Channel Scenario Extensions, Identifications, and Adaptive Modeling for 6G Wireless Communications

Wenqi Zhou, Cheng-Xiang Wang, *Fellow, IEEE*, Chen Huang, *Member, IEEE*, Zheao Li, Zhongyu Qian, Zhen Lv, *Member, IEEE*, and Yunfei Chen, *Senior Member, IEEE*

***Abstract*—To provide customized high-quality services for all users in the sixth-generation (6G) wireless communication systems, it is fundamental to study all 6G channel scenarios and establish accurate channel models for these scenarios correspondingly. However, the absence of comprehensive 6G scenario categorization and the difficulties of modeling the channels for all scenarios bring huge challenges. In this article, we aim to give a thorough overview of channel scenarios, identification algorithms, and intelligent channel modeling theories. First, different standardized scenario categorization principles are reviewed. A unified and exclusive scenario categorization method is elaborated with detailed 6G scenario definitions. Second, scenario features, feature selection principles, ML-based identification algorithms, as well as data preprocessing methods are surveyed for the benefit of accurate scenario identification. Third, the intelligent scenario adaptive channel modeling theory based on 6GPCM is specified. Statistical properties for industrial IoT and HST scenarios are simulated and compared with those from measurements. Finally, future research directions and challenges are addressed.**

***Index Terms*—Adaptive channel modeling, industrial Internet of Things (IoT), scenario categorization, scenario identification, sixth-generation (6G).**

Manuscript received 23 September 2022; revised 1 August 2023; accepted 29 August 2023. Date of publication 14 September 2023; date of current version 21 February 2024. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61960206006 and Grant 62301364; in part by the Fundamental Research Funds for the Central Universities under Grant 2242022k60006; in part by the Key Technologies Research and Development Program of Jiangsu (Prospective and Key Technologies for Industry) under Grant BE2022067 and Grant BE2022067-1; in part by EU H2020 RISE TESTBED2 Project under Grant 872172; and in part by the Young Elite Scientists Sponsorship Program by CAST under Grant 2022QNRC001. *(Corresponding authors: Cheng-Xiang Wang; Chen Huang.)*

Wenqi Zhou and Cheng-Xiang Wang are with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, China, and also with the Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing 211111, China (e-mail: wqzhou@seu.edu.cn; chxwangseu.edu.cn).

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

Zheao Li and Zhongyu Qian are with National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: li_zheao@seu.edu.cn; zyqian@seu.edu.cn).

Zhen Lv is with the Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing 211111, China (e-mail: lvzhen@pmlabs.com.cn).

Yunfei Chen is with the School of Engineering, University of Warwick, CV4 7AL Coventry, U.K. (e-mail: yunfei.chen@warwick.ac.uk).

Digital Object Identifier 10.1109/JIOT.2023.3315296

I. Introduction

WIRELESS communication systems have now evolved into the sixth-generation (6G) [1], [2], [3], thereby bringing unprecedented changes in the way of life and in the present social and economic landscape. Motivated by the ever-increasing demands for high-data rate, low latency, and high mobility, the 6G communication system will be an intelligent Internet of Everything (IoE) that aims to include all frequency bands and communication scenarios [4], [5], [6]. As stated in [7], the 6G system envisions global coverage, all spectral [8], and full applications. Considering the overwhelming geographical diversities and channel propagation characteristics in 6G communication scenarios, a comprehensive understanding of the wireless propagation environments is fundamental to extract and refine environmental features, then to construct accurate channel models and establish reliable communication network. Therefore, it is crucial to give a thorough clear categorization of existing 6G communication scenarios. In addition, assisted by the introduction of high-efficiency intelligent technologies, the intelligent scenario identification should be investigated to ensure a quick access to the matching of channel statistical model parameters and acquire the specialized channel model. Then, the wireless communication systems can be designed and optimized based on the acquired channel model [9]. In particular, with regard to the global coverage feature, 6G is expected to expand terrestrial communications to space-air-ground-sea integrated communications [10]. However, the expansion of 6G wireless communication scenarios in terms of breadth and depth poses higher requirements on system design and technical specifications.

To provide customized high-quality services for different 6G space-air-ground-sea wireless communication scenarios, new key technologies and network architectures should be employed in the future wireless communication systems. Especially, the emergence of Internet of Things (IoT) [11] scenarios and intelligent applications will bring significant uncertainty. It is expected that adaptive technologies and automatically adapted intelligent networks can be established to provide dynamic, flexible, and dedicated services to all accessed equipment, such as adaptive modulation and coding [12]. Correspondingly, wireless communication system design, deployment, and management need to be more intelligent and universal [13], [14], [15].

One of the most important parts of 6G communication system is the wireless channels. Between the transmitter (Tx) and the receiver (Rx), many uncertainties will exist due to different system setups and propagation environments. As a basis for system design, evaluation, and optimization, wireless channel modeling plays an important role [6]. Traditionally, a channel model can be established through channel measurement, multipath component (MPC) parameter estimation [16], and statistical analysis. However, the complexity of this process will be too high for 6G. First, 6G communication scenarios are manifold and more complex, especially with the usage of new technologies, such as millimeter wave (mmWave), terahertz (THz), (ultra-)massive multiple-input multiple-output (MIMO) [17], reconfigurable intelligent surface (RIS) [18], etc. Hence, it is impossible to cover all communication scenarios at all frequency bands with various system configurations. Even for a given scenario, with the expanded bandwidth, increased antenna element number, and rising speed of the terminal, the channel measurement data for storage and processing are huge. Second, high-resolution MPC parameter estimation is usually time-consuming. This is more challenging when dealing with large volume of measurement data. Especially in high-mobility communication scenarios, e.g., high-speed train (HST) [19] and vehicle-to-vehicle (V2V) [20], [21], the channel can change dramatically with very fast traveling speed. In such cases, denser channel measurements should be conducted and faster data processing should be performed to provide real-time information. Third, new channel characteristics emerge in various 6G scenarios. For example, HST channel scenarios show distinct temporal nonstationarity. How to efficiently extract channel statistical properties for specific 6G scenarios is another task that needs to be tackled. Finally, channel models that incorporating channel properties of various 6G scenarios need to be formulated. However, this may bring overwhelming complexity to channel models. Therefore, the tradeoff among accuracy, complexity, and pervasiveness is usually hard to obtain. For example, deterministic channel model, such as ray tracing (RT), very much rely on the reconstruction of the real environment, which can be highly complex [22].

One way to solve the above mentioned problems is to let the 6G system automatically identify the relevant communication scenario and adaptively matching channel model parameters to that scenario. This vision of 6G adaptive channel modeling procedure is explained in Fig. 1. Specifically, based on a comprehensive and exclusive categorization of 6G scenarios, a certain communication scenario is first identified and then model parameters for an adaptive channel model are matched. In terms of scenario extension and categorization, standardization groups, including the 3GPP TR 38.901 [23], WINNER I/II/+ [24], [25], [26], METIS [27], MiWEBA [28], QuaDRiGa [29], COST 2100 [30], [31], [32], [33], and IMT2020 [34], have proposed different categorizations. However, it will be shown later that their involved scenarios are very limited and their categorization criteria are ambiguous. They are not adequate for the engineers to select suitable standard channel models for various 6G scenarios. In the scenario identification step, as a perfect candidate to identify various 6G scenarios, machine learning (ML) has attracted much attention and can be expected to automatically identify wireless communication scenarios, process real-time measurement data/feedback signals, track cluster variances, and extract model parameters [35]. It has been widely applied in solving many wireless communication problems [36]. These include scenario identification, channel modeling, channel estimation, positioning/localization, and network management [37]. In addition to processing large amount of data sets with reduced human effort, ML can provide reliable real-time decisions with flexible adaptability. Finally, the tradeoff among accuracy, complexity, and pervasiveness/universarity of the deterministic RT, geometry-based stochastic channel models (GBSMs) [38], beam domain channel models (BDCMs) [39], [40], [41], and correlation-based stochastic channel models (CBSMs) [42], [43], [44], [45] should be considered. It is worth studying various channel modeling methodologies and choosing suitable candidates for the pervasive channel model.

There has been no work in the literature that provides a complete study on scenario extension, identification, and adaptive channel modeling. Hence, this work investigates a series of communication scenarios definitions and classifications delivered by standardized channel models, such as the 3GPP TR 38.901 [23] and WINNER I/II [24], [25], and proposes a novel scenario classification framework for 6G communications. Then, we overview the commonly used feature extraction approaches of existing ML-based communication scenario identification methods. Furthermore, the conventional nonpredictive model and ML-based predictive model are summarized. Finally, the framework of 6G adaptive channel modeling is introduced. The future research directions and challenges are also discussed. The main contributions and novelties of this work are as follows.

1) Scenario categorization methods of various standardization groups are compared and the 6G global coverage scenarios are introduced. A comprehensive and exclusive categorization method is proposed with detailed scenario definitions.
2) Features used for scenario identification and feature selection principles are presented. Identification methods for the Line-of-Sight (LoS)/non-LoS (NLoS) and multiple scenarios are analyzed. Besides, proper ML algorithms for various 6G scenario identification are analyzed.
3) The channel model parameter matching for different 6G scenarios is first illustrated using a newly proposed 6G pervasive channel model (6GPCM). The model accuracy in particular scenarios are validated by comparing the statistical properties of channel model simulations with those of channel measurements.

The remainder of this article is organized as follows. In Section II, standard scenario categorization methods are introduced and a novel scenario categorization method is proposed. In Section III, feature selection and scenario identification methods are presented. In Section IV, various channel models and the adaptive channel modeling methodology are introduced. In Section V, the model parameters and simulated

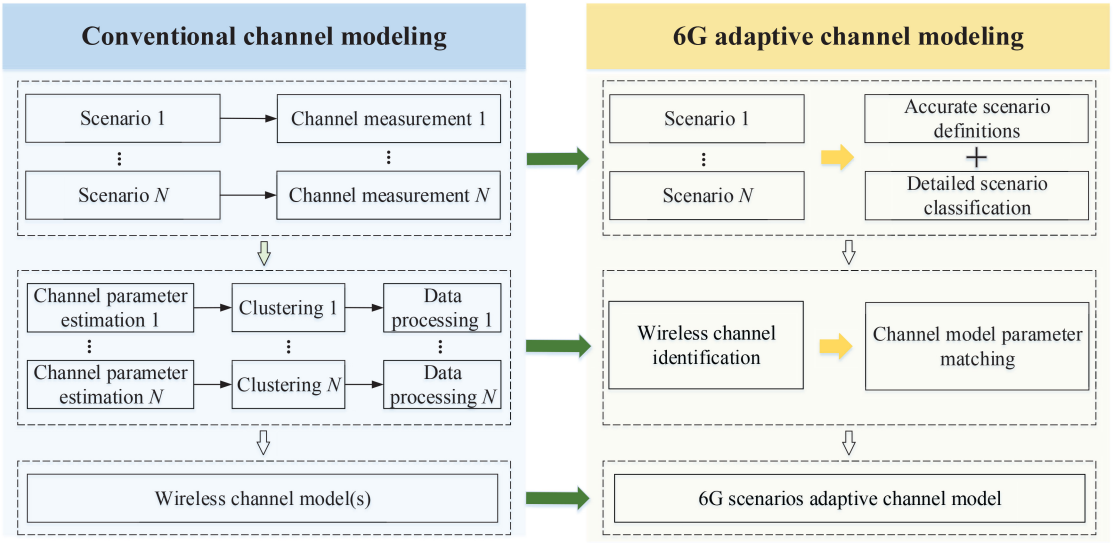


Fig. 1. Simplified wireless channel modeling procedure.

statistical properties for Industrial IoT (IIoT) and HST scenarios are illustrated. Future directions and challenges are summarized in Section VI. Finally, conclusions are drawn in Section VII.

II. Wireless Communication Scenario Categorization

Previously, in the third-generation (3G) to the fifth-generation (5G), there were several research works that focused on different sets of wireless communication scenarios and their classification based on different criteria. However, in the 6G global coverage system, scenarios exhibit increased diversity and complexity compared to those in 3G–5G. In this section, we elaborate on typical 3G–5G standard categorization methods and propose an extended categorization framework that incorporates most 6G global coverage scenarios.

A. Standard Scenario Categorizations and Definitions

Before 6G, wireless communications mainly focused on terrestrial scenarios. There were some standardization groups that had a light touch on different scenarios of interest. The involved scenarios and categorization criteria used by these standardization groups are listed in Table I. Specifically, we summarize the shortcomings of these categorization methods in a manner that is later suitable for 6G scenario categorization.

1) The 3GPP TR 38.901 standard channel model intended to accurately characterize wireless channels from 0.5 to 100 GHz. It is applicable to both link-level and system-level simulations. Considering multiple physical scenario types, the supported scenarios include indoor office, shopping mall, indoor factory, stadium, and gym. Both outdoor-to-outdoor (O2O) and outdoor-to-indoor (O2I) link topologies were included for urban microcell (UMi) and urban macrocell (UMa) cell types. There were also backhaul, device-to-device (D2D), and V2V access link types. It can be seen that these categorization criteria are quite ambiguous and the involved scenarios are limited.
2) In the WINNER I document, four physical scenario types, including in building, hotspot, metropol, and rural, were defined. Further classified by the cell type and link topology, there were small office/residence and indoor-to-outdoor (I2O) for in building scenarios; typical UMa, bad UMi, indoor, O2I, and four stationary feeders for hotspot scenarios; suburban macrocell (SMa), UMa, bad urban, O2I, and LoS feeder for metropolitan scenarios; and rural macrocell (RMa) and LoS moving networks for rural scenarios. Note that only indoor, UMa, UMi, stationary feeder, SMa, and RMa scenarios were covered in the WINNER I channel model. Here, it can be seen that the categorization criteria cannot provide exclusive categorization of different scenarios. The categorization of indoor and outdoor scenarios is not complete.
3) Evolved from the WINNER I document, the WINNER II document was delivered for both link-level and system-level simulations. A total of 18 scenarios were listed in this document. The newly covered scenarios in the WINNER II channel model were I2O, bad UMi, O2I, and rural moving networks. It supported multiantenna technologies, multiuser (MU), multicell, and polarization networks. It can be used for any wireless systems operating at 2–6 GHz. Similar to the scenario categorization in WINNER I, the categorization criteria are not exclusive or complete.
4) In WINNER +, in addition to channel measurements conducted for the validation of the WINNER II channel model, mainly four measurement campaigns were carried out for channel modeling in ultra-high frequency (UHF) and distributed antenna systems (DASs), e.g., O2I at 522 MHz, UMa at 770–790 MHz, and campus at 3.55 GHz. The main purpose of this document was to support the update of the WINNER II channel model. There are also some differences from the

TABLE I
SCENARIOS OF INTEREST IN EXISTING STANDARD DOCUMENTS

Standardized groups	Number of scenarios	Categorization criteria	Scenarios of interest	Disadvantages
3GPP TR 38.901	17	Physical scenario type, cell type, link type, link topology	**Indoor**: office, shopping mall; **Indoor industrial** scenarios; **UMi**: street canyon and open area with O2O and O2I; **UMa**: O2O and O2I; **Backhaul**: urban, street canyon; **D2D/V2V**: open area, street canyon, indoor; **Others**: stadium, gym.	• Categorization criterion is quite ambiguous; • The involved scenarios are limited.
WINNER I	17	Physical scenario type, cell type, link topology	**In building**: small office/residential, I2O; **Hotspot**: typical UMi, bad UMi, indoor, O2I, stationary feeder (4); **Metropol**: suburban, UMa, bad urban, O2I, LoS feeder; **Rural**: RMa, LoS moving networks.	• Categorization of scenarios is not complete; • The criterion can not provide exclusive categorization of different scenarios.
WINNER II	18	Physical scenario type, cell type, link topology	**In building**: office/residential, I2O; **Hotspot**: typical UMi, bad UMi, large indoor hall, O2I, stationary feeder (4), feeder link BS; **Metropol**: suburban, typical UMa, bad UMa, O2I macro-cell; **Rural**: RMa, moving networks (2).	• Detailed categorization of indoor and outdoor scenarios is not complete; • The criterion can not provide exclusive categorization of different scenarios.
WINNER +	5	UHF, DAS	**522 MHz**: urban O2I, urban pedestrian, urban vehicular; **800 MHz**: urban, O2I; **2.53 GHz**: UMa; **3.55 GHz**: DAS.	• Mainly focuses on the validation and support of WINNER II channel model.
METIS	11	Physical scenario type, link type, link topology	**Basic propagation environment**: dense urban, urban, rural, office, shopping mall, highway, open air festival, stadium; **Link type**: BS-EU, BS-BS, D2D; **Link topology**: O2O, O2I, I2I.	• The categorization of basic propagation environment is not clear.
MiWEBA	8	Link type	**Access**: open area, street canyon, hotel lobby; **Backhaul/fronthaul**: above rooftop, street canyon; **D2D**: open area, street canyon, hotel lobby.	• The involved scenarios are quite limited.
mmMAGIC	10	Physical scenario type, cell type, link topology	**UMi**: street canyon, open square; **Indoor**: office, shopping mall, airport; **O2I**: open area, street canyon, hotel lobby; **Stadium**; **Metro station**.	• The involved scenarios are quite limited.
COST2100	12	Link type, channel type, link topology	**Distributed systems**: multi-node (indoor MU, outdoor MU, O2I MU, O2I and I2I distributed nodes, outdoor relay, outdoor peer-to-peer), multi-BS; **Polarimetric channel**; **Vehicular channel**: V2V, vehicular-to-infrastructure; **UWB**.	• The involved wireless channels are quite limited.

scenario definition in WINNER II. Note that this model is generalized from 2-D to 3-D by considering elevation angles, and can be used for over-the-air (OTA) testing.

