

**UNIVERSITY OF KWAZULU-NATAL**



*Doctoral Dissertation*

**Novel Improvements of Empirical Wireless Channel Models and  
Proposals of Machine-Learning-Based Path Loss Prediction Models  
for Future Communication Networks**

*Author:*

***Mohamed Khalifa Elmezughi***

*Supervisor:*

***Prof. Thomas J. O. Afullo***

*This dissertation submitted in fulfillment of the requirements  
for the degree of Doctor of Philosophy (Ph.D.): Electronic Engineering  
in the*

School of Engineering

College of Agriculture, Engineering and Science

Durban, South Africa

***August 2022***

**UNIVERSITY OF KWAZULU-NATAL**



**Novel Improvements of Empirical Wireless Channel Models and  
Proposals of Machine-Learning-Based Path Loss Prediction Models  
for Future Communication Networks**

*Author:*

*Mohamed Khalifa Elmezughi*

*Supervisor:*

*Prof. Thomas J. O. Afullo*

School of Engineering

College of Agriculture, Engineering and Science

Durban, South Africa

*August 2022*

## *Preface*

The research contained in this dissertation was completed by the candidate *Mr. Mohamed Khalifa Elmezughi*, under the supervision of *Prof. Thomas J. O. Afullo* in the Discipline of Electrical, Electronic, and Computer Engineering, School of Engineering, College of Agriculture, Engineering, and Science, University of KwaZulu-Natal (UKZN), Howard College Campus, Durban, South Africa. This dissertation has not been submitted in any form to another university, and the results reported are due to investigations by the candidate. The entire contents of this work, except where the work of other researchers is acknowledged in the text, is the original work of the candidate.

*Mohamed Khalifa Elmezughi*

*Signed* \_\_\_\_\_

*Date* \_\_\_\_\_ *August 19<sup>th</sup>, 2022* \_\_\_\_\_

As the candidate's supervisor, I have approved this dissertation for submission.

*Prof. Thomas J. O. Afullo*

*Signed* \_\_\_\_\_

*Date* \_\_\_\_\_ *August 19<sup>th</sup>, 2022* \_\_\_\_\_

## ***Declaration 1 – Plagiarism***

I, **Mohamed Khalifa Elmezughi**, student number **219058758**, pertaining to this dissertation entitled: “*Novel Improvements of Empirical Wireless Channel Models and Proposals of Machine-Learning-Based Path Loss Prediction Models for Future Communication Networks*,” declare that:

- I. The research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;
- II. This dissertation has not been submitted for any degree or examination to any other university;
- III. This dissertation does not contain other persons’ data, pictures, graphs, or other information unless acknowledged explicitly as being sourced from other persons;
- IV. This dissertation does not contain other persons’ writing unless explicitly acknowledged as being sourced from other researches where other written sources have been quoted, then:
  - a) Their words have been re-written, but the general information attributed to them has been referenced;
  - b) Where their exact words have been used, their writing has been placed inside quotation marks and referenced;
- V. Where I have reproduced a publication of which I am the author, co-author, or editor, I have indicated in detail those parts of the publication written by myself alone and have fully referenced such publications;
- VI. This dissertation does not contain text, graphs, or tables copied and pasted from the Internet unless expressly acknowledged and the source being detailed in the References sections of the dissertation;
- VII. This dissertation is primarily a collection of material, prepared by myself; parts of this work are submitted as research articles where I am the first author.

***Mohamed Khalifa Elmezughi***

***Signed*** \_\_\_\_\_

***Date*** \_\_\_\_\_ *August 19<sup>th</sup>, 2022* \_\_\_\_\_

## ***Declaration 2 – Publications***

The following research papers **have been published or submitted for publication**, and parts of their materials are included in the dissertation:

- ***Journal articles directly related to this dissertation:***

- 1) **M. K. Elmezughi** and T. J. Afullo, “An Efficient Approach of Improving Path Loss Models for Future Mobile Networks in Enclosed Indoor Environments,” *IEEE Access*, vol. 9, pp. 110332–110345, Aug. 2021.
- 2) **M. K. Elmezughi**, O. Salih, T. J. Afullo, and K. J. Duffy, “Comparative Analysis of Major Machine-Learning-Based Path Loss Models for Enclosed Indoor Channels,” *Sensors*, vol. 22, no. 13, pp. 4967, Jun. 2022.
- 3) **M. K. Elmezughi** and T. J. Afullo, “Investigations into the Effect of High-Ordering the Log-Distance Dependency of Path Loss Models for Indoor Wireless Channels,” *International Journal on Communications Antenna and Propagation (IRECAP)*, vol. 12, no. 1, Feb. 2022.
- 4) **M. K. Elmezughi**, O. Salih, T. J. Afullo, and K. J. Duffy, “Path Loss Modeling Based on Neural Networks and Ensemble Method for Future Wireless Networks,” *Heliyon, Elsevier*, submitted for publication.

- ***Refereed conference papers directly related to this dissertation:***

- 5) **M. K. Elmezughi** and T. J. Afullo, “Proposal of a Measurement-Based Propagation Channel Model for Future Indoor Wireless Networks,” *Southern Africa 2022 Telecommunication Networks and Applications Conference (SATNAC 2022)*, Aug. 2022, pp. 1-5.
- 6) **M. K. Elmezughi** and T. J. Afullo, “Proposal of a Probabilistic Path Loss Model for Wireless Communications in Indoor Corridors,” *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), IEEE*, Dec. 2021, pp. 1–5.
- 7) **M. K. Elmezughi** and T. J. Afullo, “Evaluation of Line-of-Sight Probability Models for Enclosed Indoor Environments at 14 to 22 GHz,” *2021 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), IEEE*, Sep. 2021, pp. 1–6.

## ***Dedication***

*To my dearest, most precious possession, and beloved mother and father.*

*To all my family members.*

*To the world.*

## *Acknowledgments*

First of all, I would like to offer my full thanks, appreciation, and gratitude to ***Almighty Allah***, who created, taught, and guided me. With the blessings ***Allah*** has bestowed upon me, I am now writing my Ph.D. dissertation (and tears of joy on my cheeks) with health, family, and hope for a wonderful future. I am completely aware that without ***Allah***'s support and help, I would not have reached this level, and this research work would not have seen the light. ***Praise to Allah, Lord of the world.***

Secondly, I am profoundly and forever indebted to my parents, brothers, and sisters for their prayers, warm love, unbelievable support, continued patience, and encouragement throughout my entire life. My expressions of love and appreciation to all my family members will not be enough. They were the ones who helped me in all stages of my life from childhood when my goal in life was just to play until this age when I have changed and have the commitment and responsibility to take care of them in my whole life.

Thirdly, I would like to extend my special thanks, appreciation, and full gratitude to my academic supervisor (***Prof. Thomas J. O. Afullo***) for his valuable guidance, tremendous academic support, suggestions, patience, granting me opportunities to explore my academic interests, and encouragement throughout the whole of this Ph.D. research work.

I also want to express my fondest regards to all my friends who spent memorable times with me and helped me through any stage in my life. My appreciation and thanks especially go to ***Dr. Omran Salih*** for his support in my Ph.D. research work. Special thanks and respect to ***Dr. Ali A. Elghariani***, who supervised me for my *B.Sc.* graduation project entitled: “*Signal Detection and Channel Estimation in LTE Wireless Network*,” with continuous guidance and encouragement since that time. Furthermore, many thanks appreciation to ***Prof. Viranjay M. Srivastava, Dr. Pradeep Kumar, Dr. Naveenbalaji Gowthaman, Dr. Ahmed J. Abougarair, Dr. Nicholas O. Oyie, Dr. Gbadega Peter, Dr. Djuma Sumbiri, Dr. Dauda Ayanda, and Mrs. Florence Chelangat.***

Additionally, my profound gratitude goes to all the members of technical staff, friends, and colleagues in the Discipline of Electrical, Electronic, and Computer Engineering, Howard College Campus, University of KwaZulu-Natal (UKZN), Durban, South Africa. I am incredibly grateful to ***Mr. Nqayi Sphamandla***, the

assistant administrative officer in the Center of Radio Access and Rural Technologies (CRART), for his administrative help and good wishes throughout this work.

Last but by no means least, I would like to dedicate this dissertation to all the wonderful people in my whole life, and I am welcoming to any suggestions that will improve myself for near-future works.

*Mohamed Khalifa Elmezughi*

[m.k.elmezughi@gmail.com](mailto:m.k.elmezughi@gmail.com)

## *Abstract*

Path loss is the primary factor that determines the overall coverage of networks. Therefore, designing reliable wireless communication systems requires accurate path loss prediction models. Future wireless mobile systems will rely mainly on the super-high frequency (SHF) and the millimeter-wave (mmWave) frequency bands due to the massively available bandwidths that will meet projected users' demands, such as the needs of the fifth-generation (5G) wireless systems and other high-speed multimedia services. However, these bands are more sensitive and exhibit a different propagation behavior compared to the frequency bands below 6 GHz. Hence, improving the existing models and developing new models are vital for characterizing the wireless communication channel in both indoor and outdoor environments for future SHF and mmWave services.

This dissertation proposes new path loss and LOS probability models and efficiently improves the well-known close-in (CI) free space reference distance model and the floating-intercept (FI) model. Real measured data was taken for both line-of-sight (LOS) and non-line-of-sight (NLOS) communication scenarios in a typical indoor corridor environment at three selected frequencies within the SHF band, namely 14 GHz, 18 GHz, and 22 GHz. The research finding of this work reveals that the proposed models have better performance in terms of their accuracy in fitting real measured data collected from measurement campaigns. In addition, this research studies the impact of the angle of arrival and the antenna heights on the current and improved CI and FI models. The results show that the proposed improved models provide better stability and sensitivity to the change of these parameters. Furthermore, the mean square error between the models and their improved versions was presented as another proof of the superiority of the proposed improvement. Moreover, this research shows that shadow fading's standard deviation can have a notable reduction in both the LOS and NLOS scenarios (especially in the NLOS), which means higher precision in predicting the path loss compared to the existing standard models.

After that, the dissertation presents investigations on high-ordering the dependency of the standard CI path loss model on the distance between the transmitting and the receiving antennas at the logarithmic scale. Two improved models are provided and discussed: second-order CI and third-order CI models. The main results reveal that the proposed two models outperform the standard CI model and notable reductions in the shadow fading's standard deviation values as the model's order increases, which means that more precision is provided. This part of the dissertation also provides a trade-off study between the model's accuracy and simplicity.

Seeking higher prediction accuracy, the following part of this research work presents and evaluates the performance of several well-known machine learning methods, including multiple linear regression (MLR), polynomial regression (PR), support vector regression (SVR), as well as the methods using decision trees (DT), random forests (RF), K-nearest neighbors (KNN), artificial neural networks (ANN), and artificial recurrent neural networks (RNN). RNNs are mainly based on long short-term memory (LSTM). The models are compared based on measurement data to provide the best fitting machine-learning-based path loss prediction models. The main results obtained from this study show that the best root-mean-square error (RMSE) performance is given by the ANN and RNN-LSTM methods, while the worst is for the MLR method. All the RMSE values for the given learning techniques are in the range of 0.0216 to 2.9008 dB. Furthermore, this research shows that the models (except for the MLR model) perform excellently in fitting actual measurement data for wireless communications in enclosed indoor environments since they provide R-squared and correlation values higher than 0.91 and 0.96, respectively. The results show that these learning methods could be used as accurate and stable models for predicting path loss in the mmWave frequency regime since their precisions are much better than the standard and improved empirical models.

Motivated by the achieved high prediction accuracy of the neural networks, an ensemble-method-based neural network path loss model is proposed in the final part of this dissertation. The model is based on the ANN, RNN-LSTM, and the convolutional neural network (CNN) models. Extensive performance analysis is provided regarding prediction accuracy, stability, the contribution of input features, and the time needed to run the model. The main research finding of this work reveals that the ensemble-method-based model outperforms all the other models (ANN, RNN-LSTM, and CNN) in terms of efficiency and high prediction accuracy and could be trusted as a promising model for path loss in complex environments at high-frequency bands. The research findings of this work are vital for planning, evaluating, and optimizing future wireless communication networks.

# Table of Contents

<i>Preface</i> .....	<i>I</i>
<i>Declaration 1 – Plagiarism</i> .....	<i>III</i>
<i>Declaration 2 – Publications</i> .....	<i>IV</i>
<i>Dedication</i> .....	<i>V</i>
<i>Acknowledgments</i> .....	<i>VI</i>
<i>Abstract</i> .....	<i>VIII</i>
<i>List of Figures</i> .....	<i>XIII</i>
<i>List of Tables</i> .....	<i>XVI</i>
<i>List of Abbreviations</i> .....	<i>XVIII</i>
<b>Chapter 1: General Introduction</b> .....	<i>1</i>
1.1 Introduction.....	<i>1</i>
1.2 Problem Formulation, Motivation, and Contributions of the Dissertation.....	<i>3</i>
1.3 Scope of the Work.....	<i>5</i>
1.4 Objectives.....	<i>5</i>
1.5 Dissertation Structure.....	<i>6</i>
1.6 Contributions to Publications.....	<i>8</i>
<b>Chapter 2: Literature Review</b> .....	<i>11</i>
2.1 Introduction.....	<i>11</i>
2.2 Propagation Mechanisms.....	<i>12</i>
2.2.1 Free Space Propagation.....	<i>12</i>
2.2.2 Reflection.....	<i>13</i>
2.2.3 Diffraction.....	<i>13</i>
2.2.4 Scattering.....	<i>13</i>
2.3 Large-Scale Fading.....	<i>13</i>
2.4 Small-Scale Fading.....	<i>14</i>
2.5 Channel Modeling.....	<i>17</i>
2.6 Related Works.....	<i>17</i>
2.7 Chapter Summary.....	<i>24</i>
<b>Chapter 3: Proposals of Measurement-Based Propagation Channel Models for Wireless Communications in Enclosed Environments</b> .....	<i>26</i>
3.1 Introduction.....	<i>26</i>
3.2 Measurement Setup and Data Collection Method.....	<i>27</i>
3.3 Proposal of a Measurement-Based Path Loss Model for Enclosed Indoor Channels.....	<i>31</i>
3.4 Evaluation of Standard Line-of-Sight Probability Models.....	<i>37</i>

3.5 Proposal of a Probabilistic Path Loss Model.....	42
3.6 Chapter Summary .....	46
<b>Chapter 4: An Efficient Approach of Improving Path Loss Models for Future Mobile Networks .....</b>	<b>48</b>
4.1 Introduction.....	48
4.2 The Standard Path Loss Prediction Modeling.....	48
4.3 The Standard CI Model and Our Improved CI Path Loss Prediction Model .....	50
4.4 The Standard FI Model and Our Improved FI Path Loss Prediction Model.....	63
4.5 Chapter Summary .....	74
<b>Chapter 5: Investigations into the Effect of High-Ordering the Log-Distance Dependency of Path Loss Models .....</b>	<b>76</b>
5.1 Introduction.....	76
5.2 The Third-Order CI Path Loss Prediction Model.....	76
5.3 Results and Discussions on the Impact of High-Ordering the Log-Distance Dependency.....	77
5.4 Results and Discussions Related to Error Analysis of the Proposed Models .....	85
5.5 Chapter Summary .....	91
<b>Chapter 6: Comparative Analysis of Major Machine-Learning-Based Path Loss Models.....</b>	<b>93</b>
6.1 Introduction.....	93
6.2 Data Preparation and Models Setup .....	93
6.2.1 Data Preparation .....	94
6.2.2 Hyperparameters Tuning Setup .....	96
6.2.3 ML-Based Models' Stability.....	96
6.2.4 Evaluation Metrics.....	97
6.3 Machine-Learning-Based Models.....	98
6.3.1 Linear Regression Models .....	98
6.3.2 Support Vector Regression Model .....	99
6.3.3 Decision Tree Regression Model .....	101
6.3.4 Random Forest Regression Model.....	101
6.3.5 K-Nearest Neighbor Regression Model .....	102
6.3.6 Artificial Neural Network Model .....	103
6.3.7 Recurrent Neural Network Model.....	104
6.4 Results and Discussions of the Machine-Learning-Based Models .....	105
6.5 Chapter Summary .....	113
<b>Chapter 7: Proposal of an Ensemble-Method-Based Neural Network Path Loss Model .....</b>	<b>116</b>
7.1 Introduction.....	116
7.2 Ensemble Method for Path Loss Prediction Modeling .....	116

7.2.1 <i>The CNN Model</i> .....	117
7.2.2 <i>The Ensemble-Method-Based Model</i> .....	118
7.3 <i>Results and Discussions about Ensemble method Path Loss Modeling</i> .....	120
7.4 <i>Chapter Summary</i> .....	124
<b>Chapter 8: Conclusions and Recommendations for Related Future Works</b> .....	126
8.1 <i>Conclusions</i> .....	126
8.2 <i>Recommendations for Related Future Works</i> .....	128
<b>References</b> .....	131
<b>Appendices</b> .....	147

## *List of Figures*

Figure 2.1: Motion of the mobile station with respect to the base station. ....	16
Figure 3.1: The adopted indoor corridor environment.....	27
Figure 3.2: The channel sounder architecture. ....	29
Figure 3.3: 3D floor plan of the indoor corridor environment. ....	30
Figure 3.4: The proposed and standard path loss prediction models at 14 GHz frequency band. ....	33
Figure 3.5: The proposed and standard path loss prediction models at 18 GHz frequency band. ....	34
Figure 3.6: The proposed and standard path loss prediction models at 22 GHz frequency band. ....	35
Figure 3.7: The proposed path loss model with the measurement data at 14, 18, and 22 GHz. ....	35
Figure 3.8: Proposed model's prediction error vs. Tx-Rx separation distance.....	36
Figure 3.9: LOS probability models at 14 GHz. ....	39
Figure 3.10: LOS probability models at 18 GHz. ....	40
Figure 3.11: LOS probability models at 22 GHz. ....	41
Figure 3.12: The RMSE values of the LOS probability models. ....	42
Figure 3.13: Path loss data and models at 14, 18, and 22 GHz.....	44
Figure 3.14: PE of the probabilistic CI path loss model. ....	44
Figure 3.15: The MPE and SDE of the probabilistic model with AoA. ....	46
Figure 4.1: The air's attenuation at different frequency bands [46].....	49
Figure 4.2: Comparison of measured path loss, the fitted CI model, and the improved CI model for the LOS results at 14 GHz frequency band. ....	53
Figure 4.3: Comparison of measured path loss, the fitted CI model, and the improved CI model for the LOS results at 18 GHz frequency band. ....	54
Figure 4.4: Comparison of measured path loss, the fitted CI model, and the improved CI model for the LOS results at 22 GHz frequency band. ....	54
Figure 4.5: Directional large-scale path loss prediction models for the LOS communication scenario. ....	55
Figure 4.6: MSE between the CI and improved CI models in the LOS communication scenario.....	55
Figure 4.7: Directional large-scale path loss prediction models for the NLOS communication scenario. .	57
Figure 4.8: MSE between the CI and improved CI models in the NLOS communication scenario.....	57
Figure 4.9: The behavior of the CI model's parameter (i.e., PLE) with the AoA.....	60
Figure 4.10: The behavior of the improved CI model's parameter ( $n_1$ ) with the AoA. ....	60

Figure 4.11: The behavior of the improved CI model's parameter ( $n_2$ ) with the AoA. ....	61
Figure 4.12: The parameters of the CI and improved-CI models at two different Tx antenna heights. ....	62
Figure 4.13: Comparison of measured path loss, the fitted FI model, and the improved FI model for the LOS results at 14 GHz frequency band. ....	65
Figure 4.14: Comparison of measured path loss, the fitted FI model, and the improved FI model for the LOS results at 18 GHz frequency band. ....	65
Figure 4.15: Comparison of measured path loss, the fitted FI model, and the improved FI model for the LOS results at 22 GHz frequency band. ....	66
Figure 4.16: Directional large-scale path loss prediction models for the LOS communication scenario. ...	66
Figure 4.17: MSE between the FI and improved FI models in the LOS communication scenario. ....	67
Figure 4.18: Directional large-scale path loss prediction models for the NLOS communication scenario.	68
Figure 4.19: MSE between the FI and improved FI models in the NLOS communication scenario.....	68
Figure 4.20: The behavior of the FI model's parameter ( $\alpha$ ) with the AoA.....	69
Figure 4.21: The behavior of the FI model's parameter ( $\beta$ ) with the AoA. ....	70
Figure 4.22: The behavior of the improved FI model's parameter ( $\alpha$ ) with the AoA. ....	70
Figure 4.23: The behavior of the improved FI model's parameter ( $\beta_1$ ) with the AoA. ....	71
Figure 4.24: The behavior of the improved FI model's parameter ( $\beta_2$ ) with the AoA. ....	71
Figure 4.25: The parameters of the FI and improved-FI models at two different Tx antenna heights. ....	72
Figure 5.1: LOS Path loss models with measurement data at 14 GHz. ....	79
Figure 5.2: LOS Path loss models with measurement data at 18 GHz. ....	80
Figure 5.3: LOS Path loss models with measurement data at 22 GHz. ....	80
Figure 5.4: LOS Path loss models at 14, 18, and 22 GHz.....	81
Figure 5.5: NLOS Path loss models with measurement data at 14 GHz. ....	83
Figure 5.6: NLOS Path loss models with measurement data at 18 GHz. ....	83
Figure 5.7: NLOS Path loss models with measurement data at 22 GHz. ....	84
Figure 5.8: NLOS Path loss models at 14, 18, and 22 GHz.....	84
Figure 5.9: LOS PE of path loss models at 14 GHz.....	86
Figure 5.10: LOS PE of path loss models at 18 GHz.....	87
Figure 5.11: LOS PE of path loss models at 22 GHz.....	87
Figure 5.12: NLOS PE of path loss models at 14 GHz.....	89
Figure 5.13: NLOS PE of path loss models at 18 GHz.....	89
Figure 5.14: NLOS PE of path loss models at 22 GHz.....	90

Figure 6.1: Flow chart of the adopted ML-based path loss prediction modeling technique. ....	95
Figure 6.2: The principle of the SVR in two dimensions. ....	100
Figure 6.3: The KNN regression model with $K = 2$ . ....	102
Figure 6.4: The architecture of the proposed ANN model. ....	103
Figure 6.5: The architecture of the proposed RNN-LSTM model. ....	104
Figure 6.6: Measured and predicted path loss data for the MLR model. ....	105
Figure 6.7: Measured and predicted path loss data for the PR model. ....	105
Figure 6.8: Measured and predicted path loss data for the DT model. ....	106
Figure 6.9: Measured and predicted path loss data for the RF model. ....	106
Figure 6.10: Measured and predicted path loss data for the SVR model. ....	106
Figure 6.11: Measured and predicted path loss data for the KNN model. ....	107
Figure 6.12: Measured and predicted path loss data for the ANN model. ....	107
Figure 6.13: Measured and predicted path loss data for the RNN-LSTM model. ....	107
Figure 6.14: Prediction error of each ML-based model. ....	109
Figure 6.15: Measured vs. predicted path loss data for each model. ....	110
Figure 6.16: Training and validation loss for both ANN and RNN models. ....	111
Figure 6.17: Measured and predicted path loss and the prediction error of all the ML-based models. (a) Measured and predicted path loss; (b) Prediction error. ....	111
Figure 7.1: Flow chart of the adopted ML-based path loss prediction modeling technique. ....	119
Figure 7.2: Measured data and predicted path loss for all the neural network models. ....	121
Figure 7.3: Prediction error curves of each model. ....	121
Figure 7.4: Predicted and measured path loss values for each model. ....	122
Figure 7.5: Path loss prediction models and prediction error for all the studied models together. ....	123

## *List of Tables*

Table 3.1: Channel sounder specifications and parameters configuration.....	30
Table 3.2: Numerical values of the proposed mode's parameters and the evaluation metrics' values.....	36
Table 3.3: ITU-R, WINNER II (A1), and the proposed LOS probability models.....	38
Table 3.4: The LOS probability models with their RMSE values at 14 GHz.....	39
Table 3.5: The LOS probability models with their RMSE values at 18 GHz.....	41
Table 3.6: The LOS probability models with their RMSE values at 22 GHz.....	42
Table 3.7: The probabilistic model's MPE and SDE for the LOS Communication Scenario.....	45
Table 3.8: The NLOS results of the probabilistic model at 14 GHz frequency band. ....	45
Table 3.9: The NLOS results of the probabilistic model at 18 GHz frequency band. ....	45
Table 3.10: The NLOS results of the probabilistic model at 22 GHz frequency band. ....	46
Table 4.1: A comparison between the CI and Improved-CI models' parameters in the LOS scenario.....	56
Table 4.2: A comparison between the CI and Improved-CI models' parameters in the NLOS scenario. ...	58
Table 4.3: The behavior of the CI and Improved-CI models' parameters with the AoA at 14 GHz frequency band.....	59
Table 4.4: The behavior of the CI and Improved-CI models' parameters with the AoA at 18 GHz frequency band.....	59
Table 4.5: The behavior of the CI and Improved-CI models' parameters with the AoA at 22 GHz frequency band.....	59
Table 4.6: A comparison between the FI and Improved-FI models' parameters in the LOS scenario.....	64
Table 4.7: A comparison between the FI and Improved-FI models' parameters in the NLOS scenario. ...	67
Table 4.8: The behavior of the FI and Improved-FI models' parameters with the AoA at 14 GHz frequency band.....	73
Table 4.9: The behavior of the FI and Improved-FI models' parameters with the AoA at 18 GHz frequency band.....	73
Table 4.10: The behavior of the FI and Improved-FI models' parameters with the AoA at 22 GHz frequency band.....	74
Table 5.1: The LOS Comparative Study Results.....	79
Table 5.2: The NLOS Comparative Study Results.....	82
Table 5.3: The models' MPE and SDE for the LOS communication scenario at 14, 18, and 22 GHz.....	90

Table 5.4: The models' MPE and SDE for the NLOS communication scenario at 14, 18, and 22 GHz....	91
Table 6.1: Performance metrics' values of all the ML-based models selected.....	108
Table 6.2: Runtime comparison of the adopted ML models.....	112
Table 6.3: Performance metrics after removing the antenna height from the input features of the models. .....	113
Table 6.4: Performance metrics after removing the antenna height and the AoA from the input features. .....	113
Table 7.1: Performance metrics' values of all the ML-based models selected.....	122
Table 7.2: Performance metrics after removing the antenna height from the input features of the models. .....	122
Table 7.3: Performance metrics after removing the antenna height and the AoA from the features of the models.....	122
Table 7.4: Runtime comparison of the adopted models.....	123

## *List of Abbreviations*

<i>Abbreviation</i>	<i>Description</i>
<i>1G</i>	<i>First generation wireless communications</i>
<i>2D</i>	<i>Two-dimensions</i>
<i>2G</i>	<i>Second generation wireless communications</i>
<i>3D</i>	<i>Three-dimensions</i>
<i>3G</i>	<i>Third generation wireless communications</i>
<i>3GPP</i>	<i>Third Generation Partnership Project</i>
<i>3GPP-SCM</i>	<i>3GPP-spatial channel model</i>
<i>4G</i>	<i>Fourth generation wireless communications</i>
<i>5G</i>	<i>Fifth generation wireless communications</i>
<i>6G</i>	<i>Sixth generation wireless communications</i>
<i>ABG</i>	<i>Alpha-beta-gamma</i>
<i>ABGX</i>	<i>ABG model with cross-polarization discrimination</i>
<i>ACK</i>	<i>Acknowledgment</i>
<i>ANN</i>	<i>Artificial Neural Network</i>

<b><i>AoA</i></b>	<i>Angle of arrival</i>
<b><i>AoD</i></b>	<i>Angle of departure</i>
<b><i>AZ</i></b>	<i>Azimuth</i>
<b><i>B3G</i></b>	<i>Beyond third generation</i>
<b><i>BS</i></b>	<i>Base Station</i>
<b><i>CDF</i></b>	<i>Cumulative distribution function</i>
<b><i>CI</i></b>	<i>Close-in</i>
<b><i>CIF</i></b>	<i>CI model with frequency-dependent PLE</i>
<b><i>CIFX</i></b>	<i>CI model with frequency-dependent PLE and cross-polarization discrimination</i>
<b><i>CIX</i></b>	<i>CI model with cross-polarization discrimination</i>
<b><i>cmWave</i></b>	<i>Centimeter-wave</i>
<b><i>CNN</i></b>	<i>Convolutional Neural Network</i>
<b><i>CSI</i></b>	<i>Channel state information</i>
<b><i>CW</i></b>	<i>Continuous wave</i>
<b><i>D2D</i></b>	<i>Device-to-device</i>
<b><i>DS</i></b>	<i>Dual-slope</i>

<b><i>DT</i></b>	<i>Decision Tree</i>
<b><i>EHF</i></b>	<i>Extremely high frequency</i>
<b><i>EIRP</i></b>	<i>Effective isotropic radiated power</i>
<b><i>EL</i></b>	<i>Elevation</i>
<b><i>FI</i></b>	<i>Floating-intercept</i>
<b><i>FS</i></b>	<i>Free space</i>
<b><i>FSPL</i></b>	<i>Free space path loss</i>
<b><i>FSPLE</i></b>	<i>Free space path loss exponent</i>
<b><i>GKDE</i></b>	<i>Gaussian Kernel Density Estimation</i>
<b><i>GSCM</i></b>	<i>Geometry-based stochastic channel model</i>
<b><i>GSM</i></b>	<i>Global system for mobile communications</i>
<b><i>H2H</i></b>	<i>Human-to-human</i>
<b><i>H2M</i></b>	<i>Human-to-machine</i>
<b><i>HD</i></b>	<i>High definition</i>
<b><i>HDTV</i></b>	<i>High definition television</i>
<b><i>HPBW</i></b>	<i>Half-power beamwidth</i>
<b><i>IoT</i></b>	<i>Internet of things</i>

<b><i>ISI</i></b>	<i>Inter-symbol interference</i>
<b><i>ITU</i></b>	<i>International telecommunication union</i>
<b><i>ITU-R</i></b>	<i>ITU-Radiocommunication</i>
<b><i>KL</i></b>	<i>Kullback–Leibler</i>
<b><i>KNN</i></b>	<i>K-Nearest Neighbors</i>
<b><i>LMDS</i></b>	<i>Local multipoint distribution service</i>
<b><i>LOS</i></b>	<i>Line-of-sight</i>
<b><i>LTE</i></b>	<i>Long term evolution</i>
<b><i>M2M</i></b>	<i>Machine-to-machine</i>
<b><i>MAF</i></b>	<i>Modal attenuation factor</i>
<b><i>Massive-MIMO</i></b>	<i>Massive multiple-input multiple-output</i>
<b><i>MDP</i></b>	<i>Makrov decision process</i>
<b><i>METIS</i></b>	<i>Mobile and wireless communication enablers for twenty-twenty information society</i>
<b><i>MIMO</i></b>	<i>Multiple-input, multiple-output</i>
<b><i>ML</i></b>	<i>Machine learning</i>
<b><i>MLR</i></b>	<i>Multiple linear regression</i>

<b><i>MMSE</i></b>	<i>Minimum mean square error</i>
<b><i>mmWave</i></b>	<i>Millimeter-wave</i>
<b><i>MPC</i></b>	<i>Multipath component</i>
<b><i>MPE</i></b>	<i>Mean prediction error</i>
<b><i>MS</i></b>	<i>Mobile station</i>
<b><i>MSE</i></b>	<i>Mean square error</i>
<b><i>NIST</i></b>	<i>National institute of standards and technology</i>
<b><i>NLOS</i></b>	<i>Non-line-of-sight</i>
<b><i>OLOS</i></b>	<i>Obstructed-LOS</i>
<b><i>PAN</i></b>	<i>Personal area network</i>
<b><i>PDF</i></b>	<i>Probability density function</i>
<b><i>PDP</i></b>	<i>Power delay profile</i>
<b><i>PLE</i></b>	<i>Path loss exponent</i>
<b><i>PN</i></b>	<i>Pseudo noise</i>
<b><i>QoS</i></b>	<i>Quality of service</i>
<b><i>RAT</i></b>	<i>Radio access technology</i>
<b><i>RDA</i></b>	<i>Rotated directional antenna</i>

<b><i>ReLU</i></b>	<i>Rectified linear unit</i>
<b><i>RF</i></b>	<i>Radio frequency</i>
<b><i>RLOS</i></b>	<i>Reflected LOS</i>
<b><i>RMSE</i></b>	<i>Root mean square error</i>
<b><i>RNN</i></b>	<i>Recurrent neural network</i>
<b><i>Rx</i></b>	<i>Receiver</i>
<b><i>SCM</i></b>	<i>Spatial channel model</i>
<b><i>SC-RMSE</i></b>	<i>Spread corrected root mean square error</i>
<b><i>SDE</i></b>	<i>Standard deviation error</i>
<b><i>SF</i></b>	<i>Shadow fading</i>
<b><i>SHF</i></b>	<i>Super high frequency</i>
<b><i>SINR</i></b>	<i>Signal-to-noise plus interference ratio</i>
<b><i>SISO</i></b>	<i>Single-input, single-output</i>
<b><i>SNR</i></b>	<i>Signal-to-noise ratio</i>
<b><i>SRCM</i></b>	<i>Stochastic radio channel model</i>
<b><i>SUI</i></b>	<i>Stanford University Interim</i>
<b><i>SVR</i></b>	<i>Support vector regression</i>

<b><i>TVD</i></b>	<i>Total variation distance</i>
<b><i>Tx</i></b>	<i>Transmitter</i>
<b><i>V2V</i></b>	<i>Vehicle to vehicle</i>
<b><i>VNA</i></b>	<i>Vector network analyzer</i>
<b><i>WINNER</i></b>	<i>Wireless world initiative new radio</i>
<b><i>WLAN</i></b>	<i>Wireless local area network</i>
<b><i>XPD</i></b>	<i>Cross-polarization discrimination</i>

# CHAPTER 1

## GENERAL INTRODUCTION

---

## *Chapter 1: General Introduction*

### **1.1 Introduction**

Era after era, the demand for higher mobile data traffic is exponentially increasing due to the tremendous revolution in technologies relying totally on mobile networks and their services. In 2019, Cisco reported that by the end of 2022, the number of networked connections and devices would grow up to 28.5 *billion*, and mobile-ready applications would reach up to 12.3 *billion* of them [1], [2]. Also, it was expected that by 2023 the overall mobile data traffic would be increased to approximately 77 exabytes per month, which is a seven-fold increase over the year of 2017 [2]. Another vital issue that has led to more data needs is the circumstances that the world faces nowadays, such as the Coronavirus pandemic that made most of studying and working online, which has led to enormous use of the internet connection. Of course, this requires massive amounts of bandwidth [3]-[6].

Previously, the congestion of the spectrum below 6 GHz was enough to meet the existing systems' requirements. On the contrary, these bands will not meet the necessities of the fifth-generation (5G) cellular networks and many other applications due to their relative shortage in the bandwidth required [7]-[10]. Because of that, research has been done to adopt the frequency regime above 6 GHz as a promising solution to accomplish high peak data transmission rates up to multi gigabits per second with the contribution of complex modulation schemes and massive multiple-input multiple-output (Massive-MIMO) systems and other advanced techniques such as beamforming [12]-[16]. All these solutions together will satisfy the requirements of the 5G system. To the best of our knowledge, the research beyond 5G systems towards the sixth-generation (6G) cellular networks has already been started by escaping the current solutions of the 5G systems aiming for better performance [1], [17]. In June 2018, the first commercial 5G mobile communication standard (3GPP Release 15) was completed after many years of research and development. In the middle of 2019, some countries had already deployed 5G cellular networks. Meanwhile, the first 5G-enabled smart devices are now available on the market [1].

It is widely known that when a wireless communication system sends signals from the transmitting antenna(s), the wireless signals will have a reduction of their power as they travel through the wireless communication channel to the receiving antenna(s). This signal loss is well-known as *path loss*, which is the dominant component of the large-scale fading effects. Large-scale fading models play a vital role in optimizing base station deployments, estimating radio coverage, and characterizing the radio environment to quantify the performance of wireless communication systems [18]. Furthermore, efficient and reliable

determination of crucial factors, such as the signal field strength, carrier-to-interference (C/I) ratio, and signal-to-noise ratio (SNR), can be achieved if in-depth knowledge of propagation loss is provided [19]. As a result, for network planning and implementation of wireless communication systems, radio propagation channel characteristics in various environments are required [20].

The performance of wireless communication systems is influenced by radio propagation in physical environments since the radio waves often experience fading. The wirelessly propagated signals from any communication system's transmitting antenna(s) suffer from attenuations over distances and frequencies, which are well-known as large-scale fading and small-scale fading. In addition, the signals also experience losses because of atmospheric conditions and surrounding physical objects, leading to multipath propagation since the receiving antenna(s) receives the signals mainly from reflections, diffractions, and scattering mechanisms [21]. These multipath effects result in signal power fluctuation and increase the uncertainty of received signal power [22]. This work mainly focuses on developing large-scale path loss models that are crucial for estimating radio coverage, allocating frequencies properly, optimizing base stations, and identifying the most suitable antennas [18], [23].

Path loss is a vital component that must be modeled accurately to have reliable system design and link budget analysis. Moreover, the knowledge of the path loss provides statistically averaged (space and time) radio channel conditions. Consequently, researchers need to present more accurate path loss prediction models that can precisely describe the reduction of the wireless signal levels and accurately fit the real measured data collected in different indoor and outdoor environments over a wide range of frequency regimes. The reason behind the inability of the traditional models to be reliable models for the super-high frequency (SHF) and the millimeter-wave (mmWave) frequency bands and beyond is the significant sensitivity of the signals at these frequency bands to the propagation mechanisms in the communication channel. As an example, the mmWave signals provide substantial path loss values in the first meter away from the transmitting antenna and have a significant penetration loss through solid materials such as concrete walls [5], [24]-[27].

In general, the path loss can be modeled deterministically (theoretically), empirically (statistically), or stochastically [24]. The best understanding of the wireless channels' propagation characteristics can be done based on measurement campaigns in propagation's real environments [28]. In this dissertation, we adopted measurement-based (semi deterministic) models to predict the path loss taking into account the propagation mechanisms such as reflections and diffractions and the waveguiding effect that occurs mainly in enclosed indoor environments such as corridors.

## 1.2 Problem Formulation, Motivation, and Contributions of the Dissertation

It is worth mentioning that the mmWave and higher frequency regime have different propagation behavior from the sub-6 GHz frequencies [29]. As such, the problem of the dynamic blockage by humans was not a significant factor to be considered in communications using the sub-6 GHz. However, due to the short-wavelength, mmWave and higher frequency bands are sensitive to this blockage since it has a significant signal loss. Another issue for the mmWave bands is the poor diffraction as a propagation mechanism [30], [31]. Hence, based on these factors, and many others, accurate channel modeling over the SHF, mmWave, and sub-THz frequency bands is highly needed to design and evaluate future wireless communications networks [32]-[35].

Extensive measurement campaigns have been carried out worldwide in different outdoor and indoor environments with various communication techniques, including line-of-sight (LOS), non-line-of-sight (NLOS), and other scenarios, to provide propagation models for each specific frequency band and communication environment. Recently, most research aimed to characterize and model the wireless channel has focused on specific path loss models because of their suitability, such as the close-in (CI) free space reference distance, floating-intercept (FI), and/or alpha-beta-gamma (ABG) model [36]-[56]. The improvement of these models in the literature was based on a consideration of some factors like cross-polarization discrimination (XPD), taking into account the mismatching of the antennas' polarization as in the models named CIX and ABGX models, which are an improvement of the CI and ABG models, respectively [27], [28]. Another improvement of the CI model is by presenting the path loss exponent (PLE) term as a frequency-dependent factor, as in the CIF model [27], [50], [57]-[59]. The last two factors (frequency-dependent PLE and XPD) were considered in one model called the CIFX model [27]. A dual-slope CI path loss model was presented in [60] and [61]. This model provides higher precision in predicting the path loss than the standard CI model. Note that all these improvements can be implemented easily on the FI and ABG models. Other improved models based on other different concepts can be found in [24], [38], [58], [62]-[64]. However, the question that has motivated a part of this research is, *how can we improve the accuracy and reduce the standard deviation of the shadow fading of these standard path loss models without adding parameters that depend on something else like antennas' height or the XPD?* A basic rule states that *any linear equation is a polynomial equation with zero coefficients in higher orders*, which provides the solution to this question. The linear relationship between the path loss and the logarithmic scale of the separation distance between the transmitting and receiving antennas characterizes the common path loss models (such as the CI and FI models), as is widely known. We provide an extra parameter to make these models a function of the squared logarithm of the transmitter-receiver (Tx-Rx) separation

distance. This adopted improvement provides more precision in predicting the path loss, as will be proved in this dissertation.

Many comparative studies between the existing path loss prediction models show the CI and FI models' preference over other models like the ABG model [27], [50], [65], [66]. The CI and FI models offer a precise estimation of the large-scale path loss as a function of the 3D Tx-Rx separation distance over the SHF and mmWave frequency regimes [28], [52], [66]. Hence, we hereby propose to further improve these two models while avoiding a significant increase in the models' complexity to be used by engineers in the wireless system design and calculations of the link-budget since the total number of the improved models' parameters is within a suitable range as other well-known models such as the ABG model have. Furthermore, motivated by the need for more accurate path loss prediction models, this study presents investigations on the impact of high ordering the log-distance dependency of the path loss on the models' performance.

In addition, to the best of our knowledge, there is a research gap in understanding and modeling the LOS probability models for indoor corridor environments in the SHF band. Hence, this work tries to fill the gap by providing LOS probability models based on real measured data collected in a typical indoor corridor environment at 14, 18, and 22 GHz. Also, this research proposes a new LOS probability model that performs better than the standard ITU and WINNER II models. Furthermore, this LOS probability model will be used for proposing a probabilistic path loss prediction model to tackle the classification problem between LOS and NLOS communication scenarios. The difficulties in evaluating mobile networks at mmWave and higher frequencies due to the significant distinction between LOS and NLOS communications served as the impetus for the development of this model. Since it is generally known that the mmWave frequency bands suffer from propagation mechanisms and have higher penetration and reflection losses than the frequency regime, the difference in these frequency bands is clear. The key benefit of this proposed model, in addition to its precision and efficacy, is that it does not need knowledge of the LOS and NLOS communication conditions for the mobile equipment because the probability of having a clear LOS connection is already taken into account in the proposed LOS probability model.

To the best of our knowledge, there is a severe gap in organizing and concluding the previously conducted efforts for path loss prediction based on machine learning (ML) methods. As it is well-known, all ML-based models mainly depend on the datasets for the models' training. These datasets are brought from different environments and communication scenarios at various frequency bands for several applications. Almost all of the existing related works have proposed specific ML-based (or deep-learning-based) models and compared their performance with the traditional empirical and/or a few other ML-based path loss models to show better prediction accuracy. However, until now, there is no way to guess which

ML method is the best for a specific radio propagation environment since the real measurement data will be similar in such environments. This goal can be achieved by comparing several ML-based path loss prediction models with the same datasets and input features and running away from the thinking of comparing with the traditional linear models since ML-based models already perform better according to literature due to the ability to create a complex non-linear relationship between their inputs and outputs. Motivated by that, this work attempts to fill this gap for a typical enclosed indoor corridor environment by providing a comparative analysis of several relevant techniques used for path loss prediction. This study will give other researchers an insight into the best ML-based model for enclosed indoor small-cell communications in the SHF and mmWave frequency regimes.

The final part of this dissertation's primary goal is to offer the most reliable and accurate path loss prediction model possible. In order to do this, we propose and examine the performance of an ensemble-method-based model for path loss prediction. To the best of our knowledge, based on a literature review, this is the first effort that uses ensemble-method-based neural networks for predicting path loss. All the existing studies exploit only separate algorithms for this objective. The proposed model is built based on three neural network models: artificial neural networks (ANN), artificial recurrent neural networks (RNN) based on long short-term memory (LSTM), shortly known as RNN-LSTM, and convolutional neural networks (CNN). These research findings will contribute to the body of knowledge by providing accurate and stable path loss prediction models for future wireless communications.

### **1.3 Scope of the Work**

The ultimate goal of this research is to provide the most accurate path loss prediction modeling mainly in two ways: 1) improving the existing standard semi-deterministic models and providing new models that perform better for future wireless communications in enclosed indoor channels. 2) Compare the performance of several ML-based path loss models to propose the best-fit ML-based model using the ensemble method. Several evaluations and performance analyses will be presented in this work. These studies are mainly based on real measurement data collected in an indoor corridor environment that exists on the fifth floor of the Discipline of Electrical, Electronic, and Computer Engineering, Howard College Campus, University of KwaZulu-Natal, Durban 4001, South Africa.

### **1.4 Objectives**

The main aim of this dissertation is to propose novel methods for improving the performance of well-known standard path loss models and propose new measurement-based and machine-learning-based models for future wireless communications in enclosed indoor environments. The specific objectives of this research are as follows:

- To present, analyze, and prepare measurement data for learning and non-learning path loss modeling. The data was collected from measurement campaigns conducted at three crucial frequency bands, namely 14 GHz, 18 GHz, and 22 GHz. In addition, the LOS and NLOS communication scenarios were considered in the measurement campaigns.
- To propose a novel approach for improving the prediction accuracy of the well-known close-in (CI) free space reference distance and floating-intercept (FI) path loss prediction models.
- To provide performance evaluation of the standard and proposed models in terms of their stability and behavior with the change of vital factors such as the operating frequency, the angle of arrival (AoA), and the antenna height.
- To propose a new LOS probability model for enclosed indoor corridor channels. This model will be compared to standard LOS probability models such as ITU and WINNER II.
- To propose a probabilistic path loss prediction model has the advantage that it does not require knowledge of the LOS and NLOS communication conditions since it considers the probability of having a clear LOS path between the transmitting and receiving antennas.
- To provide an extensive comparative analysis of several machine-learning-based methods for path loss modeling. The evaluation will include the prediction accuracy, stability, complexity, run time, and the input features' contribution to the models' overall performance.
- To present a novel path loss prediction model based on the ensemble method. Three neural networks will be considered in proposing this model. Finally, an extensive performance evaluation will be performed to prove the superiority of the proposed model.

## 1.5 Dissertation Structure

After the introduction chapter, the remainder of this dissertation is structured as follows:

### *Chapter 2*

In this chapter, a brief review of the propagation mechanisms in wireless channels, large- and small-scale fadings, and the wireless channel models is presented. Moreover, the chapter discusses in detail the related works that focused on path loss and LOS probability modeling based on measurements and machine learning. The related works are thoroughly discussed to open up the vision of the research gaps in this field of study.

### *Chapter 3*

This chapter contains a complete description of propagation measurements carried out in an enclosed indoor corridor environment at the adopted frequency bands and communication scenarios. Moreover, a

detailed derivation of a proposed path loss prediction, LOS probability, and probabilistic models are presented. Finally, this chapter discusses the main results of these proposed models.

#### ***Chapter 4***

In this chapter, a novel approach for improving the well-known CI and FI path loss prediction models is provided in detail. Both the LOS and NLOS communication scenarios are considered in this research. Furthermore, the chapter presents a comprehensive comparative analysis of the standard models and the proposed improved models, taking into account vital factors such as LOS and NLOS communications, the impact of the AoA, and investigation of the antenna height on the models' parameters.

#### ***Chapter 5***

This chapter mainly focuses on investigating the impact of high-ordering the log distance dependency of the standard CI path loss model. Two models are proposed and compared with the CI model for both the LOS and NLOS conditions at all the adopted frequency bands. Also, the chapter provides a performance analysis of these models using several error metrics such as the prediction error, mean prediction error (MPE), and standard deviation error (SDE). Finally, the chapter discusses the tradeoff between the models' accuracy and complexity.

#### ***Chapter 6***

In this chapter, a comprehensive comparative study of several well-known machine-learning-based path loss prediction models is provided. More specifically, the chapter presents and evaluates the performance of path loss models using multiple linear regression (MLR), polynomial regression (PR), support vector regression (SVR), as well as the methods using decision trees (DT), random forests (RF), K-nearest neighbors (KNN), artificial neural networks (ANN), and artificial recurrent neural networks (RNN). RNNs are mainly based on long short-term memory (LSTM). The models are compared based on measurement data to provide the best fitting machine-learning-based path loss prediction models. Moreover, the run time analysis is performed on all these models as another aspect in comparing them. Finally, the chapter discusses the input features' contribution to the model's overall performance.

#### ***Chapter 7***

This chapter proposes a novel neural-networks-based path loss prediction model using the ensemble method for future indoor wireless communications. The model is evaluated and compared to the ANN, RNN-LSTM, and the convolutional neural network (CNN), which are the most accurate ML-based models, as found in chapter 6. The comparative analysis has several aspects, including the model's prediction

accuracy and stability, the impact of the input features on the overall performance, and the required run time of the model. The performance metrics adopted for evaluating the machine learning models are the R-squared, root mean squared error (RMSE), mean absolute percentage error (MAPE), mean square error (MSE), and the correlation (Corr) coefficient.

## *Chapter 8*

In this chapter, the main findings of this research work and recommendations for related future works are presented.

### **1.6 Contributions to Publications**

The following Peer-Reviewed journal and conference papers are the main contributions of this dissertation:

- ***Journal articles directly related to this dissertation:***

- 1) **M. K. Elmezughi** and T. J. Afullo, “An Efficient Approach of Improving Path Loss Models for Future Mobile Networks in Enclosed Indoor Environments,” *IEEE Access*, vol. 9, pp. 110332–110345, Aug. 2021.
- 2) **M. K. Elmezughi**, O. Salih, T. J. Afullo, and K. J. Duffy, “Comparative Analysis of Major Machine-Learning-Based Path Loss Models for Enclosed Indoor Channels,” *Sensors*, vol. 22, no. 13, pp. 4967, Jun. 2022.
- 3) **M. K. Elmezughi** and T. J. Afullo, “Investigations into the Effect of High-Ordering the Log-Distance Dependency of Path Loss Models for Indoor Wireless Channels,” *International Journal on Communications Antenna and Propagation (IRECAP)*, vol. 12, no. 1, Feb. 2022.
- 4) **M. K. Elmezughi**, O. Salih, T. J. Afullo, and K. J. Duffy, “Path Loss Modeling Based on Neural Networks and Ensemble Method for Future Wireless Networks,” *Heliyon, Elsevier*, submitted for publication.

- ***Refereed conference papers directly related to this dissertation:***

- 5) **M. K. Elmezughi** and T. J. Afullo, “Proposal of a Measurement-Based Propagation Channel Model for Future Indoor Wireless Networks,” *Southern Africa 2022 Telecommunication Networks and Applications Conference (SATNAC 2022)*, Aug. 2022, pp. 1-5.

- 6) **M. K. Elmezughi** and T. J. Afullo, “Proposal of a Probabilistic Path Loss Model for Wireless Communications in Indoor Corridors,” *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), IEEE*, Dec. 2021, pp. 1–5.
- 7) **M. K. Elmezughi** and T. J. Afullo, “Evaluation of Line-of-Sight Probability Models for Enclosed Indoor Environments at 14 to 22 GHz,” *2021 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), IEEE*, Sep. 2021, pp. 1–6.

# CHAPTER 2

## LITERATURE REVIEW

---

## *Chapter 2: Literature Review*

### **2.1 Introduction**

Mobile communication has proliferated since its debut due to its flexibility and convenience. Due to the continuous evolution of communication technologies and the exponentially increasing demand for higher mobile data traffic, research has been focused on the frequency regime above 6 GHz to overcome the congestion of the previous bands (below 6 GHz) and to cope with the requirements of the 5G wireless system and other high-speed multimedia services [67], [68].

Radio propagation models representing path loss are essential to ensure high-quality services and accurate signal coverage predictions in mobile communication networks. Accordingly, researchers have accelerated their efforts to provide reliable models for various environments and scenarios over a wide range of frequency regimes to assist network engineers in designing reliable future wireless networks and accurate link budget calculations. Moreover, accurate predictions could be beneficial in radio resource management schemes that aim to meet specific Quality-of-Service (QoS) criteria [69]-[76].

Fundamentally, traditional approaches to path loss prediction modeling are deterministic, empirical, and stochastic. Deterministic path loss prediction models are site-specific and require sufficient information about the propagation's environments. These models are often related to 3-D map propagations, such as the ray-tracing models. Moreover, these deterministic models repeat their calculations when the environment changes; therefore, they have high computational complexity. The empirical path loss models are based on measurements and observations, such as the Hata and COST 231 models. These models are easier to apply; however, they are time-consuming since they require massive measurement campaigns in specific environments and communication scenarios. Further, in terms of prediction accuracy, these models provide less accuracy than deterministic models [18], [23], [77]. Many other well-known empirical path loss prediction models are derived based on the criteria of the minimum mean square error (MMSE) between the models and the measurement data to provide the best fit of these models as a function of the separation distance between the transmitting and receiving antennas in the logarithmic scale. They include the single-frequency close-in (CI) free space reference distance model and the single-frequency floating-intercept (FI) model [78]-[82]. Stochastic path loss models have probability distributions and assumptions to be considered in the models' equations. These models suffer from limited precision because of some mathematical expressions since the communications environments are considered random variables [18].

In this chapter, the focus will be on presenting a brief review of the wireless propagation mechanisms, large-scale and small-scale fading effects, and channel modeling. Moreover, the last part of this chapter provides a comprehensive literature survey about the related works in this field of research.

## 2.2 Propagation Mechanisms

Many propagation phenomena affect the electromagnetic waves that travel through the wireless channel between the transmitters and receivers. These effects change the amplitude, phase, and direction of propagating signals. To understand and evaluate the performance of wireless communication systems, all of these consequences should be taken into account in detail. Also, the boundaries between the wireless media (between the ground and air, between the air and buildings, between space and earth, etc.) must be considered. The mobile channels are modeled as time-varying communication paths between the Tx and Rx antennas. A brief description of some of the propagation mechanisms is provided in the context of this chapter.

