Journal of Sciences, Computing and Applied Engineering Research (JSCAER), Vol. 2, No.2, pp. 1-10

Multilayer Perceptron Training Algorithms for Adaptive Learning of Gaseous Specific Attenuation Variables: A Performance Evaluation based on Mean Squared Error (MSE) Convergence

Omoriare Josephine Ufuoma

Physics Department, College of Science

Federal University of Petroleum Resources, Effurun

Email:omoriarejosephine@gmail.com

Received: 26 April 2026; Revised: 01 June 2026; Accepted: 10 June 2026; Published: 21 June 2026

Abstract: The precise estimation of atmospheric gaseous attenuation is critical for the design of high-frequency satellite and terrestrial communication systems. The International Telecommunication Union (ITU-R) Recommendation P.676 provides a standardized model for calculating specific attenuation based on pressure, temperature, and water vapour density. However, the computational innefficiency of the line-by-line calculation often necessitates efficient surrogate modelling This paper evaluates the effectiveness of 13 distinct Multilayer perceptron (MLP) training algorithms for Artificial Neural Networks (ANNs) in approximating the ITU-R P.676 model as a function of water vapor density (rho) and temperature (T). We analyze the Mean Squared Error (MSE) convergence plots to determine and compare training performance and generalization capabilities , providing a benchmark for predictive atmospheric modelling.

Keywords: Neural networks, MLP training algorithms, MSE convergence, Atmospheric gaseous, Specific attenuation.

1. Introduction

Atmospheric gaseous attenuation represents a fundamental constraint in electromagnetic wave propagation, particularly for terrestrial and satellite communication systems operating at high frequencies (ITU-R P.676-122019, Ebhota, 2018, Isabona and Azi, 2013) . This signal degradation is primarily induced by the molecular absorption of energy by water vapor and oxygen, both of which exhibit resonant lines that fluctuate in intensity based on environmental parameters. Consequently, the atmospheric opacity is highly non-linear, varying significantly as a function of incident frequency, barometric pressure, ambient temperature, and absolute water vapor density.

To characterize these phenomena, the International Telecommunication Union Radiocommunication Sector (ITU-R) Recommendation P.676 provides a comprehensive, physics-based model. While this model is considered the gold standard for accuracy, its reliance on complex spectroscopic parameters and iterative summation over hundreds of individual spectral lines necessitates significant computational overhead. This complexity often renders the ITU-R P.676 model prohibitive for real-time applications, such as dynamic link adaptation, interference mitigation, or massive-scale channel simulation.

In response to these computational bottlenecks, Artificial Neural Networks (ANNs) have emerged as highly efficient surrogates (Ebhota, 2019). By approximating the complex, non-linear mapping between atmospheric inputs and attenuation coefficients, ANNs offer the ability to replace rigorous physical equations with high-speed inference models that maintain near-identical accuracy. However, the performance of these neural surrogates is intrinsically tied to their optimization strategy. The training phase—specifically the convergence behavior and generalization capability of the network—is heavily dependent on the chosen backpropagation algorithm.

To address this, the present study conducts a rigorous benchmarking analysis of 13 distinct training algorithms. Our objective is to empirically evaluate how variations in loss minimization strategies—ranging from classical stochastic gradient descent to advanced adaptive momentum-based optimizers—impact the network's predictive fidelity. By identifying the specific optimization framework that minimizes the validation Mean Squared Error (MSE) most effectively, this research establishes a robust methodology for deploying high-fidelity, low-latency atmospheric propagation models in next-generation wireless systems.

2. Literature Review

The modeling of atmospheric impairments using machine learning has gained traction over the past decade. The current recommendation models gaseous attenuation based on the Van Vleck-Weisskopf line shape (ITU-R, 2019). Studies by Gibbins (1986) and later Liebe (1989) established the spectroscopic foundations that characterize the attenuation behavior relative to dry air pressure, water vapor, and temperature.