5) The METIS document defined a number of basic propagation environments, link types, and link topologies following the propagation mechanisms. The mentioned propagation environments were dense urban, urban, rural, indoor (office and shopping mall), highway, open air festival, and stadium. The link types of interest included base station (BS)-user equipment (EU), BS-BS, and D2D, and the link topologies included O2O, O2I, and indoor-to-indoor (I2I). The total number of involved scenarios is 11. However, the criterion to classify the propagation environment is not clear and the involved scenarios are quite limited.
6) In MiWEBA, according to the link type, there were three main categories of scenarios, including access in open area, street canyon, and hotel lobby; fronthaul/backhaul in the above rooftop and street canyon; and D2D also in open area, street canyon, and hotel lobby. The number of involved scenarios is 8. The involved scenarios are quite limited and there is a lack of transition of scenarios, such as I2O and O2I.
7) In mmMAGIC, following the signal transmission mechanisms, scenarios were classified as direct transmission and indirection transmission (scattering, ground reflection, and blockage). Then, 10 scenarios were mentioned by further dividing each scenario into indoor and outdoor cases. Nevertheless, the mentioned indoor, outdoor, and transition between indoor and outdoor scenarios are quite limited. UMa was not taken into consideration since a frequency spectrum above 6 GHz is expected to be initially used for small cell BS.
8) In COST2100, channel measurement activities were carried out for different scenarios, including distributed cooperative systems, polarimetric channels,

vehicular channels under high mobility, ultrawideband (UWB) channels, as well as mmWave and sub-mmWave systems. Here, the distributed cooperative systems included multiple nodes with a single BS and multiple user channels. There were indoor MU MIMO, outdoor MU, O2I MU MIMO, O2I, and I2I distributed nodes, outdoor relay channels, and outdoor peer-to-peer channels. The communication scenarios considered in COST2100 are also limited.

In general, existing standardized wireless communication scenario categorizations are either not complete or not clear enough. They are unable to provide useful reference for 6G system engineers to choose the dedicated channel models for various 6G scenarios.

B. 6G-Oriented Scenario Extension

1) 6G-Oriented Scenario Categorization Criterion: The proposed 6G-oriented global coverage scenario categorization framework is illustrated in Fig. 2. The criterion behind the proposed categorization scheme is from far to near and from global to specific. We aim to establish a comprehensive, subtle, and exclusive scenario categorization framework. By mapping geometrical and other physical parameters for each dedicated scenario, we expect this framework to serve as an important reference for further scenario identification.

Based on the three basic principles of exclusivity, integrity, and standardization, 6G communication scenarios are classified into space-air-ground-sea, we first classify wireless communication scenarios into 5 main categories, i.e., space, aerial, terrestrial, maritime, and transformation scenarios. In the following, we will briefly discuss each category.

2) Space Communication Scenario Categorization: The space communication scenario becomes important in providing seamless coverage for wireless communications [46]. According to the orbital altitude, space communications can be further divided into high, middle, and low Earth orbit (LEO) satellites, with the heights above 35 786 km, between 2000–35 786 km, and below 2000 km, respectively. They can be used for indoor and outdoor communications. A typical communication scenario in this category is the international space station, which operates at an altitude of 379 km. Due to the long propagation link distance of satellite scenario and the influence of meteorological conditions, such as ionosphere and cloud, rain and fog [47], it has broad coverage capability and super bandwidth connection. So the exclusivity of this kind of scenario is ensured by the height of the scenario and the characteristics of the channels. The hierarchical order and functional differentiation of low orbit, medium orbit, high orbit, and space station ensure the integrity of the classification of communication scenario.

3) Aerial Communication Scenario Categorization: Aerial communications using aircraft, unmanned aerial vehicle (UAV), balloon, missile, etc., can provide long-haul signal transmission without extremely high costs. They can not only be used for wireless communications between aerial equipments but also be used to exchange information with terrestrial or space receivers/centers. In particular, UAV plays a more important role in providing light-weight, flexible, and full-view real-time communication and tracking. In this category, the typical communication scenarios are drones and aircraft communications. Considering that the aviation channel link is short, it is mainly affected by atmospheric attenuation and other factors. The channel differences caused by different flight altitudes of aircraft are large. The high maneuverability and arbitrary trajectory of UAVs may cause channel time-domain nonstationary characteristics. The applications in mmWave and terahertz frequency bands bring spatial-temporal sparsity. The channel model has high-delay resolution under high-bandwidth conditions. Relying on the flight altitude and speed, the classification of the aerial communication scenario is exclusive. Based on the near-Earth distribution characteristics of different high-layer scatterers in the air-to-ground channel of UAV, it can ensure the integrity of the scenario classification. According to the correlation of large/small scale parameters between different layer modeling multi-links, it further guarantees the standardization of scenario classification.

4) Terrestrial Communication Scenario Categorization: Terrestrial communications still account for the majority of 6G wireless communications. It can be divided into indoor and outdoor scenarios.

For indoor scenarios, there are residence, office, education area, industry, commerce, and agriculture. The residence scenario can be further divided into multi/highrise residential, single-family house/villa, country courtyard, cave dwelling, Hakka Earth building, water building, yurt, recreational vehicle, sample room, isolation room, and resort. The typical office scenarios include open centralized office area, isolated centralized office area, small independent office, meeting room, logistics service room, emergency command center, secret room, and archives. The education area can also be further divided into students apartment, teaching building, library, canteen, stepped auditorium, stadium, laboratory, computer room, and convenience service store. The industry scenarios include warehouse, medium/large manufacturing plant, assembly plant, medium/large production line factory, power station, water treatment plant, and steel/petroleum smelter. The typical commerce scenarios are shopping mall, hotel, restaurant, cinema, culture and sports museum, cabaret, Internet cafe, bookstore, swimming pool, chess room, indoor arena, indoor attractions, airport, high-speed rail station, subway station, hospital, and bank. There are also agricultural scenarios, including field, farm, and indoor farmers market.

The outdoor scenarios are more complex. We divide them into typical urban, suburban, terrestrial transportation, in vehicle, remote area, and underground. First, in typical urban scenarios, there are open area, hotspot area, street canyon, and roof. For the open area scenarios, we further divide them into square, park, velodrome, scenic spot, ski resort, outdoor courts, airfield, golf course, and beach. Hotspot area scenarios include festival square, station square, outdoor stadium, gas/charging, outdoor parking, highway service area, and fairgrounds. Street canyons include main streets in suburb/country, rural-urban continuum, city, and outskirts, as well as city branch.

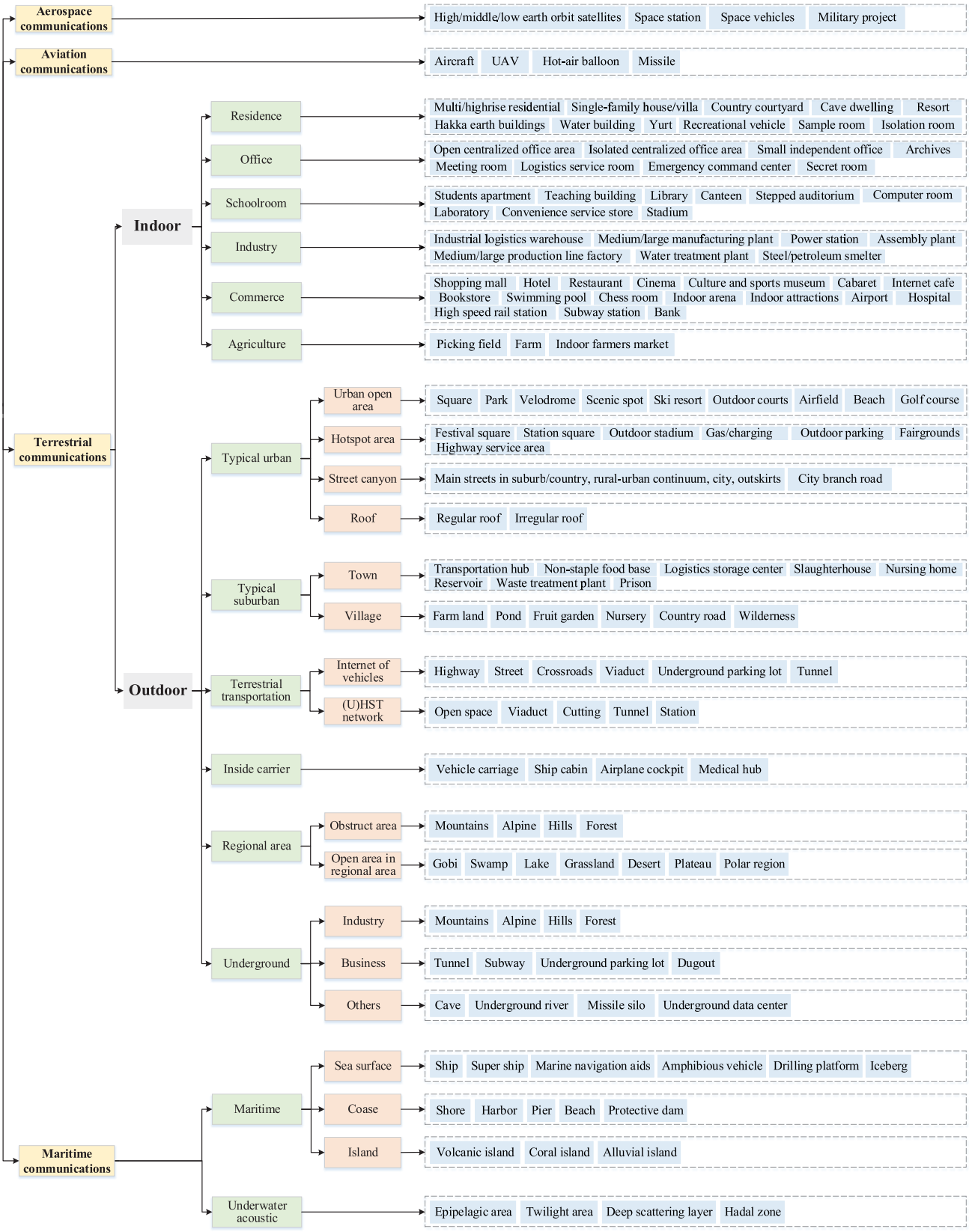


Fig. 2. 6G-oriented scenario categorization framework.

Suburbs can be further classified into town and village. In towns, there are transportation hub, nonstaple food base, logistics storage center, reservoir, nursing home, waste treatment plant, slaughterhouse, and prison. In villages, there are farm land, pond, fruit garden, nursery, country road, and wilderness. Terrestrial transportation mainly includes highway, street,

crossroads, viaduct, underground parking lot, and tunnel, as well as open space, viaduct, cutting, tunnel, and station. In vehicle scenarios can be classified as carriage, cabin, cockpit, and medical shelters. Obstructed areas and open areas are the two main forms of remote areas. The underground scenarios include industry, transportation, and others. There are mine and well drilling in industry, tunnel, subway, underground parking lot, and dugout in business. Additionally, there are other scenarios, including cave, underground river, missile silo, and underground data center.

Specific examples include residential, office, industrial, and commercial, etc. Among them, the typical multihigh-rise residential is designed according to the urban family model with rich functions and diversified structures, and shows regularity according to the regional layout. The structure of a single-family house or villa is significantly different, and the floor length/width/height is significantly different from that of an ordinary house. Offices in the urban core have a single function and structure, but a relatively random layout. Considering the production process, indoor lighting, and complex architecture used to facilitate the passage of a variety of lifting and transportation equipment, industrial buildings generally use multispan structures with large length and width. Commercial buildings can be divided into shopping malls, supermarkets, hotels, office buildings, and exhibition halls according to the function. The basestation is generally located outdoor, and the communication scenario needs to consider the moving process from outdoor to indoor scenario. The indoor scenario environment is usually closed and the signal can be reflected repeatedly. The outdoor environment is relatively open, the coverage area is wide, the propagation loss is large, and various scatters in the environment can bring obvious multipath effects. The exclusivity of the terrestrial classification is ensured by the average height of the building, the average floor spacing, the height fluctuation of the building, the regional LOS proportion, and the indoor proportion. The structure is divided into indoor and outdoor, and the integrity of the scenario classification is achieved by further refinement combined with functions. Based on the logical framework of 3GPP TR 38.901, it is extended to ensure the standardization of scenario classification.

In Table II, we have listed some definitions of typical terrestrial transportation scenarios as examples. Through parameterized scenarios, we will be able to find the most distinct features that can be used for further scenario identification. For example, the most distinct differences between Internet of Vehicles and HST network are the height and moving speed. There are also obvious variances of moving speed for detailed scenarios in Internet of Vehicles. The width and height can be associated to categorize detailed scenarios in HST network.

5) Maritime Communication Scenario Categorization: Maritime communication is a key part of accomplishing full coverage communication. It not only refers to the sea surface, coast, and island but also includes underwater acoustic communication scenarios, such as epipelagic area, twilight area, deep scattering layer, and hadal zone. The sea surface communication scenarios include ship, super ship, marine navigation aids, amphibious vehicle, drilling platform, and iceberg; the coast scenarios include shore, harbor, pier, beach, and protective dam; and the island scenarios include volcanic, coral, alluvial, mainland, and ocean islands. Above the sea surface, there are ports and shoreline on the sea. The layout of the port is similar to the residence and industry. The shoreline structure is usually narrow and long. There are certain differences between two scenarios in terms of structure. The underwater acoustic channels can be divided into shallow sea area ($\leq$100 m) communication scenario and deep sea area ($\geq$100 m) communication scenario according to the depth of the water area [48]. Due to changes in the physical properties of sea water, such as depth, salinity, and temperature, the propagation properties of sound waves in ocean communication are different in different depths. In contrast, the quality of underwater acoustic channels in deep waters is better. In shallow waters, sound waves can be reflected and scattered by the seabed and sea surface during propagation, resulting in a strong multipath effect at the receiver. In addition, compared with the ultrahigh propagation speed of electromagnetic waves in ground communication, the transmission speed of underwater acoustic waves is 1500 m/s, resulting in a very large Doppler effect at the receiver. The classification of maritime scenarios from different physical depth layers and transmission speeds can ensure the exclusivity of this type of scenario. The classification from above and below the water surface and the water bank can guarantee the integrity of scenario classification.

