ANN-based LoRaWAN Channel Propagation Model

https://doi.org/10.3991/ijim.v16i11.30095

Mohamed Hadi Habaebi[1(✉)], Ahmad Shahmi Mod Rofi[1],
Md Rafiqul Islam[1], Ahmed Basahel[2]
[1]IoT & Wireless Communication Protocols Lab, Department of Electrical and Computer Engineering, International Islamc University Malaysia (IIUM), Kuala Lumpur, Malaysia
[2]First Fuel Company, Jedah, Saudi Arabia
habaebi@iium.edu.my

Abstract—LoRaWAN wireless communication channels are often impacted by noise and interference over long-range causing loss of a received signal. One of the main drawbacks of using existing propagation models is less accurate as these models in designing the communication link are tailored to simplify the estimation. In this paper, an artificial intelligent real time path loss model is proposed. It is capable of processing complex variables over a short period of time. Providing it with enough data, the model is able to learn channel behavior and predict the path loss accurately. Results of the model are benchmarked against classical statistical curve fitting models where RMSE values are also compared and indicating that the artificial intelligent model has better accurate prediction.

Keywords—artificial neural network, LoRAWAN channel, artificially intelligent, LoRa propagation loss models

1 Introduction

Current LoRa propagation loss models are developed based on the free-space loss equation. To simplify calculation, some parameters were ignored during the development of these models resulting in less accuracy prediction.

Generally, the development of any path loss propagation model is mainly weather dependent. Thus, different environments will have different weather conditions resulting in different model parameters. In this work, LoRa propagation loss model is developed at the IIUM campus. The campus has typically tropical climate conditions. In such an environment, the propagation loss is mainly due to forest (foliage effect) and building around the campus. Therefore, the evaluation of the Neural Network (NN) is observed in the part of the experiment to develop the best fit model that can produce accurate prediction for the LoRa network at IIUM.

LoRa technology was first developed in 2009. LoRa network's operational frequency varies from one site to another depending on the region which is installed. For example, the operational frequency of LoRa in Europe is 433 MHz, 868 MHz & 915 MHz which are ISM radio bands. The Medium Access Control (MAC) protocol for LoRaWAN is allowing several receiver terminals to communicate within a gateway. Moreover, the protocol mechanism can control transmitted data packets by shared channels [1].

NN consists of many algorithms that is often produced by training a set of raw data that is modelled based on the human brain. NN are capable of processing big data depending on inputs. Furthermore, NN can also compute large amounts of complex data quickly and accurately. One of the main applications of NN is prediction of path loss propagation.

Several popular propagation loss models will be reviewed as the benchmark for the implementation of the NN in LoRa propagation, the Okumura Hata Model (OHM), Lee Model (LM), and Stanford University Interim (SUI) Model.

2 Propagation path loss models

2.1 OHM

The Okumura Model (OM) is mainly used for determining the path loss for Mobile Cellular Network (MCN). With 150 MHz to 1920 MHz frequency range, the Okumura model can operate over a range of 1 km to 100 km distances in the urban areas. Moreover, this model is able to cover a base station antenna(BSA) at a distance of 30 m to 100 m [2]. Okumura model is one among the few early models for MCN, with its simplicity and accuracy. The OM, however, has some drawbacks in the use of computational predictions. The Hata model (HM) was introduced in the place of OM in order to overcome the limitations.

The development of OM has been carried out by Hata to introduce the OHM which is widely used in the urban area. The developed model takes into account the losses due to different effects such as shadowing, reflection, diffraction, and scattering of propagated signals [3]. The HM expression is defined as below:

HM used for urban areas:

$$\begin{aligned} Lu = 69.55 + 26.16\log(f) - 13.82\log(h_B) - C_H + \\ (44.9 - 6.55\log(h_B))\log(d) \end{aligned} \tag{1}$$

- Small and Medium cities,

$$C_H = 0.8 + (1.1\log(f)\,h_M - 1.56\log(f)) \tag{2}$$

- *Large cities*,

$$C_H = \begin{cases} 8.29(\log(1.54h_M))^2 - 1.1, & f \leq 200MHz \\ 3.2(\log(11.75h_M))^2 - 4.97, & f \geq 200MHz \end{cases} \tag{3}$$

Where, h_B = the height of BSA, h_M = the height of MSA, C_H = the ACF, and d = the link distance between BSA and MSA.

