Journal of Sciences, Computing and Applied Engineering Research (JSCAER), Vol. 1, No.1, pp. 34-42

Published Online (https://jcaes.net) on May 7, 2025 by SciTech Network Press

Development of a Hybrid Machine Learning Path Loss Model for Cellular Networks in Maritime Environments Using Regression-Based Fusion

[1]Ugbeh R.N, [2]Okiemute Roberts Omasheye and [3]Abiodun I.C

[1]Department of Computer Engineering, Southern Delta State University, Ozoro, Delta State. Ugbehrn@dsust.edu.ng

[2]Department of Physics, Delta State College of Education, Mosogar

[3]Department of Physics, Federal University, Otuoke, Bayelsa State, abiodunic@fuotuoke.edu.ng

Received: 18 Feb 2025; Revised: 02 May 2025; Accepted: 06 May 2025; Published: 07 May 2025

Abstract: Accurate path loss modelling is critical for optimizing LTE network coverage and performance, particularly in maritime environments where conventional models often struggle to provide precise estimations due to dynamic environmental conditions. This study presents a hybrid path loss model that integrates machine learning (ML) with traditional empirical models using a regression-based fusion technique. The proposed approach enhances prediction accuracy by dynamically adjusting model coefficients based on real-time environmental factors. A case study of the Escravos water in Delta State, Nigeria, demonstrates the model's effectiveness compared to standard empirical models such as COST-231 and Okumura-Hata. The results indicate significant improvements in prediction accuracy, with a reduced mean squared error (MSE) and enhanced adaptability to environmental variations. The findings suggest that the proposed decision tree particle swarm optimization (DT-PSO-COST231) path loss model can serve as a valuable tool for LTE network planning in maritime environments.

Keywords: Path loss model, LTE, machine learning, regression-based fusion, maritime environments, Escravos.

1. Introduction

Reliable LTE network coverage is vital in maritime environments, as it ensures effective communication, accurate navigation, and prompt emergency operations. However, signal propagation over the open sea can be unpredictable, influenced by various factors such as multipath fading, atmospheric ducting, and humidity-induced attenuation . Conventional path loss models, such as Okumura-Hata and COST-231, provide baseline estimates for terrestrial scenarios but often fall short in maritime situations due to the unique signal behaviours that occur over water surfaces. (Akinyemi et al., 2023).

The dynamics of signal propagation over water differ significantly from those on land due to the reflective properties of water surfaces, which enhance multipath effects and lead to fluctuations in the received signal strength. Additionally, atmospheric conditions such as temperature inversions can extend radio signal ranges beyond typical diffraction limits, making predictions more complex (ITU-R P.452-16, 2015). Existing models, such as Okumura-Hata and COST-231, do not fully account for these complexities, resulting in inaccuracies in maritime environments. (Wang et al.,2015). Several models, including Okumura-Hata, COST-231, and ITU-R, have been developed to predict signal loss in both land and sea environments. However, these models often overlook specific variations that occur in maritime settings. Recent studies have investigated machine learning (ML) techniques to enhance prediction accuracy. Relying solely on ML models may limit their general applicability. Combining ML with empirical models could lead to more precise predictions (Brekke et al., 2018).

Hybrid models that integrate empirical models with ML-based adjustments have shown better results. These models dynamically adjust their predictive weights based on changing environmental conditions, leveraging the robustness of empirical models and the adaptability of ML techniques (Andreas et al., 2012). In this study, we introduce a hybrid path loss model that combines ML-driven predictions with empirical models using a regression-based fusion technique. Our approach enhances LTE path loss estimations for maritime environments.

