Journal of Sciences, Computing and Applied Engineering Research (JSCAER), Vol. 1, No.3, pp. 69-83

Published Online (https://jcaes.net) , September 12, 2025 by SciTech Network Press ISSN: 3092-8648

Artificial Intelligence Methods for Adaptive Regression Learning of Signal Propagation Loss in Cellular Communication Networks: A Review

[1]Joseph Isabona and [2]Ikechi Risi

[1]Federal University Lokoja/Department of Physics, Lokoja, 260102, Nigeria
[2] Department of Physics, Rivers State University, Port Harcourt, Rivers State, Nigeria.
Corresponding Author: joseph.isabona@fulokoja.edu.ng,

Received: 12 July 2025; Revised: 02 August 2025; Accepted: 05 August 2025; Published: 12 September 2025

Abstract: Accurate prediction of signal propagation loss is paramount for efficient planning, deployment, and optimization of cellular communication networks. Traditional methods, ranging from empirical models to deterministic ray tracing, often suffer from limited adaptability, high computational complexity, or require extensive site-specific calibration. The advent of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized this domain, offering data-driven, adaptive, and highly accurate solutions for predicting continuous values like propagation loss—a classic regression problem. This paper provides a comprehensive review of AI techniques applied to the predictive regression learning of signal propagation loss in cellular networks. We examine the evolution from conventional ML algorithms like Support Vector Machines and Ensemble Methods to advanced Deep Learning (DL) architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Graph Neural Networks (GNNs). The review categorizes existing approaches based on their underlying AI methodologies, discusses their strengths and limitations, and highlights the crucial aspects of data feature engineering and model training. Furthermore, we identify key challenges, including model interpretability, generalization across diverse environments, and computational overhead. Finally, we explore promising future research directions, such as hybrid physics-informed AI models, federated learning, explainable AI (XAI), and the integration of digital twin technology, all of which aim to enhance the robustness, accuracy, and deployability of AI-driven propagation loss prediction for future 5G and 6G cellular ecosystems.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Regression, Signal Propagation Loss, Path Loss, Cellular Networks, Wireless Communication, 5G, 6G.

1. Introduction

The rapid proliferation of wireless communication devices and the insatiable demand for high-speed, low-latency connectivity have made cellular networks indispensable to modern society. The performance, coverage, capacity, and Quality of Service (QoS) of these networks are fundamentally governed by how radio signals propagate through various environments. Signal propagation loss, often referred to as path loss, quantifies the attenuation of signal strength as it travels from a transmitter (e.g., a base station) to a receiver (e.g., a mobile device). Accurate prediction of this loss is a cornerstone for critical network tasks, including base station placement, antenna tilt optimization, interference management, handover optimization, and dynamic resource allocation [1].

Historically, predicting signal propagation loss has relied on a variety of methods. Empirical models, such as Okumura-Hata and COST 231-Hata, provide generalized formulas derived from extensive measurements, offering simplicity but often lacking site-specific accuracy [2]. Deterministic models, like ray tracing, leverage the geographical and structural details of an environment to simulate signal paths, achieving high accuracy but at the cost of immense computational power and the need for highly detailed 3D environmental data [3]. Other approaches, including hybrid models, attempt to combine the

strengths of empirical and deterministic methods. Despite their utility, these traditional techniques face inherent limitations: empirical models are often calibrated for specific environments and frequencies, leading to poor generalization; deterministic models are computationally prohibitive for large-scale, dynamic networks; and all often struggle to adapt to rapidly changing urban landscapes, mobile users, and evolving network technologies (e.g., millimeter-wave frequencies in 5G and sub-THz in 6G).
The emergence of Artificial Intelligence (AI) and Machine Learning (ML), particularly Deep Learning (DL), offers a paradigm shift in addressing these challenges. AI-driven approaches are inherently data-driven, capable of learning complex, non-linear relationships directly from measurement data or high-fidelity simulations. This ability allows them to model intricate propagation phenomena that are difficult to capture with predefined mathematical equations, adapt to diverse environments, and potentially offer a more flexible and accurate prediction framework for the continuous value of propagation loss [4].
This paper provides a detailed review of the application of AI for predictive regression learning of signal propagation loss in cellular communication networks. Our primary focus is on how various AI models are trained to output a continuous value representing the loss, given a set of input features. We aim to:

- **Contextualize** the problem of signal propagation loss prediction within cellular networks and highlight the limitations of traditional methods.
- **Introduce** the fundamental AI/ML techniques relevant for regression learning.
- **Systematically review** the state-of-the-art AI-based approaches, categorizing them by their underlying methodology (e.g., classic ML, ANNs, CNNs, RNNs, GNNs) and discussing their specific contributions and applicability to this problem.
- **Analyze** the challenges associated with deploying AI for propagation loss prediction, including data issues, interpretability, and generalization.
- **Identify** promising future research directions that can further enhance the accuracy, robustness, and practicality of AI-driven solutions for upcoming wireless generations.

The remainder of this paper is structured as follows: Section 2 provides background on signal propagation loss. Section 3 introduces the fundamentals of AI/ML for regression. Section 4 offers a comprehensive review of AI-based prediction approaches. Section 5 discusses current challenges, and Section 6 outlines future research directions. Finally, Section 7 concludes the paper.

2. Background on Signal Propagation Loss

Signal propagation loss, also known as path loss, quantifies the reduction in radio signal strength between a transmitter and a receiver. It is a fundamental parameter that directly impacts the received signal power, and consequently, the achievable data rates, coverage area, and overall network performance.

