The Current Progress and Future Prospects of Path Loss Model for Terrestrial Radio Propagation

Jian Wang¹,²,³,†, Yulong Hao¹,† and Cheng Yang¹,²,*

¹ School of Microelectronics, Tianjin University, Tianjin 300072, China; wangjian16@tju.edu.cn (J.W.); haoyulong98@tju.edu.cn (Y.H.)
² Qingdao Institute for Ocean Technology, Tianjin University, Qingdao 266200, China
³ Shandong Engineering Technology Research Center of Ocean Information Awareness and Transmission, Qingdao 266200, China
* Correspondence: ych2041@tju.edu.cn
† These authors contributed equally to this work.

Abstract: The radio channel model is a major factor supporting the whole life cycle of the terrestrial radio system, including the demonstration, design, validation, operation, and so on. To improve the spectrum sharing and spectral efficiency in terrestrial radio services, we analyze three types of path loss models in detail: deterministic, empirical, and semi-empirical models, to meet the requirements of path loss modeling for supporting traditional band expansion and reuse. Then, we conduct a comparative analysis based on the characteristics of the current models. Furthermore, a preview of the future terrestrial path loss modeling methods is provided, including intelligent modeling processes and multi-model hybridization methods. Finally, we look forward to the potential technology that can be used in future wireless communication, such as terahertz communication, reconfigurable intelligent surface technology, and integrated communication and sensing technology. The above research can provide a reference for the development of terrestrial radio channel modeling, promoting the technologies of terrestrial channel modeling. We hope this paper will stimulate more interest in modeling terrestrial radio channels.

Keywords: radio propagation; path loss; model; current progress; future prospect

1. Introduction

As the elaboration of transmission effects between the transmitter and receiver of the radio system, the radio channel is subject to many factors [1]. For instance, terrain and surface features, meteorological parameters of the troposphere and ionosphere, solar activity, etc. [2–4]. The characteristic of the channel determines the performance of the electromagnetic information system, which is also very significant in the whole life cycle, such as design, development, production, validation, operation, and maintenance [5]. The electromagnetic spectrum is the strategic resource that all radio-electronic warfare relies on and is a hub of cross-domain joint operations linking land, sea, air, and space networks. Therefore, the security of spectrum management is directly related to national security and radio system construction [6]. The rapid growth of radio systems deployment leads to a shortage of electromagnetic spectrum resources. As a result, the possibility of radio interference will increase, and the performance of radio systems will be reduced. The radio channel is strongly related to the electromagnetic spectrum [7]. To avoid the frequency utilization problem caused by the congestion of the electromagnetic spectrum, one of the most effective and frequently used methods is the utilization of channel models in frequency planning, assignment, coordination, and so on. New channel models are proposed to increase the accuracy and robustness at the frequency bands in use, reduce the mutual interference between radio systems, and supply further support for expanding and reusing the existing spectrum.
According to the modeling principle, the existing models can be categorized as deterministic, empirical, and semi-empirical. The deterministic models are derived from the electromagnetic propagation formula, and Ray Tracing (RT) \cite{8} and Parabolic Equations (PE) \cite{9} are the typical deterministic models. The empirical models (also called the statistical channel model) are based on statistical theory: the Okumura model \cite{10}, the Okumura–Hata model \cite{11}, the ITU-R P.1546 (below referred to as P.1546) model \cite{12}, and the COST-231 model \cite{13} are the typical empirical models. The semi-empirical models mix the features of deterministic and empirical channel models. The COST-259 model \cite{14} and the International Mobile Telecommunications (IMT)-Advanced model \cite{15} are typical semi-empirical models. Traditional modeling methods used the manual statistical analysis of measured data under specific scenarios and radio wave propagation conditions. In addition, the development of computer and artificial intelligence, in which machine learning (ML) is an important branch, provides an accurate and efficient modeling method with self-learning and self-adaptive capabilities. ML is a predicted and classified method that can mine hidden rules from a large number of data \cite{16}. Theoretically, many problems in radio channel modeling can be considered regression, clustering, and classification in ML. The efficiency of radio channel modeling is consistent with the capabilities of ML. Therefore, we can intelligently structure the channel model through ML to learn the propagation characteristics and hidden rules with measured data \cite{17,18}. In \cite{19}, using Artificial Neural Networks (ANNs), the authors developed a new method for multiband heterogeneous radio network scenarios. To predict path loss in the global mobile communication system band, Eichie et al. developed a Multilayer Perceptron (MLP) neural model \cite{20}. Sotiroudis et al. proposed an urban environment model, and the research results show that as long as the size of an ANN is correctly selected, the path loss model based on the ANN will operate efficiently \cite{21}. Wang et al. proposed a new method based on Backpropagation (BP) ANN for radio wave propagation prediction. Thrane et al. proposed a comprehensive model based on deep learning, supplemented by a random and RT \cite{22}.

Therefore, this paper aims to put forward the research prospect of future radio channel methods based on summarizing the research results of terrestrial radio channel models and combining the challenges faced in developing radio communication. For the above purposes, the following arrangement is made: Section 2 overviews the category of existing models based on the modeling methods, including the deterministic, empirical, and semi-empirical models. At the same time, we compare and summarize the features of the three modeling methods. Then, the outlook of the future challenge and modeling methods with the development trend of new techniques, applications, and scenarios is given in Section 3. Finally, we summarize the development of terrestrial channel modeling.

2. Current Progress of Channel Model

Ultra-short wave, microwave, and millimeter wave are the primary communication frequency bands for terrestrial radio communication. The propagation of these frequency bands works together through various propagation mechanisms. Terrestrial radio communication modes include single-input single-output (SISO) and multi-input multi-output (MIMO); MIMO includes single-input multi-output (SIMO), multi-input single output (MISO), and multi-input multi-output (MIMO). The communication scenes cover large-scale and small-scale scenes. Among them, the large-scale scenarios include land, sea, and mixed scenes, and the small-scale scenarios include indoor and street scenes. Figure 1 shows the knowledge graph of terrestrial wireless communication.
It should be noted that there are three types of radio channel models based on different modeling principles: the deterministic model, the empirical model, and the semi-empirical model, respectively. The development history is shown in Figure 2. From the figure, we can observe that models proposed based on empirical methods are the most abundant, while those derived from deterministic methods are the least prevalent. The majority of models proposed in the early and middle stages are based on empirical methods, with recent years witnessing a surge in models based on semi-empirical methods. Deterministic models still in use today include ray tracing and parabolic equation methods. It can be inferred that semi-empirical modeling methods will be the predominant approach in future modeling. Empirical methods, given their simplicity and acceptable accuracy, are likely to continue being widely employed in small-area modeling or localized adjustments to models in the future.

Figure 1. Knowledge graph of terrestrial wireless communication channels.

Next, we will analyze and compare the three modeling methods according to their typical channel models. In addition, the current status of channel models, which are mostly based on the references published in the IEEE, IET, and Hindawi between 2012 and 2022, are also described. To conclude, each modeling method will be discussed in terms of its advantages and disadvantages.
2.1. Deterministic Model

2.1.1. The Typical Model

The deterministic channel modeling method, which analyzes the radio channel propagation characteristics by solving Maxwell equations or using geometric approximation numerical methods, is derived from the radio wave propagation mechanism [23]. Figure 3 shows the typical deterministic models.