6) Transitions Between Communication Scenarios: The transitions between communication scenarios are also included in the proposed framework. For example, I2I, I2O, O2O, O2I, and other interscenario transformation. Due to space limitations in this article, we will not elaborate on them.

To sum up, existing standardized scenario classification methods are incomplete or fuzzy, the classification logic is not clear, and the corresponding granularity is rough. Moreover, it lacks multiscenario classification for 6G full coverage scenarios and scenario transformation. The proposed classification framework gives the physical definition, new planning, and scenario expansion for 6G full coverage scenarios according to environmental difference and channel characteristics. Under the guidance of comprehensive coverage architecture, the classification is carried out abiding by three criteria of exclusivity, integrity, and standardization. It provides an important basis for the research of 6G wireless communication scenario classification and the development of channel modeling.

III. Identification of Wireless Communication Scenarios

The identification of wireless communication scenarios is helpful for engineers to efficiently customize wireless communication systems and to provide high-quality services in some special communication scenarios, such as V2V and HST. In this section, scenario related features and scenario identification algorithms, especially ML-based algorithms, will be

TABLE II
DEFINITIONS OF TYPICAL TERRESTRIAL TRANSPORTATION SCENARIOS

Scenario		LoS/NLoS	Length (m)	Width (m)	Height (m)	Mobility (km/h)	Distance range (m)
Internet of vehicles	Highway	LoS/NLoS	$>$ 5,000	35∼50	Inf	90∼110	1,000∼3,000
	Street	LoS/NLoS	$>$ 200	12∼40	Inf	$\leq$ 60	50∼300
	Crossroads	LoS/NLoS	$<$ 100	12∼40	Inf	10	300
	Viaduct	LoS	$>$ 5,000	6∼26	Inf	35∼110	100∼3,000
	Underground parking lot	NLoS	50∼1,000	40∼500	3.5∼5	$\leq$ 20	15∼300
	Tunnel	NLoS	1,000∼15,000	18∼30	5∼10	100∼600	15∼2,000
HST network	Open space	LoS	$>$ 1,000	$>$ 1,000	Inf	$\leq$300	200∼3,000
	Viaduct	LoS	$>$ 5,000	6∼26	10∼30	$\leq$ 200	200∼3,000
	Cutting	LoS/NLoS	$>$ 300	48∼63 14∼19	3∼10	$\leq$200	200∼3,000
	Tunnel	NLoS	1,000∼15,000	10∼20	5∼10	$\leq$ 200	100∼3,000
	Station	LoS/NLoS	500	8∼10	10∼60	$\leq$ 80	15∼200

TABLE III
SELECTED FEATURES FOR SCENARIO IDENTIFICATION

Category	Domain	Feature	Ref.
Raw data	Time domain, frequency domain	CIR, CTF	[50], [51]
Distribution shape	Time domain	PDF, peak number, skewness, kurtosis/peakedness	[52]–[56]
Threshold comparison	Power domain	PL, RSS, Rician K-factor	[15], [57]–[60]
	Time domain	Channel gain, temporal ACF, RMS DS	[50], [54], [61]–[63]
	Spatial domain	Angular PSD, spatial CCF, RMS AS	[15], [62], [63]
	Frequency domain	Doppler PSD, FCF, RMS DPS	[51], [62]–[64]

studied. Relevant features and data-processing will be first discussed, followed by conventional manual methods. Then, new ML-based automatic methods will be examined.

A. Features and Data Preprocessing

1) Scenario-Related Features: Based on communication scenario identification, we can select several relevant features, as listed in Table III.

a) *Channel impulse response (CIR)/Channel transfer function (CTF):* One feature that contains rich information is the CIR $h(t)$ in the time domain or the CTF $H(f)$ in the frequency domain. However, the large data size induced by channel measurement snapshots and taps may hinder the timely decision, which is an issue in time-varying environments.

b) *Distribution Shapes:* The probability density function (PDF), peak number of the delay power spectral density (delay PSD), skewness, and kurtosis/peakedness are characteristics of the distribution shapes to distinguish different scenarios. Here, PDF indicates the probability distribution of the wireless channels. Usually, a LoS scenario follows the Rice distribution, while a NLoS scenario follows the Rayleigh distribution. Therefore, it can be used to distinguish LoS and NLoS scenarios. The peak number of the delay PSD can be used to identify a rich scattering environment. Delay PSD peaks with different delays represent MPCs from different scatterers. Other parameters are defined as below.

Skewness indicates to what extent the probability distribution is asymmetric. It can be calculated as

$$\epsilon = \frac{\mathbb{E}[|h(t)| - u]^3}{\sigma^3} \tag{1}$$

where u and σ are the expectation and standard deviation of $h(t)$, respectively. The skewness of a LoS scenario is usually smaller than that of a NLoS scenario [55].

Kurtosis/peakedness shows the strident of a peak. It can be calculated as

$$\kappa = \frac{\mathbb{E}[|h(t)| - u]^4}{\sigma^4}. \tag{2}$$

It is often compared with the normal distribution, which corresponds to the case $\kappa = 3$. The kurtosis in the LoS case is larger than that in the NLoS case.

c) *Power Domain Threshold Comparison:* Path loss (PL) is an important indicator of large-scale fading. It is influenced by the transmission distance d as

$$L_p = \mathrm{PL}(d_0) - 10\gamma \log_{10}\left(\frac{d}{d_0}\right) \tag{3}$$

where $\mathrm{PL}(d_0)$ is the PL at the reference distance d_0 and γ is the PL exponent in a certain scenario. PL can be used to distinguish communication scenarios with significantly different traveling distances.

Received signal strength (RSS) is the total received signal power in decibel (dB), which can be calculated as

$$\mathrm{RSS} = 10\log_{10} P \tag{4}$$

where P is the total receive signal power.

Rician K-factor is the ratio between LoS power and NLoS power in Watts as

$$K = \frac{P_{\mathrm{LoS}}}{P_{\mathrm{NLoS}}} \tag{5}$$

where P_{LoS} is the LoS power and P_{NLoS} is the NLoS power. Therefore, the Rician K-factor can be used to reflect the channel condition, namely, whether it is a LoS case or a NLoS case.

d) *Time Domain Threshold Comparison:* Channel gain is an important indicator of power distribution at different time instants. A large channel gain means that there are strong MPCs that have contributed to the Rx or means a LoS case. A small channel gain may be caused by weak MPCs or purely by noise. The channel gain can be calculated as the square of CIR.

The temporal autocorrelation function (ACF) measures how fast the channel characteristics change with time. For a time variant scenario with high mobility, the ACF can decrease very fast and the stationary duration can be very small.

The root mean squared (RMS) delay spread (DS) reflects the dispersion degree of the wireless channels in the time delay domain. For scenarios with rich scatterers, the RMS DS is usually very large.

e) *Spatial Domain Threshold Comparison:* The angular PSD can also be used for scenario identification. It can be obtained by the conventional Bartlett beamformer or calculated based on the extracted MPC angular parameters.

The spatial cross-correlation function (CCF) measures how fast the channel characteristics change at different antenna elements. For a (ultra-)massive MIMO scenario with very large antenna array, the CCF can decrease rapidly with the expansion of antenna distance and the stationary interval can be observed.

RMS angular spread (AS) reflects the dispersion degree of the channels in the angular domain. Here, the interested angle can be Azimuth Angle of Arrival (AAoA), Elevation AoA (EAoA), Azimuth Angle of Departure (AAoD), and Elevation AoD (EAoD).

f) *Frequency Domain Threshold Comparison:* The Fourier transform of the temporal ACF gives the Doppler PSD to show the power distribution over different frequency points. The frequency correlation function (FCF) shows how the channel characteristics change at different frequency points. For wireless channels with large bandwidth, this phenomenon can be significant. The Doppler effect is caused by the relative motion of Tx and Rx. The RMS doppler spread (DPS) reflects the dispersion degree of the channels in the frequency domain. For the static scenario, there is no Doppler effect; thus, its RMS DPS tends to 0.

2) Feature Selection: Efficient scenario identification relies on the features used. For different scenarios and algorithms, features that play crucial roles may be different. In [15], it was reported that the selected features for scenario identification should be informative, discriminating, and independent.

Features can be calculated and used to make decisions by comparing them with typical values/thresholds, as they assume different values in different scenarios. In [64], this method was developed for key parameter threshold-based identification. However, this method does not work well for time-varying environment, where channel properties change with time.

Multifeature fusion is also efficient for scenario recognition. Rafatnia et al. [12] proposed a multifeature fusion-based scenario recognition method for HST channels. The CIRs of some scenarios, such as rural, station, suburban, and multilink scenarios, were measured. They used the Rician K-factor, RMS DS, RMS AS, and RMS DPS. The relevance of scenarios was analyzed to show that the suburban and multilink scenarios had stronger relevance.

3) Data Preprocessing: To cope with a large volume of data and to improve identification efficiency, data preprocessing can be performed to remove redundant information. Too much redundant information will be a barrier to accurate scenario identification, but too little information may also lead to incorrect scenario identification. Therefore, appropriate preprocessing is meaningful for scenario identification. It mainly includes normalization and dimension reduction.

Normalization limits the processed data within a certain range, e.g., [0,1], thus improving classification or regression performance. The most commonly used method is z-score normalization, which unifies the processed data following the standard normal distribution. Also, batch normalization makes the input vectors of each layer follow the same distribution. Therefore, the vanishing gradient can be avoided and the convergence rate can be improved.

Dimension reduction can be used to reduce computational complexity. Principal component analysis (PCA) is widely used. Singular value decomposition (SVD) can transform correlated data into linearly uncorrelated values. The kernel method is used to handle especially complex distributed data sets [15]. The radial basis function (RBF) kernel is a typical example.

A band-pass filter can also be used to denoise the data by setting the absolute cutoff frequency slightly larger than the signal frequency. It has been verified that band-pass filter can filter as much white Gaussian noise as possible.

B. Conventional Manual Identification Method

The intuitive way to identify LoS/NLoS and other scenarios is by human inspection/manual identification. Conclusions can be drawn by observing the most distinct characteristics through synthesized analysis. For example, the LoS scenario can be easily identified if there is a dominant peak in delay PSD

with the corresponding delay equals to the time duration of the signal from the Tx to Rx linearly. However, it is usually not reliable when no distinct characteristic is available.

In 6G (ultra-)massive MIMO communication scenarios, the increase in antenna number brings massive augmentation of data. In time-varying V2V and HST scenarios, wireless channels also show very different propagation characteristics in short time periods. The complexity of data in spatial, temporal, and spectral domains is phenomenal, making manual identification impractical.

C. ML-Based Intelligent Scenario Identification

1) ML Algorithms: ML algorithms can be categorized as supervised learning, unsupervised learning, and reinforcement learning (RL) [65], [66]. Their applications in the scenario identification are summarized in the following.

a) *Unsupervised Algorithms:* This kind of algorithm does not require labels for data sets. The *k*-means is a typical clustering algorithm without prelabeling. Based on the randomly selected numbers as cluster centers, it iteratively divides data samples into a number of groups by minimizing the intracluster distance. Gaussian mixture model (GMM) can also be used to characterize features using several Gaussian distributed models.

b) *Supervised Algorithms:* They decide the relative parameters of the classifier using the training data sets. Then, this model is used for the classification of unlabeled data sets. The *k*-nearest neighbor (*k*-NN) classifies scenarios by assigning a data sample to the most common class. There is also the weighted *k*-NN algorithm to alleviate the sensitivity of the *k*-NN algorithm. Support vector machine (SVM) is a supervised ML algorithm that works well for classification and regression. It can provide accurate classification by finding the largest gaps of samples in an optimal high-dimensional hyperplane. It behaves well even with noisy and nonnormally distributed data. Random forest consists of multiple decision trees and makes final decisions based on the voting of all decision trees.

c) *RL:* RL is a representative one of the paradigms and methodologies of ML, which is used to describe and solve the problem of maximizing or achieving specific goals through learning strategies in the interaction between agent and environment. RL can solve the problems that cannot be solved by supervised learning methods in communication scenarios. The commonly used RL algorithm is the model-based algorithm, which optimizes and classifies communication networks by modeling the specialized environment. Since such algorithms are highly efficient in the training process, such intelligent methods have received more and more attention by communication communities recently.

2) ML-Based Scenario Identification: In Table IV, several existing works on LoS/NLoS and multiple scenario identification using different ML algorithms are given. Their selected features are also listed.

Most works have focused on LoS and NLoS identification, as they can benefit practical applications, such as localization [80]. With an unsupervised algorithm, the *k*-means algorithm was used to identify LoS and NLoS scenarios for UAV to BS scenarios in [58]. Many works used supervised algorithms for LoS and NLoS identification. The frequently used algorithms are SVM, random forest, artificial neural network (ANN), CNN, LSTM. In addition, semi-supervised learning, such as safe semi-supervised SVM, was widely proposed to realize NLOS identification for indoor smartphone position and navigation based on the features extracted from the channels [71].

The SVM algorithm was used for the identification of LoS and NLoS scenarios for the UAV-to-ground scenario in [59]. Then, a 3-D city map with building position and height was reconstructed. In [53], NLoS identification for the UWB channel was performed using both SVM and a mathematical convolution method. The features employed were CIR and different subsets of eight metrics, e.g., RSS, maximal amplitude, rise time, standard deviation, mean excess delay, RMS DS, kurtosis, and skewness. Different CIR data points were also used during comparison. It was concluded that the increased number of CIR data points or features may not lead to accuracy improvement. In addition, it was shown that the convolution-based method can provide comparable identification performance with fewer CIR data samples.

Using the CNN, classification of LoS and NLoS UWB propagation conditions was performed in [52]. The selected feature was downsampled delay PSD rather than CIR of indoor environment measurement data. It was shown that the proposed method has reduced computational cost while maintaining a similar performance. It was also pointed out that additional features, such as range and energy, can be included, but with an increase in training time. In [49], both SVM and CNN were used for LoS and NLoS scenario identification. With S-V model simulation data, it was pointed out that SVM performed less satisfactorily. It is worth mentioning that short-time Fourier transformation was applied to the CIR, and then the CTF images were used as the input of the CNN. Therefore, the LoS and NLoS identification task was transformed to an image recognition problem.

With the LSTM network, channel state information (CSI) along with RSS, RMS DS, kurtosis, and skewness were used as input features to identify LoS and NLoS conditions in [56]. Based on channel measurements in a meeting room, the results showed that a larger bandwidth and more features, especially the latter term, can benefit the identification accuracy.

