A Frequency Domain Predictive Channel Model for 6G Wireless MIMO Communications Based on Deep Learning

Chen Huang, *Member, IEEE,* Cheng-Xiang Wang, *Fellow, IEEE,* Zheao Li, Zhongyu Qian, Junling Li, *Member, IEEE,* Yang Miao *Member, IEEE*

***Abstract*—The development of sixth-generation (6G) wireless communication systems brings significant challenges in channel modeling. Conducting channel measurements for 6G communications is highly expensive and cannot cover all scenarios and frequency bands. Moreover, existing conventional channel models fail to accurately predict channel characteristics in unknown frequency band. As a result, predictive channel modeling has emerged as a promising solution for addressing these challenges in 6G channel modeling. In this study, we propose a frequency domain predictive channel model that combines an autoencoder with a coupling Convolution Gated Recurrent Unit (Conv-GRU) cells. The proposed model aims to predict channel characteristics in unknown frequency bands. The proposed predictive channel model is validated by using data collected from multiple frequency bands channel measurements. To evaluate its performance, several commonly used prediction networks, i.e., a general LSTM network, a GRU-based predictive network, and a Conv-LSTM-based predictive network, are conducted as benchmarks for comparison. Based on evaluation results, our proposed predictive channel model achieves the highest level of accuracy in predicting channels. Additionally, we provide a performance bound for extrapolation predictability using a Ray Tracing simulator.**

***Index Terms*—Channel characterization, channel measurements, channel modeling, deep learning, Conv-GRU.**

Manuscript received 24 April 2023; revised 22 October 2023; accepted 21 February 2024. This work was supported by the National Key R&D Program of China under Grant 2023YFB2904802, the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grants BE2022067 and BE2022067-1, the Fundamental Research Funds for the Central Universities under Grants 2242022k60006, the National Natural Science Foundation of China (NSFC) under Grants 61960206006, 62301364, and 62301151, the EU H2020 RISE TESTBED2 project under Grant 872172, Young Elite Scientists Sponsorship Program by CAST under Grant 2022QNRC001, the Start-up Research Fund of Southeast University under Grant RF1028623062. The associate editor coordinating the review of this article and approving it for publication was A. Alkhateeb. (*Corresponding Authors*: Cheng-Xiang Wang and Junling Li.)

C. Huang is with the Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing, 211111, China; and with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, 210096, China (email: huangchen@pmlabs.com.cn).

C.-X.Wang and J. Li are with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, 210096, China; and with the Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing, 211111, China (email: chxwang@seu.edu.cn, junlingli@seu.edu.cn).

Z. Li and Z. Qian are with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, 210096, China (email: li_zheao@seu.edu.cn, zyqian@seu.edu.cn).

Y. Miao is with the Mathematics and Computer Science (EEMCS), Radio Systems (RS), University of Twente, 7522 Enschede, Netherlands (email: y.miao@utwente.nl).

I. Introduction

Currently, the fifth generation (5G) wireless communication network has been rapidly deployed worldwide [1]. However, 5G cannot satisfy all the demands of future wireless communication network applications, thus the research is progressing towards the sixth-generation (6G) wireless communications, with expected initial deployments anticipated by 2030. 6G is expected to expand the existing mobile internet based on 5G, provide five sensory experiences, map the physical and virtual worlds, continuously expand the boundaries and scope of the Internet of Everything (IoE), and finally realize the intelligent connection of the "human-machine-things-environment". Based on the above application and technical demands, the vision of 6G is summarized as global coverage, all frequency spectra, full applications, and strong security [1]. To realize global coverage, 6G will expand the conventional terrestrial mobile communication to the space-air-ground-sea integrated communication network. The sub-6 GHz, millimeter wave (mmWave), terahertz (THz), and optical wireless bands will be taken into consideration to realize the large traffic and connection requirements in all frequency spectra of 6G. The vision of full applications is proposed to serve all kinds of vertical industries by combining a fusion of communications, computing, storage, control, sensing, localization, artificial intelligence (AI), and big data. Strong security is to involve constructing reliable communications, including physical layer security and network layer security, to achieve intelligent native security. Furthermore, as digitalization and intelligence continue to advance in areas of society such as medical care, transportation, and industrial manufacturing, wireless communication networks will no longer be restricted to standard application scenarios and will face greater demands.

One fundamental research aspect to enable the 6G is the studying of radio propagation channels for global coverage, all frequency spectra, and full applications. The wireless channels are the medium of signal transmission in free space and connect a transmitter (Tx) to a receiver (Rx). Channels under different frequency bands, scenarios, and systems, e.g., multiple-input multiple-output (MIMO) and (ultra) massive MIMO, usually show different characteristics, which determine the performance limits of wireless communication systems [2]. Constructing an accurate channel model is essential to analyzing the physical influence on wireless signal transmission, which lays a solid foundation for the theoretical analysis,

performance evaluation, optimization, and deployment of 6G wireless communication systems [3]. Moreover, the accuracy of the channel model directly affects the prediction and evaluation of system performance. If the channel model is inaccurate, the predicted performance parameters, such as BER and channel capacity, may deviate significantly from the actual situation. Therefore, it is crucial to ensure the accuracy of the channel model in order to make reliable predictions and design decisions. However, with the arrival of 6G, wireless channels are confronted with unprecedented challenges and difficulties, such as multiple communication scenarios, utilization of the full channel frequency bands, and processing massive communication data [4]. For conventional wireless channel modeling, various communication scenarios and full channel frequency bands make channel measurement more time-consuming and laborious. In addition, channel measurement campaigns in some extreme cases are too difficult to execute. Massive communication data make the channel estimation algorithms more complicated and increase the difficulty of channel characteristic analysis. Moreover, typical channel analysis methods can only analyze the channel characteristics under the current specific frequency band and scenario with no ability to explore complex relationships between characteristics and different frequencies or scenarios based on the massive communication data. In other words, conventional high-complexity channel models cannot predict the channel characteristics of future times, unknown frequency bands, or unknown scenarios. It is also noteworthy that mobility of user should be considered and it thus makes 6G channel modeling challenging. A pioneering work in [5] proposes the novel idea of mobility-model based non-stationary channel modeling method, and it firstly shows that dynamic motion of user leads to channel non-stationarity by introducing time-variant Doppler component. This finding points to a new direction for the theory of non-stationary channel modeling. Therefore, new bands, new scenarios, and new applications in 6G bring new challenges, making conventional channel measurement and modeling no longer applicable.

The existing channel models are mainly divided into two types: stochastic modeling and geometric modeling. The stochastic channel models use mathematical formulas derived from empirical observations to describe channel statistical characteristics, such as path loss, shadowing, and small-scale fading. On the other hand, the geometrical models focuses on the mathematical characterization of multipath components (MPCs) resulting from radio waves interacting with actual objects in the environment, and results in reflection, scattering, diffraction, and penetration. However, there are some drawbacks and challenges in the existing channel modeling approaches. First, the high-performance channel sounders used in measurement campaigns are very expensive and require extensive manpower. Besides, it is impossible to conduct channel measurements in some extreme environments, e.g., high mountain land and abysmal sea. Moreover, the existing model can only be used to analyze the known/measured channels and environment, in a specific frequency band, and cannot explore new correlations between different bands and scenarios. Finally, accurate channel models are usually complex and cannot predict the channels in unknown bands, different scenarios, and at future times, neither.

Alternative to the conventional channel modeling introduced above, AI-based channel modeling is a new solution facing the challenge of channel measurements for 6G [6]. AI is a critical technique that has been created to allow machines, particularly computer systems, to simulate human intelligence processes in various applications [7]. Machine learning (ML) is a branch of AI that allows computer systems to simulate human learning behaviors to acquire new skills to continuously improve their performance [8]. The continuous development of AI technology has been introduced in wireless communications in different applications during recent decades, [9]–[11]. Due to its advantages of regression and prediction, ML fits the channel prediction problem, therefore, ML-enabled channel modeling is regarded as a powerful tool to make up for the shortcomings of conventional wireless channel modeling methods [12]. Supervised and unsupervised ML have been widely used for channel modeling. In [13], the authors proposed an ML-based MIMO channel prediction method by using the positions of Tx and Rx, and the carrier frequency as the input of a fully feed-forward neural network (FNN) and radial basis function neural network (RBF-NN) to output channel statistical properties, such as received power, root mean square delay spread (RMS DS), and RMS angle spread (AS). An artificial neural network (ANN)-based propagation loss prediction system was presented in [14] that uses the Rx location, effective antenna height, and terrain irregularity as network inputs to estimate the received power at a different location. The work in [15] introduced an ANN-based satellite communication channel prediction method to estimate the received power, with meteorological information incorporated as network parameters, such as air temperature, humidity, rainfall rate, visibility, and relative speed. To increase prediction accuracy, the approach in [15] was improved in [16] by replacing the normal ANN with a long short-term memory (LSTM) and adding a weather condition classification procedure. An LSTM-based multimodal deep learning network (MDLN) was proposed in [17] for downlink channel prediction, which takes received signals and pilots, previous downlink channels, least-squares (LS) estimation results, user location, current uplink channel, and partial downlink channel as network inputs. In [18], the Tx and Rx locations, as well as the frequency band, were used to train an ANN network to predict the amplitude, delay, phase, and cross-polarization ratio at each location point, allowing the MIMO channels to be played back.

From the literature, most of the existing ML-based channel modeling studies focus on enhancing the channel model performance in the measured frequency and temporal-spatial scenario (i.e., non-predictive channel modeling). Nevertheless, predictive channel modeling is only discussed in a few works, e.g., [11], [19], but is an important aspect of global coverage and all-spectra 6G communications. The predictive channel modeling aims to predict the channel characteristics in unknown positions, at future times, and in unknown frequency bands. The details of the predictive channel modeling are illustrated in Section II. In addition, according to extensive existing channel measurements [20]–[23], MPCs that compose the

channels are usually distributed in groups known as clusters. In other words, the MPCs that have similar channel parameters, e.g., power, delay, and angle, are usually modeled as a cluster. In contrast, the MPCs in different clusters usually have significantly different channel parameters from each other. This thus motivates the widely use of cluster-based channel models, e.g., the 6G pervasive channel model [24], COST 259 [25], [26], COST 273 [27], COST 2100 [28], 3GPP Spatial Channel Model [29], and WINNER [30]. Nevertheless, these channel properties have not been well studied in existing ML-based channel modeling. Moreover, it is found that different type clusters, e.g., line-of-sight (LoS) clusters and non-line-of-sight (NLoS) clusters, usually show different propagation properties, [31]–[33]. Nevertheless, none of the existing ML-based channel modeling approaches consider the impact of different cluster types [11].