### 2.2.1 Free Space Propagation

Free space is an ideal propagation medium. Consider an isotropic point source fed by a transmitter of  $P_t$  watts. At an arbitrary, large distance  $d$  from the source, the radiated power is uniformly distributed over the surface area of a sphere. Thus, the received signal power  $P_r$  at distance  $d$  is given by [83]-[85]:

$$P_r = \frac{A_e G_t P_t}{4\pi d^2}; \quad (2.1)$$

where  $A_e$  is the effective area covered by the transmitter and  $G_t$  is the transmitting antenna gain. The relationship between an effective aperture and the receiving antenna gain  $G_r$ , derived in (2.1), can be given by:

$$G_r = \frac{4\pi A_e}{\lambda^2}; \quad (2.2)$$

where  $\lambda$  is the wavelength of the electromagnetic wave. By substituting  $A_e$  of Eq. (2.2) into Eq. (2.1), we obtain:

$$P_r = \frac{G_t G_r P_t}{\left(\frac{4\pi d}{\lambda}\right)^2}. \quad (2.3)$$

Free space path loss (FSL) is defined as:

$$FSL = \frac{P_t}{P_r} = \frac{1}{G_t G_r} \left(\frac{4\pi d}{\lambda}\right)^2. \quad (2.4)$$

Basically, FSL indicates the amount of power lost in the space. A larger loss implies the use of a higher transmitting power level, as the received signal strength must be at some minimal power level for correct reception at the receiving end. When  $G_t = G_r = 1$ , the free space loss is given by:

$$FSL = \left(\frac{4\pi d}{\lambda}\right)^2. \quad (2.5)$$

### 2.2.2 Reflection

Propagating wave impinges on an object that is larger as compared to its wavelength (for example, the surface of the earth, tall buildings, and large walls) [84]-[86].

### 2.2.3 Diffraction

Radio path between a transmitter and a receiver is obstructed by a surface with sharp irregular edges (for example, waves bend around the obstacle, even when line of sight (LOS) does not exist) [84]-[86].

### 2.2.4 Scattering

When objects are smaller than the wavelength of the propagating wave (for example, foliage, street signs, lamp posts), the incoming signal is scattered into several weaker outgoing signals [84]-[86].

## 2.3 Large-Scale Fading

It is vital to understand the characteristics of mobile radio propagation for accurate systems' design. As the name suggests, the large-scale fading happens over vast distances; this type of fading also called local mean attenuation. The large-scale fading is mainly because of the path loss of propagating waves as a function of distance, and the shadowing caused by large obstacles in the signal propagation path such as hills and buildings. In general, the large-scale fading is independent of frequency. This type of fading is related to the design and analysis of the link budget in telecommunication systems [83], [87].

The path loss can be defined as a reduction in the power of the propagating signals as the separation distance between the transmitting and receiving antennas increases. This significant component must be considered in the analysis and design of the link budget in telecommunication systems. The values of path loss in typical communication systems can be more than 150 dB, depends on the designed coverage area, and other factors will be covered in the next chapter [84], [85].

Path loss prediction models can be classified into two types; the first one depends on a physical anchor that catches path loss near the Tx, e.g., the CI model and CI model with frequency-dependent path loss exponent (CIF). The second one mainly depends on the mathematical curve or surface that fits the measured data, e.g., the FI and Alpha-Beta-Gamma (ABG) models [88].

## 2.4 Small-Scale Fading

It refers to the rapid changes in radio signal amplitude in a short period of time or travel distance due to the constructive and destructive interference of the multiple signal paths between the transmitter and the receiver. Small-scale multipath fading is frequency-dependent, and it is relevant to the design of reliable and efficient communication systems [89]. Each path has certain attenuation and delay; hence we can write the impulse response of the multipath channel as [83]-[87]:

$$h(t) = \sum_{k=0}^{L-1} a_k \delta(t - \tau_k); \quad (2.6)$$

where  $L$  is the number of the paths,  $a_k$  and  $\tau_k$  is the attenuation and delay of the  $k^{th}$  path respectively.

The main effects of multipath fading include the following:

- Rapid changes in signal strength over a small travel distance or time interval.
- Random frequency modulation due to varying Doppler shifts on different multipath signals.
- Time dispersion or echoes caused by multipath propagation delays [83]-[87].

There are other physical factors besides the multipath fading that influence small-scale fading in the radio propagation channel include:

- Speed of the mobile: The relative motion between the base station and the mobile results in random frequency modulation due to different Doppler shifts on each of the multipath components.
- Speed of surrounding objects: If objects in the radio channel are in motion, they induce a time varying Doppler shift on multipath components. If the surrounding objects move at a greater rate than the mobile, then this effect dominates fading.
- Transmission Bandwidth of the signal: If the transmitted radio signal bandwidth is greater than the bandwidth of the multipath channel (quantified by coherence bandwidth), the received signal will be distorted [87]-[89].

Multipath delay nature of the channel is quantified by delay spread and coherence bandwidth. The time-varying nature of the channel caused by movement is quantified by Doppler shift and coherence time. In order to compare different multipath channels and develop the design of the wireless systems, The mean excess delay, delay spread, and excess delay spread ( $X$  dB) are multipath channel parameters determined from a power delay profile and used to quantify the multipath [86]. The mean excess delay is the first moment of the power delay profile and it is defined by the following equation to be [86]:

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2}. \quad (2.7)$$

The maximum excess delay ( $X$  dB) of the power delay profile is defined to be the time delay during which multipath energy falls to  $X$  dB below the maximum.

*Delay spread* is a major parameter used to quantify the multipath. There are two metrics of delay spread, maximum delay spread ( $\sigma_{max}^{\tau}$ ), and the root mean square (RMS) delay spread ( $\sigma_{rms}^{\tau}$ ). The maximum delay spread is defined as the difference in propagation time between the longest and shortest paths. Since there is a very large number of paths and not all of these affects the system because the amplitude of these components is very low which we can neglect. RMS delay spread usually used instead of the maximum delay spread. The RMS delay spread is the square root of the second central moment of the power delay profile and is defined to be [84], [86]:

$$\sigma_{rms}^{\tau} = \sqrt{\bar{\tau}^2 - (\bar{\tau})^2}. \quad (2.8)$$

These delays are measured relative to the first detectable signal arriving at the receiver at ( $\tau_0 = 0$ ). Typical values of the delay spread are on the order of microseconds in outdoor and on the order of nanoseconds in indoor radio channels [90]-[92].

It is possible to obtain an equivalent description of the channel in the frequency domain using its frequency response characteristics similar to the delay spread parameters in the time domain, coherence bandwidth is used to characterize the channel in the frequency domain. The RMS delay spread and coherence bandwidth are inversely proportional to one another, although their exact relationship is a function of the exact multipath structure [84]-[86].

*Coherence bandwidth* is a statistical measure of the range of frequencies over which the channel can be considered "flat" (i.e., a channel which passes all spectral components with approximately equal gain and linear phase). Coherence bandwidth also defined as the range of frequencies over which any two frequency components have a strong potential for amplitude correlation [83], [92].

If the Tx and Rx are not fixed with respect to each other (i.e. there is a difference in their relative velocity), the received signal at the receiver will not have the same frequency as the transmitted one due to *Doppler phenomenon*. This can be seen clearly in mobile systems whereas the base station is fixed and the mobile stations are mobile. Therefore the frequency of the received signal will not be the same as the source, it will be different [91], [92].

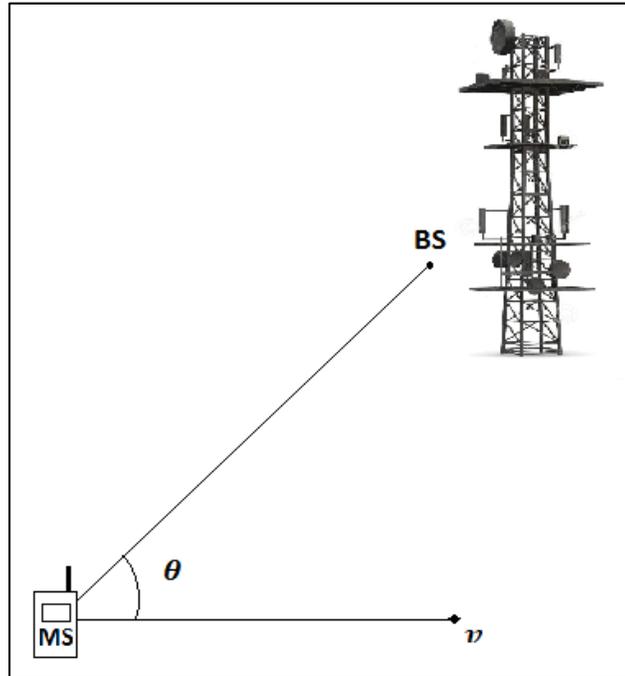
When they are moving toward each other, the frequency of the received signal is higher than that of the source. When they are moving away from each other, the received frequency will decrease. Thus, the frequency of the received signal  $f_r$  is [86]:

$$f_r = f_c - f_d; \quad (2.9)$$

where  $f_c$  is the frequency of source carrier and  $f_d$  is the Doppler shift. Doppler frequency shift is calculated using the following formula [86]:

$$f_d = \frac{v}{\lambda} \cos(\theta); \quad (2.10)$$

where  $v$  is the moving speed,  $\lambda$  is the wavelength of the carrier, and  $\theta$  is the angle between the directions of the transmitter and the receiver. The following figure illustrates the Doppler spread due to the motion of the mobile station (MS) with respect to the base station (BS).



*Figure 2.1: Motion of the mobile station with respect to the base station.*

*Doppler spread* is a measure of the spectral broadening caused by the time rate of change of the mobile radio channel and is defined as the range of frequencies over which the received Doppler spectrum is essentially non-zero. When a pure sinusoidal tone of frequency is transmitted, the received signal spectrum, called the Doppler spectrum, will have components in the range  $f_c - f_d$  to  $f_c + f_d$ . The amount of spectral broadening depends on the relative velocity of the mobile, and the angle  $\theta$  between the direction of motion of the mobile and direction of arrival of the scattered waves [84], [86].

*Coherence time* is the time domain dual of Doppler spread and is used to characterize the time varying nature of the channel in the time domain. The Doppler spread and coherence time are inversely proportional to one another. Coherence time is the duration over which the impulse response of the channel is invariant and it is a statistical quantity that quantifies the similarity of the channel response at different times [93].

## 2.5 Channel Modeling

It is known that the characteristics of the wireless signal changes as it travels from the transmitter antenna to the receiver antenna. The profile of received signal can be obtained from that of the transmitted signal if we have a model of the medium between the two. This model of the medium is called channel model. Ideally, modeling a channel is calculating all the physical processes effecting a signal from the transmitter to the receiver.

Due to reflection, refraction and scattering mechanisms of the EM waves, radio propagation can be roughly described by three nearly independent phenomena as following:

- Path loss.
- Shadow fading.
- Multipath fading.

In this research work, improvements on well-known path loss prediction models (the CI and FI models) and proposals for new path loss and LOS probability models are described and derived in the following chapters.

## 2.6 Related Works

With the continued usage of video and the growth of the Internet-of-Things (IoT), wireless data traffic has been increasingly growing at a rate of over 50% per year per subscriber, and this increase is projected to intensify over the next decade [94]-[96]. The SHF and mmWave frequency bands are promising solutions to cope with the increasing demand for higher data rates. These bands have a massive amount of bandwidth and comparable performance. However, their behavior is far from the frequency bands below 6 GHz that are used for the fourth-generation (4G) cellular system.

The SHF and mmWave bands suffer from higher path loss values, and they are affected by the propagation mechanisms more than the previous bands. In addition, the propagation at these bands has a significant penetration loss through solid materials such as concrete walls [97]-[99]. Hence, indoor limited-range communications have a lot of interest in academia and industry as vital communication environments for future mobile networks [100].

Technically speaking, from the wireless channel's perspective, indoor hotspot (IH), urban micro (UMi), urban macro (UMa), and rural macro (RMa) are the main deployment scenarios of the fifth-generation (5G) system [101]. Channel modeling is the most crucial part that is highly considered during the evolution and simulation of wireless networks. It gives a unified channel condition for each proposed scheme [101]. In

recent years, deriving large-scale networks' performance has been widely used by stochastic geometry instead of various deterministic cells [102], [103]. In order to meet several needs of modern society, academia and industry have accelerated the research to provide more advanced wireless technologies that allow users to have higher capacity and data rates of multi gigabits per second [104].

In [104], height-dependent LOS probability models for unmanned aerial vehicle (UAV) networks were studied. Their findings indicate that in UAV networks, the LOS probability function's long tail has a significant output effect on sparse networks with low base station density. A statistical LOS probability model based on the 3rd generation partnership project (3GPP) and the international telecommunication union radio communication sector (ITU-R) models is presented in [105]. The study improved the LOS probability by splitting the region into LOS and NLOS zones based on the building density. Moreover, the model was checked against visibility predictions extracted from digital building data to see how reliable it was.

Frequency-dependent LOS probability models have been investigated for built-up outdoor environments [106]. The models were derived for arbitrary sizes, heights, and orientations of buildings based on the geometry-based stochastic method. The main contribution was the simplicity of the proposed model with careful consideration of the effect of signals' wavelength and the height of the antennas. This work's main results reveal that there are higher values of the LOS probability when the frequency increases. In fact, due to the randomness of the signals' attenuation with the frequency, measurement-based investigations should be conducted to understand the LOS probability models' behavior with the frequency since this work was based on simulations using the Ray-Tracing method.

LOS Probability models for indoor office and shopping mall environments are presented in [107]. The paper discussed the ITU, WINNER II (A1), and WINNER II (B3) models. Also, they proposed a model based on improving the ITU model. The work's main research finding shows a slight preference for their proposed model over the standard models. However, this improvement is not significant compared to the increase of the model's complexity to be used by design engineers since the maximum reduction of the mean square error (MSE) was only 5.1%. In [108], LOS probability models were presented using point cloud data obtained with laser scanning. The models were derived for shopping malls, offices, and urban open square scenarios, mainly at the 63 GHz frequency band. Also, the paper considered the impact of the frequency on the LOS probability. More studies about the LOS probability modeling in indoor environments can be found in [101], [109]-[111].

The log-distance model has been widely used in the open literature for path loss prediction [112]-[119]. This empirical model assumes a linear relationship between the predicted path loss value and the Tx-Rx separation distance in the logarithmic scale. Therefore, providing the proportional parameter(s) value is based on a satisfactory linear regression analysis of the measurement dataset collected from measurement

campaigns or simulation tools. However, although the linear log-distance model is straightforward and tractable, it does not accurately predict path loss in all radio propagation situations. Therefore, more advanced modeling approaches are necessary to more precisely and flexibly depict path loss in complex and diverse surroundings. Some researchers considered improvements on the log-distance model to include the impact of shadowing and many other factors such as antennas height, operating frequency, clutter, terrain, communication scenario's category (i.e., LOS or NLOS, etc.), and the percentage of the area covered by buildings in built-up areas [113].

The mmWave spectrum has been demonstrated to deliver high throughput in the gigabit per second (Gbps) range that is required for 5G systems. The environment in which the signal propagates, however, has a significant impact on mmWave propagation. Furthermore, the wireless signals at the mmWave frequency spectrum are strongly affected by penetration loss through solid materials [120]-[121]. All these effects made the networks that used the mmWave regime require a precise LOS and NLOS classification because of the significant difference. To deal with this issue, this work proposes a probabilistic path loss prediction model for wireless communications in enclosed indoor corridor scenarios, as presented in chapter 3.

Multi-frequency path loss prediction models are receiving more attention from some researchers in recent times to develop accurate and stable path loss models for future wireless systems over a wide range of frequency regimes [18], [82]. However, these models face the same problems mentioned above as the single-frequency models. There are two main problems in using the previous empirical models: the first problem is that these models are accomplished by a large amount of measurement in a certain environment to obtain a specific model that works for a particular environment at a specific frequency band, which is obviously time-consuming. The second problem is the limited prediction accuracy (the models do not fit the measurement data with a deficient prediction error) provided by these log-distance path loss models in some specific regions. Moreover, the use of the traditional linear models for predicting path loss is not sufficient to capture the path loss behavior accurately in higher frequency bands that are adopted to cope with the emerging demands of new wireless technologies. Accordingly, innovative methods that provide reliable modeling and prediction of the wireless propagation channels are highly needed, especially for complex environments of radio wave propagation that have a severe influence on the quality of wireless communication systems.

Machine learning (ML) is a set of approaches for making predictions based on datasets and modeling algorithms. ML-based methods are now used in various disciplines, including speech recognition, image identification, natural language processing, and computer vision. In many telecommunication fields, the research based on ML of various topics such as propagation loss prediction, channel decoding, signal detection, and channel estimation has already made significant progress [82], [120]-[122]. All ML methods rely on the type of information (input features) that is used for the training. ML methods can be classified as

supervised learning and unsupervised learning. For classification or regression issues, supervised learning is used to learn a function or relationship between inputs and outputs. Unsupervised learning, on the other hand, is the process of extracting hidden rules or connections from unlabeled data. Path loss prediction can be viewed as a supervised regression problem that ML methods can handle [22], [123]-[126]. Path loss prediction models based on machine learning algorithms are promising to overcome the time consumption in traditional linear path loss models that depend mainly on measurement campaigns at new frequency bands in specific outdoor and indoor environments and communication scenarios and/or simulation-based methods, such as ray-tracing techniques [18]. ML-based algorithms have been successfully used to assist to predict path loss in several operating environments [18], [19], [22], [69], [77], [124]-[144]. Furthermore, unlike traditional models, ML-based path loss prediction models can provide reliable generalizations on the propagation environment [130], [144].

Many recent research studies have adopted the methodology of using neural networks for the prediction of path loss based on measurement data in a specific frequency band for a specific environment and communication scenario; they then compare their prediction models with the traditional models in terms of accuracy and prediction error analysis. Various supervised learning approaches, such as the ANN [136], [137], [145], [146], support vector machine (SVM) [125], [147], KNN [126], and RF [126], have been successfully used to construct path loss models. Recently, it was reported that deep learning methods such as the deep neural network (DNN) and ANN provide better prediction results compared to the traditional path loss models [20], [23]. Moreover, in [124], ANN-based path loss prediction models provided better performance than ML-based models, including the RF and SVR models. Further, the ANN-based model was proven to be superior to the log-distance model in the same study. In addition, the authors of [136], [137], [146] offered prediction models using ANN and showed more accuracy than other empirical models in terms of path loss prediction. A vision for developing real-time prediction models for path loss can be found in [146]. In [18], a DNN multi-frequency path loss model was analyzed and compared with the alpha-beta-gamma (ABG) path loss model. The results show that the DNN-based model is far better than the ABG model based on the results of prediction error analysis metrics.

Generally, any neural network is made up of nodes, which are processing components that are tightly linked and organized in layers. They have the capacity to describe any function that is given to them from the raw datasets. Consequently, setting up closed-form equations to map the input features into output target(s) is unnecessary for neural networks, unlike traditional methods. It is essential to use appropriate features as inputs to train and test ML-based models since they make the model more efficient and adaptable while reducing the solution complexity [22]. The input features of the ANN-based path loss model for unmanned aerial vehicles (UAVs) at 1800 and 2100 MHz were distance, clutter height, altitude, longitude,

latitude, and elevation, as reported in [25]. In another work, the features adopted for ML-based path loss prediction in an urban environment included the Tx–Rx separation distance, as well as building information such as height, thickness, and distance away from the antenna [48]. In [148], the input for a successful deep convolutional neural network was 2D satellite images to provide reliable LTE signal quality metrics calculations. In [149], an ANN model with 48 neurons in a single hidden layer and a hyperbolic tangent activation function (also called Tanh activation function) has provided the best performance over several empirical path loss models such as COST-231, HATA, ECC-33, and EIGI. The input features adopted for this single-layered feed-forward neural network were distance, clutter height, altitude, elevation, latitude, and longitude, while the network had one output layer, which represents the path loss. Another work adopted an ANN model with two hidden layers to predict the path loss at the 1800 MHz frequency band for smart campus environments [133]. The main result obtained from that work is that the ANN model outperforms the RF-based model for such environments. Similar results provided the ANN model's quality over the RF model since it can extract relevant input features of communication environments [142].

The author of [128] proposed a multilayer perceptron (MLP) feed-forward neural network model for predicting path loss. The model was based on 11 input features and 2 hidden layers with the use of the Tan-sigmoidal as an activation function. The 11 features selected were the Tx–Rx separation distance, operating frequency, transmitter terrain height, receiver terrain height, transmitting antenna height, average clutter height, %water, %building, %plain, %road, and %trees. At the same time, the path loss was the only output of this ANN model. The results obtained from this research reveal high prediction accuracy in predicting path loss since the degree of correlation values were higher than 0.94 for the model designed. However, the model's complexity is high since it needs comprehensive knowledge to train from many features, limiting the adoption of such models. A comparative analysis of conventional, ML, and DNN-based methods for path loss prediction proved that the latter has the best performance since higher prediction accuracy was achieved than the other methods [136]. Using many layers in deep learning provides extraction of the features from high-dimensional datasets via training; this benefit is often not possible in using traditional models [150]. Moreover, DNN-based models do not rely on a predefined mathematical formula to represent the model as the conventional models do [151]. These DNN-based methods have been applied to many communication environments, such as rural, urban, and suburban areas [18]. In [152], DNN-based models were developed for five environments under the category of urban, dense urban, suburban, dense suburban, and rural for a specific frequency regime and proved its prediction accuracy.

In contrast, in [153], DNN-based models were proposed over a wide range of frequencies from the ultra-high frequency (UHF) band to the SHF band. The previous efforts' datasets for the LOS and NLOS communication scenarios were based on measurement campaigns. Other studies considered the data from

the satellite images, such as in [20], [23], [123]. The Tx–Rx separation distance and the operating frequency have been selected as the only two features for ANN-based path loss prediction models for an urban area at 3.4, 5.3, and 6.4 GHz and for a suburban environment at 450, 1450, and 2300 MHz [77]. The proposed ANN models achieved higher prediction accuracy than the other well-known models, such as the CI model, the Gaussian process model, and the two rays model. Aside from the frequency and distance parameters, wall and floor attenuation are also utilized as input features for the ANN model to predict path loss in a multi-wall environment [78]. However, considering multi-dimensional regression to predict the path loss based on several highly correlated input features such as distance, frequency, antenna height, and other factors is still lacking in the literature. Because the associated inputs are uncertain, several candidate functions are used in the regression. Many of these characteristics, on the other hand, would be superfluous or unnecessary. Furthermore, most input features lack the capacity to discriminate for prediction. Using dimensionality reduction techniques, such as principal component analysis (PCA) or singular value decomposition (SVD), the input data could be transformed into a smaller representation set of features in order to create a decent estimate [77]. Dimensionality reduction is used to minimize the number of features in an input dataset while preserving as much helpful information as feasible.

For the ANN, some studies suggest that a neural network is a deep network with only one hidden layer. This theory supports the path loss prediction problems proposed by these studies [23], [132], [151]. Other efforts stated that having two or three hidden layers for the ANN model is enough to provide an accurate approximation of the non-linear relationship between the input features and the output target of the model [77]. Note that the complexity of the models can be reduced by adopting a small number of hidden layers taking into account the tradeoff between the models' accuracy and complexity. The diversity of these studies in the open literature justifies no specific rule to provide the optimum size (for example, the number of hidden layers and the number of neurons per hidden layer) of the ML-based models. Of course, all of that depends on the training dataset given from measurement campaigns and/or simulation tools, and this dataset depends on the specific environments and scenarios of communication. As a result, the model's hyperparameters, such as learning rate, activation function, and optimizers, are experimentally selected to provide the best model's performance. Some studies introduced hyperparameter-tuning techniques, such as grid search and random search to overcome the time consumption of manually selecting the optimum models' parameters [124], [154].

A path loss prediction model for vehicular-to-vehicular communications using a random forest has been evaluated in terms of accuracy and generalization capability. The random forest method has been successfully used in various aspects due to its simplicity and relatively high prediction accuracy. Path loss prediction using artificial neural networks for cellular networks with wireless channels between base stations

and users is discussed in [155] and [156]. In addition, the authors in [157] propose a multi-kernel-based online path loss prediction model incorporating trajectory information and user location for the downlink. Air-to-Ground path loss modeling in urban environments has been presented in [158] for unmanned aerial vehicles (UAVs) applications. The UAVs recently are of vital interest in different domains due to their ease of feasibility and mobility. The ML models proposed in [158] are based on artificial neural networks (ANN), regression trees (RT), and K-nearest-neighbors (KNN). A similar performance was observed by the three models in three selected frequency bands of 433, 900, and 5800 MHz, with a slight outperformance of the KNN model since lower values of both the root-mean-square error (RMSE) and mean absolute error (MAE) were achieved compared to the other two models.

The authors in [22] developed an RF-based path loss model benefiting from the fact that this method is based on a massive number of regression trees and that the path loss is also a regression problem, as mentioned earlier. The data used to train the model comes from four typical mobile communication terrains. The outcome of this study proved that the proposed RF model performed better than traditional wireless propagation models as well as a DNN model after constructing relevant features. The average root-mean-square error (RMSE) for the RF-based model was 6.106 dB for all types of terrains selected. A combined model for path loss and shadowing based on the ANN multilayer perceptron (ANN-MLP) was developed in [25]. The shadowing impact was analyzed based on the Gaussian process to provide the variance (or standard deviation) from the training dataset. This technique will help calculate the shadowing attributed to the shielding effect of buildings, mountains, and other objects that exist in the communication channel between the Tx and Rx. The results provided show the usefulness of the model in predicting the propagation loss.

A path loss prediction model was proposed for urban environments using the SVM method [135]. The input features selected were the Tx–Rx separation distance, vertical and horizontal antenna attenuations as system-specific parameters, as well as latitude, longitude, and terrain elevation as environment-specific parameters made up of six features. A similar work documented in [159] used the same environment-specific parameters in a deep-learning-based model. Both [135], [159] show that the proposed ML-based models provide higher efficiency than other analytical models.

Two forms of probabilistic path loss predictors for a specific communication environment are reported and investigated in [129]. The first approach utilizes Bayesian learning to get the posterior distribution of an analytical model's parameters and produces a path loss value prediction. A probabilistic neural network is used in the second technique to obtain the parameters of analytical distributions. The authors also studied the effect of changing the amount of data available for training on the ML proposed models' performance. The prediction capacity for the models was measured in terms of the total variation distance (TVD) and Kullback–Leibler (KL). It is to be noted from their results that the mixture density neural network (MDN)

model has more accuracy in describing the path loss than the Bayesian learning model. However, the latter provided better data efficiency than the MDN model. These probabilistic path loss models are beneficial since they overcome the problem of classifying the propagation path between the transmitting and receiving antennas as LOS or NLOS, given the probability of having a clear LOS connection between the Tx and Rx is already considered in the model. Hence, the knowledge of the LOS and NLOS communication scenarios is not required for such models. It is worth noting that the MDN model is basically a combination of a conventional Neural Network (NN) and a mixture model. This model has the ability to provide a distribution of the path loss values instead of point estimates. The neural network adopted for this work has only one input feature; that is, the separation distance between the Tx and Rx, one hidden layer of 64 neurons with the use of the rectified linear unit (ReLU) as an activation function, and an output layer that represents the path loss. It is clear from the previously documented works that there is an excellent opportunity to predict the path loss with the best accuracy and less time consumed by adopting ML and deep learning algorithms.

There is a severe lack of literature on predicting path loss using more advanced ML algorithms. This is because the existing research relied mainly on well-known ML methods and adjusted the models to fit measurement data with minimum errors. In comparing these ML-based and empirical models, ML methods have proved their superiority in terms of accuracy since they can provide complex equations to describe the path loss compared to the empirical models that mainly utilize a few parameters. This dissertation tries to fill the gap mentioned earlier in the literature by proposing new path loss prediction and LOS probability models and improving the performance of the existing standard models. Furthermore, an extensive comparative analysis of major ML-based path loss models will be considered. Finally, the last part of this research proposes a novel ensemble-method-based neural network path loss model for future wireless communications, as detailed in the following chapters.

## 2.7 Chapter Summary

This chapter presented in brief the main wireless propagation mechanisms, large-scale and small-scale fading effects, and wireless channel modeling. Finally, this chapter discussed in detail related works aimed at modeling the wireless communication channel based on measurement campaigns and machine learning algorithms to identify the main research gaps in this research field.

# CHAPTER 3

## **PROPOSALS OF MEASUREMENT-BASED PROPAGATION CHANNEL MODELS FOR WIRELESS COMMUNICATIONS IN ENCLOSED ENVIRONMENTS**

---

## *Chapter 3: Proposals of Measurement-Based Propagation Channel Models for Wireless Communications in Enclosed Environments*

### **3.1 Introduction**

In the last few years, the demand for significantly higher data traffic has rapidly increased due to several evolutions of communication technologies such as indoor wireless applications. These applications, exemplified by 8K ultra-high definition streaming, centimeter-level position location, and wireless cognition, are still in continuous development. Therefore, more and more data speeds will be required in the next few years, especially for meeting the requirements of the upcoming sixth-generation (6G) wireless communication system.

Because of their vast amounts of available bandwidth, the super-high frequency (SHF), millimeter-wave (mmWave), and sub-Terahertz (THz) spectra are expected to play a vital role in next-generation wireless communication systems and beyond due to the massively available bandwidth of many tens of GHz [160]. However, despite their theoretical potential for high data throughput to meet future needs, these frequency bands confront several difficulties, including significant path loss and substantial penetration loss [161]. For example, propagation of the mmWave bands for outdoor to indoor communications causes significant penetration loss values (up to 60 dB) [162]. As a result, it is preferable to segregate indoor mmWave communication networks from co-channel outdoor cellular communications [163].

Modeling the path loss is vital in designing any wireless communication network. Path loss prediction models can be achieved deterministically, stochastically, or empirically via measurement campaigns to collect actual path loss data [164]. It is notable from the literature that the most used path loss modeling is measurement-based due to its high prediction accuracy and reliability [165]-[167].

In this chapter, details about measurement campaigns that were conducted to collect the real measurement path loss data are presented. Then, new measurement-based path loss prediction and LOS probability models for wireless communications in enclosed indoor environments are presented, evaluated, and compared to well-known standard models.

## 3.2 Measurement Setup and Data Collection Method

A detailed description of radio frequency (RF) propagation measurement campaigns conducted in a typical enclosed corridor environment is provided in this section. The corridor exists on the 5<sup>th</sup> floor of the Discipline of Electrical, Electronic, and Computer Engineering, University of KwaZulu-Natal, Howard College Campus, Durban 4001, South Africa.

Before beginning the measurement campaigns, we ensured no other transmissions on the same experimental radio frequency bands existed. Also, the measurement system was carefully calibrated, and the measurements were repeated and averaged to ensure high-quality data collection.

The wireless propagation channel is a corridor environment with dimensions of 30, 1.4, 2.63 *meters* as length, width, and height. Both sides of the corridor are made of bricks and dry concrete with wooden doors to offices on one side and an elevator and a staircase on the other side. It is worth noting that these indoor corridors can be approximated as rectangular air-filled waveguides with dimensions immense compared to signals' wavelength. These environments are crucial and commonly used for many indoor applications. Figure 3.1 shows the adopted indoor corridor environment.



*Figure 3.1: The adopted indoor corridor environment.*

Directional path loss modeling is becoming extremely important for the 5G and beyond wireless communication system design. Hence, in order to analyze double-directional channels, two identical vertically polarized antennas with directional radiation patterns have been adopted for the measurements. The Tx antenna's height was 160 and 230 *centimeters* above the floor level, while the Rx height was 160 *centimeters*, which are the average antenna heights for these indoor environments that are adopted by many researchers [40], [168]-[172]. Also, these antenna heights have been chosen since they are close to the typical user height (approximately 1.6 *meters*) and the height of access points in indoor environments (around 2.3 *meters*). When the Tx antenna height was 230 *centimeters* above the floor level, we down-tilted the Tx antenna to ensure that both antennas were aligned on boresight for all the Tx-Rx measurement places in the LOS communication scenario. Three frequencies in the SHF band were considered in this work: 14, 18, and 22 GHz. Both antennas were pyramidal horn antennas with half-power beamwidth values between 13 and 19.2 *degrees* and a directional gain ranging between 19.5 and 22.1 dBi at the operating frequencies. Throughout the campaigns, the intent was to place the Tx at one end of the corridor and move the Rx away from the Tx, having a Tx-Rx separation distance of 2-24 *meters* with an incremental step of 2 *meters* a time. The reference Tx-Rx distance was 1 *meter*, as recommended by most research experts in this field [173]-[177]. Note that to satisfy the far-field requirements, the distance from the Tx should be much greater than the wavelength of the lowest operational frequency, which already exists since the wavelength of the SHF signals is in the range of *millimeters*.

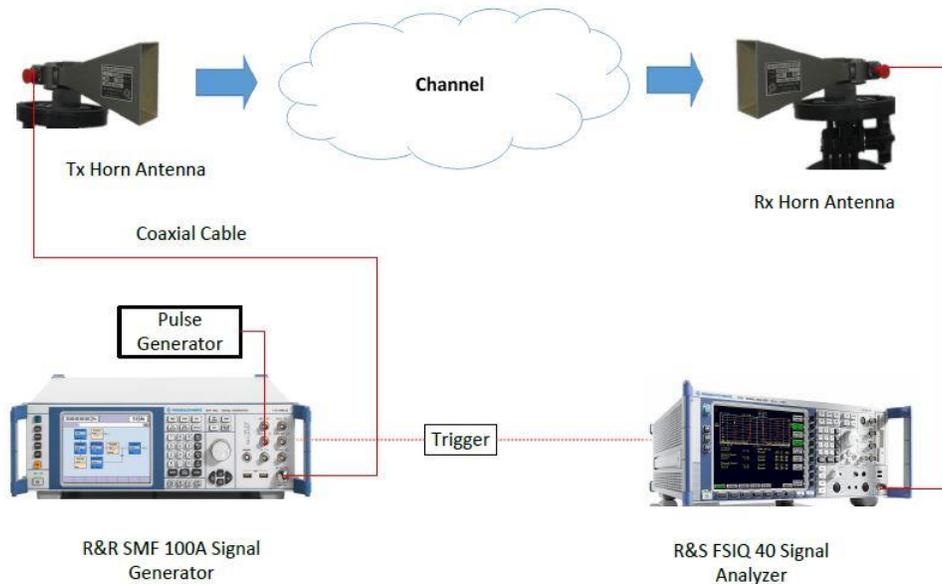
The measurements were performed under the conditions of LOS and NLOS communication scenarios. In the LOS scenario, both antennas were aligned on boresight, and there were no obstacles in the direct propagation path between them. In contrast, the Rx in the NLOS depends mainly on diffractions, reflections, and waveguiding mechanisms in the corridor environment since both antennas had no alignment on boresight. Throughout the measurements, the transmit power was fixed at 10 dBm. However, the received power level range was between -41.33 and -19.05 dBm. The angle of departure (AoD) was set at 0 *degree*, while the angle of arrival (AoA) was in the range of 0-360 *degrees* with an incremental step of 10 *degrees* at a time.

A Rohde and Schwarz SMF 100A signal generator working in continuous-wave (CW) mode was applied to feed the transmitting antenna. The signal generator can transmit signals in the frequency range between 100 KHz and 22 GHz at the Tx side. The measured results are collected by means of a Rohde and Schwarz FSIQ 40 Signal Analyzer. This receiving equipment records 500 data sets of received signal strength (RSS) per each AoA and Tx-Rx separation distance. At the Rx side, the signal analyzer can analyze the received signal in a wide frequency range between 20 Hz and 40 GHz. The maximum analysis bandwidth that can be achieved from this Rx equipment is 120 MHz. The RSS data were averaged to ensure accurate detection

of the CW signals. Both signal generator and signal analyzer were directly connected to the antennas through coaxial cables. Figure 3.2 depicts the channel sounder architecture. Since the path loss is the difference value between the Tx and Rx power, taking into account the antennas' gain and the coaxial cables' loss, the measured path loss in dB ( $PL_m[dB]$ ) is calculated by:

$$PL_m[dB] = P_t - P_r + G_t + G_r - L_{cable}, \quad (3.1)$$

where  $P_t$  is the strength of the transmitted signal in dBm,  $P_r$  is the received signal power at the Rx side in dBm,  $G_t$  and  $G_r$  are the gain of the transmitting and the receiving antennas in dBi, and  $L_{cable}$  is the total coaxial cable loss of the measurement system in dB. Figure. 3.3 represents the 3D floor plan of the indoor corridor environment. The parameters of the measurement setup used in this work are summarized in Table 3.1.



*Figure 3.2: The channel sounder architecture.*

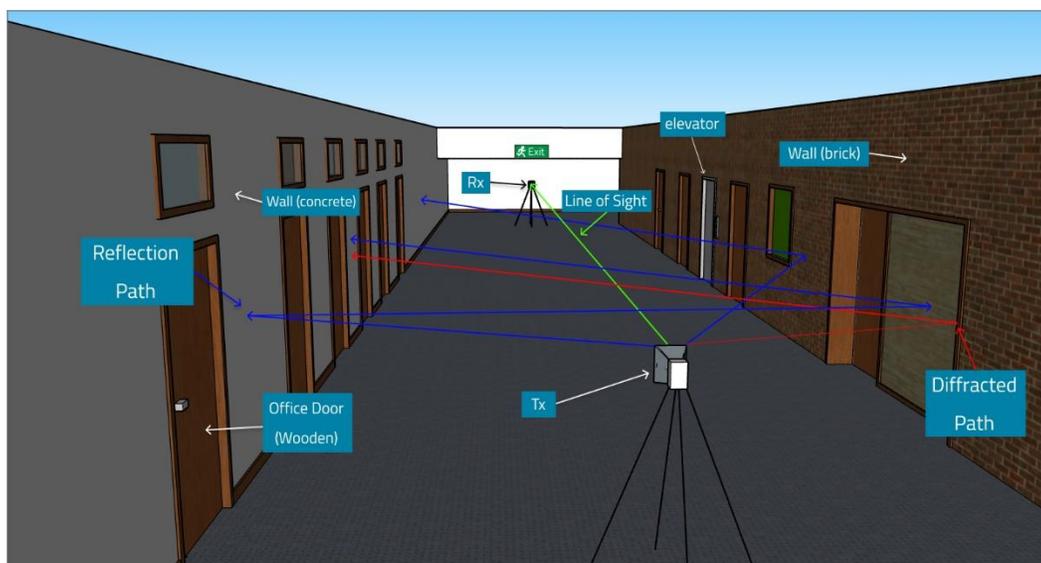


Figure 3.3: 3D floor plan of the indoor corridor environment.

Table 3.1: Channel sounder specifications and parameters configuration.

<i>Parameter</i>	<i>Configuration</i>	<i>Units</i>
<i>Center Frequencies</i>	14, 18, and 22	<i>GHz</i>
<i>Bandwidth</i>	100	<i>MHz</i>
<i>Transmission Signal</i>	Continuous Wave	-
<i>Tx and Rx Antennas</i>	Directional Horn Antennas	-
<i>Transmitted Power</i>	10	<i>dBm</i>
<i>Tx Antenna Height</i>	1.6 and 2.3	<i>m</i>
<i>Rx Antenna Height</i>	1.6	<i>m</i>
<i>Antennas Polarization</i>	Vertical	-
<i>Antennas Gain at 14 GHz</i>	19.5	<i>dBi</i>
<i>Antennas HPBW at 14 GHz</i>	Azim. 18.4°, Elev. 19.2°	Degrees
<i>Antennas Gain at 18 GHz</i>	20.95	<i>dBi</i>
<i>Antennas HPBW at 18 GHz</i>	Azim. 15.4°, Elev. 15.6°	Degrees
<i>Antennas Gain at 22 GHz</i>	22.1	<i>dBi</i>
<i>Antennas HPBW at 22 GHz</i>	Azim. 15°, Elev. 13°	Degrees

### 3.3 Proposal of a Measurement-Based Path Loss Model for Enclosed Indoor Channels

Path loss mainly governs the coverage of any wireless network. Hence, the design of trustworthy wireless systems relies on accurate path loss prediction models. Moreover, it became known that moving up the frequency to the mmWave and the sub-THz increased the data rates to meet near future demands. However, these bands are much more sensitive to the wireless channel and exhibit high path loss values due to their small wavelengths. Consequently, accurate understanding and modeling of the wireless channel are highly needed for systems planning and link-budget calculations.

This section proposes a new path loss prediction model for enclosed indoor channels. The model has a different nature from the standard CI and FI models since it assumes that the path loss (in dB) changes with the square root of the separation distance between the transmitter (Tx) and receiver (Rx) measured in meters. The motivation for proposing this model is the heightened possibility of improving the proposed equation to achieve much higher prediction accuracy and sensitivity to the wireless channel effects since, as mentioned above, moving to higher frequency bands leads to more sensitive propagation signals. The proposed model is derived and compared to the free space path loss model, CI model, and FI model at three selected frequency bands, namely 14 GHz, 18 GHz, and 22 GHz.

The proposed path loss model is measurement-based since its parameters are based on fitting the actual measurement data collected from the measurement campaigns in the selected wireless channel and frequency bands. The proposed path loss prediction model is given by:

$$PL(d)[dB] = A + B\sqrt{d} + X_{\sigma}, \quad (3.2)$$

where  $PL(d)[dB]$  is the predicted path loss value measured in dB as a function of the separation distance between the transmitting and receiving antennas.  $A$  and  $B$  are the main parameters of the proposed model, where  $A$  represents the path loss value just near the Tx antenna, and  $B$  controls the dependency of the path loss on  $\sqrt{d}$ . The parameters  $A$  and  $B$  depend on several factors such as the operating frequency of the propagating signal, the nature of the wireless channel in terms of its dimensions, and propagation mechanisms such as reflections, diffractions, and waveguiding effects.  $X_{\sigma}$  is a Gaussian random variable with zero mean and a standard deviation  $\sigma$  to represent the shadow fading (SF) effects.

The minimum mean square error (MMSE) method is adopted to provide the best-fit values of the parameters  $A$ ,  $B$ , and  $\sigma$ . The idea is to make the proposed model match the actual measured path loss achieving the minimum prediction error. Hence, the standard deviation  $\sigma$  can be written as:

$$\sigma = \sqrt{\frac{\sum X_{\sigma}^2}{N}} = \sqrt{\frac{\sum (PL(d)[dB] - A - B\sqrt{d})^2}{N}}, \quad (3.3)$$

where  $N$  is the number of the Tx-Rx separation distances (i.e., the total number of the average path loss samples recorded). Based on the content of Eq. (3.3), the parameters of the proposed model must satisfy the following:

$$\frac{\partial}{\partial A} (PL(d)[dB] - A - B\sqrt{d})^2 = 0, \quad (3.4)$$

$$\frac{\partial}{\partial B} (PL(d)[dB] - A - B\sqrt{d})^2 = 0. \quad (3.5)$$

Simplifying the previous two equations leads to:

$$NA + \sum \sqrt{d} B = \sum PL(d)[dB], \quad (3.6)$$

$$\sum \sqrt{d} A + \sum d B = \sum (PL(d)[dB] \times \sqrt{d}). \quad (3.7)$$

The parameters  $A$  and  $B$  can be expressed in their closed-form after providing the matrix forms of Eq. (3.6) and Eq. (3.7) as follows:

$$\begin{bmatrix} N & \sum \sqrt{d} \\ \sum \sqrt{d} & \sum d \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum PL(d)[dB] \\ \sum (PL(d)[dB] \times \sqrt{d}) \end{bmatrix}, \quad (3.8)$$

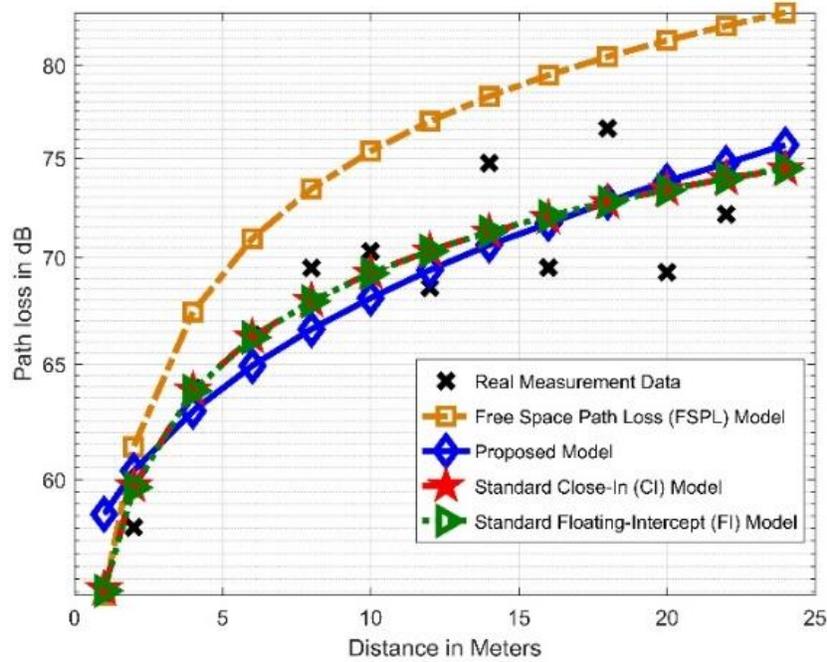
$$A = \frac{(\sum d) \times (\sum PL(d)[dB]) - (\sum \sqrt{d}) \times (\sum (PL(d)[dB] \times \sqrt{d}))}{N \sum d - (\sum \sqrt{d})^2}, \quad (3.9)$$

$$B = \frac{-(\sum \sqrt{d}) \times (\sum PL(d)[dB]) + N(\sum (PL(d)[dB] \times \sqrt{d}))}{N \sum d - (\sum \sqrt{d})^2}. \quad (3.10)$$

Finally, the value of the shadow fading's standard deviation ( $\sigma$ ) can be found from Eq. (3.3) after substituting the parameters' values.

Figure 3.4 represents the averaged actual measured path loss values, the free-space path loss model, our proposed model, and the standard CI and FI models at the 14 GHz frequency band. It is to be noted from the figure that all the models are below the curve of the FSPL model. This is because the abundance of propagation mechanisms in these enclosed indoor environments leads to constructive interference and lower path loss values than the propagation at the same frequency in open environments. Moreover, Figure 3.4 reveals the proposed model's high prediction accuracy and the precision of the standard CI and FI models since all the models' curves are close to the measurement data points.

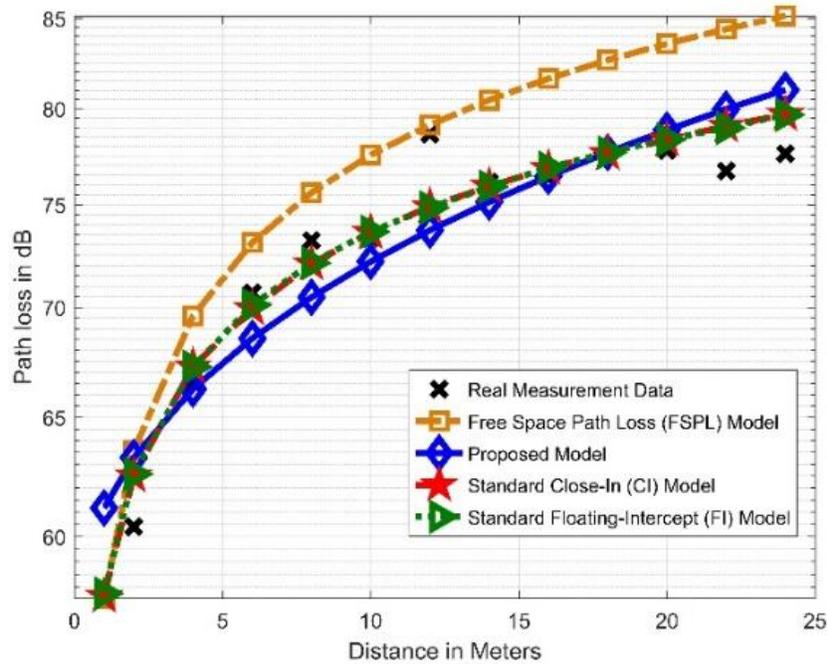
The proposed model's parameters that provide the best-fit prediction 14 GHz are  $A = 54.0353$ ,  $B = 4.3928$ , and  $\sigma = 2.7472$ . The relatively small value of sigma proves the accuracy and sensitivity of our proposed model to the propagation mechanism inside these enclosed corridor environments. Moreover, we analyzed the prediction error of the proposed model based on the real measurement data. The mean prediction error (MPE) value found is only 2.5034 dB. The standard deviation error (SDE) value is 1.1289 dB. All these values are at 14 GHz.



*Figure 3.4: The proposed and standard path loss prediction models at 14 GHz frequency band.*

The measurement data and models at 18 GHz are depicted in Figure 3.5. Again, high prediction accuracy is clear from the figure since the models follow the measurement data. Also, it is worth mentioning that the data and models' curves are closer to the FSPL curve than the case of 14 GHz since 18 GHz frequency band has higher PLE values. For our proposed model, the model's parameter at 18 GHz frequency band that provide the highest possible accuracy are  $A = 55.9049$ ,  $B = 5.0977$ , and  $\sigma = 2.5287$ . It is clear that the model's stability is high since the parameters' values have slight changes when we increase the frequency from 14 GHz to 18 GHz. Numerically, the change of  $A$ ,  $B$ , and  $\sigma$  is 3.46%, 16%, and 7.95%. The value of the shadow fading's standard deviation at 18 GHz is lower than the value at 14 GHz. This means that more precision can be achieved when we jump for higher frequency bands, which is the trend of using the

mmWave and further high frequency regime to meet the speed requirements of future communication systems.



*Figure 3.5: The proposed and standard path loss prediction models at 18 GHz frequency band.*

The MPE and SDE of the proposed model at 18 GHz are 2.0470 and 1.4911, respectively. The average MPE here is lower than the value at 14 GHz. However, the SDE is higher; which means that the signal fluctuations are more at 18 GHz. Nevertheless, the prediction error values are still lower than 14 GHz.

Figure 3.6 shows the measurement data and the models at 22 GHz. More accuracy is achieved since lower values of  $\sigma$ , MPE, and SDE are found. Table 3.2 represents the numerical values of the proposed model's parameters and the evaluation metrics' values (MPE and SDE). It is to be summarized from the results that the proposed model provides the best performance at 22 GHz frequency band compared to the results found at 14 GHz and 18 GHz. Figure 3.7 depicts the proposed model and measurement data at the three frequency bands adopted for this study. In addition, we plotted the prediction error of the proposed model as a function of the separation distance between the transmitting and receiving antennas at 14, 18, and 22 GHz frequency bands. Figure 3.8 justifies the results discussed in this section since error values are around 0.5 to 3 dB. The performance can be significantly improved by applying an efficient approach, as presented in chapter 4. However, probabilistic path loss modeling will be discussed in the following sections.

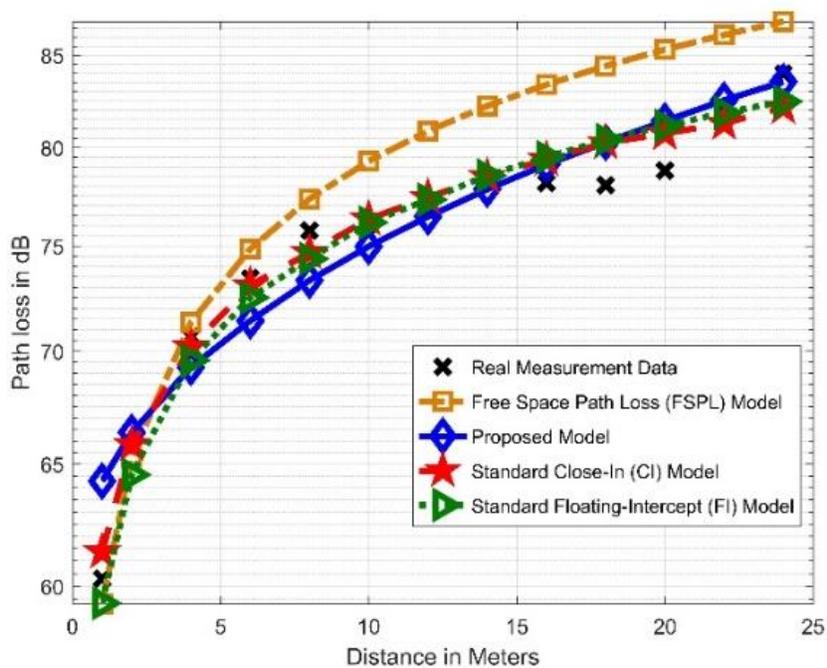


Figure 3.6: The proposed and standard path loss prediction models at 22 GHz frequency band.

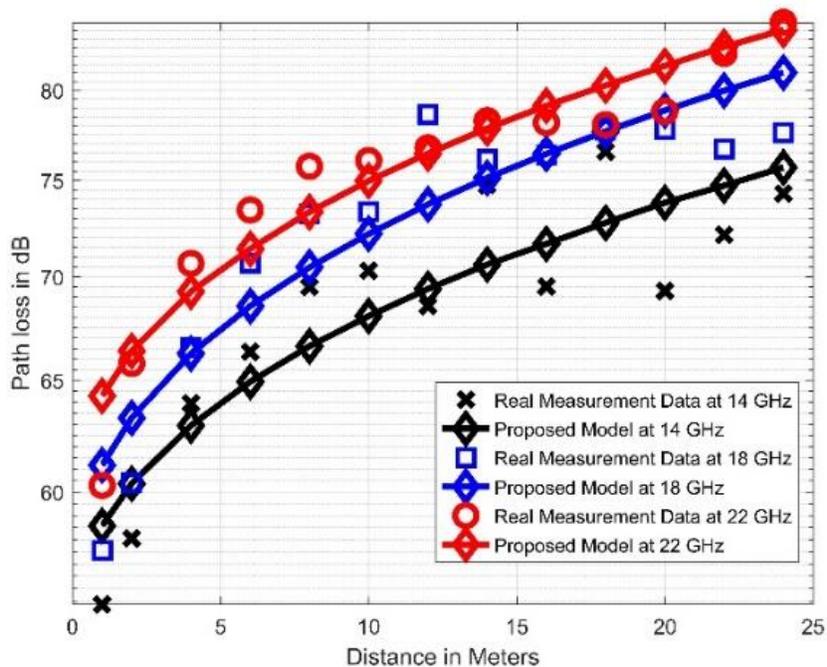


Figure 3.7: The proposed path loss model with the measurement data at 14, 18, and 22 GHz.