2.1. Mechanisms of Propagation Loss Radio signals attenuate due to several physical phenomena as they travel through space:

- **Free-Space Loss:** This is the most basic form of attenuation, occurring even in a vacuum. It is proportional to the square of the distance and the square of the frequency. The Friis transmission equation provides the theoretical free-space path loss.
- **Reflection:** Occurs when a radio wave encounters an object with large dimensions compared to its wavelength (e.g., large buildings, ground). The wave bounces off the surface.
- **Diffraction:** Occurs when a radio wave encounters an obstruction with sharp edges. The wave "bends" around the obstacle, allowing signals to propagate into shadowed areas, albeit with significant attenuation.
- **Scattering:** Occurs when a radio wave encounters objects with dimensions comparable to or smaller than its wavelength (e.g., foliage, street furniture, rough surfaces). The energy is spread in multiple directions.
- **Absorption:** Energy is absorbed by materials (e.g., walls, human bodies, atmospheric gases like oxygen and water vapor), converting electromagnetic energy into heat.

2.2. Factors Influencing Propagation Loss The magnitude of signal propagation loss is influenced by a multitude of factors, making its accurate prediction complex:

- **Distance:** Generally, loss increases logarithmically with distance.
- **Frequency:** Higher frequencies experience greater free-space loss and are more susceptible to absorption, diffraction, and blockage by obstacles. This is particularly relevant for 5G millimeter-wave (mmWave) and future sub-THz bands.
- **Environment:**
    - **Urban:** Dense buildings, high clutter (vehicles, people) lead to significant reflection, diffraction, and scattering.
    - **Suburban:** Less dense buildings, more open areas, foliage.
    - **Rural:** Open terrain, fewer obstacles, but terrain undulations can be significant.
    - **Indoor/Outdoor:** Indoor environments introduce additional wall penetration losses.
- **Antenna Heights:** Higher transmitter and receiver antennas generally reduce path loss due to fewer obstructions and clearer line-of-sight (LOS) paths.
- **Clutter:** Presence of trees, foliage, street furniture, and other small-scale objects.
- **Weather Conditions:** Rain, fog, and humidity can significantly attenuate signals, especially at higher frequencies.
- **Building Characteristics:** Material type, density, and height of buildings.

2.3. Traditional Path Loss Models Traditional models can be broadly categorized as:

- **Empirical Models:** Based on extensive measurement campaigns and statistical fitting.
    - **Free-Space Path Loss (FSPL):** A theoretical baseline, only valid for LOS in a vacuum.
    - **Log-Distance Model:** A simple extension of FSPL, incorporating an average path loss exponent.
    - **Okumura-Hata and COST 231-Hata:** Widely used for macrocell environments, but developed for specific frequency ranges (e.g., 150-2000 MHz) and environmental types. They assume a general environment and are less accurate for specific deployments or new frequency bands (e.g., mmWave).
    - **Advantages:** Simple, fast, low computational cost.
    - **Disadvantages:** Limited accuracy, poor generalization outside their calibrated environments, and lack of adaptability to dynamic changes.
- **Deterministic Models:** Rely on electromagnetic theory and detailed environmental databases.
    - **Ray Tracing/Launching:** Simulates the propagation of individual rays, accounting for reflection, diffraction, and scattering based on 3D building models and terrain data.
    - **Advantages:** High accuracy, can model complex environments.
    - **Disadvantages:** Extremely high computational complexity, requires detailed and expensive 3D geographical data, not suitable for real-time dynamic networks.
- **Hybrid Models:** Combine aspects of empirical and deterministic approaches to balance accuracy and computational efficiency. However, they often inherit some limitations of their constituent methods.

The limitations of these traditional models, particularly in the context of dense urban deployments, varied network scales (macro, micro, pico, femto), and the dynamic nature of 5G and 6G networks, underscore the need for more adaptive and accurate prediction methodologies. This is where AI and ML offer a compelling alternative.

3. Fundamentals of AI/ML for Regression Learning

Regression analysis in Machine Learning is a supervised learning task focused on predicting a continuous output variable based on one or more input features. In the context of signal propagation loss, the input features could include anything from distance and frequency to geographical coordinates and environmental characteristics, while the output is the continuous value of path loss in dB. Shown in table 1 is the comparison of the general AI regression based techniques for signal propagation loss prediction.

Table 1: Comparison of AI techniques for signal propagation loss prediction

AI Technique	**Prediction Accuracy**	**Computational Complexity**
Regression Analysis	Moderate	Low
Machine Learning	High	Moderate
Deep Learning	Very High	High

From the table, it is evident that DL algorithms outperform other AI techniques in terms of prediction accuracy. However, they also have higher computational complexity, which may limit their applicability in real-time scenarios.

3.1. Supervised Learning Paradigm In supervised learning, an algorithm learns from a labeled dataset, where each instance consists of input features (f{x}) and a corresponding target output f{y}). The goal is to learn a mapping function:

$$f(x)=\frac{1}{2}\sum_{i=1}^{n}[y_i-f(x_i,\boldsymbol{b})]^2 \quad (1)$$

$$f(x)=\frac{1}{2}\sum_{i=1}^{n}\boldsymbol{q}(\mathbf{x})^2 \quad (2)$$

where.

$$\boldsymbol{q}(\mathbf{x})=y_i-f(x_i,\boldsymbol{b}) \quad (3)$$

and

$$\boldsymbol{b}=(b_1,b_2,b_3,b_4,b_5) \quad (4)$$

The learning process involves minimizing a loss function in equation to minimize the Mean Squared Error (MSE) that quantifies the difference between the predicted and actual values.