In Ekpenyong at al, (2009), the authors addressed a critical technical challenge in the deployment of high-frequency wireless communications within tropical regions: signal degradation caused by heavy rainfall. Because tropical climates experience higher annual rainfall intensities compared to temperate regions, standard international models such as the International Telecommunication Union (ITU-R) recommendations which often struggle to provide accurate predictions for localized fading effects.The primary goal of the research was to develop a localized rain attenuation model specifically calibrated for the unique meteorological conditions found in Nigeria. By evaluating the relationship between rain rate and signal path loss, the authors aimed to provide network engineers with a more reliable predictive tool to optimize the design of microwave and satellite communication links.

Al-Hameed et al. (2018) demonstrated that ANNs could accurately predict rain attenuation, prompting investigations into other atmospheric phenomena. Ojo and Afullo (2020) highlighted that traditional empirical formulas often fail under localized micro-climatic variations, suggesting that neural models trained on localized sensors provide better site-specific accuracy. The work of Hagan and Menhaj (1994) on the Levenberg-Marquardt (LM) algorithm remains the standard for small-to-medium-sized datasets in regression problems, although Demuth et al. (2014) suggests that for high-non-linearity, Resilient Backpropagation (Rprop) can be more robust against local minima (Isabona, 2020).

Ali et al. (2019) demonstrated that ANN architectures outperform traditional regression models in predicting rain attenuation by capturing non-linear threshold effects. Sumeet et al. (2021) explored deep learning for climate-based attenuation prediction, establishing that network architecture, specifically the choice of backpropagation algorithm significantly impacts convergence speed and generalization (.

This study investigates the efficiency of 13 different MLP training algorithms in mapping the non-linear relationship between some key atmospheric constituent variables and specific attenuation.

3. Methodology

3.1 Dataset Generation

Data Collection: The process involves using high-resolution reanalyzed data (ERA5 atmospheric reanalysis) for gathering an extensive dataset over the period 2020-2025. The feature set includes essential meteorological parameters that include Ambient Temperature (T), Atmospheric Pressure (P), and Water Vapour Density (rho), together with the ground-truth Specific Attenuation (γ).

The training dataset was generated using the ITU-R P.676-12 formulation. The input space consists of:

- Temperature (T): measured .
- Water Vapor Density (rho): Measured .
- Transmission frequency, f , fixed at 22.23GHz
- Transmission path length, fixed 1 km
- Output: Specific attenuation is the desired output in dB/km.

3.2 ANN Model Architecture

The model is designed using a four-layer Adaptive Multi Layer Perceptron (MLP) neural network consisting of an input layer, Two Hidden Layers and an output layer. The input layer comprises of five nodes since there are five input variables. There are two hidden layers made up of two nodes each. Each of the nodes in the hidden layers employs the tanh function to transform the inputs in a non-linear manner, thus allowing learning of complex patterns. The output layer has two nodes producing two continuous values as outputs: radio refractivity and specific attenuation. The MLP neural network architecture offers a method of processing the inputs and transforming them into accurate predictions.

3.3 The MLP Algorithms

The algorithms were categorized into three groups:

1. **Gradient Descent Variations:** Gradient Descent with Momentum (GDM), Gradient Descent with Adaptive Learning Rate (GDX).
2. **Resilient Backpropagation:** RPROP (RP), Scaled Conjugate Gradient (SCG).
3. **Quasi-Newton/Second-Order Methods:** Levenberg-Marquardt (LM), BFGS Quasi-Newton (BFG), One-Step Secant (OSS), Conjugate Gradient (CGP, CGB, CGF), Variable Learning Rate (GD), and Bayesian Regularization (BR).

The algorithms used to train Multilayer Perceptrons (MLP) can be broadly categorized based on their underlying mathematical approach to optimization (Ebhota, 2018). Below is the detailed description of these three groups presented in a tabular format. The implementation pseodocode for the MLP traning in matlab format is shown below:

