IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 14, 2020, accepted October 19, 2020, date of publication November 2, 2020, date of current version November 12, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035209

Artificial Neural Network Based Path Loss Prediction for Wireless Communication Network

LINA WU[1,2], (Student Member, IEEE), DANPING HE[1,2], (Member, IEEE), BO AI[1,2], (Senior Member, IEEE), JIAN WANG[1,3], (Member, IEEE), HANG QI[4], KE GUAN[1,2], (Senior Member, IEEE), AND ZHANGDUI ZHONG[1,2], (Senior Member, IEEE)

[1]State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
[2]Beijing Engineering Research Center of High-speed Railway Broadband Mobile Communications, Beijing 100044, China
[3]Beijing Engineering Research Center of EMC and GNSS Technology for Rail Transportation, Beijing 100044, China
[4]China Mobile Group Design Institute Co., Ltd., Beijing 100080, China

Corresponding author: Bo Ai (boai@bjtu.edu.cn)

This work was supported in part by the Fundamental Research Funds for Central Universities under Grant 2020JBZD005; in part by the NSFC under Grant 61901029, Grant 61771036, Grant 61725101, and Grant U1834210; in part by the Royal Society Newton Advanced Fellowship under Grant 6196113039 and Grant NA191006; in part by the State Key Lab of Rail Traffic Control and Safety Project under Grant RCS2020ZT011, Grant RCS2020ZZ005, and Grant RCS2020ZT010; and in part by the Science and Technology Key Project of Guangdong Province, China, under Grant 2019B010157001.

ABSTRACT Accurate path loss (PL) prediction is essential for predicting transmitter coverage and optimizing wireless network performance. Traditional PL models are difficult to cope with the development trend of diversity, time-varying and mass wireless channels. In this work, the most widely used multilayer perceptron (MLP) neural network in artificial neural network (ANN) is employed to accurately predict PL. Three types of environmental features are defined and extracted, which describe the propagation environment only by considering limited environmental types instead of complex 3D environment modeling. Principal component analysis (PCA) is used to generate the low-dimensional environmental features, and eliminate redundant information among similar environmental types. Moreover, the information of base station (BS) and the receiver (Rx), including 3D locations, frequency, the transmitted power of BS, the antenna information, the feeder loss, and the received power of all the locations are obtained from the measurements. Different environmental features are combined with the information of BS and Rx to construct seven datasets for PL prediction models based on MLP neural networks. The impacts of the number of neurons in the hidden layer, the number of hidden layers, the number of training samples, and environmental features on PL prediction models are explored by considering the absolute value of mean error (AME), the mean absolute error (MAE), the standard deviation (STD) of error, the correlation coefficient, and the time ratio, respectively. This work aims to understand the propagation characteristics of radio waves, which can provide a theoretical basis for wireless network optimization and communication system design.

INDEX TERMS Multilayer perceptron neural network, environmental feature, principal component analysis, path loss prediction.

I. INTRODUCTION

With the rapid growth of personal communication devices and the demand of consumers for ubiquitous data access, communication systems have become an indispensable part of our daily life. In order to better optimize and plan communication systems to meet the needs of users, we must understand how radio waves propagate in realistic environments. Channel modeling is the process of characterizing the propagation law of radio waves in realistic environments, and provides strong theoretical guidance for the design, deployment and optimization of communication systems. With the rapid development of communication systems, especially since the introduction of the fifth generation mobile networks

The associate editor coordinating the review of this manuscript and approving it for publication was Jonghoon Kim.

(5G), the channel data presents the characteristics of mass, high speed and diversity, which makes the planning and optimization of 5G and future communication systems face severe challenges [1]. Therefore, it is urgent to establish accurate and efficient channel model.

Generally, traditional channel modeling methods are mainly divided into two categories: stochastic channel models and deterministic channel models [2]. Stochastic channel models mainly include geometry-based stochastic models and non-geometrical stochastic models [3]–[5]. The former is a mathematical model based on stochastic geometry, and the latter is a parametric model based on the channel measurements [6]–[8]. Stochastic channel modeling methods are based on extensive propagation measurements in various environments at different frequency bands. The typical example is that, the path loss models in 3GPP 38.901 are established through different effective stochastic channel models based on a large number of measurement campaigns [9], [10]. Since the generalization ability of stochastic channel modeling methods is relatively limited, it can not accurately predict wave propagation characteristics, so it is necessary to study different cases for each specific environment type [11]. The deterministic channel modeling uses architectural plans or topographical maps of the propagation environment to predict wave propagation characteristics based on geometric optical theory [12], [13] and electromagnetic wave propagation theory. The deterministic channel models mainly include the ray-tracing (RT) and the finite difference time domain (FDTD) [14]. Among them, RT is a typical deterministic modeling method, which can accurately predict multipath propagation with three-dimensional (3D) environment model and communication system configuration. It has been successfully applied to some related work of wireless channel analysis [15]–[21]. RT-based deterministic channel modeling approach highly depends on the accuracy of 3D environment model including information about geometry and materials. However, it is difficult to obtain accurate geometric and material information of objects in large-scale and complex environments, which limits the versatility of the deterministic channel models.

The wireless channel is a time-varying nonlinear system, which covers multi-dimensional information such as time domain, spatial domain, and frequency domain. Traditional channel modeling methods are difficult to cope with the development trend of diversity, time-varying and mass wireless channels. Machine learning has powerful learning and prediction capabilities, and can automatically learn the structural relationship between data in complex environments to approximating nonlinear systems. Moreover, machine learning is good at mining the information contained in high-dimensional data, which can significantly increase the speed of data processing. Based on these advantages, machine learning is regarded as a powerful tool for analyzing the propagation process of radio waves and constructing channel models by using measurement data. Meanwhile, with the rapid development of big data, cloud computing and artificial intelligence, intelligent communication is the mainstream trend of 5G and future wireless communication systems. The combination of machine learning and channel modeling is a challenging topic, and the research work in this field is still in the preliminary exploration stage.

Path loss (PL) is an important parameter in channel modeling, which represents the reduction in power density of a radio wave as it propagates through the channel. There are two approaches to generate PL models [22]: ones that have some anchor based on the physics of transmitted power close to the antenna, such as the close-in free space reference distance (CI) PL model [23]–[25], and ones that do a mathematical curve fitting over the dataset without any physical anchor to the transmitted power, such as the $\alpha - \beta$ (A-B) PL model [16], [17]. Artificial neural network (ANN) is an important branch of machine learning, which has the advantages of adaptability, self-learning and robustness, and can be used to predict PL. Nowadays, some achievements have been made in PL prediction based on ANN [26].

In [27], a PL prediction model based on radial basis function neural network is established, in which the height of the receiving antenna and the transmitting antenna and the distance between them are taken as the input set. In [28], different PL models are established for line-of-sight (LOS) and non-line-of-sight (NLOS). Compared with the ANN-based model established for LOS, the model for NLOS combines the diffraction loss calculated by the traditional PL model with ANN, and the comparative analysis verifies that the hybrid PL model combining ANN and the traditional PL model is more accurate. Reference [29] proposes a machine learning framework for modeling PL using a combination of three key techniques: ANN, Gaussian process, and principle component analysis (PCA) [30]. Compared with traditional PL models, the proposed model is more accurate and flexible, which will be beneficial for the site-specific design of wireless sensor network with high reliability. Reference [31] establishes a PL model based on convolutional neural networks and satellite images, and the results show that the proposed model is capable of improving PL prediction at unseen locations, compared to traditional PL models, such as RT and stochastic channel models. Sotiroudis *et al.* proposes an alternative procedure for the PL prediction in urban environments based on ANN, the procedure combines differential evolution algorithm with Levenberg-Marquardt back propagation (BP) [32] algorithm, which significantly improves the overall system performance [33]–[35]. Popoola *et al.* proposes heuristic methods, geospatial and stochastic channel models to predict PL in the VHF and UHF bands for specified routes. The comparative analysis shows that ANN-based PL model has better prediction accuracy and generalization ability than stochastic channel models [36]–[38].

In the current ANN-based PL model, signal parameters that characterize the channel characteristics are often used as the input set of ANNs, and only a few studies consider the impact of environments on channel modeling. However, environments are the most important components of radio

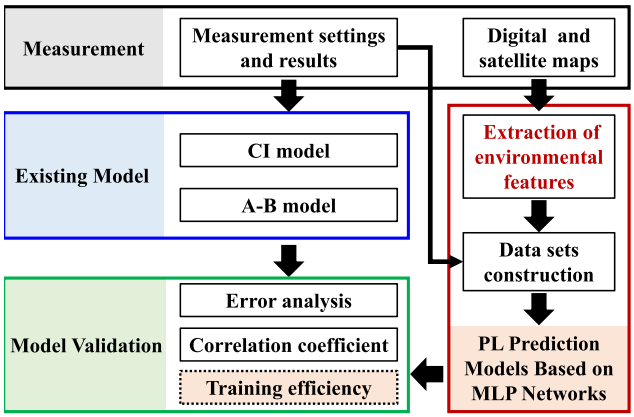


FIGURE 1. The workflow of PL prediction models.

wave propagation. In the absence of information about the geometry and materials of propagation environments, environmental types can be helpful for the description of environments [11]. Therefore, in this paper, we will conduct a comprehensive study on PL prediction models based on multilayer perceptron (MLP) neural networks and environmental types, and compare it with the CI PL model and the A-B PL model. The workflow of this paper is shown in Fig. 1, and the contributions of this work are composed as follows:

- Measurement campaigns were conducted at 2.5 GHz, in three different regions of Hangzhou, China. Based on the measurement data, the PL predicted results are obtained by the CI PL model and the A-B PL model. Then the performance evaluation of the CI PL model and the A-B PL model is achieved by comparing the measured results and the predicted results obtained by the CI PL model and the A-B PL model.
- Environmental features extracted from the limited environmental types are used to describe the radio wave propagation environment, which replaces the complex 3D environment modeling. Moreover, multiple datasets used to train and test PL prediction models are constructed by different combinations of environmental features and the information of base stations (BSs) and receiver (Rx).
- PL prediction models based on MLP neural networks with different architectures are established. Meanwhile, the impact of environmental features, ANN architectures, the dimension and percentage of training samples on PL prediction models are analyzed for model validation. PL prediction models proposed in this paper can provide a theoretical basis for predicting transmitter coverage and interference, optimizing and planning wireless network performance.

The rest of this paper is organized as follows. Section II introduces measurement campaigns and the analysis of the predicted results obtained by the CI PL model and the A-B PL model. PL prediction models based on MLP neural networks and the analysis of the predicted results obtained by these PL models are introduced in Section III. Finally, conclusions and future work are drawn in Section IV.