3.2. Traditional Machine Learning Algorithms for Regression Several classic ML algorithms have been successfully applied to regression problems, including propagation loss prediction:

- **Linear Regression:** A foundational model that assumes a linear relationship between input features and the target variable. While simple, it often serves as a baseline and can be effective when relationships are approximately linear. Polynomial regression extends this by fitting a polynomial function.
- **Support Vector Regression (SVR):** An extension of Support Vector Machines (SVMs) for regression. Instead of finding a hyperplane that separates classes, SVR finds a hyperplane that best fits the data points within a specified margin of tolerance (ϵ-insensitivity zone). It is powerful for non-linear relationships using kernel functions (e.g., Radial Basis Function - RBF) and is robust to outliers [5].
- **Decision Trees (DTs):** A non-parametric model that partitions the feature space into a set of rectangular regions. For regression, the prediction in each region is the average of the target values of the training points falling into that region. DTs are intuitive but can be prone to overfitting.
- **Ensemble Methods:** Combine multiple individual models to improve overall accuracy and robustness.

    - **Random Forests (RF):** Builds an ensemble of decision trees, each trained on a bootstrapped sample of the data and a random subset of features. Predictions are averaged

(for regression). RFs are robust, handle high-dimensional data, and can provide feature importance [6].
  - **Gradient Boosting Machines (GBM):** Sequentially builds models where each new model corrects the errors of the previous ones. Algorithms like XGBoost, LightGBM, and CatBoost are highly efficient and accurate implementations of GBM [7].

3.3. Artificial Neural Networks (ANNs) and Deep Learning (DL) Architectures ANNs are inspired by the structure and function of biological neural networks. Deep Learning is a subfield of ML that utilizes ANNs with multiple hidden layers (hence "deep") to learn hierarchical representations of data.

- **Multi-Layer Perceptrons (MLPs) / Feedforward Neural Networks:** The simplest form of ANNs, consisting of an input layer, one or more hidden layers, and an output layer. Each neuron in a layer is connected to all neurons in the subsequent layer. MLPs are universal function approximators, capable of learning complex non-linear mappings between inputs and outputs. They are trained using backpropagation and gradient descent optimization [8].
- **Convolutional Neural Networks (CNNs):** Primarily designed for processing grid-like data, such as images. CNNs employ convolutional layers that apply filters to detect local patterns and features (e.g., edges, textures) across the input. For propagation loss, CNNs can effectively process spatial data like topographical maps, building footprints, and clutter maps as input features, learning location-specific propagation characteristics [9].
- **Recurrent Neural Networks (RNNs), LSTMs, and GRUs:** Designed for sequential data, where the output at a given time step depends on previous inputs and outputs. RNNs have internal memory, allowing them to capture temporal dependencies. Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) are advanced RNN architectures that mitigate the vanishing gradient problem, making them suitable for long sequences. They can be applied to predict propagation loss in dynamic environments where path loss changes over time or across a sequence of locations [10].
- **Graph Neural Networks (GNNs):** A more recent and powerful class of neural networks designed to operate on non-Euclidean data structures, specifically graphs. In cellular networks, the topology (base stations, user equipment, connections) can naturally be represented as a graph. GNNs can learn representations of nodes and edges, capturing relationships between different network entities, which is highly relevant for modeling complex propagation environments and network-wide interactions [11].

The choice of AI/ML algorithm depends heavily on the nature of the input data, the complexity of the underlying propagation phenomena, the available computational resources, and the desired level of accuracy and interpretability.

4. AI for Propagation Loss Prediction: A Review of Approaches

The application of AI/ML for signal propagation loss prediction has gained significant traction, moving beyond traditional statistical or physical models. This section reviews various AI-based approaches, categorizing them primarily by the underlying ML/DL technique used.

4.1. Traditional Machine Learning Approaches

Early applications of AI for path loss prediction often leveraged traditional ML algorithms due to their relative simplicity, lower data requirements compared to deep learning, and sometimes better interpretability.

- **Linear and Polynomial Regression:** While simple, these models have been used as a baseline or in conjunction with feature engineering. They can capture the fundamental logarithmic relationship between distance and path loss, often enhanced by incorporating other linear effects of frequency or environment [12].
- **Support Vector Regression (SVR):** SVR has been widely explored due to its ability to handle non-linear relationships using kernel functions and its robustness to noise. Studies have shown SVR outperforming empirical models by adapting to specific environments through training data. For instance, [13] demonstrated SVR's superior accuracy for urban microcell

environments compared to Okumura-Hata and COST 231. Its effectiveness in handling high-dimensional feature spaces, including various environmental parameters and antenna characteristics, has also been noted.

- **Decision Trees and Ensemble Methods (Random Forests, Gradient Boosting):** These methods have proven highly effective due to their ability to model complex, non-linear interactions between numerous input features without requiring extensive data pre-processing or assumptions about data distribution.

    - **Random Forests (RF):** Several works, such as [6] and [14], have shown RFs providing significantly better predictions than empirical models for diverse environments (urban, suburban, indoor) and frequency bands. RFs can inherently handle mixed data types (numerical and categorical) and provide insights into feature importance, highlighting which environmental factors or network parameters contribute most to path loss.
    - **Gradient Boosting (e.g., XGBoost, LightGBM):** Advanced boosting algorithms have also been successfully applied. They often achieve state-of-the-art results due to their sequential error correction mechanism, leading to very accurate predictions for path loss [15]. These models are particularly well-suited for structured datasets with numerous features representing various propagation conditions.

Strengths of Traditional ML:

- Generally less computationally intensive to train than deep learning.
- Require less training data compared to deep neural networks.
- Some models (e.g., Decision Trees, RF feature importance) offer a degree of interpretability.
- Effective for datasets with well-defined, engineered features.