There are also works on indoor and outdoor identification, and multiple typical scenario identification. Supervised algorithms are usually used, including SVM, back-propagation neural network (BPNN), CNN, etc. In [50], based on real-time measurement data, the CTF and FCF were compared for four different environments. ML algorithms, including decision trees, SVM, and *k*-NN, were used to classify four scenarios. It was shown that the *k*-NN method with CTF and FCF as features performed the best. It was indicated that this method can be applied to real-time deployment scenarios. In [9], the BPNN was used to identify four V2V scenarios,

TABLE IV
SUMMARY OF EXISTING SCENARIO IDENTIFICATION WORKS USING ML

Identification methods	Ref.	Algorithms	Scenarios	Simulations/ Measurements	Selected features
LoS & NLoS	[59]	k-means	UAV	UMa, UMi, RMa, f_c=900 MHz	RSS
	[60]	SVM	UAV-to-ground	WINNER II, dense urban	RSS
	[55] [56]	SVM, Random forest, ANN	V2V	V2V, MIMO, f_c=5.9 GHz, B=15 MHz	Kurtosis, Rician K-factor
	[54]	SVM, convolution	Indoor, UWB	Indoor office, SISO	Energy, amplitude, rise time, RMS DS, standard deviation, mean excess delay, kurtosis, skewness
	[50] [68] [69]	SVM, CNN	Indoor, UWB	S-V model, SISO	Mean excess delay, RMS DS, number of effective rays, image of CTF
	[70]	Random forest	Indoor, WiFi	Indoor hall, SISO, f_c=5.75 GHz, B=25 MHz	Mean, standard deviation, skewness, kurtosis
	[71]	FNN	Indoor, UWB	Indoor, SISO, f_c=6489 MHz, B=499.2 MHz	Total power, first path power
	[53]	CNN	Indoor, UWB	Indoor, SISO, f_c=499.2 MHz, B=64 MHz	Delay PSD
	[52]	CNN	UAV	Urban dense, macro area, MIMO, f_c=800 MHz	Delay PSD
	[62] [63]	CNN	UMa	3GPP, UMa, massive MIMO	Configurations, space-time-frequency correlation
	[58]	RNN, LSTM, hypothesis test	Indoor, WLAN	Indoor office, SISO, f_c=2.462 GHz, B=20 MHz	CSI, RSS
	[57] [72]	LSTM, SSL	Indoor, acoustic	Meeting room, SISO, f_c=2.4 $\sim$5.4 GHz, B=600 MHz, 1.2 GHz, 3 GHz	CSI, RSS, RMS DS, kurtosis, skewness
Multiple scenarios	[51] [73] [74]	Decision tree, SVM, k-NN	Laboratory, narrow corridor, lobby, sports hall	Indoor, SISO, f_c=2.4 GHz, B=100 MHz	RSS, CTF, FCF
	[9]	BPNN	Urban, suburban, tunnel, NLoSv	V2V, SISO, f_c=5.9 GHz, B=30 MHz	Delay PSD, RMS DS, SF, Rician K-factor
	[75]	CNN	Indoor, Outdoor	Indoor, f_c=900 MHz; Outdoor, f_c=900 MHz/2.4 GHz	MPC parameters
	[76]	Rule-based classification	Highway traffic, automated vehicles	Number of scenarios=110007	Complexity metric
	[78]	Model tuning, deterministic analysis	GSM-R system	HST, SISO, f_c=932 MHz, B=100 MHz	PL, geomorphologic analysis
	[51] [64]	Correlation pattern	Indoor	Open space, lab, SISO, f_c=2.4 GHz, B=100 MHz	CTF, FCF
	[79]	Fuzzy C-means	Industrial	Tennessee Eastman, 4 scenarios	16 attributes
LoS & NLoS; Multiple scenarios	[15]	k-NN, SVM, k-means, GMM	RMa (LoS & NLoS), UMa (LoS & NLoS), HST (Viaduct &Cutting)	Simulation: QuaDRiGa, SISO, f_c=2.6 GHz, B=100 MHz Measurement: HST, SISO, f_c=2.35 GHz, B=50 MHz	PL, Rician K-factor, RMS DS, RMS AS
	[61]	Binary hypothesis testing	Expressway (LoS & NLoS), urban intersection (LoS & NLoS)	V2V, SISO, f_c=5.2 GHz, B=120 MHz	Skewness, kurtosis, RMS DS, peak-to average ratio, Rician K-factor

i.e., urban areas, highways, tunnels, and LoS blocked by vehicles (NLoSv). The input features were delay PSD, RMS DS, shadow fading (SF), and Rician K-factor. It was shown that the proposed BPNN was able to provide over 98% accuracy in identifying four scenarios. In [74], a CNN was used to distinguish indoor and outdoor scenarios at different frequency

bands. The input features were MPC parameters, such as amplitude, time delay, and Doppler frequency. The results showed the astonishing performance of a CNN in dealing with MPC parameters. In [76], scenario classification of the HST GSM-R system was studied based on geographic information systems (GISs). The model tuning and deterministic analysis were proposed to distinguish the special propagation scenarios. In [57], an RNN combined with LSTM was used for indoor office NLoS condition identification with commodity wireless local area network (WLAN) devices. The input vector consisted of the RSS using the real and imaginary parts of the CSI. The early stopping scheme was used to avoid overfitting and the simple hypothesis test was used to make the final decision. It was also found that the phase information may not be suitable for identification in a time-varying environment. In [77], the fuzzy C-means clustering algorithm was used to classify four typical industrial scenarios by using a total of 16 features. In [15], unsupervised and supervised ML algorithms were used to identify both measured HST viaduct and cutting, and simulated QuaDRiGa RMa and UMa scenarios. It was indicated that the unsupervised GMM and supervised k-NN and SVM can provide significant accuracy over 90%, while the k-means can only yield accuracy of approximately 80%. In [60], binary hypothesis testing was used for NLoS identification based on expressway and urban intersection scenarios. Several features, including skewness, kurtosis, RMS DS, peak-to-average ratio, and Rician K-factor, were considered to measure their influences on identification accuracy. It was found that the Rician K-factor performed less satisfactorily and that the RMS DS was significantly impacted by the environment.

3) Power Consumption: ML power consumption has an important influence on improving efficiency and optimizing the deployment of ML algorithms. It is affected by system hardware and software, which mainly considers floating point operation (FLOPs) as a quite important measurement index for unsupervised and supervised algorithms. Different configuration tactics for the same number of parameters will result in different FLOPs. While for high-precision or more complex algorithms, such as deep learning, FLOPs will go up a lot when the exponential growth in the number of system parameters and the size of data sets [78]. However, from the perspective of the classification application system, model simplification techniques are currently used to reduce the ML power consumption, and the overall ML performance gain should be weighed against the power configuration of the entire communication system.

4) Explainability: Currently in the field of communication scenario classification, there is an unexplainable black box problem in ML, which means that users can only see the classification results, without understanding the reasons and processes of their decisions. It is difficult to distinguish the clear logic with closed-form expressions behind artificial intelligence, so the ML explainability has become an urgent problem to be studied and solved necessarily [79]. Toward ML-based intelligent scenario identification like CNN and RL, the mainstream approach to explainability is to analyze CIR through Saliency map (importance map), or to construct a set of simple and effective convolution filters for measuring features changes during training process, which can be used to explain the learning mechanism of and visualize qualitative analysis of important features.

In summary, it can be seen that the selected features can have significant impacts on identification performance. For multiple scenario identification, supervised learning algorithms exhibit better accuracy, and features with abundant information, e.g., CTF and MPC parameters, can yield better performance. However, existing works on scenario identification mainly focus on LoS and NLoS identification, as well as indoor and outdoor classification. There is still lack of the identification of multiple scenarios. In addition, most works focused on single-input–single-output (SISO) channels at sub-6-GHz frequency band, and only a few papers have discussed MIMO channels. Thus, there is a lack of scenario identification at higher frequency band with large array for 6G.

D. Digital Map/Picture-Based Scenario Identification

In addition to the aforementioned scenario identification using on-site channel measurements, digital map/picture is another important method that can be used for efficient scenario identification. It contains abundant physical and geographic information that is sufficient to provide macroscopic identification of scenarios. For example, typical urban scenarios, including open area, hotspot, street canyon, and roof, can be easily distinguished, as they exhibit different numbers of buildings, streets, etc.

The 2-D digital map/picture can be acquired through satellite/camera, reconstruction using some software. In the future, with the development of imaging and emulation technologies, 3-D environmental information, e.g., used in virtual reality, which contains more refined environment restoration, is expected to provide better scenario identification performance. Compared with wireless channel measurement, which requires expensive equipment and exhaustive labor, this is a less expensive and more convenient method. In addition, benefiting from the powerful image processing capability of ML algorithms, real-time scenario identification can be performed in a timely and accurate manner.

Furthermore, based on accurate scenario identification, fundamental channel propagation characteristics can also be predicted using ML algorithms, even with randomly placed Tx and Rx under different frequency bands. In [81], 3-D electronic maps, including locations of Tx and Rx, distance, building density, average height, etc., were associated with real channel measurements to predict the PL of a typical hotspot area. The operated center frequencies were 700 MHz, 2.4 GHz, and 3.5 GHz. It was shown that this method can provide accurate prediction of PL even without a large amount of measurement data. It is worth noting that channel characteristics should not only include large-scale parameters, such as PL and shadowing but also include small-scale parameters. Thus, in real wireless communication, optimum system design can be performed/switched in a timely manner for a superior user experience. However, this line of work has not yet received much attention. There are also ML algorithms researches on

the identification of other special scenarios. For example, ML algorithms were used in [82], such as ANN, SVM, kNN, and RF, to classify land use land cover based on remote sensing image data provided by satellites. The ensemble ML approach combined with multisource geographical image data set to improve the classification accuracy of arid regions was proposed in [83]. Hyper spectral image about studying ML-based crop identification with temporal data information for crop phenology planning was adopted in [84]. The automating road junction ML-based identification using crowdsourcing on GPS transformed digital map data were adopted in [85].

Considering that the 6G wireless communication scenarios are usually complex and the scenario changes all the time, identifying the scenario type and formulating the corresponding parameter optimization strategy is quite difficult and critical for wireless network optimization. The conventional identification scheme relying on subjective experiences and manual decision rules has gradually exposed the disadvantages of low scalability, which is difficult to ensure the identification quality. The artificial intelligence algorithms can improve the scenario identification accuracy, and thus increase the efficiency of wireless network optimization and positioning. In addition, it can predict the channels using the multiple linear regression algorithm and provide an important reference for the adaptive optimization of subsequent communication schemes. However, how to reduce model complexity, training time, computing storage spaces and communication overhead between devices while ensuring accuracy still needs to be studied.

IV. 6G Channel Modeling Methodologies

A. Conventional Nonpredictive Wireless Channel Modeling

Conventional nonpredictive channel models mainly include deterministic and stochastic models [86]. Deterministic channel models use real channel measurements and RT. However, they are site-specific and usually have high complexity to carry out channel measurements or reconstruct the real channels.

Stochastic models use CBSM, GBSM, and BDCM. CBSM relies on independent and identically distributed (i.i.d.) Gaussian assumption. The channel coefficient is fully determined by the spatial covariance matrix. With different approximations to this matrix, one has the Kronecker-based stochastic channel model (KBSM) and Weichselberger model. They have very low complexity, but their accuracy is not satisfactory, especially the KBSM, which assumes a rich scattering environment. GBSM describes wireless channels based on the geometrical relationships among Tx, scatterers, and Rx. The MPC delay, AoD, AoA, amplitude, etc., can be derived with the aid of some empirical distributions. GBSM has very high accuracy and flexibility, but the computational complexity is also very high. BDCM, originally known as virtual channel representation (VCR), is a promising method that can provide a better tradeoff between accuracy and complexity. It characterizes the channel propagation between virtual beam pairs and can provide a performance between CBSM and GBSM. This is an emerging method for future 6G wireless channel modeling.

B. ML-Based Wireless Channel Modeling

The utilization of ML algorithms in wireless channel modeling is twofold: 1) channel characterization and 2) prediction [87]. On the one hand, to reduce the complexity of conventional channel modeling methods, researchers resort to ML algorithms to extract channel statistical parameters and to explore underlying properties. On the other hand, the many possible 6G scenarios pose high demands for channel prediction. Existing channel measurement data is expected to predict the channel statistical properties of wireless channels in the future for new scenarios, and at new frequency bands. In [88], ML-assisted channel modeling and channel estimation were introduced. There are already works on the ML-based channel characterization. They are summarized in Table V.

Many works have been performed for wireless channel modeling at sub-6 GHz and mmWave frequency bands. In [89], a generative adversarial networks-long short term memory (GAN-LSTM) framework was proposed to predict sub-6-GHz channel statistics, including RMS DS and RMS AS. In [93], CNN was used to predict the PL exponent of outdoor mmWave band channels. The impacts of building density and average distance from the Tx were analyzed using RT simulation data. In [90] and [91], a hybrid physics-based and data-driven modeling framework was proposed to show very high accuracy and great generalization ability. Based on real measurement and RT simulation data, the through-vegetation cluster parameters, including vegetation attenuation, RMS DS, and RMS AS, were predicted using an ANN. Instead of estimating channel model parameters through high-resolution MPC parameter estimation and clustering, summarized statistics that hold enough information were learned by using a ML algorithm based on approximate Bayesian computation and a DNN with two hidden layers in [94]. The summarized statistics include reflection gain, number of scatterers, probability of visibility, polarization ratio, and noise variance. The performance of the proposed methods was validated based on simulation data and real measurement data from an indoor NLoS environment at 60 GHz. In [64], with extensive training data collected from real channel measurements and GBSM simulations, channel statistical properties, such as received power, RMS DS, and RMS AS, could be exported using the feed-forward neural network (FNN) and RBF-NN. This result indicated that ML will play an important role in future channel modeling. In [100], RF and KNN were used for the prediction of PL and RMS DS. A feature selection scheme was also proposed to further improve the prediction accuracy. In [96], the PL and RMS DS were predicted for 60-GHz mmWave channels in corridor and hall. The ML methods used included back propagation (BP), SVM, and genetic algorithm (GA). It was shown that the combination of SVM with GA can provide excellent fitness with the measurement data. In [106], PL fading in an underground former gold mine at 60 GHz was estimated using MLP and RBF.

There are also works on channel modeling for other scenarios, including satellite, V2V, UAV, massive MIMO, etc. In [54], to handle the very large amount of time-varying V2V channel measurement data, ML techniques were used

TABLE V
ML-BASED WIRELESS CHANNEL CHARACTERIZATION/PREDICTION

Ref.	Scenario	ML algorithms	Statistical properties
[90]	Sub-6 GHz	GAN-LSTM	Received power, delay PSD, RMS DS, RMS AS
[65]	MmWave	FNN, RBF-NN	Received power, RMS DS, RMS AS
[91] [92]	MmWave (28 GHz, street canyon)	ANN	Vegetation attenuation, RMS DS, RMS AS
[93]	MmWave (28 GHz, urban)	Multiple linear regression	PL
[94]	MmWave (outdoor)	CNN	PL
[95]	MmWave (60 GHz, indoor)	Approximate Bayesian computation, DNN	Reflection gain, scatterers number, visibility probability, polarization ratio, noise variance
[96]	MmWave (60 GHz)	RBF-NN	RSS, PL, shadow fading
[97]	MmWave (60 GHz)	BP, SVM, GA	PL, RMS DS
[98]	VVLC	MLP-NN, RBF-NN, RF	PL
[55]	V2V	Kalman filters, particle filters; Hungarian method, Kuhn-Munkres-based method	MPC tracking, clustering
[99]	V2V	RF	PL
[100]	UAV	EBT, GPR	End-to-end loss
[101]	UAV, mmWave (60 GHz)	RF, k-NN	PL, RMS DS
[102]	High mobility massive MIMO	ST-AR, CVNN	Angel-delay domain channel
[103] [104]	Time-varying, mmWave (outdoor, 26 GHz)	RBF-NN, ANN	PL, shadow fading, RMS DS, RMS AS
[105]	Terrestrial, urban	MLP	Field strength (RSS)
[106]	Urban, 170 MHz	MLP	Field strength (RSS)
[107] [108]	Underground mine	MLP, RBF	PL
[109]	Satellite	ANN	Channel excess attenuation
[110]	DSRC	MLP-NN	PL

for LoS/NLoS identification, MPC tracking, and clustering [110]. This provides essential information for further channel modeling. In [108], an ANN was used to estimate the channel excess attenuation of a Q-band satellite channel. In [92], multiple linear regression was proposed to predict the PL model of a different operating environment using the measurement data of a certain scenario. In [111], a cluster kernel-based channel model was proposed to take advantage of both stochastic and deterministic channel models. In [97], to improve the accuracy of deterministic and stochastic models, the PL was predicted for vehicular visible light communication (VVLC), and a CTF model was proposed using ML algorithms. The vehicle mobility and environmental effect-related parameters were considered as inputs, such as distance, ambient light, Rx inclination angle, and optical turbulence. MLP-NN, RBF-NN, and decision tree-based RF algorithms were employed and compared for real measurement data to demonstrate the high accuracy of PL prediction. In [102], the RBF-NN was used to build the PL and shadowing model, and the ANN was used to build the time-varying joint small-scale channel parameters. In [112], big data enabled cluster-based channel modeling methods were summarized, including clustering techniques, cluster tracking algorithms, and different cluster-based channel models. In [97], the channel modeling of the VVLC was proposed to improve the model accuracy for pass loss, and build the channel frequency response model through consideration of vehicle mobility and environmental effects. In [113], ML-based multilayer perceptron model was

trained with statistics computed from channel realizations, to calibrate the model parameters as a regression problem involving mapping of the CTF or impulse response. In [114], a combination of the Exponential and the Generalized Gamma Distribution was proposed to model the underwater channel environment with great accuracy, and built a convolutional neural network capable of estimating the parameters from received signal.