NOTATION

$R_{Clutter}$	Raster resolution
f_c	Frequency
L_{BS}	3D location of BS
L_{Rx}	3D location of Rx
d_{3D}	Distance between BS and Rx
L_f	Feeder loss
A_z	Azimuth
D_e	Decline angle
G_{BS}	Antenna gain of BS
G_{Rx}	Antenna gain of Rx
P_{BS}	Transmitted power
P_{Rx}	Received power
S_{Rx}	Receiver sensitivity
PL	Measured results of PL
PL'	Predicted results of PL
e	Environmental type
n_e	Number of the environmental type e
$N_{Clutter}$	Total number of rasters in the corresponding area
p_e	Percentage of the environmental type e
A	Slope in the A-B PL model
B	Interception in the A-B PL model
χ_σ	Shadow fading
$H_{p\times p}$	Transformation matrix
L	Straight line
R	Rectangle
F_L	Linear environmental feature
λ_L	Low-dimensional linear environmental feature
F_{R1}	Unweighted rectangular environmental feature
ϕ_{R1}	Low-dimensional unweighted rectangular feature
F_{R2}	Weighted rectangular environmental feature
φ_{R2}	Low-dimensional weighted rectangular feature
δ	Eigenvalue
V	Eigenvector
$C_{p\times p}$	Covariance matrix
τ	Contribution rate
Γ	Cumulative contribution rate
$D_{1\sim7}$	Dataset for training and testing
AME	Absolute value of mean error
MAE	Mean absolute error
STD	Standard deviation
T	Time ratio
r	Correlation coefficient
N_h	Number of neurons in the hidden layer
N_{input}	Number of neurons in the input layer
N_{output}	Number of neurons in the output layer
lr	Learning rate
P_{train}	Percentage of training samples
f_{train}	Training function
f_{loss}	Loss function
E_{goal}	Goal error

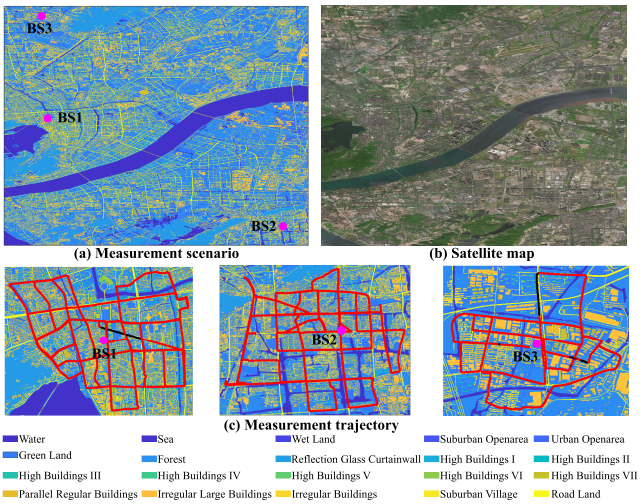


FIGURE 2. Measurement scenario and trajectory.

II. MEASUREMENT AND PATH LOSS ANALYSIS

A. MEASUREMENT CAMPAIGN

In order to obtain the first-hand information of the channels, channel measurements are conducted in three different regions of Hangzhou, China, in the 2.5 GHz frequency bands. Fig. 2(a) shows the rasterized digital map of this measurement scenario, which is obtained by dividing the digital map into many small rasters of size $R_{Clutter} \times R_{Clutter}$, where $R_{Clutter}$ is the raster resolution. Fig. 2(b) is the satellite aerial image of this scenario. Statistics show that there are 20 environmental types that affect radio propagation in the whole area, including water, sea, wet land, suburban openarea, urban openarea, green land, forest, reflection glass curtainwall, high buildings (I-VII) with different heights, parallel regular buildings, irregular large buildings, irregular buildings, suburban village and road land. Three BSs with heights of 62 m, 30 m and 41.7 m are deployed at different locations in this scenario. Rx is deployed on a vehicle, with a height of approximately 2 m. Channel measurements are carried out around three BSs, and the trajectory of the vehicle is shown in Fig. 2(c). Detailed measurement setup, such as the location of BS L_{BS}, the location of Rx L_{Rx}, the antenna pattern of BS and Rx, the feeder loss L_f, the transmitted power of BS P_{BS}, and the receiver sensitivity S_{Rx}, are shown in Table 1.

Finally, the received power P_{Rx} that characterizes the radio wave propagation characteristics is collected at 23892 different locations in the whole area, of which there are 10184, 9988, and 3720 sampling points located in the measurement areas associated with BS1, BS2, and BS3 respectively. Therefore, the measured results of PL under the above measurement setup can be calculated according to the following formula:

$$PL\,[dB] = P_{BS} - L_f + G_{BS}\,(A_z, D_e) + G_{Rx} - P_{Rx} \quad (1)$$

where A_z, D_e, and G_{BS} are the azimuth, the decline angle, and the antenna gain of BS, which are calculated based on the antenna pattern and the locations of BS and Rx. G_{Rx} and P_{Rx} are the antenna gain and the received power of Rx respectively, where $G_{Rx} = 0$ dBi.

TABLE 1. Measurement setup.

Frequency		$f_c = 2.5$ GHz
BS	3D Location	$L_{BS} = (x_{BS}, y_{BS}, z_{BS})$
	Power	$P_{BS} = 42$ dBm
	Antenna	Omnidirectional antenna Maximum antenna gain: 4 dBi
	Height	BS1∼3: $z_{BS} = [62, 30, 41.7]$ m
	Feeder Loss	BS1∼3: $L_f = [3.3, 3.3, 6.2]$ dB/m
Rx	3D Location	$L_{Rx} = (x_{Rx}, y_{Rx}, z_{Rx})$
	Antenna	Omnidirectional antenna Antenna gain: 0 dBi
	Height	$z_{Rx} = 2$ m
	Sensitivity	$S_{Rx} = $ -135 dBm

B. PATH LOSS MODEL

The CI PL model and the A-B PL model are two commonly used PL models. In this section, the CI PL model and the A-B PL model are used to fit the measurement data for predicting PL. Meanwhile, the absolute value of mean error (AME), the mean absolute error (MAE), the standard deviation (STD) of error, and the correlation coefficient r between the measured results and the predicted results are used as metrics to evaluate the performance of the CI PL model and the A-B PL model.

The CI PL model is based on fundamental principles of wireless propagation and Friis' model, which ties path loss at any frequency to the physical free space path loss (FSPL) at 1 m according to Friis' free space equation. In other words, the CI PL model accounts for the frequency dependency of PL by using one meter reference distance based on Friis' law:

$$PL^{CI}\,(f_c, d_{3D})\,[dB] = FSPL\,(f_c, 1\ m) + 10n log_{10}(d_{3D}) + \chi_\sigma^{CI} \quad (2)$$

where χ_σ^{CI} represents the shadow fading (SF) that is modeled as a zero-mean Gaussian random variable with a standard deviation σ in dB. n represents the PL exponent (PLE) obtained by minimizing the error of the measurement data, which offers insight into PL based on the environment, with a value of 2 in free space. d_{3D} is the 3D distance between BS and Rx, $d_{3D} > 1$ m, and $FSPL(f_c, 1\ m)$ is the FSPL at frequency f_c in GHz at 1 m and is calculated by [39], [40]:

$$FSPL\,(f_c, 1\ m)\,[dB] = 20 log_{10}\left(\frac{4\pi f_c}{c}\right) \quad (3)$$

where c is the speed of light, $3 \times 10^8\ m/s$.

The A-B PL model employs the best fit line to the measured data using a least-squares regression method, without consideration for the close-in free-space propagation that always occurs in practice near an antenna out in the open. Meanwhile, the A-B PL model is only applicable under certain conditions, the model parameters vary greatly when it is

applied in different frequency bands or scenarios, resulting in a lack of solid physical significance.

$$PL^{AB}\,[dB] = Alog_{10}\left(\frac{d_{3D}}{1\ m}\right) + B + \chi_{\sigma}^{AB} \tag{4}$$

where A and B represent the slope and the interception, respectively, d_{3D} is the 3D distance between BS and Rx, and χ_{σ}^{AB} is SF which can be expressed as a zero-mean Gaussian random variable with a standard deviation σ in dB.

AME is the absolute value of the mean of all errors between the measured results and the predicted results, and MAE is the mean of all absolute errors between the measured results and the predicted results. The former characterizes the performance of the prediction model for all observation samples as a whole, while the latter measures the performance of the prediction model for each observation sample. STD is a characteristic quantity which can represent the dispersion degree of all errors between the measured results and the predicted results. The correlation coefficient r is a statistical metric of the dependence or association between the measured results and the predicted results, which takes values between -1 and 1. The greater the absolute value of the correlation coefficient, the greater the correlation between the measured results and the predicted results. The absolute value of r greater than 0.8 indicates that there is a strong correlation between them. AME, MAE, STD, and r in these PL models are defined as

$$AME\,[dB] = \frac{\left|\sum_{i=1}^{N_{sample}} error_i\right|}{N_{sample}} \tag{5}$$

$$MAE\,[dB] = \frac{\sum_{i=1}^{N_{sample}} |error_i|}{N_{sample}} \tag{6}$$

$$STD\,[dB] = \sqrt{\frac{\sum_{i=1}^{N_{sample}} (error_i - \mu_{error})^2}{N_{sample}}} \tag{7}$$

$$r = \frac{\sum_{i=1}^{N_{sample}} (PL_i - \mu_{PL})\left(PL_i' - \mu_{PL'}\right)}{\sqrt{\sum_{i=1}^{N_{sample}} (PL_i - \mu_{PL})^2 \sum_{i=1}^{N_{sample}} \left(PL_i' - \mu_{PL_i'}\right)^2}} \tag{8}$$

where PL and PL' represent the measured results and the predicted results of PL, respectively. N_{sample} is the number of samples, $error = \{error_i = PL_i - PL_i', i = 1, 2, \ldots N_{sample}\}$ is the errors between the measured results PL and the predicted results PL'. μ_{error}, μ_{PL}, and $\mu_{PL'}$ are the mean values of $error$, PL and PL', respectively.