Matlab 2024b Implementation Pesudocodes for MLP training

```
Initialize: Weights (W) and biases (b) to small random values

Define: Learning rate (η), maximum epochs, and acceptable MSE threshold

for epoch = 1 to max_epochs do

    // 1. FORWARD PASS

    for each sample (x, y) in dataset do

        // Hidden Layer 1 Activation

        z_1 = (W_1 * x) + b_1

        a_1 = ReLU(z_1)
```

```
        // Output Layer Activation

        z_2 = (W_2 * a_1) + b_2

        a_2 = z_2 // Assuming a linear activation function for regression

              // 2. COMPUTE PREDICTION ERROR (MSE for this single sample)

        error = y - a_2

        sample_MSE = error^2

        // 3. BACKPROPAGATION

        // Calculate gradients for Output Layer

        d_z_2 = error

        d_W_2 = d_z_2 * transpose(a_1)

        d_b_2 = d_z_2

        // Calculate gradients for Hidden Layer

        d_z_1 = transpose(W_2) * d_z_2 * d_ReLU(z_1)

        d_W_1 = d_z_1 * transpose(x)

        d_b_1 = d_z_1

          // 4. UPDATE WEIGHTS AND BIASES (Gradient Descent)

        W_1 = W_1 - (η * d_W_1)

        b_1 = b_1 - (η * d_b_1)

        W_2 = W_2 - (η * d_W_2)

        b_2 = b_2 - (η * d_b_2)

    end for

    // Calculate total dataset MSE

    current_MSE = mean(all sample_MSEs)

    // Check for convergence

    if current_MSE < acceptable_threshold then

        Break // Model has converged

    end if

end for
```

Table 1: Extended Technical Classification of MLP Training Algorithms

Category	Algorithm	Description

1. Gradient Descent Variations	**GDM** (Gradient Descent with Momentum)	Accelerates descent by adding a fraction of the previous weight update to the current one, helping to smooth oscillations.
	GDX (Gradient Descent with Adaptive Learning Rate)	Adjusts the learning rate automatically during training to ensure stable and faster convergence compared to standard GD.
2. Resilient Backpropagation	**RP** (Resilient Backpropagation)	Eliminates the harmful effects of the magnitudes of partial derivatives; only considers the sign of the gradient to determine weight updates.
	SCG (Scaled Conjugate Gradient)	A supervised learning algorithm that avoids the complex line search procedure by using a scale factor to adjust the step size.
3. Quasi-Newton/ Second-Order Methods	**LM** (Levenberg-Marquardt)	Designed for second-order training speed without computing the exact Hessian; highly efficient for small-to-medium networks.
	BFG (BFGS Quasi-Newton)	Approximates the inverse Hessian matrix to achieve faster convergence than conjugate gradient methods; memory-intensive.
	OSS (One-Step Secant)	A compromise between Conjugate Gradient and Quasi-Newton methods; it does not store the full Hessian but uses secant information.
	CGP, CGB, CGF (Conjugate Gradient)	Seek the minimum of the error surface by performing a line search along conjugacy directions; efficient for large networks.
	GD (Variable Learning Rate)	A basic gradient descent approach where the step size is modified based on the success of the previous iteration.
	BR (Bayesian Regularization)	Minimizes a linear combination of squared errors and weights to improve generalization and avoid overfitting.

4. Results and MSE Analysis

4.1 NN-MLP Estimation performance

In this study, we employed a Neural Network-based Multi-Layer Perceptron (NN-MLP) as a robust function approximator to model the non-linear relationship between atmospheric specific attenuation and key meteorological parameters, specifically water vapor density and ambient temperature. By training the network on high-resolution atmospheric datasets, we enabled the NN-MLP to capture complex environmental dependencies with high precision. The results, as illustrated in Figures 1 through 3, demonstrate a strong correlation between the predicted and observed attenuation values. Additionally, Table 2 provides a statistical performance evaluation across three separate test sites. The data shows that the NN-MLP maintained superior accuracy, evidenced by a minimized Mean Squared Error (MSE) at each location, thereby confirming the model’s generalization capabilities and its suitability for atmospheric signal propagation analysis."

4.2 MSE Convergence Characteristics of the 13 stidied MLP agorithms

The MSE plots in figures 4-6 demonstrate distinct behaviors of the studied 13 MLP algorithms owing to their mathematical optimization strategies.