Limitations of Traditional ML:

- May struggle to automatically extract abstract features from raw, unstructured data (e.g., raw map images).
- Performance can plateau in highly complex, high-dimensional, and non-linear propagation scenarios.
- Feature engineering can be time-consuming and requires domain expertise.

4.2. Artificial Neural Networks (ANNs) and Deep Learning Approaches

The ability of deep neural networks to learn hierarchical representations and highly non-linear mappings has made them a popular choice for challenging regression problems like propagation loss prediction, especially with the availability of larger datasets and increased computational power.

(a) **Multi-Layer Perceptrons (MLPs):** As the foundation of deep learning, MLPs, also known as feedforward neural networks, have been extensively used. Studies often feed MLPs with engineered features such as transmitter-receiver distance, frequency, antenna heights, building densities, clutter types, and even latitude/longitude [8, 16]. MLPs can approximate highly complex, non-linear functions that govern radio wave propagation, often outperforming traditional empirical models and sometimes even some of the classic ML algorithms, especially when the input-output relationship is intricately non-linear.

- **Strengths:** Universal function approximators, capable of modeling highly complex relationships.
- **Limitations:** Can be "black-box" models, sensitive to hyperparameter tuning, and may require significant amounts of data to generalize well.

(b) Convolutional Neural Networks (CNNs): CNNs are particularly powerful when the input data has a spatial structure, like topographical maps, building footprints, or images representing the environment.

- **Application:** Researchers have utilized CNNs by converting environmental data (e.g., building height maps, land cover maps, digital terrain models) around the transmitter and receiver into image-like inputs [9, 17]. The convolutional layers can then automatically extract relevant spatial features (e.g., presence of high-rise buildings, density of obstacles, line-of-sight blockages) that are critical for propagation. This approach effectively automates a significant part of feature engineering.

- **Strengths:** Excellent at learning spatial features directly from raw data, reducing reliance on manual feature engineering; robust to small translations and distortions in input data.
- **Limitations:** Requires converting environmental data into a grid format, which might lose some fine-grained information; high computational cost for training large models.

(c) Recurrent Neural Networks (RNNs) and their variants (LSTMs, GRUs): These networks are designed to process sequential data, making them suitable for scenarios where propagation loss changes over time or across a sequence of locations.

- **Application:** RNNs can be used to model path loss in dynamic environments, such as for mobile users where the channel conditions evolve over time [10]. LSTMs and GRUs, in particular, can capture long-term dependencies in the channel, predicting future path loss values based on historical measurements and sequences of user movement or environmental changes. This can be crucial for predictive handover mechanisms or dynamic resource allocation.
- **Strengths:** Capable of learning from time-series data, capturing temporal dependencies.
- **Limitations:** Can be computationally intensive, and the sequence length can impact performance. Generating large, labeled time-series datasets for propagation loss can be challenging.

(d) Graph Neural Networks (GNNs): GNNs are an emerging class of neural networks designed to operate on graph-structured data. Cellular networks naturally lend themselves to graph representation, where base stations, users, and their interconnections form a graph.

- **Application:** GNNs can model propagation loss by representing the network as a graph where nodes are base stations or user equipment, and edges represent communication links. Node features might include antenna parameters, location, and environmental clutter, while edge features could include distance or initial propagation estimates. GNNs can learn how propagation loss between two nodes is influenced by their neighbors and the overall network topology [11, 18]. This is particularly promising for modeling complex multi-cell interference scenarios and network-wide channel state information.
- **Strengths:** Ideal for modeling non-Euclidean, relational data inherent in network topologies; can capture complex interactions between different network elements.
- **Limitations:** Still a relatively new area of research for path loss prediction; constructing accurate and scalable graph representations of large cellular networks can be challenging; computational complexity can be high for very large graphs.

4.3. Hybrid AI-Based Approaches

Some research explores combining different AI techniques or integrating AI with traditional models:

- **Hybrid ML/DL Models:** For example, using an RF model to identify important features, then feeding those to an MLP, or using CNNs for feature extraction from images, followed by an MLP for regression [19].
- **AI-Enhanced Ray Tracing:** Instead of fully replacing ray tracing, AI can optimize it (e.g., dynamically adjusting ray parameters) or correct its errors, leading to faster and more accurate deterministic models. AI can also be used to learn and interpolate propagation effects in areas not covered by a full ray tracing simulation [20].

4.4. Data Sources and Feature Engineering

The performance of any AI model heavily relies on the quality and quantity of its training data and the relevance of its input features.

(a) **Input Features:** Common features for path loss prediction include:

- **Geometric:** Tx-Rx distance, Tx/Rx antenna heights, geographical coordinates.
- **Frequency:** Carrier frequency.
- **Environmental:** Building heights, building density, land cover (clutter categories like urban, suburban, rural, water, forest), terrain elevation, street width, vegetation density.
- **Antenna Characteristics:** Antenna type, beamforming parameters (for active antennas).
- **Time/Weather:** Time of day, day of week, seasonal effects, atmospheric conditions.

(b) **Data Acquisition:**

- **Drive Tests/Measurement Campaigns:** Directly collected real-world data, highly accurate but expensive and time-consuming.
- **Ray Tracing Simulations:** Generate large amounts of high-fidelity data for various scenarios, but dependent on the accuracy of 3D environmental models.
- **Crowdsourcing:** Utilizing data from user devices, offering scale but with potential issues regarding data consistency and privacy.
- **Publicly Available Geospatial Data:** Open-source maps, building databases (e.g., OpenStreetMap, 3D building models), digital elevation models (DEMs), satellite imagery.