RT is the most widely used deterministic channel model method [24–26]. According to the modeling principle, RT can be classified into forward RT, backward RT, and mixed RT. Shooting and Bouncing Ray (SBR), first proposed in 1986, is the widely used forward RT [27]. Forward RT traces every ray from transmitting to receiving equipment after the initial ray is emitted. The RT process of the current ray is terminated when the ray has experienced a number of reflections and diffraction, or the ray hits the receiving model. However, the computation of forward RT will increase with time. Because of the simple principle, this method is widely used in various fields to develop RT acceleration technology and hardware computing efficiency. The core logic of backward RT and forward RT is entirely different. Backward RT is used to find the propagation path of direct, reflection, diffraction, or other combined forms by using the mirror method to traverse the basic units of triangles or other structures in the scene when the locations of transmitters and receivers are known [28]. The principle of backward RT is relatively simple, but it needs the locations of transmitters and receivers to be known. Mixed RT was first proposed to predict street radio wave propagation [29]. It is an RT method that combines the characteristics of forward RT and backward RT. The method first uses forward RT to obtain the approximate propagation path and then uses backward RT to modify and verify the obtained propagation path further.
In 1946, Lenontovich and Fock proposed PE to solve the problem of radio wave diffraction. In 1977, Tappert solved PE in acoustic waves by using the Split-Step Fourier transform (SSFT) method, which attracted people's attention and research [30]. In 1991, Kuttler and Dockery systematically deduced the PE of radio wave propagation in the troposphere for the first time. Currently, the leading solution methods of PE are the SSFT and Finite Difference (FD) methods, both of which are step algorithms. The step size is almost not limited by the wavelength of the radio wave when using SSFT to solve PE. Therefore, the result can be obtained quickly when solving large-scale problems. However, it is not easy to deal with complex boundary conditions. The calculation process of the FD method involves many matrix operations. It has significant advantages in dealing with complex boundary conditions. However, its step size is limited by the radio wave wavelength, so it is unsuitable for solving large-scale problems. It is mainly used to calculate the electromagnetic scattering characteristics of targets and radio propagation prediction in small-scale areas.

The electromagnetic calculation methods are mainly used to solve the radio wave propagation problem by calculating the Maxwell equations strictly [31]. However, these methods have not been used on a large scale due to the shortcomings of a large amount of computation and high requirements for hardware platforms.

2.1.2. Other Models

Figure 4 shows the principle and the scenario of the deterministic models in the references. Afsharinejad et al. proposed a log-distance path loss model based on Monte Carlo simulations. Shadow fading of air and leaves is considered in the model, and path loss variations are defined. Since the latter model only considers transmission distance, it can be interpreted as a simplified version of the theoretical model [32]. In [33], the authors proposed a theoretical model based on the Fresnel zone theory to solve the very near-ground path loss problem. In [34], a model based on Green’s theorem is proposed. The model is for the blending scenario. In addition, the authors propose channel models based on the standard models and propagation theories. For example, a ground reflection
model based on double rays, along with other propagation mechanisms, is proposed in [35]. In [36], they use the RT and edge peak diffraction theory to propose a model.

From the above references, it is evident that deterministic models are primarily applicable to small-scale scenarios within a prediction range of 200 m. This limitation arises because achieving precise predictions with deterministic models requires a substantial amount of environmental parameters and computational resources [37–39]. The expansion of the prediction range imposes higher demands on data collection and computing power, which is a key reason why deterministic methods are not as widely employed as the other two methods.

2.2. Empirical Model

The empirical modeling method is a channel parameters modeling method that uses mathematical statistics. The empirical model uses standardized propagation characteristics, testing cumulative data or statistical curves to interpolate or extrapolate according to equipment frequency parameters, and completes prediction analysis through a series of corrections [40]. Table 1 shows the typical empirical model parameter information.

<table>
<thead>
<tr>
<th>Model</th>
<th>Scenarios</th>
<th>Distance/km</th>
<th>Frequency/MHz</th>
<th>Transmitting Height/m</th>
<th>Receiving Height/m</th>
<th>Author</th>
<th>Proposed Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Okumura</td>
<td>Quasi-smooth urban area</td>
<td>1–20</td>
<td>150–1500</td>
<td>30–200</td>
<td>1–10</td>
<td>Okumura et al.</td>
<td>1962</td>
</tr>
<tr>
<td>Carey</td>
<td>Flat ground</td>
<td>&lt;130</td>
<td>35–460</td>
<td>30–1500</td>
<td>Average 1.8</td>
<td>FCC</td>
<td>1964</td>
</tr>
<tr>
<td>Longley-Rice</td>
<td>Ground, sea</td>
<td>1–2000</td>
<td>20–40,000</td>
<td>1–1500</td>
<td>1–9</td>
<td>Longley and Rice</td>
<td>1968</td>
</tr>
<tr>
<td>Okumura-Hata</td>
<td>Urban, suburb, rural</td>
<td>1–20</td>
<td>150–1500</td>
<td>30–200</td>
<td>1–10</td>
<td>Hata et al.</td>
<td>1980</td>
</tr>
<tr>
<td>COST-231-WI</td>
<td>Urban, dense urban</td>
<td>0.02–5</td>
<td>800–2000</td>
<td>4–50</td>
<td>1–3</td>
<td>EURO-COST</td>
<td>1991</td>
</tr>
<tr>
<td>ITU-R P1546</td>
<td>Multi-scenario</td>
<td>1–1000</td>
<td>30–4000</td>
<td>&lt;3000</td>
<td>Average 10</td>
<td>ITU</td>
<td>2001</td>
</tr>
</tbody>
</table>

2.2.1. The Typical Model

The Okumura model is one of the earliest empirical models based on field intensity measurement. Okumura was committed to studying the relationship between power density and distance in Tokyo [10]. Because the model is entirely based on statistics, its
result is a fitting curve rather than a specific formula. The Longley–Rice and Okumura models were proposed at the same time. The Longley–Rice model already has a general computer program to calculate path loss. For a given transmission, the computer calculates path loss according to frequency, path distance, polarization direction, antenna height, surface diffraction, ground radius, and ground conductivity [41]. In order to extend Okumura’s research result to areas other than Tokyo, Hata proposed the Okumura–Hata model by formulating Okumura’s fitting curve. The Okumura–Hata model is very suitable for large-scale scenes where the transmitter is higher than the receivers and surrounding buildings. The arguments of the LEE model are easily available, so it is very popular. The LEE model can be broken down into the macro and microcell sub-models. The macrocell sub-model firstly considers terrain as flat, only considers the impact of buildings, and after this, adds the effects of terrain and landform. The microcell model assumes that the attenuation of signals is highly correlated with the length of buildings on the propagation path. The LEE model also divides terrain and geomorphology influence into three situations for calculation: non-blocking situation, blocking situation, and water surface reflection situation [42]. COST-231 models were proposed by EURO-COST in 1991. Compared with the Okumura–Hata model, COST-231 models are more often used. The Okumura–Hata model is the basis of the COST-231–Hata model. The range of its frequency is extended from 1500 MHz to 2000 MHz. The COST-231–WI model combines the result of the urban environment calculated by the Walfisch–Bertoni model [43] and the corrected function of the street trend of the Ikegami model and the experimental correction. Therefore, this model is very suitable for predicting path loss of the street scene. However, this model also has some defects; when the transmitting antenna height changes slightly, the path loss jumps steeply.

Therefore, when using this model, the transmitter’s antenna should be installed at a height several meters higher than the street building. The Carely model was extended from the CCIR curve. It was included as one of the standard models by the Federal Communications Commission (FCC) of the US. This model provides the field strength–distance curve when the average height of the mobile station antenna is 1.8 m, the height of the base station antenna is 30–1500 m, and the coverage is less than 130 km on flat ground. Moreover, the ITU-R P.370 (below referred to as P.370) and P.1546 models proposed by ITU are also mainstream empirical models. The P.370 proposal is intended to forecast the radio wave propagation field strength of meter wave and decimeter wave bands under various climatic conditions for broadcast service planning engineers in various countries [44]. P.1546 is proposed based on the P.370 proposal. Although the prediction accuracy and application range are significantly improved compared with the P.370 model, both recommendations are revised based on data statistics according to empirical values, and the diffraction effect of obstacles is not strictly considered. Therefore, when diffraction is the main factor affecting propagation in mountainous areas, the prediction error of these two propagation models increases significantly [12].