Wireless channels are usually reconstructed using the clusters/MPCs generated by channel models in channel simulations. Therefore, as long as the power, delay, and angle of the clusters/MPCs are accurate, the reconstructed channel can be considered accurate. When it comes to channels operating in different frequency bands, the primary change lies in the variation of received power due to the fading of multipath in different bands. This variation is directly reflected in the angular PSD. However, the delay of multipath remains relatively constant and experiences minimal changes as the propagation distance remains unchanged. Therefore, this paper focuses on predicting the power and angle parameters of the channels, i.e., angular PSD, in different frequency bands. In this work, we aim to predict the channel characteristics at unknown frequency bands based on the obtained channel characteristics, e.g., angular power spectrum density (angular PSD), at measured frequency bands. In this work, a novel deep learning-based frequency domain predictive channel model is proposed to predict channel characteristics across various frequency bands. The proposed predictive model exploits an autoencoder network to capture large-scale fading properties and a Convolution Gated Recurrent Unit (Conv-GRU) to capture small-scale fading properties, which naturally fits the physical properties of the channel propagation process. It is noteworthy that the proposed predictive channel model is validated using channel data collected from specific frequency bands, as explained in Section IV, due to the limitations of channel measurements. However, the proposed predictive model structure can be applied in higher frequency bands such as millimeter wave by adjusting the algorithm parameters accordingly. Moreover, different types of clusters are separately considered in the learning process. The main contributions can be summarized as follows:

- We propose a novel frequency domain predictive channel model, which accurately predicts the channel characteristics in unknown frequency bands. In the proposed model, different types of clusters, i.e., LoS and NLoS clusters, are individually considered in the learning process. This naturally fits the fact that different types of clusters typically exhibit distinct channel characteristics during wireless propagation.
- To accurately model the channel fading characteristics, the proposed predictive channel model exploits an autoencoder, which is expected to capture the large-scale fading properties, and a set of Conv-GRU cells for each type of cluster, which is expected to capture the small-scale fading properties.
- To validate the proposed predictive channel model, we perform multiple frequency bands channel measurements at the 0.7 GHz, 2.3 GHz, and 3.7 GHz bands at Purple Mountain Laboratories, Nanjing, China. The synthetic channel data generated using the proposed model are compared with the channel measurement data. Furthermore, we implement the widely used LSTM network to validate the predictive performance of the proposed network structure. Meanwhile, a Convolutional-LSTM (Conv-LSTM)-based predictive network and a GRU-based predictive network are employed to validate the choice of the Conv-GRU cell in the proposed model. Furthermore, the extrapolation predictability performance is evaluated using a Ray Tracing (RT) simulator as a performance bound. According to the simulation results, the extrapolation prediction proves accurate when the frequency difference between the training and predicting frequency bands is less than 1.5 GHz.

This paper is organized as follows. Section II summarizes the problem description of the predictive channel modeling. Section III introduces the framework of the proposed frequency domain predictive channel model, where the details of the proposed network are elaborated. Next, the performed channel measurement campaigns are presented in Section IV. Section V then gives the model implementation and validation. Finally, conclusions are drawn in Section VI.

II. Problem Description

In a wireless communication system, the system transmits a known training signal (e.g., pilot signal or reference signal) $\mathbf{x}(t)$, where t denotes the transmitted time. The received signal can be written as $\mathbf{y}(t+\tau)$. The channels are thus represented by a matrix channel impulse response (CIR) $\mathbf{h}(t,\tau)$, where the elements of the matrix represent the CIR between the various Tx and Rx antenna elements. In this work, a CIR for the 6G pervasive channel model [24] is considered, and can be expressed as:

$$\mathbf{h}_{qp,f_c}(t,\tau) = \sqrt{\frac{K_R}{K_R+1}}\mathbf{h}^L_{qp,f_c}(t,\tau) + \sqrt{\frac{1}{K_R+1}}\mathbf{h}^N_{qp,f_c}(t,\tau) \tag{1}$$

where K_R is the Rician factor. $\mathbf{h}^L_{qp,f_c}(t,\tau)$ and $\mathbf{h}^N_{qp,f_c}(t,\tau)$ are the LoS component and the NLoS component, respectively, and can be expressed as (2), (3), where $\{\cdot\}^T$ is the transposition operation, f_c is the frequency band, and $F_{q/p,f_c,V}$ and $F_{q/p,f_c,H}$ are the vertical and horizontal polarization antenna patterns of A^T_p and A^R_q at f_c, respectively [24]. $\phi^R_{E/A}$ and $\phi^T_{E/A}$ are the elevation/azimuth angles of arrival and departure, respectively. $N_{qp}(t)$ is the number of clusters in the channels, $M_n(t)$ is the number of subpaths in the n-th cluster, $k_{m_n}(t)$

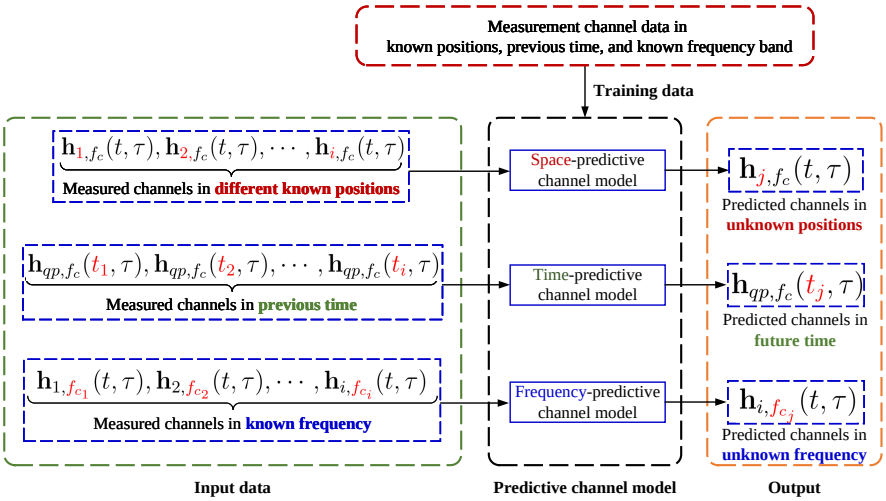


Fig. 1. Key idea of the predictive channel modeling in space, time, and frequency domains.

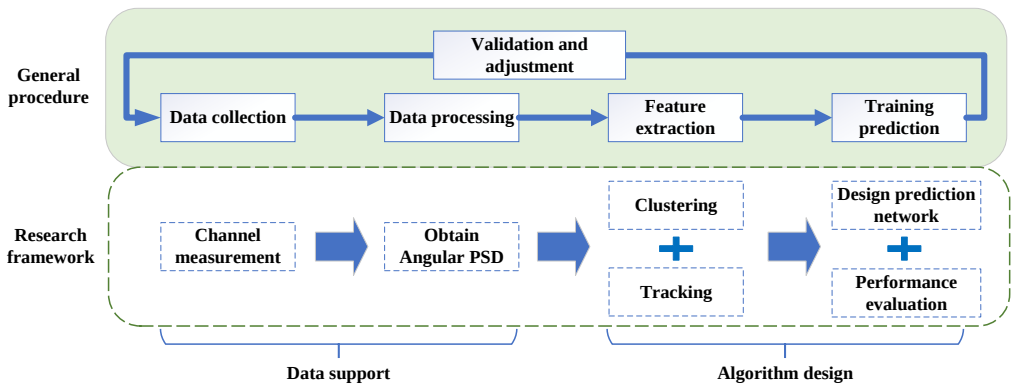


Fig. 2. General procedure of a network design and the framework of our study.

$$\mathbf{h}^{L}_{qp,f_c}(t,\tau) = \begin{bmatrix} F_{q,f_c,V(\phi^R_{E,L}(t),\phi^R_{A,L}(t))} \\ F_{q,f_c,H(\phi^R_{E,L}(t),\phi^R_{A,L}(t))} \end{bmatrix}^T \begin{bmatrix} e^{j\theta^{VV}_L} & 0 \\ 0 & -e^{j\theta^{HH}_L} \end{bmatrix} F_r \begin{bmatrix} F_{p,f_c,V(\phi^T_{E,L}(t),\phi^T_{A,L}(t))} \\ F_{p,f_c,H(\phi^T_{E,L}(t),\phi^T_{A,L}(t))} \end{bmatrix} e^{j2\pi f_c \tau^L_{qp}(t)} \delta(\tau - \tau^L_{qp}(t)) \tag{2}$$

$$\mathbf{h}^{N}_{qp,f_c}(t,\tau) = \sum_{n=1}^{N_{qp}(t)} \sum_{m=1}^{M_n(t)} \begin{bmatrix} F_{q,f_c,V(\phi^R_{E,m_n}(t),\phi^R_{A,m_n}(t))} \\ F_{q,f_c,H(\phi^R_{E,m_n}(t),\phi^R_{A,m_n}(t))} \end{bmatrix}^T \begin{bmatrix} e^{j\theta^{VV}_{m_n}} \sqrt{\mu k^{-1}_{m_n}(t)} e^{j\theta^{VH}_{m_n}} \\ \sqrt{k^{-1}_{m_n}(t)} e^{j\theta^{HV}_{m_n}} \sqrt{\mu} e^{j\theta^{HH}_{m_n}} \end{bmatrix} F_r$$
$$\begin{bmatrix} F_{p,f_c,V(\phi^T_{E,m_n}(t),\phi^T_{A,m_n}(t))} \\ F_{p,f_c,H(\phi^T_{E,m_n}(t),\phi^T_{A,m_n}(t))} \end{bmatrix} \sqrt{P_{qp,m_n,f_c}(t)} e^{j2\pi f_c \tau_{qp,m_n}(t)} \delta(\tau - \tau_{qp,m_n}(t)) \tag{3}$$

and μ are the cross-polarization power ratio and the co-polar imbalance coefficient, respectively. F_r is the rotation matrix, θ^{VV}_L, θ^{HH}_L, $\theta^{VV}_{m_n}$, $\theta^{HH}_{m_n}$, $\theta^{VH}_{m_n}$, and $\theta^{HV}_{m_n}$ are the initial phases with uniform distribution over $(0, 2\pi]$. $P_{qp,m_n,f_c}(t)$ and $\tau_{qp,m_n}(t)$ are the power and delay, respectively, of the m-th subpath in the n-th cluster in the subchannel between the p-th antenna element and the q-th antenna element.

In our previous work, we investigated the ML-based predictive channel modeling in the space domain [34] and the space-time domain [35], [36]. In this paper, we extend our work to the frequency domain and propose a frequency domain predictive channel model based on a new network structure that can provide a higher channel prediction accuracy in the frequency domain. The basic idea of the three domain predictive channel modeling is similar: build a model that can learn and remember the evolution pattern of the channels in known positions, previous time, and known frequency bands, and then predict the channels in unknown space (space predictive channel modeling), future time (time predictive channel modeling) and unknown frequency bands (frequency domain predictive channel modeling). However, due to the difference between learning and predicting objects, it requires different types of neural networks to predict the channels in space, time, and frequency domains. The general idea is shown in Fig. 1. The proposed frequency domain predictive channel model aims to learn the evolution pattern of the channels in a known frequency band, i.e., $\{\mathbf{h}_{qp,f_{c_1},1}, \mathbf{h}_{qp,f_{c_1},2}, \cdots, \mathbf{h}_{qp,f_{c_1},i}\}$ to predict the channels in another frequency band $\mathbf{h}_{qp,f_{c_2},i}$, where $\{f_{c_1,1}, f_{c_1,2}, \cdots, f_{c_1,i}\}$ denotes different frequency points in the f_{c_1} band.

