Adaptive Hybrid Neural Network Predictor over a Multi-layer Perceptron Neural Network Model for Improved Electromagnetic Signal Power Loss Prediction

Virginia C. Ebhota * and Thokozani Shongwe

Department of Electrical and Electronic Engineering Science, University of Johannesburg, Auckland Park, Johannesburg, South Africa
Email: engrchikaugwu@yahoo.com (V.C.E.); tshongwe@uj.ac.za (T.S.)
*Corresponding author

***Abstract*—This research paper presented an adaptive hybrid system predictor that employed an adaptive Linear Mean Square (LMS) filtering technique for dataset de-noising, passed through a combined Adaptive Linear Element (ADALINE) and Multi-Layer Perceptron Neural Network (MLP) for improved neural network training and prediction over a conventional MLP model. The neural network training performances of the designed filtered adaptive hybrid system predictor showed improved training and prediction results over the conventional MLP using measured data from Long-Term Evolution (LTE) microcell environment from a Line-of-Sight (LOS), termed as location-1, and a Non-Line-of-Sight (NLOS), termed as location-2. The neural network models' training results analysis was carried out using 1st order statistical performance indices, the coefficient of Regression (R), the Root Mean Square Error (RMSE), the Mean Squared Error (MSE), the Standard Deviation (SD), and the Mean Absolute Error (MAE). The statistical performance indicators measured the closeness of the prediction values to the measured values during neural network training, using two training algorithms, the Levenberg-Marquardt (LM) and the Bayesian Regularization (BR) training algorithms. The output of the designed adaptive hybrid system predictor showed superior and optimal prediction of the measured dataset over the conventional MLP for both the LOS location-1 and the NLOS location-2. The prediction results also demonstrated better prediction of the neural network models using the BR training algorithm over the LM training algorithm. However, this comes with a certain tradeoff, such as increased training time for the BR training algorithm.**

***Keywords*—dataset de-noising, least mean square filter, adaptive linear element, multi-layer perceptron neural network model, 1st-order statistical performance indices, training algorithms, signal power loss**

Manuscript received June 12, 2025; revised July 19, 2025; accepted July 28, 2025; published March 27, 2026.

I. INTRODUCTION

Planning of radio network coverage is fundamental for the reduction of operation and capital expenditure in the deployment of communication networks, as well as in the upgrade and maintenance of the existing ones [1]. Precise prediction of signal power loss between the transmitting Base Station (BS) and the receiving Mobile Station (MS) is a crucial need for enhanced Quality of Service (QoS) for enhanced communication links [1, 2].

Over the years, various Machine Learning (ML) methods have been applied to model radio system networks to enhance propagation performance and reduce errors [3]. One such ML method is Artificial Neural Network (ANN) models. These models are used to solve different functional and optimization problems [4]. One type of ANN is the Multi-layer Perceptron Neural Network (MLP), a feed-forward network that utilizes the Back-Propagation (BP) algorithm. Although MLPs offer benefits such as ease of use due to their simple and flexible architecture, they also face challenges, including difficulty in selecting an appropriate architecture for a specific problem, since choosing the right network design is often based on trial and error. This can result in either poor network generalization or overfitting during training [4].

Effective modeling and accurate signal coverage predictions are crucial for the efficient use of limited resources, driven by the increasing demand for mobile and fixed cellular telecommunication services due to the scarcity of radio frequency spectrum [5]. Reliable predictive modeling of signal power loss helps control the load on Base Station (BS) transmitters and supports the design of efficient radio network channels with minimal interference and coverage issues. Conventional neural network models, such as the MLP-ANN model, are often limited in their predictive capabilities because they rely on "trial and error" in selecting the architecture suited to the specific problem. Therefore, their effectiveness is generally restricted to the environment they were designed for and may not perform optimally in different settings.

Various conventional ANN architectural compositions have been applied in several research works for resolving different optimization and functional problems. These include Adaptive Linear Element (ADALINE), MLP, Generalized Regression (GR), Radial Basis Function (RBF), Hopfield neural networks, etc. [5−8]. Adaptive linear element ANN architectural network comprises a single linear neuron, though both ADALINE and perceptron can solve linearly separable problems, ADALINE uses the Least Mean Square (LMS) learning rule that is more powerful than the perceptron learning rule [9]. Though ADALINE can only solve linearly separable problems, it has demonstrated efficiency in adaptive modeling, system selection, and noise cancellation with notable features such as fast learning and simple usage [10, 11]. It is effective in minimizing the Mean Squared Error (MSE) and moves decision boundaries as far as possible from the training pattern. A multi-layer perceptron neural network is a feed-forward neural network model that comprises of an input layer, one or more intermediate layers, and an output layer [8, 12]. The MLP has been applied in solving many optimization problems; however, the selection of its appropriate architectural network for a particular problem is by trial and error and has left research on the best architectural composition of MLP as an open problem. Multi-layer perceptron neural network shows overlearning with the application of an inappropriate network architectural composition. Also, applying a single intermediate layer MLP for system modeling has demonstrated poor performance in de-noising datasets as it lacks competence in the handling of incoherent datasets [13, 14].

Dataset quality degradation is a result of noise during data collection, which is from electronic equipment and other sources adopted for the data collection [15]. There is always an introduction of noise during data transmission due to noisy channels and errors in the measurement process, and quantization.

Recently, the studies of data pre-processing in the enhancement of ANN training efficiency and precise prediction have been under investigation [15, 16]. Dataset de-noising ensures improvement of the quality of the dataset, and this can effectively be achieved through the application of various techniques such as filtering, wavelength analysis, and multi-fractal analysis [16]. All the techniques have their advantages and disadvantages. The multi-fractal method operates based on the Holder regularity of the corrupted dataset, while the wavelength technique is based on a threshold. Filters play an important role in the restoration process of the dataset, employing the principle of convolution and the moving window [17]. There are mainly two types of filters: linear and non-linear filters. The linear filters are made up of the LSM filters and the mean filters, while the non-linear filters are made up of the median filters [17].