Most empirical models are large-scale propagation and prediction models. The Rice and the Rayleigh models are small-scale propagation models. These two models apply to indoor or other object-intensive areas. The difference is that the Rice model applies to a situation where there is a direct path, and when there is no direct path, the Rayleigh model applies.

2.2.2. Other Models

The empirical methods can be divided into three main types. The first is based on traditional linear regression, the second is based on ML, and the last is the mixed approach. To evaluate the accuracy and improvement of these models, we introduced evaluating standards of the mean absolute error (MAE), mean relative error (MRE), root-mean-square error (RMSE), and relative root-mean-square error (RRMSE) to reflect the deviation from the measurements and the stability and robustness of the prediction model. The definitions and characteristics of these indexes are listed in Table 2.
Table 2. Equation and characteristics of the performance criteria.

<table>
<thead>
<tr>
<th>Index</th>
<th>Definition</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>(\frac{1}{n} \sum_{i=1}^{n}</td>
<td>y_i - y_i'</td>
</tr>
<tr>
<td>MRE</td>
<td>(\frac{1}{n} \sum_{i=1}^{n} \frac{</td>
<td>y_i - y_i'</td>
</tr>
<tr>
<td>RMSE</td>
<td>(\sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y_i')^2})</td>
<td>Evaluate the root-mean-square error between the predicted value and the measured value.</td>
</tr>
<tr>
<td>RRMSE</td>
<td>(\sqrt{\frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - y_i')^2}{y_i}})</td>
<td>Evaluate the relative root-mean-square error between the predicted value and measured value.</td>
</tr>
</tbody>
</table>

Figure 5 shows the principle and the scenario of the empirical models in the references. From Figure 5, we can see that almost half of the studies are based on traditional linear regression. On the one hand, some are directly modeled using linear regression based on measured data for street [45], bleeding [46–48], indoor [49–54], urban [55–57], vegetation shielding [58–60], vehicle-to-vehicle channels (V2V) [61], and station scenarios [62]. At the same time, some researchers have used linear regression to fit the standard models. For example, He et al. used linear fitting to assist the Hata model based on the measurement data for the GSM railway [63]. Nossire et al. used linear fitting to correct the Okumura model based on the measurements for an urban indoor scenario [64]. In [65], a standard Macrocell model was proposed with the rectification of measurements to predict the path loss of the blending scenario.

![Figure 5. The principle and scenario of the empirical models in the references [19,21,22,37,46–96].](image-url)

As an artificial intelligence method, ML has been widely used in modeling. These methods include Statistic Machine Learning (SML), Intelligent Computation (IC), ANN, etc. The most commonly used algorithms are the GMM, SVM, and random decision forests of SML. In [66], The GMM solves the problem of clustering in wireless propagation channels.
In [67–69], new methods based on SVM are proposed for blending and cabin scenarios. In [70,71], the authors established channel models in UAV scenarios using the random forest and KNN algorithms, and the comparative results are presented in Table 3. From the table, it can be concluded that in UAV scenarios, the predictive accuracy of the channel model established with the random forest outperforms that of the KNN algorithm.

Table 3. The result of Refs. [70,71].

<table>
<thead>
<tr>
<th>Reference</th>
<th>Scenario</th>
<th>Evaluation Parameters</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>[70]</td>
<td>UAV</td>
<td>MAE, RMSE</td>
<td>Random forest algorithm is better than KNN in channel modeling</td>
</tr>
<tr>
<td>[71]</td>
<td>UAV</td>
<td>RMSE</td>
<td></td>
</tr>
</tbody>
</table>

Moreover, the IC is also widely used in channel modeling. Cavalcanti et al. proposed a GA-based optimization method to tune constants present in empirical path loss models for urban areas at 879 MHz [72]. In [73,74], the GA was used to find the best coefficients and functions to fit the antenna measurements lower than the building and blending scenarios. In [75], a new model using PSO to tune a modified COST-231–WI propagation model was proposed in a 3G network scenario. After modification and verification, the average error gap of the PSO-based model at the test point is −1.96 dBm. In [76], using the PSO to tune the parameters of the COST-231 model improved its ability. From the above references, we can see that IC is typically employed in conjunction with other models for optimizing constants within a model. Table 4 shows the comparison of characteristics for IC.

Intelligent computing can be broadly categorized into two types: global search and local search. Genetic algorithms (GA) represent a typical global search algorithm that generates values within a specified range without directionality through “genetic” and “mutation”. Global search algorithms are suitable for scenarios with a small optimization range, but they exhibit slower optimization speeds, requiring a greater number of iterations to reach the optimal solution. Particle swarm optimization (PSO) is a typical local search algorithm. After defining the “population size” and “optimization range” for local search, at each iteration, the particle will update its position in the direction of the optimal position. Local optimization algorithms are directional and fast, making them suitable for scenarios with large optimization ranges, but they are prone to getting trapped in local optima. In [77], a comparison was conducted between the GA and PSO algorithms for optimizing the constants of the Hata model in a typical urban scenario. The results indicate that the PSO algorithm outperforms the GA algorithm in terms of prediction accuracy and convergence speed, with an RMSE difference of approximately 1 dB.

Table 4. Comparison of characteristics for IC.

<table>
<thead>
<tr>
<th>Type</th>
<th>Directionality</th>
<th>Speed</th>
<th>Range</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global search</td>
<td>Yes</td>
<td>Slow</td>
<td>Small</td>
<td>The optimal parameters can be found with enough iterations</td>
</tr>
<tr>
<td>Local search</td>
<td>No</td>
<td>Quick</td>
<td>Large</td>
<td>Easy to fall into local optimum</td>
</tr>
</tbody>
</table>

ANNs are widely used in the field of channel modeling due to their excellent self-learning and adaptive capabilities. In [74,78], the authors provided that the model based on ANFIS is more efficient, faster, and more accurate than the physical and empirical methods. In [79], the authors provided that the proposed ANFIS-based model offers desirable advantages in terms of simplicity, high prediction accuracy, and good generalization ability. As long as there are enough measurements, ANNs can be used for modeling any scenario, such as blending [80–82], indoor [83], suburban [84,85], and street scenarios [86]. The Back Propagation NN (BPNN) is used to model the blending [19,87] and cabin scenarios [88]. In particular, Ojo et al. developed two models based on the radial basis function NN (RBFNN) and the multilayer perceptron NN (MLPNN) by using the measured data as input variables.
for blending scenarios [89]. In [90], a pre-trained network called VGG-16 is proposed for modeling. They make the necessary modifications and fine-tune it with their training set. In [91], Popoola et al. use an extreme learning machine-LEM algorithm, which can reduce the training time on the premise of ensuring the accuracy of the trained model outdoors. In addition to direct modeling with ANN, it can also be used for auxiliary modeling. In [22,92], CNN extracts environmental information from satellite images to fit the path loss formula for blending and suburb scenarios for blending and suburb.