For outdoor propagation of outskirts areas, the HM can be as follows:

$$L_s = L_u - 4.78(\log(f))^2 + 18.33\log(f) - 40.94 \quad (4)$$

LM. The model is easy to use and gives an accurate result. Moreover, the Lee propagation model takes suitability into account and its antenna height correction factor (ACF) has been modified for easy use in local climate conditions [4]. The Lee path loss propagation model is expressed as follows, where L_o is the median path loss (MPL) at 1 km length, γ is the slope of the (path lose curve) PLC in dB per decade; and G is the gain in antenna.

$$P_L(d)(db) = L_o + \gamma\log(d) - 10\log(F_o) \quad (5)$$

The adjustment factor, F_o can be express as follow:

$$F_1 = \left(\frac{h_B}{30.48}\right)^2 = \left(h_b\left(\frac{ft}{100}\right)\right)^2 \quad (6)$$

$$F_2 = \frac{G_B}{4} \quad (7)$$

$$F_3 = f(x) = \begin{cases} \left(\frac{h_M}{3}\right), & x < 3 \\ \left(\frac{h_M}{3}\right)^2, & x > 3 \end{cases} \quad (8)$$

$$F_4 = \left(\frac{f}{900}\right)^{-n} \quad 2 < n < 3 \quad (9)$$

$$F_5 = \frac{1}{G_M} \quad (10)$$

2.2 SUI's model

The SUI model was introduced by a Stanford University researchers and is intended to be used for <11 GHz frequencies. And, it has three terrains that are split depending on the intensity of loss in path. The Terrain A is used for the highest impact of leafage, whereas Terrain C is utilised for lowest impact [5]. The SUI propagation model is expressed as below:

$$PL = A + 10\gamma\log\left(\frac{d}{d_o}\right) + X_f + X_h + s \quad (11)$$

$$A = 20\log\left(\frac{4\pi d_o}{\lambda}\right) \tag{12}$$

$$\gamma = a - bh_b + \frac{c}{h_b} \tag{13}$$

Where, the parameter s = log normally distributed factor that vary between 8.2 dB and 10.6 dB.

Table 1. Terrains parameter

Model Parameter	Terrain A	Terrain B	Terrain C
a	4.6	4.0	3.6
$b(m^{-1})$	0.0075	0.0065	0.005
$c(m)$	12.6	17.1	20

$$X_f = 6\log\left(\frac{f}{2000}\right) \tag{14}$$

$$X_h = \begin{cases} -10.8\log\left(\frac{h_M}{2000}\right), & for\,Terain\,A\,and\,B \\ -20.0\log\left(\frac{h_M}{2000}\right), & for\,Terain\,C \end{cases} \tag{15}$$

2.3 Propagation loss for artificial neural network (ANN)

The ANN-based model approaches combine the merits of deterministic and empirical models. In general, the combination of the two models allows a more accurate loss prediction but it will take more time to implement. The dominant strength in NN: allows ANN to calculate enormous computational capacity of massive data with resilience to different climatic zones. This is particularly due to its unique ability to model a complex nonlinear functions.

The design of ANN often follows the connections between a set of neurons with another layer. For instance, ANN feedforward, neurons in the same layer will not be able to connect to each other. A multilayer perceptron for the FF-ANN generally comprises an input layer pursued by one or more hidden layers to generate an output. The connected neurons must possess a weighting factor which describes the robustness of the connection to minimize the error.

The disadvantage of the Feed-Forward ANN occurs when the model is overtrained and often during analysing non-complex structures. This drawback is resulting in unsustainable input, which has different characteristics from the training set. To overcome such a complicated structure, the easy algorithm which called a Back-Propagation

ANN was developed. Nevertheless, data selection is still a focal point to obtain a high accurate of path loss result. According to different models that were developed with this ANN or Back-propagation ANN, the result shows the mean square error (RMSE) of 5 dB very accurately [6].

3 Methodology

The LoRaWAN RF1276 module was developed at a low cost. The module can transmit signal that covers up to 5 km distance with high-performance transparency. On top of that, it is an ultra-low-power consumption wireless module for the operating frequencies of 169 MHz, 433 MHz, 869.5 MHz, and 915 MHz.