Several previous research studies have analyzed the performance of various hybrid ANN models in signal power loss prediction for cellular networks [8]. A hybrid ANN that combines a long-distance prediction model and a multi-layer perceptron network for the prediction of signal power loss was examined in Refs. [18−20]. In Ref. [8], a hybrid neural network that combined ADALINE and a multi-layer perceptron network was applied in the prediction of signal power loss for radio networks, and a relative study of the prediction performances of the hybrid neural network of ADALINE and multi-layer perceptron network on simulated Code Division Multiple Access (CDMA) systems presented in Ref. [21]. The results from these research works demonstrate enhanced prediction power and abilities of the hybrid ANNs over the performance outputs from conventional ANNs.

In this research work, an adaptive hybrid neural network that combines an LMS filter for dataset denoising is designed in combination with ADALINE and MLP neural networks. The adaptive hybrid system predictor combined the strengths of ADALINE and MLP in its design. A similar approach has been studied in Ref. [22], but for noise filtering, noise reduction, and prediction in the Rayleigh channel, with the authors' emphasis being on efficient noise reduction with the simulated sequential dataset. In this research work, the models are empowered with the grid search-based hyperparameter tuning techniques for optimal signal power loss prediction between the BS and the MS path lengths. The hyperparameters considered include the number of neurons and the number of intermediate layers. In detail, the developed adaptive hybrid system predictor and the considered MLP-ANN prediction accuracy levels employing the LM and the BR training algorithms with tuned best values of the hyperparameters were recorded using the 1st order statistical performance indicators for the results analysis and highlighted in yellow color.

The organization of this research work is as follows: Section II discusses related works on dataset pre-processing, such as data denoising. Section III describes the concept of dataset de-noising and the components of the designed adaptive system predictor. Section IV deals with dataset measurement, collection, and materials involved. Two training algorithms were used for training the designed model: the Levenberg-Marquardt (LM) and the Bayesian Regularization (BR) training algorithms. Section V discusses these mathematical concepts of training algorithms. The training results are discussed in Section VII, while Section VIII is the conclusion.

II. Related Works on Dataset Pre-processing

The need for dataset pre-processing before neural network learning and training is for enhanced training efficiency and prediction precision. Various dataset pre-processing techniques have been utilized by researchers over the years [19, 20]. These include dataset normalization for better data access, sampling for bulk dataset representation, transformation for dataset manipulation to generate a single input, and dataset denoising for the removal of noise from the dataset [18−20]. In terms of dataset de-noising, some of the previous research works utilized different de-noising methods such as wavelength multi-resolution analysis, singular spectrum analysis, etc. [16−18]. Isabona and Srivastava [10], Anysz *et al.* [12], and Ebhota *et al.* [19] applied a wavelength de-noising technique and an ANN model in the prediction of a rainfall time series dataset. The output result demonstrated that the combined wavelength and ANN model-based prediction technique was more efficient than the application of conventional ANNs. Deme [21] and Ebhota and Shongwe [22] used a similar approach of wavelength combined with ANN models; however, these are for enhanced prediction of earthquake datasets and for underground water levels.

Three dataset pre-processing methods that involved filtering-based Singular Spectrum Analysis (SSA) and Moving Average (MA) with modular neural networks for improved prediction of rainfall datasets in China and India were examined and presented [14]. The efficiency of dynamic filtering and convolution on the prediction accuracy of video and stereo was shown. In [4], a dataset pre-processing-based modeling technique was applied for daily reservoir inflow examination. The results showed model-driven prediction accuracy advancement of the uneven flow of the reservoir on application of a pre-processed dataset. Jia *et al.* [15] investigated the combination of a linear model and Kalman filters in the prediction of missing occurrences in a time series dataset sensor stream. The findings show that the application of a linear filter model is a feasible technique for the enhancement of the prediction efficiency of the sensor dataset. A comparable linear data filtering technique was applied for the enhancement of dataset analysis and prediction for radio network coverage for the CDMA 2000 signal dataset [7]. In Ref. [7], a pre-processing method that employed vector order statistical filters was applied for enhanced adaptive trend in the stochastic noisy data prediction employing neural network models. The work output results showed information content dataset improvement through dataset denoising, and thus enhanced training process and prediction accuracy using neural network models.

The concentration of authors has been mainly on time series datasets with linear smoothing dataset filtering methods. This does not exhaustively capture the stochastic non-linearity of some of the multi-faceted special datasets. In Refs. [12, 22], a hybrid neural network that combined multi-layer perceptron neural networks and conventional log-distance predictors was investigated for signal power loss. The research output result showed enhanced signal power loss prediction, using the hybrid multi-layer perceptron in combination with a long-distance predictor over conventional adopted neural network models. The accuracy of a hybrid neural network that combined ADALINE and a multi-layer perceptron neural network was examined in Ref. [8], using the LTE dataset, with the output result showing excellent performance of the hybrid system predictor over trained and tested conventional neural networks. In Ref. [7], an enhanced adaptive hybrid predictor called the vector media filter was developed for training the stochastic noisy signal power dataset. The pre-processing of the dataset was carried out by applying a media filter before neural network training, with excellent performance output in comparison to neural network training using a conventional multi-layer perceptron neural network.

In Ref. [7], An improved adaptive predictor, termed Vector Media Filters-MLP (VMF-MLP), was developed for training stochastic noisy signal power data. The dataset was pre-processed using a media filter before training with MLP-ANN. The result shows its superior performance in comparison to conventional MLP-ANN.

Our focus in this work is to design and investigate the performance of an adaptive hybrid system predictor that combines a LMS filter, an ADALINE and a MLP model for adaptive learning and prediction of signal power loss, using a signal measured dataset from a Long Term Evolution (LTE) of two micro-cell built-up environments, Line-of-Sight (LOS), location-1, and Non-Line-of-Sight (NLOS location-2. The work designed an easy, efficient, and adaptive hybrid neural network model for signal power loss prediction and system optimization with minimal error by application of an enhanced dataset through de-noising, using an LMS filter.

Thus, the main objective of this work is to design an adaptive hybrid system predictor with a well-structured architectural implementation empowered with the grid search-based hyperparameter tuning technique for optimal signal power loss prediction and approximation between the BS and the MS path lengths. The degree of the prediction accuracy of the developed adaptive hybrid system predictor over the conventional MLP-ANN is also clearly provided using the first-order statistics performance indices in the tables and graphs.