The above reference indicates that neural networks can achieve modeling tasks in various scenarios as long as there are sufficient training data and appropriate training strategies are selected. Figure 6 shows the steps to establish a channel model using neural networks. The modeling process can be divided into two stages: preparation and modeling. In the preparation stage, an analysis of the modeling scenario is required, as different neural networks exhibit varying performances in different scenarios. In [89], modeling was built using RBFNN and MLPNN in multi-antenna scenarios. Then, the models were compared with generic models such as COST-231, Free Space, Ericson, ITU, and ECC33. The comparison results indicated that the model based on RBFNN exhibited slightly better predictive accuracy than the one based on MLPNN. Both models significantly outperformed generic models in terms of prediction consistency and accuracy, with an average absolute prediction error improvement of over 10 dB. In [93], the modeling performance of ANFIS and GRNN neural networks was compared in a tropical urban scenario. The result shows that the model based on GRNN had a reduced RMSE of 0.59 compared to the one based on ANFIS. Both models achieved approximately a 4.71 dB improvement over a linear regression method composed of COST-231–Hata and COST-231–Walfisch–Ikegami models. The choice of neural network type depends on the modeling scenario. Once the model is determined, the next steps involve analyzing the application scenario, collecting, and preprocessing the feature data. The data collection principle involves selecting differentiating data strongly correlated with radio wave propagation. For example, in flat terrain areas with minimal topographical variations, the influence of terrain factors may not be a significant consideration during model training. Data preprocessing includes removing erroneous data and normalizing data. The aim of normalizing data is to ensure consistent measurement of various feature parameters. Finally, hyperparameters of the neural network are set using optimization algorithms such as the IC algorithm, as detailed above. In the modeling phase, the collected data are proportionally divided into training and testing sets. The training set is used to train the model, while the testing set is used to evaluate the trained model’s performance. If the trained model does not meet the requirements, adjustments are made by tuning hyperparameters, changing the model, or modifying feature data.

In addition, there are mixed methods used for modeling. For instance, Sotiroudis et al. proposed a path loss propagation model based on ANN for urban scenes and used the Differential Evolution (DE) algorithm to lay out an optimal ANN for path loss prediction [21]. In [94], a model is proposed based on a fuzzy method and curve-fitting for a suburban scene. Fuzzy sets are used to distinguish between the transmission discontinuities encountered during propagation. Then, the fuzzy sets are used to fit the curve. As another example, Jo et al. first used Principal Component Analysis (PCA), which can reduce the number of features of the dataset and simplify the learning model to assist feature selection. They then built the path-loss model using ANN for the suburb scenario [95]. In [96], PCA extracted relevant features from some selected path profile attributes for NLoS cases. Then, the path-loss based on the polynomial regression method model for urban is built.
2.3. Semi-Empirical Model

The semi-empirical model combines the advantages of the deterministic and the empirical models. This type of model replaces some complex electromagnetic calculations with statistical laws to reduce computational complexity [97–99]. This model’s accuracy and calculation complexity are between the two models. Figure 7 shows typical semi-empirical models.

Figure 7. The typical semi-empirical models.

2.3.1. The Typical Model

The random geometric method (RGM), which does not need detailed environmental parameters, is one of the most popular semi-empirical modeling methods. Both the COST-259 and the IMT-Advanced models are the standard models of radio communication
based on this method. The COST-259 model is suitable for macrocell radio communication. In order to research the diversity and adaptive antenna system, the directional characteristic modeling of the channel is particularly emphasized during the development of this model [14]. The IMT-Advanced model is the beyond third-generation (B3G) mobile communication system proposed by ITU. Depending on the environment, this model can support low-mobility to high-mobility applications and a variety of data rates. The IMT-Advanced system can also provide high-quality multimedia applications that significantly improve the quality of service under a wide range of services and platforms [15]. The ITU-R P.1812 (below referred to as P.1812) and the ITU-R P.2001 (below referred to as P.2001) models proposed by ITU are also widely used as semi-empirical channel models. The P.1812 model supplements the P.1546 model. It is a propagation prediction method suitable for terrestrial point-to-surface traffic. The height of the antennas of the transmitter and receiver is up to 3000 m above the ground [100]. However, this model is unsuitable for propagation prediction of air-ground or space-ground communication. The model obtains terrain profile information through the digital map. This model considers the end feature, location variability, and building entrance loss. This model also considers the effect of LoS propagation, diffraction, tropospheric scattering, and ducting on path loss [101]. The P.2001 model is a widely used model for terrestrial propagation. It can predict the terrestrial propagation path loss caused by signal enhancement and fading in the effective range of 0% to 100% beyond the annual average. The model’s predicted value is obtained by combining the predicted values of four independent sub-models according to the statistical correlation, and the model is particularly suitable for the Monte Carlo method [102].

In addition, with the deployment of 5G and 6G communication systems, many corresponding models emerged one after another. The most important are IMT-2020, 3GPP TR 36.873, and 3GPP TR 38.901. The IMT-2020 model is the new version of the IMT model, which includes the new capabilities of the IMT that go beyond those of the IMT-Advanced. Specifically, the ITM-2020 supports frequencies up to 100 GHz and large bandwidth, three-dimensional (3D) modeling, large antenna array, blockage modeling, spatial consistency, etc. The scenarios of IMT-2020 include enhanced mobile broadband, ultra-reliable and low-latency communications, and massive machine-type communications. It can support low- to high-mobility applications and enhance data rates according to the number of users or the service requirements. IMT-2020 can not only be mobile telecommunication but also a tool for enabling massive connections for a wide range of services. It is a typical model for 5G [103]. The 3GPP TR 36.873 (study on 3D channel model for LTE) is also an important reference. This technical report is suitable for the situation of 3D beamforming and FD-MIMO. The path-loss model in this technical report can be applied in the frequency range of 2–6 GHz and for different antenna heights. The usage scenarios include the LOS and NLOS of urban micro/macro, rural macro, and indoor hotspot cells. Among them, the coverage of the base station can reach at least 3 m and up to 10,000 m [104]. The new channel model proposed in the 3GPP TR 38.901 has, to a large degree, been aligned with earlier channel models for <6 GHz, such as 3GPP TR 36.873 or IMT-Advanced. The frequency range of this model defined in this technical report is about 0.5–100 GHz. The supported scenarios are urban microcell street canyon, urban microcell, indoor office, rural microcell, and indoor factory. The bandwidth of this model is supported up to 10% of the center frequency but no larger than 2 GHz. The base station height is 10–150 m, and the mobile station height is 1–22.5 m [105].

2.3.2. Other Models

Figure 8 shows the principle and the scenario of the semi-empirical models in the references. As mentioned earlier, the semi-empirical method is a method that combines the advantages of empirical and deterministic methods. Specifically, using actual measurements to fit the parameters that are not easy to determine in the deterministic method. Several non-empirical path loss models have been built for a small-scale scenario, such as the one/two-slope, COST-231 multi-wall, RT, floating intercept (FI), and close-in free space
reference distance (CI) model. For example, the indoor models were proposed based on the CI and the FI model, with compensation factor fit by the measurement [106,107]. Khatun et al. proposed a model based on the above for indoor and outdoor airports [108]. Zhou et al. proposed a model based on ML and the deterministic model [109]. A model based on the CI and FI was proposed after clustering using the space-alternating generalized expectation-maximization algorithm. The proposed model considers the different cluster characteristics, and it also units the directional and omnidirectional path loss model into the same framework, improving the 5G mmWave channel model to some extent. In [110,111], indoor models were proposed based on the two-slope model with compensation factor fit by the measurements. At the same time, a series of models were proposed based on the COST-231 multi-wall model for indoor [112,113] and home [114]. A log-distance path loss model is the extension of the Friis model and includes the random shadow effect caused by the signal being blocked by hills, trees, buildings, etc. Therefore, it is often used to predict the obstructed scene between the receiver and transmitter, such as smart homes [115], offices [116], and subway tunnels [117]. On the other hand, Bhuvaneshwari et al. proposed a model based on the RT for indoor [118]. The model was built by modifying Fresnel’s reflection coefficient with measurement data. Heereman et al. proposed a model based on the one-slope model for a large meeting room. Through measurement, it was found that the single-slope model with a PL index varying between 1.2 and 1.7 can accurately describe PL, and in the case of human existence, the PL index increases to 2 [119]. Diago–Mosquera et al. proposed a model based on Kriging for I2O using 3.5 GHz measurement analysis [120]. This model uses Kriging-aided design to provide a modeling method to reduce cost or policies. Aldossari et al. proposed a path loss model based on the ABG model using ML for problem related to unbalanced data described in the literature regarding using weighting policies. Khatun et al. proposed the WABG model based on the 3GPP ABG model for blending scenarios [122]. Most semi-empirical models are based on these models. For example, Casillas–Perez et al. proposed a model based on the above for indoor and outdoor airports [108]. Zhou et al. proposed a model based on ML and the deterministic model [109]. A model based on the CI and FI was proposed after clustering using the space-alternating generalized expectation-maximization algorithm. The proposed model considers the different cluster characteristics, and it also units the directional and omnidirectional path loss model into the same framework, improving the 5G mmWave channel model to some extent. In [110,111], indoor models were proposed based on the two-slope model with compensation factor fit by the measurements. At the same time, a series of models were proposed based on the COST-231 multi-wall model for indoor [112,113] and home [114]. A log-distance path loss model is the extension of the Friis model and includes the random shadow effect caused by the signal being blocked by hills, trees, buildings, etc. Therefore, it is often used to predict the obstructed scene between the receiver and transmitter, such as smart homes [115], offices [116], and subway tunnels [117]. On the other hand, Bhuvaneshwari et al. proposed a model based on the RT for indoor [118]. The model was built by modifying Fresnel’s reflection coefficient with measurement data. Heereman et al. proposed a model based on the one-slope model for a large meeting room. Through measurement, it was found that the single-slope model with a PL index varying between 1.2 and 1.7 can accurately describe PL, and in the case of human existence, the PL index increases to 2 [119]. Diago–Mosquera et al. proposed a model based on Kriging for I2O using 3.5 GHz measurement analysis [120]. This model uses Kriging-aided design to provide a modeling method to reduce cost or limit measurement activities. What is more, it can also model without measured data. In addition, Zhang et al. proposed a model integrating the ML, RT, and COST-231–WI for UAVs [70]. This model used the RT to produce data for training and testing. Random forest and KNN are used to correct the COST-231–WI model. Similarly, the RT is also used to generate data for training and testing purposes [121], and ANN is used to build a model.