Effective feature engineering is crucial for traditional ML models, transforming raw data into meaningful inputs. For deep learning, especially CNNs and GNNs, the models can learn features directly from raw data representations, reducing manual effort but increasing data volume requirements.

5. Challenges and Open Issues

Despite the significant advancements, several challenges and open issues remain in the quest for fully robust and deployable AI-driven propagation loss prediction in cellular networks:

5.1. Data Availability, Quality, and Diversity:

- **Scarcity of Labeled Data:** High-quality, diverse datasets covering a wide range of frequencies, environments, antenna configurations, and measurement conditions are expensive and time-consuming to acquire. This is particularly true for emerging scenarios like mmWave and sub-THz propagation, indoor environments, or specific 3D urban topographies.
- **Data Inconsistency and Bias:** Measurement errors, device variations, and inconsistent data collection methodologies can introduce noise and bias, leading to models that generalize poorly.
- **Privacy Concerns:** Crowdsourced data, while abundant, raises significant privacy implications regarding user locations and movements.

5.2. Model Interpretability and Explainability (XAI):

- **Black-box Nature:** Deep Learning models, in particular, are often "black boxes," making it difficult to understand why a particular prediction is made. In engineering, understanding the underlying physical reasons for attenuation is crucial for network troubleshooting and design decisions.
- **Trust and Acceptance:** Lack of interpretability hinders trust among network engineers and operators, who often prefer models with clear physical justifications. Developing explainable AI (XAI) for wireless propagation is a critical area.

5.3. Generalization and Transferability:

- **Site-Specificity:** Models trained in one geographical area or environment often perform poorly when deployed in a different one, necessitating extensive retraining or fine-tuning, which is costly and time-consuming.
- **Frequency and Technology Dependence:** Models trained for one frequency band (e.g., sub-6 GHz) may not accurately predict loss for another (e.g., mmWave), due to different

propagation characteristics. The same challenge applies to different antenna technologies (e.g., MIMO, massive MIMO).

- **Dynamic Environments:** Most models struggle to adapt to real-time changes in the environment (e.g., new buildings, seasonal foliage changes, dynamic scattering from vehicles and pedestrians) without continuous retraining or recalibration.

5.4. Computational Complexity and Real-time Operation:

- **Training Time:** Deep Learning models, especially those with many layers or complex architectures (like large CNNs, GNNs), require significant computational resources (GPUs/TPUs) and time for training, especially with large datasets.
- **Inference Latency:** While inference is generally faster than training, for real-time applications like dynamic resource allocation or mobility management, even millisecond delays can be critical. Deploying complex models on resource-constrained edge devices is a challenge.

5.5. Hybridization with Physical Models:

- How to effectively combine the data-driven power of AI with the physical accuracy and interpretability of traditional electromagnetic models (e.g., ray tracing) remains an open challenge. Purely data-driven models might violate physical laws, while purely physics-based models lack adaptability.

5.6. Uncertainty Quantification:

- Most AI models provide point predictions without quantifying the uncertainty or confidence associated with those predictions. For critical wireless applications, understanding the prediction variance is important for risk assessment and robust decision-making.

5.7. Specific Literature Review: 2010-2025

This review employs a systematic approach to identify relevant literature. The review was accomplished through comprehensive survey of previous works on AI applications for signal propagation loss, combining predictive regression learning, machine learning, and neural networks modelling techniques. The specific focus was on publications from 2010 to 2025 as presented in table 2.

Table 2:Specific Literature Review on AI Application for propagation loss prediction : 2010-2025

Study	**AI Method**	**Data Source**	**Key Findings**	**Advantages**	**Limitations**
Zhang et al. (2015) [25]	Support Vector Regression (SVR)	Field measurement data	Achieved accuracy of 92% in urban environments	Robust against noise; effective in high dimensions	Needs extensive parameter tuning; requires good data quality
Isabona, and Srivastava (2016) [26]	Neural Network with log-distance path loss model	Field measurement	Achieved 95% accuracy with proposed method	High precision accuracy	Needs further field work
Ebhota et al (2018) [27]	Review of AI and non AI methods	No field data	Different AI systems achieved different accuracy	Different AI system models were stidied	No field work
Ehota et al,	Combined	Detailed field	Proposed hybrid	Proposed hybrid	Need further

Study	AI Method	Data Source	Key Findings	Advantages	Limitations
(2018)	Vector Statistics combined with AI systems	work	approach better than non hybrid method	approach achieved improved accuracy	validation
Lee et al. (2018) [29]	Artificial Neural Networks (ANN)	Simulated data	Improved prediction accuracy over traditional models by 15%	Non-linear modeling capability; generalization to unseen data	Computational complexity; risk of overfitting
Ebhota et al, (2019) [30]	ANN hyparameter s investigated	Field data	Learning rate has huge impart on ANN modes	Impact of learning rates on ANN revealed	There is need to investigated other parameters
Isabona and Igbinovia (2019) [31]	Different AI systems investigated	Field data	The radial basis function model gave better prediction results compared to the MLPNN and GRNN	The GRNN model also gave a good prediction results with marginal errors compared to the MLPNN	The complexity of each models was not investigated
Isabona (2020) [32]	ANN system studied with Wavelet	Field data	ANN method combined with wavelet processing achieved improved accuracy	The impact of pre-proccessing with ANN revieved	The complexity of hybrid method was not investigated
Chen & Wu (2020) [33]	Random Forest (RF)	Mixed datasets from various environments	Reduced prediction error by incorporating environmental factors	Handles large datasets; less sensitive to outliers	Limited interpretability of results; ensemble method complexity
Kumar et al. (2021)[34]	Deep Learning (LSTM)	Time-series signal data	Captured temporal correlation effectively, showing improved prediction	Captures complex relationships; suitable for sequential data	High computational cost; requires a massive amount of data
Isabona (2021)[35]	Joint Statistical and Machine Learning Approach	Field data	Combined hybrid approached achieved better precision accuracy	Impact of engaging a hybrid Statistical and AI revealed	May result to high computational cost
Kumar et al. (2021) [36]	Deep Learning (LSTM)	Time-series signal data	Captured temporal correlation effectively, showing	Captures complex relationships; suitable for sequential data	High computational cost; requires a massive amount of data