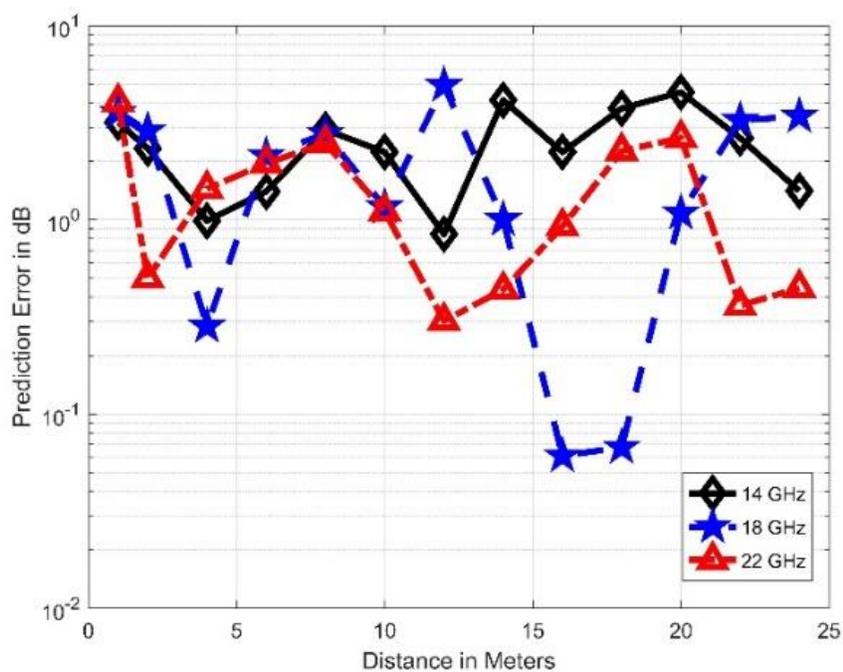


Figure 3.8: Proposed model's prediction error vs. Tx-Rx separation distance.

Table 3.2: Numerical values of the proposed mode's parameters and the evaluation metrics' values.

<i>Parameter</i>	<i>Value</i>		
	<i>14 GHz</i>	<i>18 GHz</i>	<i>22 GHz</i>
<i>A [dB]</i>	54.0353	55.9049	59.1848
<i>B</i>	4.3927	5.0977	4.9282
<i>σ [dB]</i>	2.7472	2.5286	1.8070
<i>MPE [dB]</i>	2.5034	2.0470	1.4517
<i>SDE [dB]</i>	1.1289	1.4911	1.1050

### 3.4 Evaluation of Standard Line-of-Sight Probability Models

Simulating the propagation of the wireless signals cost-effectively and reliably requires accurate channel models to design and compare radio air interfaces and to deploy the wireless systems precisely, taking into account all the possible effects and parameters such as the 3D separation distance between the Tx and Rx, bandwidth, carrier frequency, and environmental impacts [94].

Many efforts aimed to understand the channel behavior and the wave propagation's mechanisms at frequencies above 6 GHz in the SHF and mmWave frequency regimes have recently been published [170]-[177]. The mobile industry has discovered that defining the wireless signals' path loss separately for the line-of-sight (LOS) and non-line-of-sight (NLOS) situations is beneficial. As a result, mathematical models for predicting the LOS probability (the probability of having the Rx within clear LOS paths of the Tx) are highly needed. Given the higher diffraction loss at higher frequencies relative to sub-6 GHz bands where diffraction is a prevalent propagation mechanism, LOS transmission would provide more stable results for outdoor and indoor wireless communications in the SHF and mmWave frequency bands.

To the best of our knowledge, there is a research gap in understanding and modeling the LOS probability models for indoor corridor environments in the SHF band. Hence, this section tries to fill the gap by providing LOS probability models based on real measured data collected in the typical indoor corridor environment discussed above. Also, we propose a new LOS probability model that performs better than the standard ITU and WINNER II models.

It is known that LOS transmissions are more likely to occur when the Tx-Rx separation distance is small. As a result, the LOS probability is considered a monotonically decreasing transmission range function [94]. The LOS probability has a dependency on the geometry and layout of the environments and communications scenarios. Some studies have adopted a map-based approach in determining the LOS probability based on the Tx and Rx positions. [108]. From the literature, some researchers considered the LOS probability as a frequency-independent function [94], [106], [108]; on the contrary, others proved that it is frequency-dependent [106], [107]. This work focuses on ITU-R and WINNER II (A1) models and proposes a new model that accurately fits measurement data for indoor corridor environments. Table 3.3 presents the ITU-R, WINNER II (A1), and our proposed model.

Table 3.3: ITU-R, WINNER II (A1), and the proposed LOS probability models.

Model	Definition
ITU-R [94]	$P_{LOS} = \begin{cases} 1, & d \leq d_A \\ \exp\left(\frac{-d(d-d_A)}{a}\right), & d_A < d < d_B \\ b, & d \geq d_B \end{cases}$
WINNER II (A1) [94]	$P_{LOS} = \begin{cases} 1, & d \leq d_C \\ 1 - k\left(1 - (d - z \log_{10}(d))^3\right)^{\frac{1}{3}}, & d > d_C \end{cases}$
Proposed	$P_{LOS} = \begin{cases} 1, & d \leq d_D \\ 1 - \alpha\left(1 - (d - \beta \log_{10}(d))^3\right)^{\frac{1}{3}}, & d_D < d < d_E \\ \gamma \cdot \exp\left(\frac{-d(d-d_E)}{\delta}\right), & d \geq d_E \end{cases}$

For the three models, the distance  $d_A$ ,  $d_C$ ,  $d_D$  are the distances up to which the LOS condition is guaranteed (the LOS probability is 100%). ITU-R model has a constant LOS probability when the Tx-Rx distance exceeds a specific value, namely  $d_B$ . The decay of the ITU-R probability model is described as an exponential function with a decay parameter  $a$  as shown in Table 3.3. WINNER II (A1) model takes a different shape, and its decay is controlled by two factors ( $k$  and  $z$ ) as depicted. This model starts its unstopped decay after a certain distance  $d_C$ . For our model, there are two critical distances that the model changes its behavior after them, which are the reference distance (1 m away from the Tx), and the breakpoint of the environment (about 12 m for the indoor corridor studied). This model has a similar behavior as the WINNER II (A1) model before the breakpoint; however, the decay parameters are not exactly the same. After the breakpoint, the proposed model behaves as a scaled ITU-R model. Our proposed model has two decay factors ( $\alpha$  and  $\beta$ ) before the breakpoint and two other parameters ( $\gamma$  and  $\delta$ ) after it, as presented in Table 3.4.

Figure 3.9 depicts a comparison between the concrete indoor data, ITU-R model, WINNER II (A1) model, and the proposed LOS probability model at the 14 GHz frequency band. It is to be noted from the figure that all models accurately fit the real data with higher precision provided by our proposed model. Moreover, it is observed that the LOS probability is within a reasonable range for the environment studied (i.e., the indoor corridor) since the minimum probability is around 72% at the maximum Tx-Rx separation distance because of the richness of reflections and diffractions in this environment. Note that the LOS probability decay is high in the first 3 meters away from the Tx. However, it decreases slowly with the distance up to the breakpoint and more slowly after the breakpoint to the end of the corridor. The RMSE

values between the models and the actual measured data are 0.0178 for the ITU-R model, 0.0155 for the WINNER II (A1) model, and 0.0145 for our proposed model. Here, our models' parameters have lower RMSE compared with the existing models for such environments. The proposed model has also reduced the RMSE by 18.5% and 6.5% compared to the ITU-R and WINNER II (A1) models, respectively. Table 3.4 shows the three models at 14 GHz.

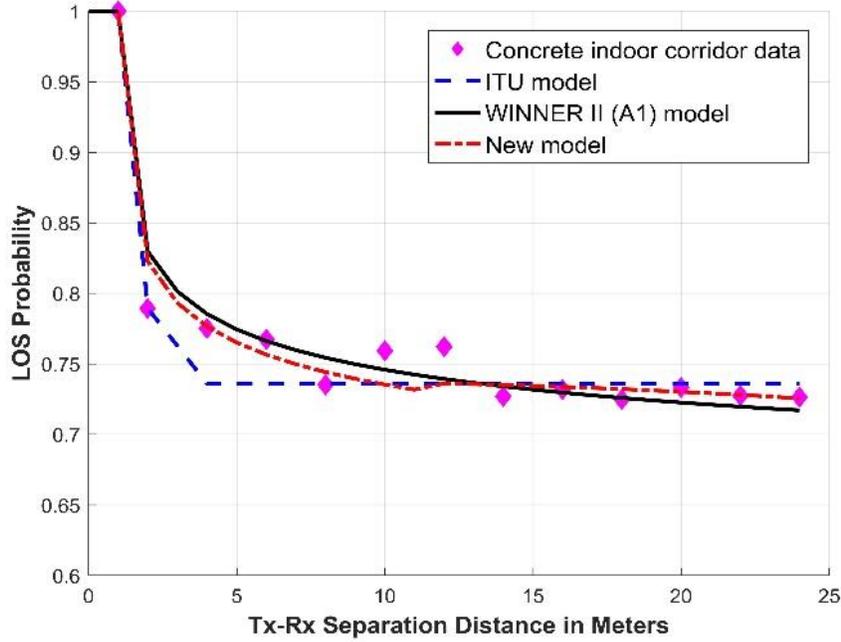
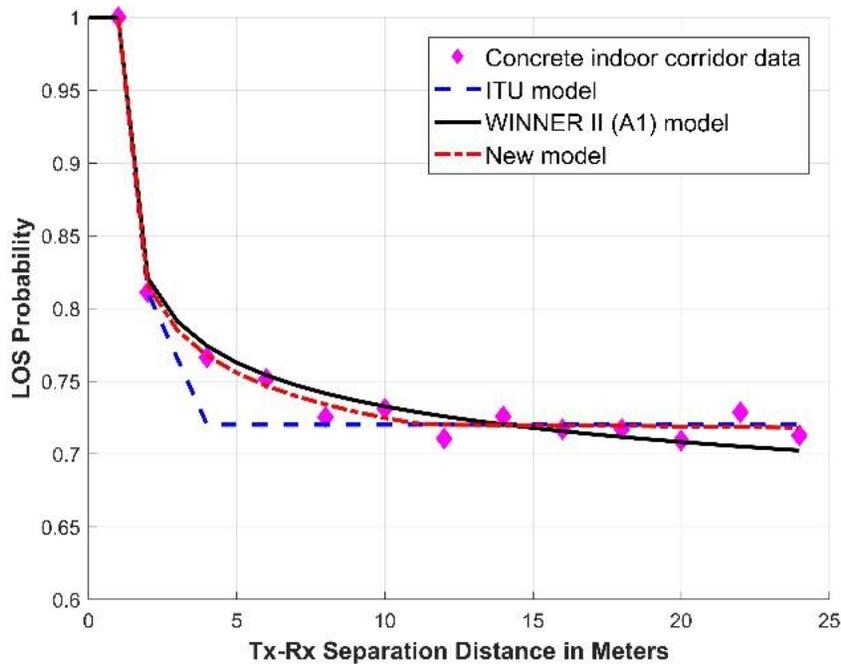


Figure 3.9: LOS probability models at 14 GHz.

Table 3.4: The LOS probability models with their RMSE values at 14 GHz.

Model	Results	RMSE
ITU-R	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ \exp\left(\frac{-d(d - d_A)}{8.5}\right), & 1 < d < 3 \\ 0.736, & d \geq 3 \end{cases}$	0.0178
WINNER II (A1)	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ 1 - 1.45 \left(1 - (d - 0.0018 \log_{10}(d))^3\right)^{\frac{1}{3}}, & d > 1 \end{cases}$	0.0155
Proposed	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ 1 - 1.47 \left(1 - (d - 0.0018 \log_{10}(d))^3\right)^{\frac{1}{3}}, & 1 < d < 12 \\ 0.72 \times \exp\left(\frac{-d(d - d_E)}{2 \times 10^4}\right), & d \geq 12 \end{cases}$	0.0145

Figure 3.10 represents the three models together with the measurement data at the 18 GHz frequency band. Again, It can be revealed that all figures fit precisely the real measured data. Note that from the figure, the LOS probability is guaranteed only when the Rx is 2 meters away from the Tx as a maximum distance for 100% LOS probability. The RMSE values at this band are smaller than the values at the 14 GHz frequency band, which means the models perform better at 18 GHz. For our proposed model, the RMSE value is 0.0063, which leads to better precision than the same model at 14 GHz by 56.6%. As a comparison between our model and the other models studied, the reduction of the RMSE is 62.5% and 38.8% from the ITU-R and WINNER II (A1) models, respectively. As a result, our model has an attractive accuracy at 18 GHz since it provides the best performance compared to the other standard models. The RMSE values and the three models' parameters at the 18 GHz frequency band can be seen in Table 3.5.



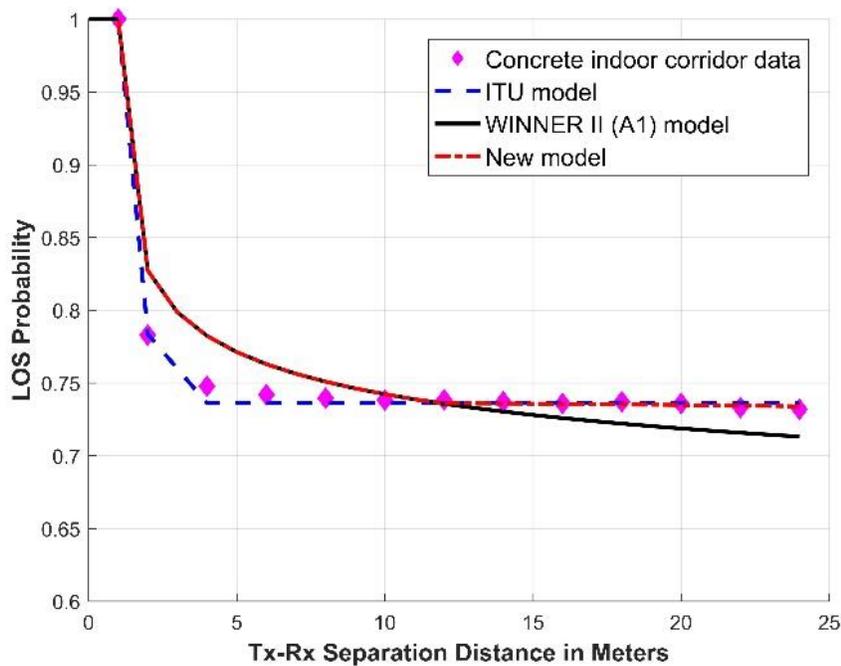
*Figure 3.10: LOS probability models at 18 GHz.*

The three models' LOS probability curves and the measurement data at the 22 GHz band are given in Figure 3.11. It can be seen that the ITU-R model has the best performance here as the RMSE value equals 0.0131, while it is 0.0197 for the WINNER II (A1) and 0.0169 for our model. However, all these values are within an acceptable range of accuracy. It is observed that the LOS probability models' accuracy depends on the operating frequency, and this dependency should be investigated to have frequency-dependent LOS probability models. Table 3.6 depicts the three models at the 22 GHz frequency band. As a summary of our results and discussions, our proposed model has the best performance in terms of its

accuracy in fitting the real measured data and its parameters' stability with the frequency. This model can be used for 4G and 5G systems in enclosed indoor environments with reliable performance. Figure 3.12 provides a graphical representation of the RMSE between the models and the measurement data at 14, 18, and 22 GHz frequency bands.

*Table 3.5: The LOS probability models with their RMSE values at 18 GHz.*

Model	Results	RMSE
ITU-R	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ \exp\left(\frac{-d(d-d_A)}{9.55}\right), & 1 < d < 3 \\ 0.72, & d \geq 3 \end{cases}$	0.0168
WINNER II (A1)	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ 1 - 1.525 \left(1 - (d - 0.0018 \log_{10}(d))^3\right)^{\frac{1}{3}}, & d > 1 \end{cases}$	0.0103
Proposed	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ 1 - 1.57 \left(1 - (d - 0.0018 \log_{10}(d))^3\right)^{\frac{1}{3}}, & 1 < d < 12 \\ 0.72 \times \exp\left(\frac{-d(d-d_E)}{10^5}\right), & d \geq 12 \end{cases}$	0.0063



*Figure 3.11: LOS probability models at 22 GHz.*

Table 3.6: The LOS probability models with their RMSE values at 22 GHz.

Model	Results	RMSE
ITU-R	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ \exp\left(\frac{-d(d-d_A)}{8.2}\right), & 1 < d < 3 \\ 0.736, & d \geq 3 \end{cases}$	0.0131
WINNER II (A1)	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ 1 - 1.532 \left(1 - (d - 0.0018 \log_{10}(d))^3\right)^{\frac{1}{3}}, & d > 1 \end{cases}$	0.0197
Proposed	$P_{LOS} = \begin{cases} 1, & d \leq 1 \\ 1 - 1.47 \left(1 - (d - 0.0018 \log_{10}(d))^3\right)^{\frac{1}{3}}, & 1 < d < 12 \\ 0.736 \times \exp\left(\frac{-d(d-d_E)}{10^5}\right), & d \geq 12 \end{cases}$	0.0169

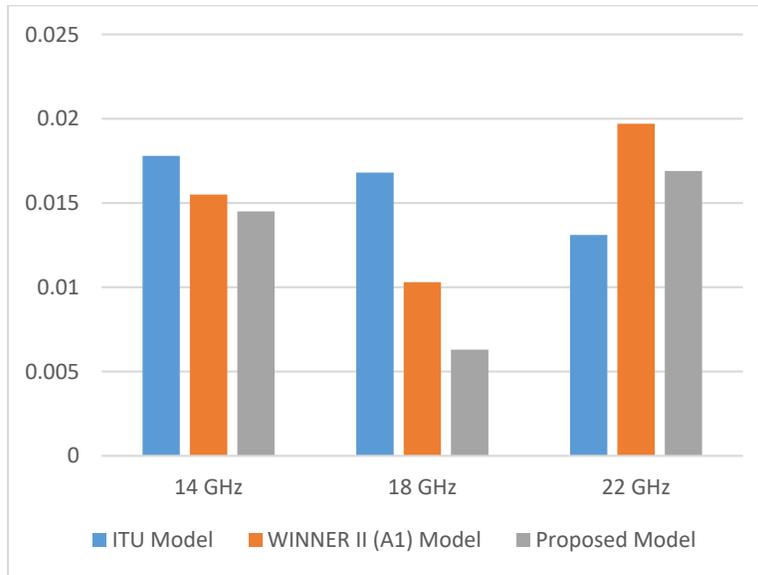


Figure 3.12: The RMSE values of the LOS probability models.

Motivated by the performance of the proposed LOS probability model, the following section provides a probabilistic path loss prediction model based on a path loss detailed in chapter 4 and the proposed LOS probability model.

### 3.5 Proposal of a Probabilistic Path Loss Model

This section proposes a probabilistic path loss prediction model for wireless communications in indoor corridor environments. The model is developed based on two evaluation studies. The first study was about

the best-fit line-of-sight (LOS) probability model. At the same time, the second one was based on the best order of log-distance dependency for the close-in (CI) free space reference distance path loss model. The evaluation of the proposed model's performance was done using the prediction error (PE), mean prediction error (MPE), and standard deviation error (SDE) between the predicted model and the measurement data. In addition, the distributions of the PE with the separation distance between the transmitting and receiving antennas are also presented.

The main benefit of using this model is the freedom from the classification problem mentioned above since only one equation can accurately predict the path loss for both the LOS and NLOS communication conditions. The path loss prediction model adopted in our proposed probabilistic model is an improved version of the standard CI model. The proposed probabilistic path loss model employs a weighting function to consider the proposed LOS probability given in previous section. Thus, the model can be written as:

$$PL_{Prob.}(d)[dB] = P_{LOS}(d) \times PL_{LOS}(d) + P_{NLOS}(d) \times PL_{NLOS}(d), \quad (3.11)$$

where  $P_{LOS}(d)$  is the LOS probability model proposed above.  $P_{NLOS}(d)$  is the NLOS probability, which is  $1 - P_{LOS}(d)$ .  $PL_{LOS}(d)$  and  $PL_{NLOS}(d)$  are given from our improved model based on substituting the parameters' values for the LOS and NLOS, respectively.

Figure 3.13 depicts the average measurement path loss data, free-space path loss model, and the proposed probabilistic path loss model as a function of the Tx-Rx distance in the log scale at 14, 18, and 22 GHz. It is clear from the figure that the measurement data and the probabilistic model curves match for all the frequency bands, which means this model achieves high prediction accuracy. Moreover, for each frequency band, the model curve is lower than the free-space path loss curve due to the fact that these corridor environments contain an abundance of propagation mechanisms such as reflections, diffractions, and waveguiding effects with low delay spread which result in constructive interference that lead to lower path loss values (the free-space path loss exponent is significantly less than 2 in our selected environment). It is also to be noted from the figure that the probabilistic CI path loss curves are not straight lines as the free-space path loss curves. The reason behind that is the consideration of the LOS probability in the model as a function of the Tx-Rx separation distance.

The PE of our proposed model as a function of the Tx-Rx separation distance is shown in Figure 3.14 for the three frequency bands. It is evident that the PE follows a random distribution with the distance between the Tx and Rx, and the values are within the range of 0.1 and 5 dB. This means that the maximum possible PE that can happen is only 5 dB at relatively far distances. This value can easily be overcome by considering this value as a margin in the design of wireless systems and link budget calculations for LOS communication scenarios. Table 3.7 provides the values of the MPE and SDE of the model at 14, 18, and

22 GHz. These values were taken based on comparing the model with the LOS path loss data. The values are within an excellent range since the maximum MPE and SDE are 2.3671 and 1.5985 dB, respectively. As a comparison between the three frequency bands, the best model performance is achieved at 18 GHz.

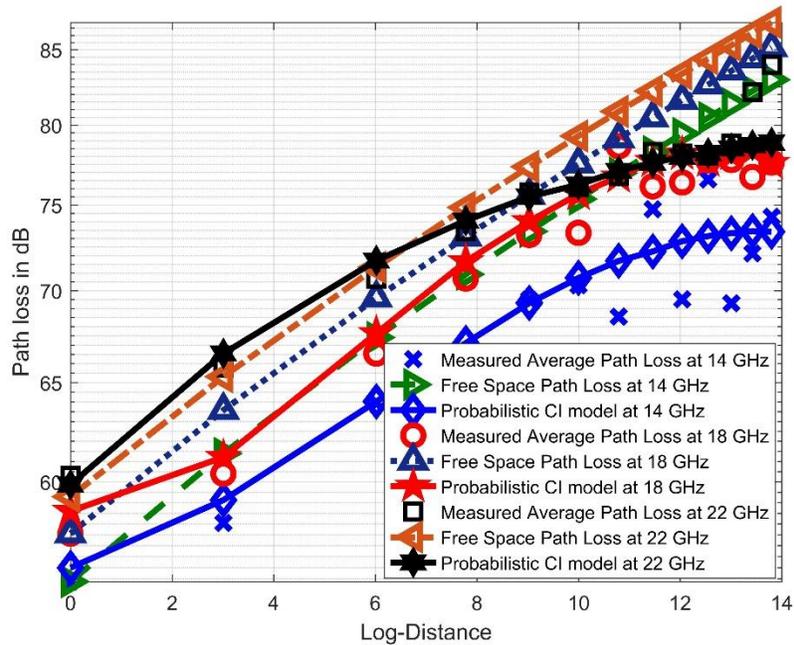


Figure 3.13: Path loss data and models at 14, 18, and 22 GHz.

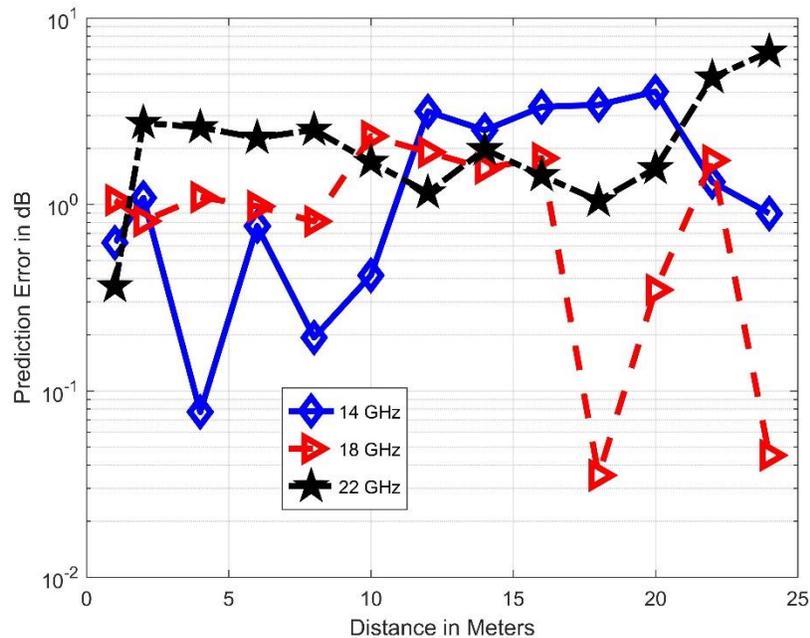


Figure 3.14: PE of the probabilistic CI path loss model.

*Table 3.7: The probabilistic model's MPE and SDE for the LOS Communication Scenario*

<i>Parameter</i>	<i>MPE [dB]</i>			<i>SDE [dB]</i>		
	<b>14 GHz</b>	<b>18 GHz</b>	<b>22 GHz</b>	<b>14 GHz</b>	<b>18 GHz</b>	<b>22 GHz</b>
<i>Probabilistic CI model</i>	1.6776	1.1510	2.3671	1.3319	0.7075	1.5985

To evaluate the effectiveness of the model's performance with the AoA in the NLOS communication scenario, we calculated the MDE and SDE of the model along with the 360 *degrees* azimuth plan with incremental steps of 30 *degrees*. Tables 3.8, 3.9, and 3.10 present the MDE and SDE with the AoA at 14, 18, and 22 GHz frequency bands, respectively. It is clear from the tables that the MDE values are generally increasing with the frequency, and the lowest relative values are when the AoA is equaled to 30 and 330 *degrees*, which are the closest values to the LOS value of 0 *degree*. Also, it is clear from the tables that the highest MPE and SDE values are when the AoA is in the range of 120 to 240 *degrees* since the receiving antenna relies mainly on the back lobes and reflections of the back and sidewalls of the corridor. However, the change from the minimum to the maximum MDE and SDE values was not significant, as can be seen from Figure 3.15 that provides the MDE and SDE values with the AoA. The MPE and SDE values are much higher for the NLOS condition than the LOS. This is due to the high probability of having a clear LOS path between the Tx and Rx in such indoor channels. For example, as discussed above, we found that the minimum LOS probability in the corridor understudy is around 70% for all the frequency bands.

*Table 3.8: The NLOS results of the probabilistic model at 14 GHz frequency band.*

	<b>30°</b>	<b>60°</b>	<b>90°</b>	<b>120°</b>	<b>150°</b>	<b>180°</b>	<b>210°</b>	<b>240°</b>	<b>270°</b>	<b>300°</b>	<b>330°</b>
<i>MPE [dB]</i>	6.8071	8.1311	8.1723	8.1414	8.2531	7.8546	8.3426	8.2019	8.1621	7.9255	6.5116
<i>SDE [dB]</i>	2.8798	4.0837	4.1332	4.1054	4.2955	4.3650	4.5215	4.2431	4.1777	3.7362	2.8732

*Table 3.9: The NLOS results of the probabilistic model at 18 GHz frequency band.*

	<b>30°</b>	<b>60°</b>	<b>90°</b>	<b>120°</b>	<b>150°</b>	<b>180°</b>	<b>210°</b>	<b>240°</b>	<b>270°</b>	<b>300°</b>	<b>330°</b>
<i>MPE [dB]</i>	6.1461	8.2669	9.1996	9.3129	9.2481	8.6972	9.4272	9.2994	9.2246	9.3876	8.5901
<i>SDE [dB]</i>	2.7961	4.1169	4.8005	5.0806	5.0437	5.1698	5.3666	5.0707	4.8355	5.4676	4.3985

Table 3.10: The NLOS results of the probabilistic model at 22 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
<i>MPE [dB]</i>	9.1977	9.8642	10.1601	10.1891	10.2001	10.0111	10.1851	10.2351	10.2151	9.8580	9.3322
<i>SDE [dB]</i>	3.5349	4.0208	4.4670	4.5241	4.6031	4.6475	4.5709	4.6127	4.5725	4.1700	3.5899

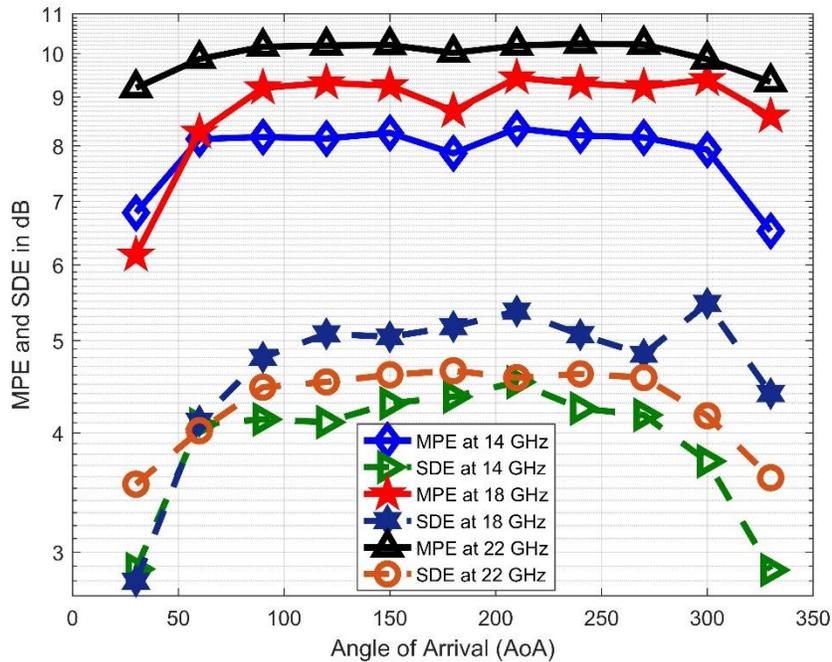


Figure 3.15: The MPE and SDE of the probabilistic model with AoA.

### 3.6 Chapter Summary

Motivated by the high demand for accurate wireless channel models, this chapter proposed new path loss prediction and LOS probability models for wireless communications in enclosed indoor environments such as corridors. The models were derived to get their parameters based on actual measurement data collected using the MMSE approach. Three frequency bands were considered for this research, namely 14 GHz, 18 GHz, and 22 GHz. The models have proved their prediction accuracy since they follow the real data for all the frequency bands selected with low SF's standard deviation values. Moreover, the proposed probabilistic path loss model was also evaluated by the concept of prediction error. It was noted that the model's MPE and SDE values are within a reasonable range, proving the proposed model's precision. Finally, these proposed models can be improved to be more sensitive to the propagation characteristics of the wireless channel, which will be the scope of the following chapter.

# CHAPTER 4

**AN EFFICIENT APPROACH OF IMPROVING  
PATH LOSS MODELS FOR FUTURE MOBILE  
NETWORKS**

---

## *Chapter 4: An Efficient Approach of Improving Path Loss Models for Future Mobile Networks*

### 4.1 Introduction

It is the norm to characterize and model the wireless communication channel in the frequency bands used (or expected to be used) to understand better and accurately deploy the upcoming systems. Therefore, many researchers focused on this area by modeling the wireless channel's behavior using different ways: models based on theories, techniques such as ray-tracing, and measurement campaigns. The latter looks promising because of its accuracy and reliability since the measurements are in real environments and communication scenarios.

After proposing our models in the previous chapter and considering the literature, a vital question has motivated the research in this chapter: *how can we improve the accuracy and reduce the standard deviation of the shadow fading of the well-known standard path loss models without adding parameters that depend on something else like antennas' height or the XPD?*

This question is answered through a fundamental principle: *any linear equation is a polynomial equation with zero coefficients in higher orders*. It is well-known that the standard path loss models (such as the CI and FI models) are a linear equation of the path loss as a function of the logarithmic scale of the separation distance between the transmitting and the receiving antennas. We incorporate an additional parameter to make these models a function of the transmitter-receiver (Tx-Rx) separation distance's squared logarithm. This adopted improvement is simple (in the improved models' equations, the proposed additional parameter that improves the standard models does not depend on anything like frequency and antenna height, etc.) and provides more precision in predicting the path loss, as will be proved in the following sections.

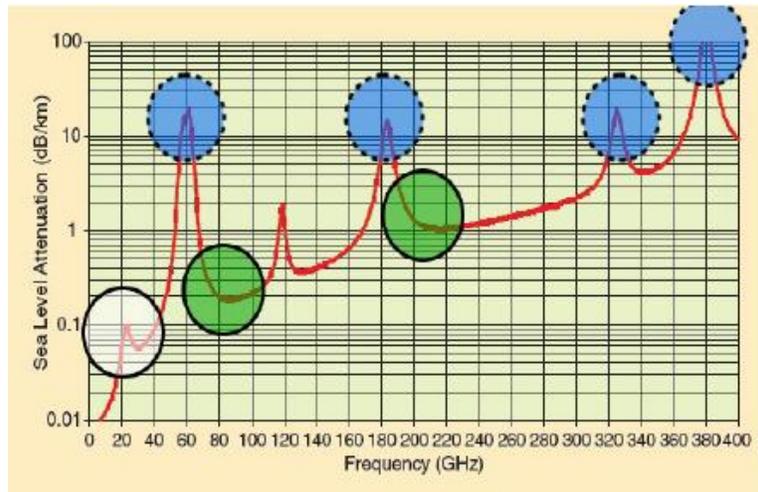
### 4.2 The Standard Path Loss Prediction Modeling

Generally, all standard path loss prediction models can be derived from Friis's equation [65], [66]:

$$FSPL(f, d) = \left( \frac{4\pi df}{c} \right)^2, \quad (4.1)$$

where  $d$  is the Tx-Rx separation distance (in *meters*),  $f$  is the frequency of the propagated signal (in Hz), and  $c$  is the speed of light in the free space, which approximately equals to  $3 \times 10^8$  m/s. This simple equation in the linear scale (absolute numbers) shows that the path loss between two isotropic antennas

aligned on boresight toward each other is mainly a function of the operating frequency and the Tx-Rx separation distance. As presented in Eq. (4.1), the path loss is distance- and frequency-dependent; the increase in the frequency or the distance will produce higher path loss values. However, this is true when the wireless channel is in free space. In reality, the wireless channel's problems such as attenuation, interference, distortion, and noise in the communication schemes have random behavior. For example, Figure 4.1 represents the attenuation in the air at different frequency bands in GHz. It is clear from the figure that the attenuation of the air does not follow a specific behavior. Because of air attenuation and atmospheric absorption, some low frequencies have large path loss values in outdoor environments. This issue and many other problems have accelerated the research to cover all frequency bands by conducting measurement campaigns in typical indoor and outdoor environments to have reliable channel models.



**Figure 4.1: The air's attenuation at different frequency bands [46]. The white circle shows a minor attenuation of the low 5G frequency bands. Attenuation levels that are similar to the 4G cmWave bands are displayed in the green circles. The blue circles show high attenuation peaks; such bands are thus ideal for indoor communications with a minimal range.**

It is more convenient to present the path loss equations in the logarithmic scale. Hence, Eq. (4.1) can be written as follows:

$$FSPL(f, d)[dB] = 32.4 + 20 \log_{10}(f) + 20 \log_{10}(d). \quad (4.2)$$

The value 32.4 comes from  $10 \log_{10} \left( \frac{4\pi \times 10^9}{c} \right)^2$ , and the value  $10^9$  is to have values of frequencies directly in GHz (i.e.,  $f$  in Eq. (4.2) is the operating frequency in GHz). The variable  $d$  in the previous equation represents the Tx-Rx separation distance in *meters*.

The existing path loss models can be categorized as single-frequency (like the CI and FI models) or multi-frequency (like the ABG model). For single-frequency models, the term  $20 \log_{10}(f)$  is constant and

can be added to the first term of Eq. (4.2). The result is a constant term that depends on the value of the frequency (single-frequency path loss models have a pure dependence on the frequency presented as a parameter). The representation of this parameter differs from one model to another. Let us name this parameter  $k_1$ . The term  $20 \log_{10}(d)$  is basically 2 times the distance in the logarithmic scale, where the value 2 is the free space path loss exponent (FSPL), which indicates that the path loss changes with the square of the Tx-Rx distance in the free space. However, this value will change significantly, depending mainly on the medium's nature between the Tx and Rx. In general, it will be easier to denote it as  $k_2$ . Hence, Eq. (4.2) can be simplified as follows:

$$PL(d)[dB] = k_1 + k_2 \times 10 \log_{10}(d), \quad (4.3)$$

where the coefficient  $k_1$  is measured in dB, and  $k_2$  is *unitless*. It is clear from the previous equation that the power of the propagated signal decreases by  $\frac{1}{d^{k_2}}$ . This means that higher values of  $k_2$  will lead to a stronger dependency of the path loss on the separation distance between the Tx and Rx. Different path loss models exist depending on the techniques used to evaluate the values of these coefficients. This research adopted from those several models in the literature two well-known path loss models, the CI and FI models, and our improvement on these models. The parameters of these semi-deterministic models are based on the real measured values of the received signal levels collected from measurement campaigns.

### 4.3 The Standard CI Model and Our Improved CI Path Loss Prediction Model

The CI model can be written from Eq. (4.3) by replacing  $k_1$  by the value of the free space path loss at the operating frequency ( $f$ ) and the reference distance ( $d_0$ ), and replacing  $k_2$  by the PLE ( $n$ ) as described in the following equation:

$$PL^{CI}(d)[dB] = FSPL(f, d_0) + 10n \log_{10}(d) + X_{\sigma}^{CI}, \quad (4.4)$$

where  $X_{\sigma}^{CI}$  is a Gaussian random variable with zero mean and a standard deviation  $\sigma$  in dB [6]. This term represents the shadow fading (SF), representing the large-scale fluctuations of the path loss values because of obstructions and other random propagation effects [177]. Having fewer values of the SF standard deviation means that the path loss models are more accurate. The SF's importance for researchers and engineers lies in the fact that it can establish standards that include large-scale fading statistical models without detailed knowledge of the characteristics of a site-specific environment [6].

In this work, we adopted the physically-based reference distance to be 1 m for the reason that the wireless signals at the frequency bands above 6 GHz exhibit significant path loss values in the first meter of the propagation away from the transmitting antenna [40]. Also, it will be easier to compare our work

with other works as most of the researches in the literature use 1 m reference distance. The free space path loss expressed in the dB scale at a reference distance  $d_0 = 1\text{ m}$  is given by:

$$FSPL(f, 1m) = 10 \log \left( \frac{4\pi f}{c} \right)^2. \quad (4.5)$$

Note that the CI model depends on one main parameter to be optimized, which is the PLE ( $n$ ). The dependency of the path loss model on the 3D Tx-Rx separation distance is characterized by this *unitless* parameter (PLE). This model depends on a physical anchor that catches path loss near the transmitting antenna. It is clear that the CI model has an intrinsic dependency on the frequency of propagation that exists in the FSPL term. This term's values vary from 48 to 82 dB when the frequency range is between 6 and 300 GHz, respectively. The CI model is suitable for single- and multi-frequency situations and can estimate the path loss from both co- and cross-polarization cases [177].

The minimum mean square error (MMSE) technique is used to optimize the CI model's parameter (i.e., the PLE). Using this approach, we can achieve the least error in fitting the real measured data by minimizing the SF standard deviation.

To predict the path loss with more accuracy and sensitivity to the small changes of the propagation environments, we add an independent parameter to the CI model's equation. The improved model has two terms that depend on the 3D Tx-Rx separation distance. This means that the path loss exponent principle exists in two parameters ( $n_1$  and  $n_2$ ), as presented in the following equation:

$$PL^{Imp. CI}(d)[dB] = FSPL(f, d_0) + 10n_1 \log_{10}(d) + 10n_2 (\log_{10}(d))^2 + X_{\sigma}^{Imp. CI}, d > 1m, \quad (4.6)$$

where  $n_1$  and  $n_2$  are the first order and second order of the PLE, respectively. This improvement of the CI will increase the opportunity to fit the real measured data collected from measurement campaigns and present more details in characterizing the wireless channel. Changing the environment where the signal can propagate, or the propagation's communication scenario (LOS or NLOS, etc.) will lead to a notable change in the values of  $n_1$  and  $n_2$ .

To have the closed-form of these parameters, let us assume that  $A = FSPL(f, d_0)$ ,  $B = PL^{Imp. CI}(d)$ ,  $D = 10 \log_{10}(d)$ , and  $E = 10 (\log_{10}(d))^2$ , then, the SF of Eq. (4.6) can be expressed as:

$$X_{\sigma}^{Imp. CI} = B - A - n_1 D - n_2 E. \quad (4.7)$$

The SF standard deviation ( $\sigma_{Imp. CI}$ ) can be determined from the experimental data using:

$$\sigma_{Imp. CI} = \sqrt{\frac{\sum (X_{\sigma}^{Imp. CI})^2}{N}}, \quad (4.8)$$

where  $N$  is the number of the Tx-Rx separation distances (i.e., the total number of the average path loss samples recorded). Now, we have to differentiate the numerator of Eq. (4.8) with respect to both  $n_1$  and  $n_2$  and equate the result to zero to have the optimum value of these parameters that will lead to the minimum value of the standard deviation as follows:

$$\frac{\partial}{\partial n_1} (\sum (B - A - n_1 D - n_2 E)^2) = 0, \quad (4.9)$$

$$\frac{\partial}{\partial n_2} (\sum (B - A - n_1 D - n_2 E)^2) = 0, \quad (4.10)$$

After the differentiation and simplification of the previous two equations, we have two linear equations which can be expressed as:

$$\sum D^2 n_1 + \sum (DE) n_2 = \sum (BD) - A \sum D, \quad (4.11)$$

$$\sum (DE) n_1 + \sum E^2 n_2 = \sum (BE) - A \sum E. \quad (4.12)$$

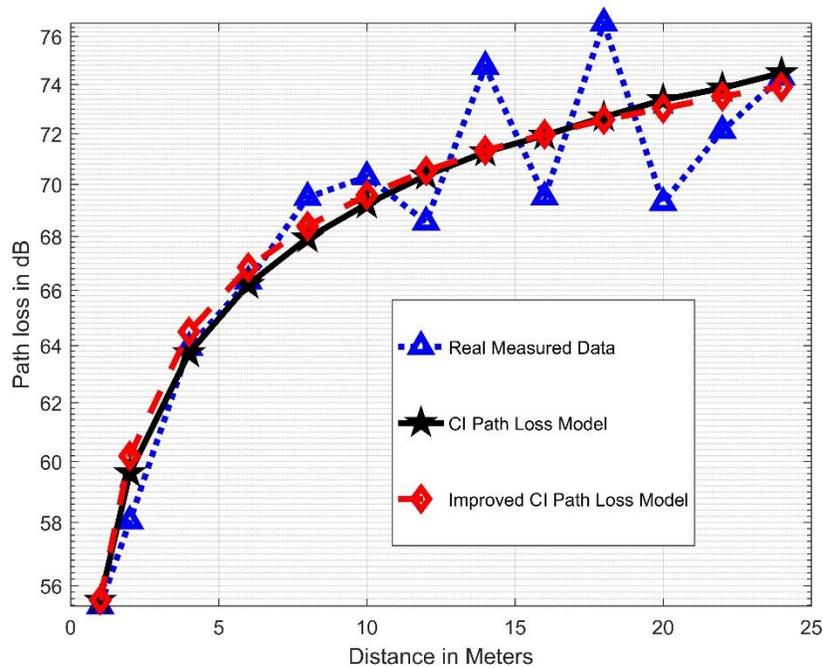
The matrix form of the equations (4.11) and (4.12) can be written easily as:

$$\begin{bmatrix} \sum D^2 & \sum (DE) \\ \sum (DE) & \sum E^2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \begin{bmatrix} \sum (BD) - A \sum D \\ \sum (BE) - A \sum E \end{bmatrix}. \quad (4.13)$$

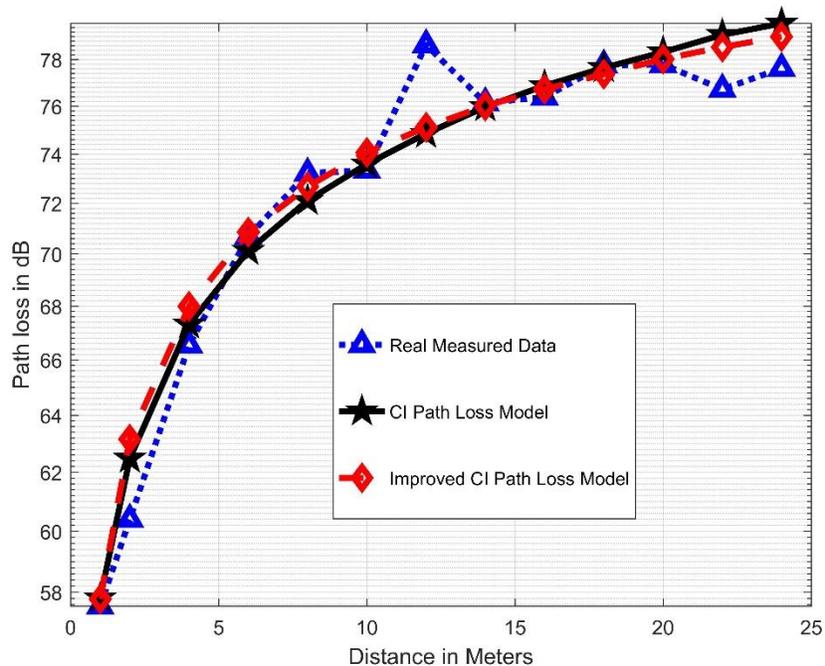
Finally, the closed-form of  $n_1$  and  $n_2$  can be found from the previous matrix.

Figures 4.2-4.4 depict the real measured path loss, the CI model, and the Improved-CI model together for the LOS communication scenario at the three frequencies. It is clear from the figures that both models fit the measured path loss adequately and both have a comparable performance with a slight preference of the improved model. This also can be noted from Table 4.1 that presents the values of the models' parameters. Our proposed model minimizes the shadow fading's standard deviation at the three frequencies by 2.3%, 5.2%, and 10.7% at 14, 18, and 22 GHz, respectively. We note that the reduction of the standard deviation becomes higher as we go for high-frequency bands. This is because the higher frequency bands suffer from many propagation effects and have higher path loss values than lower bands. From Table 4.1, for the CI model, the path loss exponent is directly proportional to the operating frequency. This leads to the general fact that higher frequency bands have higher path loss values. All the PLE's values are under the value of the FSPLE. The reason behind that is the constructive interference between the multipath signals, which makes the PLE lower than 2. For the improved-CI model, as discussed previously, the concept of the PLE has split into two parameters (i.e.,  $n_1$  and  $n_2$ ), as shown in Table 4.1. This technique

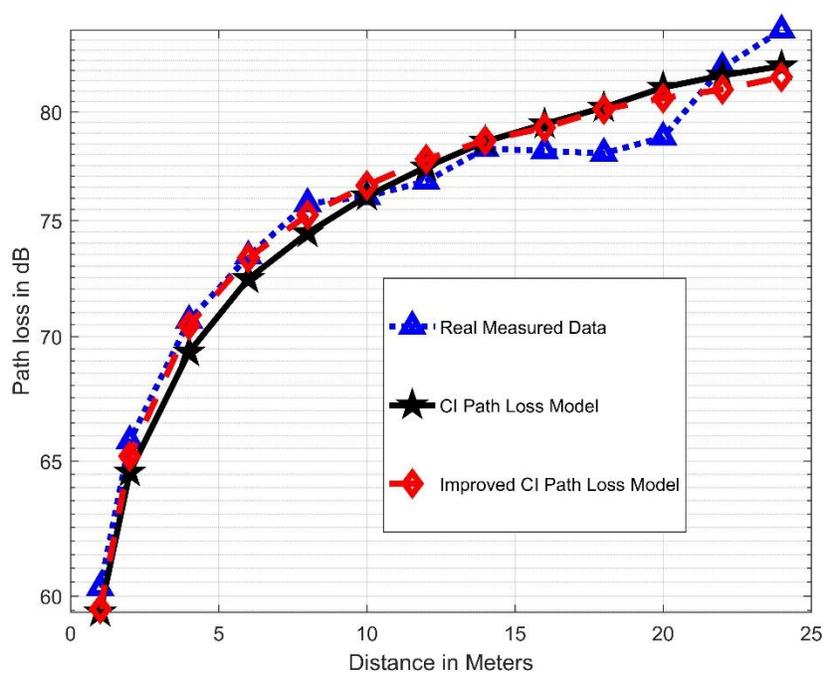
gives the model more relaxation to accurately fit the measured data. It is clear that the values of  $n_1$  are higher than the values of the CI model's PLE. However, all the values of  $n_2$  are negative, which will compensate the increase of  $n_1$  values and make the model following the measured path loss and counts all the possible signal effect. Figure 4.5 offers a graphical view of both models (CI and improved-CI) together at 14, 18, and 22 GHz for the LOS communication scenario. We plotted the MSE curves between the CI and improved CI models with respect to the separation distance between the Tx and Rx for the LOS scenario, as depicted in Figure 4.6. It is worth noting that all the MSE values are in the range of  $10^{-4}$  to  $10^{-2}$  with lower values around 14 to 16 meters (near the breakpoint of the corridor) of the Tx-Rx separation distance.



*Figure 4.2: Comparison of measured path loss, the fitted CI model, and the improved CI model for the LOS results at 14 GHz frequency band.*



*Figure 4.3: Comparison of measured path loss, the fitted CI model, and the improved CI model for the LOS results at 18 GHz frequency band.*



*Figure 4.4: Comparison of measured path loss, the fitted CI model, and the improved CI model for the LOS results at 22 GHz frequency band.*

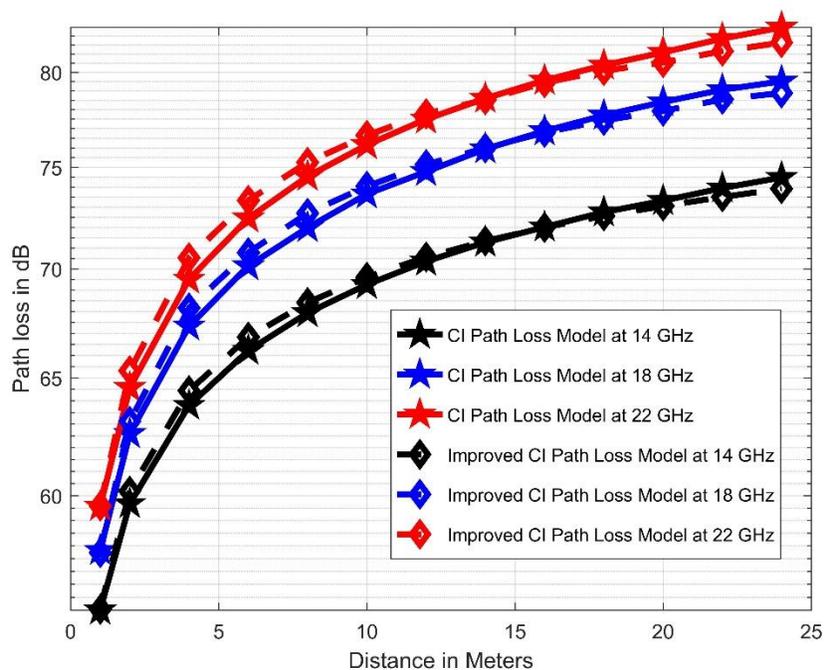


Figure 4.5: Directional large-scale path loss prediction models for the LOS communication scenario.

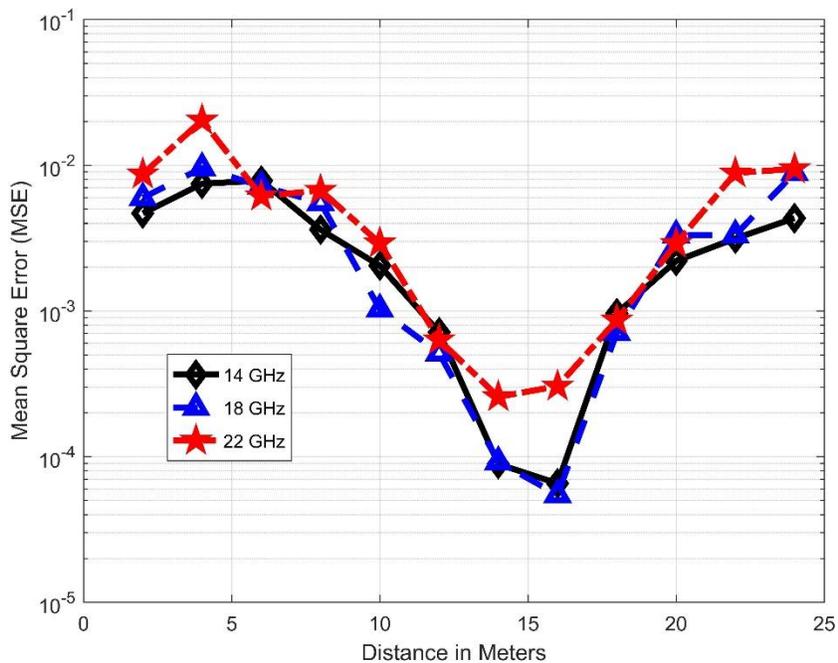


Figure 4.6: MSE between the CI and improved CI models in the LOS communication scenario.