![Figure 8. The basic model and the scenario of the semi-empirical models in the references [70,103–130].](image-url)
blending scenarios [123]. They first used regression techniques to fit the measurement data, which reduces the required number of measurements and the complexity. Then, this model was built using these processed data. Al-Samman et al. use the same method to propose a path loss model for an outdoor LoS environment [124]. In [125,126], path loss models are proposed based on the Friis model and measurement data for street scenarios and high-speed railroad communication. In [127,128], path loss models were developed based on the Uniform Geometrical Theory of Diffraction. In addition, Lee et al. also extended the ITU-R P.1411 (below referred to as P.1411) model to include NLOS, which reasonably explains propagation along street roads where low-height antennas are located [127]. Moreover, Inomata et al. used the measured data to add a correction factor to the P.1411 model to make it suitable for a street scenario [129].

Through analyzing the typical ABG and CI models, Karttunen et al. found that the estimated shadowing variance can be significantly overestimated if the “true” shadowing, the offset variations, and the slope of the straight-line fit between different street canyons are not adequately distinguished [130]. Based on the above analysis, a path loss model was proposed for the street scenario. This model overcomes the disadvantages of the traditional models. The result shows that the proposed model works well in terms of spatial consistency and is better than the existing deterministic and empirical models. Yu et al. proposed a model that uses the Okumura–Hata model to take the place of the equation transformed from the Friis in the Round Earth Loss (REL) model. The model also takes diffraction loss caused by obstacles into account. The result shows that the performance of the proposed model is better than the Okumura–Hata and RLE [131]. Moreover, Bhuvaneshwari et al. used RT to improve the path loss prediction of the COST 231–WI model [132]. In the revised model, RT and the statistically determined empirical measurements are combined. A loss term is computed using the method of images for multiple reflections. As a result, the error is considerably lowered for the proposed hybrid model. In [133], a hybrid model based on k-means and fuzzy logic is presented. The model can choose the best model for prediction from the Friis, Walfisch–Ikegami, Hata, ECC-33, Stanford University Interim, and ERICSSON models.

2.4. Model Comparison and Analysis

The channel model is generally measured by accuracy, complexity, and universality [134,135]. Accuracy reflects the difference between the measurement and the prediction; complexity reflects the calculation quantity and operation time of a model; universality reflects the usability of a model in different scenarios and environments. A perfect model should be the best compromise between accuracy, complexity, and universality. Table 5 is a comparison of three typical models. From Table 5, we can see that:

1. The principles of deterministic models are intricate, empirical models are straightforward, and semi-empirical models lie in between.
2. Correspondingly, deterministic models exhibit the lowest execution efficiency, while empirical models demonstrate the highest effectiveness.
3. The advantage of deterministic models lies in their higher precision, whereas empirical models have the lowest precision, and semi-empirical models fall in between.
4. Empirical and semi-empirical models have limited application scenarios, and their suitability is relatively constrained by distance and frequency range. In contrast, deterministic models are generally more versatile. However, there are exceptions with some empirical and semi-empirical models, such as ITU-R P.370, P.1546, P.1812, and P.2001, which have broad applicability, cover a wide frequency range, and offer good prediction accuracy. However, their computational complexity is relatively high.
Table 5. Comparison of Typical Models.

<table>
<thead>
<tr>
<th>Models</th>
<th>Key Performance Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
</tr>
<tr>
<td>Deterministic</td>
<td>RT</td>
</tr>
<tr>
<td></td>
<td>PE</td>
</tr>
<tr>
<td></td>
<td>COST-259</td>
</tr>
<tr>
<td></td>
<td>IMT-Advanced</td>
</tr>
<tr>
<td></td>
<td>IMT-2020</td>
</tr>
<tr>
<td>Semi-empirical</td>
<td>3GPP TR 36.873</td>
</tr>
<tr>
<td></td>
<td>3GPP TR 38.901</td>
</tr>
<tr>
<td></td>
<td>ITU-R P.1812</td>
</tr>
<tr>
<td></td>
<td>ITU-R P.2001</td>
</tr>
<tr>
<td></td>
<td>Okumura</td>
</tr>
<tr>
<td></td>
<td>Carey</td>
</tr>
<tr>
<td></td>
<td>Longley–Rice</td>
</tr>
<tr>
<td></td>
<td>Lee</td>
</tr>
<tr>
<td></td>
<td>Okumura–Hata</td>
</tr>
<tr>
<td></td>
<td>COST-231–Hata</td>
</tr>
<tr>
<td></td>
<td>COST-231–WI</td>
</tr>
<tr>
<td></td>
<td>ITU-R P.370</td>
</tr>
<tr>
<td></td>
<td>ITU-R P.1546</td>
</tr>
</tbody>
</table>

In conclusion, we can see that the empirical model has good complexity but bad accuracy and universality; the deterministic model has good accuracy, high complexity, and high university; the accuracy, complexity, and universality of the semi-empirical model are relatively balanced. Figure 9 shows the radar charts of these three channel models in accuracy, complexity, and universality.

![Figure 9. Performance comparison of three types of models.](image-url)
Figure 10 shows the percentage distribution of modeling methods in the surveyed literature. From the figure, it is evident that empirical methods and semi-empirical methods are currently the predominant modeling approaches, accounting for 62.00% and 32.00% of the surveyed references, respectively. Within empirical methods, the most commonly employed strategy is linear fitting, representing a substantial share of 37.50%. Artificial intelligence algorithms such as ANN, IC, and SML also hold significant proportions, at 28.57%, 14.29%, and 10.71%, respectively. Among semi-empirical methods, modeling based on the IC model, FI model, and RT method occupy the top three positions.

