

Support Vector Machine for Path Loss Predictions in Urban Environment

Robert O. Abolade[1], Solomon O. Famakinde[2], Segun I. Popoola[3,4(✉)], Olasunkanmi F. Oseni[1], Aderemi A. Atayero[3], and Sanjay Misra[3]

[1] Department of Electronic and Electrical Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria
[2] Department of Electronic and Computer Engineering, Lagos State University, Epe, Nigeria
[3] IoT-Enabled Smart and Connected Communities (SmartCU) Cluster, Covenant University, Ota, Nigeria
segun.popoola@covenantuniversity.edu.ng
[4] Department of Engineering, Manchester Metropolitan University, Manchester M1 5GD, UK

Abstract. Path Loss (PL) propagation models are important for accurate radio network design and planning. In this paper, we propose a new radio propagation model for PL predictions in urban environment using Support Vector Machine (SVM). Field measurement campaigns are conducted in urban environment to obtain mobile network and path loss information of radio signals transmitted at 900, 1800 and 2100 MHz frequencies. SVM model is trained with field measurement data to predict path loss in urban propagation environment. Performance of SVM model is evaluated using Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Standard Error Deviation (SED). Results show that SVM achieve MAE, MSE, RMSE and SED of 7.953 dB, 99.966 dB, 9.998 dB and 9.940 dB respectively. SVM model outperforms existing empirical models (Okumura-Hata, COST 231, ECC-33 and Egli) with relatively low prediction error.

Keywords: Support vector machine · Path loss · Radio propagation · Radio network planning · Machine learning

1 Introduction

Over the years, the use of mobile communication systems has continued to grow, rapidly leading to increase in network capacity [1–3]. In a bid to design an efficient wireless communications system, the random nature of the propagation channel poses a great challenge for efficient design of mobile network engineer [4–6]. Path Loss (PL) is the attenuation of radio signal power between transmitting and receiving station due to reflection, refraction and diffraction among other propagation mechanisms [7,8]. For accurate radio network design and planning, PL propagation models are important because they have effect on signal coverage and network

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12255, pp. 995–1006, 2020.
https://doi.org/10.1007/978-3-030-58820-5_71

interference [9]. Since network engineer has no control of the terrain, it is imperative to deployed accurate PL prediction model for efficient cellular communication system. PL prediction models are mathematical formulas used to characterize radio wave propagation as a function of distance, transmission frequency, antenna height and other conditions [10–14]. Radio propagation environments are categorized into rural, suburban, and urban with different and unique geographical features [7]. In previous works, Hata, COST 231, and Standard Propagation Model (SPM) models have been proposed for radio network planning at 1800 MHz [15–19]. However, signal attenuation and PL is determined by the nature of the terrain features such as high building, foliage and trees [20,21].

In previous works, Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Extreme Learning Machine (ELM) have been used to solve PL prediction problem [22]. Support Vector Machine is an algorithm than can be used to distinguish between two-groups or classes (classification) and also to obtain mathematical model for data prediction (regression) in a network. Support Vector Machine (SVM) was proposed for PL predictions in [23]. The results show that SVM gave lower computational complexity compared to that obtained using Multilayer Perceptron (MLP) neural network. The Laplacian kernel was the best among the investigated kernels. Also, the SVM algorithm using Laplacian kernel and MLP had similar performance. The authors in [24] proposed regularization of non-linear path with a modified Huber loss for the SVM. The result show that the algorithm can compute the nonlinear regularization path. SVM-based modeling technique of cabin PL prediction was also proposed in [25]. The measured path loss values points were trained inside the cabin which was used to predict the PL values of the un-measured points. The results show that modelling system is better than the curve fitting system. The authors in [26] proposed nonlinear regularization path algorithm for a class of machines learning that have quadratic penalty, which is also known as quadratic SVM loss. A nonlinear path algorithm was developed using approximation technique. The developed algorithm gave better result over conventional method. Some research activities have been carried employing machine learning in PL prediction mainly in developed countries but not in Nigeria.

In this paper, a new radio propagation model is proposed for path loss predictions in urban environment using SVM. Field measurement campaigns are conducted in urban environment to obtain mobile network and path loss information of radio signals transmitted at 900, 1800 and 2100 MHz frequencies. SVM model is trained with field measurement data to predict path loss in urban propagation environment.