*Table 4.1: A comparison between the CI and Improved-CI models' parameters in the LOS scenario.*

<i>Parameter</i>	<b>14 GHz</b>	<b>18 GHz</b>	<b>22 GHz</b>
<b><i>PLE</i> (<i>n</i>)</b>	1.37	1.58	1.66
<b><math>\sigma_{min}^{CI}</math> [dB]</b>	2.19	1.53	1.31
<b><math>n_1</math></b>	1.61	1.87	1.99
<b><math>n_2</math></b>	-0.20	-0.24	-0.28
<b><math>\sigma_{min}^{Imp. CI}</math> [dB]</b>	2.14	1.45	1.17

In the NLOS communication scenario, as it is known, the receiving antenna relies mainly on reflection, diffraction, and the effect of waveguiding in this enclosed environment for capturing the signals from the Tx. The CI model's PLE values here are 2.07, 2.38, and 2.26 at 14, 18, and 22 GHz, respectively. Note that because there is no direct dominant path from the transmitting to the receiving antennas, the values became notably high compared to the LOS results. However, they are still low compared to the communications in outdoor environments where the fluctuations of the propagated signals is much stronger. We find that in enclosed indoor environments such as corridors, the PLE values will not go much higher than the values of the FSPLE since the maximum percentage jump is in the 18 GHz band by 16%. This frequency band (i.e., 18 GHz) has a higher sensitivity to the wireless channel effects than the others for the frequencies studied. The SF standard deviation of the CI model is seen to rise to more than the double in the NLOS scenario, as shown in Table 4.2. This means that there is less precision in predicting the path loss in NLOS scenarios than in the LOS ones.

Nevertheless, the proposed model provides an attractive reduction of the standard deviation since it produces a minimum reduction of 3.24 dB (54.2% less), as can be seen in Table 4.2. This improvement is simple and highly efficient since almost all the communication methods for indoor environments are NLOS. The reason behind this improvement is the fact that the current statistical NLOS models cannot properly model the propagation mechanisms such as reflections and diffractions effects, which are captured better by the new parameter in our improved model. Again, all the values of the parameter  $n_2$  are negative, and  $n_1$  values are larger than the PLE value for the NLOS scenario as the LOS one. Figure 4.7 displays both models at the three selected frequency bands. From this figure, it is clear that the path loss curves of the improved CI model are away from the CI model's curves. The MSE between the CI and improved CI

models is outlined in Figure 4.8 for the NLOS scenario. The figure shows the increase in the MSE values compared to the LOS scenario since, for the NLOS, the MSE values are in the range of  $10^{-2}$  to  $10^0$ .

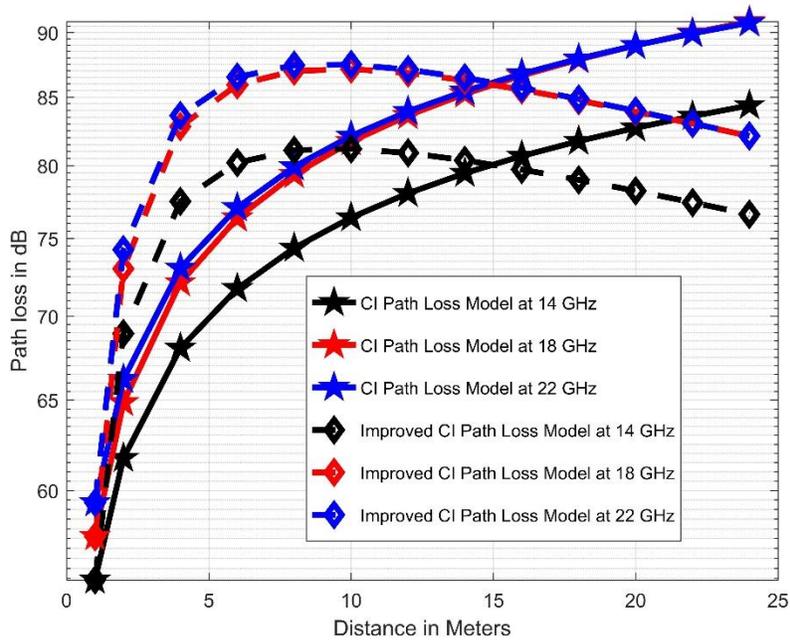


Figure 4.7: Directional large-scale path loss prediction models for the NLOS communication scenario.

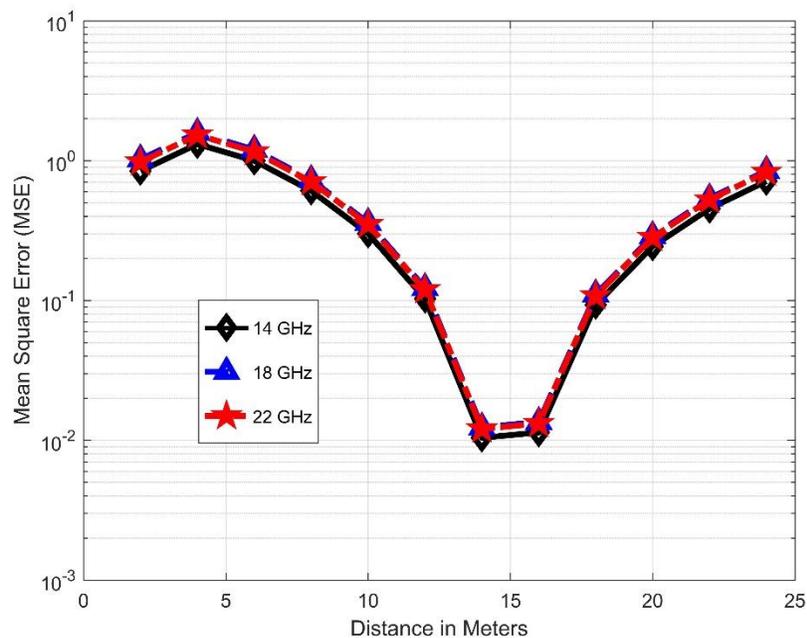


Figure 4.8: MSE between the CI and improved CI models in the NLOS communication scenario.

*Table 4.2: A comparison between the CI and Improved-CI models' parameters in the NLOS scenario.*

<i>Parameter</i>	<b>14 GHz</b>	<b>18 GHz</b>	<b>22 GHz</b>
<i>PLE (n)</i>	2.07	2.38	2.26
$\sigma_{min}^{CI}$ [dB]	5.98	6.87	6.86
$n_1$	5.09	5.86	5.69
$n_2$	-2.57	-2.97	-2.92
$\sigma_{min}^{Imp. CI}$ [dB]	2.74	3.08	3.23

The behavior of the CI and Improved-CI models' parameters with the AoA at 14, 18, and 22 GHz is presented in detail in Table 4.3, Table 4.4, and Table 4.5, respectively. From the Tables, it is clear that the PLE's minimum values occur when the AoA equals 30 and 330 *degrees*. The reason behind that is that at these AoA values, the Rx antenna is still aligned near the LOS path, and with the help of the propagation mechanisms discussed above, the PLE's value is minimized. The maximum PLE values occur at 150 and 210 *degrees* of the AoA. This shows that around these angles, when the Rx is in the opposite direction of the Tx, the propagated signal will suffer from maximum path loss values before reaching the Rx antenna. Note that when the AoA is exactly 180 *degrees*, the PLE values are within a good range compared to other AoA because of the back loops of the Rx antenna's radiation pattern, as can be seen from Figures 4.9-4.11. These findings give an insight into what will happen in reality when the Tx or Rx orientations might not be known and how the wireless signals will be affected according to this issue in such enclosed indoor corridor environments. When we compare the three frequencies together, we observe that the fluctuations of the PLE values are 6.8%, 10.4%, and 3.1% at 14, 18, and 22 GHz, respectively. This means that the 22 GHz frequency band has an attractive behavior in terms of its stability to the AoA, which leads to accurate modeling of the wireless propagation channel in the NLOS communication scenario. For the SF's standard deviation values, the difference between the maximum and minimum values are 2.04, 3.1, and 1.3 dB. It is clear that the worst performance between the three frequencies is at 18 GHz. For our proposed model, since the concept of the PLE is divided into two parameters, it is clear from the Tables that it outperforms the CI regarding the sensitivity and stability of the model's parameters with the AoA.

Table 4.3: The behavior of the CI and Improved-CI models' parameters with the AoA at 14 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
<b>PLE (n)</b>	2.0026	2.0950	2.0965	2.0954	2.1006	2.0604	2.1049	2.0968	2.0952	2.0832	1.9707
$\sigma_{min}^{CI}$ [dB]	4.73	6.18	6.25	6.20	6.40	6.37	6.63	6.37	6.30	5.80	4.59
$n_1$	4.4763	5.1908	5.2955	5.2026	5.3150	5.1780	5.3981	5.3147	5.2882	5.0508	4.3108
$n_2$	-2.105	-2.635	-2.722	-2.644	-2.736	-2.653	-2.803	-2.738	-2.717	-2.525	-1.991
$\sigma_{min}^{Imp. CI}$ [dB]	1.86	2.93	2.72	2.92	2.99	3.23	3.21	2.92	2.86	2.53	2.02

Table 4.4: The behavior of the CI and Improved-CI models' parameters with the AoA at 18 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
<b>PLE (n)</b>	2.1836	2.3478	2.4074	2.4114	2.4166	2.3503	2.4205	2.4110	2.4048	2.4198	2.3712
$\sigma_{min}^{CI}$ [dB]	4.42	6.20	7.10	7.31	7.36	7.24	7.52	7.18	6.99	7.65	6.56
$n_1$	4.5208	5.5663	6.0441	6.1024	6.1328	5.9643	6.1947	6.0495	5.9568	6.2216	5.7193
$n_2$	-1.989	-2.739	-3.095	-3.141	-3.163	-3.076	-3.212	-3.096	-3.023	-3.235	-2.849
$\sigma_{min}^{Imp. CI}$ [dB]	1.61	2.52	3.08	3.35	3.38	3.47	3.53	3.25	3.13	3.72	2.89

Table 4.5: The behavior of the CI and Improved-CI models' parameters with the AoA at 22 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
<b>PLE (n)</b>	2.2099	2.2585	2.2752	2.2747	2.2751	2.2523	2.2738	2.2775	2.2754	2.2542	2.2192
$\sigma_{min}^{CI}$ [dB]	5.93	6.58	7.03	7.13	7.22	7.23	7.18	7.21	7.20	6.70	6.03
$n_1$	5.2224	5.5827	5.7903	5.8321	5.8638	5.8427	5.8460	5.8715	5.8624	5.6136	5.3165
$n_2$	-2.564	-2.829	-2.991	-3.027	-3.054	-3.056	-3.040	-3.059	-3.053	-2.859	-2.636
$\sigma_{min}^{Imp. CI}$ [dB]	2.66	3.01	3.35	3.42	3.50	3.52	3.47	3.46	3.47	3.15	2.58

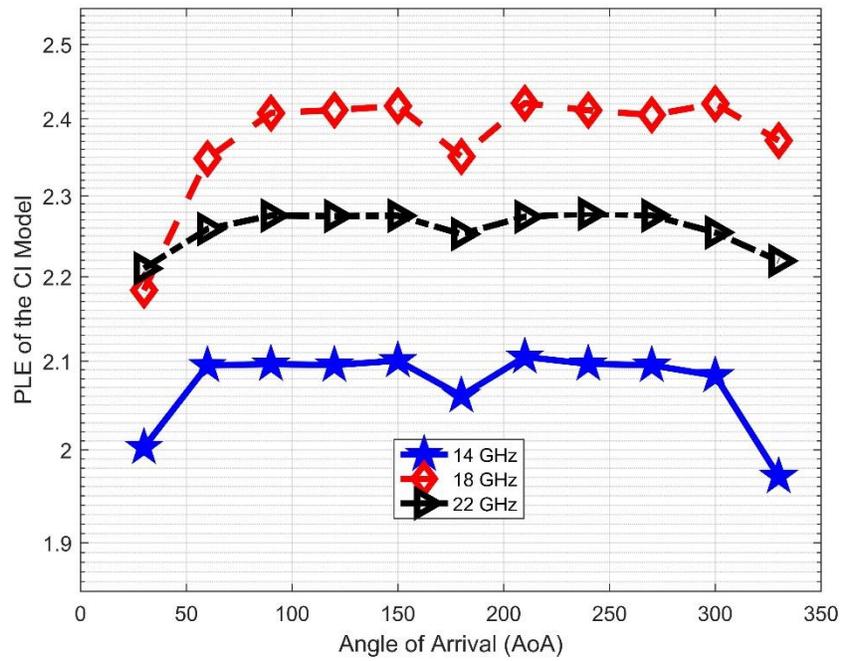


Figure 4.9: The behavior of the CI model's parameter (i.e., PLE) with the AoA.

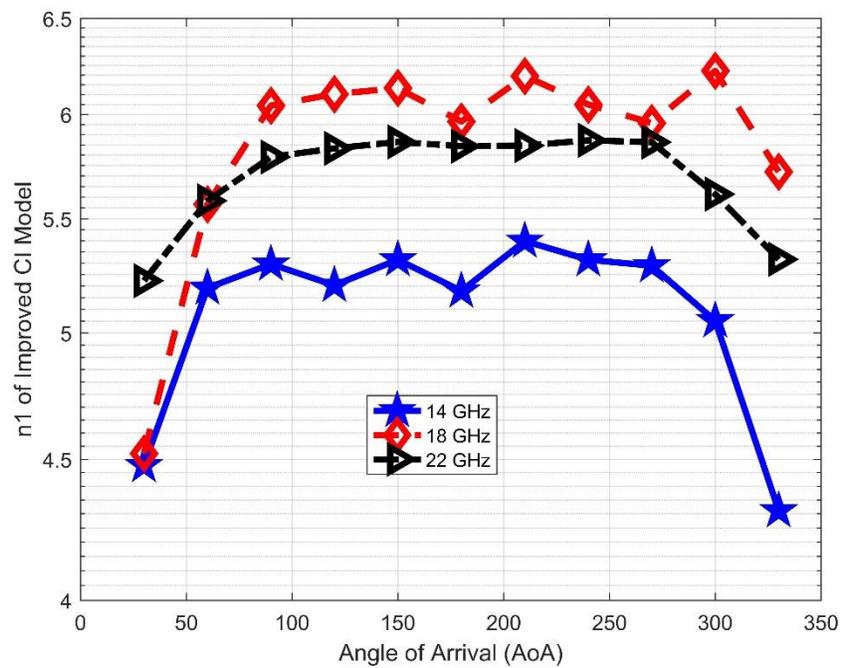
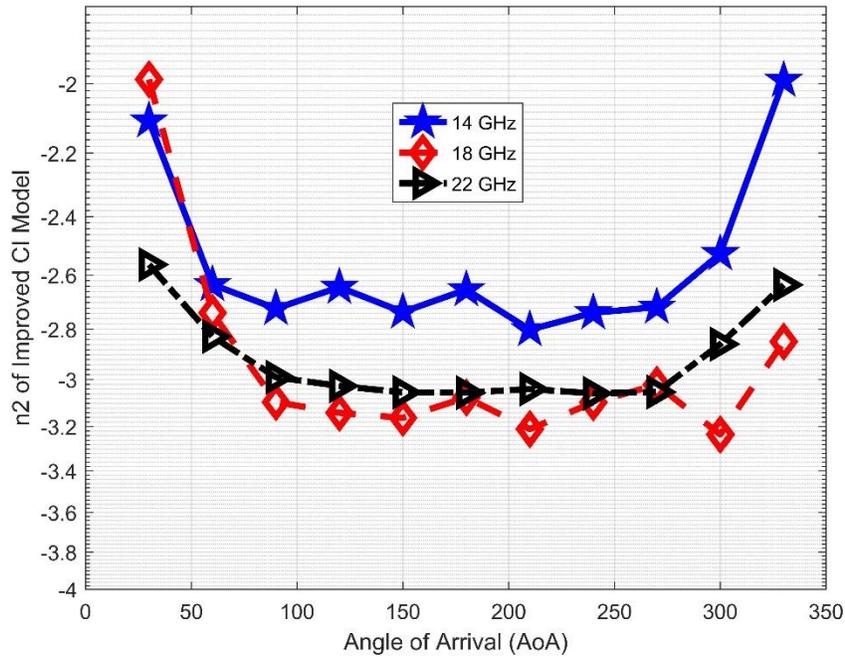
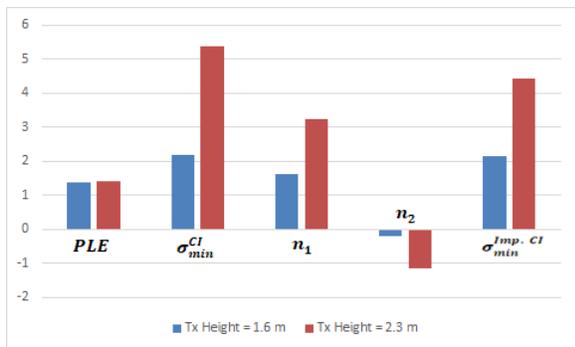


Figure 4.10: The behavior of the improved CI model's parameter ( $n_1$ ) with the AoA.

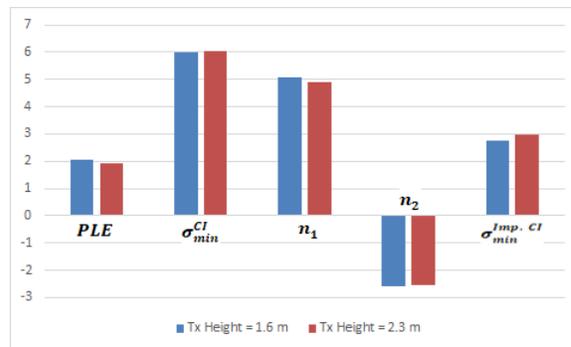


*Figure 4.11: The behavior of the improved CI model's parameter ( $n_2$ ) with the AoA.*

Figure 4.12 depicts the CI and improved-CI models' parameters at two different practical antenna heights (i.e., 1.6 and 2.3m) for both the LOS and NLOS communication scenarios. Generally, the figure shows the increase of the parameters' values when the Tx antenna height is 2.3 m compared to 1.6 Tx antenna height because of the mismatching of the antennas' heights. Moreover, it can be seen from the figure that our proposed model provides more sensitivity to the antenna height and capture more accurately the wireless propagation characteristics caused by the mismatching of the Tx and Rx antenna heights. Furthermore, it reveals that the antenna height's impact is minimum at 22 GHz and maximum at 14 GHz for both the CI model and our proposed model. However, when we look at the SF's standard deviation values, we observe that the proposed model outperforms the standard CI model depending on the antennas' heights. It is worth noting that the antenna height might not be an essential factor in the specific investigations presented in this work. However, the antenna locations, patterns, and relative orientation are significant factors, especially at high frequencies, mmWave and above. These performance studies will help engineers in designing reliable communication systems in such scenarios and have an accurate understanding and modeling of the wireless channel's behavior.



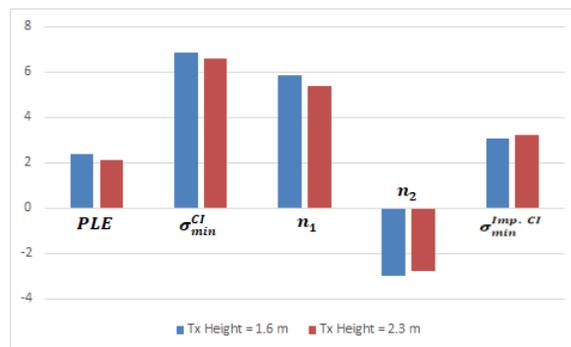
(a) 14 GHz and LOS.



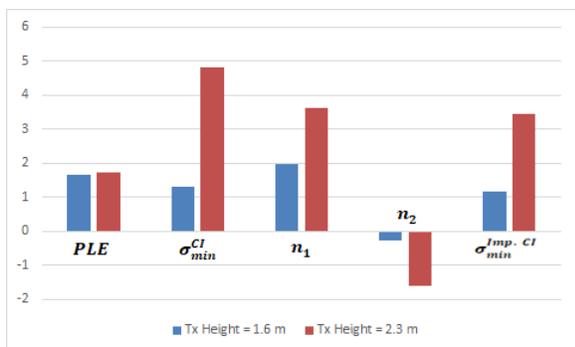
(d) 14 GHz and NLOS.



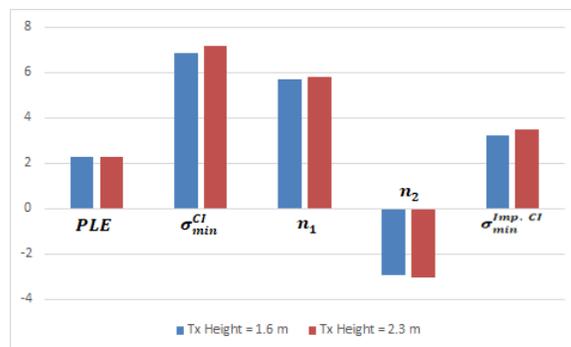
(b) 18 GHz and LOS.



(e) 18 GHz and NLOS.



(c) 22 GHz and LOS.



(f) 22 GHz and NLOS.

Figure 4.12: The parameters of the CI and improved-CI models at two different Tx antenna heights.

#### 4.4 The Standard FI Model and Our Improved FI Path Loss Prediction Model

The FI model has been widely used in 3GPP and WINNER II standards [5], [27], [P4/24], [28], [177]. It does not depend on the physical anchor point constraint that catches the path loss near the transmitting antenna. However, it depends on the mathematical curve that fits the measured path loss values. As a linear equation, the FI model has two parameters, which are the intercept (denoted by  $\alpha$ ) and slope (indicated by  $\beta$ ) of the path loss line as presented in the following equation:

$$PL^{FI}(d)[dB] = \alpha + 10\beta \log_{10}(d) + X_{\sigma}^{FI}, \quad (4.14)$$

where  $PL^{FI}(d)$  is the path loss in dB, and  $X_{\sigma}^{FI}$  is a Gaussian random variable with zero mean and a standard deviation  $\sigma_{FI}$ . In the previous expression, both  $X_{\sigma}^{FI}$  and  $\sigma_{FI}$  are in dB. It is worth noting that the FI model's parameters are unlike the CI model ( $\alpha$  is unlike FSPL, and  $\beta$  is unlike the PLE). However, the models have comparable overall performance in predicting the path loss with a preference of one over the other depending on the operating frequency as well as the environment and communication scenario of the wireless communication system [169].

For the Improved FI model, we follow the same principle as we used for the improved CI model, which is adding an independent parameter that will be the coefficient of the square of the logarithm of the 3D Tx-Rx separation distance as presented in the following equation:

$$PL^{Imp. FI}(d)[dB] = \alpha + 10\beta_1 \log_{10}(d) + 10\beta_2 (\log_{10}(d))^2 + X_{\sigma}^{Imp. FI}. \quad (4.15)$$

As seen from Eq. (4.15), this model has three parameters to be known ( $\alpha$ ,  $\beta_1$ , and  $\beta_2$ ). Using the MMSE approach and following the same derivation we used for the improved CI model, the solution matrix of these parameters can be expressed as:

$$\begin{bmatrix} N & \sum D & \sum E \\ \sum D & \sum D^2 & \sum(DE) \\ \sum E & \sum(DE) & \sum E^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum B \\ \sum(BD) \\ \sum(BE) \end{bmatrix}, \quad (4.16)$$

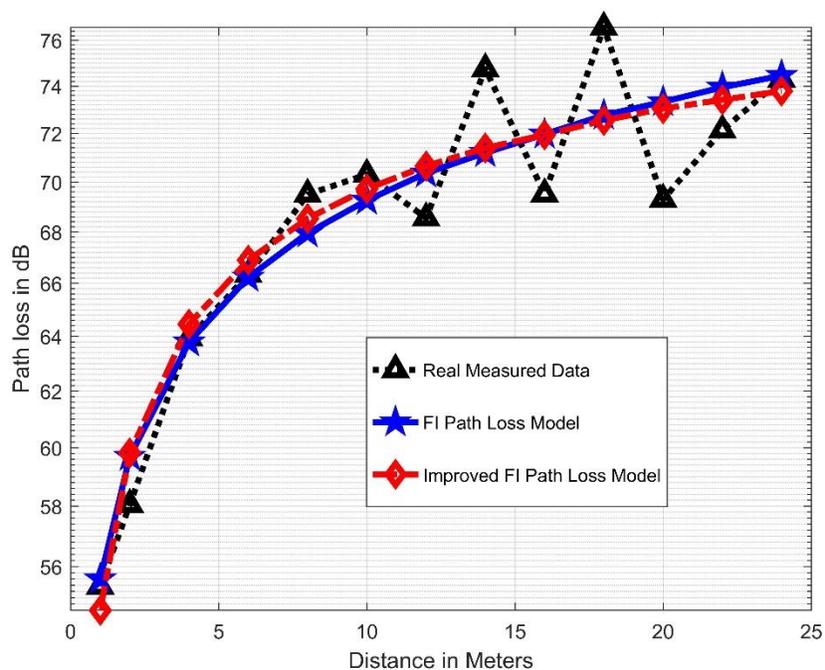
where  $B = PL^{Imp. FI}(d)$ ,  $D = 10 \log_{10}(d)$ , and  $E = 10(\log_{10}(d))^2$ . The closed-forms of the parameters are found from the previous matrix.

Figures 4.13-4.15 show the curves of the real measured data, the FI model, and the improved-FI model for the LOS communication scenario at the three frequency bands adopted for this work. The figures show that both models accurately fit the real measured data with the minimum possible MSE between the models and the data. As a comparison between the three frequencies, the best fit occurs at the 22 GHz frequency band. It must be emphasized that any path loss model will always depend on the operating frequency, no

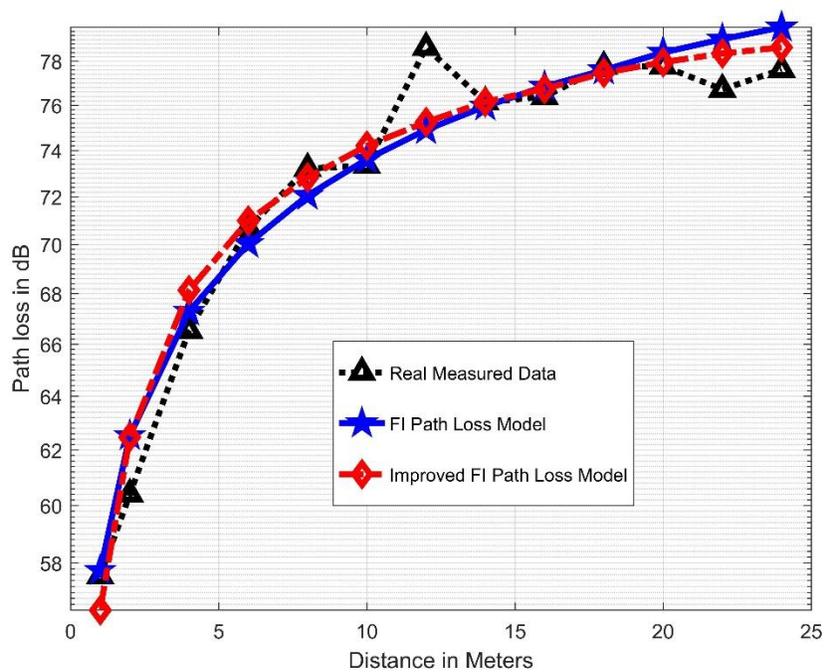
matter how they are derived. Maybe simplifications could come from the fact that the exact carrier frequency is not needed, and only the knowledge of the band (e.g., 14 GHz, 18 GHz, and 22 GHz) is enough. The models' parameters are represented in Table 4.6. It is observed that our proposed model slightly better than the standard FI model in terms of the performance since it reduces the standard deviation values by 3.2%, 13.7%, and 1.8% at 14, 18, and 22 GHz. Note that the best improvement applies at 18 GHz, contrary to what happened between the CI and improved-CI models. Also, note that the values of the parameter  $\alpha$  are not far between both models. A notable improvement can be observed from Table 4.7 since the percentage reduction goes up to 44% in the NLOS communication scenario. Upon comparing the four models, it is noted that the best model that fits the real measured data is the improved-FI model. Figure 4.16 presents both FI and improved-FI models together at the three frequencies selected. The curves of the two models in the figure are almost the same, as also confirmed from the values in Table 4.6. We have plotted the MSE between both models to show how these models behave when the Tx-Rx separation distance increases. It is noted that from Figure 4.17, the correlation between the models is high since the values of the MSE are in the range of  $10^{-6}$  and  $10^{-2}$  with minimum values around the breakpoint. Nevertheless, Figure 4.18 and Figure 4.19 show a highly notable difference between both models in terms of their performance in the NLOS scenario since the models' curves are far away from each other in Figure 4.18, and have MSE values higher than the ones we got from Figure 4.19.

*Table 4.6: A comparison between the FI and Improved-FI models' parameters in the LOS scenario.*

<i>Parameter</i>	<i>14 GHz</i>	<i>18 GHz</i>	<i>22 GHz</i>
$\alpha_{FI}$ [dB]	55.443	57.478	61.036
$\beta_{FI}$	1.365	1.590	1.503
$\sigma_{min}^{FI}$ [dB]	2.19	1.53	1.12
$\alpha_{Imp. FI}$ [dB]	54.410	56.136	60.652
$\beta_1^{Imp. FI}$	1.827	2.191	1.675
$\beta_2^{Imp. FI}$	-0.315	-0.409	-0.117
$\sigma_{min}^{Imp. FI}$ [dB]	2.12	1.38	1.10



*Figure 4.13: Comparison of measured path loss, the fitted FI model, and the improved FI model for the LOS results at 14 GHz frequency band.*



*Figure 4.14: Comparison of measured path loss, the fitted FI model, and the improved FI model for the LOS results at 18 GHz frequency band.*

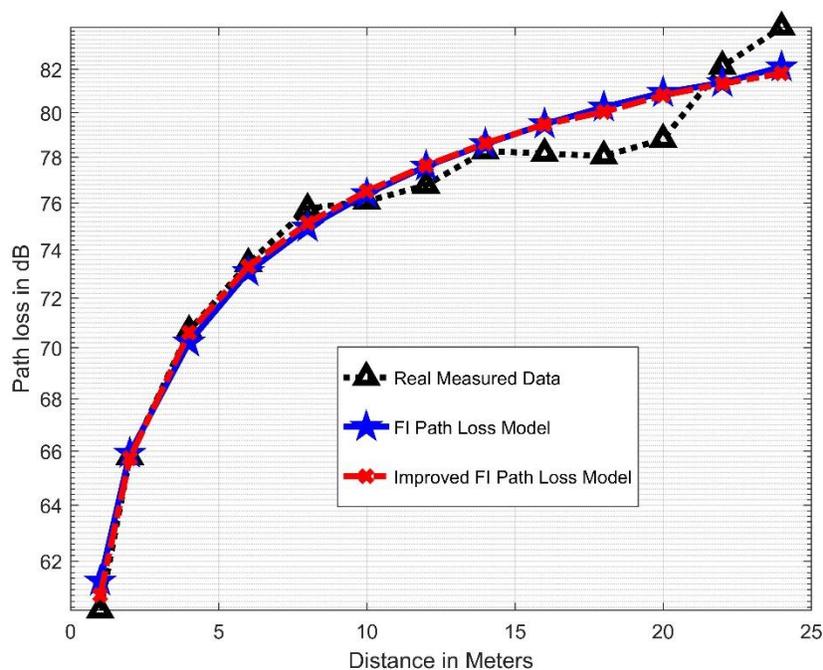


Figure 4.15: Comparison of measured path loss, the fitted FI model, and the improved FI model for the LOS results at 22 GHz frequency band.

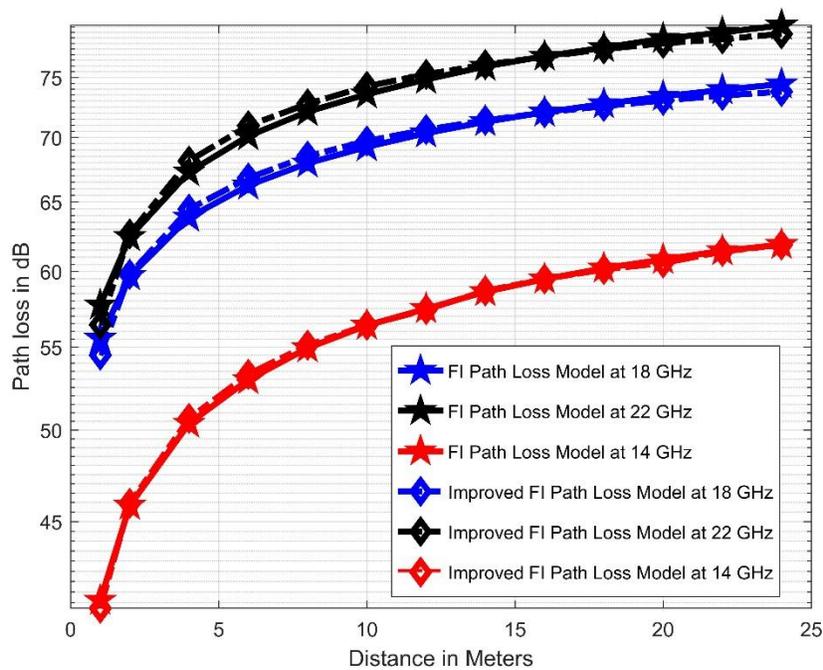


Figure 4.16: Directional large-scale path loss prediction models for the LOS communication scenario.

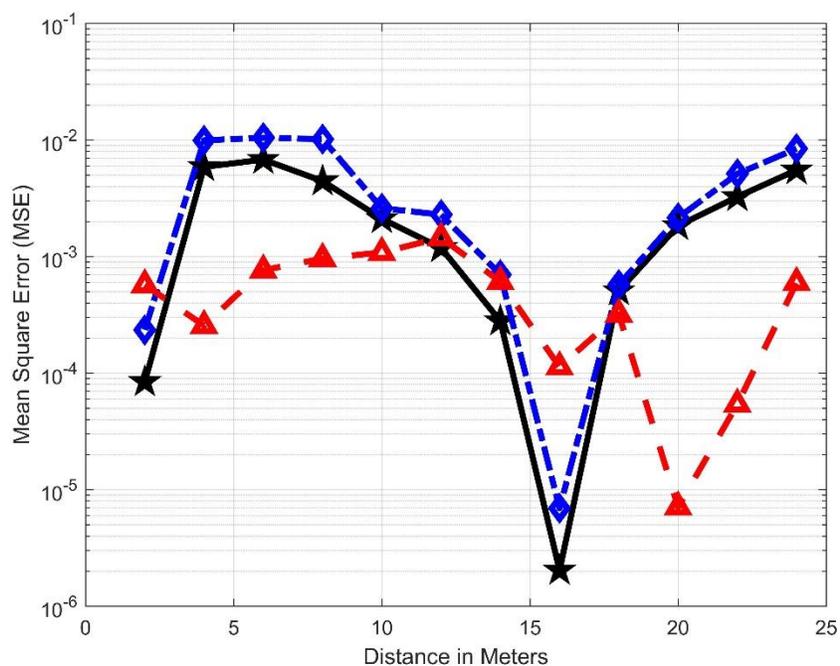


Figure 4.17: MSE between the FI and improved FI models in the LOS communication scenario.

Table 4.7: A comparison between the FI and Improved-FI models' parameters in the NLOS scenario.

Parameter	14 GHz	18 GHz	22 GHz
$\alpha_{FI}$ [dB]	67.272	71.021	73.242
$\beta_{FI}$	1.010	1.172	1.014
$\sigma_{min}^{FI}$ [dB]	3.69	4.32	4.07
$\alpha_{Imp. FI}$ [dB]	61.097	63.767	66.476
$\beta_1^{Imp. FI}$	3.772	4.430	4.039
$\beta_2^{Imp. FI}$	-1.881	-2.218	-2.061
$\sigma_{min}^{Imp. FI}$ [dB]	2.10	2.42	2.36

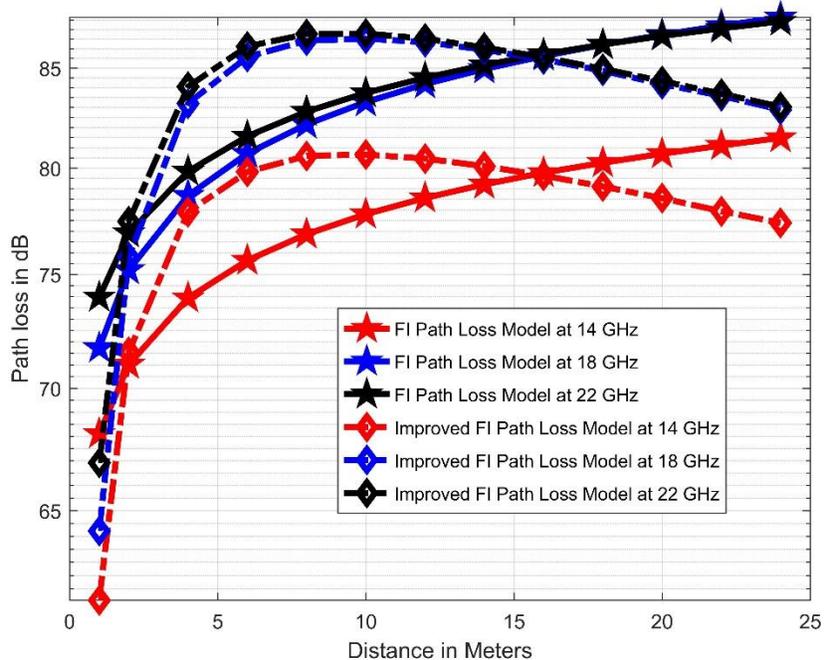


Figure 4.18: Directional large-scale path loss prediction models for the NLOS communication scenario.

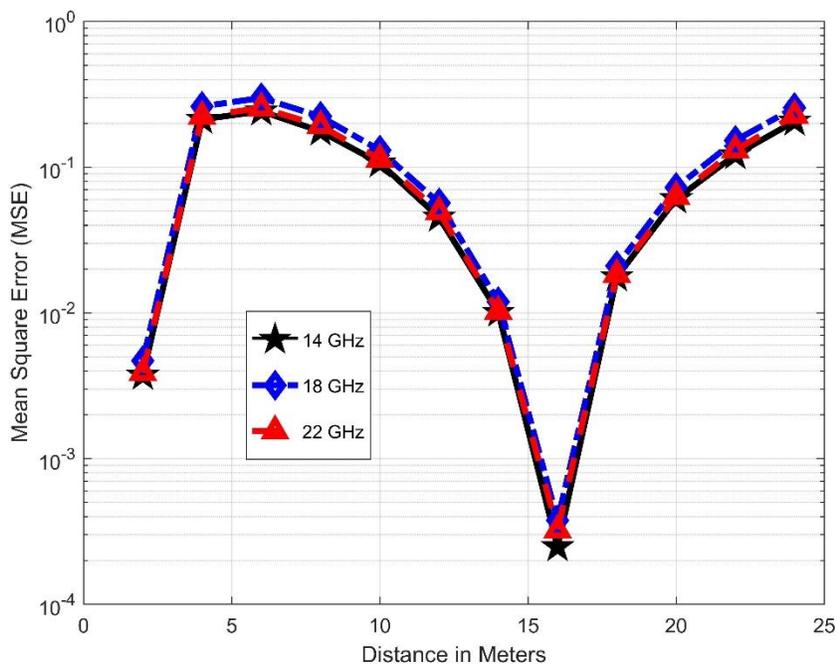
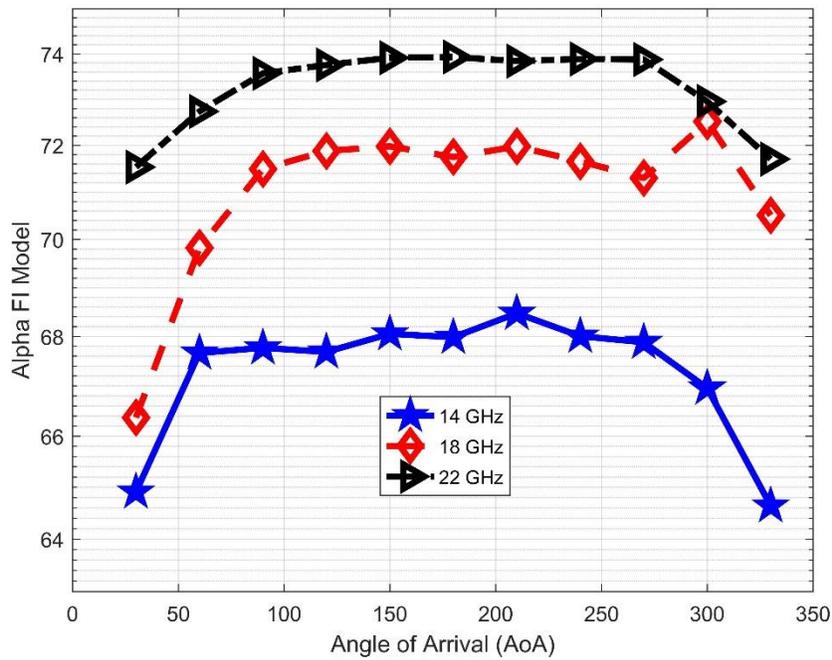


Figure 4.19: MSE between the FI and improved FI models in the NLOS communication scenario.

Figures 4.20-4.24 represent the FI and improved-FI models' parameters versus the AoA of the receiving antenna at 14, 18, and 22 GHz for the NLOS communication scenario. These figures show that while both models provide valuable stability to the change of the AoA, our proposed model shows a slight advantage. This advantage is seen clearly from Tables 4.8, 4.9, and 4.10, where the models' parameters are presented. The impact of the antenna heights on the models' parameters is presented in Figure 4.25. As for the behavior of the CI model and the proposed improved version, when there is a change in the antenna's height, there will be a corresponding change in the models' parameters, and the worst behavior is seen to occur when there is a mismatch in the antenna's heights. However, the performance of our proposed model is better than the performance of the standard FI model.



*Figure 4.20: The behavior of the FI model's parameter ( $\alpha$ ) with the AoA.*

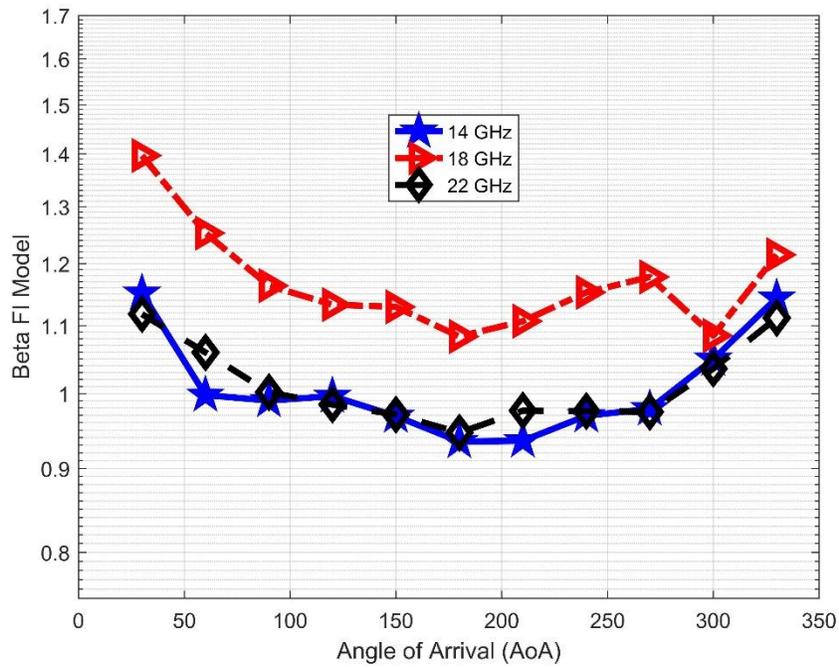


Figure 4.21: The behavior of the FI model's parameter ( $\beta$ ) with the AoA.

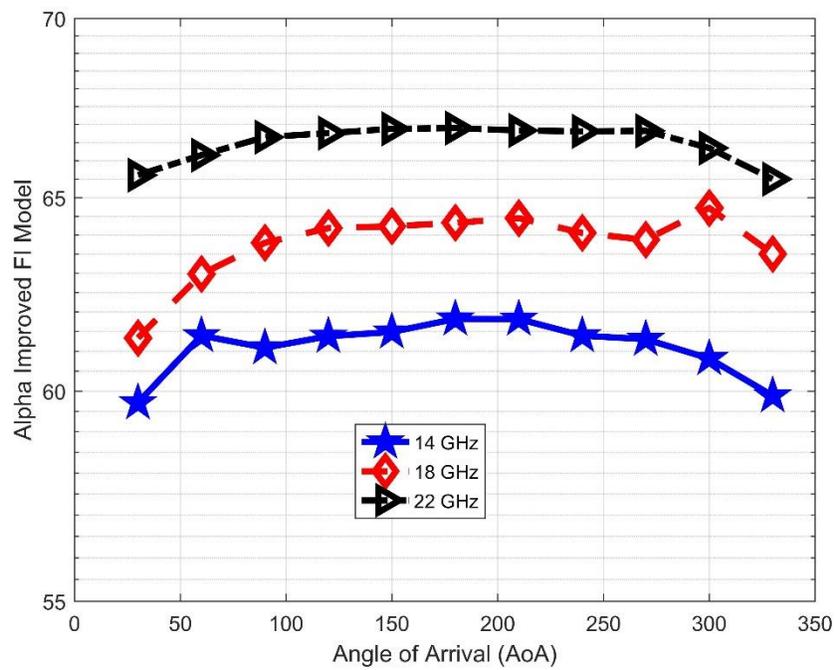


Figure 4.22: The behavior of the improved FI model's parameter ( $\alpha$ ) with the AoA.

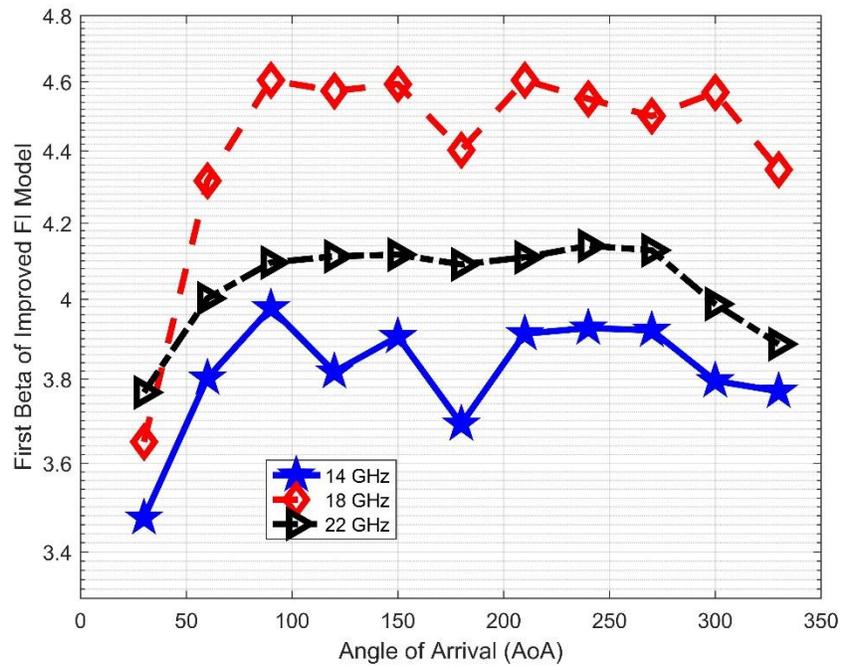


Figure 4.23: The behavior of the improved FI model's parameter ( $\beta_1$ ) with the AoA.

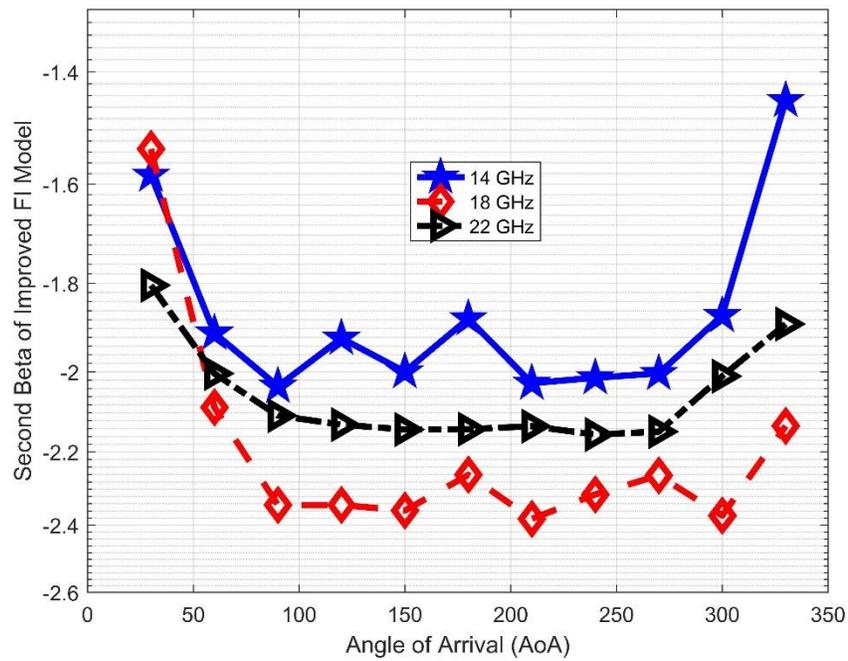
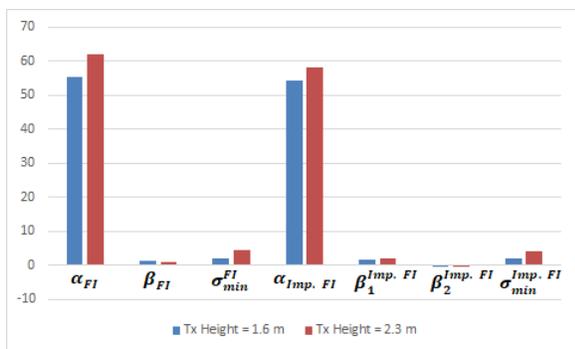
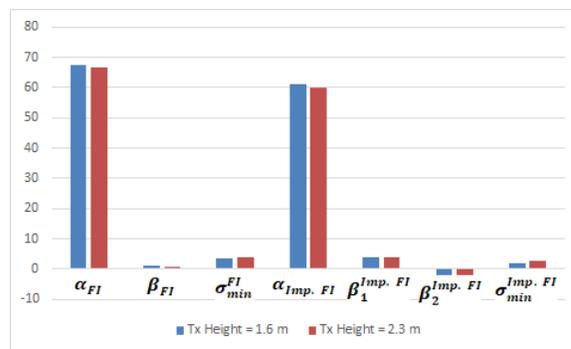


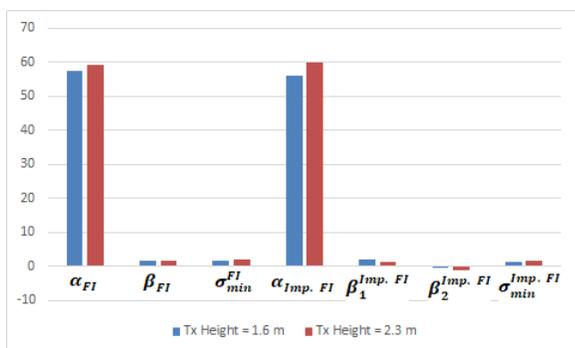
Figure 4.24: The behavior of the improved FI model's parameter ( $\beta_2$ ) with the AoA.



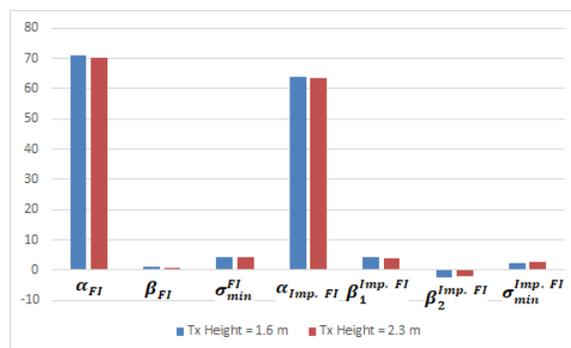
(a) 14 GHz and LOS.



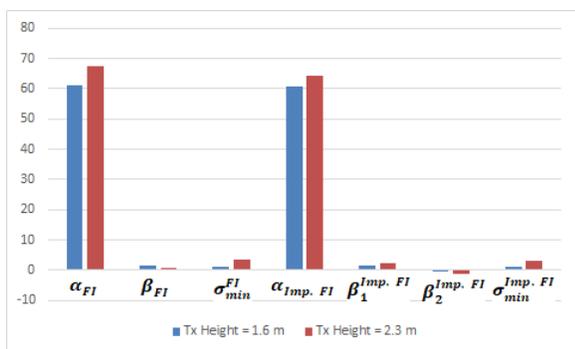
(d) 14 GHz and NLOS.



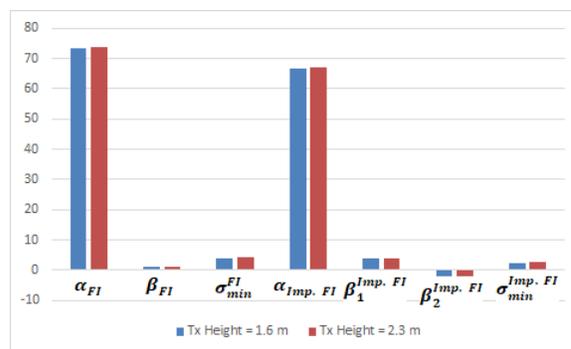
(b) 18 GHz and LOS.



(e) 18 GHz and NLOS.



(c) 22 GHz and LOS.



(f) 22 GHz and NLOS.

Figure 4.25: The parameters of the FI and improved-FI models at two different Tx antenna heights.