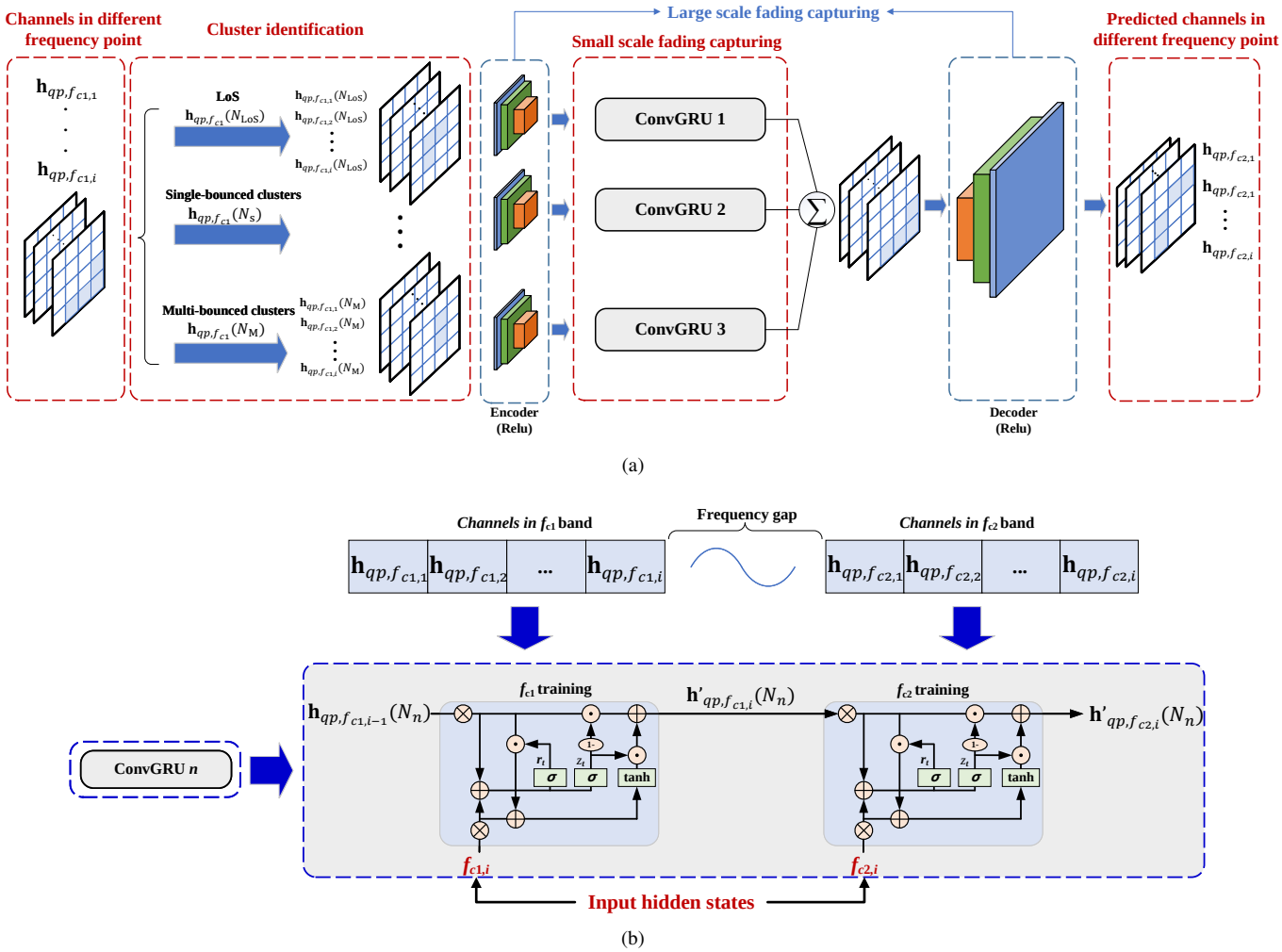


Fig. 3. Framework of the proposed prediction network. (a) The general outline of the prediction network. (b) The network structure of the conducted coupling Conv-GRU.

The general procedure for designing a neural network is given in Fig. 2. Following the procedure, we first conduct the channel measurement in different frequency bands and obtain the angular PSD as the database for training and testing. Then, we identify and track the cluster in consecutive angular PSDs and use identified clusters as training data for designing and evaluating the prediction network. Specifically, we use the angular PSD at each frequency point as training data. The prediction problem can be expressed as:

$$\left[\mathbf{h}_{qp,f_{c_1,i}}(N_n); f_{c_1,i}\ ; f_{c_2,i}\right] \Rightarrow \mathbf{h}_{qp,f_{c_2,i}}(N_n) \tag{4}$$

where $f_{c_1,i} \in [f_{c_1} - B/2, f_{c_1} + B/2]$, $f_{c_2,j} \in [f_{c_2} - B/2, f_{c_2} + B/2]$, N_n is the cluster index, and B is the bandwidth. The details of the data preprocessing and the proposed prediction network are explained in Section III.

III. A Machine Learning-Based Frequency Domain Predictive Channel Model

A. Model framework

The most significant difference between the channel characteristics in different frequency bands is the large-scale fading and small-scale fading. In particular, the difference between the large-scale fading of different frequency bands is embodied in both the LoS path and NLoS paths, whereas the small-scale fading caused by scatterers is only embodied in the NLoS paths (due to the multipath effect). Moreover, according to the propagation process, the reflected paths can be further divided into two types: *single-bounced reflection* and *multi-bounced reflection*. Due to the different propagation processes, the propagation attenuations of the single-bounce reflections and the multi-bounce reflection are expected to be significantly different. Meanwhile, numerous channel measurements show MPCs are usually distributed as clusters as introduced in Sec. I. The channel characteristics of each cluster are generally distinguished, whereas the channel characteristics inside the cluster are similar in at least one domain, e.g., delay or angle domain [20]–[22]. Inspired by this, the proposed prediction network exploits a series of Conv-GRUs [37] expected to capture the evolution pattern of each type of cluster in the frequency domain, whereas the autoencoder is exploited to learn the path loss for all clusters. It is noteworthy that, even LoS paths are usually distributed as a cluster [20].

The framework of the proposed prediction network is given in Figs. 3(a) and (b). The input $\{\mathbf{h}_{qp,f_{c_1,1}}, \mathbf{h}_{qp,f_{c_1,2}}, \cdots, \mathbf{h}_{qp,f_{c_1,i}}\}$ are the channels in

different frequency points in the f_{c_1} band; $\mathbf{h}_{qp,f_{c_1}}(N_{\mathrm{LoS}})$, $\mathbf{h}_{qp,f_{c_1}}(N_{\mathrm{S}})$, and $\mathbf{h}_{qp,f_{c_1}}(N_{\mathrm{M}})$ are the extracted clusters of the LoS cluster, single-bounced clusters, and multi-bounced clusters, respectively, where $N_{(\cdot)}$ is the index of the clusters; the output $\{\mathbf{h}_{qp,f_{c_2,1}}, \cdots, \mathbf{h}_{qp,f_{c_2,i}}\}$ are the predicted channels in different frequency points in the f_{c_2} band. Note that in order to enhance the learning efficiency of spatial features, we use the angular PSDs of the channels as input and output data, instead of directly utilizing the channel matrix. The power distribution of clusters in the angular domain is inherently similar to the distribution of pixels in images, while the channels typically exhibit a strong correlation in consecutive frequency points. Therefore, we employ the Conv-GRU network to capture both the "image pattern" and the "sequential pattern". Specifically, the frequency bands of the input channels and the predicted channels are often not continuous, meaning that there is a frequency gap between them. Considering this, we propose a coupling Conv-GRU network, as illustrated in Fig. 3(b), to mitigate the impact of the frequency gap on the accuracy of predictions. Meanwhile, the known frequency point $f_{c_1,i}$ and the predicting frequency point $f_{c_2,j}$ are set as two hidden states to the coupling Conv-GRU cells. This aims to associate the input and the output (predicted) channels with the known frequency point and the predicted frequency point, respectively. The details of the network parameter settings are given in Section IV.

B. Data preprocessing

It has been found that the performance of the neural network is strongly affected by the data preprocessing [38], [39]. This can significantly reduce the training time and improve the learning accuracy of the network if the known data pattern is extracted in advance before the training. In our case, the main procedure of the data preprocessing includes MPC parameter estimation, cluster identification, and data normalization. The MPCs parameter estimation is usually conducted by using an HRPE algorithm, e.g., space-alternating generalized expectation-maximization (SAGE) [40] or RiMAX [41], which aims to extract the specific channel parameters (such as delay, angle, and power) of discrete MPCs. Nevertheless, most of the existing HRPE algorithms are based on an iterative maximum likelihood method, which usually has a high computation complexity due to the iteration steps. Alternatively, the above channel characteristics can be obtained from angular PSD by using the Bartlett beamformer [42], which has lower computational complexity than HRPE algorithms. Hence, the Bartlett beamformer is adopted in our work to acquire the angular PSD at each frequency point fc_i, which can be expressed as:

$$\mathbf{BBF}_{fc_i}(\phi_1,\phi_2) = \frac{[\mathbf{c}_2(\Omega_{2,l}) \bigotimes \mathbf{c}_1(\Omega_{1,l})]^{\mathrm{H}}\mathbf{R}[\mathbf{c}_2(\Omega_{2,l}) \bigotimes \mathbf{c}_1(\Omega_{1,l})]}{[\mathbf{c}_2(\Omega_{2,l}) \bigotimes \mathbf{c}_1(\Omega_{1,l})]^{\mathrm{H}}[\mathbf{c}_2(\Omega_{2,l}) \bigotimes \mathbf{c}_1(\Omega_{1,l})]} \quad (5)$$

where the vectors $\mathbf{c_2}(\Omega_{2,l})$ and $\mathbf{c}_1(\Omega_{1,l})$ are the steering vectors on the Rx and Tx sides, respectively; the expressions $\Omega_{i,l} = [\cos(\phi_{i,l})\sin(\theta_{i,l}), \sin(\phi_{i,l})\sin(\theta_{i,l}), \cos(\theta_{i,l})]$, and $i = 1, 2$ determine directions/angles in a spherical coordinate system; $\phi_{2,l}, \theta_{2,l}, \phi_{1,l}, \theta_{1,l}$ is the azimuth of arrival, elevation of arrival, azimuth of departure, and elevation of departure, with the range of $(\phi_{i,l}, \theta_{i,l}) \in [-\pi, \pi) \times [0, \pi]$; the symbol $\bigotimes$ denotes the Kronecker product, and $\mathbf{R}$ is the spatial covariance matrix, which is defined as:

$$\mathbf{R} = \mathrm{E}\{\mathrm{vec}(\mathbf{h}_{qp,f_{c_i}}), \mathrm{vec}(\mathbf{h}_{qp,f_{c_i}})^{\mathrm{H}}\} \quad (6)$$

where the $\mathrm{vec}(\cdot)$ is a function that transforms the columns of a matrix into a vector.

After the channel parameter estimation, another important preprocess stage is cluster identification. The clusters are usually identified by conducting ML-based clustering algorithms. In particular, various clustering algorithms have been proposed for different scenarios, e.g., a K-power-means clustering algorithm [43] and a kernel-power-density-based clustering algorithm [44] for static channel clustering, and a trajectory-based clustering algorithm [45] and a power-angle-spectrum-based clustering algorithm [46] for time-varying channel clustering. In our work, the power-angle-spectrum-based clustering proposed in our previous work [46] is used to identify the clusters in angular PSDs. It is worth noting that other existing clustering algorithms can also be employed, the choice of clustering algorithm does not affect prediction accuracy as long as the clusters can be accurately identified for input. In addition, to efficiently capture the data pattern, the input data, i.e., angular PSDs, are normalized on the scale of the whole measurement time.

After data preprocessing, the preprocessed data are sent to the prediction network, which is introduced as follows.