2 Materials and Method

2.1 Radio Signal Measurement and Data Collection

Extensive field measurement campaign was conducted within Canaan-land, Ota, Ogun State, Nigeria. Most of these physical structures have considerable heights such that they obstruct line of sight and produce non-line of sight signal paths in

wireless communication channel at radio frequencies. Information about the geographic location and the altitude of the radio transmitters are presented in Table 1.

Table 1. Geographic locations of base station transmitters

BTS ID	Longitude	Latitude	Altitude (m)
A2GS1	3.162867	6.675068	50
A2GS2	3.162867	6.675068	50
A2GS3	3.162867	6.675068	50
A3GS1	3.162867	6.675068	50
A3GS2	3.162867	6.675068	50
A3GS3	3.162867	6.675068	50
E2GS1	3.164015	6.675253	52
E2GS3	3.164015	6.675253	52
E3GS1	3.164015	6.675253	52
E3GS3	3.164015	6.675253	52
M2GS1	3.163930	6.675245	52
M2GS3	3.163930	6.675245	52
M3GS1	3.163930	6.675245	52
M3GS3	3.163930	6.675245	52

A drive test experimental setup was designed for the field measurement campaign. The equipment, devices, and tools that constitute the experimental setup include: six commercial transceivers with fourteen (14) directional antennas, two mobile receivers, a Global Positioning System (GPS) receiver, a radio signal measurement software that runs on a Personal Computer (PC), and a motor vehicle. Ericsson RBS 2216, Ericsson RBS 2116, and Ericsson RBS 6201 base station transceivers were used for radio signal transmission at 900, 1800, and 2100 MHz respectively. Sectorial antennas of 13 dBi gain, 120° horizontally polarized sector panel were used to radiate electromagnetic signals which emanate from Ericsson RBS 2216 transmitters. 18 dBi gain, 65° vertically polarized antennas were used for radio wave transmission at 1710–1880 MHz frequency range. 17 dBi gain, 90° vertically polarized antennas were utilized for radio propagation at 2090–2290 frequency range. Two Sony Ericsson w995 mobile phones, with processing speed of 369 MHz and a removable Li-Po 930 mAh battery each, were used for radio signal reception at 900, 1800, and 2100 MHz. A Universal Serial Board (USB) magnet mount GPS receiver, BU-353-S4, was used to track mobile receiver's location at a given time. A 64-bit Windows Operating System (OS), 4 GB Random Access Memory (RAM) laptop with Intel® Core™ i5, M520 @2.40 GHz central processing capacity was used for data logging and storage.

When planning the drive test measurement survey, the area covered was initially scanned to ensure that there was no interference. The Broadcast Control Channel (BCCH) single frequency channel was obtained during each survey. There are two contiguous unused channels of a clearance of 200 kHz on either

side of the measured signal so as to ensure that the measured frequency is clean. Radio signal measurements were conducted along 14 drive test survey routes in order to adequately represent the wireless channel characteristics of a typical urban propagation environment. Received Signal Strength (RSS) from respective transmitters were measured, recorded, and stored as the mobile receivers are driven along each survey route using TEMS^{TM} Investigation software developed by InfoVista®. The amount of radio signal power transmitted by each of the transmitters was 43 dB and the selected mobile receiver has a minimum sensitivity of −100 dBm.

The empirical measurements covered six (6) commercial transceivers with fourteen (14) directional antennas namely: A2GS1, A2GS2, A2GS3, A3GS1, AW3GS2, A3GS3, E2GS1, E2GS3, E3GS1, E3GS3, M2GS1, M2GS3, M3GS1, and M3GS3. Radio signal transmission and reception were performed at 900, 1800, and 2100 MHz operating frequencies, as expected of GSM, Digital Cellular System (DCS), and UMTS wireless systems respectively, in the directions of the base station antennas. Continuous measurement of the RSS, longitude, latitude, elevation, altitude, frequency and clutter height were recorded.