III. The Hypothesis of Dataset Denoising

The linear operation in dataset de-noising requires noise addition or noise multiplication to the dataset to acquire a corrupted dataset that is subject to de-noising [17].

If *s(t)* is an input dataset that is subjected to filtering, and *y(x)* is a filtered output dataset, the output filtered dataset can be stated as [17]:

$$y(x) = \int s(t)h(t)dt \tag{1}$$

h(t) is a point spread function. The integral function stands for the convolution integral that can be stated as [17]:

$$y = w \otimes h \tag{2}$$

For a two-dimensional instance of an input with an output result, it is stated as [9]:

$$y(i,j) = \sum_{t=i-k}^{i+k} \sum_{u=j-1}^{j+1} s(t,u)h(i-t, j-u) \tag{3}$$

where $h(t, u)$ are filter's weights and (i, j, k) are weight matrix elements. The overall output results are determined through a series of shift-multiply-sum operations that shape discrete convolution.

Dataset de-noising can be attained by the application of LMS adaptive filters for non-stationary datasets and the adjustment of the parameters during dataset scanning to match the dataset generation mechanism [9, 19]. The theory of the adaptive filtering technique is the linear combination of a stationary low-pass dataset and a highpass stationary component via a weighting function. The LMS adaptive filters combine a local mean estimator that operates on the concepts as stated in the Eqs. (4−7) [22]. A window 'W' with size $m \times n$, scanned over the dataset, the mean of the window is 'μ', which is deducted from the window element to get the residual matrix 'W^r'.

Where the weighted sum is calculated correspondingly to the mean filter as described in Ref. [17]:

$$\tilde{Z} = \sum_{i,j\in} h(i,j)W^r \quad (4)$$

where $\tilde{Z}$ is the weighted sum, μ is the mean of the window under the filter that replaces the center element of the window, and h (i, j) is the weight matrix element. The subsequently modified pixel value is Ref. [17]:

$$Z = \tilde{Z} + \mu \quad (5)$$

For the succeeding iteration, there is a movement of the window over one pixel in the row, main order adjusts the matrix weight. The deviation *'e'* is calculated as the difference between the center and the matrix *'W'*.

$$e = W^r - \tilde{Z} \quad (6)$$

The Eigenvalue (λ) of the inventive window is calculated from the autocorrelation matrix of the studied window. λ application in the computation of the weight matrix modification for the next.

$$hk + 1 = hk + \eta XeXW^r \quad (7)$$

where hk is the weight matrix of the preceding iteration, and there is process continues until the window covers the entire dataset.

IV. Design of the Proposed Model

The proposed architectural model design for this research work comprised an LMS filter adopted for de-noising of the measured dataset, an ADALINE neural network model used for further noise cancellation and for fast dataset training, and an MLP model, which is a non-linear feed-forward model applied for re-training of the measured dataset. Since the ADALINE neural network only comprises linear decision boundaries, the measured de-noised dataset using the LMS filter after training using the ADALINE neural network was re-trained with an MLP, which trains using back-propagation algorithms, making use of a non-linear activation function. Fig. 1 is a flow diagram of the proposed model in comparison with the MLP-ANN training results, while Fig. 2 is the architectural combination of an LMS adaptive filter, an ADALINE neural network, a linear predictor, and an MLP-ANN, a non-linear predictor having a structural delay element Q^{-1} that permits input dataset scaling and re-sampling. In summary, the work designs:

- A distinctive adaptive hybrid system predictor that employs an adaptive Linear Mean Square (LMS) filtering technique for dataset de-noising, passed through a combined Adaptive Linear Element (ADALINE) and Multi-layer perceptron Neural Network (MLP) based signal power loss prediction model with a well-structured implementation network architecture embedded with the appropriate transfer function, learning algorithms, neuron numbers for improved neural network training and prediction, and optimal approximation of path loss between the base station and the mobile station path length.
- The developed model was tested and validated for realistic signal power loss prediction using extensive experimental signal attenuation loss datasets, which were acquired from a field test drive conducted over Long Term Evolution (LTE) in a microcell urban area, and tested using the first-order statistical performance indicators.
- The optimization of the projected model through hyperparameter tuning leverages the grid search analyses of the experimental signal power loss data.

Optimal prediction efficiency of the developed adaptive hybrid system predictor with a structured implementation architecture was compared with a conventional MLP network model using various first-order statistical performance indicators.

The ADALINE input layer is from the LMS adaptive filter applied for dataset de-noising, the output layer is $O(p)^n$ is in a way that $(Ip) = 1,2,3...,n$, and is the number of signal power samples, i.e., the dataset number for prediction, then the ADALINE output is stated as:

$$O(P)^n = \sum_{p=1}^{n} w_p\,(Ip + n) \quad (8)$$

where $(Ip + n)$ is the actual dataset input value, W_p is the weighted sum of the ADALINE linear predictor that is adaptively adjusted in the course of the training process while employing the LMS law and linear activation function. The ADALINE training process was carried out using the LMS law that applies the technique of Minimum Mean Square Error (MMSE) at the neuron input. This results in:

$$E_p = \frac{1}{P}\sum_{p=0}^{n}[w_p\,I(p+1) - O(P)^n]^2 = \frac{1}{P}\sum_{p=0}^{n} e_p^2 \quad (9)$$

The prediction results output of the ADALINE neural network, passed through an adaptive filter, $O(p)^n$ was fed as input to the MLP, with its output given as:

$$y = \sum_{p=1}^{n} O(P)^p \quad (10)$$

The MLP-ANN was trained using a nonlinear transfer function, the tangent transfer function. This is given as stated in Eq. (11) and the overall training output from the designed adaptive hybrid system predictor stated as in Eq. (12).