The antenna is connected by a coaxial cable and is used in both receiver and transmitter terminal. For a larger coverage area, a higher antenna gain is required. The used antenna should be able to operate at 869 MHz frequency with a sensitivity gain of 5 dBi.

3.1 Measurement

The transmitter module was placed at the highest point in the dormitory building of Mahallah Ruqayyah on IIUM Gombak university campus. The height of the transmitter was measured 30 meters above the ground. In contrast, the receiving antenna was placed 1.5 meters above the ground. For data collection, the receiver terminal must be placed away from its base station. The separation link between both the receiver and the transmitter terminal was measured over every meter starting at 10 meters up to 1 km [8].

Four various data sets were collected from various module designs. The frequency is set to 869 MHz by default with an output power of 25 dB at 9600 bps. Figure 1a shows the experimental setup of LoRaWAN. Figure 1b shows the configuration tool for the RF tool TTL-UART software interface for the RF1276 modem.

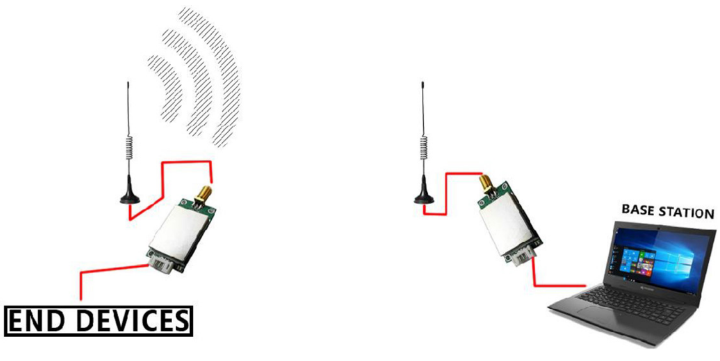


Fig. 1a. LoRaWAN hardware experimental setup

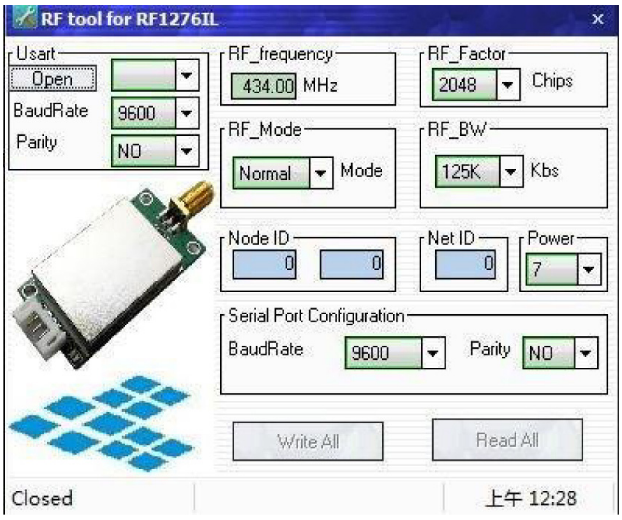


Fig. 1b. RF1276 RF tool for configuring the module settings of the LoRaWAN modem

The following Table 2 shows different LoRa module transmission settings:

Table 2. Different settings LoRa module

LoRa Module	Bandwidth (kHz)	RF Factor
Sn1	125	2048
Sn2	250	4096
Sn3	125	4096
Sn4	250	2096

3.2 Training data

There has been interestingly few studies on the use of machine learning and deep learning models with LoRa RSS datasets in the literature [9–10]. There are four separate data sets based on different RF factors and bandwidths. Only two data sets are used in training the model and the rest of datasets are used for testing the model.

To create the NN model, defining the activation function, several layers, and neurons are needed. These parameters play a fundamental role in determining the performance of the model. The selection process of the best parameter for the proposed model is certainly made using the lowest possible RMSE.

There are 3 models generated using ANN. The first model simply used two inputs. The second model used bandwidth and RF factor as inputs instead of using the Sn number. And the latter model named as a hybrid-model and it used the propagation loss model as an input.