2.2 Data Pre-processing

Data collected through drive test (i.e. RSS, Longitude, Latitude, Elevation, and Frequency) were exported from TEMS Investigation software developed by InfoVista into a spreadsheet file format. Mapping and location analysis of RSS data collected at 900, 1800, 2100 MHz radio frequencies were performed using MapInfo Pro^{TM}, produced by Pitney Bowes. Appropriate data filtering and sorting were performed using Microsoft Excel 2013 to remove data instance duplicates. The whole experimental field measurement process was accurately represented in ATOLL v3.1 radio network planning software produced by Forsk. Separation distances between base station transmitters and mobile receivers were computed for all data instances using ATOLL software.

The complete filtered and sorted data with nine variables (longitude, latitude, elevation, altitude, frequency, clutter height, distance, RSS, and PL) were randomly classified into 75% of training dataset and 25% of testing dataset for path loss model development and evaluation.

2.3 Development of SVM Model for PL Predictions

SVM was established and developed for learning theory. Moreover, excellent performances were gotten in regression and time series prediction applications with the aid of SVM regression, otherwise termed as Support Vector Regression (SVR) [27,28]. SVM consist of kernel methods which refer to a class of algorithms intended for pattern analysis. However, kernels have various conditions upon which they depend on.

Most influencing input variable attributes were selected using 10-fold validation approach. CFS Subset Evaluator and Greedy Stepwise methods were used

to search and evaluate the influence of seven independent attributes on a dependent variable (path loss). These algorithms were implemented in a Java-based machine learning software, WEKA, produced at the University of Waikato, New Zealand. Furthermore, SVM-based PL model was developed by SMOreg regression algorithm. Model parameters and kernel evaluations were obtained for PL predictions in heterogeneous urban environment.

The performance and prediction accuracy of the empirical and SVM-based PL model was evaluated using Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Standard Error Deviation (SED) with respect to PL values in both training data and testing data, respectively. MAE, MSE, RMSE and SED were calculated using Eqs. (1)–(4) respectively [29,30]:

$$MAE = \frac{1}{k}\sum_{i=1}^{k}(PL_{m,i} - PL_{p,i}), \tag{1}$$

$$MSE = \frac{1}{k}\sum_{i=1}^{k}(PL_{m,i} - PL_{p,i})^2, \tag{2}$$

$$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(PL_{m,i} - PL_{p,i})^2}, \tag{3}$$

$$SED = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(|PL_{m,i} - PL_{p,i}|) - MAE}, \tag{4}$$

where PL_m is the measured PL; PL_p is the predicted PL; and k is the number of samples in the dataset. Empirical models such as; Hata, COST 231, ECC-33 and Egli which are commonly used were employed for PL prediction based on the distance input vector provided in training and testing data sets.

3 Results and Discussion

The results obtained in this work are presented in this section. The data instances from field measurement campaign were collected and analyzed for model development. Information about the results obtained during data collection are presented in Table 2. A total of 123,985 raw data instances were logged with an average of 8,856 data instances per antenna. The remaining 18,865 unique data instances were curated for model development and evaluation after the duplicate has been removed. The mean number of unique data instances available along the survey routes of each of the fourteen sectors is 1,348. 75% of the complete RSS dataset (i.e. 14,142 unique data instances) was used for model training. The remaining 25% (i.e. 4,714 unique data instances) was used for model evaluation and testing.

Model was trained using 10-fold cross validation technique instead of dataset splitting approach. The parameters of SVM-based PL model are presented in Table 3.

Table 2. Quantitative summary of field measurement data

BTS ID	Raw data	Duplicates	Filtered data
A2GS1	2284	1626	658
A2GS2	3918	3168	750
A2GS3	4838	3388	1450
A3GS1	5551	4632	919
A3GS2	8139	7414	725
A3GS3	11555	9687	1868
E2GS1	11028	9067	1961
E2GS3	6591	4837	1754
E3GS1	24371	22274	2097
E3GS3	18319	15828	2491
M2GS1	4228	3439	789
M2GS3	6597	5052	1545
M3GS1	4123	3734	389
M3GS3	12443	10974	1469
Total	123985	105120	18865