This research focuses on Escravos, Delta State, Nigeria, as a case study to validate the effectiveness of our hybrid model. By comparing it to traditional models, we aim to demonstrate the advantages of incorporating machine learning techniques for achieving more accurate signal predictions in maritime conditions. To ensure accurate measurements, we employed state-of-the-art equipment, including: a Weather Station (WS) is set up to monitor various meteorological parameters, including temperature, humidity, wind speed, and atmospheric pressure. A Global Positioning System (GPS) provides accurate geolocation data. The setup includes a Samsung Galaxy S5 smartphone, which features a built-in Transmission Evaluation Monitoring System (TEM POCKET) for assessing mobile signal quality. Additionally, TEMS 15.2.2 investigation software is installed on a Core i5 Dell laptop for comprehensive data analysis. Both the weather parameter and the signal measurement were taken simultaneously. Field measurements were conducted during both wet and dry seasons, capturing environmental variations affecting signal transmission. (Doggett et al., 2017)

This research improves path loss modelling by combining machine learning techniques with empirical models, enhancing adaptability to real-time environmental changes. The findings aim to strengthen network planning strategies for LTE coverage in maritime environments, providing a more reliable framework for telecommunications engineers and policymakers.

1.1 Literature Review

The literature review offers a thorough overview of current research on path loss models in maritime environments, critically assessing the methodologies used and their limitations. By examining the strengths and weaknesses of various approaches, this section emphasizes the necessity for a more robust predictive model that integrates both machine learning and empirical techniques.

1.1.1 Empirical and Semi-Deterministic Models for Path Loss Prediction

Several path loss models have been developed for both terrestrial and maritime environments. Zhu et al. (2010) combined experimental data with theoretical modelling to create a tool for predicting radio wave propagation loss at sea. Their findings indicate that as the distance exceeds the line of sight (LoS), the model's accuracy decreases. This highlights the necessity for a predictive model that effectively accounts for both LoS and non-line-of-sight (NLoS) conditions in maritime environments.Wang et al. (2023) proposed a semi-deterministic model that integrates the ITU path loss model with geometric calculations to address signal reflections and diffractions. While this model effectively predicts transitions between LoS and NLoS conditions, it does not take atmospheric parameters into account. Despite this limitation, validation against channel measurement data demonstrated its accuracy in harbour scenarios.

1.1.2 Machine Learning Approaches to Path Loss Modelling

Recent studies have investigated machine learning (ML) techniques to enhance the accuracy of path loss predictions. Shen et al. (2022) developed a neural network model trained on simulated data to predict maritime path loss. This model incorporates parameters such as wind speed and antenna height to improve its accuracy. However, key atmospheric factors—such as temperature, humidity, and seasonal variations were not considered. Despite these limitations, the proposed backpropagation (BP) neural network model achieved a high coefficient of determination ($R^2 > 0.92$), significantly outperforming traditional models. This indicates the potential of ML to refine path loss predictions. In another study, Ahmed et al. (2024) introduced a novel data augmentation method to enhance ML-based path loss predictions. Their approach combined synthetic data from a cellular coverage simulator with real-world datasets collected from various environments, including farms and residential areas. By engineering channel features using Lidar-derived geographical attributes and employing a gradient-boosting algorithm, they significantly reduced the mean absolute error by approximately 12 db. These findings suggest that incorporating synthetic data can improve the adaptability and performance of ML models across different scenarios.

1.1.3 Hybrid Path Loss Models and Adaptive Techniques

Imoize et al. (2021) studied an adaptive path loss model for LTE mobile broadband networks operating at 2630 MHz, utilizing TEMS investigation tools. They improved the traditional Egli model by incorporating the Levenberg-Marquardt algorithm to tackle nonlinear propagation challenges. The enhanced model outperformed the standard version in various test locations, showcasing the effectiveness of adaptive modelling techniques. This literature review highlights the limitations of traditional path loss models and the benefits of using machine learning (ML) approaches. While ML methods have shown improved accuracy, their reliance on large datasets and the need for specific environmental conditions limit their effectiveness when used alone. To address these limitations, this study proposes the development of a hybrid path loss model that combines empirical models with ML-driven optimizations. The goal is to enhance prediction accuracy in maritime LTE networks and fill the existing gaps in the literature.