![Figure 10. Statistical chart of the proportion of modeling methods in the survey model.](image_url)

3. Future Prospects of Wireless Communication

3.1. Future Prospects of the Channel Model

With the evolution of the radio communication system, the measurement method, and intelligent technology, the current modeling method has challenged the demand for higher-quality radio communication services. The refinement and intelligence of the current channel modeling method need to be improved. In the future, the large-scale combination model suitable for large-scale scenes, the wide-area generalization model suitable for multi-scenes, and a more efficient intelligent model will be the main development direction of channel modeling. Figure 11 shows the development direction of channel modeling in the future.

3.1.1. Intelligent Modeling Process

It is challenging to balance the predicted efficiency and accuracy of the traditional channel modeling method. Modeling methods driven by artificial intelligence technologies, such as ML and IC, have excellent learning and prediction abilities, strong nonlinear fitting and adaptive abilities, good mining of complex features in high latitude and high redundancy data, and efficient performance. The ML can train channel models offline with measurement or simulation. In addition, due to its high nonlinearity, it is an ideal choice for predicting propagation parameters such as multipath fading. Moreover, it can also be used to model the specific or general model due to its high flexibility [136]. Figure 12 shows the classification and the advantages of intelligent modeling methods.
3.1. Future Prospects of the Channel Model

Figure 11. The development direction of channel modeling in the future.

Figure 12. The classification and advantages of intelligent modeling methods.

Modeling Method Based on ML

ML, which can self-learn and self-adapt, is a method that uses computers to build statistical models based on a great deal of data. We can efficiently utilize the built models to predict and analyze anonymous data. An ML algorithm can be supervised, semi-supervised, or unsupervised, depending on whether the data used for training is labeled. ML algorithms have excellent advantages in dealing with simple classification problems because of their...
simple structure and fast computing speed, making them the main algorithm for channel multipath clustering. In the field of ML, clustering belongs to unsupervised learning. K-means, the Gaussian mixture algorithm (GMA), and the density-based spatial clustering of applications with noise algorithm (DBSCAN) are the common clustering models. The K-means algorithm adopts Euclidean distance. Therefore, the model can easily achieve and converge quickly. However, it cannot capture the propagation characteristics of a multipath channel. The GMM algorithm is based on the statistical distribution of data. The DBSCAN algorithm is density-based. Both of these algorithms account for the statistical characteristics of multipath components (MPC) and can cluster dense datasets of any shape. However, when the cluster distribution is uneven, the DBSCAN algorithm has a poor clustering effect [137]. In [66], a multipath component clustering architecture was proposed based on statistical information and a compact index evaluation criterion of mean and variance. This model used the GMM model and the mean and covariance structure of channel multipath components to achieve multipath component clustering. In [138], He et al. compared a variety of multipath component clustering algorithms, including the KPM algorithm, fuzzy C-means clustering algorithm (FCM), kernel power density-based (KPD) algorithm, and DBSCAN algorithm. The first two algorithms need the number of clusters as an a priori message, and the last two algorithms can automatically generate a more reasonable number of clusters. In addition, to model the non-stationary characteristics of frequency and space through related cluster groups, the paper [139] further clusters based on clusters using the K-means clustering algorithm to obtain related cluster groups. That is, a cluster group is composed of multiple similar clusters. In addition, supervised learning is also widely used in channel modeling. In order to improve the accuracy of frequency band prediction for FM broadcasting in Beijing, Wang et al. established a high-precision model based on support vector machine (SVR) [140] and statistical machine learning algorithms [141], achieving good results and achieving high-precision modeling in Beijing.

Modeling Method Based on IC

At present, IC used in channel modeling mainly includes fuzzy inference system (FIS), ANN, evolutionary algorithm (EC), and swarm intelligence (SI). ANFIS is a kind of adaptive network equivalent to FIS in function [142]. This network solves the problem that traditional mathematical modeling tools cannot obtain satisfactory results in modeling fuzzy uncertain systems. At the same time, compared with the non-interpretability of the general ANN, ANFIS gives weight to the physical meaning of the inference parameters in the fuzzy logic, making the ANN interpretable. In [143], the author inputs the height and distance of the transceiver antenna into the fuzzy system to obtain the path loss in the areas to be predicted. The prediction results show that the path loss of the city is the largest, and the average loss increases by 10 dB every ten kilometers. The model is proven to have better performance by accurately predicting path loss. In [94], a model for predicting path loss based on the fuzzy method is proposed. This model uses free space, vegetation terrain, flat terrain, and rural terrain to establish a fuzzy set. The results make clear that the path loss index for open areas is 2.2, and for low vegetation, small towns, and high vegetation areas, the path loss index is 3.3, 4.1, and 4.7, respectively.

An ANN is composed of a number of interconnected neurons. The input layer, hidden layer, and output layer are the necessary structures of an ANN. It learns and optimizes through BP. Self-learning, self-adaptive, and nonlinear fitting are characteristics of ANNs. The ANN can effectively simulate the radio channel characteristics of real scenes by using massive measured datasets, fully training to determine the relationship between inputs and outputs. Therefore, the ANN is suitable for modeling channels with insignificant and time-varying statistical characteristics. The channel characteristics can be determined and modeled accordingly using ANNs, treating the channel as data. In [142], a model based on the ANN and adaptive evolutionary algorithm for the urban environment was built. This
model was compared with the RT. It makes clear that the prediction performance of the proposed model is better than the RT.

The main application of Evolutionary Computing (EC) is to solve combinatorial optimization problems. The process is simulated through program iteration, and the problem to be solved is regarded as an environment. The optimal solution will be found through natural selection among a population of possible solutions. In [73], a prediction model based on the microcellular area’s GA is proposed, which is applicable to the frequency within 900 MHz and the distance between 0.1 and 2 km. This model can well explain the influence of building area on path loss. Similarly, in [72], a model based on the GA was proposed. The GA is used to tune the free space mobile information system and the Ericsson model. The results show that the model test data are similar to the actual measurement data.

SI comes from research on the group behavior of social insects such as ants and bees. SI is a distributed, self-organized, and systematic collective behavior, whether natural or artificial, usually composed of a group of simple individuals without central control who interact with each other and the environment [144]. Among them, PSO is the most widely used SI in channel modeling. The PSO is a stochastic optimization algorithm based on the foraging behavior of birds, which is used to solve nonlinear problems [145]. A particle can be searched in space by combining three vectors: the personal best vector (x), the position vector (v), and the fitness value. The PSO algorithm can be applied during either the initialization phase or the iteration phase [146]. At initialization, the velocity and position of each vector are assigned randomly. In the iteration stage, the speed and position can be modified by the equation in [147] to approximate the best solution. Because the PSO does not need detailed propagation environment information, it is widely used in channel modeling. In [145], He et al. created an adaptive radial basis function (RBF) ANN model for path loss estimation using the PSO algorithm. The comparison results show that the performance of the proposed model is better than other RBF–NN-based models. Tahat and Taha applied statistical adjustment technology based on PSO to correct the path loss output of COST-231–WI. Compared with the measured values, this model has a more minor standard deviation than the traditional COST-231–WI model [75].

3.1.2. Hybridization Methods of Multi-Models

Radio channel models are often designed for different scenes, so each model has a different emphasis. There are many different scenes when predicting large-scale areas. Usually, satisfactory results will not be obtained using a single model. Therefore, we can combine different channel models to reach the goal of improving the accuracy of prediction. Multi-model fusion technology is mainly divided into subjective and objective fusion technology. In subjective fusion technology, researchers subjectively divided the predicted area of the independent sub-model according to the characteristics of each model. The typical fusion model is the advanced propagation model (APM) of the US. This model fuses the PE and the geometrical optical (GO) based models [148]. APM selects the Flat Earth (FE) model in the region where the propagation angle is greater than 5° and the propagation distance is less than 2.5 km. Because of the short propagation distance and the large transmitting angle, the earth can be approximately flat. Therefore, there are fewer obstacles in this region, so the FE model is used for prediction. Due to the propagation angle of the PE model varying with frequency, when the frequency is greater than 100 MHz, the maximum propagation angle of PE is 5°. Therefore, APM uses PE for prediction in the low-altitude region, where the altitude is high and there are almost no obstacles, a relatively simple Extended Optics (EO) model is used to predict. The EO is based on the assumption of parallel rays and is suitable for predicting high-altitude propagation without obstacles. Finally, the RO model is selected to predict propagation loss in other regions. This is because the RO model is suitable for flat surfaces and assumes that the refraction environment is horizontal and uniform. The region’s division of the APM is shown in Figure 13. The
objective fusion technology does not involve the subjective judgment of experts, and the prediction results are obtained entirely through objective formula calculation. Wang et al. used entropy weight theory and root mean square error as the evaluation standard to establish a combined prediction model by objectively weighting the ITU-R model [149].