Table 3. Models proposed

Model	No. of Input	Input
1	2	Distance, Sn Value
2	3	Bandwidth, RF Factor, Distance
3	6	Okumura-Hata (Urban), Okumura-Hata (Suburban), Lee Model, SUI Model, Distance, Sn Value

3.3 Testing proposed model

Depending on the required input to examine the model, a different dataset is then compiled. Next, the result of the propagation loss is analysed. The RMSE must be calculated for evaluation purposes. RMSE can define the error boundary as it penalizes the big error: overshoot or undershoot. RMSE can be expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Predicted\ Loss - Measured\ Loss)^2} \tag{16}$$

4 Findings

The comparison of performance evaluation of the proposed model are mainly focused on the accuracy of the data characteristics. The proposed model aims to be able to predict the loss of outdoor propagation on the IIUM campus. The obtained result of the received signal strength indicator (RSSI) was very poor, generally exceeding 100 dBm at an output power of 20 dBm. This poor performance is due to the effect of foliage in a tropical environment. To improve the propagation performance, the operating frequency is set to 868 MHz. That is the optimum frequency to avoid the interference with the GSM 900 MHz frequency.

4.1 Proposed model

The most accurate activation function was selected by comparing the value of root mean square error (RMSE). The activation function comparison test was carried out several times on various measured data sets. The test results are shown in the following table.

Table 4. Activation function comparison test

Activation Function	RMSE
Relu	13.6
Sigmoid	15.19
Tanh	12.54

To enhance this propagation model, the activation function with the lowest RMSE value was chosen. The activation function is connected to every neuron to obtain the weight based on the input. At present work, the activation function of the hyperbolic tangent (tanh) with the lowest RMSE tested 12.54 dB, was selected in order to develop the proposed model. The tanh function derived from the sigmoid function that works much better.

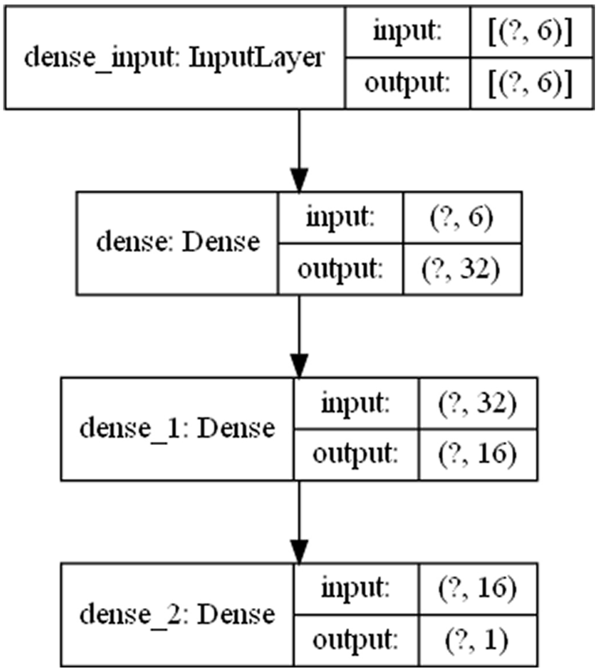
In addition, the test was carried out to estimate number of network layers and nodes needed. Based on the test result below, increasing the number of layers may greatly bring down the RMSE value. The 1st layer has 32 neurons, and the 2nd layer can be ether consist of 32 or 16 neurons to achieve a 10 dB RMSE result. However, reducing the number by using the 16 neurons instead of the 32 neurons in the 2nd layer can reduce the size of the model by 2 KBs.

Table 5. Number of layers and neurons test result

Number of Layers	Neurons	RMSE
1	8	40.14
1	16	33.78
1	32	14.52
2	32 – 8	17.87
2	32 – 16	10.65
2	16 – 32	13.95
2	32 – 32	10.25

The proposed model consists of these parameters:

1. The activation function is hyperbolic tangent.
2. Number of layers: two layers (16–32 Neurons).
3. The optimization Algorithm is Stochastic Gradient Descent.



All three models were trained for 50,000 epochs and two datasets were used. These datasets were divided into 2 – training data and validation data. Training data was used to update the network weight while validation data was used to evaluate the model performance for the current dataset.