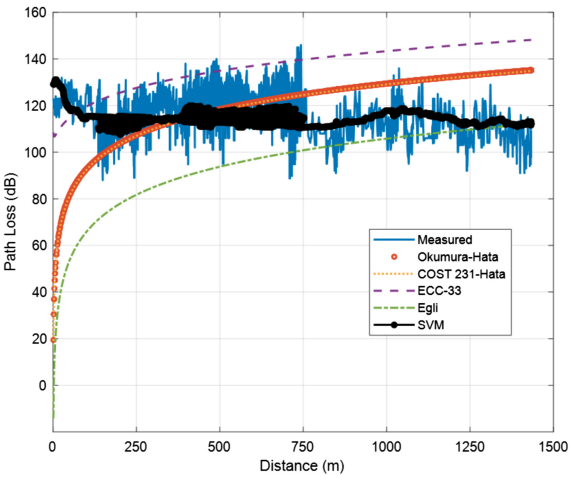

Fig. 1. Training results for path loss predictions at 900 MHz

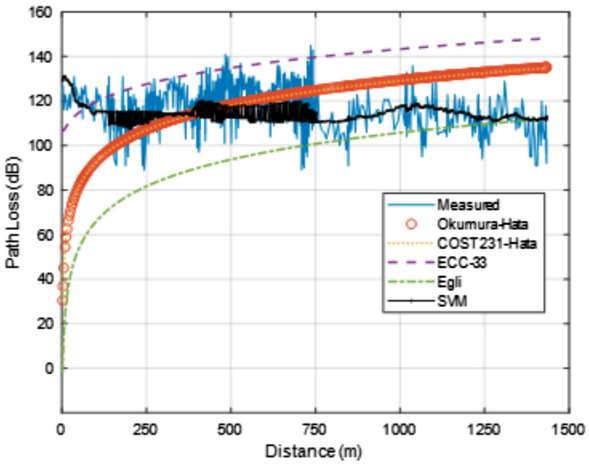

Fig. 2. Training results for path loss predictions at 1800 MHz

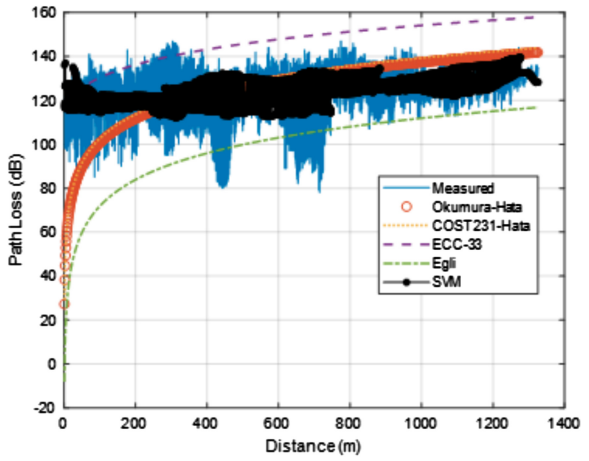

Fig. 3. Training results for path loss predictions at 2100 MHz

The developed SVM-based model and empirical models such as Okumura-Hata, COST 231, ECC-33 and Egli models were compared to the measured PL values for training and testing datasets to evaluate the prediction accuracy and generalization ability of the model. The results of the predicted model at 900, 1800 and 2100 MHz relative to the measured PL values in both training and testing datasets are graphically represented in Figs. 1, 2, 3, 4, 5 and 6 respectively.