$$\phi(p) = tanh(y(p)) \quad (11)$$

The overall output result from the neural network training using the designed adaptive hybrid system predictor on the application of a non-linear transfer function is given as:

$$\phi(p) = \tanh \sum_{p=1}^{n} O(P)^p(p) \quad (12)$$

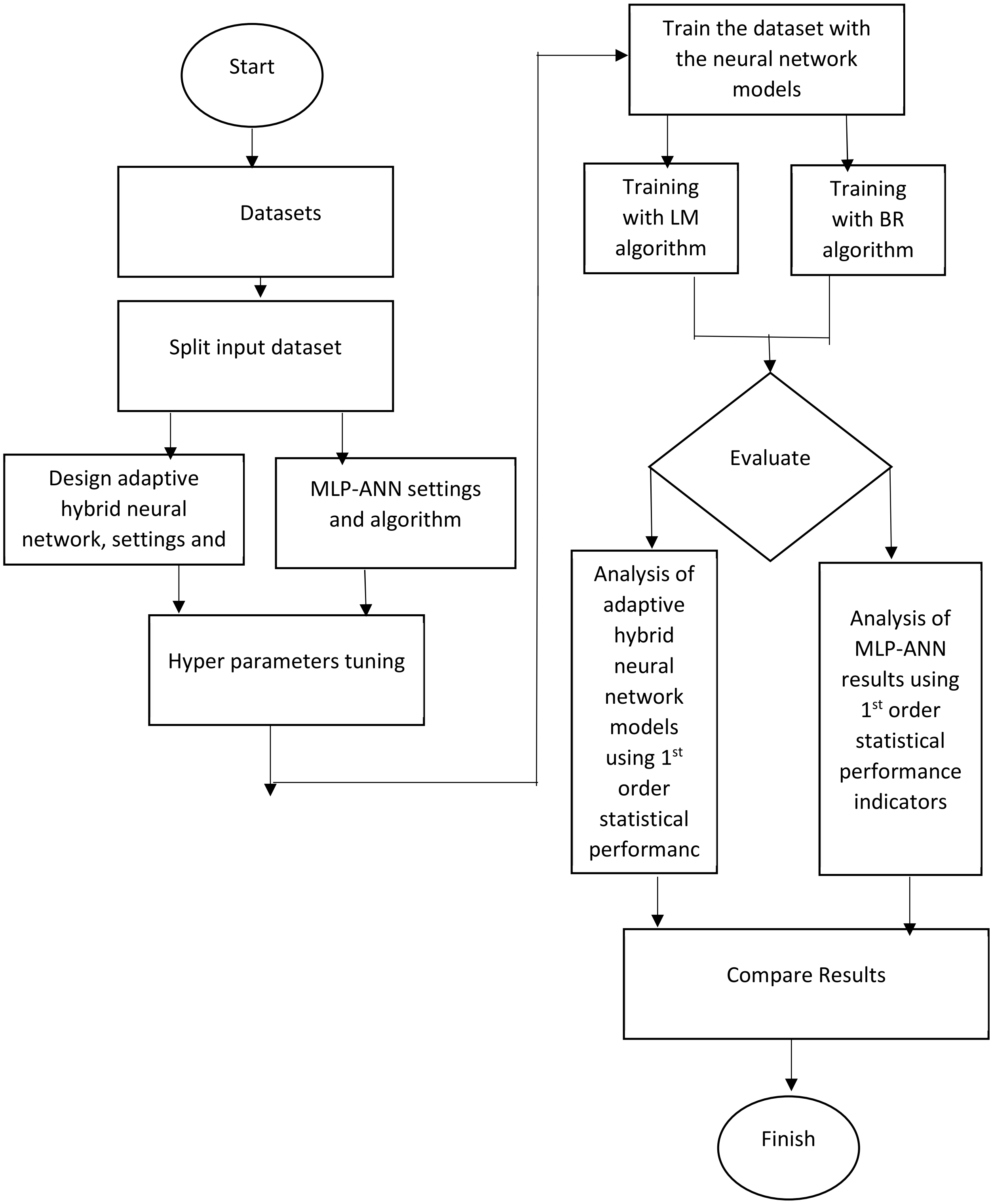


Fig. 1. Flow diagram of the de-noised dataset training of the adaptive hybrid system predictor and the MLP-ANN models.

The neural network training for the designed adaptive hybrid system predictor was carried out using two back-propagation training algorithms. The LM training algorithm and the BR training algorithm. The training of the conventional MLP model was also carried out using the same sample of datasets used for training the designed adaptive hybrid system predictor, and a comparison of the two models' training performances was made. The two training algorithms are discussed.

A. The Levenberg-Marquardt and the Bayesian Regularization Training algorithms

The LM training algorithm is an iterative training algorithm that guarantees the reduction in the performance function of all iterations in the course of network training [22]. As a result of its features, it demonstrates fastness in the training of neural network models; however, it has difficulty with memory and computational overhead due to gradient and Hessian matrix calculations [12]. The LM training algorithm updates the weight as [8]:

$$w(Ip+1) = w(Ip) - (H - \eta 1)^{-1} \Delta E(w(Ip)) \quad (13)$$

where $(J(w)Ip(op^n) = \frac{\partial e_{Ip}}{\partial wop^n}$ is the Jacobian matrix of E in correlation with the weight w.

The BR training algorithm updates weights and biases according to LM optimization, reducing squared error and weights while determining the appropriate combination for a well-generalized network [13].

$$E_{Ip} = \beta \sum_{Ip=0}^{n} e^2 + \alpha \sum_{op^n=1}^{n} w^2 \quad (14)$$

α and β are parameters for optimization.