Fig. 3 shows the measured results and the predicted results of PL associated with different BSs, respectively. Table 2 summarizes the parameters in the CI PL model and the A-B PL model, as well as the evaluation metrics, including AME, MAE, STD, and the correlation coefficient r. For the

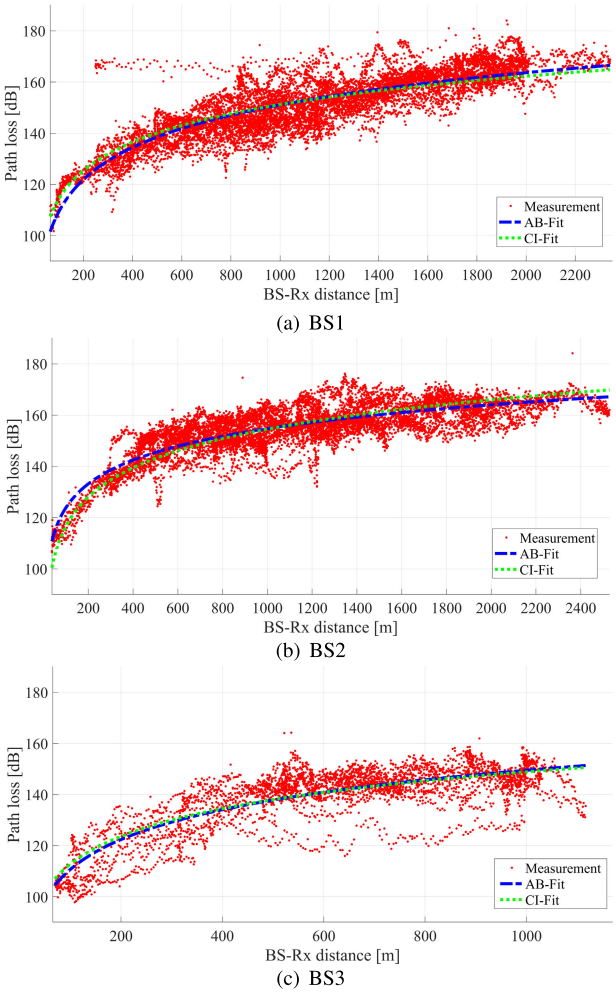


FIGURE 3. Path loss in the measurement scenario with different BS.

TABLE 2. Parameters, errors and correlation coefficients of the CI PL model and the A-B PL model.

Case		BS1	BS2	BS3
CI	n	3.69	3.80	3.61
	σ [dB]	6.72	5.79	7.21
	AME [dB]	0.11	0.18	0.11
	MAE [dB]	7.49	6.52	8.05
	STD [dB]	9.51	8.18	10.19
	r without χ_{σ}^{CI}	0.86		
	r with χ_{σ}^{CI}	(0.73, 0.75), Mean = 0.739		
A-B	A	41.77	30.85	38.75
	B	25.79	62.13	33.24
	σ [dB]	6.61	5.46	7.18
	AME [dB]	0.05	0.04	0.10
	MAE [dB]	7.33	6.14	8.00
	STD [dB]	9.35	7.72	10.14
	r without χ_{σ}^{AB}	0.87		
	r with χ_{σ}^{AB}	(0.74, 0.75), Mean = 0.749		

TABLE 3. Environmental type and percentage.

Statistical Scope		Whole Area		Area of BS1		Area of BS2		Area of BS3	
Raster Number $N_{Clutter}$		61117070		432674		467988		150136	
Type		Number n_e	Percentage p_e	Number n_e	Percentage p_e	Number n_e	Percentage p_e	Number n_e	Percentage p_e
Water		7247738	11.86%	24797	5.73%	16343	3.49%	3356	2.24%
Sea		-	-	-	-	-	-	-	-
Wet Land		-	-	-	-	-	-	-	-
Suburban Openarea		2523045	4.13%	-	-	-	-	-	-
Urban Openarea		5368406	8.78%	20482	4.73%	86405	18.46%	29590	19.71%
Green Land		27805648	45.5%	169655	39.21%	273278	58.39%	63720	42.44%
Forest		4503151	7.37%	43987	10.17%	-	-	1094	0.73%
Reflection Glass Curtainwall		33408	0.05%	1067	0.25%	-	-	-	-
High Buildings	I: Hight $\in [100, +\infty)$m	52821	0.09%	1641	0.38%	121	0.03%	166	0.11%
	II: Hight $\in [80, 100)$m	79602	0.13%	2583	0.60%	270	0.06%	242	0.16%
	III: Hight $\in [60, 80)$m	109626	0.18%	3792	0.88%	2025	0.43%	277	0.18%
	IV: Hight $\in [50, 60)$m	132569	0.22%	3927	0.91%	882	0.19%	238	0.16%
	V: Hight $\in [40, 50)$m	107223	0.18%	4514	1.04%	550	0.12%	70	0.05%
	VI: Hight $\in [30, 40)$m	183262	0.30%	6801	1.57%	626	0.13%	354	0.24%
	VII: Hight $\in [20, 30)$m	863905	1.41%	30718	7.10%	7238	1.55%	1784	1.19%
Parallel Regular Buildings		1544697	2.53%	45845	10.60%	4312	0.92%	13499	8.99%
Irregular Large Buildings		3499868	5.73%	10031	2.32%	14070	3.01%	8894	5.92%
Irregular Buildings		985107	1.61%	6434	1.49%	17466	3.73%	7157	4.77%
SubUrban Village		1220720	2.00%	-	-	-	-	-	-
Road Land		4856274	7.95%	56400	13.04%	44402	9.49%	19695	13.12%

reliability of models analysis, we use the CI PL model and the A-B PL model to fit PL for 100 times, and get the mean value of AME, MAE, and STD, and the value range of r. As shown in Table 2, all PLEs n in the CI PL model are greater than that in the FSPL model (which is 2), and all slopes A in the A-B PL model are greater than that in the FSPL model (which is 20). FSPL describes the loss in signal strength of an electromagnetic wave that result from a LOS path through free space. Therefore, n in the CI PL model and A in the A-B PL model show that there are a large number of NLOS paths in the three measurement scenarios. It can be found that, the predicted results PL' obtained by the CI PL model without χ_σ^{CI} and the A-B PL model without χ_σ^{AB} have a strong correlation with the measured results PL ($r = 0.86$ for the CI PL model and $r = 0.87$ for the A-B PL model). However, the correlation coefficient r between the measured results PL and the predicted results PL' obtained by the CI PL model with χ_σ^{CI} (belong to $0.73 - 0.75$ and the mean is 0.739) and that between the measured results PL and the predicted results PL' obtained by the A-B PL model with χ_σ^{AB} (belong to $0.74 - 0.75$ and the mean is 0.749) are less than 0.8. Moreover, although AME between PL and PL' obtained by the CI PL model and the A-B PL model are less than 1 dB, the corresponding MSE (> 6.1 dB) and STD (> 7.7 dB) are relatively large, which indicates that the predicted results of the CI PL model and the A-B PL model are not ideal.

The percentage of different environmental types in the measurement scenario is a key factor affecting radio wave propagation, which is helpful to analyze the characteristics of radio wave propagation. Based on the rasterized digital map, the percentage p_e of each environment type e in the propagation scenario can be defined as

$$p_e = \frac{n_e}{N_{Clutter}} \times 100\% \tag{9}$$

where n_e is the number of the environmental type e, which is equal to the number of rasters of the environment type e in the propagation scenario. $N_{Clutter}$ is the total number of rasters in the propagation scenario, which can be obtained by dividing the range of this scenario by the raster resolution $R_{Clutter}$.

In this paper, $R_{Clutter} = 5$ m. Hence, $N_{Clutter}$ in the whole area and the three measurement areas associated with BS1, BS2 and BS3 are 61117070, 432674, 467988, and 150136, respectively. The number of each environment type in the whole area and the three measurement areas is counted. Then, the percentage of each environment type in the corresponding area can be calculated by (9). The number and percentage of 20 environment types in the whole area and the three measurement areas are shown in Table 3.

Based on Table 3, the proportion of high buildings (I-VII) in the measurement area associated with BS1 is 12.48%, which indicates that a large number of high buildings are distributed in the measurement area associated with BS1, which further illustrates that a large number of NLOS paths exist in this scenario. According to the location of BS1 and Rxs as well as the PL results obtained by (1), it can be known that some points along the line $PL = 160$ dB in Fig. 3(a) are mainly located between high buildings, rather than on the main road, as shown by the black trajectory in the measurement area associated with BS1 in Fig. 2(c). Therefore, the occlusion of high buildings caused deep fading at these points. In contrast, the proportion of high buildings in the measurement area associated with BS3 is 2.09%, which indicates that a small number of high buildings are distributed in the measurement area associated with BS3. Some points with small PL in Fig. 3(c) are mainly located on the main road without a large number of high buildings, as shown by the black trajectory in the measurement area associated with BS3 in Fig. 2(c). The proportion of high buildings in the

measurement area associated with BS2 is 2.51%. Although there are a small number of high buildings in the measurement area associated with BS2, the lower BS height enhances the impact of high buildings on radio wave propagation, resulting in more NLOS paths in the measurement area associated with BS2, resulting in the PLE n of BS2 is greater than that of BS1 and BS3. Consequently, compared with the PL of BS1 and BS3, the PL of BS2 increases more significantly with the increase of distance.

Actually, the CI PL model and the A-B PL model fit the overall trend of the PL of all observation samples with distance as much as possible, but they cannot accurately predict the PL with large difference within the same distance range. Therefore, as shown in Table 2, the predicted results of the CI model and the A-B model for BS1 and BS3 are worse than their predicted results for BS2. Moreover, different measurement scenarios need different CI models and A-B models, which is difficult to cope with the development trend of diversity, time-varying and mass wireless channels. Therefore, it is necessary to establish a PL model to significantly help improve the modeling accuracy and efficiency.

III. PATH LOSS PREDICTION MODEL BASED ON MULTILAYER PERCEPTRON NEURAL NETWORK

In this section, three environmental features are defined and extracted by considering the limited environmental type information in the propagation scenario. PL prediction models are established based on environmental features and MLP neural networks which are the most widely used neural networks [41]. Afterwards, the performance of the PL models is evaluated in terms of environmental features, the network architectures of MLP and training time. The details for PL prediction model are as follows.

A. ENVIRONMENTAL FEATURES

Obstacles are the most important components of physical environment for radio propagation. However, for large-scale scenarios, the complexity of 3D environment modeling and the diversity of materials both lead to the decrease of accuracy and the increase of computational complexity in channel modeling. Therefore, environmental features are defined and extracted by considering the limited environmental type information, which replaces the complex environment modeling, reduces the complexity and improves the efficiency.