Table 4.8: The behavior of the FI and Improved-FI models' parameters with the AoA at 14 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
$\alpha_{FI}$ [dB]	64.907	67.663	67.766	67.683	68.051	67.981	68.470	68.003	67.879	66.954	64.638
$\beta_{FI}$	1.151	0.998	0.990	0.996	0.969	0.935	0.936	0.969	0.979	1.049	1.143
$\sigma_{min}^{FI}$ [dB]	2.86	3.82	3.88	3.83	3.97	3.96	4.13	3.95	3.91	3.56	2.76
$\alpha_{Imp. FI}$ [dB]	59.709	61.391	61.084	61.376	61.484	61.816	61.813	61.390	61.300	60.813	59.882
$\beta_1^{Imp. FI}$	3.475	3.802	3.977	3.817	3.905	3.691	3.912	3.926	3.921	3.795	3.269
$\beta_2^{Imp. FI}$	-1.58	-1.91	-2.03	-1.92	-2.01	-1.88	-2.03	-2.01	-2.01	-1.87	-1.45
$\sigma_{min}^{Imp. FI}$ [dB]	1.29	2.26	2.07	2.26	2.33	2.55	2.53	2.26	2.20	1.89	1.47

Table 4.9: The behavior of the FI and Improved-FI models' parameters with the AoA at 18 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
$\alpha_{FI}$ [dB]	66.368	69.824	71.492	71.878	71.977	71.749	71.974	71.653	71.301	72.513	70.507
$\beta_{FI}$	1.397	1.253	1.163	1.133	1.129	1.083	1.107	1.153	1.178	1.085	1.215
$\sigma_{min}^{FI}$ [dB]	2.71	3.86	4.47	4.61	4.65	4.58	4.76	4.52	4.39	4.85	4.11
$\alpha_{Imp. FI}$ [dB]	61.330	62.972	63.796	64.183	64.230	64.325	64.449	64.055	63.873	64.721	63.503
$\beta_1^{Imp. FI}$	3.649	4.316	4.604	4.573	4.592	4.402	4.604	4.549	4.499	4.568	4.346
$\beta_2^{Imp. FI}$	-1.53	-2.09	-2.34	-2.34	-2.36	-2.26	-2.38	-2.31	-2.26	-2.37	-2.13
$\sigma_{min}^{Imp. FI}$ [dB]	1.11	1.89	2.40	2.65	2.68	2.77	2.82	2.57	2.45	2.99	2.24

Table 4.10: The behavior of the FI and Improved-FI models' parameters with the AoA at 22 GHz frequency band.

	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
$\alpha_{FI}$ [dB]	71.531	72.732	73.565	73.751	73.910	73.927	73.836	73.883	73.871	72.949	71.703
$\beta_{FI}$	1.118	1.059	1.002	0.985	0.971	0.946	0.976	0.976	0.975	1.036	1.112
$\sigma_{min}^{FI}$ [dB]	3.43	3.88	4.19	4.27	4.33	4.33	4.30	4.32	4.31	3.96	3.50
$\alpha_{Imp. FI}$ [dB]	65.606	66.151	66.646	66.758	66.876	66.896	66.829	66.806	66.819	66.347	65.498
$\beta_1^{Imp. FI}$	3.767	4.001	4.095	4.111	4.115	4.090	4.108	4.139	4.127	3.987	3.886
$\beta_2^{Imp. FI}$	-1.80	-2.00	-2.10	-2.13	-2.14	-2.14	-2.13	-2.15	-2.14	-2.01	-1.88
$\sigma_{min}^{Imp. FI}$ [dB]	1.82	2.15	2.47	2.54	2.61	2.63	2.58	2.56	2.57	2.28	1.74

From the previous analysis, we observed that it is possible to provide valuable improvements to the existing wireless channel models without a notable increase in the models' complexity. In fact, with the demand for more and more data traffic, we will always need to go for higher frequency bands to meet future requirements. However, these higher frequency bands have smaller wavelengths and suffer more from the wireless propagation channel. Hence, improving the existing models and developing new models that accurately describe the wireless propagation channel will always be needed.

## 4.5 Chapter Summary

In this chapter, we proposed in detail an efficient approach for improving standard path loss models, namely the CI and FI models. The proposed improved models were derived, and extensive performance analyses were studied to prove the superiority of our improved models in terms of many factors, such as the prediction accuracy, behavior with the change of the AoA and antenna heights, and the sensitivity of the models' parameter to the change of these effects. This research has proved the efficiency of this approach since our improved models outperformed the existing models in all the studied cases. Hence, these models are promising for predicting path loss in enclosed environments for future wireless communications. The question that might come to our mind is, *is it possible to apply this approach for higher orders seeking better prediction accuracy?* This question has motivated the research presented in the following chapter.

# CHAPTER 5

**INVESTIGATIONS INTO THE EFFECT OF HIGH-  
ORDERING THE LOG-DISTANCE DEPENDENCY  
OF PATH LOSS MODELS**

---

## ***Chapter 5: Investigations into the Effect of High-Ordering the Log-Distance Dependency of Path Loss Models***

### **5.1 Introduction**

The massively available amounts of bandwidth that exist in the frequency regime beyond 6 GHz, together with the inability of the previous frequencies below 6 GHz, have attracted the attention of the research and industrial communities to consider new frequency bands such as the SHF, mmWave, and further high-frequency bands such as the THz bands to be the promising solutions to cope with the explosive demand of higher mobile data rate traffic. As a result, many research studies have reported the main wireless channel's characteristics, such as the path loss over wide ranges of the SHF, mmWave, and THz frequency bands. However, most path loss modeling research has adopted the well-known CI model to describe propagation effects of the wireless channel leaving behind the improvement concept. It is vital to estimate the path loss accurately since these high-frequency bands significantly suffer from the wireless propagation effects because of their tiny wavelength. This chapter presents investigations on high-ordering the dependency of the standard CI path loss model on the distance between the transmitting and the receiving antennas in the logarithmic scale. This work discusses two improved models: the second-order CI model (our improved CI model in the previous chapter) and the third-order CI model. The models' accuracy is validated based on the measurement data discussed previously. Finally, a trade-off study between the model's accuracy and simplicity is provided in this chapter.

### **5.2 The Third-Order CI Path Loss Prediction Model**

As mentioned in the previous chapter, the second-order CI path loss prediction model is a new simple two-parameter single-frequency model as an extension of the standard CI model. The definition of PLE is divided into two parameters. The first one controls the dependency of the model on the Tx-Rx distance in the logarithmic scale, while the second one controls the dependency on the square of the Tx-Rx distance in the logarithmic scale. These two parameters provide the CI model more sensitivity to the propagation effects such as reflections and diffractions and offer more ease in describing the frequency and environment dependency of the model. Note that the second-order CI model reverts to the standard CI model when the parameter  $n_2 = 0$ . However, this condition will not be accrued since both parameters ( $n_1$  and  $n_2$ ) are distance-dependent.

In order to provide more precision in predicting all possible propagation effects, a three-parameter distance-dependent CI path loss model, namely the third-order CI path loss model, is provided. Here, the concept of PLE is provided by three parameters that depend on the Tx-Rx log distance and its squared and cubic values. It may come into the reader's mind that having more parameters than three will provide more accuracy. However, the improvement on the model will be less than the improvement on the standard and second-order CI models since the third-order CI model already captures more propagation effects. Moreover, increasing the number of parameters will make the model's complexity higher, and communication systems' designers will not consider it since the representation of each frequency band and a specific communication environment will be described by more than three parameters. In fact, three-parameter path loss models already exist, such as the alpha-beta-gamma (ABG) path loss model. The third-order CI path loss model is expressed as:

$$PL^{\log^3 CI}(d) = FSPL(f, d_0) + 10k_1 \log_{10} \left( \frac{d}{d_0} \right) + 10k_2 \left( \log_{10} \left( \frac{d}{d_0} \right) \right)^2 + 10k_3 \left( \log_{10} \left( \frac{d}{d_0} \right) \right)^3 + X_{\sigma}^{\log^3 CI}. \quad (5.1)$$

In Eq. (5.1),  $PL^{\log^3 CI}(d)$  represents the average path loss value in dB predicted by the third-order CI model at a certain Tx-Rx separation distance in *meters*. The parameters  $k_1$ ,  $k_2$ , and  $k_3$  are the controllers of the model's dependency on the log distance powered by one, two, and three, respectively. The values of the three parameters are provided from the measurement data via the concept of the MMSE.

### 5.3 Results and Discussions on the Impact of High-Ordering the Log-Distance Dependency

This section presents and discusses the main results obtained from this research study. The comparative analysis has been performed based on the best fit of the path loss prediction models using the criteria of the MMSE. The results of both LOS and NLOS communication scenarios are provided separately in this section. Figures 5.1, 5.2, and 5.3 depict the comparison between the real measured data, the standard CI model, the second-order CI model, and the third-order CI model at 14, 18, and 22 GHz frequency bands for the LOS communication scenario. From the figures, it is clear that the three models fit the data with high accuracy. However, the third-order CI model has the best performance since it closely follows the path loss data. This can be notable clearly from Table 5.1, which presents the three models' parameters together at 14, 18, and 22 GHz frequency bands for the LOS scenario. The Table shows an attractive reduction of the shadow fading's standard deviation as the order of the model has increased. Nevertheless, this reduction is not linear, and it changes from one frequency band to another. For example, at 14 GHz, the SF's standard

deviation of the standard CI model is 2.1897, and for the second and third-order models, 2.1504 and 2.0459, respectively.

For the standard CI model, the PLE values for the LOS connection between the Tx and Rx are 1.3720, 1.5843, and 1.6584 at 14, 18, and 22 GHz, respectively. Thus, it can be observed from these values that the power degradation is more than by 2.1 dB/decade at 18 GHz relative to 14 GHz and less than 1 dB/decade at 22 GHz relative to 18 GHz. It should be noted that all the PLE values prove that the measured environment is an enclosed area since the richness of the propagation characteristics provides constructive interference between the wireless received signals making all PLE's values less than the free space path loss exponent (FSPL) value (i.e., all PLE values are less than 2). This validates the prediction accuracy since the studied environment is an indoor corridor environment with many multipath effects during the measurement campaigns. These multipath effects are mainly the reflections and the diffractions from the corridor's walls, with the waveguiding effect that provides high path gain added to the direct LOS signals. It should be noted from Table 5.1 that there is no high diversity in values of the second-order and third-order CI models. Hence, more stability is provided by these two models compared to the standard CI model.

The reduction from the first to the second-order models is 1.8%, and from the second to the third-order models, 4.9%. Consequently, the optimum model in terms of both accuracy and complexity is the third-order CI model. This is because this model has the lowest values of the standard deviation and fits with the best precision the real measured data, which means that this model can be sensitive to all the possible effects that affect the propagation of the electromagnetic wave between the Tx and Rx. In fact, more precision is always needed due to the movement to higher frequency bands such as the mmWave, Terahertz bands, and beyond due to the small values of wavelength. However, it is hard to go back to the previous microwave bands because of bandwidth scarcity. From the models' complexity point of view, the complexity of the third-order CI model is not high when the improvement of the model's performance is considered as a trade-off between complexity and accuracy. It can be noted that this model only has three parameters, which is the same number of the parameters of well-known path loss prediction models such as the alpha-beta-gamma (ABG) model. The sequence of improving the accuracy is the same for both 18 and 22 GHz bands. However, as the frequency increases, the reduction of the SF's standard deviation is increased, and higher precision is achieved.

Table 5.1: The LOS Comparative Study Results.

Parameter	14 GHz	18 GHz	22 GHz
$PLE (n)$	1.3720	1.5843	1.6584
$\sigma_{min}^{CI} [dB]$	<b>2.1897</b>	<b>1.5337</b>	<b>1.3120</b>
$n_1$	1.6070	1.8656	1.9884
$n_2$	-0.1999	-0.2394	-0.2808
$\sigma_{min}^{log^2 CI} [dB]$	<b>2.1504</b>	<b>1.4518</b>	<b>1.1778</b>
$k_1$	0.5288	0.0426	2.9882
$k_2$	2.0198	3.5139	-2.3393
$k_3$	-1.0767	-1.8207	0.9985
$\sigma_{min}^{log^3 CI} [dB]$	<b>2.0459</b>	<b>0.9246</b>	<b>1.0052</b>

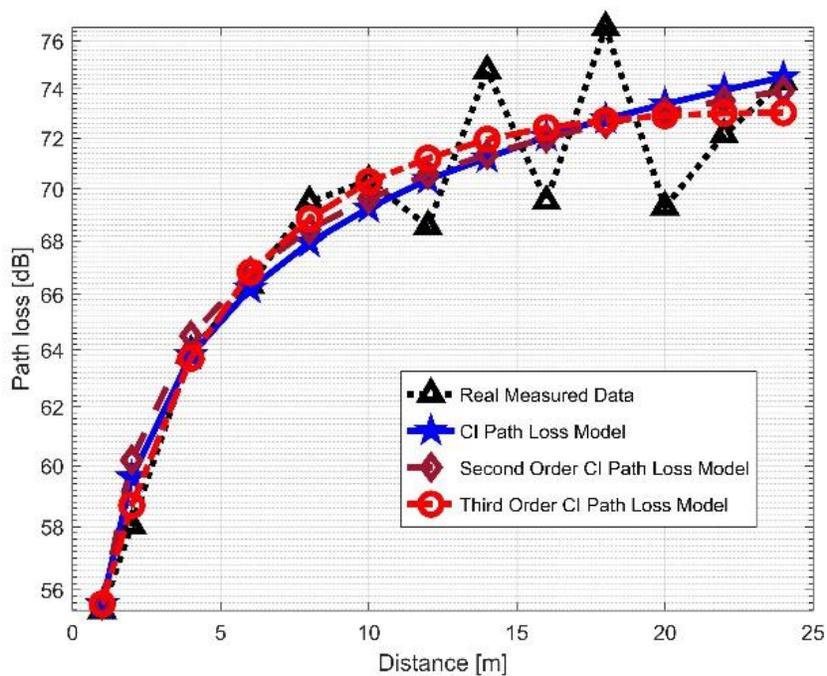


Figure 5.1: LOS Path loss models with measurement data at 14 GHz.

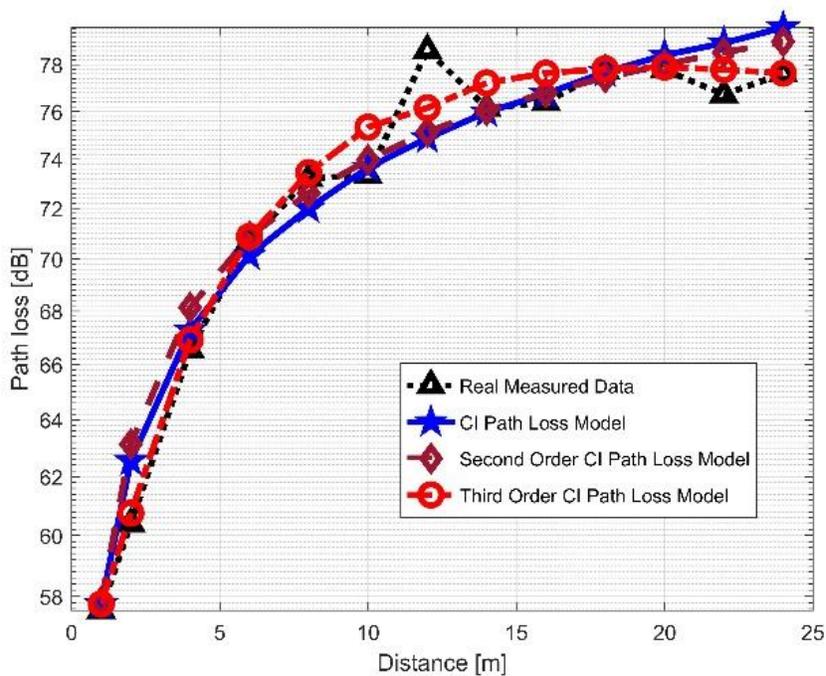


Figure 5.2: LOS Path loss models with measurement data at 18 GHz.

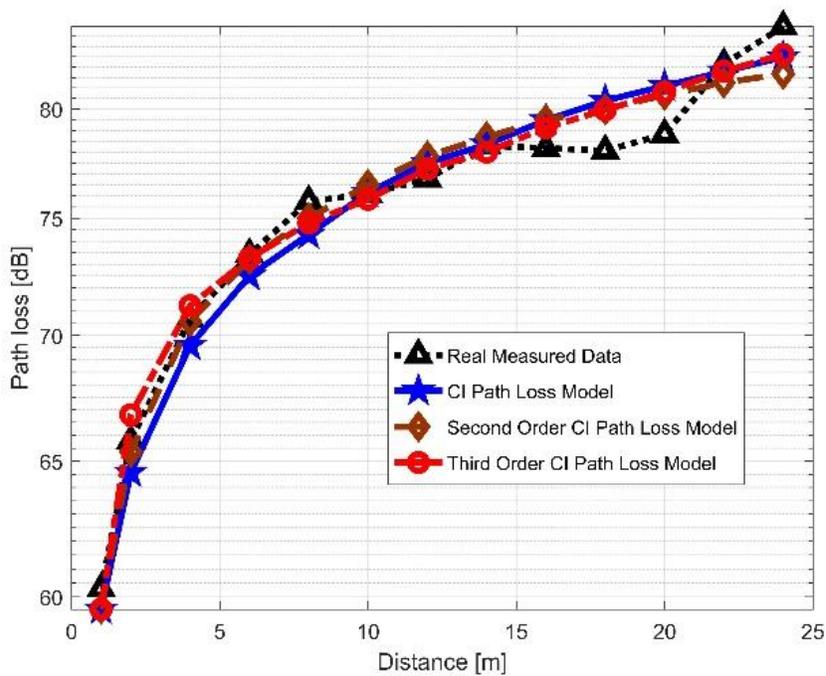


Figure 5.3: LOS Path loss models with measurement data at 22 GHz.

As a comparison between the three frequency bands together (i.e., 14, 18, and 22 GHz), the extraordinary improvement is in 18 GHz since the reduction is 5.3% and 36.3%. It is worth noting that the signals' fluctuations because of the propagation mechanisms such as reflections and diffractions depend on the type of the environment, the communication scenario, the materials used in the environment, and many other factors. Having more parameters will provide more details in describing the wireless channel and then more precision for the models to be used to calculate the coverage for systems' deployment. For example, the values of the models' parameters at 14, 18, and 22 GHz for the indoor corridor environment are presented. The materials in this corridor are dry concrete and bricks and include a staircase, elevator, and wooden doors to offices. Since higher frequency bands suffer more from the propagation mechanisms, the models' parameters will not be close to the parameters for another corridor with the same size but different materials. In other words, the sensitivity of the models is increased. Figure 5.4 provides a vision of the three models together at the three frequency bands selected for the LOS scenario. As it is known generally, higher frequency bands suffer from higher path loss values, as is to be noted from the figure.

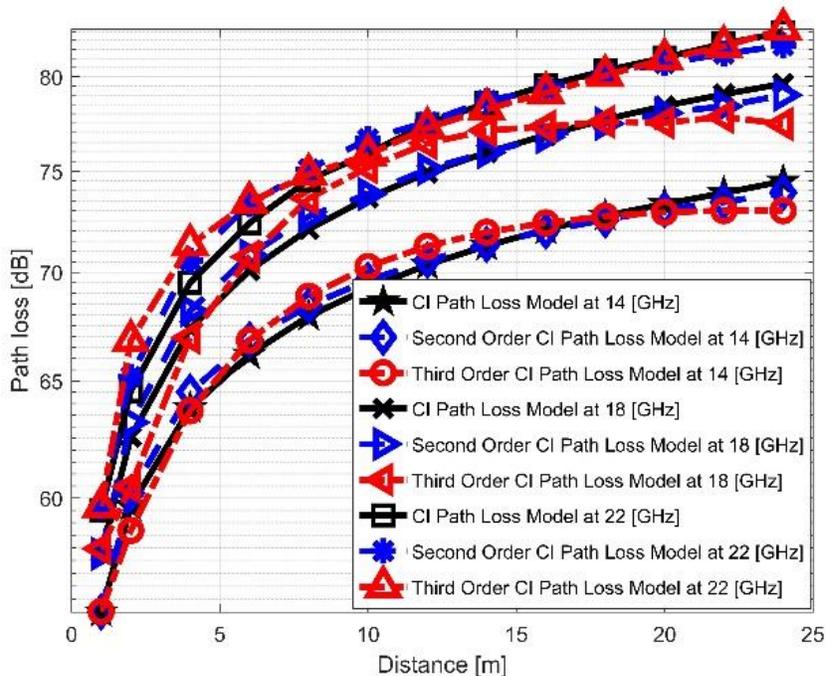


Figure 5.4: LOS Path loss models at 14, 18, and 22 GHz.

Table 5.2 provides the models' parameters at 14, 18, and 22 GHz for the NLOS scenario. It is clear from the Table that the models' parameters are higher than the values in the LOS communication scenario since the Rx relies mainly on the effects of waveguiding, reflections, and diffractions in the corridor environment. It can be noted that the values of the standard deviation are much higher than the ones of the LOS scenario. The improvement of high-ordering the CI model is more beneficial here in the NLOS scenario since the reduction exceeds half of the values. The reduction from the standard CI model to the second-order CI model is 3.2346, 3.7779, and 3.619 dB at 14, 18, and 22 GHz, respectively. The model's performance improvement is 1.3245, 1.573, and 1.5325 dB for the second to third-order CI models. All these values show an attractive and non-specific improvement of the standard CI model since this improvement does not depend on any specific factor. Figures 5.5, 5.6, and 5.7 depict a comparison between the NLOS data and the NLOS models at 14, 18, and 22 GHz, respectively. It is clear from these figures that the three models behave differently in the NLOS scenario. In addition, for the NLOS scenario, excellent convergence is seen between the measured and the presented predicted models as the distance increases. Again, the third-order CI path loss model provides the best performance of fitting the real measured data since almost a total matching between the model and the measured path loss can be seen in the figures. The three models can be seen together in Figure 5.8 at the three frequency bands selected.

*Table 5.2: The NLOS Comparative Study Results.*

<i>Parameter</i>	<b>14 GHz</b>	<b>18 GHz</b>	<b>22 GHz</b>
<i>PLE (n)</i>	2.0728	2.3768	2.2587
$\sigma_{min}^{CI}$ [dB]	<b>5.9837</b>	<b>6.8662</b>	<b>6.8582</b>
$n_1$	5.0928	5.8611	5.6949
$n_2$	-2.5699	-2.9652	-2.9242
$\sigma_{min}^{log^2 CI}$ [dB]	<b>2.7491</b>	<b>3.0883</b>	<b>3.2392</b>
$k_1$	9.4763	10.7479	10.7740
$k_2$	-11.3065	-12.7696	-13.1519
$k_3$	4.1563	4.6924	4.8962
$\sigma_{min}^{log^3 CI}$ [dB]	<b>1.4246</b>	<b>1.5153</b>	<b>1.7067</b>

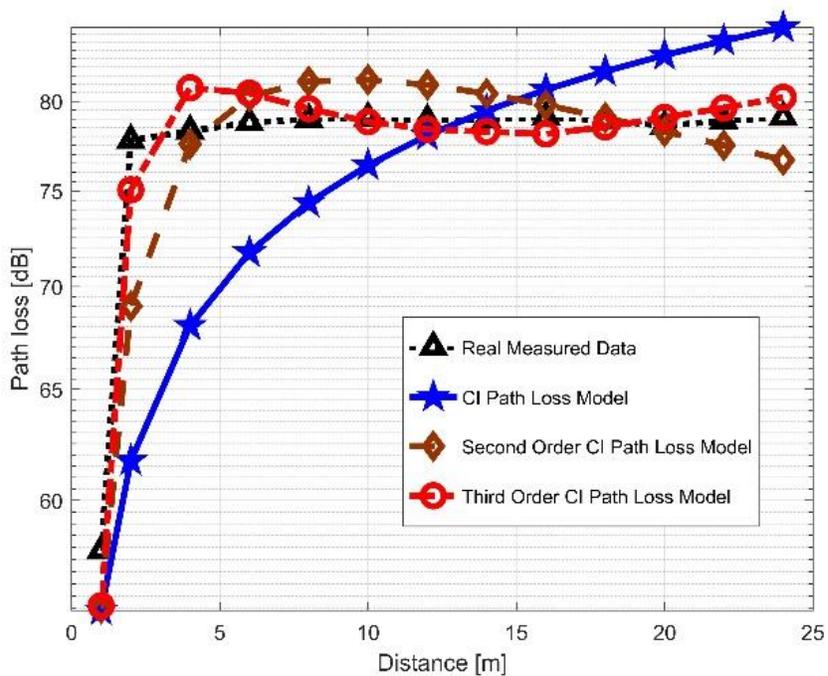


Figure 5.5: NLOS Path loss models with measurement data at 14 GHz.

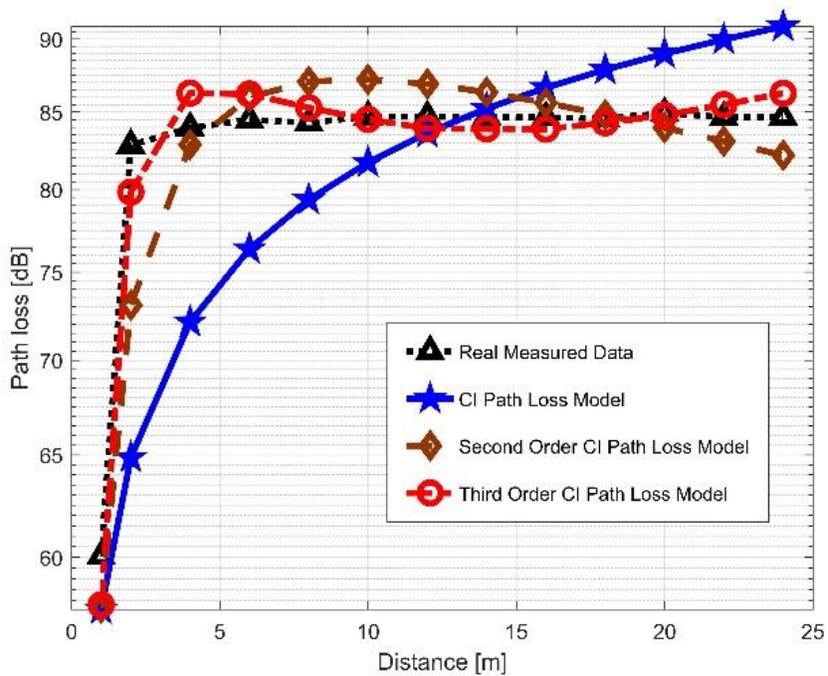


Figure 5.6: NLOS Path loss models with measurement data at 18 GHz.

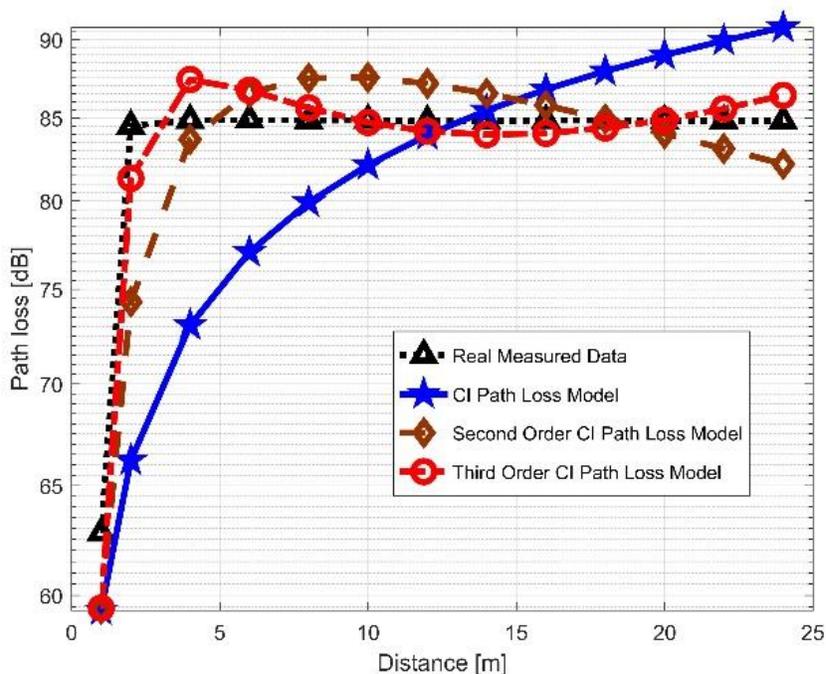


Figure 5.7: NLOS Path loss models with measurement data at 22 GHz.

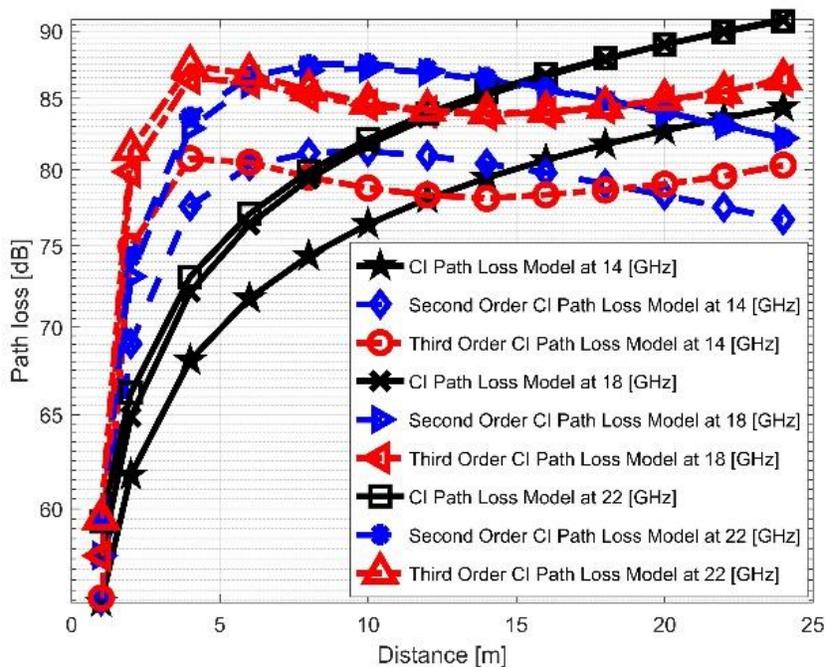


Figure 5.8: NLOS Path loss models at 14, 18, and 22 GHz.

## 5.4 Results and Discussions Related to Error Analysis of the Proposed Models

The models' performance can be further examined using valuable metrics such as mean prediction error (MPE), mean square error (MSE), standard deviation error (SDE), spread corrected root mean square error (SC-RMSE), efficiency (EF), error rate, and Gaussian kernel density estimation (GKDE) [178], [179]. These metrics measure the performance of the received signal strength predicted by the models compared to the actual received data collected from the measurement campaigns. In addition, the previous metrics can be applied directly in order to analyze the correlation between the path loss prediction models and measured path loss data. Accordingly, both the MPE and the SDE have been adopted for this research work in order to provide adequate performance evaluation of the path loss prediction models.

The prediction error (PE) can be defined as the difference between the value of the real Rx power (from the measurements) and the Rx power predicted by the path loss model as a function of the Tx-Rx separation distance as presented in the following equation:

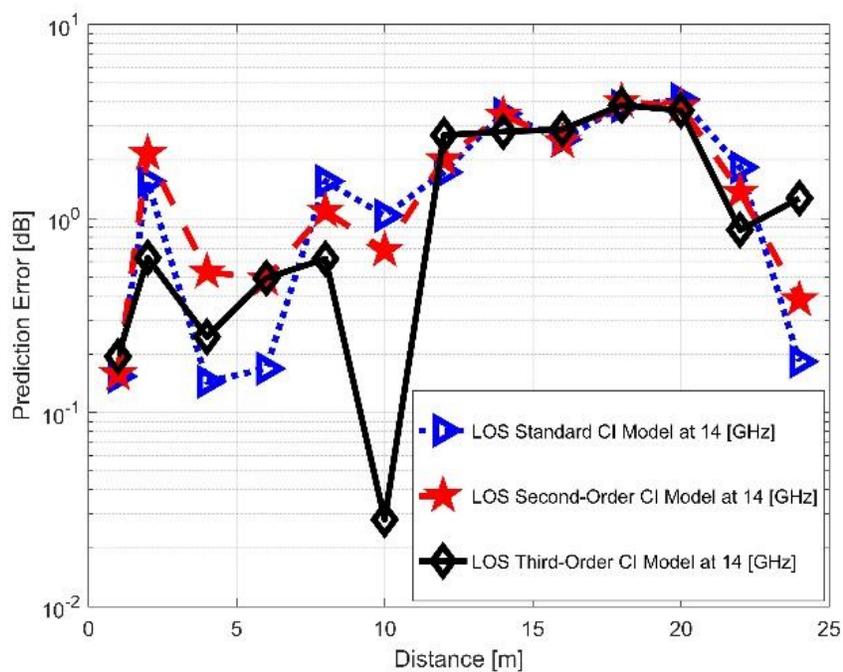
$$PE(d)[dB] = P_r^M(d)[dBm] - P_r^P(d)[dBm], \quad (5.2)$$

where  $PE(d)$  is the value of the PE (in dB) at a specific Tx-Rx separation distance in meters.  $P_r^M(d)$  and  $P_r^P(d)$  are the measured and predicted Rx power in dBm. The values of  $P_r^M(d)$  are taken directly from the measurement data, while  $P_r^P(d)$  is given by:

$$P_r^P(d)[dBm] = P_t - PL^P(d)[dB] + G_t + G_r. \quad (5.3)$$

In Eq. (5.3),  $P_t$  is the Tx power in dBm,  $G_t$  and  $G_r$  are the transmitting and receiving antenna gains in dBi, and  $PL^P(d)$  is the predicted path loss (in dB) by the models. For the predicted LOS Rx power at the receiving antenna side, the prediction errors (PEs) between the received signal strength from the measurement data and the predicted Rx power from the three models are provided in Figures 5.9, 5.10, and 5.11 at 14, 18, and 22 GHz, respectively. It can be observed from the figures that the third-order CI path loss model has a better error performance compared to the other modes since it provides the minimum PE. For the 14 GHz frequency band, the three models' MPE is 1.7138, 1.7364, and 1.5506 dB for the standard CI, the second-order CI, and the third-order CI models for the LOS condition. The third-order CI model's reduction in the MPE is 9.52% and 9.13% over the standard and the second-order models, respectively. The reduction is more attractive for the 18 GHz frequency band since the values have jumped up to 42.18% and 38.13%. This reduction reveals that the third-order CI model provides more accuracy in predicting the propagation effects that control the travel of the wireless signals between the Tx and Rx antennas. The MPE values at 22 GHz for the three models are 1.1119, 0.9743, and 0.8233 dB. Since the MPE is used to measure the bias of predictions (or whether the model is more likely to under-predict or over-predict), the three

models provide better performance and higher precision at 22 GHz than the other lower frequency bands (i.e., 14 and 18 GHz) simply because lower MPE values mean better fitness of the models to the actual measurement data.



*Figure 5.9: LOS PE of path loss models at 14 GHz.*

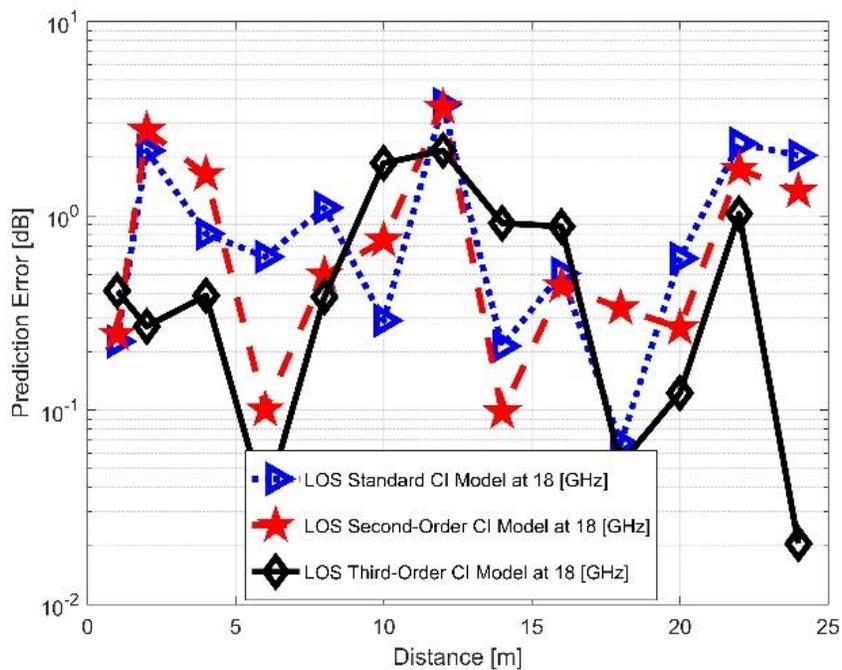


Figure 5.10: LOS PE of path loss models at 18 GHz.

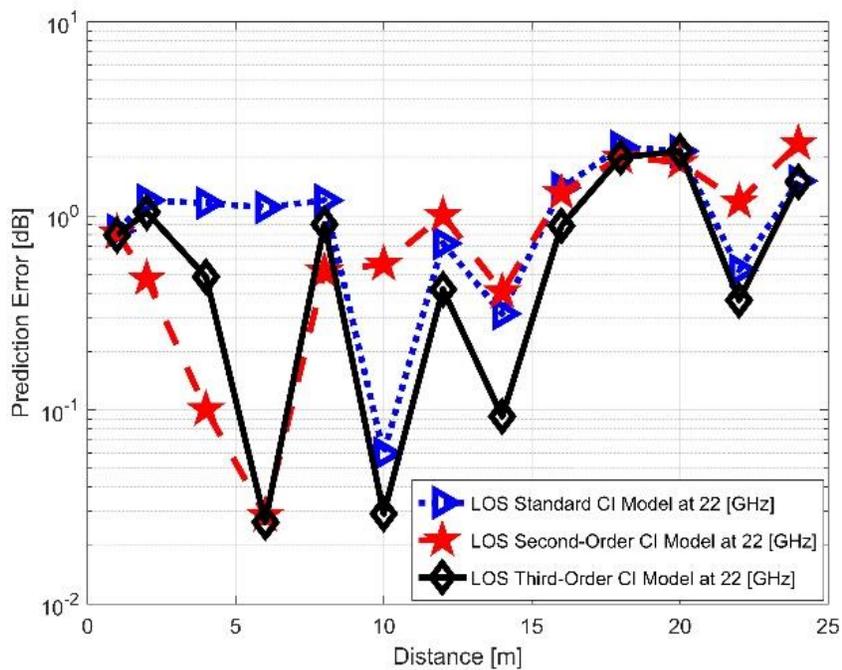


Figure 5.11: LOS PE of path loss models at 22 GHz.

The SDE is a measure of how much the errors deviate from the MPE value, and it is written as:

$$SDE[dB] = \sqrt{\frac{1}{N} \sum_{i=1}^N (PE_i - MPE)^2}, \quad (5.4)$$

where  $N$  is the number of the Tx-Rx separation distances (i.e., the total number of the average path loss samples recorded). The SDE values of the three models are 1.3585, 1.3006, and 1.3395 dB at 14 GHz, while at 18 GHz, the values are 1.0684, 1.0620, and 0.6676 dB, and they are 0.6197, 0.7136, and 0.6768 dB at 22 GHz. It can be observed from the previous values that the SDE values are generally decreasing with the frequency of propagation, and again the third-order CI model provides the best performance over the other models. It should be emphasized that the MPE is monotonically decreasing with the order of the CI path loss model, as can be seen from Table 5.3 that represents the LOS MPE and the SDE of the three models at the frequency bands adopted for this study.

Figures 5.12, 5.13, and 5.14 depict the PE of the three models for the NLOS communication scenario at 14, 18, and 22 GHz, respectively. Higher values of the MPE are observed because of the propagation effects that control the wireless signals' behavior in the channel since there is no direct LOS path for the propagated signals to reach the Rx antenna. For example, at 14 GHz, the MPE value of the standard CI model is 4.7924 dB, which is about a 180% increase over the LOS value. In contrast, only a 19.43% increase is observed from the second-order CI model. The third-order CI model reduces the MPE values over the other models by 80.83% and 54.92% over the NLOS values of the standard CI model and the second-order CI model. These achievements of the third-order CI model prove its suitability to predict accurately the path loss for future mobile networks. Table 5.4 provides the MPE and SDE values of the three models at the three frequency bands selected. It should be noted that generally, the MPE and SDE values for the NLOS scenario are very high compared to the LOS values. However, all these values are within an acceptable range in the studied enclosed indoor corridor environment. As a summary of discussing the results obtained, the optimum model in both the LOS and NLOS communication scenarios is the third-order CI model. Design engineers can apply this model to have reliable calculations of the systems' coverage for all the environments and communication scenarios. It is worth noting that the main reason of the higher error values in the NLOS communication scenario compared to the LOS results is the high probability of having a clear LOS path between the transmitting and receiving antennas.

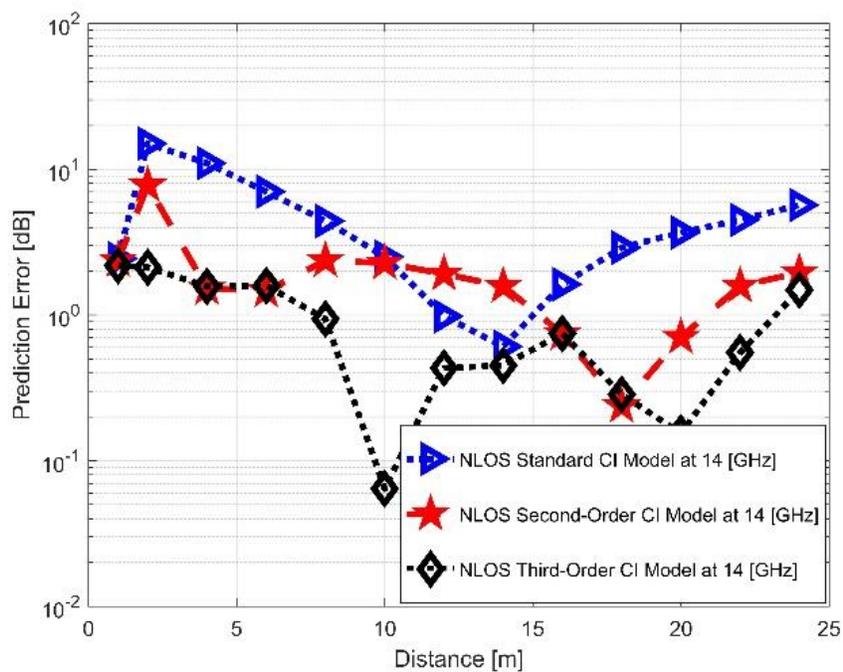


Figure 5.12: NLOS PE of path loss models at 14 GHz.

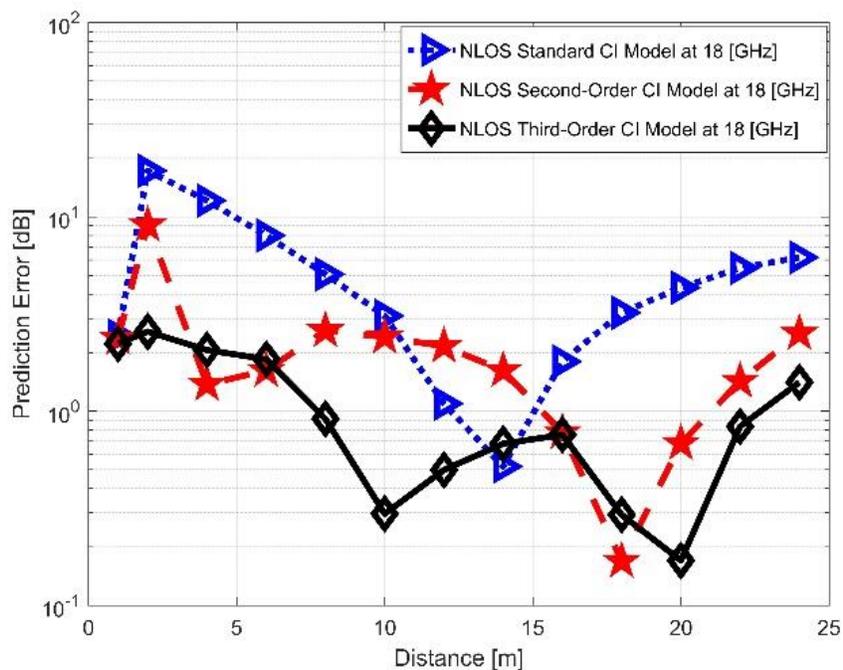


Figure 5.13: NLOS PE of path loss models at 18 GHz.

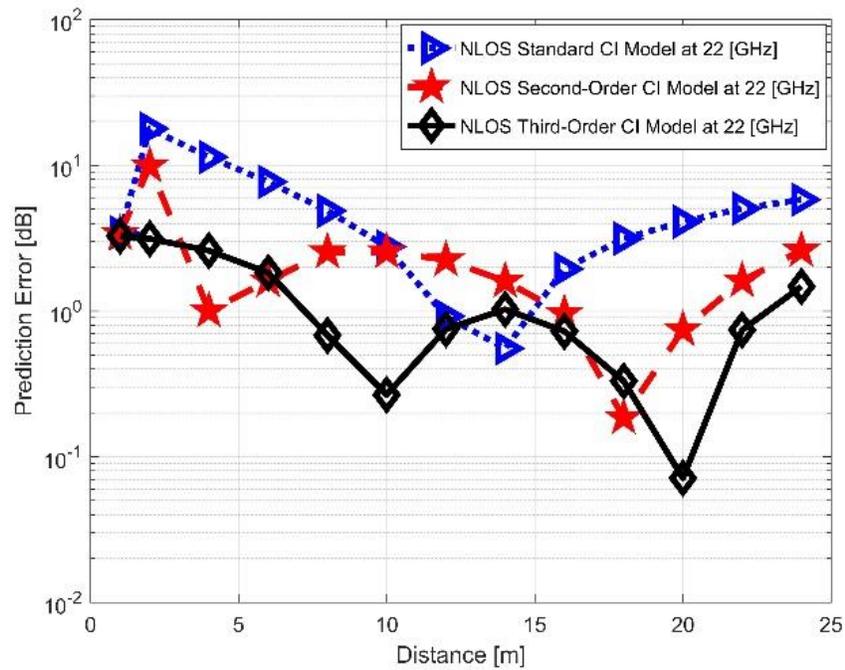


Figure 5.14: NLOS PE of path loss models at 22 GHz.

Table 5.3: The models' MPE and SDE for the LOS communication scenario at 14, 18, and 22 GHz.

Frequency (GHz)	MPE (dB)			SDE (dB)		
	14	18	22	14	18	22
Standard CI model	1.7138	1.1347	1.1119	1.3585	1.0684	0.6197
Second-order CI model	1.7064	1.0604	0.9743	1.3006	1.0620	0.7136
Third-order CI model	1.5506	0.6561	0.8233	1.3395	0.6676	0.6768

*Table 5.4: The models' MPE and SDE for the NLOS communication scenario at 14, 18, and 22 GHz.*

<i>Frequency (GHz)</i>	<i>MPE (dB)</i>			<i>SDE (dB)</i>		
	<b>14</b>	<b>18</b>	<b>22</b>	<b>14</b>	<b>18</b>	<b>22</b>
<i>Standard CI model</i>	4.7924	5.4283	5.3419	3.9831	4.5495	4.5462
<i>Second-order CI model</i>	2.0380	2.2130	2.3760	1.7680	2.1110	2.3302
<i>Third-order CI model</i>	0.9187	1.1187	1.2984	0.7067	0.7805	1.0428

## 5.5 Chapter Summary

This chapter investigated the impact of high-ordering the CI model's accuracy dependency on the Tx-Rx separation distance in the logarithmic scale. Both the LOS and NLOS communication scenarios were considered. The main result of this chapter is that, since there are more parameters that depend on the Tx-Rx distance in the logarithmic scale, higher accuracy is achieved. The optimum number of parameters is three since the third-order CI model fits the data perfectly in both the LOS and NLOS scenarios in all the frequency bands selected. In addition, the model's complexity is not high since other standard models are used with the same number of parameters. Finally, this chapter showed that this principle can be applied to any other path loss prediction model to achieve more precision in predicting the path loss and provide more details in describing the wireless propagation channel. Future research related to this chapter could be directed towards implementing measurements in more environments and scenarios and providing other improvements to acquire reliable channel models that can accommodate various environments and optimize existing models' parameters.

# CHAPTER 6

## COMPARATIVE ANALYSIS OF MAJOR MACHINE-LEARNING-BASED PATH LOSS MODELS

---

## *Chapter 6: Comparative Analysis of Major Machine-Learning-Based Path Loss Models*

### **6.1 Introduction**

Unlimited access to information and data sharing wherever and at any time for anyone and anything is a fundamental component of the 5G wireless communication and beyond. Therefore, as discussed previously, it has become inevitable to exploit the SHF and mmWave frequency bands for future wireless networks due to their attractive ability to provide extremely high data rates because of the availability of vast amounts of bandwidth. However, due to the characteristics and sensitivity of wireless signals to the propagation effects in these frequency bands, more accurate path loss prediction models are vital for planning, evaluating, and optimizing future wireless communication networks. This chapter presents and evaluates the performance of several well-known machine learning methods, including multiple linear regression (MLR), polynomial regression (PR), support vector regression (SVR), as well as the methods using decision trees (DT), random forests (RF), K-nearest neighbors (KNN), artificial neural networks (ANN), and artificial recurrent neural networks (RNN). RNNs are mainly based on long short-term memory (LSTM). The models are compared based on measurement data to provide the best fitting machine-learning-based path loss prediction models. Several performance metrics are adopted for this research. This chapter also proves that these learning methods could be used as accurate and stable models for predicting path loss in the mmWave and higher frequency regime.

### **6.2 Data Preparation and Models Setup**

There are many distinct types of machine learning algorithms, each with its own structure. Our objective is to see if these models can offer reliable prediction results at mmWave frequencies for a specific environment, e.g., a typical indoor corridor that can be viewed as an air-filled rectangular waveguide with huge dimensions compared to the wireless signals' wavelength. To the best of our knowledge, this is the first effort in predicting the path loss at frequency bands higher than 6 GHz for typical indoor corridor environments based on several ML methods. Path loss prediction models with the highest possible accuracy are vital for such environments since the trend is to rely on indoor channels for future wireless networks. Therefore, the primary motivation of this work is to evaluate the feasibility and prediction accuracy of various machine-learning-based models for predicting path loss in indoor corridor environments.

This section provides information on the preparation of the raw data for ML-based path loss prediction. Moreover, the hyperparameters tuning method selected to achieve the best models' performance is presented in this section. Furthermore, the justification of the models' stability is discussed. Finally, this section presents the evaluation metrics adopted to measure and compare the models' performance.

### 6.2.1 Data Preparation

The adopted real measured data were collected in a typical indoor corridor environment, as detailed in chapter 3. The average total number of samples collected from the measurement campaigns is 865 samples, considering all the operating frequencies (i.e., 14, 18, and 22 GHz), Tx–Rx separation distances (2–24 m with an incremental step of 2 m), antenna height values (i.e., 1.6 and 2.3 m), and AoA values (0–360 *degrees* with an incremental step of 10 *degrees*).

The raw datasets collected from the measurements were analyzed and cleaned to provide one reliable path loss value for each Tx–Rx separation distance, frequency, AoA, and Tx antenna height. Accordingly, the best features that lead to the optimum performance for the ML-based models adopted in this work are: (1) The distance between the transmitting and receiving antennas, which is the most crucial input feature that significantly affects the path loss values. (2) The operating frequency to provide multi-frequency path loss prediction models for the frequency range between 14 and 22 GHz. (3) The AoA of the Rx antenna to has LOS and NLOS characterizations of the communication's condition. (4) The Tx antenna height allows for more generalization of the target models.

The datasets were cleaned since this step is a vital part of any aspect of modeling based on ML algorithms. Working with impure datasets can lead to several significant challenges. On the other hand, cleaned and high-quality datasets can cause reliable models to provide outstanding results. Although there are many data cleaning methods, for this research, we adopted the method of removing irrelevant values and taking care of some missing values using the interpolation between the nearest two values. Hence, the datasets are ready for the next step after cleaning and analyzing.

After carefully selecting the input features, normalization of the data was applied for some of the ML methods. This was performed by first computing the mean value for each feature, then subtracting this mean value over the entire dataset feature to centralize the data, and finally, calculating the standard deviation and dividing the subtracted values by the standard deviation. After that, the processed data were applied to each model and divided into training and testing datasets based on a reliable hyperparameter tuning technique and cross-validating the developed model to provide accurate and stable results. More specifically, of the 865 collected samples, 80% of these samples were used for training and cross-validating each model by dividing these datasets 7-fold to evaluate the model's stability, as to be detailed in Section

6.2.3. The other 20% of the datasets were used for validating the models' prediction accuracy. Figure 6.1 depicts the flow chart of the adopted ML-based path loss prediction strategy.

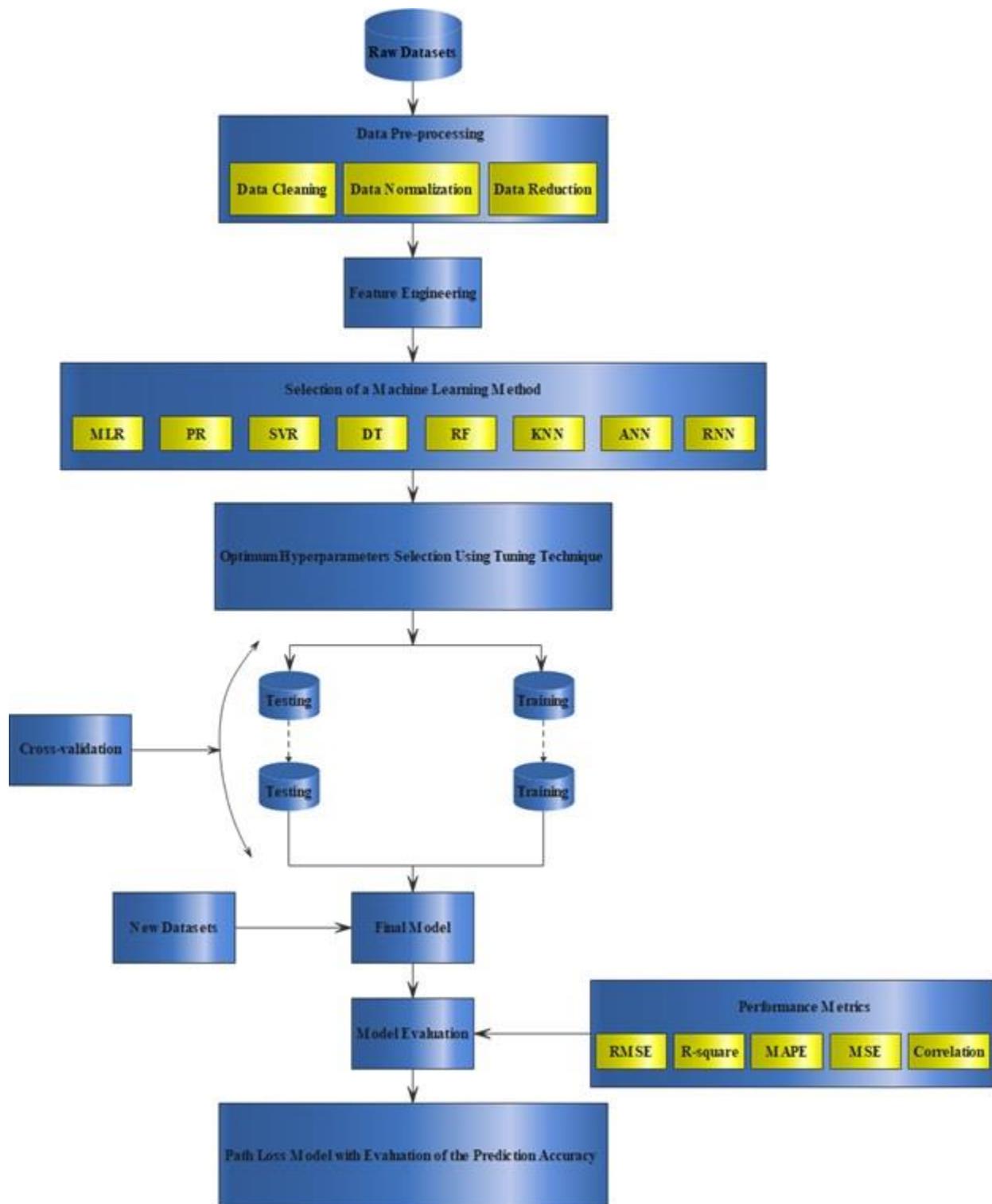


Figure 6.1: Flow chart of the adopted ML-based path loss prediction modeling technique.