Study	AI Method	Data Source	Key Findings	Advantages	Limitations
			improved prediction		
Ituabhor et al, (2022) [37]	Cascade forward neural networks	Time-series signal data	Captured the time-series data correlation	Accurate in adaptive learning of time series data	Limited to time series data
Ituabhor et al, (2022) [38]	Hybrid Empirical and Machine Learning	Non-time series data	Achieved a high RMSE accuracy	High predictive accuracy	Computational cost to studied
Smith et al. (2023) [39]	Gradient Boosting Machines (GBM)	Real-world network performance data	Achieved a high R^2 of 0.95 for urban scenarios	High predictive accuracy; flexible handling of various data types	Increased likelihood of overfitting; may require cross-validation
Olukanni et al (2023) [40]	Hybrid Empirical and ANN	Real-world network performance data	Achieved a high R^2 of 0.95 for complex urban area	Achieved High predictive accuracy of various data	Further validation study needed
E. F. Ramirez et al, (2024) [41].	Hybrid AI Systems	Combine multiple AI methods for accuracy	Achieved some level of accuracy	Best of both worlds with respect to accuracy	Increased complexity in system integration
Ugbeh R.N et al, (2025) [42]	Hybrid AI model	Real-time data	Developed Hybrid Machine model very efficient	Achieved high precision accuracy	Complexity of developed model
kechi Risi, and Konyeha, (2025) [43]	Deep AI model	Real-time data	Deep AI very efficient with large data	Achieved high accuracy with large datasets	Achieved low accuracy with small datasets
Emughedi, (2025) [44]	Machine Learning with Exhaustive GMM Clustering Algorithm	Real-time data	Non hybrid machine clustering algorithm	Achieved precised clustering performance	Further validation of approach needed

In analyzing the studies, several patterns emerge. Machine learning techniques such as SVR, ANN, RF, LSTM, and GBM show varied success in predicting signal propagation loss. Key insights from the literature highlight:

- **Variability in Data Sources**: Many studies use simulated data, which may not translate perfectly to real-world scenarios. There is a need for more field-tested approaches that combine both simulated and empirical data.
- **Model Complexity vs. Interpretability**: More complex AI models like deep learning offer superior prediction capabilities but at the cost of interpretability. Stakeholders in

telecommunications may find simpler models like RF or SVR more useful despite some loss in accuracy.

- **Emerging Trends**: The recent trend towards utilizing deep learning techniques, particularly LSTM, indicates a growing recognition of the importance of temporal and spatial correlations in signal loss data.
- **Challenges Ahead**: While AI methods show promise, the necessity of extensive datasets for training and validation remains a critical barrier, especially in forming generalizable models applicable in diverse environments.

6. Future Directions

The field of AI for wireless propagation loss prediction is rapidly evolving. Addressing current challenges and capitalizing on emerging AI paradigms will pave the way for more robust, accurate, and practical solutions.

6.1. Hybrid Physics-Informed AI Models:

- **Integrating Domain Knowledge:** Future research will increasingly focus on "Physics-Informed Machine Learning (PIML)" [21]. This involves incorporating known physical laws (e.g., Friis equation, reflection/diffraction principles, conservation of energy) directly into the AI model's architecture, loss function, or regularization. This can improve generalization, reduce data requirements, and enhance interpretability by ensuring predictions adhere to fundamental physical principles. For instance, an AI model could learn correctional factors for a ray-tracing engine rather than predicting path loss from scratch.

6.2. Federated Learning for Privacy-Preserving Data Collection:

- **Distributed Training:** Federated Learning (FL) allows multiple organizations (e.g., network operators) or devices (e.g., user equipment) to collaboratively train a shared AI model without exchanging their raw local data [22]. This is crucial for overcoming data privacy concerns and aggregating diverse datasets from various network deployments, leading to more robust and generalizable models.

6.3. Explainable AI (XAI) for Wireless Communications:

- **Demystification** Developing XAI techniques tailored for wireless propagation models is essential for building reliable and enabling practical deployment. This includes methods to visualize what features an AI model prioritizes, identify critical propagation paths, or pinpoint environmental factors driving specific attenuation levels.

6.4. Digital Twin Technology and Real-time Prediction:

- **Virtual Network Replica:** The concept of a "digital twin" – a virtual replica of a physical network synchronized with real-time data – offers a powerful framework. AI propagation models can be a core component of this digital twin, providing highly accurate, real-time path loss predictions that adapt to dynamic environmental changes, user mobility, and network configurations. This enables proactive optimization, planning, and predictive maintenance [23].

6.5. Advanced Graph Neural Networks and Transformers:

- **End-to-End Network Modeling:** Further research into GNNs for cellular networks holds immense promise, moving beyond simple node-to-node prediction to model complex network-level interactions, interference, and resource allocation decisions that depend on propagation. Transformer architectures, initially successful in natural language processing and computer vision, are also being explored for their ability to model long-range dependencies and complex interactions in large, unstructured datasets, potentially applicable to propagation environments [24].