1.2 THEORETICAL BACKGROUND

1.2.1 Okumura-Hata Model

The COST-231 is an empirical model extended from the Okumura-Hata model for higher frequency scenarios, and it is expressed thus: (Isabona et al., 2021).

$$PL = 69.55 + 26.16log10(f_c) - 13.82log(h_{b1}) - (ahm) +$$

$$(44.9 - 6.55\log(h_{b1}))log10(x) \quad (1)$$

$$\square h_{\square}=3.2(\log(11.75h_{\square}))^2-4.97 \quad (2)$$

where,
PL=pathloss in decibel(dB)

f=frequency in Gigahertz (GHz), h_b=Base station antenna height in meters(m) h_m=Mobile station antenna height in meters(m).
d= Distance between transmitter and receiver antennas in meters(m), ah_m = mobile station antenna correction height.

1.2.2 Cost231 Model

The COST 231 model uses empirical and deterministic methods for path loss estimation :

$$PL = 43.6 + 33.9log10(f_c) + \left(44.9 - 6.55log10(h_{b1})\right)\log(X) - 13.82log10(h_{b1}) - ahm - 2\left(log10\left(\frac{f_c}{28}\right)\right)2 - C_m) \quad (3)$$

where,

PL=pathloss in decibel(dB) f = frequency in (MHz)

X = Distance between transmitter and receiver antennas in meters (m) h_b = Base station antenna height in meters (m) h_m=Mobile station antenna height in meters (m)=1.5m
C_m =5.4dB for the maritime environment ah_m= ahm = $3.2(\log10(11.75hm))^2$- 4.97 (Mobile antenna height correction factor)

1.2.3 ITU-R M2414 MODEL

The ITU-R M.2414 model is a key standard in telecommunications for predicting path loss, especially in 5G networks. It is part of the International Telecommunication Union's recommendations and ensures reliable communication in diverse scenarios such as Urban Macro, Urban Micro, indoor hotspots, and maritime environments. The model accounts for environmental factors like Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions, offering realistic signal propagation insights. It covers a broad frequency range from 100 MHz to 100 GHz, which is essential for millimetre-wave communications in 5G. Under the ITU-R M.2414 model, the primary path loss (PL) formula is expressed as:

$$PL = 22.7 + 26log10(f_c) + 36.7log10(X) - C$$
(4)

where:

X, f, and C represents the distance between the transmitter and receiver in meter, the frequency in MHz and C accounts for a correction factor that accommodates specific environmental conditions

ITU-R M.2414 is a comprehensive model incorporating various real-world factors into path loss predictions, making it a cornerstone in designing next-generation wireless communication systems.

From the measured RSRP, signal loss values were calculated using (Rappaport, 2002 & Seybold, 2005):

$$PL(dB) = EIRP(dBm) - RSRP(dBm) \tag{5}$$

where EIRP is the adequate isotropic radiated power expressed as;

$$EIRP = P_t + G_r + G_t - L_r - L_t \tag{6}$$

where □□ and □□ are the receiver and transmitter antenna gains, respectively, □□ and □□ are transmitter and receiver cable losses in dB, and P_t is transmitter power.

Then Eqn (5) is expressed as

$$PL = P_t + G_r + G_t - L_t - L_r(dBm) - RSRP(dBm) \tag{7}$$

1.3 Model Formulation:

The COST 231-Hata model is divided into offset (P1), system (P2) parameter, and the slope (P3).

Hence, from equation 3 the formulation of the adaptive model for the Cost231 model was deduced as follows:

The COST-231-Hata model is initially computed using the given formulas:

Offset parameters, P1 = 46.3 − 13.82 log(ht) − □(hr) + □□ (8)

Slope of the model curve, P2 = [44.90 − 6.550 log(ht)]log(d) (9)

System design parameter, P3 = 33.9 log(□) (10)

Therefore;

PLcost (dB) = P1 + P2log (d) + P3 log (f) (11)

where ht, a(hr) and Cm indicate the base station height, mobile station correction factor, and the environmental correction factor, with P1, P2, and P3 parameters typically fixed in COST-231.

Gradient Boosting Regression (GBR) is employed to optimize the fixed parameters of the COST231 model.