## 6.2.2 Hyperparameters Tuning Setup

Hyperparameter tuning techniques play an essential role in searching for the best hyperparameter for machine learning applications. Machine learning algorithms are based on complex hyperparameters that create complicated black boxes and lead to optimization challenges. Moreover, determining the hyperparameters' values can be time-consuming if we try all possible combinations of hyperparameters. However, several techniques have been proposed to select the best hyperparameters for a particular model, including Bayesian Optimization Automate Hyperparameter Tuning (Hyperpot) [180], Spearmint Bayesian optimization [181], Sequential Model-Based Optimization (know as SMAC) [182], Autotune: A derivative-free optimization [183], Google viezier [184], Genetic Algorithm [185], and Optuna Approach [186]. These approaches aim to select the best hyperparameter that minimizes the mean square error and maximizes the accuracy, such as R-square. This is performed by training the machine learning techniques in all the hyperparameter possibilities, then selecting the ones that lead to the objective (i.e., the best performance). The tuning techniques mainly assist in looping through predefined hyperparameters and fitting the estimator (model) on the training set.

Since machine learning techniques require hyperparameter tuning to achieve both the efficiency of the model training process and the resulting model accuracy, the best hyperparameter of machine learning algorithms requires determination. This work has applied the Optuna approach to select the best hyperparameter that leads to suitable data fitting. The Optuna approach is used in our work for multiple reasons. Optuna is a software framework for automating the optimization process of these hyperparameters. It automatically finds optimal hyperparameter values using various samplers such as grid search, random [187], Bayesian, and evolutionary algorithms [186]. It has eager search spaces that use automated searches for the optimal hyperparameter using Python conditionals, loops, and syntax. It is also a state-of-the-art algorithm that efficiently searches large spaces and prunes unpromising trials for faster results. In addition, the Optuna can parallelize hyperparameter searches over multiple threads or processes without modifying code. Finally, the best machine learning techniques depend on identifying the proper hyperparameters, avoiding both overfitting and underfitting, which is the trend of this research to achieve objective prediction accuracy.

## 6.2.3 ML-Based Models' Stability

In this work, the *k-folds* cross-validation technique plays a significant role in assessing the machine learning model stability. The technique first divides the measured data into  $k = k_1, k_2, \dots, k_7$  subsets, named folds. Then it trains the model on the data using  $k_1$  to  $k_6$  of the folds and evaluates the model's performance on the  $k_7$  data. The technique continues repeating the same approach by training the model based on *six-folds* and evaluating the model's performance according to the *seventh (testing) fold*. In each cross-validation

training, the error estimation is averaged over all  $k$  trials to get the total effectiveness of the model. As can be seen, every data point gets to be in a validation set exactly once and in a training set  $k - 1$  times. This significantly reduces the bias as the model uses most of the data for fitting and reduces variance as most of the measured data are also being used in the validation set. The use of cross-validation adds reliable stability to the effectiveness of the machine learning model since the interchange of the measured data is applied between the training and testing sets.

#### 6.2.4 Evaluation Metrics

This subsection presents the performance metrics used to evaluate and compare the performance of the proposed ML-based path loss models. The performance metrics adopted are (1) R-squared (or  $R^2$ ). (2) Root mean squared error (RMSE). (3) Mean absolute percentage error (MAPE). (4) Mean square error (MSE). (5) Correlation (Corr) coefficient. The main reason for choosing these five well-known metrics is to compare our results with the values obtained by other researchers for similar and different environments since most works in the literature adopt these metrics. The performance metrics are expressed as [19]:

$$R^2 = 1 - \frac{\sum_{i=1}^Q (PL_i - \widehat{PL}_i)^2}{\sum_{i=1}^Q (PL_i - \overline{PL})^2}, \quad (6.1)$$

$$RMSE = \sqrt{\frac{1}{Q} \sum_{i=1}^Q (PL_i - \widehat{PL}_i)^2}, \quad (6.2)$$

$$MAPE = \frac{1}{Q} \sum_{i=1}^Q \left| \frac{PL_i - \widehat{PL}_i}{PL_i} \right|, \quad (6.3)$$

$$MSE = \frac{1}{Q} \sum_{i=1}^Q (PL_i - \widehat{PL}_i)^2, \quad (6.4)$$

$$Corr = \frac{\sum_{i=1}^Q (PL_i - \overline{PL})(\widehat{PL}_i - \overline{\widehat{PL}})}{\sqrt{\sum_{i=1}^Q (PL_i - \overline{PL})^2} \sqrt{\sum_{i=1}^Q (\widehat{PL}_i - \overline{\widehat{PL}})^2}}, \quad (6.5)$$

where  $Q$  is the total number of samples used for the calculation of the performance metrics,  $PL$  is the empirical path loss value,  $\widehat{PL}$  is the predicted path loss value, and  $\overline{PL}$  and  $\overline{\widehat{PL}}$  are the mean values of  $PL$  and  $\widehat{PL}$ , respectively.

## 6.3 Machine-Learning-Based Models

This section presents various ML-based models for predicting Path Loss for future indoor wireless communications. Eight prediction models are adopted in this work, namely MLR, PR, DT, RF, SVR, KNN, ANN, and RNN-LSTM. The following subsection will represent these models in detail.

### 6.3.1 Linear Regression Models

Linear Regression Models are essential techniques for addressing the regression challenges in machine learning using data modeling. The linear regression models contain various types of models and depend upon several elements [188]-[191]. These elements incorporate the type of target variable, the shape of the regression line, and the number of independent variables. This work adopted two types of linear regression models to predict the path loss for the selected enclosed indoor environment. These models are MLR and PR. The MLR model is a predictive model that considers more than one input to predict the target. This model identifies the correlation between the various features (dependent variables) and the target (independent variable) to find the best fit for the measured data. The MLR assumes that the inputs are  $X_1, X_2, X_3, \dots, X_n$ , and seeks to predict the target real-value  $Y$ . The MLR model has the form:

$$Y = f(X) = \beta_0 + \sum_{i=1}^n \beta_i X_i + \epsilon, \quad (6.6)$$

where  $\beta_i$  are unknown coefficients (model coefficients) and  $\epsilon$  defines the errors or noise. The most famous estimator that is used to estimate the coefficients model is the Least Squares (LS), in a way that it picks the coefficients  $\beta_0, \beta_1, \beta_2, \dots, \beta_n$  to minimize the mean squares error (MSE) presented in the following equations:

$$MSE = \sum_{j=1}^M (Y_j - \beta_0 - f(X_j))^2, \quad (6.7)$$

$$= \sum_{j=1}^M (Y_j - \beta_0 - \sum_{i=1}^N X_{ji} \beta_j)^2. \quad (6.8)$$

However, if basic expansion is made to Eq. (6.6) by substituting  $X_2 = X_1^2, X_3 = X_1^3, \dots$ , this leads to a new form of model called *polynomial regression* representation. A polynomial regression model can be defined as a new function that takes the form:

$$Y = f(X) = \gamma_0 + \sum_{n=1}^N \gamma_n X^n + \epsilon, \quad (6.9)$$

where  $n$  is a polynomial degree of the PR model and  $\gamma_n$  represents the model coefficients of PR, where the LS method estimation is applied to estimate the PR model coefficient using the measured data. The polynomial regression model attempts to generate a polynomial function that estimates the measured data

points. It determines the best-fit curve that passes through the entire measured data to minimize the predicted error.

The PR model is a modified version of the MLR model where the relationship between the independent and dependent variables is defined by the  $n$ -th degree. The best fit curve in polynomial regression passes through all the data points, depending on the power of  $X$  (or the value of  $n$ ). It is recommended to analyze the turn towards the end as the higher polynomials can give undesired results.

The number of iterations adopted using the Optuna technique was 100 for both the MLR and PR methods. Moreover, the number of degrees for the PR model was 6. Finally, it is worth noting that the datasets were normalized before training and testing these models to achieve the best performance.

### 6.3.2 Support Vector Regression Model

Vapnik [192] proposed SVM algorithms for the binary classification problem. Later on, they worked on both multiclassification and regression problems, known as support vector classification (SVC) and SVR algorithms. The SVR applies a similar concept as the SVC algorithm with some changes. A few changes include that the target values are real numbers, the infinite possibility of which became challenging to predict using the same SVC. However, the SVR selects a boundary distance  $\{-\epsilon, \epsilon\}$  from the original hyperplane to predict the real numbers. This boundary distance is the margin of tolerance that takes only data points within this boundary. Therefore, the main goal is always similar: minimizing the prediction error and individualizing the hyperplane to maximize the margin [193].

As aforementioned, the SVR applies a similar concept to SVC, but a target variable is a real number  $Y \in R$ . As stated by Huang and Tsai [194] and Patel *et al.* [195], the SVR seeks the linear regression function as an alternative to finding the hyperplane in the SVC by Eq. (6.10). This can be achieved by selecting a threshold error  $\epsilon$ , which is defined to minimize the expression in Eq. (6.11). This expression is called the  $\epsilon$ -insensitivity loss error function. The SVR regression process, therefore, seeks to minimize  $\epsilon$  in Eq. (6.11) and  $\|W\|^2$  in the expression of  $R$ . The target value of the SVR method is given by:

$$Y = W^T X + b, \quad (6.10)$$

where  $Y$  is the target,  $W$  is the coefficient,  $X$  is the input feature, and  $b$  is a constant. We define:

$$|\hat{Y} - Y|_{\epsilon} = \begin{cases} \text{Zero}, & |\hat{Y} - Y| \leq \epsilon \\ |\hat{Y} - Y| - \epsilon, & \text{otherwise} \end{cases} \quad (6.11)$$

$$R = \frac{1}{\epsilon} \|W\|^2 + C \left( \sum_{i=1}^N |\hat{Y}_i - Y_i|_{\epsilon} \right). \quad (6.12)$$

Tolerance variables are also introduced, defined as the value in excess of  $\epsilon$  and  $\xi$  to limit the value to the regression target. Thus, the minimization of Eq. (6.12) is updated to Eq. (6.13), under the conditions of Equations (6.14) and (6.15) for  $\xi_i$  and  $\xi_i^* \leq 0$  and  $i = 1, 2, 3, \dots, N$ . That is:

$$R = \frac{1}{\epsilon} \|W\|^2 + C \left( \sum_{i=1}^N |\xi_i - \xi_i^*|_{\epsilon} \right), \quad (6.13)$$

$$(W^T X_i + b) - Y_i \leq \epsilon - \xi_i, \quad (6.14)$$

$$Y_i - (W^T X_i + b) \leq \epsilon + \xi_i^*. \quad (6.15)$$

The standard kernel functions are considered in this study, given explicitly by the linear, radial, and polynomial functions in Equations (6.16)–(6.18), respectively.

$$K(X_i, X_j) = X_i^T X_j, \quad (6.16)$$

$$K(X_i, X_j) = e^{-\lambda \|X_i - X_j\|}, \lambda > 0, \quad (6.17)$$

$$K(X_i, X_j) = (X_i^T X_j + 1)^d. \quad (6.18)$$

Note that the shape of the kernel function directly influences the values obtained by the SVR regression. Similarly, the constant  $C$  in Eq. (6.12) and the parameters  $\lambda$  and  $d$  in Equations (6.17) and (6.18) should be optimized. For this purpose, the Optuna technique is applied to choose the optimal parameters for  $C$ ,  $\lambda$ , and  $d$ , considering the lowest RMSE. According to our developed SVR model, the values of the primary hyperparameters obtained from the Optuna hyperparameter tuning technique are: (1) The kernel adopted was the radial basis function (RBF). (2) The kernel coefficient for RBF equals 0.001. (3)  $C = 995.2783$ . Figure 6.2 shows the principle of using the SVR in two dimensions.

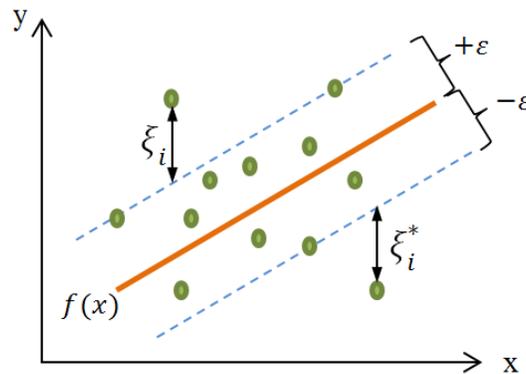


Figure 6.2: The principle of the SVR in two dimensions.

### 6.3.3 Decision Tree Regression Model

Decision tree (DT) learning plays a critical role in solving classification and regression problems. The classification and regression accuracy for its performance, when compared to existing techniques, is sufficient. The classification model learned through these techniques is represented as a tree and is called a decision tree. ID3Q [196], C4.5Q [197], and CART [198] are decision tree learning algorithms. More details can be found in [199].

The proposed decision tree creates a regression model that uses the tree structure form. It breaks down the measured dataset into small subsets while the corresponding decision tree is progressively developed, with the final output being a tree with decision nodes and leaf nodes. The decision node (an input) contains feature branches (e.g., Tx–Rx distance, AoA, Tx Height, frequency), each representing values for the attribute tested. The leaf node (e.g., the path loss value) represents a decision on the numerical target. The top-most decision node in a tree that associates with the best predictor is known as the root node. The DT regression model identifies ways to split the measured data via an algorithmic approach into smaller subsets. This approach is repeated several times until the best results are obtained. The optimum rules that lead to the best results are obtained by using variance reduction as a measure of impurity. These results are used to calculate the variance reduction for each output. A higher variance leads to a higher impurity, meaning that the corresponding conditions should be chosen as the optimum conditions for the model. Based on our model, the selected hyperparameters' values are the MAE as a variance calculation function, and the tree has 69 nodes.

### 6.3.4 Random Forest Regression Model

Random forest (RF) is one of the learning algorithms that uses the tree as a base learner. The RF is introduced since a single regressor is not enough to predict the correct fit. The reason is that, based on sample data, the regressor cannot distinguish between noise and pattern, so it performs sampling with changes such that the given  $n$  trees to be learned are based on the dataset samples after taking the averages. The RF model sets up several trees to address the regression challenge, where each tree contains a root node, leaf nodes, and internal nodes. The root node has a set of training samples, and leaf nodes correspond to the final result. The internal nodes are split by features, and the criterion used to obtain features or split nodes is the MSE. Moreover, in the proposed setup, each tree is learned using four features selected randomly. After creating  $n$  trees, when the testing data is used, the decision regarding the majority of trees that come up is considered the final output. The number of trees adopted for this work in our RF model was 1090, with a maximum of 17 tree nodes.

### 6.3.5 K-Nearest Neighbor Regression Model

The K-nearest neighbors algorithm is a supervised machine learning technique used to solve regression and classification problems. It is simple to understand the concept of the KNN algorithm and its application. However, it has a significant drawback of becoming significantly slow if the data size in use is increased. The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other. The KNN algorithm works by calculating the distances between a query and all the points in the data, choosing a particular number of neighbors ( $K$ ) closest to the query, and then voting for the most frequent label (in the case of classification) or averaging the labels (in the case of regression). Selecting the correct number of neighbors ( $K$ ) can lead to the best fit in the case of regression or classification, which can be performed by applying various  $K$ 's and selecting the one that gives the best results. For this hyperparameter, the best fit obtained was  $K = 2$ , and the distance adopted was based on *Minkowski's* formula. This was achieved by using the Optuna technique using different numbers and selecting the one that best fitted the measured data. Figure 6.3 represents how the KNN method predicts a query using two neighbors.

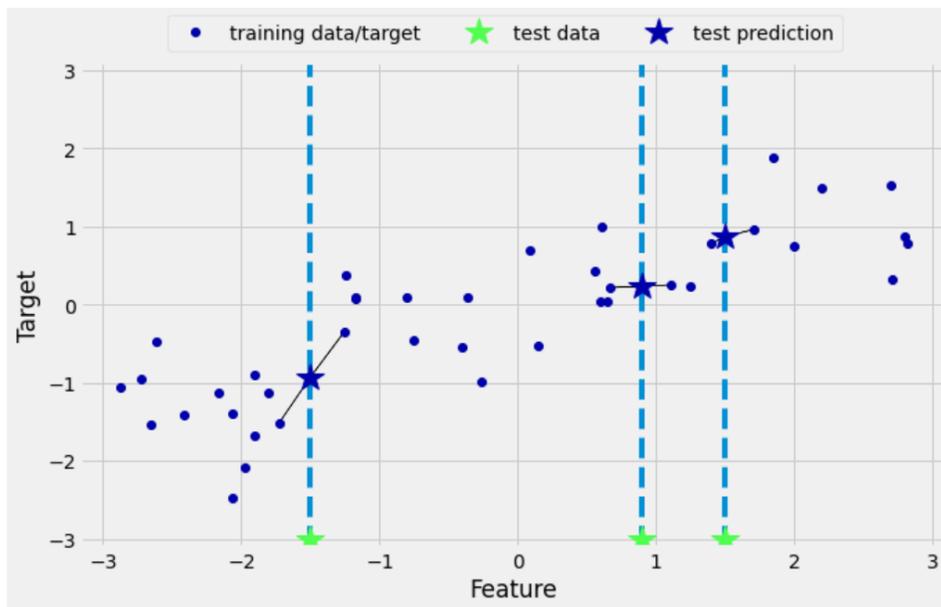


Figure 6.3: The KNN regression model with  $K = 2$ .

The KNN calculation uses the average of the numerical target of the  $K$  nearest neighbors, applying one of the following distance functions:

$$\text{Euclidean: } D = \sqrt{\sum_{i=1}^K (X_i - Y_i)^2}, \quad (6.19)$$

$$\text{Manhattan: } D = \sum_{i=1}^K |X_i - Y_i|, \quad (6.20)$$

$$\text{Minkowski: } D = \left( \sqrt[q]{\sum_{i=1}^K (|X_i - Y_i|^q)} \right)^{\frac{1}{q}}, \quad q \geq 1, \quad (6.21)$$

where  $X$  and  $Y$  are the original path loss and the predicted path loss, respectively.

### 6.3.6 Artificial Neural Network Model

Artificial neural networks have been developed based on biological neural network functionality. The ANNs are a network that contains a group of neurons, various layers, and activation functions, all of which get activated based on inputs. The proposed model is based on ANN architecture that includes four hidden layers of feed-forward neural networks. The first two hidden layers contain 96 neurons, and the last two hidden layers have 32 neurons, with each hidden layer followed by the well-known *ReLU* activation function. The input for the network accepts four features from the preprocessed data. The output layer is a single neuron with a linear activation function as the transfer function that leads to the predicted value. Figure 6.4 depicts the architecture of our proposed ANN model. The output result is a real value representing the path loss. The hyperparameter values of the proposed ANN (number of layers, number of neurons, activation functions, learning rate = 0.001) are obtained using the Optuna technique, which gives the best hyperparameters that fit the measured data.

The proposed ANN model learns to move up or down depending on the trend feature extraction from the data, giving the fit curve. The parameter weight at every epoch is adjusted using the gradient descent with momentum to reach the global minimum error. The proposed model uses a comprehensive hyperparameter setup to identify the best weights of the parameters for the path loss prediction.

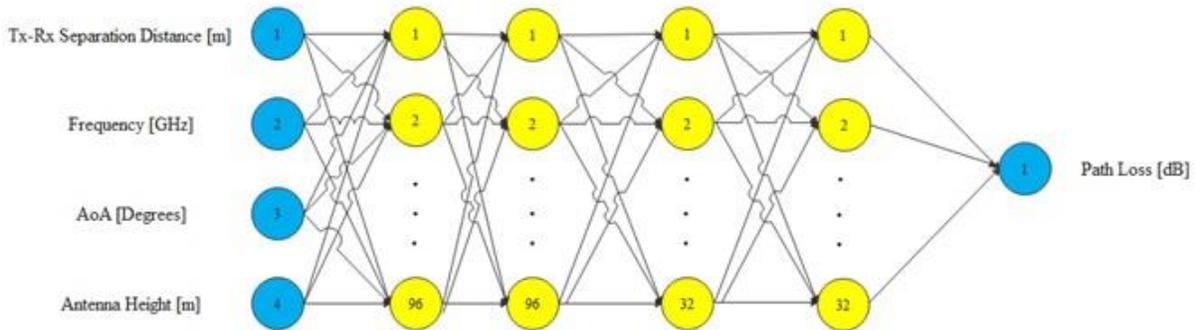


Figure 6.4: The architecture of the proposed ANN model.

### 6.3.7 Recurrent Neural Network Model

The recurrent neural network based on the long-short-term memory layer is a type of ANN where the links between nodes form a directed graph along a temporal sequence. This makes it exhibit temporal dynamic behavior. Derived from feed-forward neural networks, RNNs can use their internal state (memory) to process variable-length sequences of inputs [200], which makes them applicable to tasks such as unsegmented, connected handwriting recognition, or speech recognition. While the ANN output layers depend on the previous layer output to train the neural network, the RNN requires both the previous layer output and the internal state of the neural network. The internal state is defined as the output of each hidden neuron when processing the previous input observations. They are thus well-suited to process time series of data and capture their time dependencies. On the other hand, it considers the current input and the output that it has learned from the previous input for making a decision. The proposed RNN architecture is expected to extract feature representations that encode some aspects of the path loss. This new way of learning gives the RNN-LSTM model significant performance on several applications, which motivated us to select the RNN-LSTM for our work. The architecture of the proposed model is the same as the original version of RNN-LSTM; the only changes are on the hyperparameters setup.

The proposed RNN-LSTM model involves the hyperparameters of two hidden layers of internal state (LSTM) and one feed-forward neural network layer. The first and last LSTM hidden layers have 128 and 32 neurons, respectively. The feed-forward neural network layer contains 48 neurons. Each hidden layer is followed by the *ReLU* activation function. The model uses a learning rate of 0.0001 to train the model. The input for the network accepts four features from the preprocessed data. The output layer is a single neuron with a linear activation function as the transfer function. The output result is a real value representing the path loss. The cost function is used to optimize the mean square error. Figure 6.5 shows the architecture of the proposed RNN-LSTM model. It should be noted that the RNN-LSTM model is more complex compared to the ANN model and has more parameters to train [201], [202].

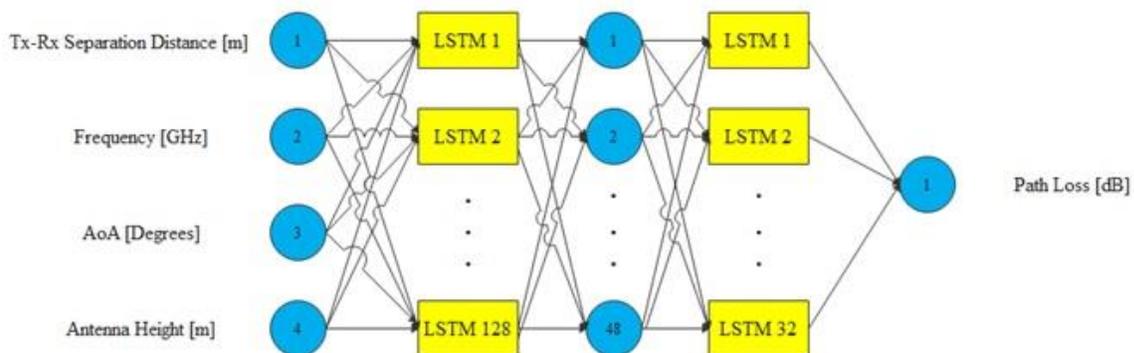
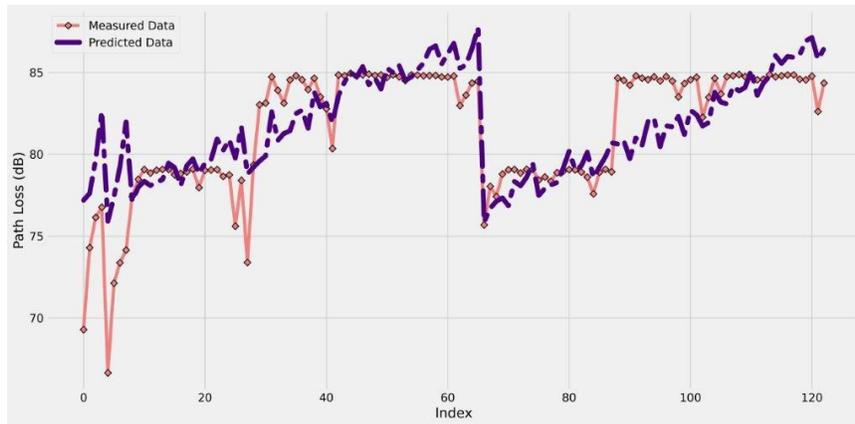


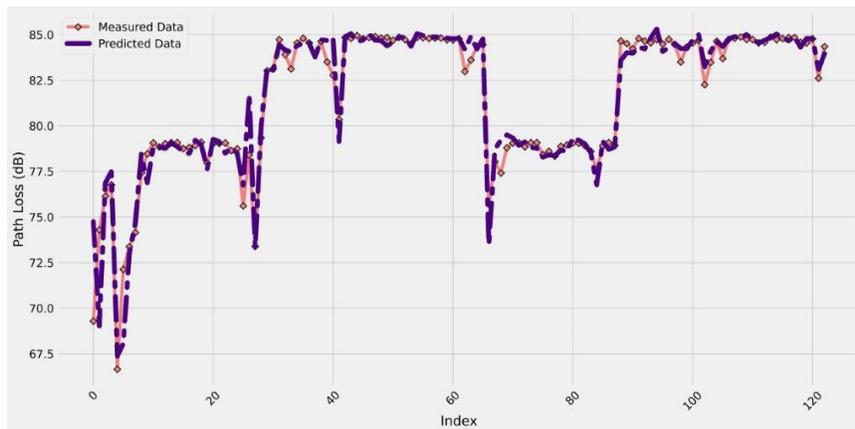
Figure 6.5: The architecture of the proposed RNN-LSTM model.

## 6.4 Results and Discussions of the Machine-Learning-Based Models

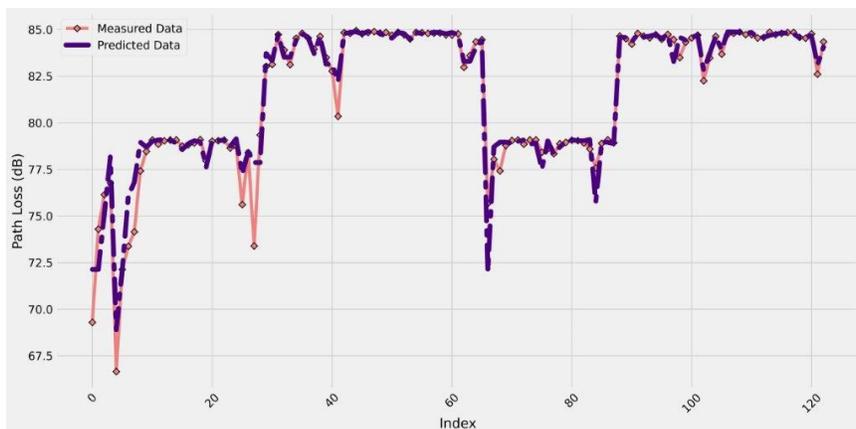
Figures 6.6-6.13 depict the measured data and ML-based path loss prediction models for the enclosed indoor wireless channel selected in this work. These results are for 20% of the measurement data used for testing the performance of the models. As described above, the input features of all the ML-based models adopted for this study are the Tx–Rx distance, operating frequency, AoA, and the Tx antenna height. From Figures 6.6-6.13, all the models (except the MLR) fit the real measurement data accurately. Furthermore, it is clear from the figures that there is a significant match between the measurement data and the predicted models, which means a high prediction accuracy was provided by these models. Based on numerical analysis, the R-squared values of all the models fall between 0.4704 and 0.9798, while the RMSE values are in the range of 0.0216 to 2.9008 dB.



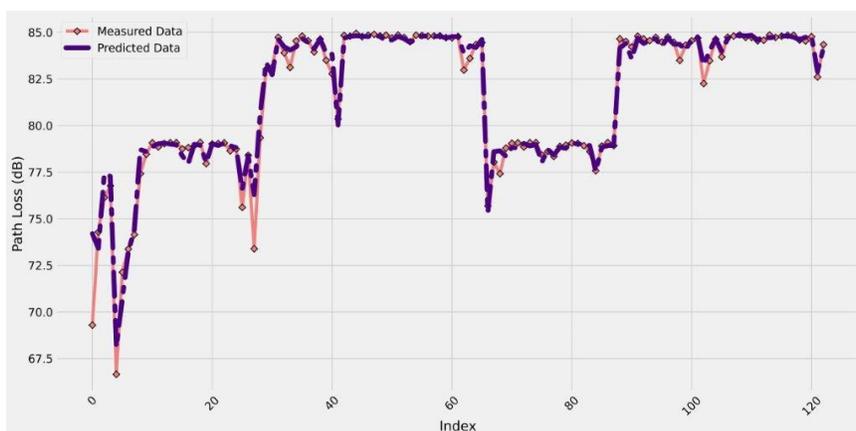
*Figure 6.6: Measured and predicted path loss data for the MLR model.*



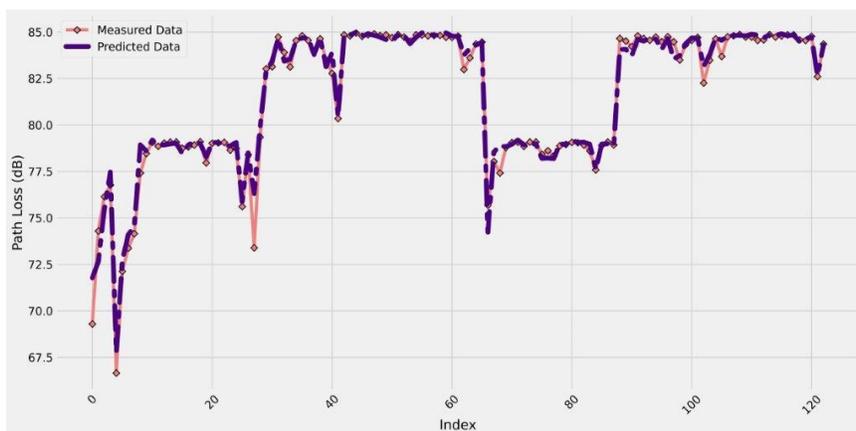
*Figure 6.7: Measured and predicted path loss data for the PR model.*



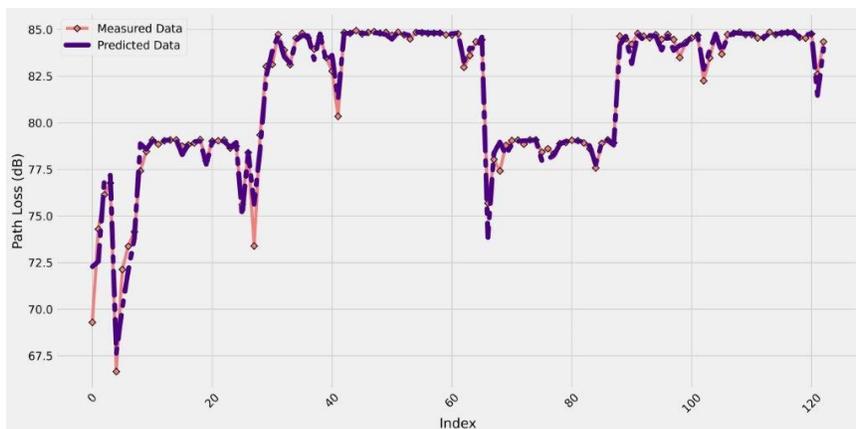
*Figure 6.8: Measured and predicted path loss data for the DT model.*



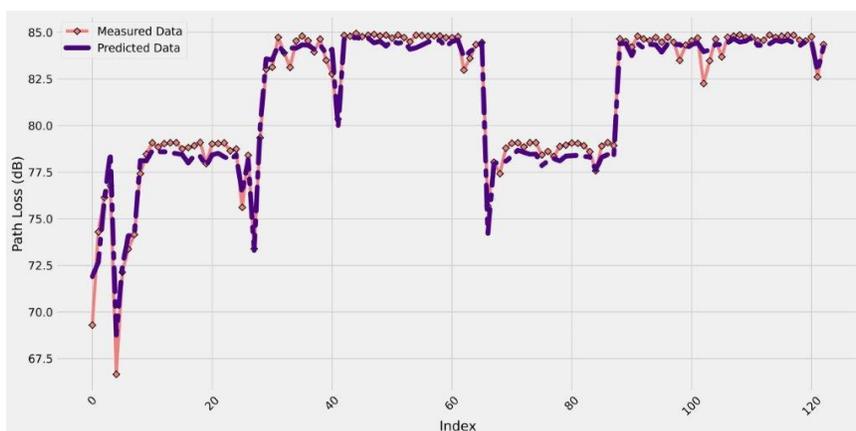
*Figure 6.9: Measured and predicted path loss data for the RF model.*



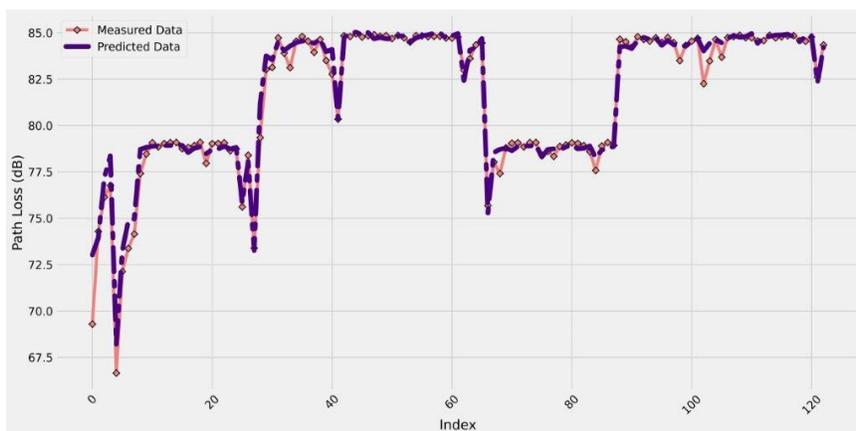
*Figure 6.10: Measured and predicted path loss data for the SVR model.*



*Figure 6.11: Measured and predicted path loss data for the KNN model.*



*Figure 6.12: Measured and predicted path loss data for the ANN model.*



*Figure 6.13: Measured and predicted path loss data for the RNN-LSTM model.*

Moreover, the MAPE values are between 0.37% and 6.94%, and for the correlation factor, the values are in the range of 69.45% to 99.07%. Thus, all the previous metrics results show the quality of the model's predictions for such environments. The reasons behind the accuracy of these models are: (1) The availability of training data since 80% of the measurement data was used to train the models. (2) Efficient input feature selection that considers crucial factors, such as the AoA and the Tx antenna's height for these indoor environments, in addition to essential factors, such as the Tx–Rx separation distance and the multi-frequency operating range of 14 to 22 GHz. (3) The use of the hyperparameter tuning technique, namely Optuna, to choose the best values of the hyperparameters (for example, type of activation function, number of layers, number of neurons, learning rate, number of trees) instead of choosing them manually, which leads to a minimum prediction error. (4) The preprocessing of the data for some models, such as the ANN and RNN-LSTM.

As a comparison between the models, the ANN model provides the best average RMSE value, while the worst is for the MLR model. However, the RNN-LSTM, KNN, RF, DT, SVR, and PR show their ability to predict path loss since their average RMSE values were less than 1.1653 dB. This can be validated from the R-squared and the correlation coefficient values that show a minimum of 0.8690 and 0.9342, respectively, close to the ideal value of 1. Table 6.1 provides the performance metric values of the selected ML-based models. It is worth noting that the table provides three values of each metric, the minimum, average, and maximum value. These values came from the cross-validation technique adopted for this study that divided the measurement data into seven folds to evaluate the stability of each model. It can be seen from the table that the models provide highly stable results since the deviation of each metric from its average to maximum or minimum is small. Furthermore, the results displayed in Table 6.1 are based on all values of the frequency bands selected for this research (i.e., our ML path loss models are multi-frequency), which means that these models have the ability to accurately predict the propagation loss at the adopted frequency regime.

*Table 6.1: Performance metrics' values of all the ML-based models selected.*

Models	R <sup>2</sup>			RMSE			MAPE			MSE			Corr		
	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
MLR	0.4704	0.5220	0.5711	2.5191	2.7713	2.9008	0.0220	0.0239	0.0252	6.3461	7.6943	8.4144	0.6945	0.7274	0.7573
PMR	0.8743	0.9286	0.9556	0.8721	1.0480	1.2856	0.0059	0.0072	0.0087	0.7605	1.1248	1.6528	0.9550	0.9679	0.9781
SVR	0.9054	0.9414	0.9798	0.5369	0.9252	1.1586	0.0037	0.0054	0.0064	0.2882	0.8952	1.3423	0.9520	0.9731	0.9907
DT	0.8690	0.9125	0.9473	0.8678	1.1653	1.5263	0.0052	0.0068	0.0085	0.7531	1.3988	2.3296	0.9342	0.9606	0.9743
RF	0.9189	0.9461	0.9688	0.6680	0.9145	1.2534	0.0041	0.0056	0.0070	0.4463	0.8688	1.5709	0.9669	0.9761	0.9853
KNN	0.9068	0.9443	0.9756	0.5909	0.9175	1.2163	0.0040	0.0052	0.0058	0.3492	0.8787	1.4794	0.9524	0.9747	0.9877
ANN	0.9017	0.9352	0.9755	0.0270	0.0355	0.0435	0.0220	0.0312	0.0587	0.0007	0.0013	0.0019	0.9530	0.9732	0.9878
RNN-LSTM	0.8889	0.9160	0.9762	0.0216	0.0418	0.0592	0.0161	0.0390	0.0694	0.0005	0.0019	0.0035	0.9433	0.9666	0.9882

Figure 6.14 represents the test set prediction error of the ML-based models used in this work. In addition, the figure provides insights into the range of the difference between the measured and ML-predicted path loss models. These curves observe random distributions with average error values around zero with error impulses of 6 dB, as shown in the worst case of the MLR model. This means that the predicted ML-based path loss value has a maximum of only a 6 dB difference from the real measured path loss value. Moreover, the predicted and actual path loss values are given in Figure 6.15. Again, the figure proves what we have discussed in this section: *all the models have a significant prediction performance with the best accuracy provided by all the models since there are clear straight lines, except the MLR model with an inaccurate performance.*

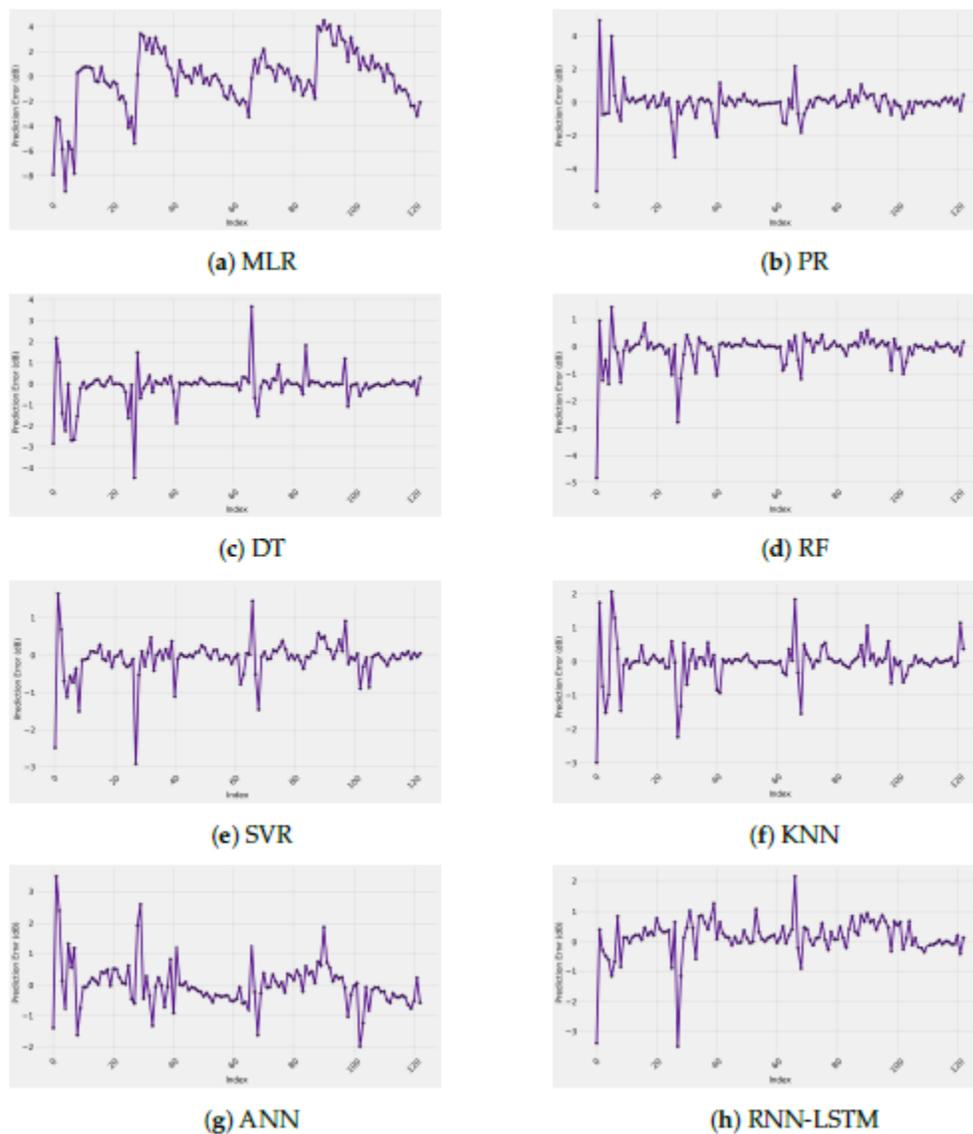
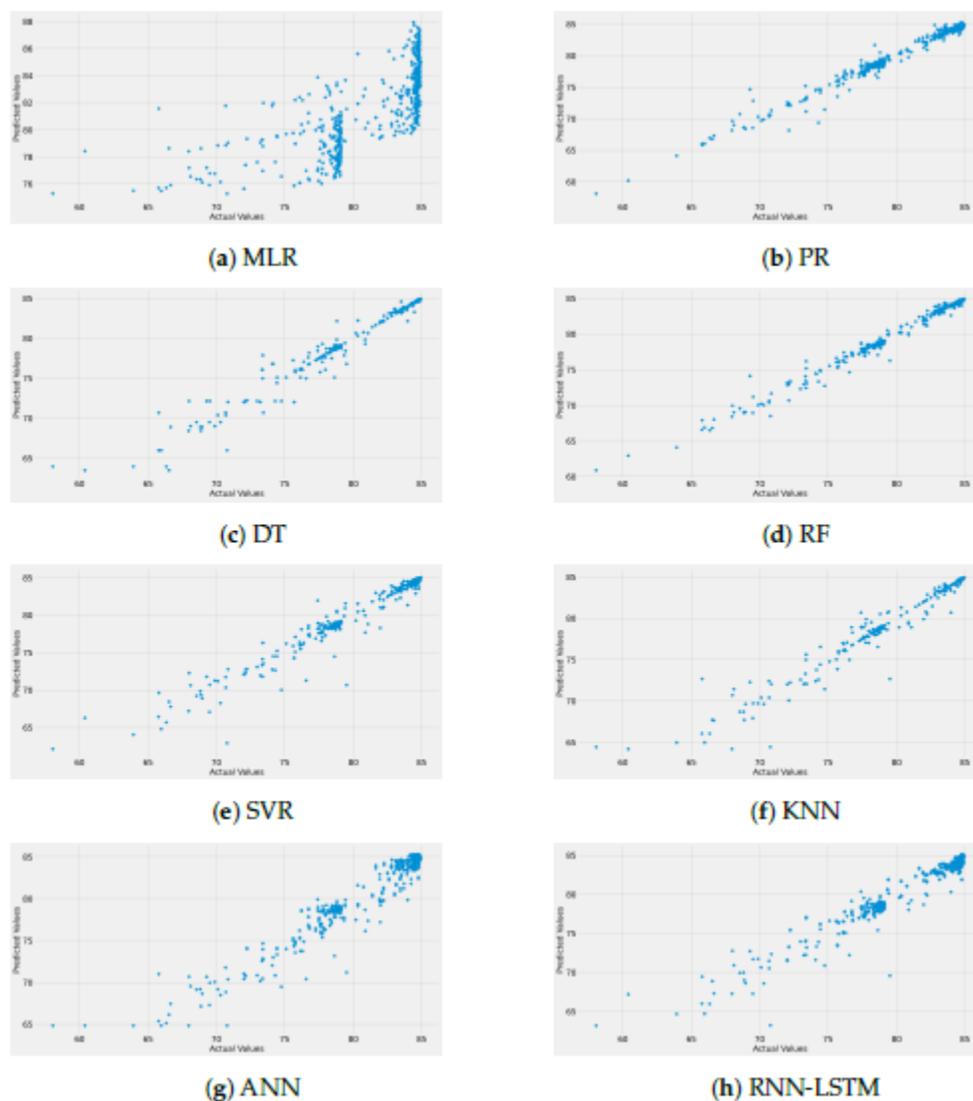
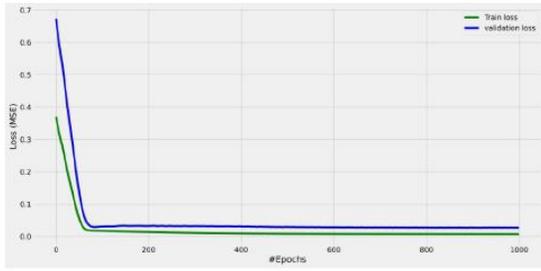


Figure 6.14: Prediction error of each ML-based model.

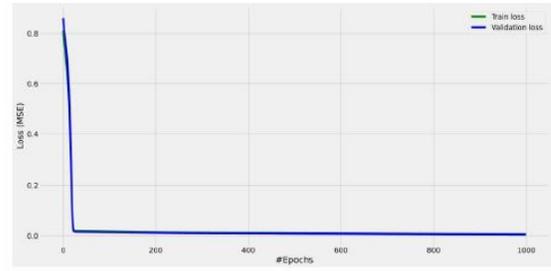
The validation and training loss in Figure 6.16 shows the excellent fitting of the ANN and RNN models since the validation curve is slightly higher than the testing curve in the case of the ANN model, while the validation curve matches the testing curve across all the epoch values in the case of the RNN-LMTS model. The results reveal that the structure of these neural networks provides high precision in fitting the measurement data without underfitting or overfitting issues. Figure 6.17 depicts the measured and predicted path loss and the prediction error of all the ML-based models selected in this work.



*Figure 6.15: Measured vs. predicted path loss data for each model.*

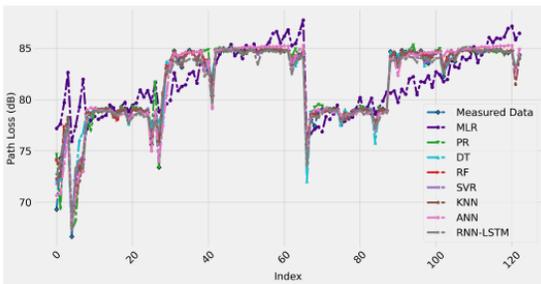


(a) ANN

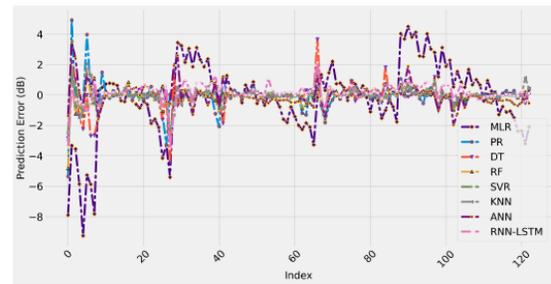


(b) RNN-LSTM

**Figure 6.16: Training and validation loss for both ANN and RNN models.**



(a)



(b)

**Figure 6.17: Measured and predicted path loss and the prediction error of all the ML-based models. (a) Measured and predicted path loss; (b) Prediction error.**

Runtime analyses of ML models are essential for understanding the complexity of machine learning algorithms. It is crucial for algorithm selection in specific tasks and vital for successful implementation. Therefore, it is always a good practice to do runtime analysis and comprehend the complexity of ML algorithms. Runtime analysis can be seen from two directions: time complexity and space complexity. Time complexity measures how fast or slow a model performs the task, while space complexity is the amount of memory required to execute the task. In this work, time complexity analysis is performed to study the comparison between the adopted ML models. Table 6.2 represents the runtime of each model where some models take less than a second to complete the task, such as MLR, PL, DT, and KNN, while others take more time to finish the task. More specifically, as depicted in Table 6.2, the minimum runtime was achieved by the MLR model, 21.3 *ms*, which means that the MLR model has the least complexity among the other models. Nevertheless, the MLR model has the worst performance according to the results presented in Table 6.1. The highest runtime was observed from the run of the RNN-LSTM model since it took almost 125 s to train the model. In general, the runtime obtained by our ML models is comparatively short, which indicates that the adopted models can fit the path loss problems with relatively low complexity. However, the runtime varies according to the performance of the computer used.

*Table 6.2: Runtime comparison of the adopted ML models.*

Model	RunTime (seconds)
MLR	0.0213
PR	0.5251
SVR	3.8647
DT	0.1379
RF	11.2477
KNN	0.0782
ANN	59.1485
RNN-LSTM	124.9184

The experimental platform is on a PC with an Intel Core i7 processor, Gen (10) 1.20 GHz cores, 64-bit operating system, and an x64-based processor. It also has 1 TB shared memory and 16.0 GB RAM. The software used for the model implementation includes Python Version 3.5.2, Tensor flow backend 1.1.0, and TFlearn 0.3. The adopted algorithms have been used from two build-in python libraries called Scikit-learn (Sklearn) and Tensorflow (TF). These are the most useful and robust libraries for machine learning in Python. It provides efficient tools for machine learning modeling, such as classification, regression, and clustering. Training and testing times for the results are provided in Table 6.2 for each model.

To investigate the impact of choosing the antenna height as an input feature, we removed the Tx antenna height column from the datasets' input features. Table 6.3 provides the performance metrics' values when the adopted ML-based models have three input features: the Tx–Rx separation distance, operating frequency, and the AoA. It is clear from the results that the overall performance of the models became worse than when the models had all four input features. For example, the average R-squared and the correlation coefficient values are reduced by approximately 3% and 4%, respectively, while the average RMSE value was increased by 1.6942 dB. For the RNN-LSTM model, the average RMSE was also increased by 1.3731 dB. The results presented in Table 6.3 indicate the importance of considering the Tx antenna as an input feature.

Furthermore, the impact of only having the separation distance and frequency as the two input features (removing the Tx height and the AoA) is investigated. Table 6.4 shows the values of the performance metrics after removing both the Tx height and the AoA from the input features. Again, the accuracy of the models became worse compared to the results provided in Tables 6.1 and 6.3. For instance, the average R-squared value and the correlation coefficient reductions are approximately 9.3% and 5.2%, respectively. In addition, the average RMSE value was increased by 1.9673 dB for the ANN model, while for the RNN-LSTM model, the matter was increased by 1.9808 dB. Again, the results reveal the effectiveness of having the four input features together.