![Diagram of APM and Modified APM](image)

**Figure 13.** The propagation region’s division of APM and future modification.

Based on the previous analysis, the propagation region is re-divided based on APM. Figure 13 gives the new region’s division for modifying APM. The region division is changed in the near-ground, where the propagation distance is less than 10 km, and the propagation height is less than 100 m. Dense buildings characterize the region. Therefore, RT can realize better prediction accuracy and operation speed by combining ML. Beyond this region, we can select a mixed method combining PE, semi-empirical, and empirical models, which can also be assisted by machine learning. This region is characterized by sizeable topographic relief and many ground obstacles. Because this region is close to the ground, we can easily obtain the radio environment. Therefore, it is easy to see that ML can effectively combine traditional methods for support modeling. As mentioned earlier, the propagation angle of PE varies with frequency. When the frequency reaches 20 GHz, the maximum propagation angle of PE is only 0.4°. Therefore, as the carrier frequency continues to increase in the future, the PE method will no longer be suitable for prediction in this area.

3.2. Potential Technology for Future Communication

Future radio communication will be a highly heterogeneous four-dimensional “space, air, ground, and sea” communication [150]. The communication enables comprehensive, multi-level, and three-dimensional coverage to meet communication applications in diverse environments and under various requirements. To achieve four-dimensional communication, the current wireless networks require even faster transmission rates and lower latency to support intelligent coordination among various communication systems. Longer transmission distances are necessary to facilitate long-distance transmission scenarios, such as those in air, space, and sea. The degree of sensory integration needs to be further enhanced to empower emerging applications like autonomous driving and digital twinning. Furthermore, there is a need to increase channel capacity and spectrum utilization further to address the explosive growth of wireless services and frequency devices. In response to the requirements mentioned above, new technologies such as terahertz (THz) communication, reconfigurable intelligent surfaces (RIS), integrated communication and sensing
ICAS, orbital angular momentum (OAM), and ultra-large antenna arrays technology (ULAs) will emerge as crucial enabling technologies driving the future development of wireless communication. Figure 14 shows the evolution from current communication to future communication.

**Figure 14.** The evolution from current communication to future communication.

### 3.2.1. The Application Technology of THz

Because millimeter wave has the advantages of continuity and wide bandwidth, it is the critical application of 5G [151]. Due to the increasing data volume and the pursuit of a high data transmission rate, millimeter wave has yet to meet the current development, which causes people to focus on the THz band with higher frequency and wider bandwidth [152]. The THz frequency band, which ranges from 0.1 to 10 THz, has rich bandwidth and spectrum resources. It is one of the key technologies of future communication. Combined with multi-antenna technology, THz is expected to achieve Tbps-level high-speed mobile communication. However, with the increase in frequency, the wavelength is shorter and shorter. The frequency coverage will significantly decrease, which will increase the networking cost. Meanwhile, there may be apparent changes in propagation characteristics. For example, because the wavelength of THz is close to molecules and small particles in the atmosphere (such as water vapor and oxygen), THz’s absorption or scattering effect will be more obvious [153,154]. In addition, narrow beams in THz communication are prone to misalignment, leading to a decrease in communication link capacity and even link interruption, which is also a major challenge for THz communication applications [155].

Fortunately, the beamforming technology based on MIMO can improve the transmission distance of THz communication. In [156], Enahoro et al. explored the performance of the least mean square algorithm (LMS) beamforming algorithm. The result shows that the LMS algorithm increases signal quality by eliminating interfering signals and noise by nulling them, while sending maximum signal (beams) to the desired direction. In [157], the prospect of ultra-massive multi-input multi-output (MIMO) technology to combat the distance problem at the THz band is considered. They achieved high gain and directionally...
adjustable narrow beam transmission based on MIMO technology. It also can significantly improve coverage by using RIS. ICAS can effectively solve the problem of narrow beam misalignment. RIS and ICAS will be discussed in the following text.

3.2.2. Reconfigurable Intelligent Surface Technology

As mentioned earlier, the high attenuation and small coverage of THz communication have become one of the key issues in its deployment. In the past, we usually improved signal coverage by establishing relay stations, but this will greatly increase the cost of deployment [158]. RIS can solve the problems of low coverage and low-frequency efficiency of THz communication. RIS integrates a large number of low-cost passive reflection elements on a planar surface, and each element independently reflects radio waves by controlling their amplitude and phase, thereby achieving the reconfiguration of the propagation environment [159]. Compared with traditional relay stations, RIS-assisted wireless networks can significantly improve communication quality and reduce networking costs. In the future, RIS can play the following role in wireless communication.

Act as a network repeater to increase coverage: By deploying RIS to establish a virtual line-of-sight link and receiver, the coverage range of millimeter wave communication can be greatly improved, and communication quality can be improved [160]. Wu et al. proposed a RIS-aided single-cell wireless system where one RIS is deployed to assist in the communications between a multi-antenna access point (AP) and multiple single-antenna users. The result shows that the received SNR increases quadratically with the number of reflecting elements of the RIS, which is more cost-efficient than the conventional massive MIMO or multi-antenna AF relay [161]. At the same time, Peng et al. utilized RIS to assist in UAV communication; thus, the communication quality and flexibility of the ground-to-air network were improved [162].

Improve the security of wireless communication: The physical layer security is a fundamental issue in wireless communications [163]. RIS can adaptively adjust the phase shifts of its reflecting units to strengthen the desired signal and/or suppress the undesired signal. Yang et al. utilized RIS to maximize the SNR of the legitimate user path, thereby cutting the eavesdropper path. They achieved the aim of improving the security of wireless communication [163]. Cui et al. maximized the secrecy rate of the legitimate communication link by jointly designing the AP’s transmit beamforming and the RIS’s reflect beamforming [164]. Chen et al. proposed a method of adjusting the reflecting coefficients of the RIS to change the attenuation and scattering of the incident electromagnetic wave so that it can propagate in the desired way toward the intended receiver [165].

In addition, RIS can also be used as the signal reflection hub. In this scenario, it can realize the communication from device to device and achieve communication interference suppression while supporting low-power transmission [159].

3.2.3. Integrated Communication and Sensing Technology

ICAS is also one of the key technologies in future communication. On the one hand, it is because many future application scenarios, such as unmanned driving, digital twins, and virtual reality, require both high-performance sensing and wireless communication [166]. On the other hand, communication and sensing systems have long been studied as two independent systems. However, they have great similarities in spectrum occupation, system architecture, antenna composition, and signal processing. Through joint design, such as signal transmission integration and hardware architecture integration, communication and sensing functions can be achieved simultaneously on the same set of devices and frequency spectrum to alleviate the shortage of spectrum resources, reduce equipment interference, and reduce equipment costs [167].

The integration of communication and perception realizes the integrated design of communication, perception, and computing functions through means such as air interface and protocol joint design, spectrum resource sharing, and software and hardware device sharing, resulting in the fusion and symbiosis of communication and perception.
functions. While transmitting information, communication and perception integration can also analyze the characteristics of radio waves such as direct, reflected, and scattered waves to locate, measure distance, measure speed, image, detect, identify, reconstruct the environment, and other target or environmental information [168].