Definition 1 (Linear Environmental Features): Suppose there are s environmental types in a given scenario, which are denoted as $e_1, e_2, \ldots, e_s$, respectively. From the measurements, locations of all Rxs and the associated BSs are known, which are denoted as L_{Rx} and L_{BS}, respectively. Hence, the projection points of Rx and BS on the digital map can be obtained, denoted as L'_{Rx} and L'_{BS}, respectively. As shown in Fig.4 (a), the straight line L can be uniquely determined by L'_{Rx} and L'_{BS}, then the ***linear environmental feature*** is defined as a s-dimensional vector F_L:

$$F_L = (l_1, l_2, \ldots, l_s) \tag{10}$$

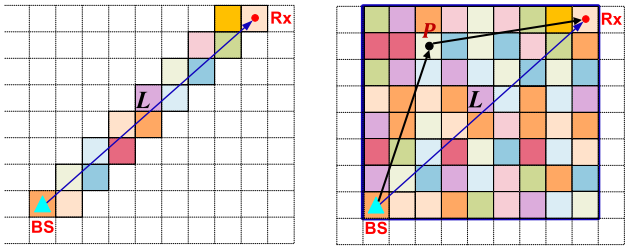


FIGURE 4. Environmental features.

where l_i is the length of the straight line L passing through the corresponding environmental type e_i, $i = 1, 2, \ldots, s$.

Definition 2 (Unweighted and Weighted Rectangular Environmental Features): Suppose there are s environmental types in a given scenario, which are denoted as $e_1, e_2, \ldots, e_s$, respectively. From the measurements, locations of Rx and the associated BS are known, which are denoted as L_{Rx} and L_{BS}, respectively. Hence, the projection points of Rx and BS on the digital map can be obtained, denoted as L'_{Rx} and L'_{BS}, respectively. As shown in Fig.4 (b), a rectangle R with the diagonal line L connecting L'_{Rx} and L'_{BS} can be determined, then the ***unweighted rectangular environmental feature*** is defined as a s-dimensional vector F_{R1}:

$$F_{R1} = (N_1, N_2, \ldots, N_s) \tag{11}$$

where N_i is the number of the corresponding environmental type e_i in the rectangle R.

Similarly, the ***weighted rectangular environmental feature*** is defined as a s-dimensional vector:

$$F_{R2} = (\omega_1, \omega_2, \ldots, \omega_s) \tag{12}$$

where ω_i is a weighted accumulation of the number of the corresponding type e_i in the rectangle R, expressed as

$$\omega_i = \sum_{P \in R, R_P = e_i} \frac{dis\left(L'_{BS}, L'_{Rx}\right)}{dis\left(P, L'_{Rx}\right) + dis\left(P, L'_{BS}\right)} \tag{13}$$

where $dis\left(L'_{BS}, L'_{Rx}\right)$ is the distance between L'_{BS} and L'_{Rx}. Similarly, $dis\left(P, L'_{BS}\right)$ and $dis\left(P, L'_{Rx}\right)$ represent the distance between the sampling point P and L'_{BS} and between the sampling point P and L'_{Rx}, respectively.

B. PRINCIPAL COMPONENT ANALYSIS OF ENVIRONMENTAL FEATURES

Considering the correlation between similar environmental types, the extracted environmental features are transformed to a set of uncorrelated variables by PCA, which eliminates redundant information among environmental features, reduces the feature dimension, and improves the predicted efficiency.

1) PRINCIPAL COMPONENT ANALYSIS

Fundamentally speaking, PCA is a simple dimensionality reduction technique, or feature combination technique [30].

The central idea of this technique is to find an orthogonal transformation to transform a group of variables that may have correlation into a group of linearly unrelated variables, while preserving all the information in initial variables as much as possible. The transformed variables are called principal components.

Suppose there are p variables, and each variable has m observation samples, denoted as $X_i \in \{(x_{i1}, x_{i2}, \ldots, x_{im})', i = 1, 2, \ldots, p\}$. Loosely, PCA can be expressed as

$$\left[Y_1, \ldots, Y_p\right] = \left[X_1, \ldots, X_p\right] H_{p\times p} \tag{14}$$

where $H_{p\times p} \in \left\{\left(h_{ij}\right)_{p\times p}, i, j = 1, 2, \ldots, p\right\}$ is the transformation matrix. The transformed variable $Y_i \in \left\{(y_{i1}, y_{i2}, \ldots, y_{im})', i = 1, 2, \ldots, p\right\}$ is the $i-th$ principal component, which is obtained by the weighted additive combination of all initial variables. The relationship between the components of initial variables and the components of principal components is as follows.

$$y_{ij} = \sum_{k=1}^{p} h_{ki} x_{kj}, \quad i = 1, 2, \ldots, p, j = 1, 2, \ldots, m \tag{15}$$

where $i = 1, 2, \ldots, p; j = 1, 2, \ldots, m$.

Moreover, the transformation matrix $H_{p\times p}$ must satisfy the following conditions.

- The first principal component Y_1 represents more information of all initial variables compared to the second principal component Y_2. Likewise, Y_2 represents more than the third principal component Y_3, and so on. Variance, which measures the dispersion of variables, can be used to represent the information of variables. Therefore, the variance of principal components follows the rule:

$$Var\left(Y_1\right) \geqslant Var\left(Y_2\right) \geq \ldots \geq Var\left(Y_p\right) \tag{16}$$

  where $Var\left(Y_i\right)$ represents the variance of the principal component $Y_i, i = 1, 2, \ldots, p$.

$$\begin{cases} Var\left(Y_i\right) = \dfrac{\sum_{j=1}^{m}\left(y_{ij} - \overline{Y_i}\right)^2}{m-1} \\ \overline{Y_i} = \dfrac{\sum_{j=1}^{m} y_{ij}}{m} \end{cases} \tag{17}$$

- The principal components are independent of each other. In other words, there is no redundant information between different principal components. Covariance provides a measure of the strength of the correlation between two or more sets of random variables. If the variables are correlated in some way, then their covariance will be non-zero. Conversely, if the variables are not correlated, their covariance will be zero. Therefore, the covariance between different principal components is zero, that is,

$$\begin{cases} Cov\left(Y_i, Y_j\right) \equiv 0, \quad i \neq j, \quad i, j = 1, \ldots, p \\ Cov\left(Y_i, Y_j\right) = \dfrac{\sum_{k=1}^{m}\left(y_{ik} - \overline{Y_i}\right)\left(y_{jk} - \overline{Y_j}\right)}{m-1} \\ \overline{Y_i} = \dfrac{\sum_{k=1}^{m} y_{ik}}{m} \\ \overline{Y_j} = \dfrac{\sum_{k=1}^{m} y_{jk}}{m} \end{cases} \tag{18}$$

  where $Cov\left(Y_i, Y_j\right)$ is the covariance between the principal component Y_i and the principal component Y_j.

So to sum up, PCA is to reduce the number of variables without much loss of information. The solving process of principal component analysis includes the following four parts:

- **Standardization of initial variables:** PCA is quite sensitive to the variances of initial variables. If there are large differences between the ranges of initial variables, those variables with larger ranges will play a leading role in the analysis, ignoring the influence of those variables with small ranges on the principal components, leading to deviations in the analysis results. Therefore, we must standardize the range of initial variables so that each one of them contributes equally to the analysis. Mathematically, the standardization of the initial variable X_i is expressed as

$$\begin{cases} \widetilde{X_i} \in \left\{(\widetilde{x_{i1}}, \widetilde{x_{i2}}, \ldots, \widetilde{x_{im}})'\right\} \\ \widetilde{x_{ij}} = \dfrac{x_{ij} - \overline{X_i}}{\sqrt{\frac{\sum_{k=1}^{m}\left(x_{ik} - \overline{X_i}\right)^2}{m-1}}} \\ \overline{X_i} = \dfrac{\sum_{k=1}^{m} x_{ik}}{m} \end{cases} \tag{19}$$

  where $i = 1, 2, \ldots, p$, and $j = 1, 2, \ldots, m$.
- **Covariance matrix computation:** Once the standardization is done, all initial variables will be transformed to the same scale. The next step is to analyze the correlations between the standardized variables. The covariance matrix which can summary the correlations between all the possible pairs of the standardized variables, is expressed as

$$C_{p\times p} = \left(Cov\left(\widetilde{X_i}, \widetilde{X_j}\right)\right)_{p\times p} \tag{20}$$

  where $i, j = 1, 2, \ldots, p$.
- **Eigenvectors and eigenvalues of covariance matrix computation:** Based on the relationship between variance and information: the greater the variance, the larger the dispersion of the variable, and the richer the information contained in this variable. Geometrically speaking, principal components represent the direction of those variables that explain a maximal amount of variance.

The eigenvectors of the covariance matrix are actually the directions with the largest variance. Therefore, the principal components can be obtained by solving the eigenvectors of the covariance matrix $C_{p\times p}$. Meanwhile, the principal components in order of significance can be obtained by ranking the corresponding eigenvectors in order of their eigenvalues, highest to lowest.

In other words, suppose the eigenvalues of the covariance matrix $C_{p\times p}$ are $\delta_1, \delta_2, \ldots, \delta_p$, and the corresponding eigenvectors are $V_1, V_2, \ldots, V_p$, respectively. If the eigenvalues satisfy: $\delta_1 > \delta_2 > \ldots > \delta_p$, the $i-th$ principal component Y_i will be obtained by

$$Y_i = \sum_{j=1}^{p} \widetilde{X}_j V_i \tag{21}$$

which is equivalent to

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{im} \end{bmatrix} = \begin{bmatrix} \widetilde{x_{11}} & \widetilde{x_{21}} & \ldots & \widetilde{x_{p1}} \\ \widetilde{x_{12}} & \widetilde{x_{22}} & \ldots & \widetilde{x_{p2}} \\ \ldots & \ldots & \ddots & \ldots \\ \widetilde{x_{1m}} & \widetilde{x_{2m}} & \ldots & \widetilde{x_{pm}} \end{bmatrix} \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{ip} \end{bmatrix} \tag{22}$$

where $V_i = \left(v_{i1}, v_{i2}, \ldots, v_{ip}\right)', i = 1, 2, \ldots, p$.

- **Dimensionality reduction based on contribution rate:** In this step, we need to reduce the dimension of variables by discarding the unimportant components which are those components with less information. The percentage of information accounted for by each component, that is, the contribution rate τ of each component representing the initial variables, can be computed by

$$\tau_i = \frac{\delta_i}{\sum_{j=1}^{p} \delta_j} \tag{23}$$

where $i = 1, 2, \ldots, p$, and the cumulative contribution rate Γ_i of the first i principal components is

$$\Gamma_i = \frac{\sum_{j=1}^{i} \delta_j}{\sum_{j=1}^{p} \delta_j} \tag{24}$$

Given a threshold G, if $\Gamma_i \geq G$, we choose to reduce the dimension of the p-dimensional variable $\{Y_1, Y_2, \ldots, Y_p\}$ to the i-dimensional variable $\{Y_1, Y_2, \ldots, Y_i\}$, which is the main component.