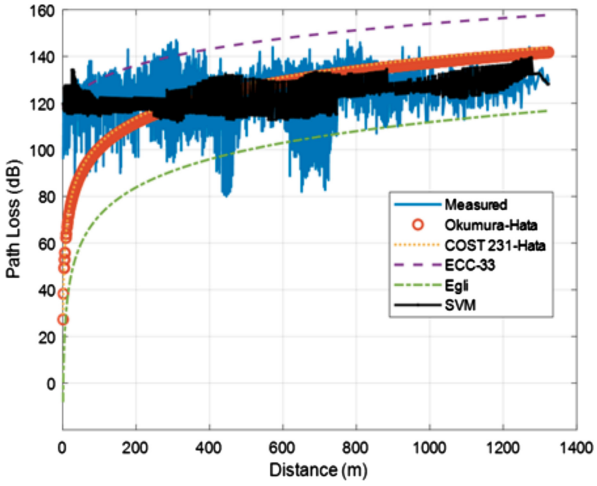

Fig. 4. Testing results for path loss predictions at 900 MHz

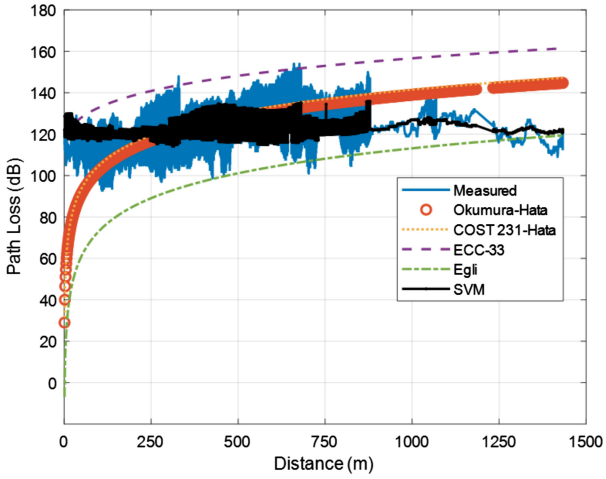

Fig. 5. Testing results for path loss predictions at 1800 MHz

Egli model produced the highest prediction error with MAE, MSE, RMSE, and SED of 27.000 dB, 969.657 dB, 31.139 dB, and 16.384 dB, respectively when compared to the measured PL values in training dataset. The performance evaluation results of the empirical models and SVM PL model based on the training data are presented in Table 4. The performance evaluation results of the empirical models and SVM PL model based on the testing data are presented in Table 5. The generalization ability demonstrated by SVM-based PL model (MAE, MSE, RMSE, SED of 7.933 dB, 98.773 dB, 9.938 dB, and 9.878 dB

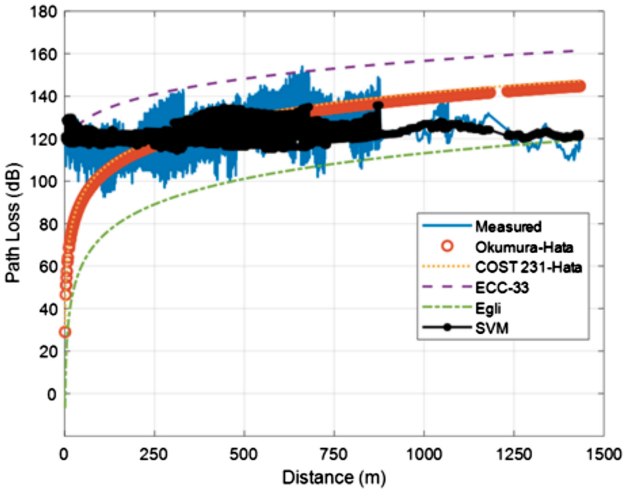

Fig. 6. Testing results for path loss predictions at 2100 MHz

Table 3. SVM model parameters

Attribute (normalized)	Weight
Constant	0.4487
Longitude	0.1939
Latitude	−0.2789
Elevation	−0.2531
Altitude	0.4655
Frequency	0.1207
Clutter height	−0.0149
Distance	0.1413

respectively) is much relatively better than those of all the empirical models. Egli model demonstrated the least generalization ability with MAE, MSE, RMSE, and SED of 27.044 dB, 974.318 dB, 31.214 dB, and 16.429 dB, respectively when compared to the measured PL values in testing dataset.

The prediction outputs of the developed SVM-based model, and popular empirical models (i.e. Okumura-Hata, COST 231, ECC-33, and Egli) were compared to the measured path loss values in both training and testing datasets to evaluate the prediction accuracy and generalization ability of the path loss models. The prediction error produced by SVM-based path loss model (MAE, MSE, RMSE, and SED values of 7.953 dB, 99.966 dB, 9.998 dB, and 9.940 dB respectively) is much relatively lower than those of all the empirical models.