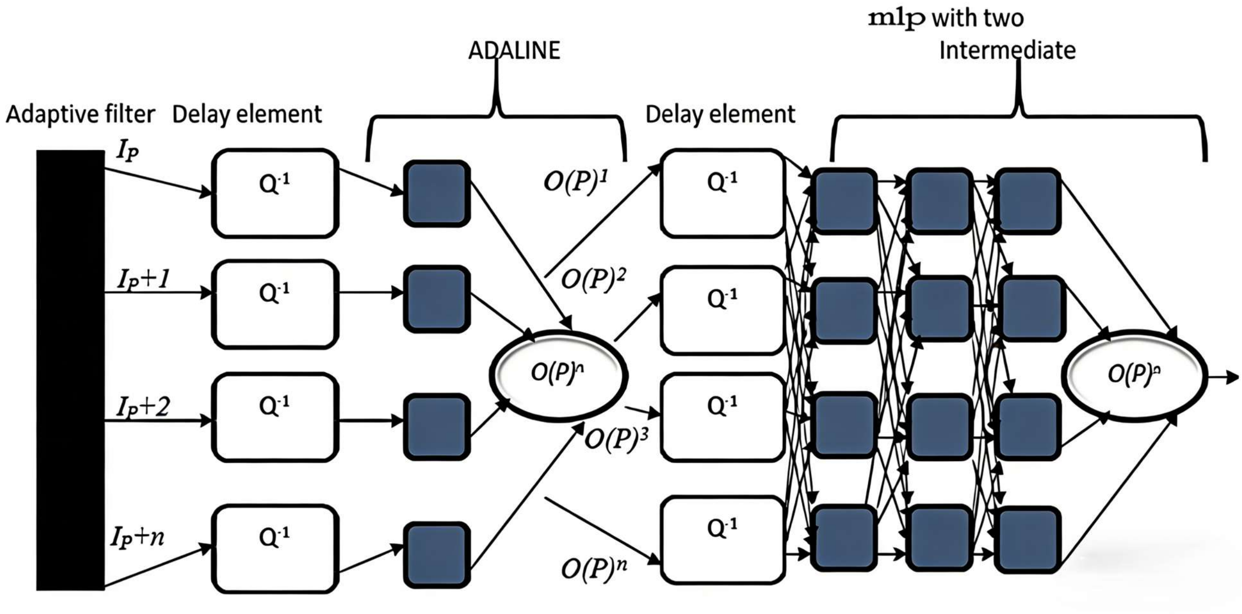


Fig. 2. Architectural composition of the proposed adaptive hybrid system predictor.

V. Dataset Collection, Exerted Tools, and Training Methodology

The experimental dataset was collected during a field measurement carried out in an LTE microcell environment with a 1900 MHz operating frequency. The datasets are the signal power, which are the inputs to the neural network models used for training purposes and comprise 2,220 datasets from the LOS and 2,500 datasets from the NLOS locations, respectively. A drive test was employed for data access during data collection from the BS transmitting antenna measurement points. There is a need for an appropriate training set selection from the real propagation path from which the models will learn to calculate received power, which is the most crucial factor in the training phase. For training optimization, the training set involves measurement data from different routes with different characteristics of propagation, such as reflection, diffraction, reflection, direct rays, etc. The selected routes also include received positions that show various ranges of the input parameter. Hence, the network can learn to behave in different situations, and thus make an appropriate generalization for application to new cases.

The field measurement obtained and employed for the training of the proposed adaptive hybrid predictor was acquired as a live signal around the LTE transceiver base station antenna (NodeB) for one year (12 months). The measurement took such a long period to cater to the location's seasonal variations for easy scalability to larger and more complex datasets. The transceiver base station antennas are sectionalized with 17.5dBi gain and 43dBm transmit power. To define the measurement data location and the Node B location, the Global Positioning System (GPS) equipment was applied.

Some of the measurement tools were a laptop engraved with Test Mobile Software (TEMS), LTE proficient mobile handsets engraved with TEMS, a network scanner, a Global Positioning System (GPS) device, a compass, etc. These tools were used for matching up User Equipment (UE) measurements, i.e., the MS locations corresponding to the BS transmitter and the field test environment. Map Info and an Excel spreadsheet were used for the dataset extraction and normalization. A deep learning toolbox in the Matrix Laboratory was used for training, testing, and validation of the neural network models. These models were the designed hybrid adaptive system predictor and the conventional MLP model, used for comparison purposes. The dataset extracted from the log file is the Reference Signal Received Power (RSRP) (dBm). The TEMS software, engraved in both the laptop and the handset, enabled optimized access in the recording and extraction of the signal power dataset along the test routes. The measurement of the dataset was carried out in two different routes from the selected BS transmitter, a LOS route and an NLOS route.

The first important step in the training process is the appropriate measurement points characterization in the training route according to their type of dominant path. The choice of training routes was a planned process, and a balanced number of measured data points representing different propagation conditions was supplied. The neural network was trained using different numbers of neurons in the intermediate layers, which vary from 5 to 70 neurons, to extract better training points.

A written neural network program was used in the training of the measured dataset via the deep learning neural network toolbox in MATLAB 2022b. An early stopping training technique to minimize over-fitting during network training was employed. The data set division on the application of the early stopping training technique was in the ratio of 70%:15%: 15%, for training, testing, and validation. This helps with improvement in the accuracy of the network training [19]. The normalization of the training dataset was carried out by using an Excel spreadsheet to improve network generalization ability. The dataset normalization was done by exerting the expression in Eq. (16) [19].

$$N_d = \frac{(N_0 - N_{min}}{(Nmin_{max}} \qquad (15)$$

N_d and N_0 are the value of the normalized parameters and the value of the initial parameters respectively. N_{max} and N_{min} are the maximum and minimum parameter values, respectively.

The 1st-order statistical performance indicators were applied for the results analysis. These are the Coefficient of Regression (R), Root Mean Square Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error, and Standard Deviation (SD). The R is a measure of the strength of the relationship between the measured and the predicted values. The RMSE suggests the mean error magnitude between the measured and the predicted values, while the MAE suggests the closeness of the predicted and the measured values. The MSE is the mean average of the squared difference between the measured and the predicted values, while the SD is the dispersion between the measured and the predicted values. These measurement performance indices are given as [12, 13]:

$$R = \frac{K_{measured}(\sum t_k - y_k) - (\sum t_k)(\sum y_k)}{\sqrt{[K_{measured}(\sum t_k^2 - (\sum t_k)^2][K_{meausred}(\sum y_k^2) - (\sum y_k)^2]}} \qquad (16)$$

$$RMSE = \sqrt{MSE} = \frac{1}{K_{measured}} \sqrt{\sum_{k=1}^{k_{measured}} [t_k - y_k]^2} \qquad (17)$$

$$MSE = \frac{1}{k_{measured}} \sum_{k=1}^{k_{measured}} (t_k - y_k)^2 \qquad (18)$$

$$MAE = \frac{1}{K_{measured}} \sum_{k=1}^{K_{measured}} |t_k - y_k| \qquad (19)$$

$$SD = \sqrt{(\frac{1}{K_{measured}} \sum_{k=1}^{K_{measured}} |t_k - y_k| - MAE)^2} \qquad (20)$$

where t_k is the measured values, while y_k is the predicted values, and K = 1, 2, 3…… are the measured signal power numbers.