C. AutoEncoder

As shown in Fig. 3(a), a series of autoencoders are expected to capture the large-scale fading. The autoencoder is a typical architecture of feedforward neural network, which generally consists of 3 basic components: encoder, code, and decoder. The input of the autoencoders is the same as the output that can efficiently represent input data through unsupervised learning, as demonstrated in Fig. 4.

- Encoder: Compresses the inputs and produces the code, consisting of a series of progressively decreasing layers.
- Code: Its dimension is generally much smaller than the inputs, so the autoencoder can be used for dimensionality reduction.
- Decoder: Often a mirror image of the encoding architecture, consisting of a series of gradually decreasing layers, and the decoder can reconstruct the inputs only using this code.

An autoencoder exploits a deterministic mapping, which can be expressed as:

$$\mathbf{y}_{\mathrm{out}} = f_\theta(\mathbf{x}_{\mathrm{in}}) = s(\mathbf{W}\mathbf{x}_{\mathrm{in}} + \mathbf{b}) \quad (7)$$

The input vector $\mathbf{x}_{\mathrm{in}} \in [0,1]^d$ is mapped to a hidden representation $\mathbf{y}_{\mathrm{out}} \in [0,1]^{d'}$, parameterized by $\theta = \{\mathbf{W}, \mathbf{b}\}$. $\mathbf{W}$ is a $d' \times d$ weight matrix, and $\mathbf{b}$ is a bias vector. The resulting latent representation $\mathbf{y}_{\mathrm{out}}$ is then mapped again to a

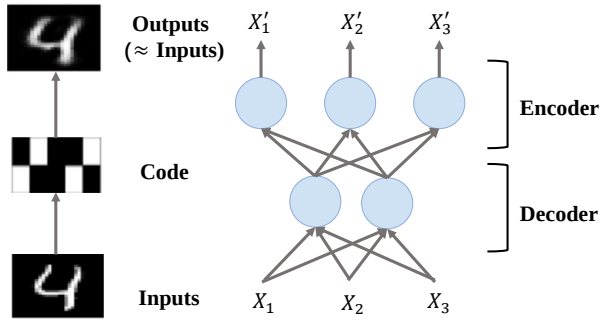


Fig. 4. Network structure of autoencoder.

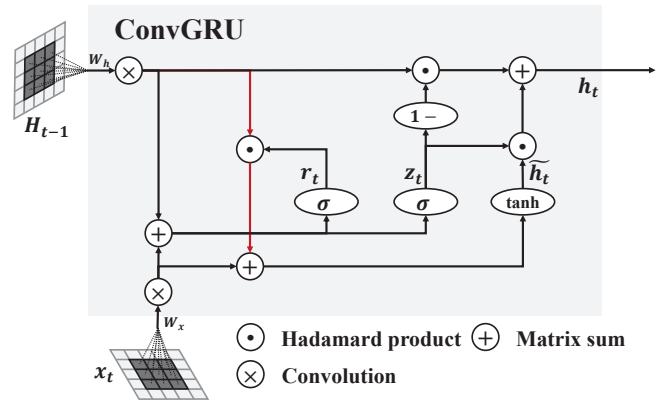


Fig. 5. Network structure of the Conv-GRU model.

reconstructed vector $\mathbf{z_{map}} \in [0,1]^d$ in the input space, which can be expressed as:

$$\mathbf{z}_{\text{map}} = g_{\theta'}(\mathbf{y}_{\text{out}}) = s\left(\mathbf{W}'\mathbf{y}_{\text{out}} + \mathbf{b}'\right) \tag{8}$$

where $\theta' = \{\mathbf{W}', \mathbf{b}'\}$. The weight matrix $\mathbf{W}'$ of the reverse mapping may optionally be constrained by $\mathbf{W}' = \mathbf{W}^T$. Each training $\mathbf{x}_{\text{in}}^{(i)}$ is thus mapped to a corresponding $\mathbf{y}_{\text{out}}^{(i)}$ and a reconstruction $\mathbf{z}_{\text{map}}^{(i)}$. The parameters of this model are optimized to minimize the average reconstruction error:

$$\begin{aligned}\theta^\star, \theta'^\star &= \arg\min_{\theta,\theta'} \frac{1}{n}\sum_{i=1}^{n} L\left(\mathbf{x}_{\text{in}}^{(i)}, \mathbf{z}_{\text{map}}^{(i)}\right) \\ &= \arg\min_{\theta,\theta'} \frac{1}{n}\sum_{i=1}^{n} L\left(\mathbf{x}_{\text{in}}^{(i)}, g_{\theta'}\left(f_\theta\left(\mathbf{x}_{\text{in}}^{(i)}\right)\right)\right)\end{aligned} \tag{9}$$

where L is a loss function. In this study, the mean absolute error (MAE) is adopted as the loss function, which can be expressed as:

$$L(\mathbf{z}_{\text{map}}, \mathbf{z}^*_{\text{map}}) = \frac{\sum_n^{i=1} \left|\left|\mathbf{z}^*_{\text{map}} - \mathbf{z}_{\text{map}}\right|\right|}{n} \tag{10}$$

The ability of the autoencoder to efficiently represent data makes it an expert in data dimensionality reduction and feature extraction. Compared to another widely used feature extraction method, principal component analysis (PCA), the autoencoder can learn the nonlinear characteristics of the training data due to its neural network structure. Thus, the features extracted by the autoencoder can be used to efficiently train the following neural networks. In other words, the autoencoder shows good performance in the pretraining process for the whole network [47]. Moreover, as we expect the network to capture spatial properties from the angular PSDs, we have utilized convolution layers in constructing the AutoEncoder, rather than standard fully connected layers. The detailed network structure and parameters are given in Section V-A.

D. Conv-GRU

As shown in Fig. 3(b), a set of ConvGRUs is individually exploited for each type of cluster. According to existing studies, the LSTM [48] and GRU [49], [50] are good candidates used in various applications [48] for sequence-related data prediction and classification. Besides, the LSTM and the GRU also show good robustness facing the problems of gradient vanishing and exploding due to their concise network structure [51]. Particularly, the GRU is the overwhelming version of the LSTM, which is found to have better performance in terms of convergence in CPU time, parameter updating efficiency, and generalization [49], [50].

On the other hand, neither the conventional fully connected LSTM nor the fully connected GRU with one dimension training vector can effectively extract the multi-dimension data features in the spatiotemporal sequence-related prediction problem, such as video and satellite/radar image datasets. Therefore, a Conv-LSTM is proposed in [52] that replaces normal matrix multiplication with a convolution operation during input-to-state and state-to-state transitions of each gate in an LSTM module. Hence, a Conv-LSTM can well process multi-dimension data as the input to extract spatial features by performing convolution operations. Inspired by [52], a Conv-GRU is proposed in [37] to directly encode the locality and temporal smoothness prior feature of multi-dimension data, e.g., images or angular PSD.

Similar to any GRU, the Conv-GRU is designed to solve the sequence prediction problem and avoid the long-term dependency problem by using gate mechanisms to allow each recurrent unit to adaptively capture dependencies of different time scales. Note that, the time sequence in Conv-GRU is the data input sequence rather than the actual time of the measured channels. As demonstrated in Fig. 5, the Conv-GRU removes or transmits data features to the cell state by updating the gate z_t and resetting the gate r_t, without separating memory cells. The active signal h_t of the Conv-GRU is determined by the previous state h_{t-1}. The updated gate z_t controls how many active signals will be updated during the next iteration. The reset gate r_t allows the ConvGRU to drop (forget) the previously computed state. The Conv-GRU state transitions are expressed as:

$$z_t = \sigma(W_{xz} \otimes x_t + W_{hz} \otimes h_{t-1}) \tag{11}$$

$$r_t = \sigma(W_{xr} \otimes x_t + W_{hr} \otimes h_{t-1}) \tag{12}$$

$$\tilde{h_t} = \tanh(W_{xh} \otimes x_t + r_t \odot (W_{hh} \otimes h_{t-1})) \tag{13}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h_t} \tag{14}$$

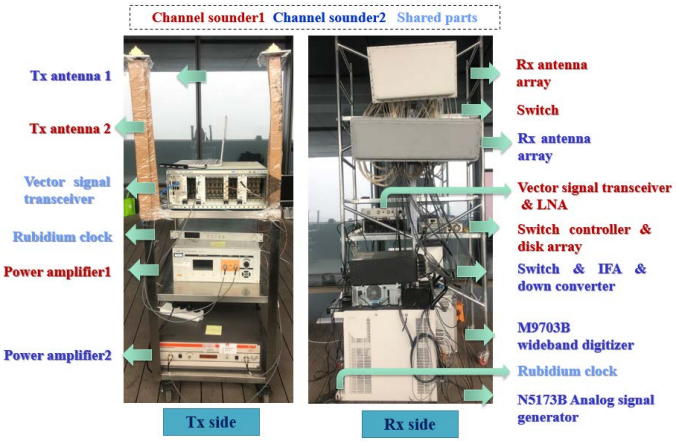


Fig. 6. Photos of the channel sounders used for measurements.

where $\odot$ is the Hadamard product $\otimes$ is the convolution operation. The current moment input of the Conv-GRU unit is x_t. W_{xz}, W_{xr}, W_{xh}, W_{hz}, W_{hr}, and W_{hh} are the weight matrices of the current input x_t and external state h_t, respectively. All Conv-GRU units have the same form of a chain of repeating modules of a neural network to realize spatiotemporal sequence-related data prediction. The detailed network structure and parameters are given in Section V-A.

IV. Channel Measurement Campaigns

In this work, both the training data and testing data are collected from the channel measurements. We conduct channel measurement campaigns by using self-built MISO channel sounders. Specifically, two independent channel sounding systems are implemented at the Rx side to measure the channels in different frequency bands. A National Instrument time-domain channel sounder consisting of a vector signal transceiver and low-noise amplifier at the Rx side is adopted to measure the 0.7 GHz and 2.3 GHz bands channels, whereas a Keysight broadband digital receiver (M9703B), an analog signal generator (N5173B), a down converter, and intermediate frequency amplifier (IFA) are adopted at the Rx side to measure the 3.7 GHz band channels. Meanwhile, a vector signal generator and a power amplifier are exploited at the Tx side. A pair of Rubidium Clocks are used to synchronize the sounding system at both the Tx and Rx sides. In addition, the Rubidium Clock at the Rx side is used to synchronize two Rx sounding systems.

Fig. 6 shows a photo of the channel sounders in our measurement campaigns. An omnidirectional antenna is at the Tx side, whereas two 4×8 switch plane arrays are adopted at both Rx sides. Notably, a larger size of antenna typically offers angular domain higher resolution and provides larger data volumes for training. Consequently, this could potentially result in higher prediction accuracy. The transmission signals were executed in snapshots, where the time duration of each snapshot was 800 μs, and the bandwidth was 100 MHz for all the 0.7 GHz, 2.3 GHz, and 3.7 GHz channels. The details of the antenna arrays and the measurement campaign can be found in [53]. Table. I gives the system setups of the measurement.