2) DIMENSIONALITY REDUCTION OF ENVIRONMENTAL FEATURES

It can be seen from Table 3 that there are no four types of environment: sea, wet land, suburban openarea, and suburban village in the three measurement areas. Therefore, 16-dimensional linear, unweighted rectangular, and weighted rectangular environmental features can be extracted based on the definition of environmental features in the section III-A and the remained 16 environmental types, including water, urban openarea, green land, forest, reflection glass curtainwall, high buildings (I-VII), parallel regular buildings, irregular large buildings, irregular buildings and road land, in Table 3. Then, PCA is used to select the main component from these environmental features. As shown in Table 4, the cumulative contribution of the first ten components of the linear environmental features is 99.47% ($\geq G = 99\%$), that is, these principal components $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{10}$ already contain 99.47% of the information in the initial linear environmental features. Therefore, it is feasible to simplify the environmental feature F_L into a 10-dimensional vector $\lambda_L = (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{10})$.

TABLE 4. Contribution rate and cumulative contribution rate of each principal component in linear environmental features.

Principal Component	*Contribution*	*Cumulative Contribution*
λ_1	68.56%	68.56%
λ_2	12.53%	81.09%
λ_3	6.26%	87.35%
λ_4	3.39%	90.74%
λ_5	2.65%	93.39%
λ_6	1.86%	95.25%
λ_7	1.66 %	96.91%
λ_8	1.16%	98.07%
λ_9	0.91%	98.98%
λ_{10}	0.49%	**99.47%**
λ_{11}	0.14%	99.61%
λ_{12}	0.11%	99.72%
λ_{13}	0.04%	99.81%
λ_{14}	0.07 %	99.88%
λ_{15}	0.07%	99.95%
λ_{16}	0.05%	100%

TABLE 5. Contribution rate and cumulative contribution rate of each principal component in unweighted rectangular environmental features.

Principal Component	*Contribution*	*Cumulative Contribution*
ϕ_1	89.66%	89.66%
ϕ_2	6.94%	96.60%
ϕ_3	1.69%	98.29%
ϕ_4	0.61%	98.90%
ϕ_5	0.45%	**99.35%**
ϕ_6	0.29%	99.64%
ϕ_7	0.16 %	99.80%
ϕ_8	0.09%	99.89%
ϕ_9	0.07%	99.96%
ϕ_{10}	0.02%	99.98%
ϕ_{11}	0.01%	99.99%
ϕ_{12}	0.01%	100%
ϕ_{13}	0%	100%
ϕ_{14}	0 %	100%
ϕ_{15}	0%	100%
ϕ_{16}	0%	100%

Similarly, the cumulative contribution rates of the components of the unweighted and weighted rectangular environmental features are shown in Table 5 and Table 6, respectively. Then, retaining more than 99% of the information in initial environmental features, the environmental features F_{R1} and F_{R2} are reduced to 5-dimensional and 10-dimensional feature vectors, namely $\phi_{R1} = (\phi_1, \phi_2, \phi_3, \ldots, \phi_5)$ and $\varphi_{R2} = (\varphi_1, \varphi_2, \varphi_3, \ldots, \varphi_{10})$, respectively.

TABLE 6. Contribution rate and cumulative contribution rate of each principal component in weighted rectangular environmental features.

Principal Component	*Contribution*	*Cumulative Contribution*
φ_1	42.39%	42.39%
φ_2	23.9%	66.29%
φ_3	11.2%	77.49%
φ_4	7.53%	85.02%
φ_5	4.3%	89.32%
φ_6	3.31%	92.63%
φ_7	2.39 %	95.02%
φ_8	2.09%	97.11%
φ_9	1.3%	98.41%
φ_{10}	0.75%	**99.16%**
φ_{11}	0.47%	99.63%
φ_{12}	0.16%	99.79%
φ_{13}	0.12%	99.91%
φ_{14}	0.06 %	99.97%
φ_{15}	0.03%	100%
φ_{16}	0%	100%

TABLE 7. Data sets with different environmental features.

Data set	Input		Target
D_1	$L_{BS}, L_{Rx}, d_{3D}, A_z,$ $D_e, G_{BS}, P_{BS}, f_c, L_f$		PL
D_2		F_L	
D_3		λ_L	
D_4		F_{R1}	
D_5		ϕ_{R1}	
D_6		F_{R2}	
D_7		φ_{R2}	

C. DATA SETS CONSTRUCTION

After extracting environmental features, it is necessary to combine the measurement data with environmental features to construct datasets for establishing PL prediction models. From the section II-A, the information of BS and Rx is obtained by measurement and subsequent processing, including L_{BS}, L_{Rx}, f_c, L_f, A_z, D_e, G_{BS}, P_{BS}, and PL at all measurement positions. The distance d_{3D} between BS and Rx is a key factor affecting radio wave propagation, so the distance d_{3D} is also used as an input parameter for PL prediction models.

Environmental features of the measurement data are extracted by the method described in section III-A, and the corresponding low-dimensional features are also generated by using PCA. In order to analyze the impact of environmental features on PL, seven datasets containing different information are constructed as shown in Table 7. The dataset without environmental features is expressed as D_1, where the set $\{L_{BS}, L_{Rx}, d_{3D}, A_z, D_e, G_{BS}, P_{BS}, f_c, L_f\}$ forms the input set of MLP neural networks, and $\{PL\}$ is the target set. D_2, D_3, D_4, D_5, D_6, and D_7 are the datasets with environmental features, where their input sets are the combination of the input set of the dataset without environmental features and the corresponding environmental features, and $\{PL\}$ is their target set.

D. PATH LOSS PREDICTION MODEL

Compared with many neural network structures, MLP neural network which is trained by the BP algorithm, is more mature in theory and performance, and has the advantages of simple structure, stable state, and easy implementation. In this section, MLP networks with different architectures are employed to establish PL prediction models. Meanwhile, the impacts of ANN architectures, the percentage of training samples, and environmental features on the predicted results are evaluated by AME, MAE, STD, the correlation coefficients r, and the time ratio T. The time ratio T can reflect the training efficiency of different MLP neural networks established to analyze the impact of the same parameter, such as ANN architectures, the percentage of training samples, and environmental features, on the PL model, which is defined as

$$T = \frac{T_{training}}{max(T_{training})} \tag{25}$$

where $T_{training} = [t_1, t_2, \ldots, t_{N_{net}}]$, $max(T_{training})$ is the maximum value of the components in $T_{training}$, N_{net} is the number of MLP neural networks, and t_i is the training time of the $i-th$ MLP neural network, $i = 1, 2, \ldots, N_{net}$.

TABLE 8. Network architectures and hyperparameter settings.

Architectures	Hidden Layer Node	Activation Function
One hidden layer	$8^{(1)} - 14^{(1)}$	tansig/purelin
Two hidden layer	$\{10^{(1)}, 5^{(2)}\}$	tansig/purelin
Three hidden layer	$\{10^{(1)}, 5^{(2)}, 3^{(3)}\}$	tansig/purelin
lr: 0.05	P_{train}: 70%	f_{train}: L-M
f_{loss}: MSE	E_{goal}: 0.00001	Epoch: 1000

1) NUMBER OF NEURONS

The number of neurons in the hidden layer is a key factor affecting the performance of MLP networks. If the number of neurons is too small, ANN does not have the necessary learning and information processing capabilities. Conversely, too many neurons will reduce the learning rate and easily lead to overfitting. Today, there is no clear theoretical method on how to determine the number of neurons. However, the number of neurons in a MLP network with one hidden layer can be roughly determined by empirical formula.

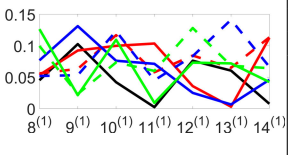
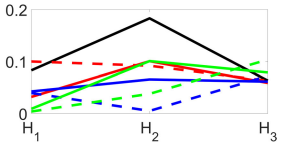
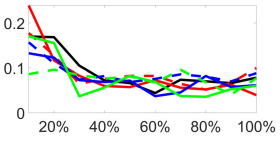
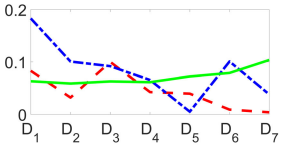
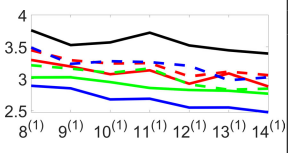
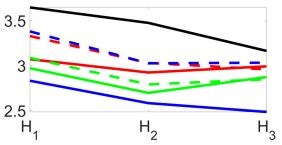
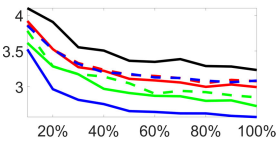
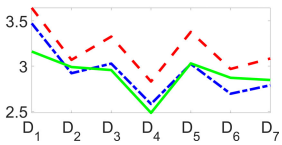
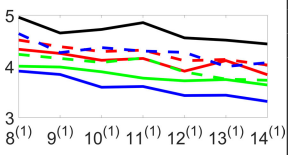
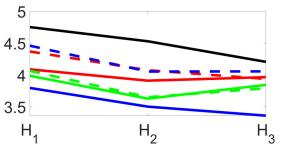
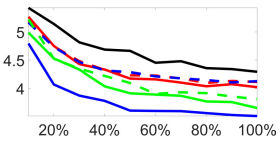
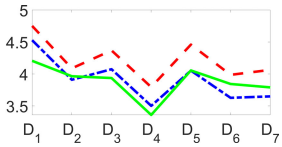
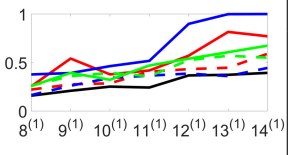
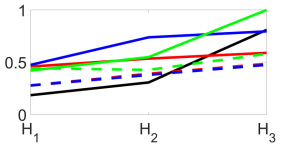
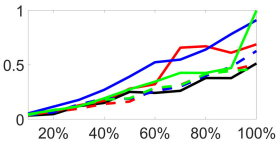
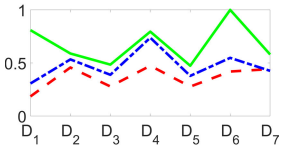
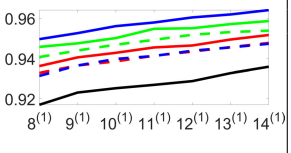
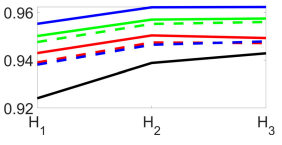
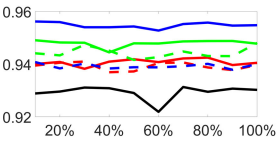
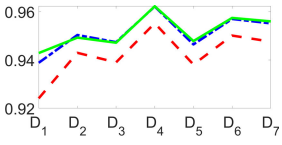
In this work, the number of neurons in the hidden layer is determined as follows:

$$N_h = \sqrt{N_{input} + N_{output}} + \alpha \tag{26}$$

where N_{input} and N_{output} are the number of neurons in the input and output layers, respectively. N_h is the number of neurons in the hidden layer, and the factor α can take a value between 1 and 10.