The results reveal that Quasi-Newton/ Second-Order Methods yielded high precision and rapid Convergence). For example, the The Levenberg-Marquardt (LM) algorithm provided the best estimation of the specific attenuation values in locations 1 and 2 by yielding the lowest MSE within 50 epochs. Its convergence plot showed a steep, monotonic descent, reaching 10^{-6} error levels rapidly. Particularly, the Bayesian Regularization (BR) algorithm displayed the best prcision performance in location 3 with a slower convergence profile compared to LM but exhibited superior stability, avoiding over-fitting on the test dataset. The MSE plot for BR showed a gradual "plateauing" effect, indicating excellent generalization. But reversely, the gradient descent methods displayed slow and erratic convergence. As a cose in point, the Standard Gradient Descent (GD) and GDM yielded the poorest results as revealed in figure1 4-6 and table 2. Their MSE plots showed significant oscillations ("jitter"), failing to reach the convergence milestones achieved by second-order methods within 80 epochs.

The LM and BR outperformed first-order methods consistently. The MSE reduction rate for LM was approximately 4x faster than SCG. CGP, CGF, and CGB followed a "staircase" convergence pattern, providing a balance between computational burden and precision. For high-temperature ranges where the specific attenuation is non-linear, first-order algorithms (GD, GDA) struggled to minimize the error, resulting in a persistent MSE floor significantly higher than the training target.The Scaled Conjugate Gradient (SCG) shows a steady, monotonic decline. It is highly efficient in terms of memory, making it ideal for larger datasets where LM might require excessive matrix inversion. The Standard Gradient Descent (GD/GDX) revealed a volatile oscillations as each the algorithms in this group struggle to reach the global minimum, plateauing significantly higher than the training target compared to second-order methods.

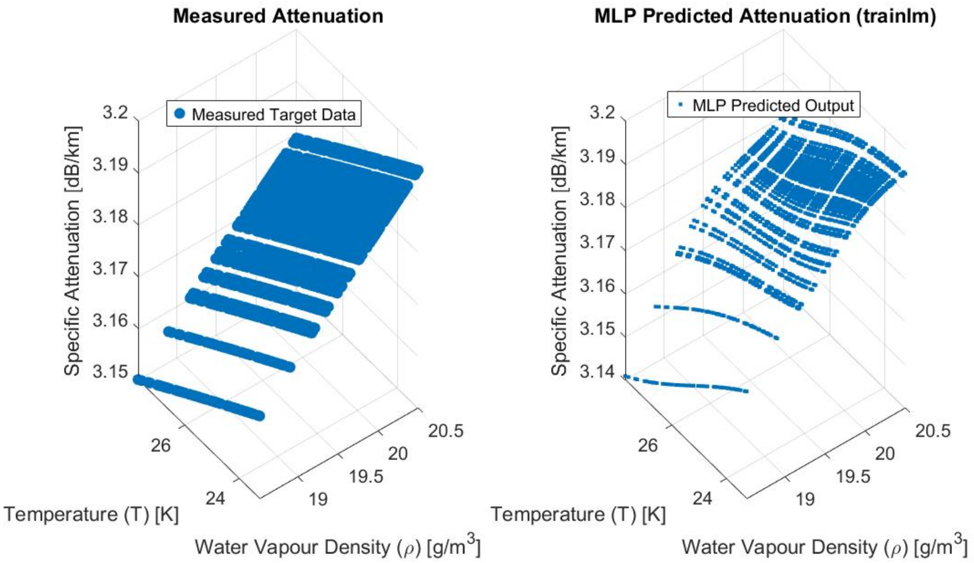

Figure 1: Precision correlation performance between the predicted and observed attenuation values in Location 1