The measurement is conducted at Purple Mountain Laboratories, Nanjing, China. Purple Mountain Laboratories is a typical urban communication scenario, i.e., the buildings are located on the roads with an average height of 35 m, the roadside green belt is distributed along the roads, and the interval between each building is approximately 30–40 m. The Rx is placed on the eighth floor of the main building (approximately 30 m height, nearly the top of the building), whereas the Tx moves along different routes, as shown in Fig. 7. Routes L1 and L2 denote LoS scenarios whereas routes N1–N4 denote NLoS scenarios (the LoS path is blocked by other buildings). Figs. 8(a)–(c) show two samples of angular PSD, which include the azimuth of arrival (AoA) and elevation of arrival (EoA), obtained from channels at 0.7 GHz, 2.3 GHz, and 3.7 GHz, respectively. A slight difference can be found among the angular distributions of the angular PSDs of the 0.7 GHz, 2.3 GHz, and 3.7 GHz channels, whereas the receiving powers of the channels at the three bands are significantly different. This is because of the similar propagation environment and distinctively different path loss fading

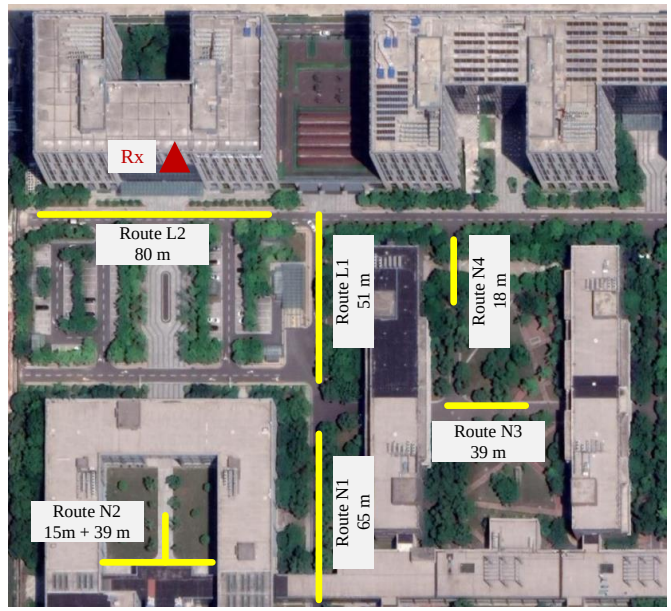
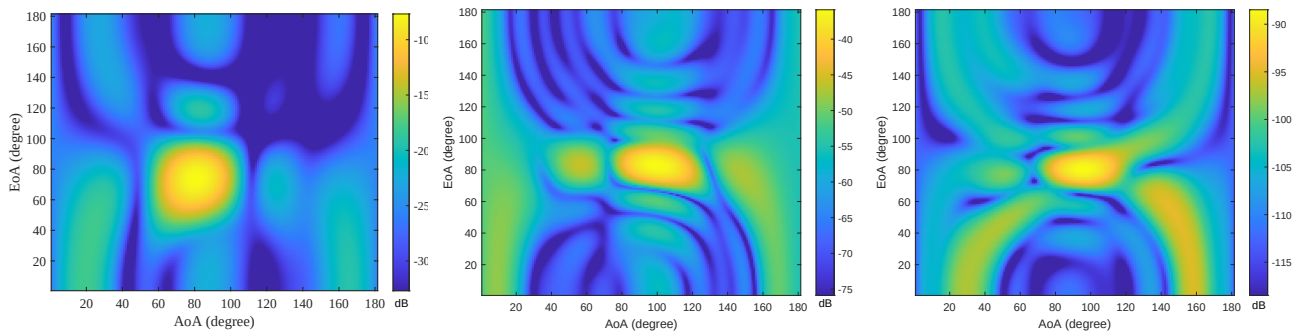


Fig. 7. Top view of measurement routes.

TABLE I
System Setup of Multiple Frequency Bands Measurements.

Parameter	Value		
Carrier frequency	0.7 GHz	2.3 GHz	3.7 GHz
Number of Rx antennas	4×4	4×8	
Number of Tx antennas	1		
Bandwidth	100 MHz		
Time duration per snapshot	800 μs		
Transmit power	40 dBm		
Number of frequency points	1023		



(a) (b) (c)

Fig. 8. Angular PSD samples of (a) 0.7 GHz channels, (b) 2.3 GHz channels, and (c) 3.7 GHz channels.

and frequency-selective channel property of the different band channels.

V. Model Implementation and Validation

In this section, the proposed predictive channel model is validated by using the measurement data. The model implementation is first presented, which provides the data size of the training data and testing data. The network parameter is presented next, which provides the detailed parameters of the predictive network in the proposed predictive channel model. Then, the prediction accuracy of the proposed predictive model is validated by comparing the predicted channels and measured channels.

A. Model Implementation

In the proposed predictive channel model, there are 3 conventional layers of the encoder network with kernel size of (5,5,5), (3,3,16), (3,3,144) in each layer, and 3 conventional layers of the decoder networkwith kernel size of (3,3,144), (3,3,16), (5,5,5) in each layer as well. The linear rectification function (ReLU) is adopted as the active function. The encoder transforms the data feature size from 180×180 into 36× 36×5, 12× 12× 16, and 4× 4×144. The max epoch is set to 500. For the Conv-GRU, there is 1 layer network with 2 Conv-GRU cells for each type of cluster. The kernel size of each Conv-GRU is set to 3×3 with the SAME padding (padding value: 0); the batch size and stride are set to 32 and 1, respectively. Table II lists the details of the network layout. According to the network structure in Fig. 3 and network layout in Table II, the total number of parameters in the proposed model is 620,705.

B. Validation Method

We apply the channel measurement data to our proposed frequency domain predictive channel model. There are a total of 12350 data samples collected from the channel measurement, which contain 8620 data samples for LoS scenarios and 3730 data samples for NLoS. Depending on the selection of training data and the predicting data, there are two different training and prediction strategies: i) *interpolation prediction*, which uses the lower and higher frequency band channels as training data and predicts the middle frequency band channels, as illustrated in Fig. 9(a); and ii) *extrapolation prediction*, which uses the lower frequency band channels as training data and predicts the higher frequency band channels, as illustrated in Fig. 9(b).

TABLE II
Network Layout.

Parameters	Value
Layers of Encoder and Decoder	3,3
Number of Conv-GRU layer	1
Number of Conv-GRU	2
Kernel size (AutoEncoder)	(5,5,4), (3,3,16), (3,3,144)
Kernel size (Conv-GRU)	3×3
Loss function	MAE
L2 weight regularization	0.004
Active function	ReLU
Sparsity regularization	6
Sparsity proportion	0.2
Maximum epochs	500
Learning rate	0.05
Padding	SAME
Padding value	0
Stride	1
Batch size	32

To evaluate the prediction accuracy of the proposed predictive channel model, the root means square error (RMSE) and mean absolute percentage error (MAPE) are used to measure the difference between the predicted angular PSDs and the measured angular PSDs. Both RMSE and MAPE have been widely used for sequence data prediction evaluation. Specifically, RMSE measures the *absolute deviation* between the predicted value and the actual value, whereas MAPE measures the *relative deviation* between the predicted value

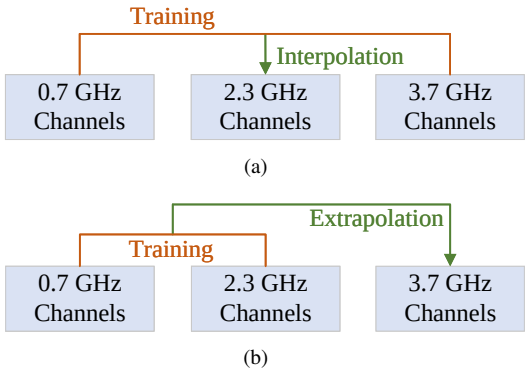


Fig. 9. Two different training and prediction strategies. (a) Interpolation prediction. (b) Extrapolation prediction.

and the actual value. The RMSE ($\Re$) can be expressed as:

$$\Re = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\hat{\mathbf{h}}_{qp,f_{c_i,j}} - \mathbf{h}_{qp,f_{c_i,j}}\right)^2} \tag{15}$$

where N is the total number of predicted data, and $\hat{\mathbf{H}}(f_{c_i}, j)$ and $\mathbf{H}(f_{c_i}, j)$ are the angular PSDs of predicted channels and measured channels, respectively. The MAPE ($\Im$) can be expressed as:

$$\Im = \frac{1}{N}\sum_{j=1}^{N}\left|\frac{\hat{\mathbf{h}}_{qp,f_{c_i,j}} - \mathbf{h}_{qp,f_{c_i,j}}}{\mathbf{h}_{qp,f_{c_i,j}}}\right| \times 100\% \tag{16}$$

C. Model Accuracy

Fig. 10 gives the angular PSD prediction results by using the interpolation prediction strategy, where (a) and (b) are measured and predicted angular PSDs of 2.3 GHz channels, respectively. It can be found that the proposed predictive channel model can well predict the power distribution in the angular domain. This indicates that the proposed predictive channel model can well capture the channel spatial characteristics at different frequency band channels. Particularly, the clusters with high power, e.g, LoS cluster, are predicted more accurately compared with the clusters with lower power. This may be because the clusters with higher power acquire more attention from the network, i.e., higher feedback weights, during the learning progress. Moreover, it can be found that the average power of the predicted channel matches the average power of the measured channels well. This indicates that the proposed predictive channel model can well capture the large-scale fading difference between different frequency band channels.

As a comparison to the interpolation prediction strategy, the angular PSD predictions of 0.7 GHz and 3.7 GHz channels by using the extrapolation prediction strategy are simulated. Figs. 11(a) and (b) show the measured angular PSDs of the 0.7 GHz and 3.7 GHz channels, respectively, whereas Figs. 11(c) and (d) give the corresponding predicted angular PSDs of the 0.7 GHz and 3.7 GHz channels, respectively. By comparing the prediction results in Fig. 10 and Fig. 11, it can be found that the prediction accuracy of the interpolation prediction is significantly higher than that of the extrapolation prediction.

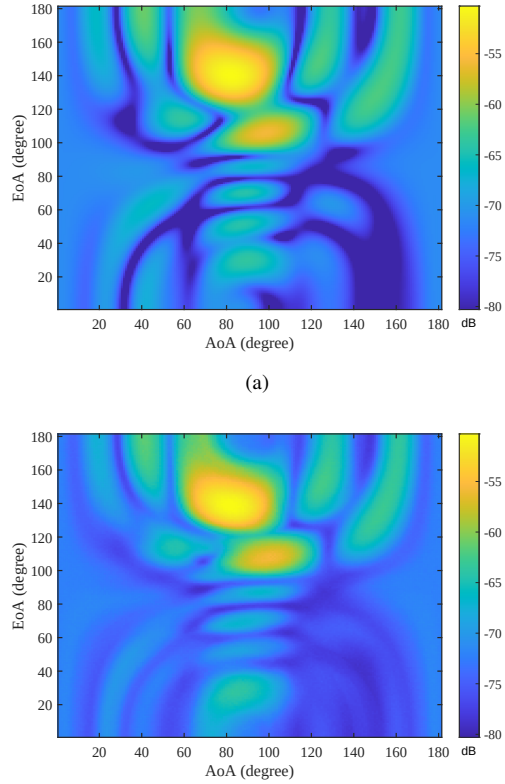


Fig. 10. Angular PSD comparison between the measurement channels and the interpolation predicted channels. (a) Measured angular PSD of 2.3 GHz channels. (b) Predicted angular PSD of 2.3 GHz channels by using the proposed predictive channel model with interpolation prediction strategy.

Although, the extrapolation prediction can still well predict the clusters with a higher power, but cannot accurately predict the clusters with a relatively lower power. This may be because of the *anchoring effect* of the interpolation: the trained lower and higher frequency band channels give a clear boundary for predicting middle frequency band channels. On the other hand, it is open-ended for extrapolation prediction strategy, which is more challenging to accurately predict the channels in unknown frequency bands. In addition, the accuracy of the extrapolation prediction of 0.7 GHz and 3.7 GHz channels are similar, where the quantified evaluation is given below.

