selection and parameterisation by considering scenario-specific factors and leveraging appropriate evaluation metrics.

The study emphasises the significance of accurate path loss prediction for wireless communication systems and shows how machine learning models enhance prediction accuracy.

Moreover, the study demonstrated that the selection of a path loss model should be approached with meticulous consideration, factoring in elements such as model complexity, frequency-dependent behaviour, LOS and NLOS scenarios, environmental circumstances, data excellence, and the specific type of model utilised.

5. Limitations and recommendations

While the suggested AHT-SEML model performs well in path loss prediction, there are a few drawbacks. The assessment is based on simulated path loss data from a particular propagation model, which may not fully depict real-world variations and complexities. Secondly, the study is limited to mmWave systems in certain Nigerian cities, and the model's application to other areas and environments requires more exploration. Furthermore, the ensemble model's base regressors and hyperparameters are chosen using automated techniques, which might introduce biases or limits in optimisation. More study and validation with real-world data and various scenarios are required to assess the model's resilience and applicability in practical scenarios. Also, future research may focus on refining model parameters and exploring novel modelling approaches to enhance accuracy and robustness in wireless communication systems.

CRediT authorship contribution statement

Johnson O. Afape: Writing – review & editing, Writing – original

draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Alexander A. Willoughby:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Modupe E. Sanyaolu:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Formal analysis. **Obiseye O. Obiyemi:** Writing – review & editing, Visualization, Validation, Resources, Investigation, Funding acquisition. **Katleho Moloi:** Writing – review & editing, Visualization, Validation, Resources, Investigation, Funding acquisition. **Janet O. Jooda:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation. **Oluropo F. Dairo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors thank Redeemer's University for the privilege given and the anonymous reviewers.

References

- [1] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G.K. Karagiannidis, E. Björnson, K. Yang, I. Chih-Lin, A. Ghosh, Millimeter wave communications for future Mobile networks, IEEE J. Sel. Area. Commun. 35 (9) (2017) 1909–1935, <https://doi.org/10.1109/JSAC.2017.2719924>.
- [2] A.N. Uwaechia, N.M. Mahyuddin, A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges, IEEE Access 8 (2020) 62367–62414, <https://doi.org/10.1109/ACCESS.2020.2984204>.
- [3] G.O. Ajayi, E.B.C. Ofoche, Some tropical rainfall rate characteristics at Ile-Ife for microwave and millimeter wave applications, J. Clim. Appl. Meteorol. 23 (4) (1984) 562–567, [https://doi.org/10.1175/1520-0450\(1984\)023<0562:STRCA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<0562:STRCA>2.0.CO;2).
- [4] K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li, T. Balercia, S. Lee, Y. Kim, A. Ghosh, T. Thomas, T. Nakamura, Y. Kakishima, T. Imai, H. Papadopoulos, A. Ghosh, 5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments, in: IEEE Vehicular