*Table 6.3: Performance metrics after removing the antenna height from the input features of the models.*

Models	R <sup>2</sup>			RMSE			MAPE			MSE			Corr		
	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
MLR	0.4072	0.5227	0.5903	4.4745	5.0024	5.2940	0.5178	0.5409	0.6281	20.0212	25.024	28.0272	0.6091	0.6998	0.7703
PMR	0.6479	0.7324	0.8261	3.2093	3.3903	3.8639	0.4212	0.324	0.5472	10.2993	11.4943	14.9299	0.6503	0.7477	0.8451
SVR	0.8022	0.8432	0.8792	2.3015	2.6190	2.7020	0.0936	0.0753	0.1063	5.2969	6.8592	7.3010	0.8502	0.8734	0.8904
DT	0.7882	0.8161	0.8577	2.1743	2.2942	2.6169	0.0851	0.1066	0.1380	4.7276	5.2635	6.8480	0.8300	0.8606	0.8789
RF	0.8837	0.9048	0.9384	1.4678	1.9786	2.2862	0.0825	0.0946	0.1064	2.1545	3.9147	5.2266	0.8785	0.9028	0.9451
KNN	0.8636	0.9111	0.9532	1.5418	1.7558	2.2647	0.0742	0.0857	0.0969	2.3772	3.0829	5.1288	0.9048	0.9281	0.9567
ANN	0.8528	0.9093	0.9374	1.4144	1.7297	1.9756	0.0797	0.0801	0.0894	2.0006	2.9918	3.9030	0.9187	0.9382	0.9685
LSTM	0.7537	0.8509	0.8971	1.4146	1.4149	1.4160	0.0793	0.0893	0.0969	2.0011	2.0021	2.0051	0.8845	0.9062	0.9441

*Table 6.4: Performance metrics after removing the antenna height and the AoA from the input features.*

Models	R <sup>2</sup>			RMSE			MAPE			MSE			Corr		
	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
MLR	0.3580	0.4367	0.5379	4.9492	5.4129	5.9942	0.6272	0.7240	0.8212	24.4943	29.2993	35.9299	0.4983	0.5658	0.6692
PMR	0.5707	0.6344	0.7385	3.3198	4.0033	4.0034	0.5978	0.6429	0.7181	11.0212	13.024	16.0272	0.6098	0.6343	0.7044
SVR	0.7562	0.7732	0.7943	2.9204	3.2991	3.5713	0.1097	0.1443	0.1720	8.5288	10.8839	12.7539	0.7970	0.8264	0.8650
DT	0.6648	0.7298	0.7696	2.5871	2.9862	3.2512	0.1198	0.1443	0.1697	6.6935	8.9175	10.5706	0.6633	0.7250	0.8088
RF	0.8211	0.8512	0.9092	1.8355	2.1076	2.6163	0.1064	0.1101	0.1345	3.3690	4.4421	6.8452	0.8447	0.8780	0.9026
KNN	0.7937	0.8455	0.8832	2.1794	2.2420	2.8016	0.1009	0.1272	0.1324	4.7498	5.0264	7.8486	0.8419	0.8814	0.9150
ANN	0.8089	0.8482	0.9052	1.9763	2.0028	2.4527	0.0988	0.1049	0.1364	3.9056	4.0114	6.0158	0.8973	0.9224	0.9306
LSTM	0.7456	0.8240	0.8976	1.7334	2.0226	2.4280	0.0839	0.0998	0.1291	3.0049	4.0910	5.8954	0.8686	0.9278	0.9499

Based on the results provided, it is clear that the careful selection of the input features for training the models can ensure high prediction accuracy without considering more features, which leads to less complexity in using the proposed models for such environments. However, there is a high demand for developing precise mechanisms capable of generalizing the results for data far different (representing other communication environments) from the data used for training the models to overcome the need to conduct extensive measurement campaigns in all the possible communication scenarios. Furthermore, these generalizing mechanisms will allow for adaptive data-driven prediction models that accurately represent the channel characteristics for future mobile networks relying on fewer training samples, which leads to faster and more cost-effective planning of the wireless systems. As a suggestion for expanding the datasets, our improved CI and FI models can be used in this manner since they provide higher precision than the standard models in terms of fitting the real measurement data for path loss prediction.

## 6.5 Chapter Summary

Due to the characteristics of the wireless signals at the SHF and mmWave frequency bands compared with today's sub-6 GHz frequency regime, providing accurate and stable path loss prediction models is a challenging problem. Motivated by that, this chapter considered an extensive comparative analysis to evaluate the performance of the most widespread and used machine learning methods, namely the MLR, PR, RF, DT, SVR, KNN, ANN, and RNN-LSTM. The input features used to train the models were carefully selected: the Tx–Rx separation distance, frequency, Tx height, and the AoA. To ensure reliable and stable

results, we used a cross-validation technique to divide the data into seven folds of training and testing datasets and provide the minimum, average, and maximum results. Also, we utilized a hyperparameters tuning method to select the optimum hyperparameters of the model and avoid the time consumption of the manual selection. Furthermore, five performance metrics were applied to evaluate the models: R-squared, RMSE, MAPE, MSE, and the correlations coefficient. The main results obtained from this chapter reveal that all the adopted models (except the MLR) have accurate and stable performances in predicting the path loss for enclosed indoor environments, such as corridors. Moreover, as the comparison between the models, the best-fit models according to the minimum RMSE and high R-squared and correlation factor are the ANN and RNN-LSTM. This work shows that these ML-based models could be promising solutions with higher precision for predicting path loss for future indoor wireless communication networks.

The following research will be directed toward developing a ML-based path loss model that provides better accuracy and stability than other well-known methods. Moreover, using the results from this study that the best performance was achieved by the neural networks (i.e., ANN and RNN-LSTM), prediction based on deep neural networks such as using a convolution neural network (CNN) will be investigated. Finally, all these research findings will also be considered to provide the most accurate path loss model based on the ensemble method of neural networks, as to be discussed in the next chapter.

# CHAPTER 7

## **PROPOSAL OF AN ENSEMBLE-METHOD-BASED NEURAL NETWORK PATH LOSS MODEL**

---

## ***Chapter 7: Proposal of an Ensemble-Method-Based Neural Network Path Loss Model***

### **7.1 Introduction**

A jump in the need for higher frequency bands has accelerated lately due to the evolution of technologies that require greater data speeds. In recent years, various efforts have developed path loss prediction models for 5G-and-beyond communication networks in the regime of millimeter-wave and sub-terahertz frequencies. However, more advanced models are needed for enhanced flexibility and precision, especially in complex environments. These advanced models will help in deploying wireless networks with the guarantee of covering communication environments with optimum quality of service. This chapter presents path loss prediction models based on neural networks, namely artificial neural network (ANN), artificial recurrent neural network (RNN) based on long short term memory (LSTM), shortly known as RNN-LSTM, and convolutional neural network (CNN). Moreover, an ensemble-method-based neural network path loss model is proposed in this chapter. Finally, an extensive performance analysis of the four models is provided regarding prediction accuracy, stability, the contribution of input features, and the time needed to run the model.

To the best of our knowledge, based on a literature review, this is the first effort that uses ensemble-method-based neural networks for predicting path loss. All the existing studies exploit only separate algorithms for this objective. Moreover, the model run time is used in evaluating the performance of the proposed model and comparing it with other standard ML algorithms. Also, the relative contribution of the models' input features to prediction accuracy can be used to carefully consider when collecting the raw data via measurement campaigns, drive tests, or simulation tools, as well as in building any ML-based model for similar environments.

### **7.2 Ensemble Method for Path Loss Prediction Modeling**

If incorporating machine learning methods in telecommunication or any other operation, the question that arises is *which machine learning methods should I use?* due to the increased number of current machine learning architectures and the improvements made therein. Each alternative method has its advantages and drawbacks. The motivating idea of this study is to combine the advantages of three models based on the ANN, RNN-LSTM, and CNN architectures in one method instead of selecting only one of these. This method was hypothesized to achieve a more optimal model with better accuracy than the individual models.

The proposed ensemble method considers the advantages of these three ML-based models by training them with the same dataset to predict the path loss. In the ensemble step, we used a novel probability aggregation method based on applying a coefficient to the output of each model. The probability coefficients of the models were determined during the validation phase. Figure 7.1 shows a flowchart of the ensemble method, which starts with pre-processing steps to prepare the dataset and select the best features as an input. Training the selected neural network models is performed. After that, the probability coefficients of these models are determined by searching for the best rate for each model. Finally, the ensemble method combines the sum of selected models multiplied by their appropriate probability rate.

### 7.2.1 The CNN Model

Convolution neural networks (CNNs) are well-known methods used to solve supervised and unsupervised problems [203], [204]. They have been shown to significantly improve solving challenges in various sectors, including medical imaging, telecommunication, etc. The CNN is a class of artificial neural networks used to detect, extract, and analyze complex features. It consists of neurons and comprises a set of layers that accept and process inputs of high dimensional vectors as input. The CNN has several hyperparameters essential in achieving optimum model performance. These hyperparameters can be divided into two types:

- a) Hyperparameter that determines the network structure such as:
  - *Kernel Size* - presents the size of the filter.
  - *Kernel Type* - shows values of the actual filter (e.g., edge detection, sharpen).
  - *Stride* - is the rate at which the kernel passes over the input image.
  - *Padding* add layers of 0's to make sure the kernel passes over the image's edge.
  - *Hidden layer* - is layers between input and output layers.
  - *Number of neurons* - tells the number of neurons in each hidden layer.
  - *Activation functions* - allow the model to learn nonlinear prediction boundaries.
- b) Hyperparameter that determines the network trained such as:
  - *Learning rate* - regulates the weight update at the end of each batch.
  - *Momentum* - regulates the value to let the previous update influence the current weight update.
  - *A number of epochs* - is the iterations of the entire training dataset to the network during training.
  - *Batch size* is the number of patterns shown to the network before the weights are updated.

Optimization of CNN hyperparameters is challenging due to the large number of hyperparameters in a standard CNN's architecture [205], and finding the optimized CNN model can improve CNN performance. This work adopted an optimized CNN model using the Optuna technique, which is shown to effectively

select the optimal hyperparameters to achieve path loss prediction. The hyperparameters of our CNN model include two hidden layers. The first layer is a one-dimensional convolution layer with 128 filters and a kernel size of 2; the second is a feed-forward neural network layer containing 64 neurons. The *ReLU* activation function follows each hidden layer. The input for the network accepts four features from the preprocessed data, which are the distance, frequency, antenna height, and AoA. The output layer is a single neuron with a linear activation function as the transfer function that leads to the predicted value. More hyperparameter tuning, such as *learning rate* = 0.001, *batch size* = 32, and *number of epochs* = 100, are obtained using the Optuna technique, which gives the optimum values for such hyperparameters that accurately fit the measured data.

### 7.2.2 The Ensemble-Method-Based Model

The proposed ensemble method has combined the benefits of the adopted ANN, RNN-LSTM, and CNN models into one model to predict the path loss for future wireless communication systems. The processed data is divided into training and testing datasets. After that, the Optuna technique selects the best hyperparameters for the adopted models [206]. This is done by training all models in several parameters and choosing the best hyperparameters that fit the measured data and minimize the prediction error. The best hyperparameters are used to train and test the models. The optimum results were obtained using a training size of 80%, while the test data size was 20%.

In regression applications, combining the predictions of a solution set is named ensemble learning, and the aggregation technique used is called the ensemble method. The proposed ensemble method combines and analyzes the results of three models: ANN, RNN-LSTM, and CNN. An averaging technique can be the most straightforward answer to combining the outputs of various models. The average result is a simple and effective method preferred in solution communities with close success scores. However, the proposed ensemble method requires that a robust method must have a greater impact on the final prediction when combining various models. In this study, this is achieved by giving more weight to the best model.

The ensemble method has been designed more efficiently here to predict path loss. Predicting path loss is challenging due to the reasons mentioned above. It is possible to solve these complex challenges with the weighted sum technique. The proposed ensemble method has achieved the goal by combining the outputs of each model with an optimization technique where the optimum weighted parameters can be obtained.

The proposed ensemble method multiplies the probability weight of each model with its model accuracy to predict path loss. These calculations can show the prediction of the ensemble's path loss using Eq. (7.1) where the  $P_i$  notation is the probability weight of the  $i$  model. Since the softmax activation function is present in the last output layer of the three models, the sum of all probabilities given to each model is equal

to one. Path loss predictions made by the three models are multiplied with their corresponding weighted value to give the ensemble method output. These multiplication results are then summed to produce the final results, as:

$$\text{Ensemble Output} = \sum_{i=1}^3 P_i M_i. \quad (7.1)$$

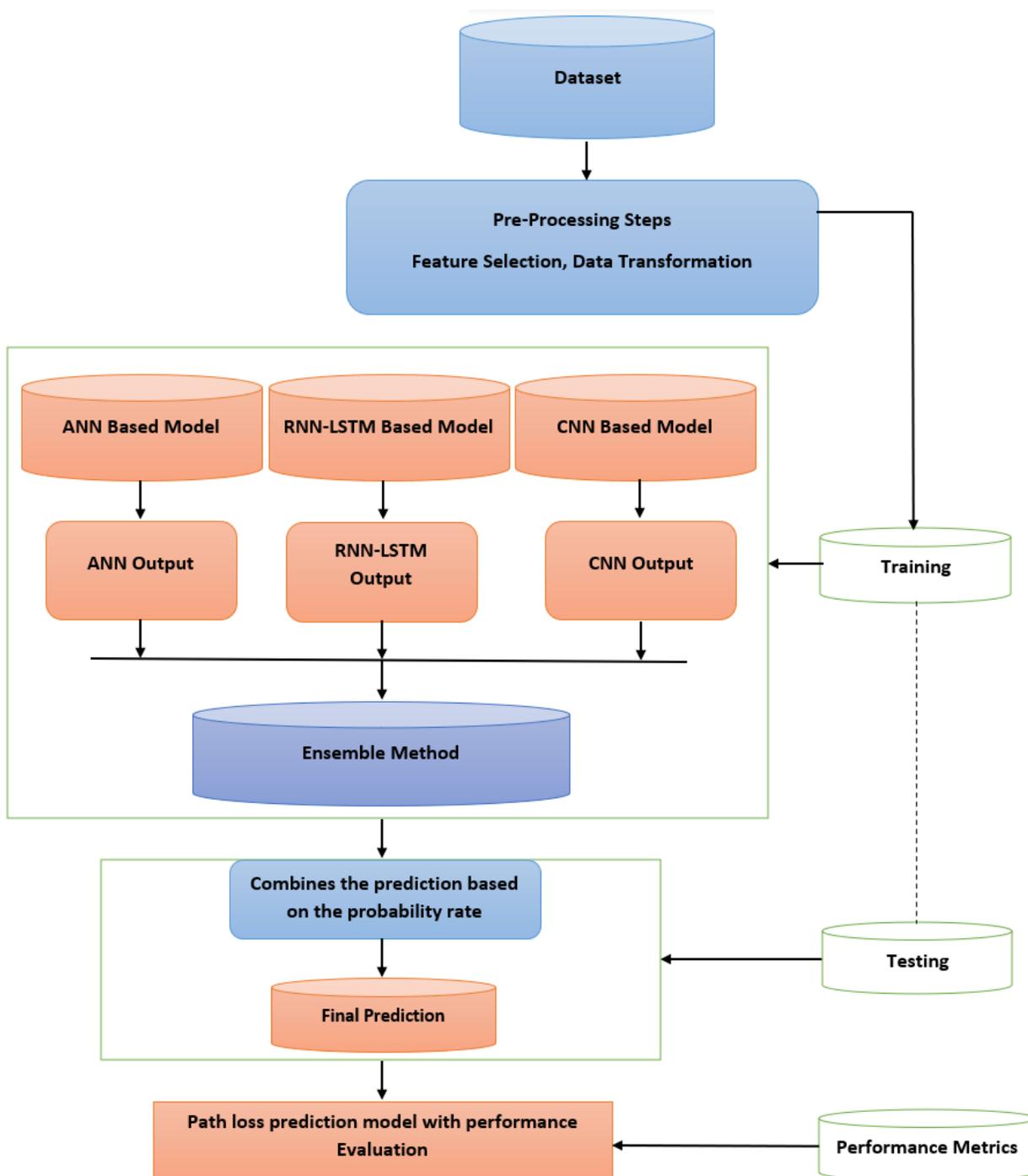


Figure 7.1: Flow chart of the adopted ML-based path loss prediction modeling technique.

### 7.3 Results and Discussions about Ensemble method Path Loss Modeling

This section presents and discusses the main research finding of the chapter. The results are based on plotting the predicted neural network models with the actual datasets individually to observe the behavior of each model in terms of fitting the measurement data. Moreover, the prediction models are plotted together with real data to provide a comparative analysis between the models. The comparison is also studied using the performance metrics mentioned above. Furthermore, the run time and the input features' contribution to the accuracy of the models are numerically provided in this section.

Measured data and the predicted path loss model results are given in Figure 7.2 using the ANN, RNN-LSTM, CNN, and ensemble method with four input features (distance, frequency, antenna height, and AoA). Generally, all the models accurately follow the measurement data. No overfitting and underfitting issues were found for all the models studied, and the training and testing of the models were performed several times to ensure the accuracy and stability of the results. As a justification of the results displayed in Figure 7.2, the numerical results of the five performance metrics are presented in Table 7.1 for all the models with the four input features. For R-square values, a value of 0.9753 is found when the ensemble method model is used, which is the highest value and close to the ideal value of 1. The other models have achieved R-square values of 0.9352, 0.9160, and 0.9543 for the ANN, RNNLSTM, and CNN models. This means that the R-square has improved by 4.3%, 6.5%, and 2.2% over the three models, respectively.

The best performance for the RMSE, MAPE, and MSE was achieved using the proposed ensemble since it provides the lowest values of these error metrics. However, all the models provide accurate path loss predictions since the maximum error value is less than 0.1 dB.

For the correlation coefficient, the best value is found to be 0.9884, achieved by the ensemble method model, whereas the worst value is 0.9666 when the RNN-LSTM model is used. Therefore, as a summary of the results presented in Figure 7.2 and Table 7.1, the best prediction accuracy is achieved by the proposed ensemble method model. As a justification of the results, Figure 7.3 shows the behavior of the prediction error of all the four models selected. It is clear from the figure that the models exhibit relatively low values of prediction error since the distribution is mainly between -3 and 3 dB, with the proposed ensemble method performing best. This is easily observable from Figure 7.4, which depicts the predicted path loss values as a function of the real measured path loss values. Generally, all models provide a straight-line shape, meaning that the prediction path loss values are extremely close to the actual path loss values. The predicted models with the measurement data are represented together in Figure 7.5. These results are much better than using standard and improved empirical path loss models summarized in the previous chapters.

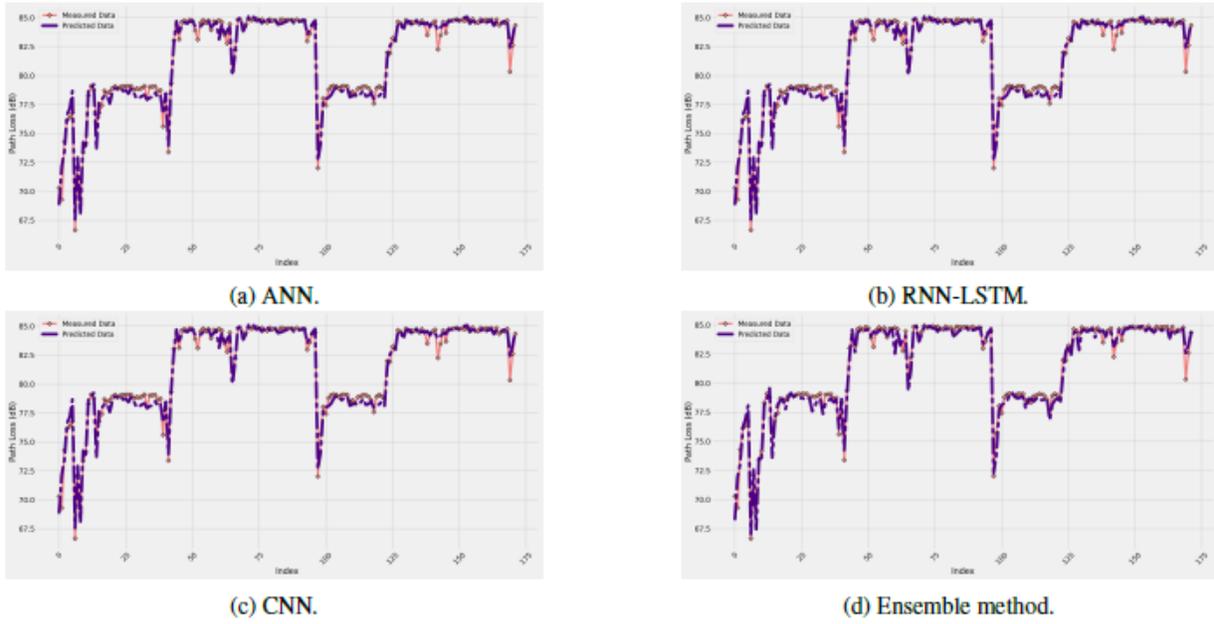


Figure 7.2: Measured data and predicted path loss for all the neural network models.

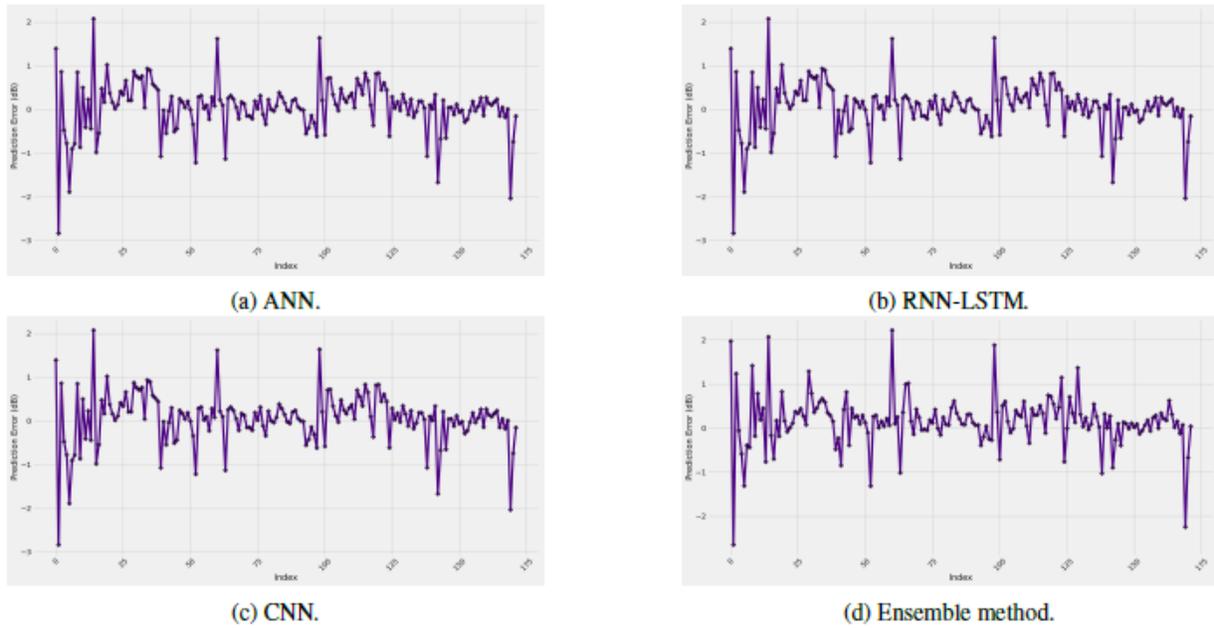


Figure 7.3: Prediction error curves of each model.

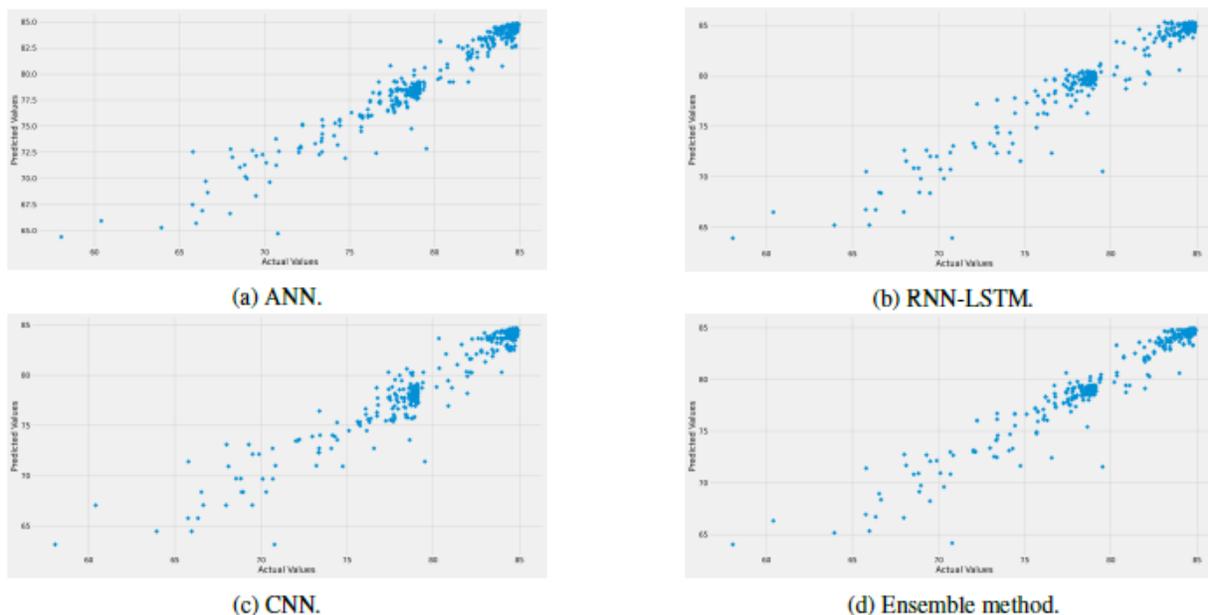


Figure 7.4: Predicted and measured path loss values for each model.

Table 7.1: Performance metrics' values of all the ML-based models selected.

Models	$R^2$	RMSE	MAPE	MSE	Corr
ANN	0.9352	0.0355	0.0312	0.0013	0.9732
RNN-LSTM	0.9160	0.0418	0.0390	0.0019	0.9666
CNN	0.9543	0.0310	0.0263	0.0010	0.9781
<b>Ensemble method</b>	<b>0.9753</b>	<b>0.0228</b>	<b>0.0204</b>	<b>0.0005</b>	<b>0.9884</b>

Table 7.2: Performance metrics after removing the antenna height from the input features of the models.

Models	$R^2$	RMSE	MAPE	MSE	Corr
ANN	0.9093	1.7297	0.0801	2.9918	0.9382
RNN-LSTM	0.8509	1.4149	0.0893	2.0021	0.9062
CNN	0.9323	0.0347	0.0342	0.0012	0.9693
<b>Ensemble method</b>	<b>0.9685</b>	<b>0.0243</b>	<b>0.0243</b>	<b>0.0006</b>	<b>0.9787</b>

Table 7.3: Performance metrics after removing the antenna height and the AoA from the features of the models.

Models	$R^2$	RMSE	MAPE	MSE	Corr
ANN	0.8482	2.0028	0.1049	4.0114	0.9224
RNN-LSTM	0.8240	2.0226	0.0998	4.0910	0.9278
CNN	0.9099	0.1019	0.6763	0.0107	0.8356
<b>Ensemble method</b>	<b>0.9249</b>	<b>0.0993</b>	<b>0.0684</b>	<b>0.0098</b>	<b>0.9401</b>

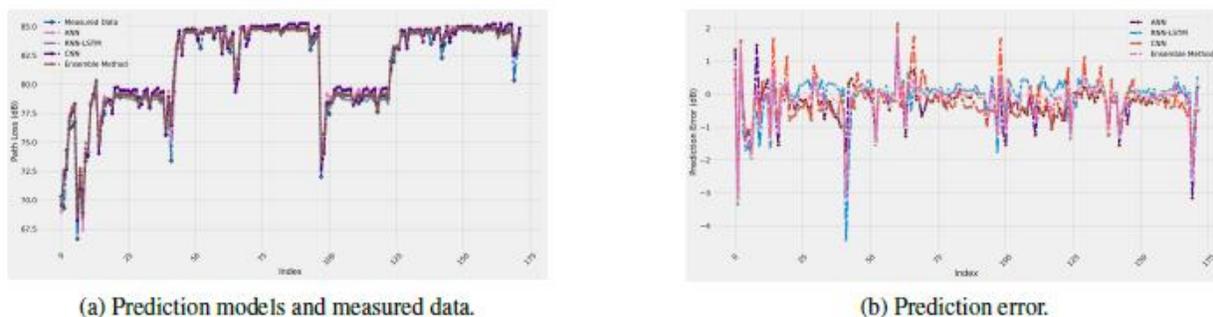


Figure 7.5: Path loss prediction models and prediction error for all the studied models together.

Table 7.4: Runtime comparison of the adopted models.

Models	RunTime [seconds]
ANN	59.1485
RNN-LSTM	124.9184
CNN	92.6543
Ensemble method	276.7636

To evaluate the contribution of the antenna height as an input feature to the overall prediction accuracy of the neural network models, we removed it and kept the other three features (i.e., distance, frequency, and antenna height). As was expected, the prediction accuracy of all models is decreased. Specifically, the R-square value decreased by 4.5% for the ensemble method model, whereas the reduction is 6.7%, 3.2%, and 2.4% for the ANN, RNN-LSTM, and CNN models. This means that the contribution of the antenna height to the overall model accuracy is not significant for enclosed indoor corridor environments where the richness of propagation mechanisms exists. Table 7.2 summarizes the performance metrics values after removing the antenna height from the input features of the models. From the Table, the metric values do not have significant changes, which indicates the superiority of the distance, frequency, and AoA as the main contributors to overall model accuracy.

We removed the antenna height and AoA from the models' input features. The performance metrics' results are presented in Table 7.3. Here, a notable change in the values reveals that the AoA contributes considerably to the models' prediction accuracy. Generally, the best performance is still observed when the ensemble model is used. However, as it depends on all the other three neural network models, the ensemble method model has a longer run time of 276.7636 seconds, almost 4.7 times the time required to run the ANN model. This ANN model requires a minimum time of 59.1485 seconds, among the others. Table 7.4 summarizes the run time required for each model.

The complexity and run time of the model are increased compared to the other models. However, as a tradeoff between the model's complexity and accuracy, the ensemble method model provides the highest accuracy and stability among all the other three models and the other well-known ML-based models. The proposed ensemble method model can be a trustful model to predict path loss for high-frequency bands in complex environments where the propagation signal suffers from several effects in the wireless channel.

## 7.4 Chapter Summary

In this chapter, an ensemble method is proposed based on combining three machine learning techniques. The ensemble approach results have shown that the architecture can identify and predict a high number of the measured path loss data and minimize the mean square error. This conclusion was based on 20% testing performed on the measured dataset. It achieved a high accuracy percentage of the correlation and R-square metrics while reducing the mean square error, absolute mean square error, and the root square error to a significant error. These results indicate that the proposed ensemble method can provide an accurate path loss prediction model, which can play a significant role in modeling wireless communication channels for 5G systems and beyond.

# CHAPTER 8

## **CONCLUSIONS AND RECOMMENDATIONS FOR RELATED FUTURE WORKS**

---

## *Chapter 8: Conclusions and Recommendations for Related Future Works*

### **8.1 Conclusions**

It is well-known that each radio receiver needs a specific minimum received signal power (sensitivity) to decode the received information successfully. Accurate determination of the received signal strength demands advanced path loss prediction models. Hence, these models are of extreme importance in designing and planning any mobile communication network, especially for locating base stations, coverage analysis, and link budget calculations.

The first part of this dissertation proposed a new measurement-based path loss prediction model for wireless communications in enclosed indoor environments. The model was derived, evaluated, and compared with two well-known standard models: the close-in (CI) free space reference distance model and the floating-intercept (FI) model. In addition, the free space path loss (FSPL) model was also considered in this part. The main research findings reveal that the proposed model provides high prediction accuracy in all the frequency bands selected since it well-fits the measurement data. Furthermore, the mean prediction and standard deviation errors between the proposed model and the actual data were between 1.1 and 2.5 dB, proving the model's precision.

After that, a new LOS probability model was presented and compared with two standard models, which are ITU-R and WINNER II (A1). The models' parameters are derived based on the MMSE concept between the models and measurement data. This study's main findings reveal that our proposed model has the best performance at 14 and 18 GHz, while the ITU-R model has a slight preference at 22 GHz. Besides, the impact of the operating frequency is investigated. The results show that the LOS probability models' accuracy depends on the frequency, and this dependency does not follow a specific behavior since the WINNER II (A1) and our model have the best accuracy at 18 GHz compared to 14 and 22 GHz bands. Moreover, this work shows that the LOS probability in indoor environments can be modeled precisely through our proposed model. The reason behind proposing this accurate LOS probability model was to propose a probabilistic path loss prediction model to tackle the classification problem between LOS and NLOS communication scenarios. The motivation for proposing this model was the challenges in analyzing mobile networks at mmWave and higher frequencies because of the big difference between LOS and NLOS communications. The difference is evident in these frequency bands since it is well-known that the mmWave frequency bands suffer from the propagation mechanisms and have higher penetration and

reflection losses than the sub-6 GHz frequency regime. The main advantage of this probabilistic model, in addition to its accuracy and effectiveness, is that it does not require the knowledge of the mobile equipment's LOS and NLOS communication conditions since the probability of having a clear LOS connection is already considered in the model. Furthermore, the results show that the proposed probabilistic path loss model is accurate and stable since it follows the measurement data at the three frequencies adopted with acceptable MPE and SDE values compared to both the LOS and NLOS actual data.

Motivated by the need for higher accuracy and improving standard models, in chapter 4, an efficient improvement of two well-known path loss prediction models, namely the CI and FI models, was presented and discussed in detail. The validation of the models' performance was given by applying the CI and FI models and their improvement to fit real measured data. The data was collected in a typical indoor corridor environment at three frequencies in the SHF band, which are 14, 18, and 22 GHz. Both the LOS and NLOS communication scenarios were considered in this research. The main findings of this work are that our proposed models generally outperform the existing standard models in terms of several factors, such as the accuracy of predicting the path loss with the lowest possible value of the MSE, minimizing the SF's standard deviation for both LOS and NLOS conditions (with better improvements in the NLOS scenario), and providing better sensitivity and stability of the models' parameters with the change of the AoA and antenna height. The improvement of the models was effected through a valuable approach. There is no notable increase in the models' complexity to be used by the planning engineers for wireless systems' deployment and link budget calculations. Motivated by the achieved results in chapter 4, the impact of high-ordering the log distance dependency of the CI model on the Tx-Rx separation distance in the logarithmic scale was investigated in chapter 5. As was expected, higher prediction accuracy is achieved by the third-order CI model since it has more parameters depending on the Tx-Rx distance in the logarithmic scale. Therefore, it can be concluded that the optimum number of parameters is three since the third-order CI model fits the data perfectly in both the LOS and NLOS scenarios in all the frequency bands selected. In addition, the model's complexity is not high since other standard models are used with the same number of parameters. It is worth noting that the principle proposed in chapter 5 can be applied to any other path loss prediction model to achieve more precision in predicting the path loss and provide more details in describing the wireless propagation channel.

The final part of this dissertation provided an extensive comparative analysis to evaluate the performance of the most widespread and used machine learning methods, namely the MLR, PR, RF, DT, SVR, KNN, ANN, and RNN-LSTM. The input features used to train these ML-based models were carefully selected: the Tx-Rx separation distance, frequency, Tx height, and the AoA. Moreover, we utilized a hyperparameters tuning method to select the optimum hyperparameters of the model and avoid the time

consumption of the manual selection. Five performance metrics were applied to evaluate the models: R-squared, RMSE, MAPE, MSE, and the correlations coefficient. The main results obtained from this research reveal that all the adopted models (especially the ANN and RNN-LSTM) have accurate and stable performances in predicting the path loss for enclosed indoor environments, such as corridors. Motivated by that and seeking higher prediction accuracy, a novel ensemble-method-based neural network path loss model was proposed. The ensemble method integrates the advantages of combined models to overcome the shortcomings of each technique when used individually. This has been done to resolve issues related to one another when used for path loss prediction challenges. It also applies the Optuna technique or hyperparameters-tuning, which can look for the best hyperparameters values and the optimum probability weight for each model. Also, this architecture focuses on improving the limitation of neural networks in path loss prediction. Among all the proposed and studied models, the best performing path loss model is the ensemble method since it outperforms all the other methods. The proposed method's best performance is primarily because of the combination of the three neural network models (i.e., ANN, RNN-LSTM, and CNN).

Finally, this dissertation shows that the proposed models can be trusted as accurate and reliable models for predicting the path loss at frequency bands above 6 GHz in enclosed indoor environments such as corridors since they provide better accuracy, sensitivity, and stability than well-known standard path loss prediction models.

## 8.2 Recommendations for Related Future Works

The research findings of this dissertation motivate further investigations and performance studies based on measurements. The following points will be considered in future works:

- 1) Future research could be directed towards implementing measurements in more cities and scenarios and providing other improvements to acquire reliable channel models that can accommodate various environments and optimize existing models' parameters.
- 2) The ensemble-method-based neural network path loss model proposed in this work can be evaluated using different measurement datasets to verify whether they are limited to data or can be generalized. To achieve that, measurement campaigns in higher frequency bands for different indoor and outdoor environments will be conducted to collect the required data for training and testing the ML-based models.
- 3) The proposed ensemble method model can combine several machine learning methods in one model integration to detect the path loss with a high accuracy rate. Thus, it is worth investigating the integration of deep learning techniques with different models based on ensemble theory.

- 4) The channel sounder used in this research work was limited to the 22 GHz frequency band. Future works should cover higher frequency bands in the mmWave regime and sub-terahertz that will be utilized for high-speed multimedia services.
- 5) This study focused on indoor corridor environments; future studies should consider other essential indoor environments that are lacking in literature such as mosques, clinics, shopping malls, gym centers, restaurants, and cafes. Moreover, the modeling and characterization of path loss considering outdoor environments such as stadiums and over rooftops to streets are limited in the literature; future works should cover these crucial scenarios for reliable communications everywhere.
- 6) After covering several indoor environments, it becomes necessary to propose a mathematical formula of the PLE and our proposed models' parameters as a function of the structure dimensions and materials used. This can make the estimation of the signal behavior at further indoor environments more reliable and straightforward with the help of machine learning algorithms.
- 7) The penetration loss and characterizations of the reflection and transmission coefficients will be considered in future works to make our proposed models more precise and trustworthy for indoor to outdoor communications at the 6G and beyond because the signals at the mmWave frequency regime are susceptible to the structure and materials of the buildings. These studies are of importance to achieve the best path loss models that can be reliable for wireless systems' planning and link budget calculations.

# REFERENCES

---

## References

- [1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless Communications through Reconfigurable Intelligent Surfaces," *IEEE Access*, vol. 7, pp. 116753-116773, 2019.
- [2] (Feb. 2019). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017-2022. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/serviceprovider/visualnetworking/index-dex-vni/white-paper-c11-738429.pdf>
- [3] N. O. Oyie and T. J. O. Afullo, "Measurements and Analysis of Large-Scale Path Loss Model at 14 and 22 GHz in Indoor Corridor," *IEEE Access*, vol. 6, pp. 17205-17214, 2018.
- [4] X. Wu, C. X. Wang, J. Sun, J. Huang, R. Feng, Y. Yang, and X. Ge, "60-GHz Millimeter-Wave Channel Measurements and Modeling for Indoor Office Environments," *IEEE Trans. Antennas Propag.*, vol. 65, no. 4, pp. 1912-1924, Apr. 2017.
- [5] T. S. Rappaport, G. R. Maccartney, M. K. Samimi, and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029-3056, Sep. 2015.
- [6] P. F. M. Smulders, "Statistical Characterization of 60-GHz Indoor Radio Channels," *IEEE Trans. Antennas Propag.*, vol. 57, no. 10, pp. 2820-2829, Oct. 2009.
- [7] L. Pometcu and R. D'Errico, "An Indoor Channel Model for High Data-Rate Communications in D-Band," *IEEE Access*, vol. 8, pp. 9420-9433, 2020.
- [8] A. Osseiran, J. F. Monserratand, and P. Marsch, "5G Mobile and Wireless Communications Technology," *Int. J. Satell. Commun. Networ.*, vol. 34, no. 3, pp. 351-360, May 2016.
- [9] M. Aborahama, A. Zakaria, M. H. Ismail, M. El-Bardicy, M. El-Tarhuni, and Y. Hatahet, "Large-Scale Channel Characterization at 28 GHz on a University Campus in the United Arab Emirates," *Telecommun. Syst.*, vol. 1, pp. 1-15, Jan. 2020.
- [10] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335-349, 2013.
- [11] A. Zhou, J. Huang, J. Sun, Q. Zhu, C.-X. Wang, and Y. Yang, "60 GHz Channel Measurements and Ray Tracing Modeling in an Indoor Environment," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2017, pp. 3029-3056.
- [12] G. R. MacCartney, H. Yan, S. Sun, and T. S. Rappaport, "A Flexible Wideband Millimeter-Wave Channel Sounder with Local Area and NLOS to LOS Transition Measurements," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1-7.
- [13] C.-L. Cheng, S. Kim, and A. Zajic, "Comparison of Path Loss Models for Indoor 30 GHz, 140 GHz, and 300 GHz Channels," in *Proc. 11th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2017, pp. 716-720.

- [14] A. M. Al-samman, T. A. A. Rahman, M. H. Azmi, and S. A. Al-Gailani, "Millimeter-Wave Propagation Measurements and Models at 28 GHz and 38 GHz in a Dining Room for 5G Wireless Networks," *Measurement*, vol. 130, pp. 71-81, Dec. 2018.
- [15] F. Hossain, T. Geok, T. Rahman, M. Hindia, K. Dimyati, and A. Abdaziz, "Indoor Millimeter-Wave Propagation Prediction by Measurement and Ray Tracing Simulation at 38 GHz," *Symmetry*, vol. 10, no. 10, p. 464, Oct. 2018.
- [16] T. Kleine-Ostmann, C. Jastrow, S. Priebe, M. Jacob, T. Kurner, and T. Schrader, "Measurement of Channel and Propagation Properties at 300 GHz," in *Proc. Conf. Precis. Electromagn. Meas.*, Jul. 2012, pp. 258-259.
- [17] A. Gatherer. What Will 6G be. [Online]. Available: <https://www.comsoc.org/publications/ctn/what-will-6g-be>
- [18] Nguyen, C.; Cheema, A. A., "A Deep Neural Network-Based Multi-Frequency Path Loss Prediction Model from 0.8 GHz to 70 GHz," *Sensors* 2021, 21, 5100.
- [19] Wen, J.; Zhang, Y.; Yang, G.; He, Z.; Zhang, W. "Path Loss Prediction Based on Machine Learning Methods for Aircraft Cabin Environments." *IEEE Access* 2019, 7, 159251–159261.
- [20] Ates, H. F.; Hashir, S. M.; Baykas, T.; Gunturk, B.K. "Path Loss Exponent and Shadowing Factor Prediction from Satellite Images using Deep Learning." *IEEE Access* 2019, 7, 101366–101375.
- [21] Huang, J. I. E.; Cao, Y.; Raimundo, X.; Cheema, A. "Rain Statistics Investigation and Rain Attenuation Modeling for Millimeter-Wave Short-Range Fixed Links." *IEEE Access* 2019, 7, 156110–156120.
- [22] He, R.; Gong, Y.; Bai, W.; Li, Y.; Wang, X. "Random Forests Based Path Loss Prediction in Mobile Communication Systems." In *Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, Chengdu, China, 11–14 December 2020; pp. 3029–3056.
- [23] Faruk, N.; Popoola, S.I.; Surajudeen-Bakinde, N.T.; Oloyede, A.A.; Abdulkarim, A.; Olawoyin, L.A.; Ali, M.; Calafate, C. T.; Atayero, A. A. "Path Loss Predictions in the VHF and UHF Bands within Urban Environments: Experimental Investigation of Empirical, Heuristics and Geospatial Models." *IEEE Access* 2019, 7, 77293–77307.
- [24] M. K. Samimi, T. S. Rappaport, and G. R. MacCartney, Jr., "Probabilistic Omnidirectional Path Loss Models for Millimeter-Wave Outdoor Communications," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 357-360, Aug. 2015.
- [25] N. A. Muhammad, P. Wang, Y. Li, and B. Vucetic, "Analytical Model for Outdoor Millimeter Wave Channels using Geometry-Based Stochastic Approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 912-926, Feb. 2017.
- [26] Z. Pi and F. Khan, "An Introduction to Millimeter-Wave Mobile Broadband Systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101-107, Jun. 2011.
- [27] G. R. Maccartney, T. S. Rappaport, S. Sun, and S. Deng, "Indoor Office Wideband Millimeter-Wave Propagation Measurements and Channel Models at 28 and 73 GHz for Ultra-Dense 5G Wireless Networks," *IEEE Access*, vol. 3, pp. 2388-2424, 2015.

- [28] A. M. Al-Samman, M. N. Hindia, and T. A. Rahman, "Path Loss Model in Outdoor Environment at 32 GHz for 5G System," in *Proc. IEEE 3rd Int. Symp. Telecommun. Technol. (ISTT)*, Nov. 2016, pp. 9-13.
- [29] S. Ju, Y. Xing, O. Kanhere, and T. S. Rappaport, "Millimeter Wave and Sub-Terahertz Spatial Statistical Channel Model for an Indoor Office Building". *IEEE J. Selected Areas Commun.*, vol. 39, no. 6, pp.1561-1575, Apr. 2021.
- [30] G. R. MacCartney, T. S. Rappaport, and S. Rangan, "Rapid Fading due to Human Blockage in Pedestrian Crowds at 5G Millimeter-Wave Frequencies," in *Proc. GLOBECOM-IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–7.
- [31] T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, and S. Deng, "Small Scale, Local Area, and Transitional Millimeter Wave Propagation for 5G Communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6474–6490, Dec. 2017.
- [32] G. J. M. Janssen, P. A. Stigter, and R. Prasad, "Wideband Indoor Channel Measurements and BER Analysis of Frequency Selective Multipath Channels at 2.4, 4.75, and 11.5 GHz," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1272–1288, Oct. 1996.
- [33] P. F. M. Smulders and A. G. Wagemans, "Wide-Band Measurements of mmWave Indoor Radio Channels," in *Proc. 3rd IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Oct. 1992, pp. 329–333.
- [34] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. JSAC-5, no. 2, pp. 128–137, Feb. 1987.
- [35] T. Manabe, Y. Miura, and T. Ihara, "Effects of antenna directivity and polarization on indoor multipath propagation characteristics at 60 GHz," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 441–448, Apr. 1996.
- [36] M. N. Hindia, A. M. Al-Samman, and T. Bin Abd Rahman, "Investigation of Large-Scale Propagation for Outdoor-Parking Lot Environment for 5G Wireless Communications," in *Proc. IEEE 3rd Int. Symp. Telecommun. Technol. (ISTT)*, Nov. 2016, pp. 14-18.
- [37] D. Chizhik, J. Du, R. Feick, M. Rodriguez, G. Castro, and R. A. Valenzuela, "Path Loss and Directional Gain Measurements at 28 GHz for Non-Line-of-Sight Coverage of Indoors with Corridors," *IEEE Trans. Antennas Propag.*, vol. 68, no. 6, pp. 4820-4830, Jun. 2020.
- [38] S. Sun, T. S. Rappaport, S. Rangan, T. A. Thomas, A. Ghosh, I. Z. Kovacs, I. Rodriguez, O. Koymen, A. Partyka, and J. Jarvelainen, "Propagation Path Loss Models for 5G Urban Micro- and Macro-Cellular Scenarios," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1-6.
- [39] A. M. Al-Samman, T. A. Rahman, T. Al-Hadhrami, A. Daho, M. N. Hindia, M. H. Azmi, K. Dimiyati, and M. Alazab, "Comparative Study of Indoor Propagation Model below and above 6 GHz for 5G wireless networks," *Electronics*, vol. 8, no. 1, p. 44, Jan. 2019.
- [40] M. B. Majed, T. A. Rahman, O. A. Aziz, M. N. Hindia, and E. Hanafi, "Channel Characterization and Path Loss Modeling in Indoor Environment at 4.5, 28, and 38 GHz for 5G Cellular Networks," *Int. J. Antennas Propag.*, vol. 2018, pp. 1-14, Sep. 2018.

- [41] A. I. Sulyman, A. Alwarafy, H. E. Seleem, K. Humadi, and A. Alsanie, "Path Loss Channel Models for 5G Cellular Communications in Riyadh City at 60 GHz," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1-6.
- [42] N. R. Zulkefly, T. A. Rahman, M. H. Azmi, and O. A. Aziz, "6.5 GHz and 10.2 GHz Path Loss Measurements and Modeling for 5G Communications System Prediction," *Int. J. Res. Eng. Technol.*, vol. 6, no. 11, pp. 6-11, Nov. 2017.
- [43] M. Khalily, M. Ghorraishi, S. Taheri, S. Payami, and R. Tafazolli, "Millimeter-Wave Directional Path Loss Models in the 26 GHz, 32 GHz, and 39 GHz Bands for Small Cell 5G Cellular System," in *Proc. 12th Eur. Conf. Antennas Propag. (EuCAP)*, 2018, pp. 5-17.
- [44] A. M. Al-Samman, T. A. Rahman, N. Hindia, and J. Nasir, "Path Loss Model for Indoor Emergency Stairwell Environment at Millimeter Wave Band for 5G Network," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 26, no. 6, pp. 3024-3032, Nov. 2018.
- [45] A. Al-Samman, T. Rahman, M. Hindia, A. Daho, and E. Hanafi, "Path Loss Model for Outdoor Parking Environments at 28 GHz and 38 GHz for 5G Wireless Networks," *Symmetry*, vol. 10, no. 12, p. 672, Nov. 2018.
- [46] G. R. Maccartney, T. S. Rappaport, M. K. Samimi, and S. Sun, "Millimeter-Wave Omnidirectional Path Loss Data for Small Cell 5G Channel Modeling," *IEEE Access*, vol. 3, pp. 1573-1580, 2015.
- [47] A. M. Al-samman, T. Abd Rahman, and M. H. Azmi, "Indoor Corridor Wideband Radio Propagation Measurements and Channel Models for 5G Millimeter Wave Wireless Communications at 19 GHz, 28 GHz, and 38 GHz Bands," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1-12, Mar. 2018.
- [48] H. A. Obeidat, R. Asif, N. T. Ali, Y. A. Dama, O. A. Obeidat, S. M. R. Jones, W. S. Shuaieb, M. A. Al-Sadoon, K. W. Hameed, A. A. Alabdullah, and R. A. Abd-Alhameed, "An Indoor Path Loss Prediction Model using Wall Correction Factors for Wireless Local Area Network and 5G Indoor Networks," *Radio Sci.*, vol. 53, no. 4, pp. 544-564, Apr. 2018.
- [49] T. O. Olasupo, C. E. Otero, L. D. Otero, K. O. Olasupo, and I. Kostanic, "Path Loss Models for Low-Power, Low-Data Rate Sensor Nodes for Smart Car Parking Systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 6, pp. 1774-1783, Jun. 2018.
- [50] S. Sun, T. S. Rappaport, T. A. Thomas, A. Ghosh, H. C. Nguyen, I. Z. Kovács, I. Rodriguez, O. Koymen, and A. Partyka, "Investigation of Prediction Accuracy, Sensitivity, and Parameter Stability of Large-Scale Propagation Path Loss Models for 5G Wireless Communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 2843-2860, May 2016.
- [51] A. Al-Hourani and K. Gomez, "Modeling Cellular-to-UAV Path-Loss for Suburban Environments," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 82-85, Feb. 2018.
- [52] D. Casillas-Perez, C. Camacho-Gomez, S. Jimenez-Fernandez, J. A. Portilla-Figueras, and S. Salcedo-Sanz, "Weighted ABG: A General Framework for Optimal Combination of ABG Path-Loss Propagation Models," *IEEE Access*, vol. 8, pp. 101758-101769, Jun. 2020.

- [53] S. Kaddouri, M. E. Hajj, G. Zaharia, and G. E. Zein, "Indoor Path Loss Measurements and Modeling in an Open-Space Office at 2.4 GHz and 5.8 GHz in the Presence of People," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1-7.
- [54] C. U. Kumari, P. Kora, K. Meenakshi, and S. K., "Short Term and Long Term Path Loss Estimation in Urban, Suburban and Rural Areas," in *Proc. 3rd Int. Conf. Comput. Methodolog. Commun. (ICCMC)*, Mar. 2019, pp. 114-117.
- [55] J.-H. Lee, J.-S. Choi, and S.-C. Kim, "Cell Coverage Analysis of 28 GHz Millimeter Wave in Urban Microcell Environment using 3-D Ray Tracing," *IEEE Trans. Antennas Propag.*, vol. 66, no. 3, pp. 1479-1487, Mar. 2018.
- [56] J. R. Perez, R. P. Torres, L. Rubio, J. Basterrechea, M. Domingo, V. M. R. Penarrocha, and J. Reig, "Empirical Characterization of the Indoor Radio Channel for Array Antenna Systems in the 3 to 4 GHz Frequency Band," *IEEE Access*, vol. 7, pp. 94725-94736, 2019.
- [57] K. Haneda *et al.*, "5G 3GPP-Like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1-7.
- [58] K. Haneda *et al.*, "Indoor 5G 3GPP-Like Channel Models for Office and Shopping Mall Environments," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, May 2016, pp. 694-699.
- [59] S. Sun, G. R. MacCartney, and T. S. Rappaport, "Millimeter-Wave Distance-Dependent Large-Scale Propagation Measurements and Path Loss Models for Outdoor and Indoor 5G Systems," in *Proc. 10th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2016, pp. 1-5.
- [60] M. Ozpolat, K. Bhargava, E. Kampert, and M. D. Higgins, "Multi-Lane Urban mmWave V2V Networks: A Path Loss Behaviour Dependent Coverage Analysis," *Veh. Commun.*, vol. 30, pp. 2096-2214, Mar. 2021.
- [61] N. O. Oyie and T. J. O. Afullo, "A Comparative Study of Dual-Slope Path Loss Model in Various Indoor Environments at 14 to 22 GHz," in *Proc. Prog. Electromagn. Res. Symp.*, Aug. 2018, pp. 121-128.
- [62] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. MacCartney, Jr., T. S. Rappaport, and A. Alsanie, "Radio Propagation Path Loss Models for 5G Cellular Networks in the 28 GHz and 38 GHz Millimeter-Wave Bands," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 78-86, Sep. 2014.
- [63] G. R. MacCartney and T. S. Rappaport, "Rural Macrocell Path Loss Models for Millimeter Wave Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1663-1677, Jul. 2017.
- [63] S. Hur, S. Baek, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, "Proposal on Millimeter-Wave Channel Modeling for 5G Cellular System," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 454-469, Apr. 2016.
- [64] M. Cheffena and M. Mohamed, "Empirical Path Loss Models for Wireless Sensor Network Deployment in Snowy Environments," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2877-2880, Sep. 2017.