For channel prediction, in [109], a real-time PL prediction method was proposed based on the MLP for dedicated short-range communications (DSRCs). A higher prediction accuracy was achieved in comparison to the statistical method. In [98], a nonparameterized data-driven approach RF was used for the PL prediction of V2V communication. The contributions of different features were also discussed to achieve accurate performance. In [101], both unsupervised and supervised ML techniques were leveraged for massive MIMO-OFDM high-mobility wireless channel prediction. Through 3GPP NLoS scenario simulation, the proposed spatio-temporal autoregressive (ST-AR) and complex-valued NN (CVNN)-based channel prediction methods showed enhanced channel prediction performance. In [99], the prediction of end-to-end loss (including SF) was formulated as a supervised regression problem. The authors used ensemble bagged trees (EBTs) and exponential Gaussian process regression (GPR) methods to process raw data and processed data, respectively, to show accurate performance. In [95], the RBF-NN was used to investigate the receive power, PL, and SF of a 60-GHz mmWave channel. It showed that the RBF-NN outperformed the BP network.

It can be seen that, NNs are widely used for wireless channel modeling. The most interesting statistical properties are PL, SF, RMS DS, and RMS AS. In channel prediction works, most output features are PL. There is still lack the prediction of more channel statistical information. In both channel modeling and prediction, the involved scenarios are very limited.

C. Pervasive Channel Modeling

Although the above mentioned conventional nonpredictive and ML-based predictive channel models have different merits in terms of accuracy and complexity, there is still a lack of unified channel modeling framework that is pervasive for all frequency bands and all scenarios. It is desirable to analyze system performance with different technologies, frequency bands, array sizes [115], etc., under the same channel modeling framework. The pervasive channel model should be beneficial to figure out the relationships among model parameters, propagation characteristics, as well as system performance. It is especially important for the 6G channel model standardizations and the investigation of fundamental theories and technologies of 6G networks. In addition, with the adjusted model parameters, the pervasive channel model can be degenerated to the simplified channel models for specific scenarios.

As shown in Fig. 3, a pervasive channel modeling theory can be established with a unified modeling methodology, a unified CIR, and integrated statistical properties of 6G channels for all spectra and all scenarios. One intuitive option is to utilize the pervasiveness of the GBSM and to fit most channel characteristics of various scenarios into the GBSM. In [116], the 6GPCM was proposed based on the GBSM. It combined most channel characteristics into a unified framework and derived a unified CIR for all frequency bands and all scenarios. The impacts of frequency bands, scenarios, mobile velocities, and antenna array sizes were analyzed based on the 6GPCM. Another option is to use the GBSM as the main framework. Since RT has high accuracy and can provide MPC parameters in very complicated environments, while ML can predict channel characteristics in future time, new frequency band, and unknown environments, they can be used as supportive methods for the main framework, where GBSM is the core, to provide extra credits of the pervasive channel model.

The pervasive model uses a unified cluster-based geometric stochastic channel modeling method, a unified CIR expression, and a comprehensive consideration of the statistical characteristics of 6G all-band all-scenario channels. It can be simplified to a target channel model for a specific frequency band and a specific scenario by adjusting the parameters of the channel model. Through the analysis of 6G universal channel model, the complex mapping relationship between channel model parameters, channel statistical characteristics and communication system performance, frequency bands and scenarios, can be studied. As a unified channel model framework, it is crucial to promote 6G channel model standardization, 6G generic theory, and system fusion construction.

V. 6G Scenario Adaptive Channel Modeling

The 6G scenario adaptive channel modeling can be achieved through intelligent scenario identification and automatic channel model parameter matching, as shown in Fig. 4. After identifying the targeted communication scenario, the relative channel model parameters should be automatically matched and the dedicated channel model should be abstracted.

In this section, we will use the 6GPCM as an example to serve as the core of 6G scenario adaptive channel modeling [116]. Detailed procedures and all the 6GPCM related parameters are summarized in Fig. 4, including the user defined and wireless channel parameters. As for the user defined parameters, system setup, array configuration, and static/mobility related parameters need to be determined in advance. The large scale fadings (LSFs), cluster related parameters, and MPC related parameters should be matched to the targeted scenario. Under the simplified 6GPCM framework, dedicated CIR/CTF for a certain scenario can then be derived.

In the following, we will show how to use of 6GPCM, model parameter matching, and its abbreviations for several typical space-air-ground-sea communication scenarios.

A. Complex Channel Matrix

The complex channel matrix of the 6GPCM is given as

$$\mathbf{H} = [PL \cdot SH \cdot BL \cdot WE \cdot AL]^{1/2} \cdot \mathbf{H_s} \tag{6}$$

where PL, SH, BL, WE, and AL denote the LSFs, i.e., path loss, shadowing, blockage loss, weather effect loss, and

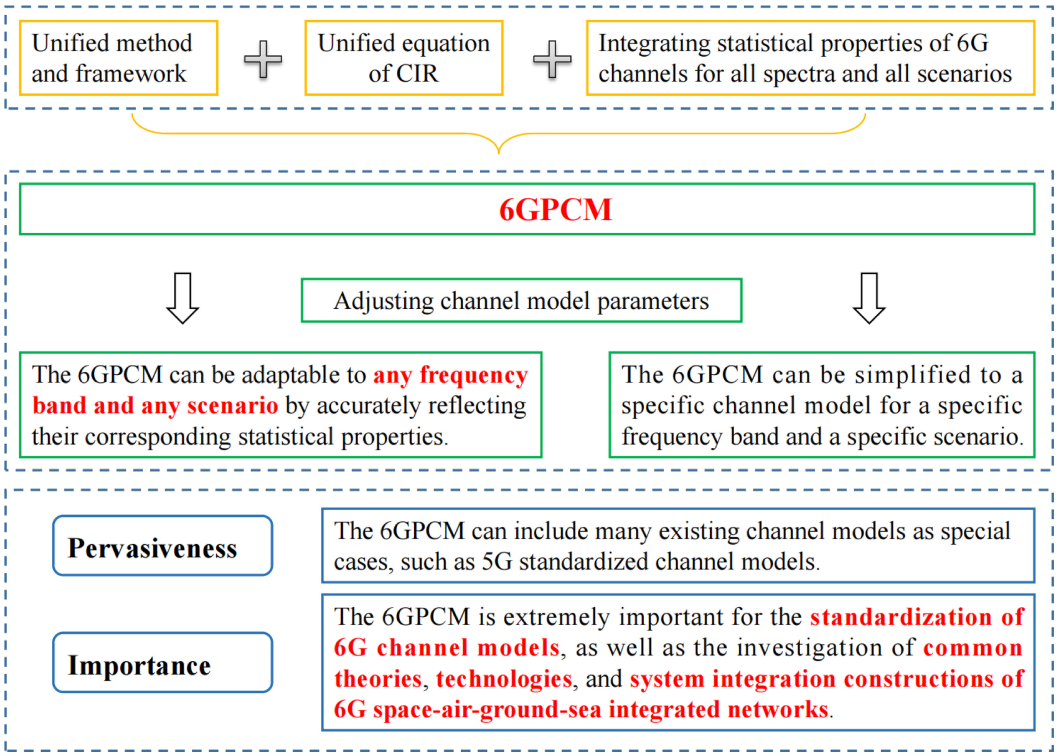


Fig. 3. Pervasive channel modeling theory [116].

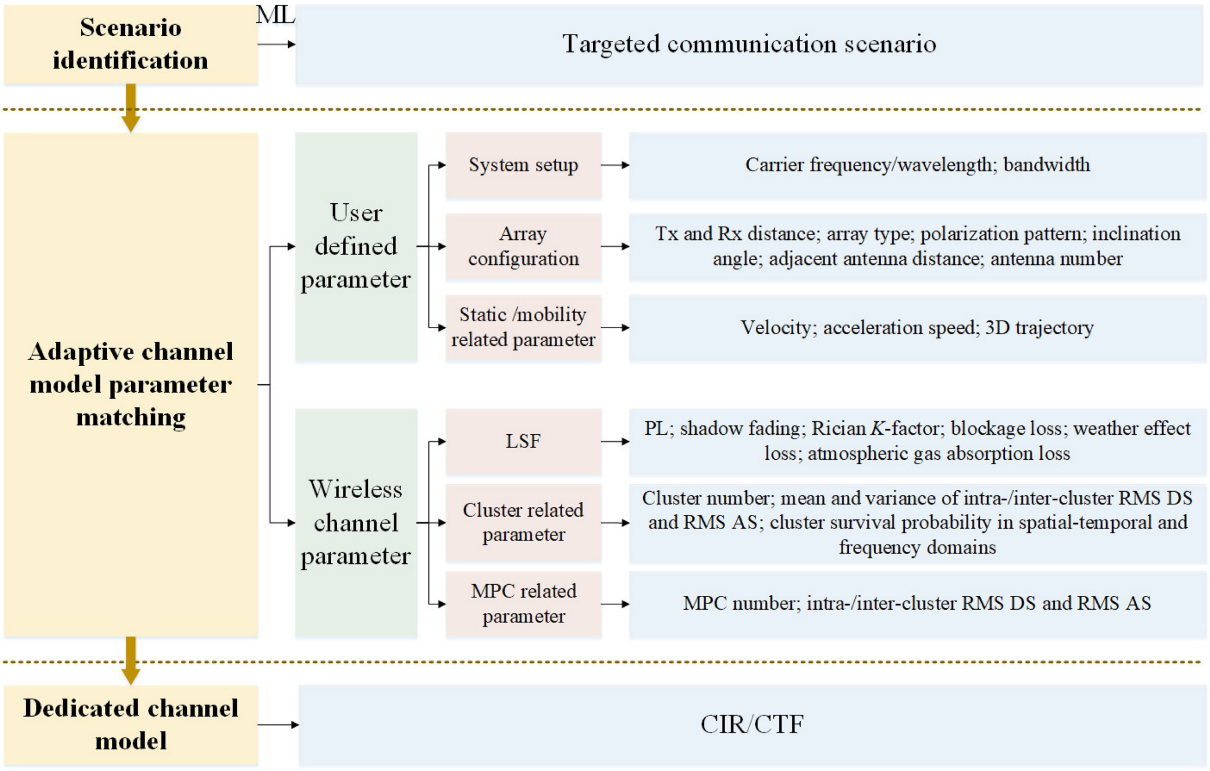


Fig. 4. Detailed procedures and parameter matching of scenario adaptive channel modeling.

atmospheric gas absorption loss, respectively. The small-scale fading (SSF) is represented by $\mathbf{H_s} = [h_{qp,f_c}(t,\tau)]_{M_R \times M_T}$, where $p = 1, \ldots, M_T$ and $q = 1, \ldots, M_R$, M_T and M_R are Tx and Rx antenna element numbers, respectively. The CIR $h_{qp,f_c}(t,\tau)$ is derived as the sum of LoS and NLoS components, as illustrated in (7)–(9), shown at the bottom of p. 7302. Here, $K(t)$ is the Rician factor at time instant t. There are antenna patterns, Faraday rotation, and amplitude, delay, AAoA, AAoD, EAoA, and EAoD for each MPC of clusters. More detailed symbol interpretations and derivations of the

TABLE VI
DEFINITIONS OF SIGNIFICANT PARAMETERS

Parameters	Definitions
A_p^T, A_q^R	The pth Tx antenna element and qth Rx antenna element
δ_T, δ_R	Antenna spacings of Tx and Rx arrays
C_n^A, C_n^Z	The first-bounce and last-bounce clusters of nth path
$v^T(t), v^R(t), v^{A_n}(t), v^{z_n}(t)$	Speeds of the Tx, Rx, cluster C_n^A, and cluster C_n^Z at time t
$\alpha_A^T(t), \alpha_A^R(t), \alpha_A^{A_n}(t), \alpha_A^{Z_n}(t)$	Azimuth angles of movements of the Tx, Rx, cluster C_n^A, and cluster C_n^Z at time t
$\alpha_E^T(t), \alpha_E^R(t), \alpha_E^{A_n}(t), \alpha_E^{Z_n}(t)$	Elevation angles of movements of the Tx, Rx, cluster C_n^A, and cluster C_n^Z at time t
$\phi_{A,m_n}^T, \phi_{E,m_n}^T$	Azimuth angle of departure (AAoD), elevation angle of departure (EAoD) from $A_1^{(T)}$ to the mth scatterer in $C_n^{(A)}$ at initial time
$\phi_{A,m_n}^R, \phi_{E,m_n}^R$	azimuth angle of arrival (AAoA), and elevation angle of arrival (EAoA) from $A_1^{(R)}$ to the mth scatterer in $C_n^{(Z)}$ at initial time
$\phi_{A,L}^T, \phi_{E,L}^T, \phi_{A,L}^R, \phi_{E,L}^R$	AAoD, EAoD, AAoA, and EAoA from A_1^T to A_1^R at initial time
$d_{m_n}^T, d_{m_n}^R$	Distance from $A_1^{T(R)}$ to the mth scatterer in $C_n^{A(Z)}$at initial time
$d_{p,m_n}^T(t), d_{q,m_n}^R(t)$	Distance from $A_1^{T(R)}$ to the mth scatterer in $C_n^{A(Z)}$at time t
$\widehat{d}_{p,m_n}^T(t), \widehat{d}_{q,m_n}^R(t)$	Unit vector from $A_1^{T(R)}$ to the mth scatterer in $C_n^{A(Z)}$at time t
$D, D_{qp}(t)$	Distance from A_1^T to A_1^R at initial time and time T, respectively
$f_{D,qp}^L(t), f_{D,qp,m_n}(t)$	Doppler shift from A_p^T to A_q^R of the LOS path and the NLOS path at time t
$P_{qp,m_n}(t), \tau_{qp,m_n}(t)$	Power and delay of the ray from A_p^T through the mth scatterer in C_n^A and the mth scatterer in C_n^Z to A_q^R at time t
$F_{p(q),V}, F_{p(q),H}$	Antenna patterns of $A_p^T(A_q^R)$ for vertical and horizontal polarizations

LoS and NLoS components can be found in [116]. All parameters are time variant and change with the cluster birth-death in the spatial, temporal, and frequency domains, $h_{qp,f_c}(t,\tau)$, the CIR between $A_p^T$and A_q^R at the carrier frequency f_c, is derived as the sum of LoS and NLoS components, as illustrated in (7). Here, $K(t)$ is the Rician factor at time instant t, and the calculation of $h_{qp,f_c}^{\mathrm{LoS}}(t,\tau)$ and $h_{qp,f_c}^{\mathrm{NLoS}}(t,\tau)$ can be represented as (8) and (9), respectively, $F_{p(q),f_c,V}(\cdot)$ and $F_{p(q),f_c,V}(\cdot)$ are the antenna patterns of Tx (Rx) antenna for vertical and horizontal polarizations at corresponding carrier frequency f_c, respectively, $\kappa_{m_n}(t)$ is the cross polarization power ratio, μ is co-polar imbalance, $\theta_{m_n}^{VH}$, $\theta_{m_n}^{HV}$, and $\theta_{m_n}^{HH}$ are initial phases modeled as random variables uniformly distributed over (0, 2π]. Besides, F_r represents Faraday rotation referring to the rotation of the polarization plane caused by the propagation of electromagnetic waves through the ionosphere in LEO satellite scenario, and $\Psi_{m_n} = 108/f_c^2$ is the Faraday rotation angle, where f_c is in GHz. Otherwise, in scenarios without considering the influence of ionosphere, we can set $\Psi_{m_n} = 0$. Additionally, $P_{qp,m_n,f_c}(t)$ and $\tau_{qp,m_n,f_c}(t)$ is the power and delay of the mth ray in the nth cluster between A_p^T and A_q^R at time instant t. Also, θ_L^{VV} and θ_L^{HH} denote random phase in (0, 2π], $\tau_{qp}^L(t)$ is the time delay of LoS path at time instant t, and given the speed of light c, $\tau_{qp}^L(t)$ can be calculate as $\tau_{qp}^L(t) = \|\vec{A}_q^R(t) - \vec{A}_p^T(t)\|/c = D_{qp}(t)/c$, where $\|\cdot\|$ calculates the Frobenius norm. Just to be clear, Table VI summarizes the definitions of significant parameters.