The objective of Gradient Boosting Regression (GBR) is to identify the optimal values of P1, P2, and P3 that minimize the discrepancy between the measured path loss (PL measured) and the predicted path loss. Instead of using fixed values for these parameters, we enable GBR to obtain a new value which is most effective for the terrain. This is achieved through the following approach: The GBR model learns the correlation between measured path loss and COST-231 predictions. Each new tree added to the gradient boosting model is designed to mitigate residual errors by dynamically updating the parameters. The GBR fine-tunes the weight of each parameter to improve overall accuracy.

1.3.1 Fusion Model Equation

The fusion model modifies COST-231 parameters by integrating decision tree particle Swarm Optimization (DT-PSO) predictions. This is achieved through a regression model that minimizes the error between the predicted and measured path loss which in turn optimize the COST231 parameters.

$$PL(fused) = w1PL(COST231) + w2PL(DT - PSO) + b \quad (12)$$

where:

$PL_{(fused)}$, $PL_{(COST231)}$ and $PL_{(DT-PSO}$ difine the Final fused path loss., Path loss predicted by COST-231 and the Path loss predicted by DT-PSO.

The fusion model equation after optimization becomes:

$$PL(optimized) = P_1^1 + P_2^1 \log(d) + P_3^1 \log(f)$$
(13)

where: P_1^1, P_2^1, P_3^1 are the optimized adjustable parameters learned from GBR.

MATERIALS AND METHODS

2.1 **Experimental Site and Data Collection.**

The study was conducted in Delta State, Nigeria, focusing on the Escravos waters. A total of eighty transmission stations were surveyed for signal strength across four maritime locations: the Okerenghigho River, Opuraja River, Ogidigben River, and Escravos River. The examined environments included linear settlements, freshwater and saltwater bodies, shipyards, islands, and dense mangrove forests. The survey spanned 22.46 NM (42.34 km) in the Old Forcados River, covering an area of 1,542 km², and was bounded by the coordinates 5°00.151'N, 5°45'E and 5°04.5'N, 5°15'E. The significance of these locations is underscored by their economic value and distinctive geographical conditions, which may affect network performance.

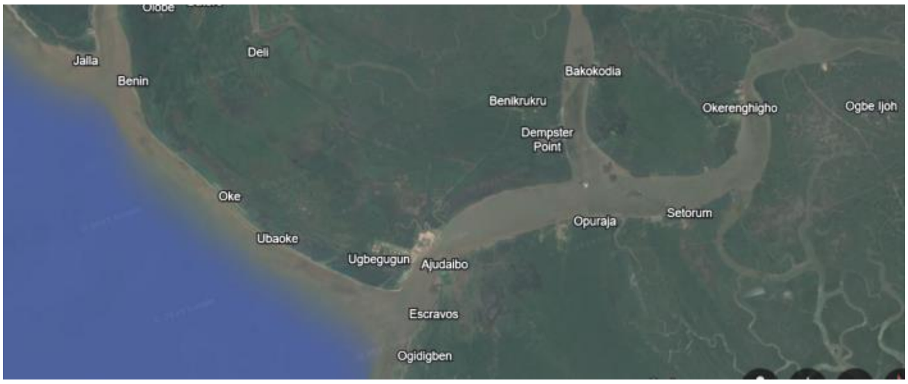

Fig. 1 Map of the surveyed Areas (Landsat Copernicus data, 2015)

2.2 Methodology

The survey collected data from July 2021 to November 2023, measuring a signal transmitted at 800 MHz using LTE 4G technology provided by AIRTEL Nigeria. A handheld weather station and data from the Nigerian Meteorological Agency (NIMET) were used to monitor atmospheric parameters like temperature, humidity, wind speed, and pressure. RSRP measurements were taken at 0.02 km intervals with TEM pocket equipment simultaneously with the weather parameters while travelling at an average speed of 20 km/h on a speedboat. Data was transferred to a laptop pre-installed with Transmission Evaluation Monitoring System (TEMS) software version 15.2.2. Simulations utilized the conventional radio-wave propagation model (Okumura-Hatta), ITU, and the proposed Hybrid Machine Language Algorithm (HMLA), which combines the Decision Tree Algorithm, Particle Swarm Optimization, and the Cost231 model using the Gradient Boosting Regression (GBR). The Cost231 was choosing based on its best RMSE prediction against other tested models. The machine learning technique builds an ensemble of weak learners (decision trees) to correct prediction errors progressively. It is well-suited for fusing the DT-PSO model with COST-231 to enhance path loss prediction accuracy. A link budget computation and performance assessment were performed using first-order statistical indicators; Mean Absolute