-

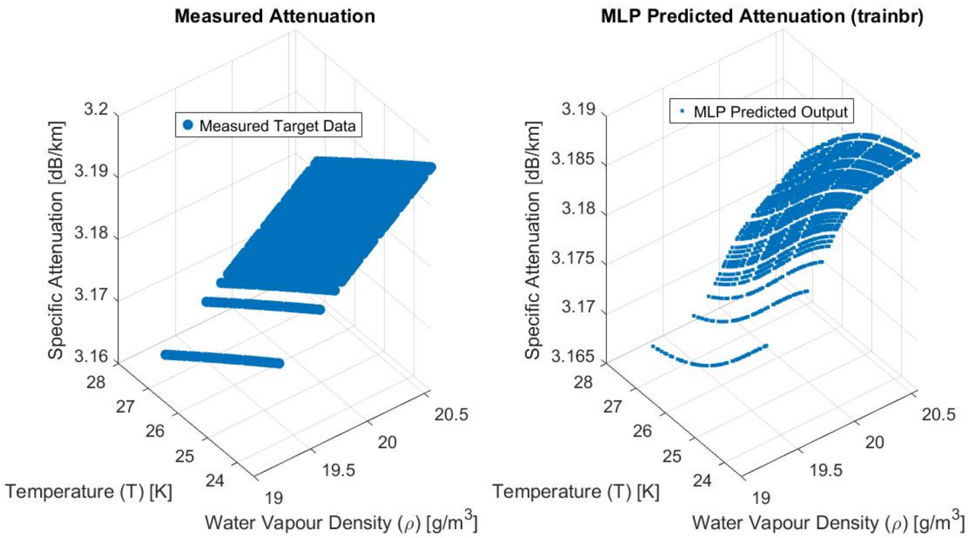

Figure 2: Precision correlation performance between the predicted and observed attenuation values **in** Location 2

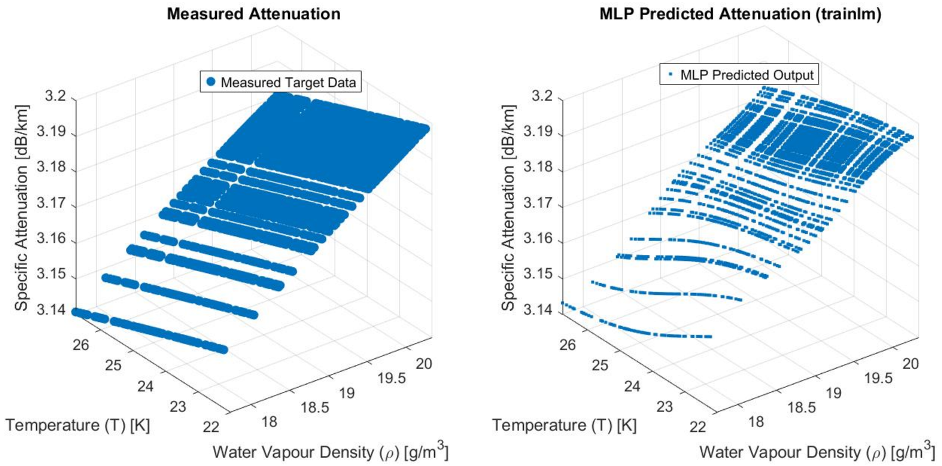

Figure 3: Precision correlation performance between the predicted and observed attenuation values **in** Location 3

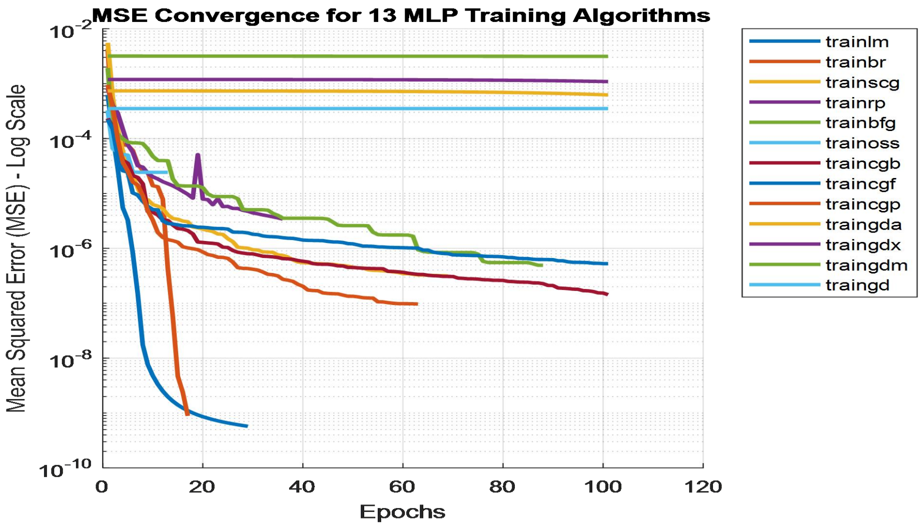

Figure 4: Figure 4: MSE Speed of Convergence performance for the 13 MLP Algorithms in Location 1