4.2 Model 1

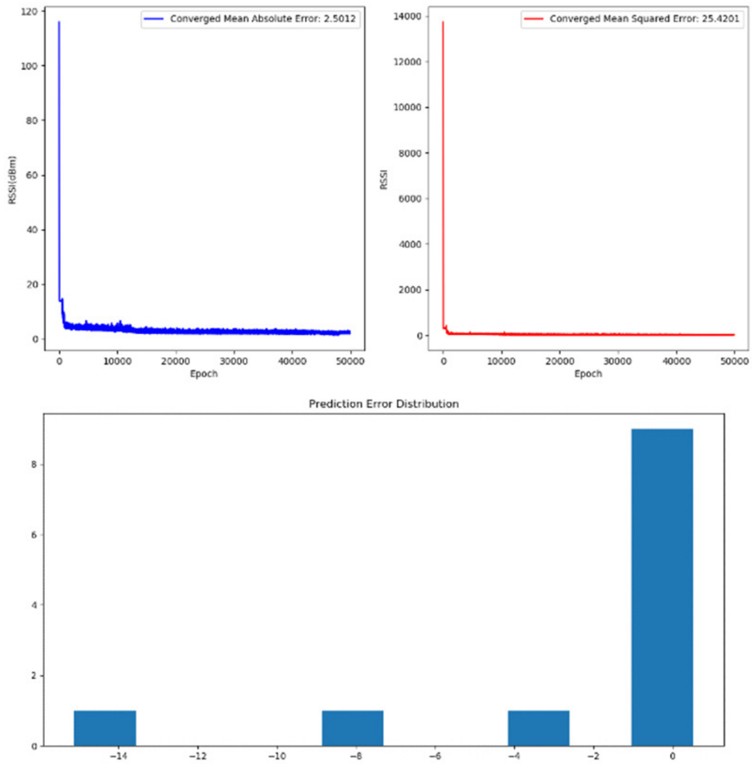


Fig. 2. Model 1 testing graph and validation error distributions

The first model was designed with 2 inputs to predict the same RSSI. The model was trained for 50,000 epochs. The training result fluctuated around 20–25 RMSE. Figure 2 above shows the training process plot that took place over time and the error distribution for the validation set. The most frequent error value falls below 1 dBm, which is good. However, some errors are very high, which occur 3 times with a value above 8 dBm.

4.3 Model 2

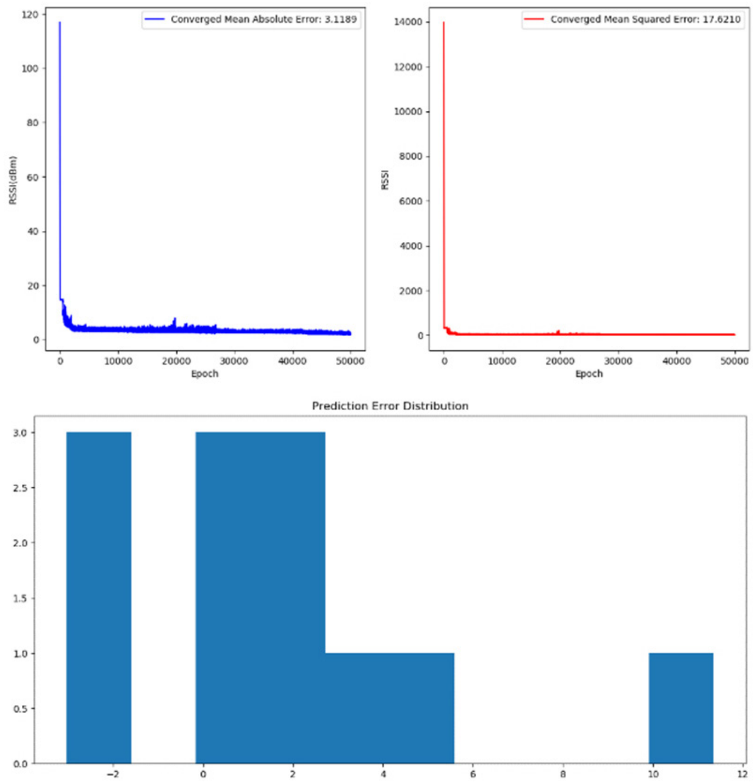


Fig. 3. Model 2 testing graph and validation error distributions

The second model accepts 3 inputs, which is expected to work much better as it can differentiate between the RF and bandwidth differences. The training process gives very consistent results between 23–26 RMSE. Compared to the first model, the current model's performance does not exceed the expectation as it produces about similar error deviation.