- Technology Conference, 2016-July, 2016, pp. 1–7, <https://doi.org/10.1109/VTCSpring.2016.7503971>.
- [5] S. Sun, T.S. Rappaport, S. Rangan, T.A. Thomas, A. Ghosh, I.Z. Kovacs, I. Rodriguez, O. Koymen, A. Partyka, J. Jarvelainen, Propagation path loss models for 5G urban micro-and macro-cellular scenarios, in: IEEE Vehicular Technology Conference, 2016-July, 2016, pp. 1–6, <https://doi.org/10.1109/VTCSpring.2016.7504435>.
 - [6] S. Hur, S. Baek, B. Kim, J. Park, A.F. Molisch, K. Haneda, M. Peter, 28 GHz channel modeling using 3D ray-tracing in urban environments, in: 2015 9th European Conference on Antennas and Propagation, EuCAP, 2015, pp. 1–5, 2015.
 - [7] R.D.A. Timoteo, D.C. Cunha, G.D.C. Cavalanti, A proposal for path loss prediction in urban environments using support vector regression, *Advanced International Conference on Telecommunications, AICT, 2014-July(July)*, 2014, pp. 119–124.
 - [8] T.S. Rappaport, G.R. MacCartney, M.K. Samimi, S. Sun, Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design, *IEEE Trans. Commun.* 63 (9) (2015) 3029–3056, <https://doi.org/10.1109/TCOMM.2015.2434384>.
 - [9] S. Salous, V. Degli Esposti, F. Fuschini, R.S. Thomae, R. Mueller, D. Dupleich, K. Haneda, J.M. Molina Garcia-Pardo, J. Pascual Garcia, D.P. Gaillot, S. Hur, M. Nekovee, Millimeter-wave propagation: characterisation and modeling toward fifth-generation systems. [Wireless corner], *IEEE Antenn. Propag. Mag.* 58 (6) (2016) 115–127, <https://doi.org/10.1109/MAP.2016.2609815>.
 - [10] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, T. Kürner, The design and applications of high-performance ray-tracing simulation platform for 5G and beyond wireless communications: a tutorial, *IEEE Communications Surveys and Tutorials* 21 (1) (2019) 10–27, <https://doi.org/10.1109/COMST.2018.2865724>.
 - [11] K.G. Lee, S.J. Oh, J.S. Woo, K.T. Lee, Propagation characteristics of suburban environments using hybrid ray-tracing simulation, in: IEEE Vehicular Technology Conference, 2016, pp. 1–5, <https://doi.org/10.1109/VTCFall.2016.7881181>, 0.
 - [12] O. Ahmadien, H.F. Ates, T. Baykas, B.K. Gunturk, Predicting path loss distribution of an area from satellite images using deep learning, *IEEE Access* 8 (2020) 64982–64991, <https://doi.org/10.1109/ACCESS.2020.2985929>.
 - [13] B. Xiang, D. Hu, J. Wu, Deep learning-based downlink channel estimation for FDD massive MIMO systems, *IEEE Wireless Communications Letters* 12 (4) (2023) 699–702, <https://doi.org/10.1109/LWC.2023.3240512>.
 - [14] Y. Yang, F. Gao, G.Y. Li, M. Jian, Deep learning-based downlink channel prediction for FDD massive MIMO system, *IEEE Commun. Lett.* 23 (11) (2019) 1994–1998.
 - [15] U. Masood, H. Farooq, A. Imran, A machine learning based 3D propagation model for intelligent future cellular networks, in: 2019 IEEE Global Communications Conference, GLOBECOM 2019 - Proceedings, 2019, pp. 1–6, <https://doi.org/10.1109/GLOBECOM38437.2019.9014187>.
 - [16] W. Jin, H. Kim, H. Lee, A novel machine learning scheme for mmWave path loss modeling for 5G communications in dense urban scenarios, *Electronics* 11 (12) (2022) 1809, <https://doi.org/10.3390/electronics11121809>.
 - [17] S.P. Sotirodakis, S.K. Goudos, K. Siakavara, Deep learning for radio propagation: using image-driven regression to estimate path loss in urban areas, *ICT Express* 6 (3) (2020) 160–165, <https://doi.org/10.1016/j.icte.2020.04.008>.
 - [18] Y. Zhang, J. Wen, G. Yang, Z. He, X. Luo, Air-to-Air path loss prediction based on machine learning methods in urban environments, *Wireless Commun. Mobile Comput.* (2018), <https://doi.org/10.1155/2018/8489326>, 2018.
 - [19] S.I. Popoola, E. Adetiba, A.A. Atayero, N. Faruk, C.T. Calafate, Optimal model for path loss predictions using feed-forward neural networks, *Cogent Engineering* 5 (1) (2018) 1444345, <https://doi.org/10.1080/23311916.2018.1444345>.