VI. Hyperparameter Settings for Model Training

The utilized hyperparameters are discussed:

A. *Influence of the Number of Neurons on the Designed Model Architecture*

The determination of the number of neurons in the intermediate layer of the adaptive hybrid model remains a vital integral in the neural network model architectural composition. An inadequate number of neurons or too many neurons in the intermediate layer can result in underfitting or overfitting, respectively. Under-fitting results when the intermediate layer neurons are not enough to satisfactorily detect the signals, mainly in a multi-faceted dataset, while over-fitting is a result of too much processing information in the network. This can also lead to excess required neural network training time.

Thus, there is a need for reasonable consideration in choosing the number of neurons in the intermediate layer. The two training algorithms applied, the LM and the BR algorithms, were trained and tested with 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 neurons. This is to find out the effect of the neuron increment on their performance. These effects were measured using the various error statistics.

B. *Influence of Transfer Function*

The transfer function is a singular monotonically increasing and differentiable function that is applied to input data signals translation produce the final output signals of the neuron. The transfer function is a vital distinct concept of neural networks, mainly for these reasons: without the activation functions, the overall neural network model organization will be similar to a normal linear function, which cannot learn non-linear relationships. Also, the transfer function method polishes the main computation accomplished by neural networks. Finally, the transfer function tends to boost the learning rate and the formation pattern of the datasets. The choice of the right transfer function has a positive influence on the neural network training algorithm's performance. A non-linear transfer function, the tangent transfer function, was employed in this study to ascertain the stability of the proposed neural network model and its performance in terms of the statistical performance indicator.

C. *Influence of the Number of the Intermediate Layer and the Training Algorithm*

The intermediate layer number as well as the utilized training algorithm have a major effect on the performance and the success of the neural network training and prediction. Thus, the choice of the intermediate layer number is one of the most vital challenges when investigating the performance of a neural network architecture for predictive modeling. The application of too many intermediate layers results in poor generalization and rigorous neural network model training [24]. According to [14], two intermediate layers in combination with "m" output neurons are proper for a neural network to study "N" data samples and produce just minimal error. Thus, our previous works [16, 20] studied the impact of one, two, and three intermediate layer numbers. Results show that the two-layered neural network model has superior performance over one-layered and three-layered neural network models for all investigated algorithms. The

LM and the BR training algorithm performances were also assessed, with the results showing fast model training applying the LM algorithms; however, the model training showed better generalization using the BR training algorithm. The BR algorithm is applied in weight update during network training in agreement with the LM algorithm and has demonstrated better training by linear permutation of squared error and weight variables [20]. The algorithm uses back propagation and modifies all variables following the function approximation method.

D. Influence of Learning Rate

Learning rate is a parameter for neural network training that controls the size of the weight and the bias changes during network training. It is simply the fastness of a network in abandoning old beliefs for a new one. Generally, there is a need for a learning rate that is small enough for useful convergence of the network and also high enough that the training of the network doesn't take so long [22].

Therefore, in updating the weight w_{ik} by gradient descent, a learning rate η has to be selected. The huge Mu of the BR training algorithm is compensated by the use of a low training learning rate of 0.1 during the neural network models training to limit it from running past optimal with vast steps. The change in weight added to the old weight is equivalent to the product of the learning rate and the gradient descent multiplied by −1.

$$\Delta w_{ik} = \eta \frac{\partial E}{\partial w_{ik}} = \eta O_k \delta_k \tag{21}$$

The essence of the −1 is for updating in the minimum direction of the error function and not the maximum. Hence, to prevent oscillation, such as connection weight alternation inside the network, and improve convergence rate, an adaptive learning rate is used as a modification of gradient descent back propagation.

VII. Results Tabulation and Analysis Using 1st Order Statistical Performance Indices

The results output is tabulated and discussed.

A. Results Tabulation

The signal power loss prediction results for the designed adaptive hybrid system predictor and the conventional MLP model are analyzed using the 1st order statistical performance indices employed during the neural network models' training. The comparison with the MLP model was chosen, and not with ADALINE, because ADALINE can only solve linearly separable problems, while MLP can proficiently solve non-linear separable problems. The neural network training was carried out in two phases using measured data collected from LOS location-1 and NLOS location-2 LTE microcell environments. The results are shown in Tables I and II for LOS and NLOS locations. The trainings were carried out using the LM and the BR training algorithms.

B. Coefficient of Regression Results Analysis

The coefficient of regression results from the trained adaptive hybrid system predictor and the conventional MLP-ANN models are recorded in Tables I−II for LOS location 1 and NLOS 2 location, respectively. For location-1, R are 0.9922 and 0.9938 on training the designed adaptive hybrid system predictor with LM and BR training algorithms, respectively and 0.9797 and 0.9792 on training the conventional MLP-ANN model with LM and BR training algorithms, respectively. For NLOS location-2, R of 0.9612 and 0.9636 were recorded on training the designed adaptive hybrid system predictor with LM and BR training algorithms, respectively and R of 0.9285 and 0.9619 were recorded on training the conventional MLP-ANN model with LM and BR training algorithms, respectively. These values clearly show the superiority in the prediction of the measured signal power dataset by the adaptive hybrid system predictor over the conventional MLP-ANN model on both training with the LM and the BR training algorithms. The training results of the designed adaptive system predictor for location-1 LOS and location-2 NLOS, as shown in Tables I and II, are highlighted with red color for LOS location-1 and blue color for NLOS location-2, respectively. As R measures the closeness of the predicted values to the measured values, with values closer to +1 showing strong closeness of the predicted values to the measured values, the designed adaptive hybrid system predictor shows stronger prediction of the measured dataset over the conventional MLP-ANN model. Also, in comparison of the training performances of the LM and the BR training algorithms from the tabulated results, the results show superior performances of the BR training algorithm over the LM training algorithm. This is because the BR training algorithm reduces squared error, thereby resulting in a better-generalized network.