The validation test result has shown a very consistent RMSE evaluation. From Figure 3 above, the error frequently distributed below 4 dBm and only one reading that gives a very high error above 10 dBm. The RMSE value for the validation test is 17.62 dB, which is much lower compared to the training process. This indicates that the constructed model works best for this set of data tests.

4.4 Model 3

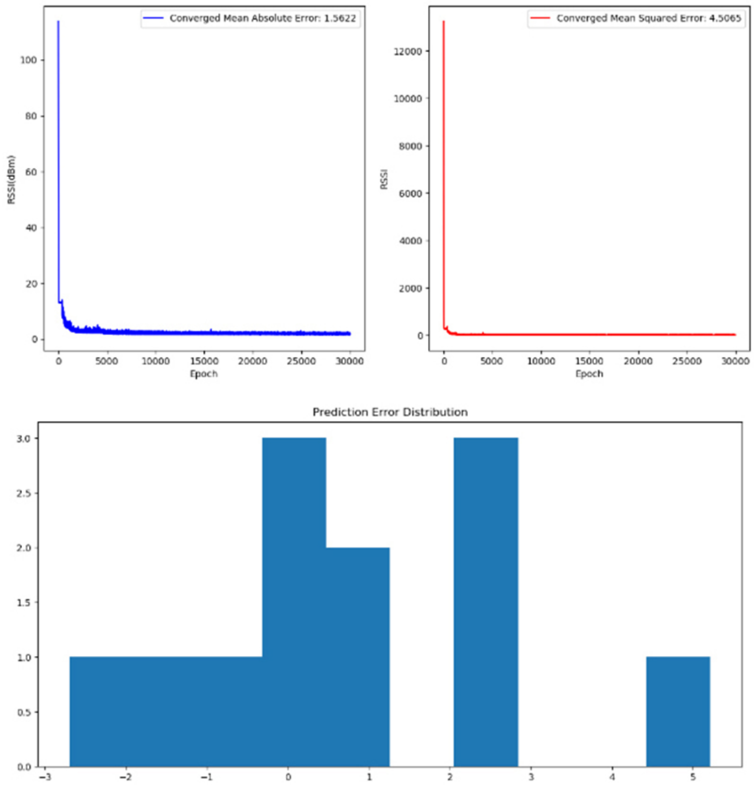


Fig. 4. Model 3 testing graph and validation error distributions

The third model (hybrid model) was designed to accept 6 inputs, and 4 of these inputs were taken from the benchmark's path loss prediction. The benchmarks considered are the OHM, the SUI model and the LM. This model should take advantage of the benchmark model to improve the proposed model.

This model improves significantly according to Figure 4 as the training error deviated from 15–16 RMSE. This model produces the least error compared to the first and second models. The following validation test shows a very good result, 4.51 RMSE. The error distribution did not exceed 5 dBm. Most validation results show that the error is typically falls below 3 dBm.

In general, the training result for these three models is between 15 and 26 dB RMSE. This error value is considered a bit high due to the random multipath spikes of the measured RSSI dataset value (training output). As a result, the train network cannot suggest the appropriate weight that contributed to this value. However, the model can be further enhanced by providing inputs that can respond to this error.

4.5 Benchmark

To assess the capabilities of this model, a test was conducted to find out how it responds to other data sets. To test data with distinct, the rf factor and bandwidth were used. This method is easy to do because the size of the model is only about 21.2 Kbytes and no additional steps are required.