Considering the dimension of the datasets constructed in the section III-C, it is finally determined that the number of neurons in the MLP network with one hidden layer can take a value between 8 and 14. In order to explore the impact of neurons on PL prediction models, seven MLP neural networks with 8 to 14 neurons in the hidden layer are established, and 70% training samples were randomly selected from the above seven datasets. Table 8 shows the different architectures and hyperparameter settings of MLP networks, where $n^{(i)}$ means

TABLE 9. Error analysis of PL prediction models under different network settings.

	MLP Networks with Different Numbers of Neurons in the Hidden Layer	MLP Networks with Different Numbers of Hidden Layers	MLP Network $H_2 = \{10^{(1)}, 5^{(2)}\}$ Based on Different Training Percentages	MLP Networks Obtained by Different Envirnomental Features
Legend	D_1 D_2 D_3 D_4 D_5 D_6 D_7	D_1 D_2 D_3 D_4 D_5 D_6 D_7	D_1 D_2 D_3 D_4 D_5 D_6 D_7	$10^{(1)}$ $10^{(1)},5^{(2)}$ $10^{(1)},5^{(2)},3^{(3)}$
AME [dB]				
	[0.002 dB, 0.141 dB]	[0.004 dB, 0.183 dB]	[0.036 dB, 0.239 dB]	[0.004 dB, 0.183 dB]
MAE [dB]				
	[2.476 dB, 3.76 dB]	[2.483 dB, 3.644 dB]	[2.575 dB, 4.1 dB]	[2.483 dB, 3.644 dB]
STD [dB]				
	[3.317 dB, 4.91 dB]	[3.36 dB, 4.752 dB]	[3.507 dB, 5.429 dB]	[3.36 dB, 4.752 dB]
Time Ratio				
	$max(T_{training})$ = 19.294 s	$max(T_{training})$ = 19.254 s	$max(T_{training})$ = 15.814 s	$max(T_{training})$ = 19.254 s
Correlation				
	[0.917, 0.964]	[0.924, 0.962]	[0.922, 0.956]	[0.924, 0.962]

that the number of neurons in the $i-th$ hidden layer is n. The learning rate lr is 0.05, the percentage of training samples to total samples P_{train} is 70%, the training function f_{train} is Levenberg-Marquardt (L-M), the loss function f_{loss} is mean-square error (MSE), the goal error E_{goal} is 0.00001, and epoch which defines the number times that the learning algorithm will work through the entire training dataset, is 1000.

The second column in Table 9 shows the analysis of the error between the measured results PL and the predicted results PL' obtained by the MLP networks with one hidden layer and different numbers of neurons in the hidden layer.

It can be seen that, whether it is based on the dataset without environmental features or the dataset with environmental features, the AME of MLP-based PL prediction models is less than 0.15 dB. Meanwhile, the MAE and STD of these PL prediction models are less than 3.76 dB and 4.91 dB, respectively, which are better than the predicted results of the CI PL model and the A-B PL model. Moreover, as the number of neurons in the hidden layer increases, the STD of MLP-based PL prediction models shows a downward trend. This phenomenon is mainly because the increase of neurons improves the nonlinear characterization ability of MLP neural networks, which makes PL prediction models more stable and STD smaller. However, the increase of neurons will lead to the increase of training time and affect the efficiency of PL prediction models. As shown in the fourth figure in the second column of Table 9, the time ratio T increases as the number of neurons in the hidden layer increases. It takes the longest training time to establish the MLP neural network with 14 neurons using the dataset with unweighted environmental features, which is 19.294 s.

The correlation coefficient r between the measured results and the predicted results obtained by the MLP-based PL models are greater than 0.91, which are significantly higher than that between the measured results and predicted results obtained by the CI PL model and the A-B PL model. We find that the correlation coefficient between the predicted results obtained by MLP neural networks based on the datasets with environmental features and the measured results is greater than that between the predicted results obtained by MLP neural networks based on the datasets without environmental features and the measured results, which shows the effectiveness of environmental features and also conforms to the law that environment has an important influence on radio wave propagation. Further, no matter how many neurons are in the hidden layer, the correlation coefficient between the predicted results obtained by MLP neural networks based on the dataset with unweighted environmental features and the measured results is the largest, which reveals that the unweighted environmental features can more effectively describe the measurement scenario.

In summary, compared with the CI PL model and the A-B PL model, PL prediction models based on MLP networks can achieve high-accuracy PL prediction. Meanwhile, the introduction of environmental features, especially the unweighted environmental features, helps MLP neural networks to learn the relationship between radio propagation and environmental types automatically, and improve the stability of PL prediction models.

2) NUMBER OF HIDDEN LAYERS

The number of hidden layers affects the nonlinear mapping capability of ANN. The more hidden layers, the stronger the nonlinear mapping capability. In order to explore the impact of the number of hidden layers on PL prediction models, we established three MLP networks with one, two and three hidden layers, expressed as $H_1 = \{10^{(1)}\}$, $H_2 = \{10^{(1)}, 5^{(2)}\}$, and $H_3 = \{10^{(1)}, 5^{(2)}, 3^{(3)}\}$. Meanwhile, these MLP networks are trained with 70% samples selected randomly from the above seven datasets. The architecture and hyperparameter settings of these MLP networks are shown in Table 8.

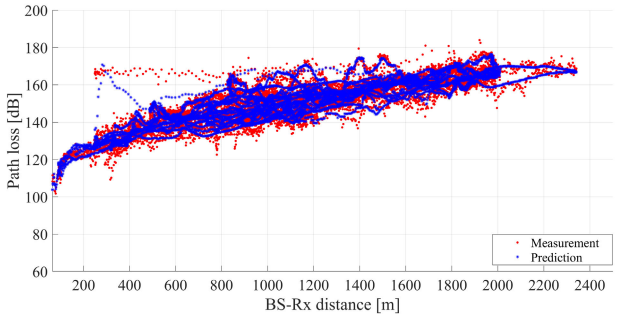
The third column in Table 9 shows AME, MAE, STD, the correlation coefficient, and the time ratio of PL prediction models based on MLP neural networks with different number of hidden layers. It can be seen that the maximum value of AME is less than 0.2 dB, and the MAE and STD of PL prediction models based on MLP neural networks are less than 3.644 dB and 4.752 dB, which are better than the predicted results of the CI PL model and the A-B PL model. As the number of hidden layers increases, the MAE and STD of MLP-based PL prediction models show a downward trend. However, a large number of hidden layers lead to overfitting and increase the training time, which will affect accuracy and efficiency of these PL prediction models. As shown in the fourth figure in the third column of Table 9, compared with the training time of the MLP neural network with two hidden layers, the training time of the MLP network with three hidden layers increases significantly, and the training time of the MLP neural network with three hidden layers based on the dataset with weighted rectangular environmental features is the longest, which is 19.254 s.

The correlation coefficients between the measured results and the predicted results obtained by MLP neural networks are greater than 0.924, which are higher than that between the measured results and the predicted results obtained by the CI PL model and the A-B PL model. Compared with the predicted results obtained by MLP neural networks based on the datasets with linear environmental features and weighted rectangular environmental features, the predicted results obtained by MLP neural networks based on the dataset with unweighted environmental features has the strongest correlation with the measured results, which reveals that the unweighted rectangular environmental features are more helpful to improve the accuracy and stability of PL prediction models. Moreover, the predicted results obtained by MLP neural networks based on the datasets with environmental features are better than that obtained by MLP neural networks based on the datasets without environmental features, which further indicates the effectiveness of environmental features.

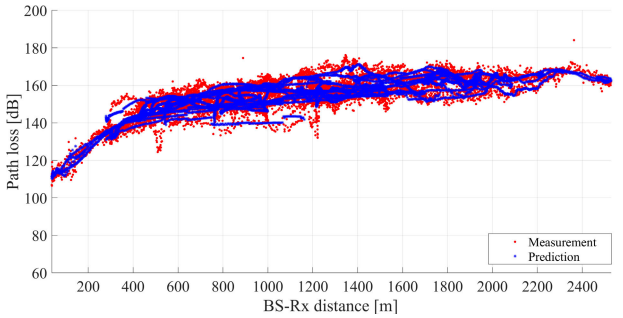
On balance, the accurate and effective prediction of PL can be achieved by adding hidden layers in a certain range and introducing environmental features. In this study, based on the unweighted rectangular environmental features, the MLP neural network with two hidden layer $H_2 = \{10^{(1)}, 5^{(2)}\}$ realized the PL prediction with both accuracy and efficiency (see Fig. 5).

3) TRAINING PERCENTAGE

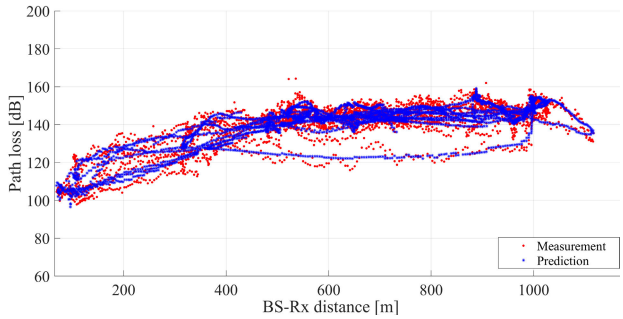
ANN has high requirements on the number of training samples. A small number of samples leads to overfitting and affects the generalization ability of ANN. A large number of samples leads to long training time and reduces the efficiency of ANN. In order to analyze the impact of the number of



(a) BS1



(b) BS2



(c) BS3

FIGURE 5. The predicted results of path loss based on the model $H_2 = \{10^{(1)}, 5^{(2)}\}$.

training samples on PL prediction models, the predicted results of the MLP-based PL model $H_2 = \{10^{(1)}, 5^{(2)}\}$ under different numbers of training samples are compared. The training samples are randomly selected with different percentages (10% - 100%) from the training dataset which accounted for 70% of the total dataset. The architecture and hyperparameter settings of the PL prediction model are shown in Table 8.