B. Parameterization and Modeling of 6G Communication Scenarios

In [116, Table IV], the configurations and model parameters of the 6GPCM at different frequency bands and scenarios were listed, including center frequency, antenna number, SF, RMS DS, RMS AS, etc. However, only indoor scenarios in the THz band, UAV-to-ground scenario, and ultramassive MIMO scenario were discussed. More dedicated parameter settings for all scenarios will be added in the future.

In this work, we emphasize the parameters that may be sensitive to wireless communication scenarios and system settings. Taking the space satellite, UAV, terrestrial V2V and HST, and maritime communication scenarios as examples, we list the most distinct channel characteristics, parameters and modeling methods, and key parameter adjustments in Table VII. Only the single-link and single-frequency case is considered.

In satellite channels, the ionosphere effect and rain attenuation should be considered. We characterize them by introducing the Faraday rotation matrix $\mathbf{F}_{\mathrm{r}}$, which represents the rotation of the polarization plane caused by electromagnetic wave propagation through the ionosphere, and the rain attenuation factor RA. The space scenario can be characterized by defining in the 6GPCM $AL = BL = 1$, the co-polar imbalance $\mu = 1$, the time-invariant MPC number in the nth cluster $M_n(t) = M_n$, the cluster survival probability at integral multiples of frequency and space intervals $P_{\mathrm{surv}}(\Delta f) = P_{\mathrm{surv}}(\Delta r) = 1$, the frequency dependent factor $\gamma_{m_n} = 0$ because of the limited bandwidth, and the 2-D spatial lognormal process $\xi_n(p,q) = 1$, which simulates the smooth power variation at the pth transmit and qth receive antennas.

The aerial scenario must include the 3-D movement property. This can be done by introducing elevation angles for the Tx, Rx, and scatterers. The LSPs should be related to the UAV height. The simplified parameters for this specific channel are $AL = BL = WE = 1$, $\mu = 1$, $M_n(t) = M_n$,

TABLE VII
SIMPLIFIED MODELS OF THE 6GPCM AND PARAMETER ADJUSTMENTS

Scenarios	Distinct channel characteristics	Parameters and modeling methods	Other key parameter adjustments
Space (LEO satellite) [118]	Ionosphere effect	Faraday rotation matrix ($\mathbf{F}_{\mathrm{r}}$)	$AL=1$, $BL=1$, $\mu=1$, $M_n(t)=M_n$, $P_{\mathrm{surv}}(\Delta f)=1$, $\gamma_{m_n}=0$, $P_{\mathrm{surv}}(\Delta r)=1$, $\xi_n(p,q)=1$
	Rain attenuation	Model rain attenuation (RA)	
Aerial (UAV) [119]	3D movement	Movements of Tx, Rx, and scatterers all have elevation angles	$AL=1$, $BL=1$, $WE=1$, $\mu=1$, $\psi_{m,n}=0$, $M_n(t)=M_n$, $P_{\mathrm{surv}}(\Delta f)=1$, $\gamma_{m_n}=0$, $\xi_n(p,q)=1$
	LSPs relate to the height	LSPs relate to the height	
Terrestrial (IIoT) [120]	Existing of DMCs	Model DMCs	$WE=1$, $\mu=1$, $\psi_{m,n}=0$, $M_n(t)=M_n$, $P_{\mathrm{surv}}(\Delta f)=1$, $\gamma_{m_n}=0$, $h^{\mathrm{NLoS}}_{qp,f_c}(t,\tau)=h^{\mathrm{NLoS_{SC}}}_{qp,f_c}(t,\tau)+h^{\mathrm{NLoS_{DMC}}}_{qp,f_c}(t,\tau)$
Terrestrial ((U)HST) [121]	Large Doppler shift/spread	Doppler frequency of MPC is time-variant	$WE=1$, $\mu=1$, $\psi_{m,n}=0$, $M_n(t)=M_n$, $v^{A_n}=0$, $v^{Z_n}=0$, $v^{T}=0$, $P_{\mathrm{surv}}(\Delta f)=1$, $\gamma_{m_n}=0$, $P_{\mathrm{surv}}(\Delta r)=1$, $\xi_n(p,q)=1$
	Time domain non-stationarity	Channel parameters are time-variant, clusters are distributed on the inner wall of the vacuum tube and birth-death in time domain	
	Waveguide effect	$N_{qp}(t)$ relates to waveguide effects in different positions	
Maritime [122]	Location dependence	$h^{\mathrm{LoS}}_{qp,f_c}(t,\tau)$, $h^{\mathrm{NLoS}_1}_{qp,f_c}(t,\tau)$, and $h^{\mathrm{NLoS}_2}_{qp,f_c}(t,\tau)$ will appear or disappear in the channel	$AL=1$, $BL=1$, $WE=1$, $\mu=1$, $\psi_{m,n}=0$, $M_n(t)=M_n$, $h_{qp,f_c}(t,\tau)=\sqrt{\frac{K(t)}{K(t)+1}}h^{\mathrm{LoS}}_{qp,f_c}(t,\tau)+\sum_{i=1}^{2}\sqrt{\frac{S_i}{K(t)+1}}h^{\mathrm{NLoS}_i}_{qp,f_c}(t,\tau)$, $P_{\mathrm{surv}}(\Delta f)=1$, $\gamma_{m_n}=0$, $\xi_n(p,q)=1$
	Fluctuation of sea waves	Pierson-Moskowitz (P-M) spectrum	

$P_{\mathrm{surv}}(\Delta f)=1$, $\gamma_{m_n}=0$, $\xi_n(p,q)=1$, and the Faraday rotation angle $\psi_{m,n}=0$ in the 6GPCM for this scenario.

In terrestrial scenarios, we take the IIoT and ultra-HST (UHST) scenarios as examples. First, for the IIoT channel, dense multipath is the key channel characteristic. We can address this characteristic by analyzing the model dense MPCs. Other parameters that can be fixed and the suggested values are: $WE=1$, $\mu=1$, $\psi_{m,n}=0$, $M_n(t)=M_n$, $P_{\mathrm{surv}}(\Delta f)=1$, and $\gamma_{m_n}=0$. In Fig. 5, we give an example of 6GPCM to fit the RMS DS of IIoT channel measurement data in [122]. Two cases with different antenna heights are compared. In cases 1 and 2, we set the height of antenna arrays to 0.84 and 1.6 m, respectively. Here, most of the surrounding objects are higher than the Tx and Rx antenna arrays in case 1, while the objects are generally lower than the antenna arrays in case 2. As shown in Fig. 5, the larger the height of the antenna arrays, the larger the RMS DS will be. The simulation result is consistent with the measurement data, since there are less multipath components that are blocked by surrounding objects to reach the Rx side in case 2.

Then, for the (U)HST channel, there is a very large Doppler shift. Distinct time domain nonstationarity brings cluster birth-death in the time domain and time-variant channel parameters. In addition, clusters are distributed on the inner wall of the

$$h_{qp,f_c}(t,\tau)=\sqrt{\frac{K(t)}{K(t)+1}}h^{\mathrm{LoS}}_{qp,f_c}(t,\tau)+\sqrt{\frac{1}{K(t)+1}}h^{\mathrm{NLoS}}_{qp,f_c}(t,\tau). \tag{7}$$

$$h^{\mathrm{LoS}}_{qp,f_c}(t,\tau)=\begin{bmatrix}F_{q,f_c,V}\left(\phi^R_{E,\mathrm{L}}(t),\phi^R_{A,\mathrm{L}}(t)\right)\\ F_{q,f_c,H}\left(\phi^R_{E,\mathrm{L}}(t),\phi^R_{A,\mathrm{L}}(t)\right)\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}e^{j\theta^{VV}_{\mathrm{L}}} & 0\\ 0 & e^{j\theta^{HH}_{\mathrm{L}}}\end{bmatrix}\mathbf{F}_{\mathrm{r}}\begin{bmatrix}F_{p,f_c,V}\left(\phi^T_{E,\mathrm{L}}(t),\phi^T_{A,\mathrm{L}}(t)\right)\\ F_{p,f_c,H}\left(\phi^T_{E,\mathrm{L}}(t),\phi^T_{A,\mathrm{L}}(t)\right)\end{bmatrix}\cdot e^{j2\pi f_c\tau^L_{qp}(t)}\delta\left(\tau-\tau^L_{qp}(t)\right) \tag{8}$$

$$h^{\mathrm{NLoS}}_{qp,f_c}(t,\tau)=\sum_{n=1}^{N_{qp}(t)}\sum_{m=1}^{M_n(t)}\begin{bmatrix}F_{q,f_c,V}\left(\phi^R_{E,m_n}(t),\phi^R_{A,m_n}(t)\right)\\ F_{q,f_c,H}\left(\phi^R_{E,m_n}(t),\phi^R_{A,m_n}(t)\right)\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}e^{j\theta^{VV}_{mn}} & \sqrt{\mu\kappa^{-1}_{m_n}(t)}e^{j\theta^{VH}_{mn}}\\ \sqrt{\kappa^{-1}_{m_n}(t)}e^{j\theta^{HV}_{mn}} & \sqrt{\mu}e^{j\theta^{HH}_{mn}}\end{bmatrix}\mathbf{F}_{\mathrm{r}}$$
$$\begin{bmatrix}F_{p,f_c,V}\left(\phi^T_{E,m_n}(t),\phi^T_{A,m_n}(t)\right)\\ F_{p,f_c,H}\left(\phi^T_{E,m_n}(t),\phi^T_{A,m_n}(t)\right)\end{bmatrix}\sqrt{P_{qp,m_n,f_c}(t)}\cdot e^{j2\pi f_c\tau_{qp,m_n}(t)}\cdot\delta\left(\tau-\tau_{qp,m_n}(t)\right). \tag{9}$$

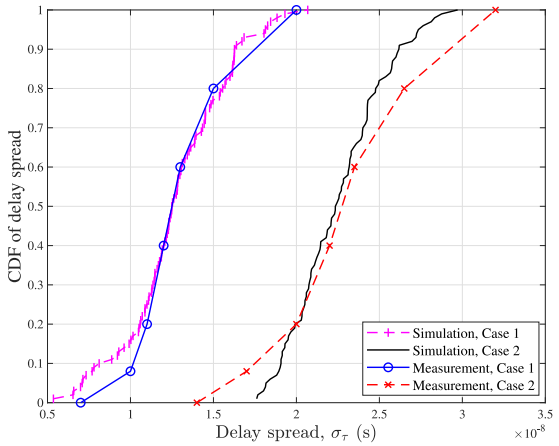


Fig. 5. CDFs of RMS DSs in IIoT scenarios with different antenna heights ($f_c = 28$ GHz, $M_R = M_T = 1$, $v^R = v^T = 0$ m/s, $N(t_0) = 15$).

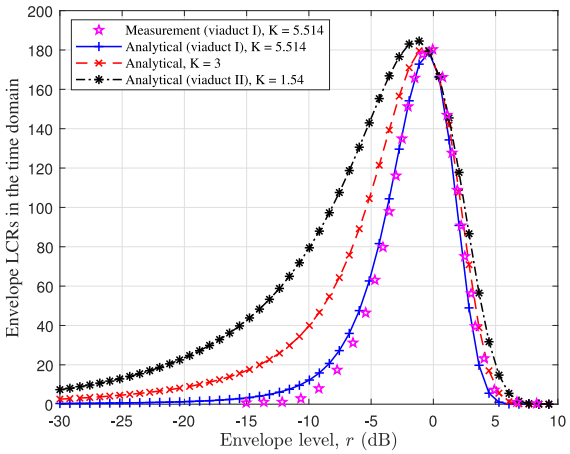


Fig. 6. LCRs in HST viaduct scenarios ($f_c = 930.2$ MHz, $K = 5.514$ in viaduct I, $K = 1.54$ in viaduct II, $v^{Zn} \sim U(0, 5)$ m/s).

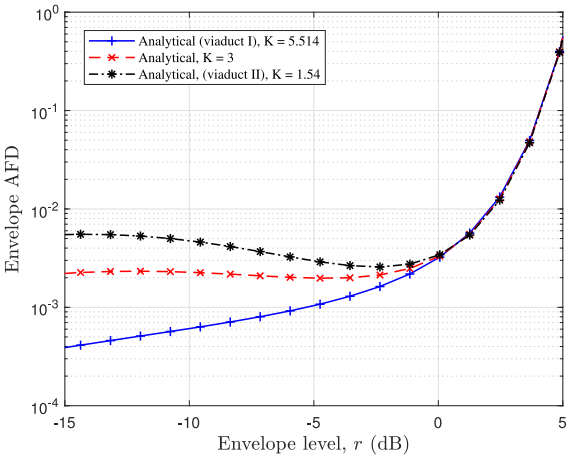


Fig. 7. AFDs in HST viaduct scenarios ($f_c = 930.2$ MHz, $K = 5.514$ in viaduct I, $K = 1.54$ in viaduct II, $v^{Zn} \sim U(0, 5)$ m/s).

vacuum tube, and the waveguide effect must be characterized by setting the number of clusters between the pth and qth antenna in different positions. In this scenario, the channel has: $WE = 1$, $\mu = 1$, $\psi_{m,n} = 0$, $M_n(t) = M_n$, $v^{A_n} = v^{Z_n} = v^T = 0$, $P_{\text{surv}}(\Delta f) = P_{\text{surv}}(\Delta r) = 1$, $\gamma_{m_n} = 0$, and $\xi_n(p, q) = 1$. In Figs. 6 and 7, the level crossing rate (LCR) and average fading duration (AFD) are simulated using the 6GPCM. The fitted two groups of viaduct channel measurement data with different height and also different scatterer heights are referred from [123]. Among them, the biggest difference between viaduct I and viaduct II scenarios is the difference of Rician factor K, i.e., the proportion of LoS component, caused by the height difference of the Tx relative to the surrounding scatterers (trees). From the simulated and measurement results shown in Figs. 6 and 7, we can see that the LCR at the envelope level of 5 dB below the median value in viaduct II is a little more than twice of the value in viaduct I, while the AFD varies slightly.

In the maritime scenario, the channel is location dependent. So that the LoS and NLoS MPCs in the rough sea surface and over the sea surface evaporation waveguide will appear and disappear in the channel. The fluctuation of sea waves can be characterized by the Pierson-Moskowitz spectrum. Other simplified parameters are $AL = BL = WE = 1$, $\mu = 1$, $\psi_{m,n} = 0$, $M_n(t) = M_n$, $P_{\text{surv}}(\Delta f) = 1$, $\gamma_{m_n} = 0$, and $\xi_n(p, q) = 1$.