To quantitatively evaluate the prediction accuracy, we use the RMSE and MAPE with two validation cases, as introduced earlier, for two validation cases to quantify the deviation between the measured channels and the corresponding predicted channels. A lower RMSE and MAPE indicate smaller difference between measured and predicted channels, which ultimately indicates the higher prediction accuracy. In addition, to further compare the predictive network performance with existing and widely used machine learning methods, LSTM [48] and GRU [49], [50], are conducted as comparison methods. It is noteworthy that, in the comparison experiments, the network structure remains the same, but the Conv-GRU cell is replaced by the Conv-LSTM cell and GRU cell. Furthermore, to verify the performance of the predictive network structure, a general LSTM network, where the autoencoder is removed, is adopted as another comparison method. A fully connected

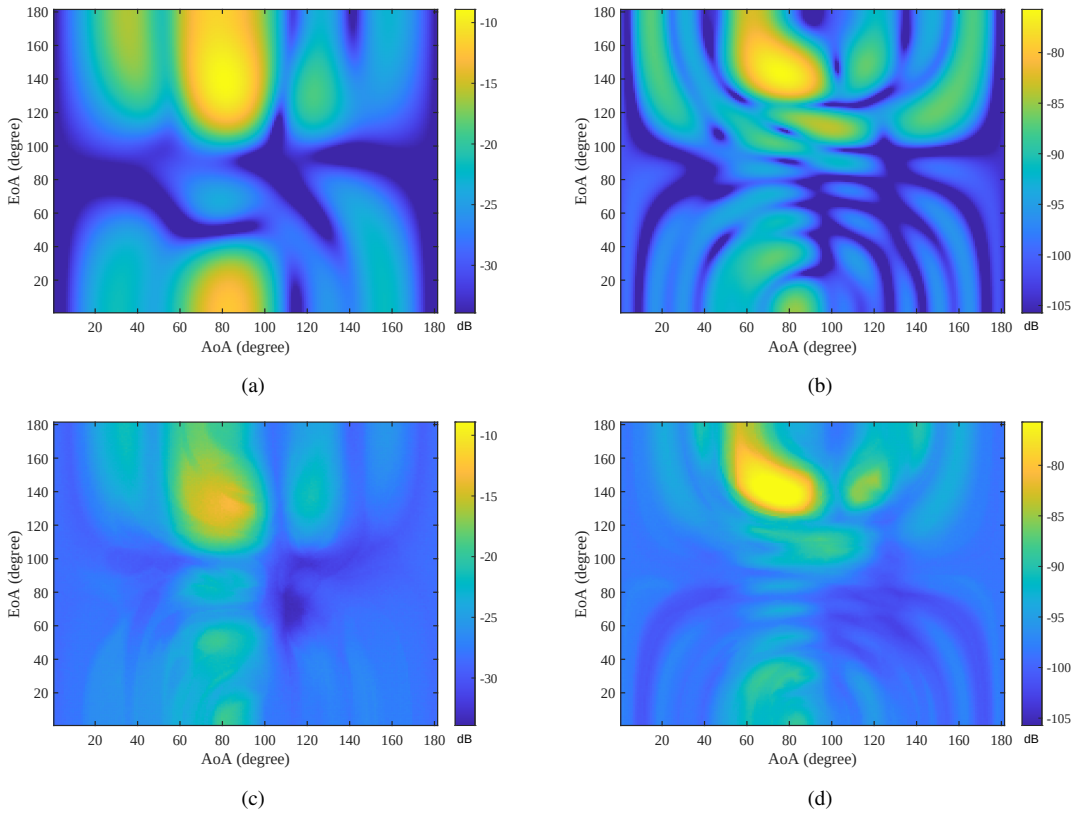


Fig. 11. Angular PSD comparison between the measurement channels and the extrapolation predicted channels. (a) Measured angular PSD of 0.7 GHz channels. (b) Measured angular PSD of 3.7 GHz channels. (c) Predicted angular PSD of 0.7 GHz channels. (d) Predicted angular PSD of 3.7 GHz channels.

network is used before the general LSTM/GRU for data size transformation.

Table III lists the prediction performance comparison results, where the most accurate prediction results are emphasized in bold. According to the comparison, the proposed predictive channel model, GRU-based network, and Conv-LSTM-based network achieve significantly higher accuracy than the general LSTM network. This indicates that the proposed predictive network structure has a better ability to capture the evolution channel characteristics and predict them in the frequency domain compared to a general predictive network. Compared with other methods, the proposed predictive channel model achieves the highest prediction accuracy in most cases with both interpolation and extrapolation prediction strategies. While it may not always outperform other methods in every single case, it consistently demonstrates superior performance in most scenarios. Particularly, the GRU-based network achieves slightly higher prediction accuracy compared with the Conv-LSTM-based network. This indicates that the Conv-GRU in the predictive channel model can better capture the spatial feature of the angular PSD and ultimately improve the prediction accuracy.

Specifically, due to the anchoring effect, all four prediction methods achieve significantly higher prediction accuracy by using an interpolation prediction strategy. This is also consistent with the observation from Fig. 10 and Fig. 11, and gives the intuition that the interpolation prediction strategy is more recommended for the practical implementation of channel prediction.

D. Model Robustness

According to the existing studies [38], [39], the selection of training and testing data shows a significant impact on the ML algorithm performance. Therefore, in this work, we use two different validation cases:

1) Validation Case 1: We first conduct a widely used K-fold cross-validation [54]:

i. Each of the databases $\mathbf{h}_{qp,f_{c_i}}$ is divided into K disjoint sets, $\mathbf{h}_{qp,f_{c_i,j}} = \{\mathbf{h}_{qp,f_{c_i,1}} \cup \mathbf{h}_{qp,f_{c_i,2}} \cup \cdots \cup \mathbf{h}_{qp,f_{c_i,K}}\}$, with $\mathbf{h}_{qp,f_{c_i,k_1}} \cap \mathbf{h}_{qp,f_{c_i,k_2}} = \emptyset$, for $k_1 \neq k_2$).
ii. We conduct the training based on the set $\mathbf{h}_{qp,f_{c_i}}$ after removing the samples of set $\mathbf{h}_{qp,f_{c_i,j}}$, and then use the set $\mathbf{h}_{qp,f_{c_i,j}}$ for validation.
iii. We then revalidate the results for each $j = 1, \cdots, K$. In our case, we set $K = 10$ while considering the trade-off between computational complexity and accuracy.

2) Validation Case 2: The K-fold cross-validation iteratively evaluates the training performance over the whole datasets, it thus shows the overall performance and robustness of the proposed algorithm. However, in practice, it is too expensive and exhausting to measure the channels of all environments in the city and preprocess them as training data. Therefore, it is more practical to collect channel data from typical scenarios and use them to train a predictive channel model for other similar scenarios. Hence, we then use the dataset collected from different positions for the training and testing process for a more practical validation.

Table IV summarizes the details of the validation layout

TABLE III
PREDICTION PERFORMANCE COMPARISON

Method	Prediction Strategy	RMSE	MAPE (%)	NMSE
General LSTM network	Interpolation	13.97	12.63	0.0159
	Extrapolation	19.96	19.42	0.0218
GRU-based network	Interpolation	7.99	6.65	0.0099
	Extrapolation	12.78	11.69	0.0148
Conv-LSTM-based network	Interpolation	8.06	6.88	0.0101
	Extrapolation	10.45	11.62	0.0125
Proposed predictive	Interpolation	**6.95**	**6.17**	**0.0091**
	Extrapolation	**10.75**	**9.85**	**0.0128**

TABLE IV
VALIDATION LAYOUT PARAMETERS.

Data samples	Number of data samples
Total data samples (LoS/NLoS)	12350 (8620/3730)
Training data for Case 1* (LoS/NLoS)	11115 (7758/3357)
Validating data for Case 1* (LoS/NLoS)	1235 (862/373)
Training data for Case 2 (LoS/NLoS)	8645 (6034/2611)
Validating data for Case 2 (LoS/NLoS)	3705 (2586/1119)

*Using K-fold cross validation, where $K = 10$.

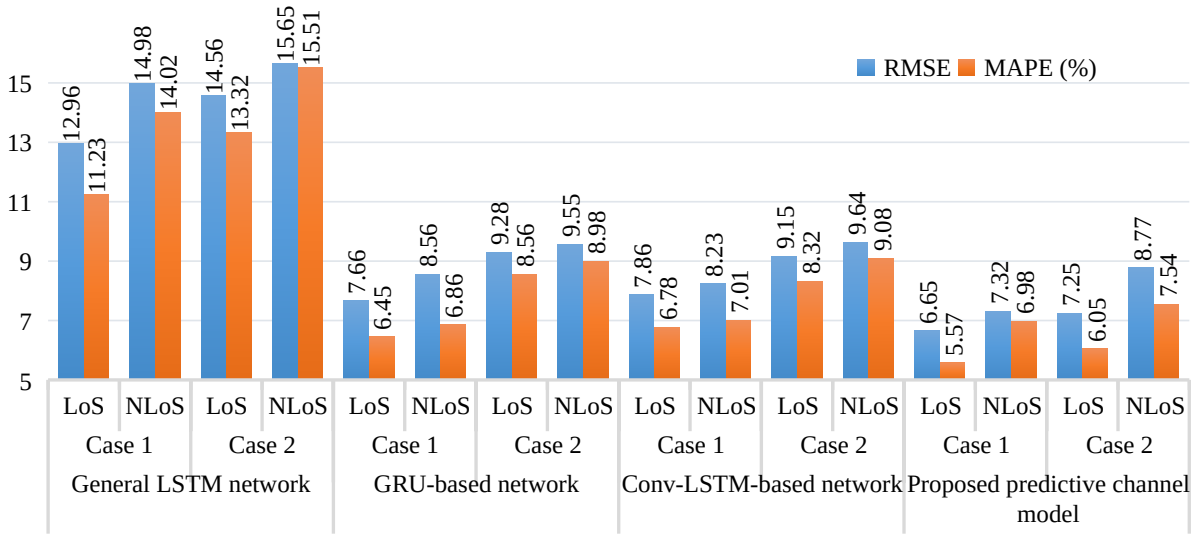


Fig. 12. Prediction performance comparison between different validation cases.

parameters. Note that, all simulations are conducted by using the *interpolation prediction strategy*. Fig. 12 shows the prediction performance of all four prediction methods in different validation cases. Consisting with the simulation results in Table III, the proposed predictive channel model outperforms the other prediction methods in most cases. Meanwhile, all four prediction methods achieve better prediction accuracy in LoS scenarios. This is because the complicated environment in NLoS scenarios leads to more complicated channel characteristics and ultimately brings more challenges to predicting the unknown band channels. Furthermore, the accuracy of all prediction methods decreases in Case 2 simulations, compared with the prediction accuracy in Case 1 simulation. This is caused by the different selections of testing and validating data. Although exclusively training and validating data are used in both two validation cases' simulations, the validating