The fourth column in Table 9 shows AME, MAE, STD, the correlation coefficient, and the time ratio of MLP-based PL prediction models with different training samples. The MAE and STD of the MLP-based PL prediction models decrease and the training time increases with the increase of training samples. We find that the AME, MAE and STD of the MLP-based PL models built with a percentage of 60% and above training samples are very close, which means that 60% of the training samples is enough to obtain the stable prediction models. As shown in the fourth figure in the fourth column of Table 9, the training time of the MLP neural network $H_2 = \{10^{(1)}, 5^{(2)}\}$ based on all training samples with weighted rectangular environmental features is the longest, which is 15.814 s. The time ratios of the PL prediction models established by 60% of the training samples corresponding to seven datasets ($D_1 \sim D_7$) with different environmental features are 0.24, 0.32, 0.27, 0.52, 0.26, 0.34, and 0.29, respectively.

To be brief, MLP neural networks obtained by 60% training samples can achieve high-accuracy PL prediction, which significantly reduces the training time and improves the efficiency of PL prediction models. Meanwhile, the introduction of environmental features, especially unweighted environmental features, improve the accuracy and stability of PL prediction models.

4) EFFECTIVENESS OF ENVIRONMENTAL FEATURES

Based on the above analysis, it can be seen that the introduction of environmental features helps to establish the stable PL prediction model, and the unweighted rectangular environmental features are more effective than the other two environmental features. In order to further explore the effectiveness of environmental features and their impacts on PL prediction models, the predicted results of the PL models $H_1 = \{10^{(1)}\}$, $H_2 = \{10^{(1)}, 5^{(2)}\}$, and $H_3 = \{10^{(1)}, 5^{(2)}, 3^{(3)}\}$ under different environmental features are compared in this section. The architecture and hyperparameter settings of these MLP neural networks are shown in Table 8.

The fifth column in Table 9 shows AME, MAE, STD, the correlation coefficient, and the time ratio of PL prediction models based on different environmental features. The MAE, STD and the correlation coefficient r of the datasets with environmental features is smaller than that of the dataset without environmental features, which indicates that the environmental features improve the performance of PL prediction models and verifies the effectiveness of the environmental features proposed in this work. Meanwhile, MAE, STD, and the correlation coefficient r bsaed on the datasets with different environmental features satisfy the following formula:

$$\begin{cases} MAE_{D_1} \geqslant MAE_{D_2} \geqslant MAE_{D_6} \geqslant MAE_{D_4} \\ STD_{D_1} \geqslant STD_{D_2} \geqslant STD_{D_6} \geqslant STD_{D_4} \\ r_{D_1} \leqslant r_{D_2} \leqslant r_{D_6} \leqslant r_{D_4} \end{cases} \tag{27}$$

From (27), it can be seen that the unweighted environmental feature is the most effective among the three types of environmental features, which can be answered from the definition of environmental features. The linear environmental feature only considers the environmental types on the LOS path. However, due to the complexity of the propagation environment, multiple propagation mechanisms produced by the surrounding environment, such as reflection, diffraction, and scattering, exist in radio propagation. Therefore, the effectiveness of linear environmental features is lower than that of rectangular environmental features.

Because the height information of the radio propagation environment is not considered, a large number of environ-

mental types which do not affect radio propagation are added in the extraction of rectangular environmental features. If the environmental type that does not affect radio propagation is close to the LOS path, the weighted rectangular environmental feature will enhance the interference of this environmental type. Therefore, compared with the weighted environmental features, the unweighted environmental features are more effective when the height information is not considered. In addition, although some information of environmental features are lost after dimensionality reduction, and the performance of PL prediction models is reduced, but it can still achieve high accuracy prediction of PL. Compared with PL prediction models based on the initial environmental features, the training time of PL prediction models based on the low-dimensional environmental features are reduced by 20% on average.

IV. CONCLUSION AND FUTURE WORK

In this paper, PL prediction models based on MLP neural networks with different architectures are established. Three environmental features are defined by considering limited environmental type information and propagation mechanism, which avoid the cumbersome process of 3D environment modeling and improves the efficiency of PL prediction models. Meanwhile, PCA is used to eliminate the linear correlation between similar environmental types. Seven different datasets are constructed by different combinations of the measurement data and environmental features. PL prediction models based on MLP neural networks with different architectures are trained by these datasets. The impact of environmental features and network hyperparameters, such as the number of neurons in the hidden layer, the number of hidden layers, and the number of training samples on PL prediction models are studied and analyzed.

The predicted results show that: (1) compared with the CI PL model and the A-B PL model, the PL models based on MLP neural networks can achieve high-accuracy PL prediction. (2) The introduction of environmental features improves the accuracy and stability of these PL prediction models, and the effectiveness of the unweighted rectangular environmental features are more significant. (3) The accuracy of PL prediction models increase with the complexity of the network architecture. However, an overly complex network architecture is prone to overfitting, resulting in a decrease in the accuracy and efficiency of PL prediction models. (4) 60% of the training samples from training sets is enough for obtaining a high efficiency and stable PL prediction model.

In summary, in the absence of 3D environment modeling, MLP neural network uses limited environmental types to reflect the propagation characteristics, and establishes PL prediction models, which can provide a theoretical basis for wireless network optimization and communication system design. However, due to the lack of height information, the environmental features can not accurately reflect the propagation characteristics, which affects the accuracy of the PL prediction models. RT can extract multipath information reflecting propagation characteristics. Therefore, RT and ANN can be combined in future work to further improve the accuracy of PL prediction models.

REFERENCES

[1] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, "Planning wireless cellular networks of future: Outlook, challenges and opportunities," *IEEE Access*, vol. 5, pp. 4821–4845, 2017.
[2] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Kurner, "The design and applications of high-performance ray-tracing simulation platform for 5G and beyond wireless communications: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 10–27, 1st Quart., 2019.
[3] L. Zeng, X. Cheng, C.-X. Wang, and X. Yin, "A 3D geometry-based stochastic channel model for UAV-MIMO channels," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–5.
[4] X. Cheng and Y. Li, "A 3-D geometry-based stochastic model for UAV-MIMO wideband nonstationary channels," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1654–1662, Apr. 2019.
[5] R. Jia, Y. Li, X. Cheng, and B. Ai, "3D geometry-based UAV-MIMO channel modeling and simulation," *China Commun.*, vol. 15, no. 12, pp. 64–74, Dec. 2018.
[6] H. Chang, J. Bian, C.-X. Wang, Z. Bai, W. Zhou, and E.-H.-M. Aggoune, "A 3D non-stationary wideband GBSM for low-altitude UAV-to-ground V2V MIMO channels," *IEEE Access*, vol. 7, pp. 70719–70732, 2019.
[7] J. Zhang, C. Pan, F. Pei, G. Liu, and X. Cheng, "Three-dimensional fading channel models: A survey of elevation angle research," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 218–226, Jun. 2014.
[8] Z. . Huang, X. Cheng, and N. Zhang, "An improved non-geometrical stochastic model for non-WSSUS vehicle-to-vehicle channels," *ZTE Commun.*, vol. 17, no. 4, pp. 62–71, Dec. 2019.
[9] U. Karabulut, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, "Spatial and temporal channel characteristics of 5G 3D channel model with beamforming for user mobility investigations," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 38–45, Dec. 2018.
[10] X. Cheng, B. Yu, L. Yang, J. Zhang, G. Liu, Y. Wu, and L. Wan, "Communicating in the real world: 3D MIMO," *IEEE Wireless Commun.*, vol. 21, no. 4, pp. 136–144, Aug. 2014.
[11] L. Wu, D. He, K. Guan, B. Ai, C. Briso-Rodriguez, T. Shui, C. Liu, L. Zhu, and X. Shen, "Received power prediction for suburban environment based on neural network," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2020, pp. 35–39.
[12] K. Chamberlin and R. Luebbers, "An evaluation of Longley-Rice and GTD propagation models," *IEEE Trans. Antennas Propag.*, vol. 30, no. 6, pp. 1093–1098, Nov. 1982.
[13] J. B. Andersen, "UTD multiple-edge transition zone diffraction," *IEEE Trans. Antennas Propag.*, vol. 45, no. 7, pp. 1093–1097, Jul. 1997.
[14] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, vol. 14, no. 3, pp. 302–307, May 1966.
[15] W. Fan, I. Carton, P. Kyosti, and G. F. Pedersen, "Emulating ray-tracing channels in multiprobe anechoic chamber setups for virtual drive testing," *IEEE Trans. Antennas Propag.*, vol. 64, no. 2, pp. 730–739, Feb. 2016.
[16] K. Guan, B. Peng, D. He, J. M. Eckhardt, S. Rey, B. Ai, Z. Zhong, and T. Kurner, "Channel characterization for intra-wagon communication at 60 and 300 GHz bands," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5193–5207, Jun. 2019.
[17] K. Guan, B. Peng, D. He, J. M. Eckhardt, S. Rey, B. Ai, Z. Zhong, and T. Kurner, "Measurement, simulation, and characterization of train-to-infrastructure inside-station channel at the terahertz band," *IEEE Trans. THz Sci. Technol.*, vol. 9, no. 3, pp. 291–306, May 2019.
[18] J. Chen, X. Yin, L. Tian, and M.-D. Kim, "Millimeter-wave channel modeling based on a unified propagation graph theory," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 246–249, Feb. 2017.
[19] K. Guan, B. Ai, B. Peng, D. He, G. Li, J. Yang, Z. Zhong, and T. Kurner, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part I: Paradigm, significance analysis, and scenario reconstruction," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9112–9128, Oct. 2018.
[20] K. Guan, B. Ai, B. Peng, D. He, G. Li, J. Yang, Z. Zhong, and T. Kurner, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part II: Case study for paradigm implementation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9129–9144, Oct. 2018.