VI. Future Research Directions and Challenges

A. Effective Feature Selection for Various 6G Scenario Identification

In addition to global coverage, 6G will be an integrated intelligent network that covers all spectra and full applications. The exploited frequency resource will move from conventional sub-6 GHz to mmWave, THz, and optical wireless. Other applications, such as (ultra-)massive MIMO, RIS, and IIoT, should also be included. Therefore, scenario identification is a challenging task for 6G. How can we fully consider and make better use of the relative characteristics to improve the accuracy of scenario identification? The features commonly used are CIR, CSI, delay PSD, kurtosis, skewness, Rician K-factor, RMS DS/AS, etc. To handle more complex scenario identification tasks, more general and effective features should be proposed. It should not only use the pure environmental features but also use features that are representative of different frequency bands and applications. For example, the improved PL and coherent bandwidth should be considered, as PL is usually larger at higher frequency bands and there exists frequency nonstationarity for a larger bandwidth. The stationary interval of (ultra-)massive MIMO, scatterer height and density, etc., should also be considered.

B. Efficient ML Algorithm for Grandiose Data and Propagation Channel Digital Map Processing

With the rapid development and deployment of 6G and the optimization of network performance, ML and digital map are gradually received more attention in communication scenarios. Digital map extends from simple architectural map to new intelligent maps containing complex information [124]. First, digital map evolves from static to dynamic because the mobility with intelligent terminals. Second, the digital map contains not only the conventional building environment information but also the intelligent terminal information or network information, as well as the frequency band and bandwidth used in each scenario. Generally, the new digital map

has large data volume, high complexity, and high-processing cost. The improvement of digital map model structure, data dimension, and dynamic response will bring massive communication data to be processed. In this case, conventional data processing methods are no longer suitable for the new intelligent digital map processing. Thus, new ML methods is adopted to further extract and analyze the features of intelligent digital map for identifying channel classification information. In addition, it is difficult to uniformly analyze and process the physical environment features in different communication scenarios without standardization of digital map. Therefore, it is necessary to establish a unified framework and standards for digital map.

C. 6G Pervasive Channel Model and RT for Adaptive Channel Modeling

6GPCM also needs to be further studied and expanded in the specific segmentation scenarios of space-air-ground-sea, and more scenario measurement data is needed for verifying and refining. First of all, in terms of the full frequency band, how to solve the problem of spectrum resource congestion, research, such as short-wave, infrared and ultraviolet spectrum characteristics and comprehensive utilization is a future challenge. Second, how to build an integrated network of space-air-ground-sea in the full coverage scenario channels, and study global deep coverage and channel modeling, including underwater communication and underground communication, is another challenge. Third, how to study orbital angular momentum communication and other channel characteristics analysis, scenario classification and identification, scenario parameter adaptive matching, etc. is also a key challenge in the future in terms of all-application scenario channels and the key communication technologies spawned.

It is an important direction for RT simulation to help building high-precision channel information database, and provide effective data supports for adaptive wireless channel modeling. However, for mobile communication scenarios, such as vehicle networking and drones in 6G scenarios, RT simulation will meet huge computational complexity and processing problems, and how to effectively model channel properties to hold better high accuracy when solving complexity problems is a certain challenge.

The complexity of RT calculation depends on the complexity of the scenario and the maximum number of reflections allowed by the propagation mechanism. It determines the number of intersections between rays and surfaces, which is also the most time-consuming part of RT calculation. Currently, preprocessing algorithms can usually be used to reduce the number of faces and optimize the number of rays emitted, to reduce the number of intersections. With the development of computer hardware resources, optimization can also be achieved through high-performance GPU parallel acceleration. Considering the computational complexity of large-scale complex scenarios, such as the Industrial IoT or mobile scenarios, high-performance computing resources may be required, and the simulation can be deployed on servers with GPU acceleration [125], [126].

VII. Conclusion

In this work, a comprehensive framework that integrates scenario extension, identification, and adaptive channel modeling has been proposed for 6G. By comparing scenario categorization in existing standardization documents, it has been shown that the scenario categorization needs to be extended to cover most 6G scenarios. Furthermore, the scenario identification technique has been comprehensively investigated. The scenario features, ML-based algorithms, and data preprocessing methods have been carefully analyzed and compared. The novel scenario adaptive channel modeling has been proposed, in which the model parameter matching for particular 6G scenarios has been illustrated. Based on the 6GPCM, statistical properties, such as IIoT and HST scenarios, have been simulated. The RMS DSs, LCRs, and AFDs of the channel model simulations have shown great agreement with those of the channel measurement data. It has also been indicated that the proposed scenario categorization method can serve as an important prerequisite for the accurate identification of 6G communications scenarios and adaptive modeling of 6G wireless channels. Finally, future research directions and challenges have been given.

References

[1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[2] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.

[3] G. Liu et al., "Vision, requirements and network architecture of 6G mobile network beyond 2030," *China Commun.*, vol. 17, no. 9, pp. 92–104, Sep. 2020.

[4] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, 2nd Quart., 2023.

[5] X.-H. You et al., "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, Jan. 2021, Art. no. 110301.

[6] C.-X. Wang, J. Huang, H. Wang, X.-Q. 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.

[7] H. Zhou, L. Jiao, S. Zheng, L. Yang, W. Shen, and X. Yang, "Generative adversarial network-based electromagnetic signal classification: A semi-supervised learning framework," *China Commun.*, vol. 17, no. 10, pp. 157–169, Oct. 2020.

[8] T. S. Rappaport et al., "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78729–78757, 2019.

[9] M. Yang et al., "Machine-learning-based scenario identification using channel characteristics in intelligent vehicular communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3961–3974, Jul. 2021.

[10] H. Cui et al., "Space–air–ground integrated network (SAGIN) for 6G: Requirements, architecture and challenges," *China Commun.*, vol. 19, no. 2, pp. 90–108, Feb. 2022.

[11] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.

[12] T. Zhou, Y. Wang, C.-X. Wang, S. Salous, L. Liu, and C. Tao, "Multi-feature fusion based recognition and relevance analysis of propagation scenes for high-speed railway channels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8107–8118, Aug. 2020.

[13] K. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[14] P. Zhang, L. Li, K. Niu, Y. Li, G. Lu, and Z. Wang, "An intelligent wireless transmission toward 6G," *Intell. Converg. Netw.*, vol. 2, no. 3, pp. 244–257, Sep. 2021.

[15] J. Zhang, L. Liu, Y. Fan, L. Zhuang, T. Zhou, and Z. Piao, "Wireless channel propagation scenarios identification: A perspective of machine learning," *IEEE Access*, vol. 8, pp. 47797–47806, 2020.

[16] 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.

[17] S. Wu, C.-X. Wang, E.-H. Aggoune, M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1207–1218, Jun. 2014.

[18] M. Jian et al., "Reconfigurable intelligent surfaces for wireless communications: Overview of hardware designs, channel models, and estimation techniques," *Intell. Converg. Netw.*, vol. 3, no. 1, pp. 1–32, Mar. 2022.

[19] C.-X. Wang, A. Ghazal, B. Ai, Y. Liu, and P. Fan, "Channel measurements and models for high-speed train communication systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 974–987, 2nd Quart., 2015.

[20] Y. Yuan, C.-X. Wang, X. Cheng, B. Ai, and D. I. Laurenson, "Novel 3D geometry-based stochastic models for non-isotropic MIMO vehicle-to-vehicle channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 298– 309, Jan. 2014.

[21] J. Bian et al., "A 3D wideband non-stationary multi-mobility model for vehicle-to-vehicle MIMO channels," *IEEE Access*, vol. 7, pp. 32562–32577, 2019.

[22] K. Guan et al., "Channel sounding and ray tracing for train-to-train communications at the THz band," in *Proc. EuCAP*, Krakow, Poland, Apr. 2019, pp. 1–5.

[23] *Study on Channel Model for Frequencies From 0.5 to 100 GHz*, 3GPP, Sophia Antipolis, France, Rep. TR 38.901, Jun. 2018.

[24] D. Baum et al., "Final report on link level and system level channel models," CTH, CTTC, PUT, SM, UU/CTH, document IST-2003-507581, Nov. 2005. [Online]. Available: http://www.ist-winner.org/deliverables.html

[25] P. Kyösti et al., "WINNER II channel models," EBITG, TUI, UOULU, CU/CRC, NOKIA, Oulu, Finland, document IST-4-027756, Apr. 2008. [Online]. Available: http://www.ist-winner.org/deliverables.html

[26] J. Meinila et al., "WINNER+ final channel models," Int. Telecommun. Union Radio, Geneva, Switzerland, document CELTIC/CP5-026 D5.3, Jun. 2010. [Online]. Available: http://projects.celtic-initiative.org/winner+/

[27] P. Popovski et al., "Scenarios, requirements and KPIs for 5G mobile and wireless system," METIS, Boston, MA, USA, document ICT-317669-METIS/D1.1, Apr. 2013. [Online]. Available: https://www.metis2020.com/documents/deliverables.html

[28] A. Maltsev et al., "Channel modeling and characterization," MiWEBA, Bayern, Germany, document MiWEBA D5.1 v1.0, Jun. 2014. [Online]. Available: http://www.miweba.eu/data/MiWEBA_D5.1_v1.01.pdf

[29] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[30] A. Costa and S. Hayking, *Multiple-Input Multiple-Output Channel Models: Theory and Practice*. Hoboken, NJ, USA: Wiley, 2001.

[31] L. Correia, *Mobile Broadband Multimedia Networks*. New York, NY, USA: Academic, 2006.

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

[33] R. Verdone and A. Zanella, Eds., *Pervasive Mobile and Ambient Wireless Communications: COST Action 2100*. London, UK: Springer, 2012.

[34] *IMT Vision-Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond*, document ITU-R standard M.2412-0, ITU, Geneva, Switzerland, Oct. 2017.

[35] 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.

[36] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[37] 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.

[38] J. Wang et al., "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.

[39] C. Sun, X. Q. Gao, S. Jin, Z. Ding, and C. Xiao, "Beam division multiple access transmission for massive MIMO communications," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2170–2184, Jun. 2015.

[40] L. You, X.-Q. Gao, G. Li, X.-G. Xia, and N. Ma, "BDMA for millimeter-wave/Terahertz massive MIMO transmission with per-beam synchronization," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1550– 1563, Jul. 2017.

[41] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[42] H. Özcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of 'kronecker' MIMO radio channel model," *Electron. Lett.*, vol. 39, no. 16, pp. 1209–1210, Aug. 2003.

[43] W. Weichselberger, H. Özcelik, M. Herdin, and E. Bonek, "A novel stochastic MIMO channel model and its physical interpretation," in *Proc. WPMC*, Yokosuka, Japan, Oct. 2003, pp. 1–6.

[44] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation at both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[45] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation, " *IEEE J. Sel. Areas Commun.*, vol. SAC-5, no. 2, pp. 128–137, Feb. 1987.

[46] S. Chen, S. Sun, and S. Kang, "System integration of terrestrial mobile communication and satellite communication—The trends, challenges and key technologies in B5G and 6G," *China Commun.*, vol. 17, no. 12, pp. 156–171, Dec. 2020.

[47] X. Lin, S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, "5G from space: An overview of 3GPP non-terrestrial networks," *IEEE Commun. Stand. Mag.*, vol. 5, no. 4, pp. 147–153, Dec. 2021.

[48] M. Domingo, "Overview of channel models for underwater wireless communication networks," *Phys. Commun.*, vol. 1, no. 3, pp. 163–182, Sep. 2008.

[49] F. Wang, Z. Xu, R. Zhi, J. Chen, and P. Zhang, "LOS/NLOS channel identification technology based on CNN," in *Proc. NICS*, Hanoi, Vietnam, Dec. 2019, pp. 200–203.

[50] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "A machine learning approach for the classification of indoor environments using RF signatures," in *Proc. GlobalSIP*, Nov. 2018, pp. 1060–1062.

[51] D. Scazzoli, M. Magarini, L. Reggiani, Y. L. Moullec, and M. M. Alam, "A deep learning approach for LoS/NLoS identification via PRACH in UAV-assisted public safety networks," in *Proc. PIMRC*, London, U.K., Sep. 2020, pp. 1–6.

[52] V. B. Vales, T. Domínguez-Bolaño, C. J. Escudero, and J. A. García-Naya, "Using the power delay profile to accelerate the training of neural network-based classifiers for the identification of LOS and NLOS UWB propagation conditions," *IEEE Access*, vol. 8, pp. 220205–220214, 2020.

[53] Z. Zeng, S. Liu, and L. Wang, "NLOS identification for UWB based on channel impulse response," in *Proc. ICSPCS*, Dec. 2018, pp. 1–6.

[54] C. Huang, A. F. Molisch, R. He, R. Wang, P. Tang, and Z. Zhong, "Machine-learning-cased data processing techniques for vehicle-to-vehicle channel modeling," *IEEE Commun. Mag.*, vol. 57, no. 11, pp. 109–115, Nov. 2019.

[55] C. Huang et al., "Machine learning-enabled LOS/NLOS identification for MIMO systems in dynamic environments," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3643–3657, Jun. 2020.

[56] T. Chang, S. Jiang, Y. Sun, A. Jia, and W. Wang, "Multi-bandwidth NLOS identification based on deep learning method," in *Proc. EuCAP*, Dusseldorf, Germany, Mar. 2021, pp. 1–5.

[57] J. Choi, W. Lee, J. Lee, J. Lee, and S. Kim, "Deep learning based NLOS identification with commodity WLAN devices," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3295–3303, Apr. 2018.

[58] J.-L. Wang, Y.-R. Li, A. B. Adege, L.-C. Wang, S.-S. Jeng, and J.-Y. Chen, "Machine learning based rapid 3D channel modeling for UAV communication networks," in *Proc. IEEE CCNC*, Jan. 2019, pp. 1–5.

[59] O. Esrafilian and D. Gesbert, "3D city map reconstruction from UAV based radio measurements," in *Proc. IEEE GLOBECOM*, Singapore, Dec. 2017, pp. 1–6.

[60] J. Xie, W. Wang, and Z. Xu, "Identification of NLOS condition in different scenarios based on wireless propagation features," in *Proc. IEEE ICEICT*, Xi'an, China, Aug. 2021, pp. 568–571.

[61] T. Zeng, Y. Chang, Q. Zhang, M. Hu, and J. Li, "CNN-based LOS/NLOS identification in 3-D massive MIMO systems," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2491–2494, Dec. 2018.

[62] J. Li, Y. Chang, T. Zeng, and Y. Xiong, "Channel correlation based identification of LOS and NLOS in 3D massive MIMO systems," in *Proc. WCNC*, Marrakesh, Morocco, Apr. 2019, pp. 1–6.

[63] M. I. AlHajri, N. Alsindi, N. T. Ali, and R. M. Shubair, "Classification of indoor environments based on spatial correlation of RF channel fingerprints," in *Proc. APSURSI*, Jul. 2016, pp. 1447–1448.

[64] J. Huang et al., "A big data enabled channel model for 5G wireless communication systems," *IEEE Trans. Big Data*, vol. 6, no. 2, pp. 211–222, Jun. 2020.

[65] V. Gupta, V. K. Mishra, P. Singhal, and A. Kumar, "An overview of supervised machine learning algorithm," in *Proc. SMART*, Moradabad, India, Dec. 2022, pp. 87–92.

[66] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for IoT applications: A machine learning approach," *IEEE Antennas Wireless Propag. Lett.* , vol. 17, no. 12, pp. 2164–2168, Dec. 2018.

[67] J. B. Kristensen, M. M. Ginard, O. K. Jensen, and M. Shen, "Non-line-of-sight identification for UWB indoor positioning systems using support vector machines," in *Proc. IEEE IWS*, Guangzhou, China, 2019, pp. 1–3.

[68] A. Kirmaz, D. S. Michalopoulos, I. Balan, and W. Gerstacker, "LOS/NLOS classification using scenario-dependent unsupervised machine learning," in *Proc. IEEE PIMRC*, Helsinki, Finland, 2021, pp. 1134–1140.

[69] M. Ramadan, V. Sark, J. Gutierrez, and E. Grass, "NLOS identification for indoor localization using random forest algorithm," in *Proc. WSA*, Bochum, Germany, 2018, pp. 1–5.

[70] A. Olejniczak, O. Blaszkiewicz, K. K. Cwalina, P. Rajchowski, and J. Sadowski, "Deep learning approach for LOS and NLOS identification in the indoor environment," in *Proc. URSI*, Warsaw, Poland, Oct. 2020, pp. 104–107.