 - [20] Y. Zhang, J. Wen, G. Yang, Z. He, J. Wang, Path loss prediction based on machine learning: principle, method, and data expansion, *Appl. Sci.* 9 (9) (2019) 1908, <https://doi.org/10.3390/app9091908>.
 - [21] H. Cheng, S. Ma, H. Lee, CNN-based mmWave path loss modeling for fixed wireless access in suburban scenarios, *IEEE Antenn. Wireless Propag. Lett.* 19 (10) (2020) 1694–1698, <https://doi.org/10.1109/LAWP.2020.3014314>.
 - [22] C. Nguyen, A.A. Cheema, A deep neural network-based multi-frequency path loss prediction model from 0.8 ghz to 70 ghz, *Sensors* 21 (15) (2021) 5100, <https://doi.org/10.3390/s21155100>.
 - [23] S. Ojo, M. Akkaya, J.C. Sopuru, An ensemble machine learning approach for enhanced path loss predictions for 4G LTE wireless networks, *Int. J. Commun. Syst.* 35 (7) (2022) e5101, <https://doi.org/10.1002/dac.5101>.
 - [24] S.P. Sotirodakis, G. Athanasiadou, G.V. Tsoulos, C. Christodoulou, S.K. Goudos, Ensemble learning for 5G flying base station path loss modelling, in: 2022 16th European Conference on Antennas and Propagation, EuCAP 2022, 2022, pp. 1–4, <https://doi.org/10.23919/eucap53622.2022.9768903>.
 - [25] N. Moraitsis, L. Tsipi, D. Vouyioukas, A. Gkioni, S. Louvros, On the assessment of ensemble models for propagation loss forecasts in rural environments, *IEEE Wireless Communications Letters* 11 (5) (2022) 1097–1101, <https://doi.org/10.1109/LWC.2022.3157520>.
 - [26] O. Sagi, L. Rokach, Ensemble learning: a survey, *Wiley Interdisciplinary Reviews: Data Min. Knowl. Discov.* 8 (4) (2018) e1249.
 - [27] Z.-H. Zhou, Ensemble learning, in: *Machine Learning*, Springer, Singapore, 2021, pp. 181–210, https://doi.org/10.1007/978-981-15-1967-3_8.
 - [28] J. Sill, G. Takacs, L. Mackey, D. Lin, Feature-weighted linear stacking, *ArXiv Preprint ArXiv:0911.0460* (2009). <http://arxiv.org/abs/0911.0460>.
 - [29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *Scikit-learn: machine learning in {Python}*, *J. Mach. Learn. Res.* 12 (2011) 2825–2830.
 - [30] F. Qamar, N. Hindia, K. Dimyati, K.A. Noordin, M.B. Majed, T.A. Rahman, I. S. Amiri, Investigation of future 5g-iot millimeter-wave network performance at 38 GHz for urban microcell outdoor environment, *Electronics* 8 (5) (2019) 495, <https://doi.org/10.3390/electronics8050495>.
 - [31] T.T. Oladimeji, P. Kumar, M.K. Elmezoughi, Path loss measurements and model analysis in an indoor corridor environment at 28 GHz and 38 GHz, *Sensors* 22 (19) (2022) 7642.
 - [32] U.R. Kamboh, M.R. Shahid, H. Aldabbas, A. Rafiq, B. Alouffii, M.A. Habib, U. Ullah, Radio network forensic with mmWave using the dominant path algorithm, *Secur. Commun. Network.* (2022), <https://doi.org/10.1155/2022/9692892>, 2022.
 - [33] F.E. Shaibu, E.N. Onwuka, N. Salawu, S.S. Oyewobi, K. Djouani, A.M. Abu-Mahfouz, Performance of path loss models over mid-band and high-band channels for 5G communication networks: a review, *Future Internet* 15 (11) (2023) 362.
 - [34] A.I. Suliman, A.T. Nassar, M.K. Samimi, G.R. MacCartney, T.S. Rappaport, A. Alsanie, Radio propagation path loss models for 5G cellular networks in the 28 GHz and 38 GHz millimeter-wave bands, *IEEE Commun. Mag.* 52 (9) (2014) 78–86.
 - [35] MathWorks, Ray Tracing for Wireless Communications, Natick, Massachusetts: The MathWorks Inc. <https://www.mathworks.com/help/antenna/ug/ray-tracing-for-wireless-communications.html>, 2023.
 - [36] MathWorks, Urban Link and Coverage Analysis Using Ray Tracing, Natick, Massachusetts: The MathWorks Inc. <https://www.mathworks.com/help/comm/ug/urban-channel-link-analysis-and-visualization-using-ray-tracing.html>, 2023.
 - [37] S.I. Popoola, S. Misra, A.A. Atayero, Outdoor path loss predictions based on Extreme learning machine, *Wireless Pers. Commun.* 99 (1) (2018) 441–460, <https://doi.org/10.1007/s11277-017-5119-x>.