[21] B. Ai, A. F. Molisch, M. Rupp, and Z.-D. Zhong, "5G key technologies for smart railways," *Proc. IEEE*, vol. 108, no. 6, pp. 856–893, Jun. 2020.
[22] 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.
[23] 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.
[24] 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.
[25] S. Sun, T. S. Rappaport, T. A. Thomas, A. Ghosh, H. C. Nguyen, I. Z. Kovacs, 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.
[26] J. Ahrens, L. Ahrens, and H. D. Schotten, "A machine learning method for prediction of multipath channels," *ZTE Commun.*, vol. 17, no. 4, pp. 12–18, Dec. 2019.
[27] P.-R. Chang and W.-H. Yang, "Environment-adaptation mobile radio propagation prediction using radial basis function neural networks," *IEEE Trans. Veh. Technol.*, vol. 46, no. 1, pp. 155–160, 1997.
[28] I. Povescu, I. Nafomita, P. Constantinou, A. Kanatas, and N. Moraitis, "Neural networks applications for the prediction of propagation path loss in urban environments," in *Proc. IEEE VTS 53rd Veh. Technol. Conf.*, vol. 1, Spring 2001, pp. 387–391.
[29] 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*, vol. 20, no. 7, p. 1927, Mar. 2020, doi: 10.3390/s20071927.
[30] H. Yu and M. Bennamoun, "1D-PCA, 2D-PCA to nD-PCA," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 4, Aug. 2006, pp. 181–184.
[31] J. Thrane, D. Zibar, and H. L. Christiansen, "Model-aided deep learning method for path loss prediction in mobile communication systems at 2.6 GHz," *IEEE Access*, vol. 8, pp. 7925–7936, 2020.
[32] N. A. Al-Sammarraie, Y. M. H. Al-Mayali, and Y. A. Baker El-Ebiary, "Classification and diagnosis using back propagation artificial neural networks (ANN)," in *Proc. Int. Conf. Smart Comput. Electron. Enterprise (ICSCEE)*, Jul. 2018, pp. 1–5.
[33] S. P. Sotiroudis, S. K. Goudos, K. A. Gotsis, K. Siakavara, and J. N. Sahalos, "Application of a composite differential evolution algorithm in optimal neural network design for propagation path-loss prediction in mobile communication systems," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 364–367, 2013.
[34] S. P. Sotiroudis, S. K. Goudos, K. A. Gotsis, K. Siakavara, and J. N. Sahalos, "Modeling by optimal artificial neural networks the prediction of propagation path loss in urban environments," in *Proc. IEEE-APS Topical Conf. Antennas Propag. Wireless Commun. (APWC)*, Sep. 2013, pp. 599–602.
[35] S. P. Sotiroudis, S. K. Goudos, and K. Siakavara, "Neural networks and random forests: A comparison regarding prediction of propagation path loss for NB-IoT networks," in *Proc. 8th Int. Conf. Modern Circuits Syst. Technol. (MOCAST)*, May 2019, pp. 1–4.
[36] N. Faruk, S. I. Popoola, N. T. Surajudeen-Bakinde, A. A. Oloyede, A. Abdulkarim, L. A. Olawoyin, M. Ali, C. T. Calafate, and A. A. Atayero, "Path loss predictions in the VHF and UHF bands within urban environments: Experimental investigation of empirical, heuristics and geospatial models," *IEEE Access*, vol. 7, pp. 77293–77307, 2019.
[37] S. I. Popoola, A. Jefia, A. A. Atayero, O. Kingsley, N. Faruk, O. F. Oseni, and R. O. Abolade, "Determination of neural network parameters for path loss prediction in very high frequency wireless channel," *IEEE Access*, vol. 7, pp. 150462–150483, 2019.
[38] M. A. Salman, S. I. Popoola, N. Faruk, N. T. Surajudeen-Bakinde, A. A. Oloyede, and L. A. Olawoyin, "Adaptive neuro-fuzzy model for path loss prediction in the VHF band," in *Proc. Int. Conf. Comput. Netw. Informat. (ICCNI)*, Oct. 2017, pp. 1–6.
[39] 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.
[40] H. T. Friis, "A note on a simple transmission formula," *Proc. IRE*, vol. 34, no. 5, pp. 254–256, May 1946.
[41] J. Singh and R. Banerjee, "A study on single and multi-layer perceptron neural network," in *Proc. 3rd Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Mar. 2019, pp. 35–40.

LINA WU (Student Member, IEEE) received the B.S. degree in information and computing science from Ludong University, in 2014, and the M.S. degree in applied mathematics from the Beijing Information Science and Technology University, in 2018. She is currently pursuing the Ph.D. degree with Beijing Jiaotong University, Beijing, China. Her research activities are mainly in image processing, machine learning, and ray-tracing technologies.

DANPING HE (Member, IEEE) received the B.E. degree from the Huazhong University of Science and Technology, in 2008, the M.Sc. degree from the Universite Catholique de Louvain (UCL) and Politecnico di Torino (PdT), in 2010, and the Ph.D. degree from the Universidad Politécnica de Madrid, in 2014. In 2012, she was a visiting scholar with the Institut national de recherche en informatique et en automatique, France. She has worked in Huawei Technologies as a Research Engineer, from 2014 to 2015. From 2016 to 2018, she conducted postdoctoral research with Beijing Jiaotong University, where she is currently an Associate Professor with Beijing Jiaotong University. She has authored/coauthored more than 40 papers, three patents, and one IEEE standard. Her papers received five Best Paper Awards. Her current research interests include radio propagation and channel modeling, ray-tracing technologies, and wireless communication algorithm design. She also received the 2019 Applied Computational Electromagnetics Society (ACES)-China Young Scientist Award.

BO AI (Senior Member, IEEE) received the master's and Ph.D. degrees from Xidian University, China. He graduated from Tsinghua University with the honor of an Excellent Postdoctoral Research Fellow at Tsinghua University, in 2007. He was a Visiting Professor with the Electrical Engineering Department, Stanford University, in 2015. He is currently working at BJTU as a Full Professor and a Ph.D. candidate advisor. He has authored/coauthored eight books and has published over 300 academic research papers in his research area. He holds 26 invention patents. He has been the research team leader for 26 national projects and has won some important scientific research prizes. He has been notified by the Council of Canadian Academies (CCA) that, based on the Scopus database, he has been listed as one of the Top 1 percent of authors in his field all over the world. His interests include the research and applications of channel measurement and channel modeling, and dedicated mobile communications for rail traffic systems. He is an IET Fellow and an IEEE VTS Distinguished Lecturer. He has received many awards such as the Outstanding Youth Foundation from the National Natural Science Foundation of China, the Qiushi Outstanding Youth Award by the Hong Kong Qiushi Foundation, the New Century Talents by the Chinese Ministry of Education, the Zhan Tianyou Railway Science and Technology Award of the Chinese Ministry of Railways, and the Science and Technology New Star of the Beijing Municipal Science and Technology Commission. He has been the Co-Chair or a Session/Track Chair for many international conferences. He is an Editor of the IEEE TRANSACTIONS ON CONSUMER ELECTRONICS and an Editorial Committee member of the *Wireless Personal Communications* journal. He has been the Lead Guest Editor for Special Issues of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE ANTENNAS AND PROPAGATIONS LETTERS, and the *International Journal on Antennas and Propagation*.

JIAN WANG (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Beijing Jiaotong University, Beijing, China, in 2000, 2003, and 2007, respectively. He was a Lecturer with the School of Electronics and Information Engineering, Beijing Jiaotong University, from 2007 to 2010, where he is currently a Professor. His current research interests include collaborative vehicle-road system research, computerized simulation of train control systems, new GNSS applications in railway, and other ITS technologies.

HANG QI received the B.S. degree from Zhengzhou University, in 2014, and the M.S. degree from the Beijing University of Posts and Telecommunications, in 2017. Since 2017, he has been working with the China Mobile Group Design Institute Company, Ltd., as a Consulting Designer. His research activities are mainly in wireless network communication systems and new technology applications. He has authored/co-authored three papers and two patents.

KE GUAN (Senior Member, IEEE) received the B.E. and Ph.D. degrees from Beijing Jiaotong University, in 2006 and 2014, respectively. He is currently a Professor with the State Key Laboratory of Rail Traffic Control and Safety, School of Electronics and Information Engineering, Beijing Jiaotong University. In 2015, he has been awarded a Humboldt Research Fellowship for Postdoctoral Researchers. In 2009, he was a Visiting Scholar with the Universidad Politécnica de Madrid, Spain. From 2011 to 2013, he was a Research Scholar with the Institut für Nachrichtentechnik (IfN), Technische Universität Braunschweig, Germany. From September 2013 to January 2014, he was invited to conduct joint research in Universidad Politécnica de Madrid, Spain. He has authored/coauthored two books and one book chapter, more than 200 journal and conference papers, and four patents. His current research interests include measurement and modeling of wireless propagation channels, high-speed railway communications, vehicle-to-x channel characterization, and indoor channel characterization for high-speed short-range systems, including future terahertz communication systems. He is the pole leader of the European Railway Research Network of Excellence (EURNEX). He serves as the Publicity Chair in PIMRC 2016, the Publicity Co-Chair in ITST 2018, the Track Co-Chair in EuCNC, the Session Convener of EuCAP 2015-2019, and a TPC Member for many IEEE conferences, such as Globecom, ICC, and VTC. He has been a delegate in 3GPP and a member of the IC1004 and CA15104 initiatives. He was a recipient of the 2014 International Union of Radio Science (URSI) Young Scientist Award. His papers received eight Best Paper Awards, including the IEEE Vehicular Technology Society 2019 Neal Shepherd Memorial Best Propagation Paper Award. He is an Editor of IEEE Access, the *IET Microwave, Antenna and Propagation*, *Physical Communication*, and a Guest Editor of the IEEE Transactions on Vehicular Technology and the *IEEE Communication Magazine*.

ZHANGDUI ZHONG (Senior Member, IEEE) is a Professor and Advisor of Ph.D. candidates with Beijing Jiaotong University, Beijing, China, where he is currently the Director of the School of Computer and Information Technology and a Chief Scientist of the State Key Laboratory of Rail Traffic Control and Safety. He is also the Director of the Innovative Research Team of Ministry of Education, and a Chief Scientist of the Ministry of Railways in China. He is an Executive Council Member of the Radio Association of China, and a Deputy Director of the Radio Association of Beijing. He has authored/coauthored seven books, five invention patents, and more than 200 scientific research papers in his research area. His research interests include wireless communications for railways, control theory and techniques for railways, and GSM-R systems. His research has been widely used in railway engineering, such as Qinghai-Xizang railway, Datong-Qinhuangdao Heavy Haul railway, and many high-speed railway lines of China.

Dr. Zhong was a recipient of the MaoYiSheng Scientific Award of China, the ZhanTianYou Railway Honorary Award of China, and the Top ten Science/Technology Achievements Award of Chinese Universities.