Table 4. Performance of SVM and empirical PL models on training dataset

Model	MAE (dB)	MSE (dB)	RMSE (dB)	SED (dB)
Okumura-Hata	11.51	236.93	15.393	15.391
COST 231	11.778	241.055	15.526	15.374
ECC-33	21.884	609.75	24.693	11.948
Egli	27	969.657	31.139	16.384
SVM	7.953	99.966	9.998	9.94

Table 5. Performance of SVM and empirical PL models on testing dataset

Model	MAE (dB)	MSE (dB)	RMSE (dB)	SED (dB)
Okumura-Hata	11.507	237.888	15.424	15.424
COST 231	11.765	241.847	15.551	15.409
ECC-33	21.831	607.061	24.639	11.896
Egli	27.044	974.318	31.214	16.429
SVM	7.933	98.773	9.938	9.878

Egli model produced the highest prediction error with MAE, MSE, RMSE, and SED values of 27.000 dB, 969.657 dB, 31.139 dB, and 16.384 dB, respectively when compared to the measured path loss values in training dataset.

4 Conclusion

In this paper, SVM model was developed for path loss predictions in urban propagation environment. Field measurement campaigns were conducted to obtain RSS values and path loss values at varying longitude, latitude, altitude, elevation, clutter height, distance, and available radio frequencies (900, 1800, and 2100 MHz) within Canaaland, Ota, Ogun State, Nigeria. SVM model was trained with the network parameters to predict path loss. The performance of SVM model was compared with empirical models (Hata, COST 231, ECC-33, and Egli). Results from experimentation showed that SVM model gave the best output with MAE, MSE, RMSE, SED of 7.953 dB, 99.966 dB, 9.998 dB, 9.940 dB respectively. Comparative analysis showed that SVM model achieved high prediction accuracy with better generalization ability.

Acknowledgement. This work was carried out under the IoT-Enabled Smart and Connected Communities (*SmartCU*) research cluster of the Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria. The research was fully sponsored by Covenant University Centre for Research, Innovation and Development (CUCRID), Covenant University, Ota, Nigeria.

References

1. Berezdivin, R., Breinig, R., Topp, R.: Next-generation wireless communications concepts and technologies. IEEE Commun. Mag. **40**, 108–116 (2002)
2. Ghosh, A., Ratasuk, R., Mondal, B., Mangalvedhe, N., Thomas, T.: LTE-advanced: next-generation wireless broadband technology. IEEE Wirel. Commun. **17**, 10–22 (2010)
3. Wang, C.-X., Haider, F., Gao, X., You, X.-H., Yang, Y., Yuan, D., et al.: Cellular architecture and key technologies for 5G wireless communication networks. IEEE Commun. Mag. **52**, 122–130 (2014)
4. Laiho, J., Wacker, A., Novosad, T. (eds.): Radio Network Planning and Optimisation for UMTS, vol. 2. Wiley, New York (2002)
5. Mishra, A.R. (eds.): Advanced Cellular Network Planning and Optimisation: 2G/2.5G/3G... Evolution to 4G. Wiley, New York (2007)
6. Mishra, A.R. (eds.): Fundamentals of Network Planning and Optimisation 2G/3G/4G: Evolution to 5G. Wiley, New York (2018)
7. Rappaport, T.S.: Wireless Communications: Principles and Practice, vol. 2. Prentice Hall PTR, New Jersey (1996)
8. Tse, D., Viswanath, P.: Fundamentals of Wireless Communication. Cambridge University Press, Cambridge (2005)
9. Nawrocki, M., Aghvami, H., Dohler, M.: Understanding UMTS Radio Network Modelling, Planning and Automated Optimisation: Theory and Practice. Wiley, New York (2006)
10. Oseni, O.F., Popoola, S.I., Abolade, R.O., Adegbola, O.A.: Comparative analysis of received signal strength prediction models for radio network planning of GSM 900 MHz in Ilorin, Nigeria. Int. J. Innov. Technol. Exploring Eng. **4**, 45–50 (2014)
11. Obot, A., Simeon, O., Afolayan, J.: Comparative analysis of path loss prediction models for urban macrocellular environments. Niger. J. Technol. **30**, 50–59 (2011)
12. Popoola, S.I., Atayero, A.A., Popoola, O.A.: Comparative assessment of data obtained using empirical models for path loss predictions in a university campus environment. Data Brief **18**, 380–393 (2018)
13. Faruk, N., Ayeni, A., Adediran, Y.: Characterization of propagation path loss at VHF/UHF bands for Ilorin city, Nigeria. Niger. J. Technol. **32**, 253–265 (2013)
14. Salman, M.A., Popoola, S.I., Faruk, N., Surajudeen-Bakinde, N., Oloyede, A.A., Olawoyin, L.A.: Adaptive neuro-fuzzy model for path loss prediction in the VHF band. In: International Conference on Computing Networking and Informatics (ICCNI) 2017, pp. 1–6 (2017)
15. Al Salameh, M.S., Al-Zu'bi, M.M.: Prediction of radiowave propagation for wireless cellular networks in Jordan. In: 2015 7th International Conference on Knowledge and Smart Technology (KST), pp. 149–154 (2015)
16. Faruk, N., Ayeni, A., Adediran, Y.A.: On the study of empirical path loss models for accurate prediction of TV signal for secondary users. Prog. Electromagn. Res. **49**, 155–176 (2013)
17. Ibhaze, A.E., Ajose, S.O., Atayero, A.A.-A., Idachaba, F.E.: Developing smart cities through optimal wireless mobile network. In: IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech) 2016, pp. 118–123 (2016)
18. Nimavat, V.D., Kulkarni, G.: Simulation and performance evaluation of GSM propagation channel under the urban, suburban and rural environments. In: 2012 International Conference on Communication, Information & Computing Technology (ICCICT), pp. 1–5 (2012)