6.6. Unsupervised and Self-Supervised Learning:

- **Reducing Label Dependence:** Given the challenge of acquiring labeled data, research into unsupervised and self-supervised learning methods is critical. These approaches can learn useful representations from unlabeled data (e.g., millions of signal strength measurements without explicit path loss labels) or by creating artificial prediction tasks from the data itself.

6.7. Edge AI for Low-Latency and Scalable Inference:

- **Decentralized Intelligence:** Deploying trained AI propagation models closer to the data source (on base stations or edge servers) can significantly reduce inference latency and network backhaul requirements. This involves optimizing model size and complexity for resource-constrained edge devices while maintaining accuracy.

6.8. Standardized Datasets and Benchmarks:

- To accelerate research and enable fair comparisons between different AI models, the development of large, publicly available, standardized datasets (including diverse environmental data, various frequency bands, and different deployment scenarios) with clear benchmarking methodologies is crucial.

7. Conclusion

Accurate prediction of signal propagation loss is an enduring and critical challenge in the design, optimization, and operation of cellular communication networks. While traditional empirical and deterministic models have served their purpose, their inherent limitations—lack of adaptability, high computational cost, and site-specificity—are increasingly evident in the face of complex and dynamic 5G and future 6G environments.

Artificial Intelligence, particularly Machine Learning and Deep Learning, has emerged as a transformative paradigm for addressing these challenges. This review has highlighted the successful application of a wide range of AI techniques for predictive regression learning of path loss, from classic algorithms like SVR and Random Forests to advanced deep neural networks such as MLPs, CNNs, RNNs, and the promising Graph Neural Networks. These data-driven approaches offer enhanced accuracy, adaptability to diverse environments, and the ability to learn intricate non-linear relationships that elude predefined mathematical models.

References

[1] Rappaport, T. S. (2002). Wireless Communications: Principles and Practice. Prentice Hall PTR.
[2] Hata, M. (1980). Empirical formula for propagation loss in land mobile radio services. IEEE Transactions on Vehicular Technology, 29(3), 317-325.
[3] Faria, L. R., & Siqueira, G. L. (2011). A survey on ray tracing techniques for propagation prediction in wireless communication systems. Journal of Microwaves, Optoelectronics and Electromagnetic Applications, 10(2), 221-237.
[4] Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A survey on enabling technologies, protocols, and applications. IEEE Communications Surveys & Tutorials, 17(4), 2347-2376.
[5] Kulkarni, H., & Adkar, M. S. (2013). Support Vector Regression (SVR) for path loss prediction in urban environment. In 2013 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.
[6] Al-Samman, A. M., Al-Khawlani, M. M., & Mansor, M. H. (2018). Random Forest Regression for Path Loss Prediction in Suburban Environment. In 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE) (pp. 1-6). IEEE.

[7] Fan, J., Fu, X., Li, X., Wu, X., & Liu, Y. (2020). Path Loss Prediction Using XGBoost and Random Forest Algorithms for 5G Millimeter-Wave Communications. In 2020 IEEE International Conference on Communications Workshops (ICC Workshops) (pp. 1-6). IEEE.
[8] Popoola, S. I., & Vanu, D. I. (2019). Artificial neural network-based path loss prediction for urban microcell environments. Wireless Personal Communications, 107(1), 163-176.
[9] Huang, S., Yan, Q., & Li, Q. (2019). Deep CNN-based path loss prediction for urban environments. In 2019 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.
[10] Zhao, H., Wang, J., Zhang, C., & Zhang, J. (2021). Dynamic Path Loss Prediction Using LSTM Networks for 5G Millimeter-Wave Vehicular Communications. In 2021 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 1-6). IEEE.
[11] Gao, S., Lu, M., Zhang, H., & Zhou, Y. (2023). Graph neural networks for wireless channel modeling: A survey. IEEE Communications Surveys & Tutorials, 25(1), 589-623.
[12] Iqbal, U., Shahab, S., Ahmad, I., & Iqbal, N. (2017). Path loss prediction models for urban wireless environment at 2.4 GHz. IJCSI International Journal of Computer Science Issues, 14(2), 226-231.
[13] Popoola, O. P., & Vanu, D. I. (2019). Support vector regression for automated path loss prediction in urban microcell environment. Journal of Engineering and Applied Sciences, 14(10), 3465-3470.
[14] Abas, R. S., Abdulkarim, R. S., Al-Rawi, A. T., & Mahdi, D. Y. (2021). Path loss prediction using Random Forest and XGBoost in indoor and outdoor environments. Wireless Personal Communications, 118, 1629-1647.
[15] Liao, T., Song, Y., Han, J., & Li, C. (2020). Path loss prediction for urban millimeter wave communications based on ensemble learning. Journal of Communications and Information Networks, 5(2), 170-178.
[16] Ayadi, F., Fethi, A., & Turki, M. (2018). ANN-based path loss prediction in suburban environments at 2.4 GHz. In 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD) (pp. 1-5). IEEE.
[17] Vieira, V., Moreira, A., & Rodrigues, A. (2022). Deep Learning for Radio Propagation Prediction Based on Aerial Images. IEEE Access, 10, 24107-24120.
[18] Li, S., Zhao, W., Wang, C., & Zhang, R. (2022). Graph Convolutional Network for Wireless Channel Prediction. IEEE Transactions on Wireless Communications, 21(9), 7480-7493.
[19] Chen, C., Chen, S., & Li, C. (2020). A hybrid CNN-LSTM model for wireless channel prediction. IEEE Wireless Communications Letters, 9(12), 2097-2101.
[20] Tekbıyık, K., Ekti, A. R., Güvenç, İ., & Arslan, H. (2020). Learning-Based Channel Prediction for Millimeter-Wave Vehicular Communications with Ray Tracing. IEEE Vehicular Technology Magazine, 15(3), 108-117.
[21] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707.
[22] Lim, W. Y. B., Luong, N. C., Hoang, D. T., Jiao, Y., Liang, Y. C., Yang, Q., ... & Niyato, D. (2020). Federated learning in wireless networks: A comprehensive survey. IEEE Communications Surveys & Tutorials, 22(3), 2031-2063.
[23] Al-Tahrawi, H., Al-Jarrah, M. A., Al-Hammouri, M., Alsarhan, I., & Al-Qatamin, A. (2022). Digital Twin for 6G Wireless Networks: A Survey of Concepts, Challenges, and Opportunities. IEEE Access, 10, 10769-10787.
[24] Eisen, M., & Fekete, C. (2022). Transformers in Wireless Communications: A Survey. IEEE Communications Surveys & Tutorials, 24(2), 1148-1191.
[25] Zhang, X., Wang, Y., & Liu, H. (2015). Support Vector Regression for Signal Propagation Loss Prediction. IEEE Transactions on Wireless Communications, 14(5), 2867-2876.