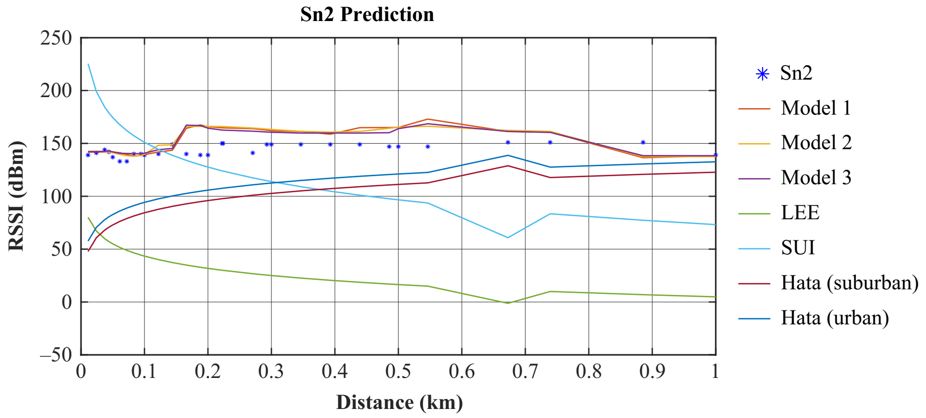


Fig. 5. Sn2 dataset prediction comparison

Figure 5 shows the first tests results. The proposed model shows good results. Of all the models shown below, the suggested model and the Okumura-Hata are in close agreement to actual value. Results showed the output power of proposed model ranges from 140 dB to 175 dB. Conversely, the performance of the OHM fluctuates from 100 to 130 dB. This observation indicates that the ANN model is far more efficient than the classic OHM since the measured data is quite close to the 150 dB limit.

Then the build model was tested with the second data set. The results of this test are shown in Figure 6 below. Overall, the design performance outperforms the other models.

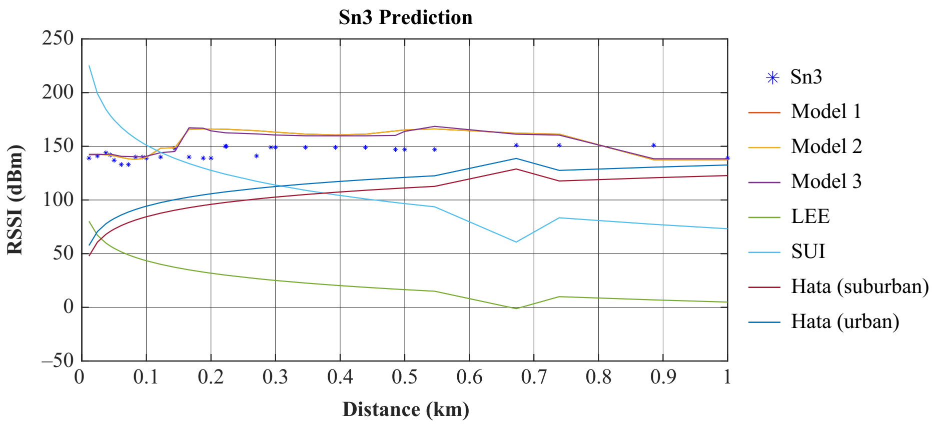


Fig. 6. Prediction comparison for Sn3 data

As a result, the proposed ANN model predicts the path loss to vary from 140 to 174 dB. Figure 6 shows that the OHM is closer to the measured data, Whereas the Lee and SUI models had a huge gap. This is due to evere degradation during channel propagation that has attenuated the signal with great loss.

4.6 RMSE

For every PLM, the RMSE was calculated to for the evaluation these models. The RMSE value can determine the percentage of overshoot and undershoot over a wide range, as it penalizes large errors. Therefore, it is more suitable to use in evaluating the performance of the path loss.

Table 6. Different LoRa modules settings

Model	Model 1	Model 2	Model 3	SUI's
RMSE	13.742	13.742	12.892	44.227
Model	Hata's (Urban)	Hata's (Suburban)	LEE's	
RMSE	51.884	42.637	112.812	

Table 7. Different LoRa module settings

Model	**Model 1**	**Model 2**	**Model 3**	**SUI's**
RMSE	13.963	13.741	12.901	44.234
Model	**Hata's (Urban)**	**Hata's (Suburban)**	**LEE's**	
RMSE	51.922	42.675	112.845	

From Tables 6 and 7 above, few points need to be noted. The proposed model performed well as compared to other models. The RMSE value obtained is less than 14. Conversely, error values of other models are very high. As a matter of fact, the lower values of RMSE, the better the performance is. The typical RMSE values are between 12 to 15 dB [7].

5 Conclusion

In conclusion, the obtained results indicate that the ANN proposed propagation model provides accurate predictions compared to the other path loss models. As mentioned earlier, predicted models can predict outperform than benchmarks by a wide margin. The developed model is able to predict losses as low as 12 dB. However, obtained results of the ANN training as well as the validation process showed losses values as low as 7 dB. This induce that the proposed model may be valid for a similar dataset. A larger set of training data will yield to a better result. In communication system, accurate propagation loss is critical. Thus, developing a PLM which is able to predict the loss in path accurately and quickly is an essential task to determine the range of communication and select the optimum base station. The proposed model might be a useful method for ANN LoRa communication networks mainly in a tropical climate condition.