C. The Standard Deviation, Root Mean Square Error, Mean Squared Error, and Mean Absolute Error Results Analysis

The results as recorded from training the designed adaptive hybrid system predictor in Tables I−II show that for location-1 LOS, SD is 1.3425, RMSE is 1.7038, MSE is 2.9029, and MAE is 1.0491 on training with LM training algorithm and on training with BR training algorithm for the same location-1, SD is 1.0871, RMSE is 1.5036, MSE is 2.2609, and MAE is 1.0389. For the same location-1, training the conventional MLP-ANN model gives SD as 1.6016, RMSE as 2.7756, MSE as 7.7040, and MAE as 2.2668 with the LM training algorithm, and on training with the BR training algorithm, SD is 1.5465, RMSE is 2.7357, MSE is 7.4842, and MAE is 2.2566.

The training results of the adaptive hybrid system predictor for location-2 NLOS using the LM training algorithm give SD of 1.6890, RMSE of 2.0853, MSE of 4.384, and MAE of 1.2230. Standard deviation of 1.5707, RMSE of 2.0173, MSE of 4.0696, and MAE of 1.2659 were recorded on training the same designed adaptive hybrid system predictor with the BR training algorithm. Training the conventional MLP-ANN with the LM algorithm for location-2 NLOS gave SD of 2.1014, RMSE of 2.8.716, MSE of 8.2460, and MAE of 1.9570, while training the same conventional MLP for location-2 with the BR algorithm gave SD of 1.4658, RMSE of 2.0932, MSE of 4.3813, and MAE of 1.4942. These results show clearly a better prediction of the measured dataset using the designed adaptive hybrid system predictor, and also a superior prediction using the BR training algorithm.

Minimal prediction errors were recorded on the application of both the designed adaptive hybrid system predictor and the BR training algorithm for both location-1 LOS and location-2 NLOS in comparison to the conventional MLP-ANN model and the LM training algorithm, respectively.

TABLE I. SIGNAL POWER LOSS ERROR ANALYSIS FOR LOS LOCATION-1

Neural Network Models Training with Training Algorithms	1st Order Statistical Performance Indices				
	R	SD	RMSE	MSE	MAE
Conventional MLP trained with the LM algorithm	0.9787	1.6018	2.7756	7.7040	2.2668
Designed Filtered Adaptive Hybrid System, Predictor trained with LM algorithm	0.9922	1.3425	1.7038	2.9029	1.0491
Conventional MLP trained with the BR algorithm	0.9792	1.5465	2.7357	7.4842	2.2566
Designed Filtered Adaptive Hybrid System, Predictor trained with BR algorithm	0.9938	1.0871	1.5036	2.2609	1.0389

TABLE II. SIGNAL POWER LOSS ERROR ANALYSIS FOR NLOS LOCATION-2

Neural Network Models Training with Training Algorithms	1st Order Statistical Performance Indices				
	R	SD	RMSE	MSE	MAE
Conventional MLP with trained LM algorithm	0.9285	2.1014	2.8716	8.2460	1.9570
Designed Filtered Adaptive Hybrid System, Predictor trained with LM algorithm	0.9612	1.6890	2.0853	4.3484	1.2230
Conventional MLP trained with the BR algorithm	0.9618	1.4658	2.0932	4.3813	1.4942
Designed Filtered Adaptive Hybrid System, Predictor trained with BR algorithm	0.9636	1.3107	2.0173	4.0696	1.2659

TABLE III. ANALYSIS OF NEURON VARIATION IN MLP-ANN WITH ONE INTERMEDIATE LAYER FOR LOS USING LM AND BR TRAINING ALGORITHM

Training Algorithm/ Neuron Numbers	Training time (s)		Coefficient of Correlation (r)	
	LM	BR	LM	BR
5	00:00:02	00:00:09	0.8100	0.9712
10	00:00:08	00:00:15	0.8175	0.9740
15	00:00:11	00:00:17	0.9150	0.9745
20	00:00:12	00:00:22	0.9250	0.9747
25	00:00:15	00:00:24	0.9550	0.9755
30	00:00:16	00:00:29	0.9700	0.9762
35	00:00:20	00:00:31	0.9755	0.9782
40	00:00:22	00:00:33	0.9777	0.9785
45	00:00:23	00:00:39	0.9770	0.9780
50	00:00:26	00:00:40	0.9710	0.9779
55	00:00:28	00:00:44	0.9705	0.9776
60	00:00:28	00:00:46	0.9700	0.9767
65	00:00:33	00:00:49	0.9695	0.9765
70	00:00:35	00:00:56	0.9690	0.9762

TABLE IV. ANALYSIS OF NEURON VARIATION IN MLP-ANN WITH TWO INTERMEDIATE LAYERS FOR LOS USING LM AND BR TRAINING ALGORITHMS

Training Algorithm/ Neuron Numbers	Training time		Coefficient of Correlation (r)	
	LM	BR	LM	BR
[5, 10]	00:00:02	00:00:24	0.9705	0.9783
[15, 20]	00:00:04	00:00:24	0.9710	0.9785
[25, 30]	00:00:10	00:00;24	0.9770	0.9788
[35, 40]	00:00:12	00:00:24	0.9787	0.9792
[45, 50]	00:00:18	00:00:24	0.9781	0.9784
[55, 60]	00:00:20	00:00:24	0.9730	0.9747
[65, 70]	00:00:20	00:00:24	0.9725	0.9740

TABLE V. ANALYSIS OF NEURON VARIATION IN ADAPTIVE HYBRID PREDICTOR WITH ONE INTERMEDIATE LAYER FOR LOS USING LM AND BR TRAINING ALGORITHMS

Training Algorithm/ Neuron Number	Training time (s)		Coefficient of Correlation (r)	
	LM	BR	LM	BR
5	00:00:01	00:00:04	0.9905	0.9911
10	00:00:02	00:00:06	0.9907	0.9913
15	00:00:04	00:00:10	0.9908	0.9914
20	00:00:07	00:00:20	0.9009	0.9014
25	00:00:08	00:00:20	0.9910	0.9917
30	00:00:10	00:00:20	0.9913	0.9918
35	00:00:12	00:00:20	0.9916	0.9919
40	00:00:13	00:00:20	0.9917	0.9920
45	00:00:15	00:00:20	0.9915	0.9919
50	00:00:16	00:00:20	0.9910	0.9918
55	00:00:39	00:00:21	0.9910	0.9918
60	00:00:39	00:00:21	0.9909	0.9918
65	00:00:44	00:00:21	0.9908	0.9917
70	00:00:46	00:00:22	0.9907	0.9916

The architectural complexity of the adaptive hybrid system predictor and the conventional MLP-ANN, and their computational efficiency tradeoffs from the neural network training results, considering the LM and the BR training algorithms, are shown in Tables III−VI.