- [65] A. M. Al-Samman, T. A. Rahman, M. Hadri, and M. N. Hindia, "Path Loss and RMS Delay Spread Model for 5G Channel at 19 GHz," in *Proc. IEEE 13th Int. Colloq. Signal Process. Appl. (CSPA)*, Mar. 2017, pp. 49-54.
- [66] Shen, Y.; Shao, Y.; Xi, L.; Zhang, H.; Zhang, J. Millimeter-Wave Propagation Measurement and Modeling in Indoor Corridor and Stairwell at 26 and 38 GHz. *IEEE Access* 2021, 9, 87792–87805.
- [67] Pometcu, L. An Indoor Channel Model for High Data-Rate Communications in D-Band. *IEEE Access* 2019, 8, 9420–9433.
- [68] Ramya, P.M.; Boban, M.; Zhou, C.; Stanczak, S. Using Learning Methods for V2V Path Loss Prediction. In *Proceedings of the 2019 IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakesh, Morocco, 15–18 April 2019; pp. 1–6.
- [69] Yi, L.; Tao, L.; Jun, S. RSSI Localization Method for Mine Underground based on RSSI Hybrid Filtering Algorithm. In *Proceedings of the 2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN)*, Guangzhou, China, 6–8 May 2017; pp. 327–332.
- [70] Hakim, G.P.N.; Alaydrus, M.; Bahaweres, R.B. Empirical Approach of Ad Hoc Path Loss Propagation Model in Realistic Forest Environments. In *Proceedings of the 2016 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, Jakarta, Indonesia, 3–5 October 2016; pp. 139–143.
- [71] Goldoni, E.; Prando, L.; Vizziello, A.; Savazzi, P.; Gamba, P. Experimental Data Set Analysis of RSSI-Based Indoor and Outdoor Localization in LoRa Networks. *Internet Technol. Lett.* 2019.
- [72] Ren, J.; Wang, Y.; Bai, W.; Niu, C.; Meng, S. An Improved Indoor Positioning Algorithm based on RSSI Filtering. In *Proceedings of the 2017 IEEE 17th International Conference on Communication Technology (ICCT)*, Chengdu, China, 27–30 October 2017; pp. 1136–1139.
- [73] Li, S.; Yang, X.; Zhao, R.; Liu, Y.; Zhou, X.; Zhang, L. An Indoor Positioning Method Based on RSSI Probability Distribution. *IOP Conf. Ser. Mater. Sci. Eng.* 2019, 490, 042054.
- [74] Kaibi, Z.; Yangchuan, Z.; Subo, W. Research of RSSI Indoor Ranging Algorithm based on Gaussian-Kalman Linear Filtering. In *Proceedings of the 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Xi'an, China, 3–5 October 2016; pp. 1628–1632.
- [75] Phunthawornwong, M.; Pengwang, E.; Silapunt, R. Indoor Location Estimation of Wireless Devices Using the Log-Distance Path Loss Model. In *Proceedings of the TENCON 2018—2018 IEEE Region 10 Conference*, Jeju, Korea, 28–31 October 2018; pp. 499–502.
- [76] Jo, H.S.; Park, C.; Lee, E.; Choi, H.K.; Park, J. Path Loss Prediction Based on Machine Learning Techniques: Principal Component Analysis, Artificial Neural Network and Gaussian Process. *Sensors* 2020, 20, 1927.
- [77] Kim, M.D.; Liang, J.; Lee, J.; Park, J.; Park, B. Path Loss Measurements and Modeling for Indoor Office Scenario at 28 and 38 GHz. In *Proceedings of the 2016 International Symposium on Antennas and Propagation (ISAP)*, Okinawa, Japan, 24–28 October 2016; pp. 64–65.

- [78] Ambroziak, S.J.; Katulski, R.J. An Empirical Propagation Model for Mobile Radio Links in Container Terminal Environment. *IEEE Trans. Veh. Technol.* 2013, 62, 4276–4287.
- [79] Laskowski, M.; Ambroziak, S.J.; Correia, L.M.; Swider, K. On the Usefulness of the Generalised Additive Model for Mean Path Loss Estimation in Body Area Networks. *IEEE Access* 2020, 8, 376873–376882.
- [80] Nurminen, H.; Talvitie, J.; Ali-Löytty, S.; Müller, P.; Lohan, E. S.; Piché, R.; Renfors, M. Statistical Path Loss Parameter Estimation and Positioning Using RSS Measurements in Indoor Wireless Networks. *In Proceedings of the 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Sydney, Australia*, 13–15 November 2012; pp. 13–15.
- [81] Bose, A.; Chuan, H.F. A Practical Path Loss Model for Indoor WiFi Positioning Enhancement. *In Proceedings of the 2007 6th International Conference on Information, Communications & Signal Processing, Singapore*, 10–13 December 2007; pp. 1–5.
- [82] Soltani, M.; Pourahmadi, V.; Mirzaei, A.; Sheikhzadeh, H. Deep Learning-Based Channel Estimation. *IEEE Commun. Lett.* 2019, 23, 652–655.
- [83] D. Tse and V. Pramod. “Fundamentals of Wireless Communication”. *Cambridge university press*, 2005.
- [84] S. R. Saunders and A. Alejandro. “Antennas and Propagation for Wireless Communication Systems,”. John Wiley & Sons, 2007.
- [85] X. Yin and C. Xiang. “Propagation Channel Characterization, Parameter Estimation, and Modeling for Wireless Communications,”. *John Wiley & Sons*, 2016.
- [86] X. Wang, and H. V. Poor. “Wireless Communication Systems: Advanced Techniques for Signal Reception”. *Prentice Hall Professional*, 2004.
- [87] D. P. Agrawal, and Z. Qing-An. “Introduction to Wireless and Mobile Systems,”. *Cengage learning*, 2015.
- [88] C. L. Cheng, S. Kim, and A. Zaji’c, “Comparison of Path Loss Models for Indoor 30 GHz, 140 GHz, and 300 GHz Channels,” *2017 11<sup>th</sup> European Conf. Antennas Propag. (EUCAP)*, Mar. 2017, pp. 716–720.
- [89] S. Sun *et al.* “Propagation Models and Performance Evaluation for 5G Millimeter-Wave Bands.” *IEEE Transactions on Vehicular Technology* 67, no. 9, 2018: 8422–8439.
- [90] H. Yang. “Introduction to Digital Wireless Communications,”, *IET*, vol. 72, 2017.
- [91] D. Tse and V. Pramod. “Fundamentals of Wireless Communication”. *Cambridge university press*, 2005.
- [92] N. Morinaga, K. Ryuji, and S. Seiichi. “Wireless Communication Technologies: New MultiMedia Systems: New Multimedia Systems,”, *Springer Science & Business Media*, 2000.
- [93] E. Dahlman, P. Stefan, and S. Johan. “4G, LTE-advanced Pro and the Road to 5G,”. *Academic Press*, 2016.

- [94] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," *Future Generat. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.
- [95] T. S. Rappaport, "Spectrum frontiers: The New World of Millimeter-Wave Mobile Communication," *Invited Keynote Presentation, The Federal Communications Commission (FCC) Headquarters*, Mar. 2016.
- [96] N. A. Muhammad *et al.*, "Analytical Model for Outdoor Millimeter Wave Channels using Geometry-Based Stochastic Approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 912–926, Jun. 2016.
- [97] Z. Pi and F. Khan, "An Introduction to Millimeter-Wave Mobile Broadband Systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
- [98] M. K. Samimi, T. S. Rappaport, and G. R. Maccartney, "Probabilistic Omnidirectional Path Loss Models for Millimeter-Wave Outdoor Communications," *IEEE Wirel. Commun. Lett.*, vol. 4, no. 4, pp. 357–360, Mar. 2015.
- [99] M. Shafi *et al.*, "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Apr. 2017.
- [100] S. Hur *et al.*, "Proposal on Millimeter-Wave Channel Modeling for 5G Cellular System," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 454–469, Feb. 2016.
- [101] W. Li, L. Tian, J. Zhang, and Y. Cheng, "Analysis of Base Station Deployment Impact on LOS Probability Model for 5G Indoor Scenario," *In 2017 IEEE/CIC Int. Conf. Commun. China (ICCC)*, Oct. 2017, pp. 1–5.
- [102] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A Tractable Approach to Coverage and Rate in Cellular Networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.
- [103] M. Ding, P. Wang, D. Lopez-Perez, G. Mao, and Z. Lin, "Performance Impact of LOS and NLOS Transmissions in Dense Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2365–2380, Mar. 2016.
- [104] Z. Meng, Y. Chen, M. Ding, and D. López-Pérez, "A New Look at UAV Channel Modeling: A Long Tail of LoS Probability," *IEEE 30th Annu. Int. Symp. Personal, Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–6.
- [105] R. Aleksiejunas, A. Cesiul, and K. Svirskas, "Statistical LOS/NLOS Channel Model for Simulations of Next Generation 3GPP Networks," *Elektronika Ir Elektrotechnika*, vol. 24, no. 5, pp. 74–79, Oct. 2018.
- [106] Z. Cui, K. Guan, C. Briso-Rodríguez, B. Ai, and Z. Zhong, "Frequency-Dependent Line-of-Sight Probability Modeling in Built-up Environments," *IEEE IoT J.*, vol. 7, no. 1, pp. 699–709, Oct. 2019.
- [107] K. Haneda *et al.*, "Indoor 5G 3GPP-like Channel Models for Office and Shopping Mall Environments," *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, May 2016, pp. 694–699.
- [108] 5GCM, "5G Channel Model for Bands up to 100 GHz," Tech. Rep., Oct. 2016. [Online]. Available: <http://www.5gworkshops.com/5GCM.html>

- [109] L. Tian, J. Zhang, P. Tang, F. Huang, and Y. Zheng, "Delay Characteristics for Directional and Omni-Directional Channel in Indoor Open Office and Shopping Mall Environments at 28 GHz," *In 2016 IEEE 27th Annu. Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2016, pp. 1–4.
- [110] R. Miao, L. Tian, Y. Zheng, P. Tang, F. Huang, and J. Zhang, "Indoor Office Channel Measurements and Analysis of Propagation Characteristics at 14 GHz," *IEEE 26th Annu. Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Aug. 2015, pp. 2199–2203.
- [111] J. Li, "Los Probability Modeling for 5G Indoor Scenario," *IEEE Int. Symp. Antennas Propag. (ISAP)*, Oct. 2016, pp. 204–205.
- [112] S. Van Brandt, R. Van Thielen, J. Verhaevert, T. Van Hecke, and H. Rogier, "Characterization of Path Loss and Large-Scale Fading for Rapid Intervention Team Communication in Underground Parking Garages," *Sensors*, vol. 19, no. 11, Jan. 2019.
- [113] S. Y. Seidel, T. S. Rappaport, S. Jain, M. L. Lord, and R. Singh, "Path Loss, Scattering, and Multipath Delay Statistics in Four European Cities for Digital Cellular and Microcellular Radiotelephone," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 721–730, Nov. 1991.
- [114] M. G. Nilsson, C. Gustafson, T. Abbas, and F. Tufvesson, "A Path Loss and Shadowing Model for Multilink Vehicle-to-Vehicle Channels in Urban Intersections," *Sensors*, vol. 18, no. 12, pp. 1–19, Dec. 2018.
- [115] H. S. Jo and J. G. Yook, "Path Loss Characteristics for IMT-Advanced Systems in Residential and Street Environments," *IEEE Antennas Wirel. Propag. Lett.*, vol. 9, pp. 867–871, Aug. 2010.
- [116] J. Bi *et al.*, "Fast Radio Map Construction by Using Adaptive Path Loss Model Interpolation in Large-Scale Building," *Sensors*, vol. 19, no. 3, pp. 1–19, Jan. 2019.
- [117] H. Masui, M. Ishii, K. Sakawa, H. Shimizu, T. Kobayashi, and M. Akaike, "Microwave Path-Loss Characteristics in Urban LOS and NLOS Environments," *IEEE Veh. Technol. Conf. (VTC)*, May 2001, pp. 395–398.
- [118] W. Tang, X. Ma, J. Wei, and Z. Wang, "Measurement and Analysis of Near-Ground Propagation Models Under Different Terrains for Wireless Sensor Networks," *Sensors*, vol. 19, no. 8, Jan. 2019.
- [119] A. Osseiran, J. F. Monserratand and P. Marsch, "5G Mobile and Wireless Communications Technology," *Int. J. Satell. Commun. and Networ.*, vol. 34, no. 3, pp. 351–360, May. 2016.
- [120] M. Marcus and B. Pattan, "Millimeter Wave Propagation: Spectrum Management Implications," *IEEE Microwave Magazine*, vol. 6, no. 2, pp. 54–62, Aug. 2005.
- [121] Ye, H.; Li, G.Y.; Juang, B.H. Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems. *IEEE Wirel. Commun. Lett.* 2017, 7, 114–117.
- [122] Gruber, T.; Cammerer, S.; Hoydis, J.; Brink, S.T. On Deep Learning-Based Channel Decoding. *In Proceedings of the 2017 51st Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD, USA, 22–24 March 2017*; pp. 1–6.

- [123] Thrane, J.; Zibar, D.; Christiansen, H.L. Model-Aided Deep Learning Method for Path Loss Prediction in Mobile Communication Systems at 2.6 GHz. *IEEE Access* 2020, 8, 7925–7936.
- [124] Zhang, Y.; Wen, J.; Yang, G.; He, Z.; Wang, J. Path Loss Prediction Based on Machine Learning: Principle, Method, and Data Expansion. *Appl. Sci.* 2019, 9, 1908.
- [125] Piacentini, M.; Rinaldi, F. Path loss Prediction in Urban Environment Using Learning Machines and Dimensionality Reduction Techniques. *Comput. Manag. Sci.* 2011, 8, 371–385.
- [126] Oroza, C.A.; Zhang, Z.; Watteyne, T.; Glaser, S.D. A Machine-Learning-Based Connectivity Model for Complex Terrain Large-Scale Low-Power Wireless Deployments. *IEEE Trans. Cogn. Commun. Netw.* 2017, 3, 576–584.
- [127] Tahat, A.; Edwan, T.; Al-Sawwaf, H.; Al-Baw, J.; Amayreh, M. Simplistic Machine Learning-Based Air-to-Ground Path Loss Modeling in an Urban Environment. In *Proceedings of the 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC), Paris, France, 20–23 April 2020*; pp. 158–163.
- [128] Bolli, S. Propagation Path Loss Model Based on Environmental Variables. In *Proceedings of the 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 6–8 October 2020*; pp. 368–373.
- [129] Saleh, T.; Petrov, D.; Tirkkonen, O.; Raisanen, V. Probabilistic Path Loss Predictors for mmWave Networks. In *Proceedings of the 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), Helsinki, Finland, 25–28 April 2021*; pp. 1–6.
- [130] Aldossari, S.; Chen, K.C. Predicting the Path Loss of Wireless Channel Models Using Machine Learning Techniques in MmWave Urban Communications. In *Proceedings of the 2019 22nd International Symposium on Wireless Personal Multimedia Communications (WPMC), Lisbon, Portugal, 24–27 November 2019*; pp. 1–6.
- [131] Eichie, J.O.; Oyedum, O.D.; Ajewole, M.O.; Aibinu, A.M. Comparative Analysis of Basic Models and Artificial Neural Network Based Model for Path Loss Prediction. *Prog. Electromagn. Res. M* 2017, 7, 133–146.
- [132] Ayadi, M.; Zineb, A.B.; Tabbane, S. A UHF Path Loss Model Using Learning Machine for Heterogeneous Networks. *IEEE Trans. Antennas Propag.* 2017, 65, 3675–3683.
- [133] Singh, H.; Gupta, S.; Dhawan, C.; Mishra, A. Path Loss Prediction in Smart Campus Environment: Machine Learning-based Approaches. In *Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 25–28 May 2020*; pp. 1–4.
- [134] Yang, G.; Zhang, Y.; He, Z.; Wen, J.; Ji, Z.; Li, Y. Machine-Learning-Based Prediction Methods for Path Loss and Delay Spread in Air-to-Ground Millimetre-Wave Channels. *IET Microwaves Antennas Propag.* 2019, 13, 1113–1121.
- [135] Timoteo, R.D.A.; Cunha, D.C.; Cavalcanti, G.D.C. A Proposal for Path Loss Prediction in Urban Environments Using Support Vector Regression. In *Proceedings of the AICT2014: The Tenth Advanced International Conference on Telecommunications, Paris, France, 20–24 July 2014*; pp. 119–124.

- [136] Ostlin, E.; Zepernick, H.J.; Suzuki, H. Macrocell Path-Loss Prediction Using Artificial Neural Networks. *IEEE Trans. Veh. Technol.* 2010, 59, 2735–2747.
- [137] Popescu, I.; Nikitopoulos, D.; Constantinou, P.; Nafornta, I. ANN Prediction Models for Outdoor Environment. In *Proceedings of the 2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications, Helsinki, Finland*, 11–14 September 2006; pp. 366–371.
- [138] Zhao, X.; Hou, C.; Wang, Q. A New SVM-Based Modeling Method of Cabin Path Loss Prediction. *Int. J. Antennas Propag.* 2013, 2013, 279070.
- [139] Popescu, I.; Nafornta, I.; Constantinou, P. Comparison of Neural Network Models for Path Loss Prediction. In *Proceedings of the IEEE International Conference on Wireless and Mobile Computing, Networking And Communications, Montreal, QC, Canada*, 4–22 August 2005; pp. 44–49.
- [140] Cavalcanti, B.J.; Cavalcante, G.A.; de Mendonça, L.M.; Cantanhede, G.M.; de Oliveira, M.M.M.; D’Assunção, A.G. A Hybrid Path Loss Prediction Model Based on Artificial Neural Networks using Empirical Models for LTE and LTE-A at 800 MHz and 2600 MHz. *J. Microwaves Optoelectron. Electromagn. Appl.* 2017, 16, 708–722.
- [141] Zhang, Y.; Wen, J.; Yang, G.; He, Z.; Luo, X. Air-to-Air Path Loss Prediction Based on Machine Learning Methods in Urban Environments. *Wirel. Commun. Mob. Comput.* 2018, 2018, 8489326.
- [142] Sotiroudis, S.P.; Goudos, S.K.; Siakavara, K. Neural Networks and Random Forests: A Comparison Regarding Prediction of Propagation Path Loss for NB-IoT Networks. In *Proceedings of the 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCASST), Thessaloniki, Greece*, 13–15 May 2019; pp. 1–4.
- [143] Moraitis, N.; Tsipi, L.; Vouyioukas, D. Machine Learning-Based Methods for Path Loss Prediction in Urban Environment for LTE Networks. In *Proceedings of the 2020 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Thessaloniki, Greece*, 12–14 October 2020; pp. 4–9.
- [144] Egi, Y.; Otero, C.E. Machine-Learning and 3D Point-Cloud Based Signal Power Path Loss Model for the Deployment of Wireless Communication Systems. *IEEE Access* 2019, 7, 42507–42517.
- [145] H. S. Jo, C. Park, E. Lee, H. K. Choi, and J. Park, “Path Loss Prediction Based on Machine Learning Techniques: Principal Component Analysis, Artificial Neural Network and Gaussian Process,” *Sensors (Switzerland)*, vol. 20, no. 7, Jan. 2020.
- [146] Park, C.; Tettey, D.K.; Jo, H.S. Artificial Neural Network Modeling for Path Loss Prediction in Urban Environments. arXiv 2019, arXiv:1904.02383.
- [147] Zhang, T.; Liu, S.; Xiang, W.; Xu, L.; Qin, K.; Yan, X. A Real-Time Channel Prediction Model Based on Neural Networks for Dedicated Short-Range Communications. *Sensors* 2019, 19, 3541.
- [148] Uccellari, M.; Facchini, F.; Sola, M.; Sirignano, E.; Vitetta, G.M.; Barbieri, A.; Tondelli, S. On the Application of Support Vector Machines to the Prediction of Propagation Losses at 169 MHz for Smart Metering Applications. *IET Microw. Antennas Propag.* 2018, 12, 302–312.

- [149] Ahmadien, O.; Ates, H.F.; Baykas, T.; Gunturk, B.K. Predicting Path Loss Distribution of an Area from Satellite Images Using Deep Learning. *IEEE Access* 2020, 8, 64982–64991.
- [150] Popoola, S.I.; Adetiba, E.; Atayero, A.A.; Faruk, N.; Calafate, C.T. Optimal Model for Path Loss Predictions using Feed-Forward Neural Networks. *Cogent. Eng.* 2018, 5, 1444345.
- [151] Nguyen, D.C.; Cheng, P.; Ding, M.; Lopez-Perez, D.; Pathirana, P.N.; Li, J.; Seneviratne, A.; Li, Y.; Poor, H.V. Enabling AI in Future Wireless Networks: A Data Life Cycle Perspective. *IEEE Commun. Surv. Tutor.* 2020, 23, 553–595.
- [152] Ferreira, G.P.; Matos, L.J.; Silva, J.M.M. Improvement of Outdoor Signal Strength Prediction in UHF Band by Artificial Neural Network. *IEEE Trans. Antennas Propag.* 2016, 64, 5404–5410.
- [153] Popoola, S.I.; Jefia, A.; Atayero, A.A.; Kingsley, O.; Faruk, N.; Oseni, O.F.; Abolade, R.O. Determination of Neural Network Parameters for Path Loss Prediction in Very High Frequency Wireless Channel. *IEEE Access* 2019, 7, 150462–150483.
- [154] Olajide, O.Y.; Yerima, S.M. Channel Path-Loss Measurement and Modeling in Wireless Data Network (IEEE 802.11n) Using Artificial Neural Network. *Eur. J. Electr. Eng. Comput. Sci.* 2020, 4, 1–7.
- [155] M. Ayadi, A. Ben Zineb, and S. Tabbane, “A UHF Path Loss Model Using Learning Machine for Heterogeneous Networks,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 7, pp. 3675–3683, May 2017.
- [156] H. Suzuki, “Macrocell Path-Loss Prediction Using Artificial Neural Networks,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, May 2010.
- [157] M. Kasparick, R. L. G. Cavalcante, S. Valentin, S. Stańczak, and M. Yukawa, “Kernel-Based Adaptive Online Reconstruction of Coverage Maps With Side Information,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5461–5473, Jul. 2015.
- [158] A. Tahat, T. Edwan, H. Al-Sawwaf, J. Al-Baw, and M. Amayreh, “Simplistic Machine Learning-Based Air-to-Ground Path Loss Modeling in an Urban Environment,” *2020 5th Int. Conf. Fog Mob. Edge Comput. FMEC 2020*, Apr. 2020, pp. 158–163.
- [159] Ribero, M.; Heath, R.W.; Vikalo, H.; Chizhik, D.; Valenzuela, R.A. Deep Learning Propagation Models over Irregular Terrain. In *Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 12–17 May 2019; pp. 4519–4523.
- [160] X. Wu, C. X. Wang, J. Sun, J. Huang, R. Feng, Y. Yang, and X. Ge, “60-GHz Millimeter-Wave Channel Measurements and Modeling for Indoor Office Environments,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 4, pp. 1912–1924, Apr. 2017.
- [161] N. A. Muhammad, P. Wang, Y. Li, and B. Vucetic, “Analytical Model for Outdoor Millimeter Wave Channels using geometry-based Stochastic Approach,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 912–926, Feb. 2017.
- [162] K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li, and T. Balercia, “5G 3GPP-like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments,” *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–7.

- [163] F. D. Diba, M. A. Samad, and D. Y. Choi, "Centimeter and Millimeter-Wave Propagation Characteristics for Indoor Corridors: Results From Measurements and Models," *IEEE Access*, vol. 9, pp.158726-158737, Nov. 2021.
- [164] S. Y. Seidel, and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. Antennas Propag.*, vol. 40, no. 2, pp. 207–217, Feb. 1992.
- [165] N. Yoza-Mitsuishi and R. Sun, "Path Loss Characterization of 7-GHz Terrestrial Propagation Channel in Urban Environment", *2019 IEEE 90th Veh. Techn. Conf. (VTC2019-Fall)*, Sep. 2019, pp. 1-5.
- [166] T. Zwick, and T. J. Beukema, and H. Nam, "Wideband channel sounder with measurements and model for the 60 GHz indoor radio channel", *IEEE Trans. Veh. Tech.*, vol. 54, no. 4, pp. 1266–1277, Aug. 2005.
- [167] S. Geng, J. Kivinen, X. Zhao, and P. Vainikainen, "Millimeter-Wave Propagation Channel Characterization for Short-Range Wireless Communications", *IEEE Trans. Veh. Tech.*, vol. 58, no. 1, pp. 3–13, May. 2008.
- [168] S. Deng, G. R. MacCartney, and T. S. Rappaport, "Indoor and Outdoor 5G Diffraction Measurements and Models at 10, 20, and 26 GHz," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1-7.
- [169] M. Khatun, H. Mehrpouyan, and D. Matolak, "60-GHz Millimeter-Wave Path Loss Measurements in Boise Airport," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2018, pp. 1276-1280.
- [170] H. Zhao, Q. Wang, and K. Shi, "Analysis on Human Blockage Path Loss and Shadow Fading in Millimeter-Wave Band," *Int. J. Antennas Propag.*, vol. 2017, pp. 1-6, Jan. 2017.
- [171] O. Ahmadien, H. F. Ates, T. Baykas, and B. K. Gunturk, "Predicting Path Loss Distribution of an Area from Satellite Images using Deep Learning," *IEEE Access*, vol. 8, pp. 64982-64991, 2020.
- [172] A. Alwarafy, A. I. Sulyman, A. Alsanie, S. A. Alshebeili, and H. M. Behairy, "Path-Loss Channel Models for Receiver Spatial Diversity Systems at 2.4 GHz," *Int. J. Antennas Propag.*, vol. 2017, pp. 1-12, Jan. 2017.
- [173] M. Laskowski, S. J. Ambroziak, L. M. Correia, and K. Swider, "On the Usefulness of the Generalised Additive Model for Mean Path Loss Estimation in Body Area Networks," *IEEE Access*, vol. 8, pp. 176873-176882, 2020.
- [174] L. Rubio, J. Reig, H. Fernández, and V. M. Rodrigo-Peñarrocha, "Experimental UWB Propagation Channel Path Loss and Time-Dispersion Characterization in a Laboratory Environment," *Int. J. Antennas Propag.*, vol. 2013, pp. 1-7, Mar. 2013.
- [175] L. Wu, D. He, B. Ai, J. Wang, H. Qi, K. Guan, and Z. Zhong, "Artificial Neural Network Based Path Loss Prediction for Wireless Communication Network," *IEEE Access*, vol. 8, pp. 199523-199538, 2020.
- [176] Y. Xing and T. S. Rappaport, "Propagation Measurements and Path Loss Models for Sub-THz in Urban Microcells," 2021, arXiv:2103.01151. [Online]. Available: <http://arxiv.org/abs/2103.01151>

- [177] Y. Xing, T. S. Rappaport, and A. Ghosh, "Millimeter Wave and Sub-THz Indoor Radio Propagation Channel Measurements, Models, and Comparisons in an Office Environment," *IEEE Commun. Lett.*, early access, Jun. 10, 2021.
- [178] N. Faruk *et al.*, "Clutter and Terrain Effects on Path Loss in the VHF/UHF Bands," *IET Microwaves, Antennas Propag.*, vol. 1s2, no. 1, pp.69-76, Jan. 2018.
- [179] N. Faruk *et al.*, "Large-Scale Radio Propagation Path Loss Measurements and Predictions in the VHF and UHF Bands," *Heliyon*, vol. 15, e07298, Jan. 2021.
- [180] Bergstra, J.; Komer, B.; Eliasmith, C.; Yamins, D.; Cox, D.D. Hyperopt: A Python Library for Model Selection and Hyperparameter Optimization. *Comput. Sci. Discov.* 2015, 8, 014008.
- [181] Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems 25 (NIPS 2012)*; Curran Associates, Incorporated: New York, NY, USA, 2012.
- [182] Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Sequential Model-Based Optimization for General Algorithm Configuration. In *Proceedings of the International Conference on Learning and Intelligent Optimization, Rome, Italy, 17–21 January 2011*; Volume 6683, pp. 507–523.
- [183] Koch, P.; Golovidov, O.; Gardner, S.; Wujek, B.; Grin, J.; Xu, Y. Autotune: A Derivative-Free Optimization Framework for Hyperparameter Tuning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining July, London, UK, 19–23 August 2018*; pp. 443–452.
- [184] Golovin, D.; Solnik, B.; Moitra, S.; Kochanski, G.; Karro, J.; Sculley, D. Google Vizier: A Service for Black-Box Optimization. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017*; pp. 1487–1495.
- [185] Wicaksono, A.S.; Supianto, A.A. Hyper Parameter Optimization Using Genetic Algorithm on Machine Learning Methods for Online News Popularity Prediction. *Int. J. Adv. Comput. Sci. Appl.* 2018, 9, 263–267.
- [186] Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-Generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019*; pp. 2623–2631.
- [187] Zahedi, L.; Mohammadi, F.G.; Rezapour, S.; Ohland, M.W.; Amini, M.H. Search Algorithms for Automated Hyper-Parameter Tuning. *arXiv 2021*, arXiv:2104.14677.
- [188] Maulud, D.; Abdulazeez, A.M. A Review on Linear Regression Comprehensive in Machine Learning. *J. Appl. Sci. Technol. Trends* 2020, 1, 140–147.
- [189] Chatterjee, S.; Hadi, A.S. Regression Analysis by Example; *John Wiley Sons: Hoboken, NJ, USA*, 2013.
- [190] Hans, C. Bayesian Lasso Regression. *Biometrika* 2009, 96, 835–845.
- [191] Gauraha, N. Introduction to the LASSO. *Resonance* 2018, 23, 439–464.

- [192] Vapnik, V.N. An Overview of Statistical Learning Theory. *IEEE Trans. Neural Netw.* 1999, 10, 988–999.
- [193] Xu, X.; Zhou, C.; Wang, Z. Credit Scoring Algorithm Based on Link Analysis Ranking with Support Vector Machine. *Expert Syst. Appl.* 2009, 36, 2625–2632.
- [194] Huang, C.L.; Tsai, C.Y. A hybrid SOFM-SVR with a Filter-Based Feature Selection for Stock Market Forecasting. *Expert Syst. Appl.* 2009, 36, 1529–1539.
- [195] Patel, J.; Shah, S.; Thakkar, P.; Kotecha, K. Predicting Stock and Stock Price Index Movement Using Trend Deterministic Data Preparation and Machine Learning Techniques. *Expert Syst. Appl.* 2015, 42, 259–268.
- [196] Quinlan, J.R. Induction of decision trees. *Mach. Learn.* 1986, 1, 81–106.
- [197] Quinlan, J.R. *Programs for Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2014.
- [198] Loh, W.Y. Classification and Regression Trees. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 2011, 1, 14–23.
- [199] Han, J.; Pei, J.; Kamber, M. *Data Mining: Concepts and Techniques*; Elsevier: Amsterdam, The Netherlands, 2011.
- [200] Dupond, S. A Thorough Review on the Current Advance of Neural Network Structures. *Annu. Rev. Control.* 2019, 14, 200–230.
- [201] Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term Dependencies with Gradient Descent is Difficult. *IEEE Trans. Neural Netw.* 1994, 5, 157–166.
- [202] Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* 1997, 9, 1735–1780.
- [203] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet Classification with Deep Convolutional Neural Networks,” *Advances in Neural Information Processing Systems*, 2012.
- [204] Y. Zhu *et al.*, “Converting Tabular Data into Images for Deep Learning with Convolutional Neural Networks,” *Scientific Reports*, vol. 11, no. 1, pp.1-11, May 2021.
- [205] A. S. Wicaksono, and A. A. Supianto, “Hyper Parameter Optimization using Genetic Algorithm on Machine Learning Methods for Online News Popularity Prediction,” *Int. J. Advanced Computer Science Applications*, vol. 9, no. 12, pp. 263–267, 2018.
- [206] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next Generation Hyperparameter Optimization Framework,” *In Proceedings 25th ACM SIGKDD Int. conf. Knowledge Discovery Data Mining*, Jul. 2019, pp. 2623-2631.

# APPENDICES

---

## *Appendices*

### **Appendix A: Averaged LOS and NLOS real measured data at 14 GHz frequency band**

Table (A.1): LOS 14 GHz Averaged Real Measured Data

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-19.0455	-31.7818
<b>4</b>	-24.936	-26.9733
<b>6</b>	-27.3373	-26.7662
<b>8</b>	-30.4923	-27.6502
<b>10</b>	-31.3021	-29.9632
<b>12</b>	-29.5482	-29.8268
<b>14</b>	-35.7433	-29.1092
<b>16</b>	-30.5105	-30.8837
<b>18</b>	-37.5614	-34.3761
<b>20</b>	-30.2908	-29.0022
<b>22</b>	-33.1304	-33.1269
<b>24</b>	-35.2936	-33.978

Table (A.2): NLOS 14 GHz Averaged Real Measured Data at 30° and 60° AoAs

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power at 30° AoA (dBm)</i>		<i>Received Power at 60° AoA (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-30.515	-30.515	-36.1756	-36.1756
<b>4</b>	-34.1816	-34.1816	-35.9726	-35.9726
<b>6</b>	-34.5907	-34.5907	-37.1009	-37.1009
<b>8</b>	-35.2939	-35.2939	-37.1307	-37.1307
<b>10</b>	-37.2066	-37.2066	-37.5506	-37.5506
<b>12</b>	-36.5375	-36.5375	-37.3632	-37.3632
<b>14</b>	-37.4586	-37.4586	-37.4265	-37.4265
<b>16</b>	-35.9147	-35.9147	-37.4195	-37.4195
<b>18</b>	-37.2895	-37.2895	-37.5741	-37.5741
<b>20</b>	-36.283	-36.283	-37.6117	-37.6117
<b>22</b>	-37.1934	-37.1934	-37.502	-37.502
<b>24</b>	-36.9682	-36.9682	-37.5715	-37.5715

Table (A.3): NLOS 14 GHz Averaged Real Measured Data at 90° and 120°AoAs

<i><b>Tx-Rx Separation</b></i> <i><b>Distance (m)</b></i>	<i><b>Received Power at 90° AoA (dBm)</b></i>		<i><b>Received Power at 120° AoA (dBm)</b></i>	
	<i><b>Tx Antenna height (1.6 m)</b></i>	<i><b>Tx Antenna height (2.3 m)</b></i>	<i><b>Tx Antenna height (1.6 m)</b></i>	<i><b>Tx Antenna height (2.3 m)</b></i>
<i><b>2</b></i>	-35.2333	-35.2333	-36.2064	-36.2064
<i><b>4</b></i>	-37.5457	-37.5457	-35.8982	-35.8982
<i><b>6</b></i>	-37.2579	-37.2579	-37.1117	-37.1117
<i><b>8</b></i>	-37.2703	-37.2703	-37.3807	-37.3807
<i><b>10</b></i>	-37.4	-37.4	-37.5137	-37.5137
<i><b>12</b></i>	-37.5321	-37.5321	-37.2579	-37.2579
<i><b>14</b></i>	-37.3498	-37.3498	-37.3873	-37.3873
<i><b>16</b></i>	-37.5503	-37.5503	-37.57	-37.57
<i><b>18</b></i>	-37.3502	-37.3502	-37.5747	-37.5747
<i><b>20</b></i>	-37.5277	-37.5277	-37.5984	-37.5984
<i><b>22</b></i>	-37.5875	-37.5875	-37.5836	-37.5836
<i><b>24</b></i>	-37.1856	-37.1856	-37.4058	-37.4058

Table (A.4): NLOS 14 GHz Averaged Real Measured Data at 150° and 180° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 150° AoA (dBm)</i>		<i>Received Power at 180° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-36.2369	-36.2369	-36.9212	-36.9212
<i>4</i>	-37.4688	-37.4688	-36.0121	-36.0121
<i>6</i>	-37.5272	-37.5272	-37.1118	-37.1118
<i>8</i>	-36.8349	-36.8349	-36.8384	-36.8384
<i>10</i>	-37.131	-37.131	-36.6527	-36.6527
<i>12</i>	-37.5901	-37.5901	-37.1991	-37.1991
<i>14</i>	-37.5121	-37.5121	-36.3807	-36.3807
<i>16</i>	-37.5799	-37.5799	-36.3189	-36.3189
<i>18</i>	-37.3451	-37.3451	-37.2481	-37.2481
<i>20</i>	-37.4834	-37.4834	-37.0615	-37.0615
<i>22</i>	-37.5862	-37.5862	-37.5396	-37.5396
<i>24</i>	-37.5828	-37.5828	-37.5101	-37.5101

Table (A.5): NLOS 14 GHz Averaged Real Measured Data at 210° and 240° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 210° AoA (dBm)</i>		<i>Received Power at 240° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-37.3756	-37.3756	-36.2775	-36.2775
<i>4</i>	-37.3178	-37.3178	-36.8062	-36.8062
<i>6</i>	-37.3735	-37.3735	-37.2766	-37.2766
<i>8</i>	-37.2047	-37.2047	-37.4697	-37.4697
<i>10</i>	-37.5879	-37.5879	-37.5444	-37.5444
<i>12</i>	-37.5645	-37.5645	-37.4199	-37.4199
<i>14</i>	-37.4512	-37.4512	-37.4546	-37.4546
<i>16</i>	-37.4752	-37.4752	-37.5939	-37.5939
<i>18</i>	-37.1328	-37.1328	-37.412	-37.412
<i>20</i>	-37.5963	-37.5963	-37.0945	-37.0945
<i>22</i>	-37.6117	-37.6117	-37.3795	-37.3795
<i>24</i>	-37.5645	-37.5645	-37.5823	-37.5823

Table (A.6): NLOS 14 GHz Averaged Real Measured Data at 270° and 300° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 270° AoA (dBm)</i>		<i>Received Power at 300° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-36.0783	-36.0783	-34.2597	-34.2597
<i>4</i>	-36.459	-36.459	-36.1293	-36.1293
<i>6</i>	-37.4478	-37.4478	-36.8925	-36.8925
<i>8</i>	-37.4003	-37.4003	-36.8149	-36.8149
<i>10</i>	-37.5201	-37.5201	-37.4536	-37.4536
<i>12</i>	-37.5377	-37.5377	-37.2316	-37.2316
<i>14</i>	-37.5842	-37.5842	-37.1453	-37.1453
<i>16</i>	-37.4671	-37.4671	-37.5256	-37.5256
<i>18</i>	-37.2049	-37.2049	-37.6351	-37.6351
<i>20</i>	-37.1702	-37.1702	-37.242	-37.242
<i>22</i>	-37.428	-37.428	-37.6062	-37.6062
<i>24</i>	-37.566	-37.566	-37.5762	-37.5762

Table (A.7): NLOS 14 GHz Averaged Real Measured Data at 330°AoA

<i><b>Tx-Rx Separation Distance (m)</b></i>	<i><b>Received Power at 330° AoA (dBm)</b></i>	
	<i><b>Tx Antenna height (1.6 m)</b></i>	<i><b>Tx Antenna height (2.3 m)</b></i>
<i><b>2</b></i>	-30.7288	-30.7288
<i><b>4</b></i>	-32.9187	-32.9187
<i><b>6</b></i>	-34.1064	-34.1064
<i><b>8</b></i>	-35.7479	-35.7479
<i><b>10</b></i>	-36.0053	-36.0053
<i><b>12</b></i>	-34.7062	-34.7062
<i><b>14</b></i>	-37.2426	-37.2426
<i><b>16</b></i>	-37.0074	-37.0074
<i><b>18</b></i>	-37.4113	-37.4113
<i><b>20</b></i>	-35.2386	-35.2386
<i><b>22</b></i>	-36.9572	-36.9572
<i><b>24</b></i>	-36.9005	-36.9005

**Appendix B: Averaged LOS and NLOS real measured data at 18 GHz frequency band**

Table (B.1): LOS 18 GHz Averaged Real Measured Data

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<i>2</i>	-19.3961	-26.9725
<i>4</i>	-25.5372	-27.8724
<i>6</i>	-29.6856	-29.0984
<i>8</i>	-32.2146	-29.8421
<i>10</i>	-32.3361	-32.3736
<i>12</i>	-37.6352	-33.4309
<i>14</i>	-35.1406	-34.1056
<i>16</i>	-35.3779	-36.7387
<i>18</i>	-36.7532	-39.8744
<i>20</i>	-36.7763	-40.981
<i>22</i>	-35.714	-40.0876
<i>24</i>	-36.6121	-39.1943

Table (B.2): NLOS 18 GHz Averaged Real Measured Data at 30° and 60° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 30° AoA (dBm)</i>		<i>Received Power at 60° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-29.8907	-29.8907	-35.894	-35.894
<i>4</i>	-34.3654	-34.3654	-38.4292	-38.4292
<i>6</i>	-35.8463	-35.8463	-39.5216	-39.5216
<i>8</i>	-38.4858	-38.4858	-40.216	-40.216
<i>10</i>	-37.8956	-37.8956	-40.42	-40.42
<i>12</i>	-39.948	-39.948	-40.803	-40.42
<i>14</i>	-38.4363	-38.4363	-40.1877	-40.1877
<i>16</i>	-39.3119	-39.3119	-41.1549	-41.1549
<i>18</i>	-39.2954	-39.2954	-41.0129	-41.0129
<i>20</i>	-41.0584	-41.0584	-41.2644	-41.2644
<i>22</i>	-38.1502	-38.1502	-40.0415	-40.0415
<i>24</i>	-39.7308	-39.7308	-40.6857	-40.6857

Table (B.3): NLOS 18 GHz Averaged Real Measured Data at 90° and 120° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 90° AoA (dBm)</i>		<i>Received Power at 120° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-38.6229	-38.6229	-39.6282	-39.6282
<i>4</i>	-40.7194	-40.7194	-40.9947	-40.9947
<i>6</i>	-40.8449	-40.8449	-40.7162	-40.7162
<i>8</i>	-40.9071	-40.9071	-41.1488	-41.1488
<i>10</i>	-41.2632	-41.2632	-41.0192	-41.0192
<i>12</i>	-41.1755	-41.1755	-41.2354	-41.2354
<i>14</i>	-41.1765	-41.1765	-40.8786	-40.8786
<i>16</i>	-41.2922	-41.2922	-41.2576	-41.2576
<i>18</i>	-41.1123	-41.1123	-41.1483	-41.1483
<i>20</i>	-41.1063	-41.1063	-41.2996	-41.2996
<i>22</i>	-40.9683	-40.9683	-41.0559	-41.0559
<i>24</i>	-41.1526	-41.1526	-41.2313	-41.2313

Table (B.4): NLOS 18 GHz Averaged Real Measured Data at 150° and 180° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 150° AoA (dBm)</i>		<i>Received Power at 180° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-39.8307	-39.8307	-39.699	-39.699
<i>4</i>	-40.9772	-40.9772	-40.4033	-40.4033
<i>6</i>	-41.0138	-41.0138	-40.014	-40.014
<i>8</i>	-41.1019	-41.1019	-39.6162	-39.6162
<i>10</i>	-41.159	-41.159	-39.7837	-39.7837
<i>12</i>	-41.1586	-41.1586	-40.9684	-40.9684
<i>14</i>	-41.0506	-41.0506	-41.0327	-41.0327
<i>16</i>	-41.2463	-41.2463	-41.2906	-41.2906
<i>18</i>	-41.2133	-41.2133	-39.9996	-39.9996
<i>20</i>	-41.3279	-41.3279	-40.0858	-40.0858
<i>22</i>	-41.147	-41.147	-40.3796	-40.3796
<i>24</i>	-41.2297	-41.2297	-40.5655	-40.5655

Table (B.5): NLOS 18 GHz Averaged Real Measured Data at 210° and 240° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 210° AoA (dBm)</i>		<i>Received Power at 240° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-40.6287	-40.6287	-39.3284	-39.3284
<i>4</i>	-40.8201	-40.8201	-40.4492	-40.4492
<i>6</i>	-41.042	-41.042	-40.9888	-40.9888
<i>8</i>	-41.2369	-41.2369	-40.7909	-40.7909
<i>10</i>	-41.3826	-41.3826	-41.2029	-41.2029
<i>12</i>	-41.3826	-41.3826	-41.2029	-41.2029
<i>14</i>	-41.0698	-41.0698	-41.1843	-41.1843
<i>16</i>	-41.3008	-41.3008	-41.1357	-41.1357
<i>18</i>	-41.0495	-41.0495	-41.0801	-41.0801
<i>20</i>	-41.0194	-41.0194	-41.4114	-41.4114
<i>22</i>	-41.3366	-41.3366	-41.1636	-41.1636
<i>24</i>	-41.2606	-41.2606	-41.1775	-41.1775

Table (B.6): NLOS 18 GHz Averaged Real Measured Data at 270° and 300° AoAs

<i>Tx-Rx Separation</i> <i>Distance (m)</i>	<i>Received Power at 270° AoA (dBm)</i>		<i>Received Power at 300° AoA (dBm)</i>	
	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>	<i>Tx Antenna</i> <i>height (1.6 m)</i>	<i>Tx Antenna</i> <i>height (2.3 m)</i>
<i>2</i>	-38.7568	-38.7568	-41.2604	-41.2604
<i>4</i>	-39.9769	-39.9769	-41.0348	-41.0348
<i>6</i>	-40.7907	-40.7907	-41.1611	-41.1611
<i>8</i>	-40.7317	-40.7317	-40.957	-40.957
<i>10</i>	-41.1396	-41.1396	-41.1884	-41.1884
<i>12</i>	-41.1973	-41.1973	-41.117	-41.117
<i>14</i>	-41.052	-41.052	-41.1188	-41.1188
<i>16</i>	-41.1663	-41.1663	-41.2401	-41.2401
<i>18</i>	-41.1325	-41.1325	-41.14	-41.14
<i>20</i>	-41.3751	-41.3751	-41.3028	-41.3028
<i>22</i>	-40.9978	-40.9978	-41.0822	-41.0822
<i>24</i>	-41.3209	-41.3209	-41.3378	-41.3378

Table (B.7): NLOS 18 GHz Averaged Real Measured Data at 330°AoA

<i><b>Tx-Rx Separation Distance (m)</b></i>	<i><b>Received Power at 330° AoA (dBm)</b></i>	
	<i><b>Tx Antenna height (1.6 m)</b></i>	<i><b>Tx Antenna height (2.3 m)</b></i>
<i><b>2</b></i>	-37.246	-37.246
<i><b>4</b></i>	-39.6336	-39.6336
<i><b>6</b></i>	-40.0052	-40.0052
<i><b>8</b></i>	-39.2768	-39.2768
<i><b>10</b></i>	-40.9589	-40.9589
<i><b>12</b></i>	-41.1891	-41.1891
<i><b>14</b></i>	-40.8349	-40.8349
<i><b>16</b></i>	-40.828	-40.828
<i><b>18</b></i>	-40.8703	-40.8703
<i><b>20</b></i>	-41.0182	-41.0182
<i><b>22</b></i>	-41.2438	-41.2438
<i><b>24</b></i>	-40.7897	-40.7897

**Appendix C: Averaged LOS and NLOS real measured data at 22 GHz frequency band**

Table (C.1): LOS 22 GHz Averaged Real Measured Data

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-20.5883	-34.3158
<b>4</b>	-25.4706	-28.8864
<b>6</b>	-28.2274	-28.9466
<b>8</b>	-30.5564	-31.0983
<b>10</b>	-30.8748	-32.7576
<b>12</b>	-31.5676	-33.0946
<b>14</b>	-33.0764	-33.9076
<b>16</b>	-32.9663	-35.6658
<b>18</b>	-32.8518	-33.5877
<b>20</b>	-33.6026	-37.0341
<b>22</b>	-36.9381	-34.2235
<b>24</b>	-38.7736	-32.2008

Table (C.2): NLOS 22 GHz Averaged Real Measured Data at 30° and 60° AoAs

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power at 30° AoA (dBm)</i>		<i>Received Power at 60° AoA (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<i>2</i>	-32.6456	-32.6456	-34.8072	-34.8072
<i>4</i>	-34.5147	-34.5147	-35.7021	-35.7021
<i>6</i>	-35.7216	-35.7216	-36.713	-36.713
<i>8</i>	-36.2883	-36.2883	-36.7372	-36.7372
<i>10</i>	-36.3463	-36.3463	-37.0305	-37.0305
<i>12</i>	-36.2224	-36.2224	-37.0168	-37.0168
<i>14</i>	-35.9854	-35.9854	-37.1399	-37.1399
<i>16</i>	-36.9318	-36.9318	-37.0905	-37.0905
<i>18</i>	-36.8798	-36.8798	-36.8996	-36.8996
<i>20</i>	-36.5678	-36.5678	-37.1595	-37.1595
<i>22</i>	-37.0281	-37.0281	-37.095	-37.095
<i>24</i>	-36.9846	-36.9846	-37.1348	-37.1348

Table (C.3): NLOS 22 GHz Averaged Real Measured Data at 90° and 120°AoAs

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power at 90° AoA (dBm)</i>		<i>Received Power at 120° AoA (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-36.308	-36.308	-36.6666	-36.6666
<b>4</b>	-36.7196	-36.7196	-36.812	-36.812
<b>6</b>	-37.0411	-37.0411	-37.1717	-37.1717
<b>8</b>	-37.0575	-37.0575	-37.0418	-37.0418
<b>10</b>	-37.1566	-37.1566	-37.1693	-37.1693
<b>12</b>	-37.1193	-37.1193	-37.0319	-37.0319
<b>14</b>	-37.1709	-37.1709	-37.1233	-37.1233
<b>16</b>	-37.1016	-37.1016	-37.1411	-37.1411
<b>18</b>	-37.0692	-37.0692	-37.0729	-37.0729
<b>20</b>	-37.1505	-37.1505	-37.1802	-37.1802
<b>22</b>	-37.1408	-37.1408	-37.0323	-37.0323
<b>24</b>	-37.2454	-37.2454	-37.1466	-37.1466

Table (C.4): NLOS 22 GHz Averaged Real Measured Data at 150° and 180° AoAs

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power at 150° AoA (dBm)</i>		<i>Received Power at 180° AoA (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-37.0483	-37.0483	-36.9013	-36.9013
<b>4</b>	-36.8506	-36.8506	-37.1494	-37.1494
<b>6</b>	-37.1392	-37.1392	-36.677	-36.677
<b>8</b>	-37.0764	-37.0764	-36.8977	-36.8977
<b>10</b>	-37.1955	-37.1955	-36.9753	-36.9753
<b>12</b>	-37.1243	-37.1243	-36.7631	-36.7631
<b>14</b>	-37.1134	-37.1134	-36.575	-36.575
<b>16</b>	-37.0899	-37.0899	-36.993	-36.993
<b>18</b>	-37.0796	-37.0796	-36.8423	-36.8423
<b>20</b>	-37.0437	-37.0437	-36.8407	-36.8407
<b>22</b>	-37.1217	-37.1217	-36.9036	-36.9036
<b>24</b>	-37.0826	-37.0826	-36.6772	-36.6772

Table (C.5): NLOS 22 GHz Averaged Real Measured Data at 210° and 240° AoAs

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power at 210° AoA (dBm)</i>		<i>Received Power at 240° AoA (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-36.8729	-36.8729	-36.7995	-36.7995
<b>4</b>	-36.8797	-36.8797	-37.1478	-37.1478
<b>6</b>	-36.9942	-36.9942	-37.1737	-37.1737
<b>8</b>	-37.1636	-37.1636	-37.134	-37.134
<b>10</b>	-37.1578	-37.1578	-37.1491	-37.1491
<b>12</b>	-37.1174	-37.1174	-37.0926	-37.0926
<b>14</b>	-37.0289	-37.0289	-37.0891	-37.0891
<b>16</b>	-37.1069	-37.1069	-37.1432	-37.1432
<b>18</b>	-37.078	-37.078	-37.1426	-37.1426
<b>20</b>	-37.03	-37.03	-37.1092	-37.1092
<b>22</b>	-37.0536	-37.0536	-37.1148	-37.1148
<b>24</b>	-37.1719	-37.1719	-37.1077	-37.1077

Table (C.6): NLOS 22 GHz Averaged Real Measured Data at 270° and 300° AoAs

<i>Tx-Rx Separation Distance (m)</i>	<i>Received Power at 270° AoA (dBm)</i>		<i>Received Power at 300° AoA (dBm)</i>	
	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>	<i>Tx Antenna height (1.6 m)</i>	<i>Tx Antenna height (2.3 m)</i>
<b>2</b>	-36.7863	-36.7863	-35.2777	-35.2777
<b>4</b>	-37.1623	-37.1623	-35.9093	-35.9093
<b>6</b>	-37.1564	-37.1564	-36.7163	-36.7163
<b>8</b>	-37.0198	-37.0198	-36.644	-36.644
<b>10</b>	-37.1072	-37.1072	-36.8838	-36.8838
<b>12</b>	-37.1498	-37.1498	-36.8681	-36.8681
<b>14</b>	-37.0156	-37.0156	-36.9396	-36.9396
<b>16</b>	-37.1263	-37.1263	-37.1297	-37.1297
<b>18</b>	-37.0454	-37.0454	-36.8425	-36.8425
<b>20</b>	-37.0771	-37.0771	-37.069	-37.069
<b>22</b>	-37.1447	-37.1447	-37.1474	-37.1474
<b>24</b>	-37.1294	-37.1294	-37.0182	-37.0182

Table (C.7): NLOS 22 GHz Averaged Real Measured Data at 330° AoA

<i><b>Tx-Rx Separation Distance (m)</b></i>	<i><b>Received Power at 330° AoA (dBm)</b></i>	
	<i><b>Tx Antenna height (1.6 m)</b></i>	<i><b>Tx Antenna height (2.3 m)</b></i>
<i><b>2</b></i>	-32.6274	-32.6274
<i><b>4</b></i>	-34.9093	-34.9093
<i><b>6</b></i>	-36.0243	-36.0243
<i><b>8</b></i>	-36.2	-36.2
<i><b>10</b></i>	-36.6493	-36.6493
<i><b>12</b></i>	-36.365	-36.365
<i><b>14</b></i>	-36.8624	-36.8624
<i><b>16</b></i>	-36.9767	-36.9767
<i><b>18</b></i>	-36.4959	-36.4959
<i><b>20</b></i>	-37.0468	-37.0468
<i><b>22</b></i>	-36.7535	-36.7535
<i><b>24</b></i>	-36.7015	-36.7015