Error (MAE), Standard Deviation (STD), MEAN Absolute Percentage Error (MAPE), Root Mean Square (RMSE) and the coefficient of determination (R^2).

3.Results

Table1: Optimized Cost231 parameters from the Research Locations

Location	P^1_1	P^1_2	P^1_3
Ogidigben River	549.78	24.42	-141.96
Okerenghigho River	501.72	29.52	-127.47
Escravos River	179.66	26.34	-16.57
Opuraja River	-150.62	28.77	92.97
Mean	**270.14**	**27.26**	**-48.26**

From equation (13);

$$PL\ (dB) = \mathbf{P^1_1} + \mathbf{P^1_2}\log(d) + \mathbf{P^1_3}\log(f)$$

Therefore,

$$PL(optimized)\ (dB) = 270.14+27.26\log(d)-48.26(fc) \quad (14)$$

Table 2: Measured and Predicted Pathloss values.

Km	COST231 dBm	OKUMURA dBm	ITU-RM2414 dBm	MEASURED dBm	DT-PSO-COST231 Model dBm
0.32	99.67614	86.556578	79.880	119.0642	133.949512
0.34	101.2072	88.087651	68.815	119.0642	134.675401
0.36	102.5989	89.479333	87.846	119.0642	135.359787
0.38	103.8745	90.754902	66.959	119.0642	136.007161
0.40	105.0518	91.932249	86.139	119.0642	136.621321
0.42	106.1450	93.025425	85.379	120.9705	137.205510
0.44	107.1652	94.045667	84.669	127.1618	137.762517
0.46	108.1217	95.002101	84.003	130.1611	138.294760
0.46	109.0218	95.902241	83.377	129.2163	138.294760
0.48	109.8719	96.752346	72.785	129.2163	138.804348
0.50	110.6772	97.557688	72.225	131.7452	139.293130
0.52	111.4423	98.322745	61.693	132.7440	139.762739
0.54	112.1709	99.051357	61.186	132.7440	140.214623
0.56	112.8664	99.746839	60.702	133.0028	140.650071
0.58	113.5316	100.41207	70.239	133.0028	141.070237
0.60	114.1691	101.04958	99.795	126.3327	141.476157
0.62	114.7811	101.66158	89.370	126.3327	141.868766
0.64	115.3696	102.25004	99.960	126.3327	142.248909
0.66	115.9363	102.81669	98.566	131.0002	142.617353
0.68	116.4827	103.36311	88.186	143.2997	142.974798
0.70	115.0102	103.89067	77.819	132.7837	143.321880