TABLE V
TRAINING AND PREDICTING STRATEGY BASED ON RT SIMULATED CHANNELS

	Frequency Band of Training Channels	Frequency Band of Predicting Channels
Case 1	2.5 GHz	3 GHz, 3.5 GHz, 4 GHz, 4.5 GHz, 5 GHz, 5.5 GHz, 6 GHz
Case 2	2.5 GHz, 3 GHz	3.5 GHz, 4 GHz, 4.5 GHz, 5 GHz, 5.5 GHz, 6 GHz
Case 3	2.5 GHz, 3 GHz, 3.5 GHz	4 GHz, 4.5 GHz, 5 GHz, 5.5 GHz, 6 GHz
Case 4	2.5 GHz, 3 GHz, 3.5 GHz, 4 GHz	4.5 GHz, 5 GHz, 5.5 GHz, 6 GHz
Case 5	2.5 GHz, 3 GHz, 3.5 GHz, 4 GHz, 4.5 GHz	5 GHz, 5.5 GHz, 6 GHz

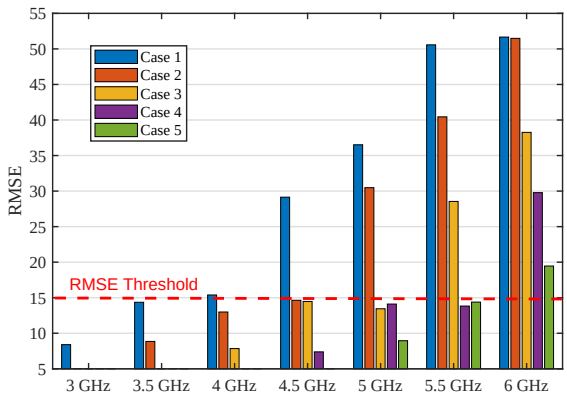


Fig. 13. RMSE between the predicted channels and RT simulated channels by using the training and prediction strategies given in Table V.

data and training data used in Case 1 simulations are collected from the same street, whereas the training and validating data used in Case 2 simulations are collected from different streets. This proves that the selection of training data significantly impacts prediction accuracy, and the collection of channel measurement data is still one of the crucial issues for predictive channel modeling.

E. Prediction Boundary Evaluation

To determine the extrapolation prediction boundaries of the proposed predictive channel model, a series of indoor communication simulations were executed using the RT software Wireless InSite® (WI) for a comprehensive assessment. Within WI, an indoor office scenario measuring $46\ m \times 17\ m$ was modeled, encompassing both LoS and NLoS scenarios. The Tx is positioned at the corner of the room, while the Rxs are randomly distributed across 100 positions within the room. The bandwidth is set to 100 MHz (simulated by 100 frequency points), enabling the generation of synthetic channels from 2.5 GHz to 6 GHz. Specifically, channels from 2.5 GHz, 3 GHz, 3.5 GHz, 4 GHz, and 4.5 GHz are employed as training data to predict channels at 3.5 GHz, 4 GHz, 4.5 GHz, 5 GHz, 5.5 GHz, and 6 GHz in different combinations, as listed in Table V.

Fig. 13 shows the RMSE between the predicted channels and the RT simulated channels by using the training and prediction strategies given in Table V. According to the evaluation of the existing AI-based channel modeling accuracy [11], we set a threshold (RMSE $\leq$ 15) to decide whether a prediction result is accurate or not, as plotted in Fig. 13. It can be found that the RMSE between the predicted channels and the RT simulated channels generally increases when the frequency band of the predicted channels increases. Specifically, in Case 1, by using the 2.5 GHz channels as training data, the predictive channel model can achieve a fairly good accuracy when predicting the 3 GHz channels, but lost credibility when predicting the channels at higher frequency bands. And this tendency can be found in other simulated cases. This is mainly because the correlation between the training and predicting channels decreases with the increasing frequency difference between the training and predicting channels, and the lower the correlation, the more difficult it is to predict. Moreover, according to the RMSE threshold, the RMSE between the predicted channels and the RT simulated channels are significantly increased when the frequency gap between the training and predicting frequency band is larger than 1.5 GHz. Therefore, it gives the intuition that the extrapolation prediction accuracy of the proposed channel is generally promised when the frequency gap between the training and predicting frequency band is smaller than 1.5 GHz. It is noteworthy that, the predicted limitation of 1.5 GHz is relative and is influenced by the frequency non-stationarity of the channels. For wireless channels, the coherent bandwidth typically expands with higher carrier frequency band [55]. Consequently, the frequency non-stationarity reduces at higher frequency bands within the same bandwidth. In another word, the difference between sub-channels is less pronounced at higher frequency bands within the same bandwidth. For instance, the channel variation between 100 GHz and 101.5 GHz is smaller than that between 5 GHz and 6 GHz. This notably reduces the prediction challenges, allowing for a relatively larger prediction limitation for frequency gaps.

Besides, compared to other cases, Case 5 has a larger training database. Nevertheless, the larger training database fails to translate into a higher prediction accuracy for Case 5. Moreover, the RMSE of Case 5 experiences a significant increase when the frequency gap between the training and predicting frequency band exceeds 1.5 GHz. This is because the training channel data has a limited impact on prediction

when there is a significant difference in frequency between the training and prediction channels.

Overall, according to our experiments, we can conclude that: i) for sub-6 GHz band, the extrapolation prediction can be only trusted when the frequency difference between the training and predicting frequency bands is smaller than 1.5 GHz; ii) the training data that covers larger frequency bands does not lead to a higher channel prediction accuracy.

VI. Conclusion

In this paper, a deep learning-based novel frequency domain predictive channel model has been proposed. In the proposed model, a set of Conv-GRU cells have been exploited to capture the channel characteristics of different types of clusters, i.e., the LoS cluster, the single-bounced clusters and the multi-bounced clusters, whereas an autoencoder is employed to capture the large-scale fading feature. To validate our predictive channel model, we have conducted channel measurements at 2.3 GHz and 3.7 GHz bands in urban scenarios. Then, we have implemented the proposed predictive channel model using the measurement data, where two different prediction strategies, i.e., interpolation prediction and extrapolation prediction, have been compared and analyzed. Furthermore, several widely used ML approaches have been conducted as comparisons, i.e., a general LSTM, a GRU-based and an Conv-LSTM-based predictive network. Besides, the boundary of the extrapolation predictability has been analyzed based on the RT simulator. The results have shown that the proposed predictive channel model can outperform the other models in most cases. Particularly, the proposed predictive channel can achieve the best prediction performance by using the interpolation strategy, e.g., the RMSE, MAPE, and NMSE between the channel measurement data and prediction data are around 7, 6% and 0.01, respectively. The a performance bound for extrapolation predictability has been investigated. It can be concluded from the experiments that: i) the predictive channel modeling in NLoS scenarios is more challenging than it is in LoS scenarios; ii) building a comprehensive channel measurement database is crucial for accurate predictive channel modeling; iii) increasing the frequency band coverage of the training data does not necessarily result in higher accuracy in channel prediction, and iv) accurate extrapolation prediction can be achieved for sub-6 GHz band channels when the frequency difference between the training and predicting frequency bands is less than 1.5 GHz.

References

[1] X.-H. You, C.-X. Wang, J. Huang, et al., "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 1–74, Jan. 2021.

[2] A. F. Molisch, "Wireless Communications," *John Wiley & Sons*, 2012.

[3] C.-X. Wang, J. Huang, H. Wang, X. Gao, X.-H. You, and Y. Hao, "6G wireless channel measurements and models: Trends and challenges," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 22–32, Dec. 2020.

[4] C.-X. Wang, J. Bian, J. Sun, W. Zhang, and M. Zhang, "Survey of 5G channel measurements and models," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3142–3168, Aug. 2018.

[5] Ruisi He, Bo Ai, Gordon L. Stüber, and Zhangdui Zhong, "Mobility model-based non-stationary mobile-to-mobile channel modeling," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4388–4400, July 2018.

[6] Z. Zhang, J. Zhang, Y. Zhang, L. Yu and G. Liu, "AI-based time-, frequency-, and space-domain channel extrapolation for 6G: Opportunities and challenges," *IEEE Veh. Technol. Mag.*, vol. 18, no. 1, pp. 29–39, March 2023.

[7] S. Bi, R. Zhang, Z. Ding, and S. Cui, "Wireless communications in the era of big data," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 190–199, Oct. 2015.

[8] T. G. Dietterich, "Machine-learning research," *AI Mag.*, vol. 18, no. 4, pp.97–97, Dec. 1997.

[9] C.-X. Wang, M. D. Renzo, S. Stanczak, S. Wang, and E. G. Larsson, "Artificial intelligence enabled wireless networking for 5G and beyond: Recent advances and future challenges," *IEEE Wireless Commun.*, vol. 27, no. 1, pp.16–23, Feb. 2020.

[10] C. Huang et al., "Artificial intelligence enabled radio propagation for communications–Part I: Channel characterization and antenna-channel optimization," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3939–3954, June 2022.

[11] C. Huang et al., "Artificial intelligence enabled radio propagation for communications–Part II: Scenario identification and channel modeling," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3955–3969, June 2022.

[12] S. Zhang, Y. Zhang, Z. Zhang, J. Zhang, L. Xia and T. Jiang, "Deep learning-based time-varying channel prediction for MIMO systems," *Proc. IEEE VTC2022-Spring*, Helsinki, Finland, 2022, pp. 1-5.

[13] J. Huang, C.-X. Wang, L. Bai, et al., "A big data enabled channel model for 5G wireless communication systems," *IEEE Trans. Big Data*, vol. 6, no. 2, pp. 211–222, June 2020.

[14] A. R. Ozdemir, M. Alkan, M. Kabak, et al., "The prediction of propagation loss of FM radio station using artificial neural network," *J. Electromagn. Anal. Appl.*, vol. 6, no. 11, pp. 358–365, Sept. 2014.

[15] L. Bai, C.-X. Wang, Q. Xu, S. Ventouras, and G. Goussetis, "Prediction of channel excess attenuation for satellite communication systems at Q-band using artificial neural network," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2235–2239, Nov. 2019.

[16] L. Bai, Q. Xu, S. Wu, S. Ventouras, and G. Goussetis, "A novel atmosphere-informed data-driven predictive channel modeling for B5G/6G satellite-terrestrial wireless communication systems at Q-Band," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14225-14237, Dec. 2020.

[17] Y. Yang, F. Gao, Z. Zhong, B. Ai, and A. Alkhateeb, "Deep transfer learning-based downlink channel prediction for FDD massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7485–7497, Dec. 2020.

[18] X. Zhao et al., "Playback of 5G and beyond measured MIMO channels by an ANN-based modeling and simulation framework," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1945–1954, Sept. 2020.

[19] K. Mao, Q. Zhu, F. duan, Y. Qiu, M. Song, W. Fan, and Y. Miao, "A2G channel measurement and characterization via TNN for UAV multi-scenario communications," in *Proc. GlobeCom' 2022*, Rio de Janeiro, Brazil, Dec. 2021.

[20] C. Huang et al., "Geometry-cluster-based stochastic MIMO model for vehicle-to-vehicle communications in street canyon scenarios," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 755–770, Feb. 2021.

[21] M. Yang et al., "Measurements and cluster-based modeling of vehicle-to-vehicle channels with large vehicle obstructions," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5860–5874, Sept. 2020.

[22] R. He, B. Ai, G. Wang, M. Yang, C. Huang and Z. Zhong, "Wireless channel sparsity: Measurement, analysis, and exploitation in estimation," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 113–119, Aug. 2021.

[23] C. Huang et al, "Geometry-cluster-based stochastic MIMO model for vehicle-to-vehicle communications in street canyon scenarios," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 755-770, Feb. 2021.