[71] X. Bai, L. Zhang, T. Yang, and Z. Hu, "Semi-supervised learning based acoustic NLOS identification for smartphone indoor positioning," in *Proc. IEEE ICSPCC*, Dalian, China, 2019, pp. 1–6.

[72] M. Zhao, Y. Yu, C. Yi, H. Wang, and S. Gao, "Machine-learning-assisted scenario classification using large-scale fading characteristics and geographic information," in *Proc. IEEE ICCC*, Sanshui, China, 2022, pp. 43–48.

[73] K. Abdullah et al., "A machine learning-based technique for the classification of indoor/outdoor cellular network clients," in *Proc. IEEE CCNC*, Las Vegas, NV, USA, 2020, pp. 1–2.

[74] H. Li, Y. Li, S. Zhou, and J. Wang, "Wireless channel feature extraction via GMM and CNN in the tomographic channel model," *J. Commun. Inf. Netw.*, vol. 2, no. 1, pp. 41–51, Mar. 2017.

[75] T. Ponn, M. Breitfuß, X. Yu, and F. Diermeyer, "Identification of challenging highway-scenarios for the safety validation of automated vehicles based on real driving data," in *Proc. EVER*, Sep. 2020, pp. 1–10.

[76] J. Ding, H. Wei, Z. Zhong, and B. Ai, "Study on the propagation scenario classification of the high-speed railway GSM-R system based on GIS," in *Proc. ICWMMN*, Beijing, China, Nov. 2013, pp. 306–310.

[77] S. Fan, Y. Li, M. Zhang, D. Feng, Q. Chen, and Y. Jiang, "An anomaly detection and scenario classification scheme based on fuzzy C-means clustering," in *Proc. CAC*, Shanghai, China, Nov. 2020, pp. 5223–5228.

[78] D. Li, X. Chen, M. Becchi, and Z. Zong, "Evaluating the energy efficiency of deep convolutional neural networks on CPUs and GPUs," in *Proc. IEEE BSS*, Oct. 2016, pp. 477–484.

[79] S. Zheng and C. Ding, "A group lasso based sparse KNN classifier," *Pattern Recognit. Lett.*, vol. 131, pp. 227–233, Mar. 2020.

[80] 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, Jun. 2022.

[81] S. Geng, W. Hu, H. Ding, Z. Qian, and X. Zhao, "Research on urban path loss model by machine learning based on 3D electronic maps and channe measurements," *J. Electron. Inf. Technol.*, vol. 44, pp. 1–8, Nov. 2021.

[82] R. Saini and S. Rawat, "Land use land cover classification in remote sensing using machine learning techniques," in *Proc. IHCSP*, 2023, pp. 99–104.

[83] H. Du, M. Li, Y. Xu, and C. Zhou, "An ensemble learning approach for land use/land cover classification of arid regions for climate simulation: A case study of Xinjiang, Northwest China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 2413–2426, 2023.

[84] R. Sharma, D. Pantola, S. D. Kalony, and R. Agarwal, "Analysis machine learning approach and model on hyper spectral (Sentinel-2) images for land cover classification: Using SVM," in *Proc SMART*, 2021, pp. 680–684.

[85] C. Djouvas, I. Despotis, and C. Christodoulou, "Automating road junction identification using crowdsourcing and machine learning on GPS transformed data," in *Proc. SMAP*, Corfu, Greece, 2021, pp. 1–6.

[86] R. Feng, C.-X. Wang, J. Huang, X. Gao, S. Salous, and H. Hass, "Classification and comparison of massive MIMO propagation channel models," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23452–23471, Dec. 2022.

[87] 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, Jun. 2022.

[88] S. Aldossari and K.-C. Chen, "Machine learning for wireless communication channel modeling: An overview," *Wireless Pers. Commun.*, vol. 106, no. 1, pp. 41–70, Mar. 2019.

[89] 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 VTC Spring*, Helsinki, Finland, Jun. 2022, p. 5.

[90] P. Zhang, C. Yi, and H. Wang, "Machine-learning-assisted modeling of millimeter-wave channels," in *Proc. APS/URSI*, Singapore, Dec. 2021, pp. 233–234.

[91] P. Zhang, C. Yi, B. Yang, H. Wang, C. Oestges, and X. You, "Predictive modeling of millimeter-wave vegetation-scattering effect using hybrid physics-based and data-driven approach," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4056–4068, Jun. 2022.

[92] S. Aldossari and K.-C. Chen, "Predicting the path loss of wireless channel models using machine learning techniques in mmWave urban communications," in *Proc. WPMC*, Lisbon, Portugal, Nov. 2019, pp. 1–6.

[93] J.-Y. Lee, M. Kang, and S.-C. Kim, "Path loss exponent prediction for outdoor millimeter wave channels through deep learning," in *Proc. IEEE WCNC*, Marrakesh, Morocco, Apr. 2019, pp. 1–5.

[94] A. Bharti, R. Adeogun, and T. Pedersen, "Learning parameters of stochastic radio channel models from summaries," *IEEE Open J. Antennas Propag.*, vol. 1, pp. 175–188, 2020.

[95] W. Hu, S. Geng, and X. Zhao, "Mm-wave 60 GHz channel fading effects analysis based on RBF neural network," in *Proc. ICCCS*, Shanghai, China, May 2020, pp. 589–593.

[96] Y. Wen, W. Hu, S. Geng, and X. Zhao, "Machine learning based mm-wave 60 GHz channel modeling for 5G wireless communication systems," in *Proc. IEEE/CIC ICCC*, Chengdu, China, Dec. 2019, pp. 1005–1010.

[97] B. Turan and S. Coleri, "Machine learning based channel modeling for vehicular visible light communication," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 9659–9672, Oct. 2021.

[98] P. Ramya, M. Boban, C. Zhou, and S. Stanczak, "Using learning methods for V2V path loss prediction," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Marrakesh, Morocco, Apr. 2019, pp. 1–6.

[99] Y. Aksan, M. Yuksel, and A. Ozbayoglu, "Channel characterization for aircraft integrated antennas via machine learning," in *Proc. IEEE WCNC*, Nanjing, China, Mar. 2021, pp. 1–6.

[100] G. Yang, Y. Zhang, Z. He, J. Wen, Z. Ji, and Y. Li, "Machine learning-based prediction methods for path loss and delay spread in air-to-ground millimetre-wave channels," *IET Microw. Antennas Propag.*, vol. 13, no. 8, pp. 1113–1121, Apr. 2019.

[101] C. Wu, X. Yi, Y. Zhu, W. Wang, L. You, and X. Gao, "Channel prediction in high-mobility massive MIMO: From spatio-temporal autoregression to deep learning," *IEEE J. Sel. Areas Commun.*, hboxvol. 39, no. 7, pp. 1915–1930, Jul. 2021.

[102] X. Zhao et al., "Neural network and GBSM based time-varying and stochastic channel modeling for 5G millimeter wave communications," *China Commun.*, vol. 16, no. 6, pp. 80–90, Jun. 2019.

[103] 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, Sep. 2020.

[104] B. Gschwendtner and F. Landstorfer, "An application of neural networks to the prediction of terrestrial wave propagation," in *Proc. ICAP*, Edinburgh, U.K., Mar. 1993, pp. 804–807.

[105] T. Balandier, A. Caminada, V. Lemoine, and F. Alexandre, "170 MHz field strength prediction in urban environment using neural nets," in *Proc. PIMRC*, Sep. 1995, pp. 120–124.

[106] N. Zaarour, S. Affes, N. Kandil, and N. Hakem, "Comparative study on a 60 GHz path loss channel modeling in a mine environment using neural networks," in *Proc. IEEE ICUWB*, Oct. 2015, pp. 1–4.

[107] N. Zaarour, N. Kandil, and N. Hakem, "An accurate neural network approach in modeling an UWB channel in an underground mine," in *Proc. IEEE APS-URSI*, Jul. 2013, pp. 1608–1609.

[108] 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.
[109] T. Zhang, S. Liu, W. Xiang, L. Xu, K. Qin, and X. Yan, "A real-time channel prediction model based on neural networks for dedicated short range communications," *Sensors*, vol. 19, no. 16, pp. 35–47, Aug. 2019.
[110] C. Huang, A. F. Molisch, Y. 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.
[111] J. Zhang, "The interdisciplinary research of big data and wireless channel: A cluster-nuclei based channel model," *China Commun.*, vol. 13, no. S2, pp. 14–26, 2016.
[112] R. He et al., "Clustering enabled wireless channel modeling using big data algorithms," *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 177–183, May 2018.
[113] R. Adeogun, "Calibration of stochastic radio propagation models using machine learning," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 12, pp. 2538–2542, Dec. 2019.
[114] A. Al-Amodi, M. Masood, and M. Khan, "Underwater wireless optical communication channel characterization using machine learning techniques," in *Proc. IEEE OGC*, Shenzhen, China, Dec. 2022, pp. 50–54.
[115] J. Bian, C.-X. Wang, X. Q. Gao, X.-H. You, and M. Zhang, "A general 3D non-stationary wireless channel model for 5G and beyond," *IEEE Trans. Wireless Commun.*, vol. 39, no. 4, pp. 3211–3224, May 2021.
[116] C.-X. Wang, Z. Lv, X. Gao, X.-H. You, Y. Hao, and H. Haas, "Pervasive 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, Sep. 2022.
[117] Z. Li, C.-X. Wang, J. Huang, and Y. Zheng, "A non-stationary GBSM for 6G LEO satellite communication systems," in *Proc. ICCC*, Xiamen, China, Jul. 2021, pp. 1–6.
[118] H. Chang et al., "A novel nonstationary 6G UAV-to-ground wireless channel model with 3-D arbitrary trajectory changes," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9865–9877, Jun. 2021.
[119] Y. Li, C.-X. Wang, and Y. Liu, "A 3D non-stationary geometry-based stochastic model for industrial automation wireless communication systems," in *Proc. PIMRC*, Helsinki, Finland, Sep. 2021, pp. 1–6.
[120] Y. Xu, K. Yu, L. Li, X. Lei, L. Hao, and C.-X. Wang, "A 3D non-stationary mmWave channel model for vacuum tube ultra-high-speed train wireless communication systems," in *Proc. IEEE WCNC*, Nanjing, China, Apr. 2021, pp. 1–6.
[121] Y. He et al., "A novel 3D non-stationary maritime wireless channel model," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 2102–2116, Mar. 2022.
[122] Y. Wang et al., "Measurement-based analysis and modeling of channel characteristics in an industrial scenario at 28 GHz," in *Proc. IEEE VTC Fall*, Sep. 2021, pp. 1–5.
[123] L. Gao, Z. Zhong, B. Ai, L. Xiong, and H. Zhang, "Analysis and emulation of the small-scale fading characteristics in the high-speed rail scenarios," in *Proc. ICST*, Harbin, China, Aug. 2011, pp. 1181– 1185.
[124] Y. Zeng and X. Xu, "Toward environment-aware 6G communications via channel knowledge map," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 84–91, Jun. 2021.
[125] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Krner, "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., 2018.
[126] Z. Yun and M. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015.

Wenqi Zhou received the B.Sc. degree in electrical engineering and automation from Shandong Jianzhu University, Jinan, China, in 2001, and the M.Eng. degree in control theory and control engineering from Shandong University, Jinan, in 2006. He is currently pursuing the Ph.D. degree with Southeast University, Nanjing, China.

His research interests include 6G wireless communications and AI-based predictive channel modeling.

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

He was a Research Assistant with 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 a Professor with Southeast University, Nanjing, China, since 2018, where he is currently the Executive Dean of the School of Information Science and Engineering. He is also a Professor with the Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing. He has authored four books, three book chapters, and over 520 papers in refereed journals and conference proceedings, including 27 highly cited papers. He has also delivered 24 invited keynote speeches/talks and 16 tutorials in international conferences. His current research interests include wireless channel measurements and modeling, 6G wireless communication networks, and electromagnetic information theory.

Dr. Wang received 15 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, and IEEE/CIC ICCC 2022. He is a Highly Cited Researcher recognized by Clarivate Analytics from 2017 to 2020. He is currently an Executive Editorial Committee Member of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was an IEEE Communications Society Distinguished Lecturer in 2019 and 2020. He has served as an Editor for over ten international journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2007 to 2009, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, from 2011 to 2017, and the IEEE TRANSACTIONS ON COMMUNICATIONS, from 2015 to 2017. He was a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on *Vehicular Communications and Networks* (Lead Guest Editor), Special Issue on Spectrum and Energy Efficient Design of Wireless Communication Networks, and Special Issue on Airborne Communication Networks. He was also a Guest Editor for the IEEE TRANSACTIONS ON BIG DATA, Special Issue on Wireless Big Data, and is a Guest Editor for the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, Special Issue on Intelligent Resource Management for 5G and Beyond. He has served as a TPC Member, a TPC Chair, and a General Chair for more than 30 international conferences. He is a member of the Academia Europaea (The Academy of Europe), a Member of the European Academy of Sciences and Arts, a Fellow of the Royal Society of Edinburgh, IET, and China Institute of Communications.

Chen Huang (Member, IEEE) received the Ph.D. degrees from Beijing Jiaotong University, Beijing, China, in 2021.

From April 2021 to April 2023, He was a Postdoctoral Research Associate with Pervasive Communication Research Center, Purple Mountain Laboratories (PML), Nanjing, China, and also with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University (SEU), Nanjing. From 2018 to 2020, he has been a Visiting Scholar with the University of Southern California, Los Angeles, CA, USA, and with the Université Catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium. Since April 2023, he is an Research Assistant Professor with Pervasive Communication Research Center, PML, and an Extramural Supervisor with National Mobile Communications Research Laboratory, School of Information Science and Engineering, SEU. He has authored/coauthored one book chapters, more than 40 journal and conference papers, as well as several patents. His research interests include channel characterization, clustering and tracking, channel modeling, and applications of machine learning-based techniques on propagation channel modeling.

Dr. Huang is selected in the Young Elite Scientists Sponsorship Program by China Association for Science and Technology and the Outstanding Postdoctoral Fellow Program in Jiangsu, received two times the Best Paper Award at international conferences, and serves as the Technical Program Committee Member for several conferences, including GlobeCom, ICC, VTC-Fall, and VTC-Spring.

Zheao Li received the B.E. degree in information engineering from Southeast University, Nanjing, China, in 2020, where he is currently pursuing the M.Sc. degree with National Mobile Communications Research Laboratory, School of Information Science and Engineering.

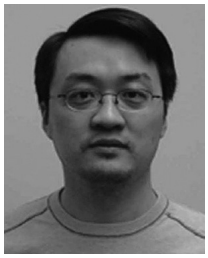
He is also an Intern with Purple Mountain Laboratories, Nanjing. His research interest is ML-based predictive channel modeling.

Zhongyu Qian received the B.E. degree in communication engineering from Dalian University, Dalian, China, in 2019. He is currently pursuing the M.Sc. degree with National Mobile Communications Research Laboratory, Southeast University-Monash University Joint Graduate School, Southeast University, Nanjing, China.

He is also an Intern with Purple Mountain Laboratories, Nanjing. His research interest is ML-based predictive channel modeling.

Zhen Lv (Member, IEEE) received the B.Sc. and M.Sc. degrees in information and communication engineering from Shandong University, Jinan, China, in 2015 and 2018, respectively.

She is currently a Wireless Channel Engineer with Purple Mountain Laboratories, Nanjing, China. Her research interests include 6G channel modeling and wireless big data.



Yunfei Chen (Senior Member, IEEE) received the B.E. and M.E. degrees in electronics engineering from Shanghai Jiao Tong University, Shanghai, China, in 1998 and 2001, respectively, and the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 2006.

He is currently working as an Associate Professor with The University of Warwick, Coventry, U.K. His research interests include wireless communications, cognitive radios, wireless relaying, and energy harvesting.