As the key technology for future communication, ICAS has been extensively researched by the academic community [169]. Carlos et al. explored the full duplex ICAS system from a signal-processing perspective [170]. In [171,172], the effectiveness of communication and the bounds of the sensing are discussed. In [173,174], design schemes for wireless communications and radar sensing shared waveforms were proposed. In [175,176], new frameworks for ICAS were proposed. Preeti et al. proposed an adaptive preamble design to balance the positioning accuracy and communication rate of the ICAS system [177].

3.2.4. Orbital Angular Momentum Technology

The explosive growth of wireless services and devices is a major issue that needs to be considered in future communications. However, due to the strict limitations of spectrum resources in traditional multiplexing technologies (such as time division multiplexing, frequency division multiplexing, and code division multiplexing), current multiplexing technologies cannot meet the growing demand for ultra-high communication capacity from users. OAM can provide a new degree of multiplexing freedom for wireless communication systems and is recognized as one of the potential key technologies in future communication [178].

Compared with traditional planar waves, the vortex electromagnetic waves carrying OAM have a helical phase structure in the wavefront and a helical phase factor. In addition, different modalities theoretically satisfy strict orthogonality. Thus, in OAM wireless communication, multiple signals can be modulated to multiple OAM modes for transmission. This enables universal data transmission under limited resources and improves channel capacity. Research on OAM wireless communication has begun. Chen et al. proposed a misalignment–robust receiving model by extracting the phases of channel state information and implementing joint phase compensation and signal detection. Compared with the conventional model, the proposed model is more robust [179]. Jing et al. proposed a simple channel–independent beamforming model with fast symbol-wise maximum likelihood detection. Though the bit error rate of the proposed model is the same as that of conventional models, the numbers of complex additions and complex multiplications are much smaller than those of the traditional model [180]. Liang et al. proposed a hybrid orthogonal division multiplexing model using OAM multiplexing and orthogonal frequency division multiplexing in conjunction [181]. Hu et al. proposed an OAM-based independent analog beam selection by using the divergence to solve the problem that the interferences among different users degrade the system performance.

3.2.5. Ultra-Large Antenna Arrays Technology

Radio communication systems have already adopted MIMO and have demonstrated significant performance improvements [182]. With the increased RF, the spacing between antennas is reduced so that a more extensive antenna array can be constructed [183,184]. In 2010, Marzetta first proposed providing large-scale MIMO antenna arrays for the base station to improve users’ service [185]. In addition, using 3D antenna arrays can increase system capacity [186]. According to the channel measurement result, it has been verified that the 3D MIMO technology achieves 33% channel capacity gain compared to 2D MIMO technology [187]. With the increase in antennas, 3D MIMO channels become dispersed in the spatial domain, significantly increasing the channel capacity [188]. Therefore, large-scale MIMO technology is a crucial technology of 5G [189]. For future communication, as the frequency increases, ULAA can be used. However, many problems are still to be solved in modeling channels with the ULAA. For example, the measurement results of using virtual $40 \times 40$ planar antenna arrays show that the channel characteristics are non-stationary on the array [190]. Channel models should, therefore, take into account
the non-stationarity of antenna arrays. In addition, more antennas will produce narrower beams, so more accurate multipath propagation angles (MPCs) are required to make the radio wave propagation beam point accurately to users. In addition, it is also a challenge to use ULAAs for channel measurement from the perspective of channel measurement. On the one hand, applying ULAAs in HF channel transmitters is difficult, especially in the THz band. On the other hand, with the increase in antenna array size, the amount of measurement data increases exponentially, and more computing costs are required.

4. Conclusions

This paper systematically overviews the existing terrestrial radio channel models. Specially, these models are sorted into deterministic, empirical, and semi-empirical models. Then, we analyzed the features of the three kinds and summarized the semi-empirical model. The results show that the deterministic model has the highest accuracy, does not need a large number of measured data, and is suitable for various scenarios, but its calculation and simulation are time-consuming and not suitable for large-scale scenes. The empirical model has features of high universality, low complexity, simple calculation, poor accuracy, a large amount of measured data required, poor scene applicability, and suitable for small-scale scenes. The semi-empirical model has a moderate amount of computation, adaptation to large-scale scenes, low complexity, and higher accuracy. Namely, the three typical models are relatively balanced in universality, accuracy, and complexity.

At the same time, development trends in the future are also analyzed: the large-scale combination model suitable for large-scale scenes, the wide-area generalization model suitable for multi-scenes, and a more efficient intelligent model will be the main development direction of channel modeling. After that, for the requirement of intelligence and refinement, we propose to use the intelligent method based on the ML or the IC to realize the intelligent modeling process and use the fusion multi-model technology to achieve refined modeling. Finally, we look forward to the potential technologies of future communication, which include the THz application technology, reconfigurable intelligent surface technology, integrated communication and sensing technology, orbital angular momentum technology, and ultra-large antenna array technology. We hope this paper will stimulate more interest in modeling terrestrial radio channels.


Funding: This research was funded by the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (No. CEMEE2022G0201).

Conflicts of Interest: The authors declare no conflict of interest.

References

22. Thrane, J.; Zibar, D.; Christiansen, H.L. Model-aided deep learning method for path loss prediction in mobile communication systems at 2.6 GHz. IEEE Access 2020, 8, 7925–7936. [CrossRef]
35. Li, W.; Hu, X.; Jiang, T. Path loss models for IEEE 802.15. 4 vehicle-to-infrastructure communications in rural areas. IEEE Internet Things J. 2018, 5, 3865–3875. [CrossRef]


47. Zhao, X.; Geng, S.; Coulibaly, B.M. Path-loss model including LOS-NLOS transition regions for indoor corridors at 5 GHz [wireless corner]. *IEEE Antennas Propag. Mag.* 2013, 55, 217–223. [CrossRef]


64. Nossire, Z.; Gupta, N.; Almazaydeh, L.; Xiong, X. New empirical path loss model for 28 GHz and 38 GHz millimeter wave in indoor urban under various conditions. Appl. Sci. 2018, 8, 2122. [CrossRef]


68. Wang, A.; Lei, S. Research on modeling of underground path loss prediction based on SA-SVM. J. Phys. Conf. Ser. 2022, 2245, 012016. [CrossRef]


81. Nguyen, C.; Cheema, A.A. A deep neural network-based multi-frequency path loss prediction model from 0.8 GHz to 70 GHz. Sensors 2021, 21, 5100. [CrossRef]


90. Ahmad, O.; Ates, H.F.; Baykas, T.; Gunturk, B.K. Predicting path loss distribution of an area from satellite images using deep learning. *IEEE Access* 2020, 8, 64982–64991. [CrossRef]


92. Cheng, H.; Ma, S.; Lee, H.; Cho, M. Millimeter wave path loss modeling for 5G communications using deep learning with dilated convolution and attention. *IEEE Access* 2021, 9, 62867–62879. [CrossRef]


96. Jiang, R.T. Path loss modelling based on path profile in urban propagation environments. *IET Commun.* 2022, 16, 685–694. [CrossRef]


105. 3GPP TR 38.901; Study on Channel Model for Frequencies from 0.5 to 100 GHz. 3GPP: Sophia Antipolis, France, 2022.

106. Alwarafy, A.; Sulyman, A.I.; Alsanie, A.; Alshebeili, S.A.; Behairy, H.M. Path-loss channel models for receiver spatial diversity systems at 2.4 GHz. *Int. J. Antennas Propag.* 2017, 2017, 6795014. [CrossRef]


120. Diago-Mosquera, M.E.; Aragón-Zavala, A.; Rodríguez, M. Testing a 5G communication system: Kriging-aided O2I path loss modeling based on 3.5 GHz measurement analysis. Sensors 2021, 21, 6716. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.