19. Rath, H.K., Verma, S., Simha, A., Karandikar, A.: Path loss model for Indian terrain-empirical approach. In: Twenty Second National Conference on Communication (NCC) 2016, pp. 1–6 (2016)
20. Oseni, O.F., Popoola, S.I., Enumah, H., Gordian, A.: Radio frequency optimization of mobile networks in Abeokuta, Nigeria for improved quality of service. Int. J. Res. Eng. Technol. **3**, 174–180 (2014)
21. Mitra, A., Reddy, B.: Handbook on Radio propagation for tropical and subtropical countries (1987)
22. Goldberg, D.E., Holland, J.H.: Genetic algorithms and machine learning. Mach. Learn. **3**, 95–99 (1988)
23. Benmus, T.A., Abboud, R., Shatter, M.K.: Neural network approach to model the propagation path loss for great Tripoli area at 900, 1800, and 2100 MHz bands. In: 2015 16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), pp. 793–798 ((2015)
24. Ostlin, E., Zepernick, H.-J., Suzuki, H.: Macrocell path-loss prediction using artificial neural networks. IEEE Trans. Veh. Technol. **59**, 2735–2747 (2010)
25. DalkiliÇ, T.E., Hanci, B.Y., Apaydin, A.: Fuzzy adaptive neural network approach to path loss prediction in urban areas at GSM-900 band. Turkish J. Electri. Eng. Comput. Sci. **18**, 1077–1094 (2010)
26. Ayadi, M., Zineb, A.B., Tabbane, S.: A UHF path loss model using learning machine for heterogeneous networks. IEEE Trans. Antennas Propag. **65**, 3675–3683 (2017)
27. Stitson, M., Gammerman, A., Vapnik, V., Vovk, V., Watkins, C., Weston, J.: Support vector regression with ANOVA decomposition kernels. In: Soentpiet, R., (ed.) Advances in Kernel Methods–Support Vector Learning, pp. 285–292 (1999)
28. Müller, K.-R., Smola, A.J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., Vapnik, V.: Predicting time series with support vector machines. In: Gerstner, W., Germond, A., Hasler, M., Nicoud, J.-D. (eds.) ICANN 1997. LNCS, vol. 1327, pp. 999–1004. Springer, Heidelberg (1997). https://doi.org/10.1007/BFb0020283
29. Faruk, N., Popoola, S.I., Surajudeen-Bakinde, N.T., Oloyede, A.A., Abdulkarim, A., Olawoyin, L.A., et al.: Path loss predictions in the VHF and UHF bands within urban environments: experimental investigation of empirical, heuristics and geospatial models. IEEE Access **7**, 77293–77307 (2019)
30. Popoola, S.I., Jefia, A., Atayero, A.A., Kingsley, O., Faruk, N., Oseni, O.F., et al.: Determination of neural network parameters for path loss prediction in very high frequency wireless channel. IEEE Access **7**, 150462–150483 (2019)