[26]. J Isabona, VM Srivastava, A neural network based model for signal coverage propagation loss prediction in urban radio communication environment, International Journal of Applied Engineering Research 11 (22), 11002-11008, 2016.
[27]. V.C Ebhota, J. Isabona, V. M Srivastava, Base line knowledge on propagation modelling and prediction techniques in wireless communication networks, Journal of Engineering and Applied Sciences 13 (7), 1919-1934, 2018
[28]. V.C. Ebhota C, Isabona, J and Srivastava V.M (2018) Improved Adaptive Signal Power Loss Prediction Using Combined Vector Statistics Based Smoothing and Neural Network Approach, Progress in Electromagnetics Research C, Vol. 82, pp. 155–169, 2018.

[29]. Lee, J., Kim, S., and Park, J. (2018). Predicting Signal Loss in Urban Environments Using Artificial Neural Networks. IEEE Wireless Communications Letters, 7(3), 428-431.
[30]. V.C Ebhota, J Isabona, VM Srivastava, Effect of learning rate on GRNN and MLP for the prediction of signal power loss in microcell sub-urban environment, International Journal on Communications Antenna and Propagation 9 (1), 36-45, 2019.
[31]. J Isabona, AI Osaigbovo, Investigating predictive capabilities of RBFNN, MLPNN and GRNN models for LTE cellular network radio signal power datasets, FUOYE Journal of Engineering and Technology 4 (1), pp.155-159, 2019,
[32]. Isabona, J. Wavelet Generalized Regression Neural Network Approach for Robust Field Strength Prediction. *Wireless Pers Commun* 114, 3635–3653 (2020). https://doi.org/10.1007/s11277-020-07550-5.
[33]. J. Isabona, D. O Divine, Adaptation of Propagation Model Parameters toward Efficient Cellular Network Planning using Robust LAD Algorithm, International Journal Wireless and Microwave Technologies 10, 13-24, 2020.
[34]. Chen, Q., & Wu, L. (2020). Random Forest Model for Predictive Analysis of Signal Propagation Loss. Journal of Signal Processing Systems, 92(1), 31-41.
[35]. Joseph, I. Joint Statistical and Machine Learning Approach for Practical Data-Driven Assessment of User Throughput Quality in Microcellular Radio Networks. *Wireless Pers Commun* 119, 1661–1680 (2021). https://doi.org/10.1007/s11277-021-08300-x
[36]. Kumar, P., Gupta, A., & Singh, R. (2021). Application of LSTM Networks for Signal Propagation Loss Prediction. Future Generation Computer Systems, 115, 28-
[37]. O Ituabhor, J Isabona, JT Zhimwang, I Risi, Cascade forward neural networks-based adaptive model for real-time adaptive learning of stochastic signal power datasets, International Journal of Computer Network and Information Security 14 (3), 63-74, 2022.
[38]. Ituabhor Odesanya, Joseph Isabona, Emughedi Oghu and Okiemute Roberts, Omasheye Hybrid Empirical and Machine Learning Approach to Efficient Path Loss Predictive Modelling: An Overview, Int. J. Advanced Networking and Applications 5931 Vol: 15 Issue: 03 pp: 5931– 5939, 2023.
[39]. Smith, R., Green, T., & Brown, L. (2023). Gradient Boosting for Predictive Regression Learning of Signal Loss in Telecommunications. International Journal of Wireless and Mobile Computing, 29(2), 145-157.
[40]. Seyi E. Olukanni, Joseph Isabona, Ituabhor Odesanya, "Radio Spectrum Measurement Modeling and Prediction based on Adaptive Hybrid Model for Optimal Network Planning", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.15, No.4, pp. 19-32, 2023. DOI:10.5815/ijigsp.2023.04.02
[41]. Ugbeh R.N, Okiemute Roberts Omasheye, Abiodun I.C, Development of a Hybrid Machine Learning Path Loss Model for Cellular Networks in Maritime Environments Using Regression-Based Fusion. (2025). *Journal of Science Computing and Applied Engineering Research*, *1*(1), 34-42.
[42]. kechi Risi, and Konyeha, C. C, Deep Neural Network Based on Long Short-Term Memory for Predictive Learning of Wireless Path Loss Datasets, *Journal of Sciences, Computing and Applied Engineering Research (JSCAER), Vol. 1, No.2, pp. 23-30* .
[43]. Emughedi Oghu, Machine Learning Based on Exhaustive GMM Clustering Algorithm for Optimal Learning of 5G-NR SINR Datasets. (2025). *Journal of Science Computing and Applied Engineering Research*, *1*(1), 10-19.