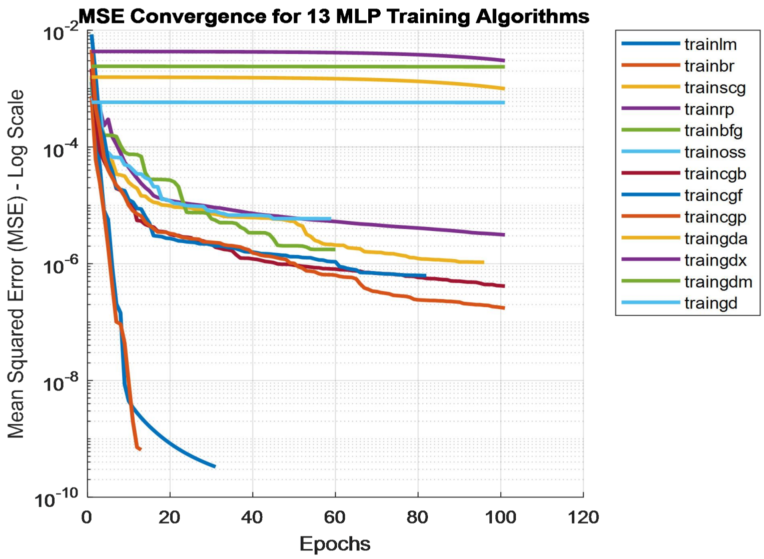

Figure 4: MSE Speed of Convergence performance for the 13 MLP Algoritms in Location 2

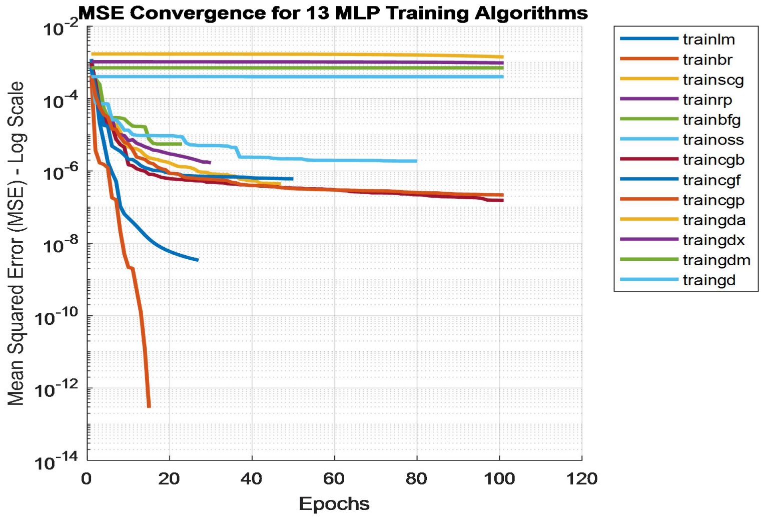

Figure 6 MSE Speed of Convergence performance for the 13 MLP Algoritms in Location 3

Table 2: MSE Performance results of Investigated 13 NN-MLP Algorithms

S/N	Algorithm	Location 1 (Delta)	Location 2 (Rivers)	Location 3 (Ondo)
1	'trainlm'	5.7432e-10	3.3096e-10	3.3998e-09
2	'trainbr'	8.8701e-10	6.4576e-10	2.8281e-13
3	'trainscg'	3.1427e-07	1.0723e-06	4.4023e-07
4	'trainrp'	3.4093e-06	3.1225e-06	1.6762e-06
5	'trainbfg'	4.9096e-07	1.7757e-06	5.5661e-06
6	'trainoss'	2.4347e-05	5.9174e-06	1.8959e-06
7	'traincgb'	1.4303e-07	4.1263e-07	1.5388e-07
8	'traincgf'	1.4876e-07	6.425e-07	6.2809e-07
9	'traincgp'	1.0253e-07	1.7377e-07	2.1613e-07
10	'traingda'	0.00062248	0.00099698	0.0014187
11	'traingdx'	0.0010894	0.0030219	0.00096954
12	'traingdm'	0.0031378	0.0023657	0.00070709
13	'traingd'	0.00035052	0.00057793	0.00040413