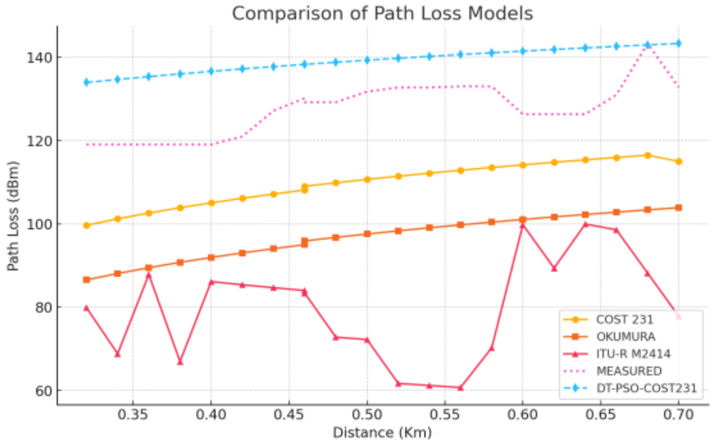

Fig 2: proposed hybrid model compared with the conventional models

Table 3: Precision Fitting Table for All Models

KPI	COST-231	OKUMURA	ITU-R-M2414	DT-PSO-COST231
MAE (Mean Absolute Error**)**	17.91 dB	30.94 dB	47.70 dB	**11.51 dB**
STD (Standard Deviation)	3.84	3.87	14.29	**4.45**
MAPE (Mean Absolute Percentage Error)	13.98%	24.21%	37.15%	**9.20%**
RMSE (Root Mean Square Error)	18.32 dB	31.18 dB	49.80 dB	**12.34 dB**
Coefficient of Determination (R^2)	0.943836	0.894175	0.827348	0.99887

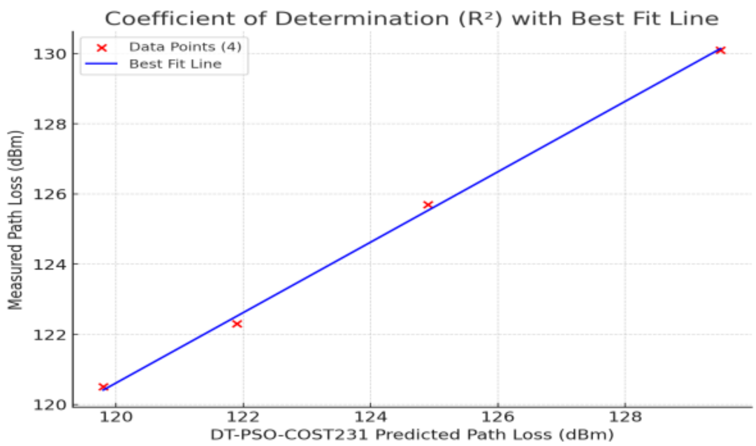

Fig. 3: Best fit line showing a correlation between path loss of the proposed model and the measured value.

4.1 Discussion

The results shown in Table 3 provide a detailed comparison of the performance of various path loss models, specifically focusing on COST-231, Okumura, ITU-R M2414, and the proposed hybrid model (DT-PSO-COST231). It was found that the hybrid DT-PSO-COST231 model exhibits the lowest error rates across all metrics (MAE: 11.51 dB, RMSE: 12.34 dB, MAPE: 9.20%, and R^2: 0.99887), indicating that it offers the best approximation to the measured values. The COST-231 model ranks second but still has higher errors than DT-PSO-COST231 (MAE: 17.91 dB, RMSE: 18.32 dB, R^2: 0.943836). In contrast, the ITU-R M2414 model shows the highest error rates, making it the least reliable option in this dataset, with an MAE of 47.70 dB and an RMSE of 49.80 dB. This observation aligns with findings in the literature; Zhu et al. (2010) noted that standard models face difficulties predicting path loss in non-line-of-sight (NLoS) conditions, similar to the elevated errors seen in the ITU-R M2414 in this study. Additionally, research by Shen et al. (2022) has demonstrated that machine-learning models can outperform empirical models by adapting to varying environmental conditions. The DT-PSO-COST231 model adheres to this trend, showcasing enhanced accuracy. Furthermore, the results from Ahmed et al. (2024) support the notion that combining machine learning techniques with existing models significantly improves prediction accuracy, reinforcing the integration of DT-PSO with COST-231 in this study.