[24] C.-X. Wang, Z. Lv, X. Q. Gao, X.-H. You, Y. Hao, and H. Haas, "Pervasive wireless channel modeling theory and applications to 6G GBSMs for all frequency bands and all scenarios," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9159–9173, Sept. 2022.

[25] A. F. Molisch, H. Asplund, R. Heddergott, M. Steinbauer, and T. Zwick, "The COST259 directional channel model-part I: overview and methodology," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3421–3433, Dec. 2006.

[26] H. Asplund, A. A. Glazunov, A. F. Molisch, K. I. Pedersen, and M. Steinbauer, "The COST 259 directional channel model–Part II: Macrocells," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3434– 3450, Dec. 2006.

[27] H. Hofstetter, A. F. Molisch, and N. Czink, "A twin-cluster MIMO channel model," in *Proc. 2006 1st Eur. Conf. Antennas Propag.*, Nice, 2006, pp. 1–8.

[28] L. Liu *et al.*, "The COST 2100 MIMO channel model," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.

[29] 3GPP TR 25.996, "Spatial channel model for multiple input multiple output (MIMO) simulations," Jul. 2018.
[30] J. Meinila, P. Kyosti, T. Jamsa, and L. Hentila, "WINNER II channel models," in *Proc. Radio Technol. Concepts IMT-Adv*, 2009, pp. 39–92.
[31] J. Karedal *et al.*, "A geometry-based stochastic MIMO model for vehicle-to-vehicle communications," *IEEE Trans, Wireless Commun.*, vol. 8, no. 7, pp. 3646–3657, Jul. 2009.
[32] O. Renaudin, V. Kolmonen, P. Vainikainen, and C. Oestges, "Wideband measurement-based modeling of inter-vehicle channels in the 5-GHz band," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3531–3540, Oct. 2013.
[33] C. Huang, R. Wang, C.-X. Wang, P. Tang, A. F. Molisch, "A geometry-based stochastic model for truck communication channels in freeway scenarios" *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5572–5586, Aug. 2022.
[34] L. Yu and C.-X. Wang, "Prediction of wireless mmWave massive MIMO channel characteristics based on graph attention networks," in *Proc. ICETC'20*, Tokyo, Japan, pp. 1–6, Dec. 2020.
[35] Z. Li, C.-X. Wang, J. Huang, W. Zhou, and C. Huang, "A GAN-GRU based space-time predictive channel model for 6G wireless communication systems," *IEEE Trans. Veh. Technol.*, submitted for publication.
[36] Z. Li, C. -X. Wang, J. Huang, W. Zhou and C. Huang, "A GAN-LSTM based AI framework for 6G wireless channel prediction," in *Proc. IEEE VTC2022-Spring*, Helsinki, Finland, 2022, pp. 1–5.
[37] N. Ballas, L. Yao, C. Pal, and A. Courville, "Delving deeper into convolutional networks for learning video representations," *arXiv:1511.06432*, Nov. 2015. [Online]. Available: http://arxiv.org/abs/1511.06432
[38] C. Huang et al., "Machine-learning-based data processing techniques for vehicle-to-vehicle channel modeling," *IEEE Commun. Mag.*, vol. 57, no. 11, pp. 109–115, Nov. 2019.
[39] C. Huang et al., "Machine learning-enabled LOS/NLOS identification for MIMO systems in dynamic environments," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 3643–3657, June 2020.
[40] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.
[41] A. Richter, "Estimation of radio channel parameters: Models and algorithms," Ph.D. dissertation, Techn. Univ. Ilmenau, Ilmenau, Germany, May 2005. [Online]. Available: http://www.db-thueringen.de
[42] M. S. Babtlett, "Smoothing periodograms from time-series with continuous spectra," *Nature*, vol. 161, no. 4096, pp. 686–687, May 1948.
[43] N. Czink, P. Cera, J. Salo, E. Bonek, J. P. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric MIMO channel data including path powers," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2006, pp. 1–5.
[44] R. He et al., "A kernel-power-density-based algorithm for channel multipath components clustering," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7138–7151, Nov. 2017.
[45] C. Huang, A. F. Molisch, Y.-A. Geng, R. He, B. Ai, and Z. Zhong, "Trajectory-joint clustering algorithm for time-varying channel modeling," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1041–1045, Jan. 2020.
[46] C. Huang et al., "A power-angle-spectrum based clustering and tracking algorithm for time-varying radio channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 291–305, Jan. 2019.
[47] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. ICML'08*. ACM Press, Helsinki, Finland, July, 2008.
[48] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020.
[49] J. Chung, C. Gulcehre, K. Cho et al., "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv:1412.3555*, Dec. 2014.
[50] K. Cho, et al. "On the properties of neural machine translation: Encoder-decoder approaches." *arXiv:1409.1259*, Oct. 2014.
[51] M. Ravanelli, P. Brakel, M. Omologo and Y. Bengio, "Light gated recurrent units for speech recognition," *IEEE Trans. Emerging Top. Comput. Intell.*, vol. 2, no. 2, pp. 92–102, April 2018.
[52] X. Shi, Z. Chen, H. Wang, et al., "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Adv. NIPS*, pp. 802–810, Dec. 2015.
[53] X. Chen, Z. Zhou, L. Zhang, C.-X. Wang, J. Huang, "Comparison and modeling of multi-frequency wideband channels at sub-6 GHz bands," in *Proc. ICCC' 2021*, Xiamen, China, July 2021.
[54] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Berlin, Germany: Springer, 2006.
[55] J. Wang, C.-X. Wang, J. Huang, H. Wang, X. Gao, X. You, and Y. Hao, "A novel 3D non-stationary GBSM for 6G THz ultra massive MIMO wireless systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12312–12324, Dec. 2021.

Chen Huang (S'17-M'21) received the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2021. From 2018 to 2020, he has been a Visiting Scholar with the University of Southern California (USC), Los Angeles, CA, USA and with the Universite Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium. From April 2021 to April 2023, he was a Postdoctoral Research Associate in the Pervasive Communication Research Center, Purple Mountain Laboratories (PML), Nanjing, 211111, China, and also in the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University (SEU), Nanjing, 210096, China. Since April 2023, he has been an associate professor in the Pervasive Communication Research Center, PML, and an extramural supervisor in the National Mobile Communications Research Laboratory, School of Information Science and Engineering, SEU, China.

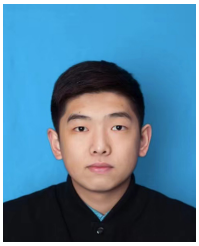
Dr. Huang was selected in Young Elite Scientists Sponsorship Program by China Association for Science and Technology and Outstanding Postdoctoral Fellow Program in Jiangsu, received 3 times the Best Paper Award from IEEE ICCT2023, WCSP 2018, IEEE/CIC ICCC 2018, and serves as the Technical Program Committee (TPC) member for several conferences, including GlobeCom, ICC, VTC-fall, VTC-spring, etc. His research interests include 6G channel measurements, characterization, and modeling, machine learning-based channel prediction, and localization. He has authored/co-authored 1 book chapters, more than 60 journal and conference papers, as well as 17 patents.

Cheng-Xiang Wang (Fellow, IEEE) received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2004.

He was a Research Assistant with the Hamburg University of Technology, Hamburg, Germany, from 2000 to 2001, a Visiting Researcher with Siemens AG Mobile Phones, Munich, Germany, in 2004, and a Research Fellow with the University of Agder, Grimstad, Norway, from 2001 to 2005. He was with Heriot-Watt University, Edinburgh, U.K., from 2005 to 2018, where he was promoted to a professor in 2011. He has been with Southeast University, Nanjing, China, as a professor since 2018, and he is now the Executive Dean of the School of Information Science and Engineering. He is also a professor with Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing, China. He has authored 4 books, 3 book chapters, and over 540 papers in refereed journals and conference proceedings, including 27 highly cited papers. He has also delivered 27 invited keynote speeches/talks and 18 tutorials in international conferences. His current research interests include wireless channel measurements and modeling, 6G wireless communication networks, and electromagnetic information theory.

Dr. Wang is a Member of the Academia Europaea (The Academy of Europe), a Member of the European Academy of Sciences and Arts (EASA), a Fellow of the Royal Society of Edinburgh (FRSE), IEEE, IET and China Institute of Communications (CIC), an IEEE Communications Society Distinguished Lecturer in 2019 and 2020, a Highly-Cited Researcher recognized by Clarivate Analytics in 2017-2020. He is currently an Executive Editorial Committee Member of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has served as a TPC Member, a TPC Chair, and a General Chair for more than 30 international conferences. He received 16 Best Paper Awards from IEEE GLOBECOM 2010, IEEE ICCT 2011, ITST 2012, IEEE VTC 2013 Spring, IWCMC 2015, IWCMC 2016, IEEE/CIC ICCC 2016, WPMC 2016, WOCC 2019, IWCMC 2020, WCSP 2020, CSPS2021, WCSP 2021, IEEE/CIC ICCC 2022, and IEEE ICCT 2023.

Zheao Li received the B.E. and M.Sc. degree in information engineering from Southeast University, China, in 2020 and 2023, respectively. He is currently working toward the Ph.D. degree with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include predictive channel modeling, digital twin, and machine learning.



Zhongyu Qian received the B.E. degree in communication engineering from Dalian University, China, in 2019, and the M.Sc. degree from Southeast University-Monash University Joint Graduate School, Southeast University, in 2024. He is currently pursuing the Ph.D. degree with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, China. His research interests include predictive channel modeling, digital twin, and machine learning.

Junling Li (Member IEEE) received the B.S. degree from Tianjin University, Tianjin, China, and the M.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2013 and 2016, respectively. In 2020, she received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. She was a Joint Postdoctoral Research Fellow at Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS), University of Waterloo, and the Chinese University of Hong Kong, Shenzhen from 2020 to 2022. She is currently an Associate Professor in the National Mobile Communications Research Laboratory at Southeast University, Nanjing, China. Dr. Li received the Best Paper Award at the IEEE ICCC 2019 and ICCT 2023. Her research interests include digital twin, channel modeling, game theory, and machine learning-based channel prediction.

Yang Miao received the M.Sc. and Ph.D. degrees from the Radio Propagation Laboratory, Mobile Communications Research Group, Tokyo Institute of Technology, Japan, in 2012 and 2015, respectively. From 2010 to 2015, she was a Research Assistant with the Takada Laboratory, Tokyo Institute of Technology. From 2015 to 2018, she was a Post-Doctoral Researcher with the Institute of Information and Communication Technologies, Electronics, and Applied Mathematics, Universite Catholique de Louvain, Louvain-la-Neuve, Belgium, and IMEC, Wireless, Acoustics, Environment, and the Expert Systems Laboratory, Ghent University, Ghent, Belgium. From 2017 to 2018, she was a part-time Senior Antenna Engineer with Jaguar Radio Wave Corporation, Shenzhen, China. From 2018 to 2019, she was a Research Assistant Professor with the Southern University of Science and Technology, Shenzhen, China. Since Aug. 2019, she has been an Assistant Professor with the Radio Systems Group, University of Twente, 4TU, The Netherlands. From Nov. 2021, she is also affiliated in part-time with KULeuven as a Marie Curie Individual Fellow. Her current research focuses on Joint Communication and Sensing, incorporating mobility and human factors. She currently serves as a co-chair in COST action subgroup ISAC.