6 Acknowledgment

This work is funded by the IIUM Publication Research Initiative Grant Scheme P-RIGS18-003-0003.

7 References

[1] B. Chaudhari, and S. Borkar, "Design Considerations and Network Architectures for Low-Power Wide-Area Networks," *LPWAN Technol. IoT M2M Appl.*, pp. 15–35, 2020, https://doi.org/10.1016/B978-0-12-818880-4.00002-8

[2] A. Mandal, and M. K. Nigam, "Comparative Path Loss Analysis Of Okumura And COST 231 Models For Wireless Mobile Communication Using MATLAB Simulation," vol. 1, no. 10, pp. 46–54, 2018.

[3] A. Deme, D. Dajab, B. Bajoga, M. Mu, and D. Choji, "Hata-Okumura Model Computer Analysis for Path Loss Determination at 900 MHz for Maiduguri, Nigeria," vol. 3, no. 3, pp. 1–10, 2013.

[4] J. Chebil, A. K. Lwas, R. Islam, and A. Zyoud, "Adjustment of Lee Path Loss Model for Suburban Area in Kuala Lumpur-Malaysia," vol. 5, pp. 252–257, 2011.

[5] V. S. Abhayawardhana, I. J. Wassellt, D. Crosby, M. P. Sellars, and M. G. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," *IEEE Veh. Technol. Conf.*, vol. 61, no. 1, pp. 73–77, 2005, https://doi.org/10.1109/VETECS.2005.1543252

[6] Y. Zhang, J. Wen, G. Yang, Z. He, and J. Wang, "Path loss prediction based on machine learning: Principle, method, and data expansion," *Appl. Sci.*, vol. 9, no. 9, 2019, https://doi.org/10.3390/app9091908

[7] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the practical error of path loss models," *Int. J. Antennas Propag.*, vol. 2012, 2012, https://doi.org/10.1155/2012/754158

[8] A. S. Mod Rofi, M. H. Habaebi, M. R. Islam, and A. Basahel, "LoRa Channel Propagation Modelling using Artificial Neural Network," 2021 8th International Conference on Computer and Communication Engineering (ICCCE), 2021, pp. 58–62, https://doi.org/10.1109/ICCCE50029.2021.9467234

[9] K. Saraubon, N. Wiriyanuruknakon, and N. Tangthirasunun, "Flashover Prevention System using IoT and Machine Learning for Transmission and Distribution Lines," *International Journal of Interactive Mobile Technology (iJIM)*, vol. 15, no. 11, pp. 34–48, 2021. https://doi.org/10.3991/ijim.v15i11.20753

[10] S. El Abkari, A. Jilbab, and J. El Mhamdi, "RSS-based Indoor Positioning Using Convolutional Neural Network," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 16, no.12, pp. 82–93, 2020. https://doi.org/10.3991/ijoe.v16i12.16751

8 Authors

Mohamed Hadi Habaebi is a professor with the department of electrical and computer engineering, International Islamic University Malaysia. His interests are in FSO, radio channel propagation, IoT and AI. Email: habaebi@iium.edu.my

Ahmad Shahmi Mod Rofi is a Bachelor of Science in Enginering student at the department of electrical and computer engineering. His research iterests are in IoT technologies, LoRaWAN and AI. Email: shahi.rof97@gmail.com

Md Rafiqul Islam is a professor with the department of electrical and computer engineering, International Islamic University Malaysia. His interests are in FSO, radio channel propagation and IoT. Email: rafiq@iium.edu.my

Ahmed Abdullah Basahel received his PhD from International Islamic University Malaysia (IIUM) in 2017. He is currently working as Research Associate at the department of Electrical and Computer Engineering (IIUM). His research areas are but not limited to availability/reliability of communications systems, optical wireless communications including FSO, AI for next generation networks & hybrid networks integration. Email: ba_sahal@hotmail.com

Article submitted 2022-02-09. Resubmitted 2022-03-03. Final acceptance 2022-03-04. Final version published as submitted by the authors.