5. Conclusion

This study evaluated the effectiveness of conventional path loss models, including COST-231, Okumura, and ITU-R M2414, in predicting LTE signal propagation in maritime environments. The results showed that these models are not accurate enough for this application due to the complex and dynamic nature of maritime settings. To address these limitations, the study introduced a hybrid model called DT-PSO-COST231, which combines machine learning techniques with empirical models to improve prediction accuracy. The DT-PSO-COST231 model significantly enhances the accuracy of path loss predictions compared to traditional models. It achieved lower Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) values, demonstrating superior performance in maritime network planning. These results confirm that integrating machine learning with empirical path loss models provides better adaptability to varying environmental conditions, leading to more reliable LTE coverage predictions in maritime settings. Integrating machine learning into path loss modeling presents new opportunities for optimizing LTE network planning. The proposed model enhances real-time adaptability, which is crucial for maritime communication systems. The research findings can assist telecommunications engineers and network operators in creating more accurate coverage models for coastal and offshore environments. Ultimately, this advancement will improve connectivity for maritime operations, emergency response, and navigation. By increasing the accuracy of path loss predictions, this study contributes to the efficient deployment of LTE networks in maritime regions, helping to reduce costs related to ineffective network planning and signal loss. Additionally, the hybrid model can be adapted for future wireless communication technologies, including 5G and beyond, making it relevant for evolving network infrastructures.

REFERENCES

1. Ahmad, I., Khan, A. R., Jabbar, A., Alquraan, M., Mohjazi, L., Rehman, M. U., ... & Hussain, S. (2024). Proactive blockage prediction for UAV-assisted handover in future wireless networks. *arXiv preprint* arXiv:2402.04332. https://arxiv.org/abs/2402.04332

2. Akinyemi, L. A. (2019). Towards developing path loss models for dryland and wetland environments. *2019 AFRICON*. IEEE. https://doi.org/10.1109/AFRICON46755.2019.9134041

3. Bakare, B. I., & Ekolama, M. S. (2023). Application of artificial intelligence (AI) to GSM operations. *European Journal of Science, Innovation and Technology, 3*(6), 482–495.

4. Doggett, M. K. (1997). *An atmospheric sensitivity and validation study of the variable terrain radio parabolic equation model (VTRPE)* [Unpublished doctoral dissertation]. Institution not specified.

5. Imoize, A. L., & Ogunfuwa, T. E. (2019). Propagation measurements of a 4G LTE network in a lagoon environment. *Nigerian Journal of Technological Development*, *16*(1), 1–9. https://doi.org/10.4314/njtd.v16i1.1

6. Isabona, J., & Imoize, A. L. (2021). Terrain-based adaptation of propagation model loss parameters using non-linear square regression. *Journal of Engineering and Applied Science, 68*, Article 33. https://doi.org/10.1186/s44147-021-00033-1

7. Ishimaru, A., Jaruwatanadilok, S., Ritcey, J. A., & Kuga, Y. (2009). A MIMO propagation channel model in a random medium. *IEEE Transactions on Antennas and Propagation, 58*(1), 178–186. https://doi.org/10.1109/TAP.2009.2034582

8. Lu, J., Zhu, G., & Ai, B. (2010, September). Radio propagation measurements and modeling in railway viaduct area. In *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)* (pp. 1–5). IEEE. https://doi.org/10.1109/WICOM.2010.5601255

9. Shen, S., Zhang, W., Zhang, H., Ren, Q., Zhang, X., & Li, Y. (2022). An accurate maritime radio propagation loss prediction approach employing neural networks. *Remote Sensing, 14*(19), 4753. https://doi.org/10.3390/rs14194753

10.Shon, T., & Moon, J. (2007). A hybrid machine learning approach to network anomaly detection. *Information Sciences, 177*(18), 3799–3821. https://doi.org/10.1016/j.ins.2007.03.025

11. Slater, L. J., Arnal, L., Boucher, M. A., Chang, A. Y. Y., Moulds, S., Murphy, C., ... & Zappa, M. (2023). Hybrid forecasting: Blending climate predictions with AI models. *Hydrology and Earth System Sciences*, *27*(9), 1865–1889. https://doi.org/10.5194/hess-27-1865-2023

12. Wang, H., Du, W., & Chen, X. (2015). Evaluation of radio over sea propagation based on ITU-R Recommendation P.1546-5. *Journal of Communications, 10*(4), 231–237. https://doi.org/10.12720/jcm.10.4.231-237

13. Wang, J., Hao, Y., & Yang, C. (2023). The current progress and prospects of path loss model for terrestrial radio propagation. *Electronics*, *12*(24), 4959. https://doi.org/10.3390/electronics12244959