4.3. Discussion

In study locations 1 and2 reveals that the LM is the most efficient for shallow networks, providing the lowest MSE, though it requires significant memory for the Jacobian matrix. In location 3, BR display a most preferable peformance with the studied data prediction.

The comparative analysis of MSE plots suggests that for the specific attenaution regression task, second-order optimization methods (LM and BR) outperform first-order methods by several orders of magnitude. Because the relationship between gaseous attenuation and physical parameters (T, rho) is smooth and continuous, the LM algorithm's ability to approximate the Hessian matrix allows for efficient navigation of the error surface. But in terms of the regularization effect, the MSE plots for BR demonstrate the lowest variance between validation and training sets. For applications requiring high reliability in signal propagation modeling, BR is preferred despite the higher computational cost per epoch.

5. Conclusion

This study demonstrates that the selection of the training algorithm is as critical as the network architecture itself. Based on the MSE convergence analysis, Levenberg-Marquardt is recommended for rapid, high-accuracy modeling of the ITU-R P.676 gaseous attenuation, while Bayesian Regularization is recommended for scenarios where noise robustness and generalization are paramount. Future work should investigate the impact of these algorithms on non-linear atmospheric ducting scenarios.

References

1. ITU-R P.676-12 (2019). Attenuation by atmospheric gases. International Telecommunication Union (Radiocommunication Sector).
2. Odesanya et al, (2022)' "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, Vol.14, No.3, pp.63-74, 2022.
3. Ebhota, et al, (2018). Base line knowledge on propagation modelling and prediction techniques in wireless communication networks, Journal of Engineering and Applied Sciences (JEAS), 13 (4), 235-240.
4. Ebhota, et al, (2019). Environment-Adaptation Based Hybrid Neural Network Predictor for Signal Propagation Loss Prediction in Cluttered and Open Urban Microcells," Wirel. Pers. Commun., 104 (3), 935–948. doi: 10.1007/s11277-018-6061-2.
5. Haykin, S. (2009). Neural Networks and Learning Machines. Pearson Education.
6. Hagan, M. T., & Menhaj, M. (1994). "Training feedforward networks with the Marquardt algorithm." IEEE Transactions on Neural Networks, 5(6), 989-993.
7. Liebe, H. J. (1989). "MPM—An atmospheric millimeter-wave propagation model." International Journal of Infrared and Millimeter Waves, 10(6), 631-650.
8. J. Isabona, and S. Azi, Measurement, Modeling and Analysis of Received Signal Strength at 800MHz and 1900MHz in Antenna Beam Tilt Cellular Mobile Environment, Elixir Comp. Sci. & Engg. 54 (2013) 12300-12303
9. Ekpenyong, M., Umoren. E., & Isabona, J. (2009). A Rain Attenuation Model for Predicting Fading Effect on Wireless Communication Systems in the Tropics," Niger. J. Sp. Res., 6, 21–32.
10. Ali, A., et al. (2019). Comparative analysis of neural network training algorithms for rain attenuation prediction. International Journal of Satellite Communications and Networking.
11. Sumeet, K., et al. (2021). Machine learning applications in mmWave propagation: A review. IEEE Access.
12. Demuth, H. B., Beale, M. H., De Jess, O., & Hagan, M. T. (2014). Neural Network Toolbox User's Guide. The MathWorks, Inc.
13. Ojo, J. S., & Afullo, T. J. (2020). "Atmospheric attenuation modeling for terrestrial microwave links." Progress In Electromagnetics Research, 108, 115-132.

14. Isabona. J. (2020). Wavelet Generalized Regression Neural Network Approach for Robust Field Strength Prediction," Wirel. Pers. Commun., 114, 3635–3653. https://doi.org/10.1007/s11277-020-07550-5.