Training results from LOS location-1 are shown in Tables III and IV for the MLP-ANN using a single intermediate layer and two intermediate layers, respectively. The variations of the intermediate layer neurons were from 5 to 70 during the neural network training using the LM and the BR training algorithms. On application of 40 neurons in the intermediate layer, the highest correlation coefficient was recorded as highlighted in yellow colour. This indicates the closest of the predicted values to the actual dataset on application of 40 neurons. As the neuron numbers in the MLP-ANN intermediate layer increase, there is a rapid decrease in the correlation coefficient and other statistical performance indicators, inferring that the MLP-ANN performs excellently in predicting systems that are not too complex. The proposed adaptive hybrid predictor, however, predicts excellently as the number of neurons increases, showing its capability to efficiently predict both simple and complex systems, and performs optimally for both a single intermediate layer as well as two intermediate layers, as shown in Tables V−VI, respectively. However, the application of two intermediate layers has the advantage of faster training of the network model. The LM and BR performance abilities were also noted as can be seen in the tables, that although the BR training algorithm gives better prediction results, it requires more training time when compared to the LM training algorithm. The same training result pattern was recorded for NLOS location-2 for both the MLP-ANN and the adaptive hybrid predictor.

TABLE VI. ANALYSIS OF NEURON VARIATION IN ADAPTIVE HYBRID PREDICTOR WITH TWO INTERMEDIATE LAYERS FOR LOS USING LM AND BR TRAINING ALGORITHMS

Training Algorithm/NEURON Numbers	Training Time		Coefficient of Correlation (r)	
	LM	BR	LM	BR
[5, 10]	00:00:44	00:00:10	0.9910	0.9916
[15 ,20]	00:00: 42	00:00:10	0.9915	0.9919
[25, 30]	00:00:41	00:00:10	0.9918	0.9919
[35, 40]	00:00:40	00:00:10	0.9920	0.9930
[45, 50]	00:00:39	00:00:10	0.9921	0.9935
[55, 60]	00:00:38	00:00:10	0.9921	0.9934
[65, 70]	00:00:38	00:00:10	0.9922	0.9938

D. *Result Analysis of the Neural Network Training States*

A major objective of the designed filtered adaptive hybrid system predictor is for optimized choices of weights and biases during the neural network training by minimizing the sum of squared error on application of the sharpest descent technique over the conventional MLP model. Therefore, the problem associated with input-output mapping can be optimized by the location of a function that offers minimum error in the course of neural network model training. There is an establishment of values that tends to local minima by gradient descent on the application of all data instances for adjustment of weight determination, while the epoch is defined. i.e., training times for all the training vectors exerted for weight update.

The training states of the neural network models show small gradient values for the designed hybrid system predictor in comparison to the conventional MLP model for both LOS location-1 and NLOS location-2. This is hugely because the LMS adaptive filter adopted in the design of the adaptive hybrid system predictor for de-noising of the dataset is a type of stochastic gradient descent algorithm; as such, the value of the gradient descent at which it tends to local minima is less in comparison with the conventional MLP-ANN model.

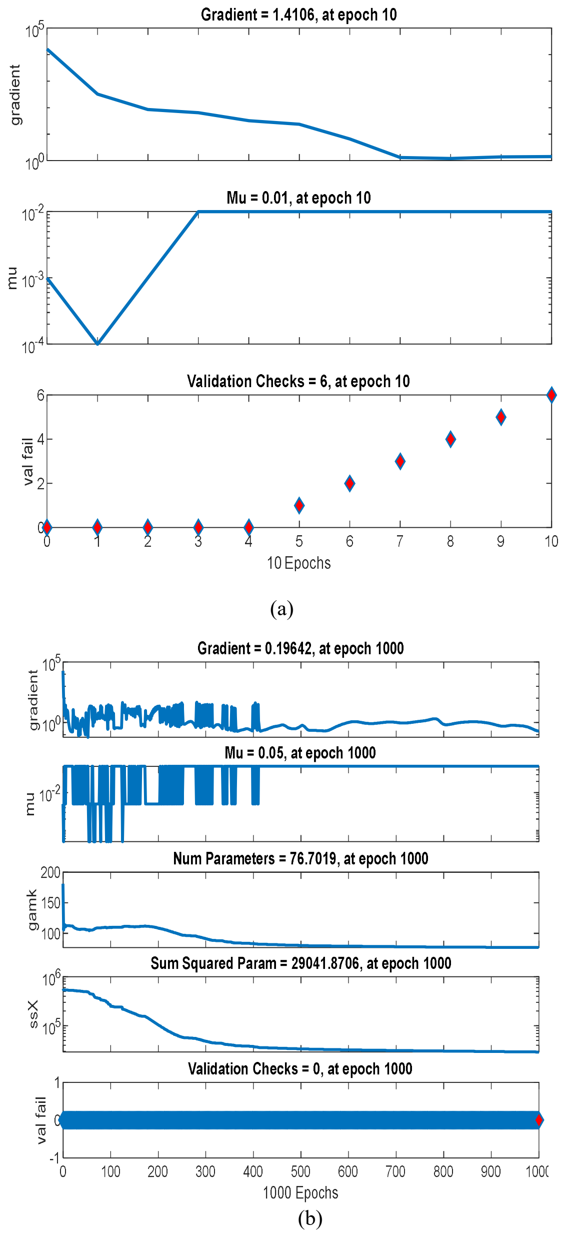
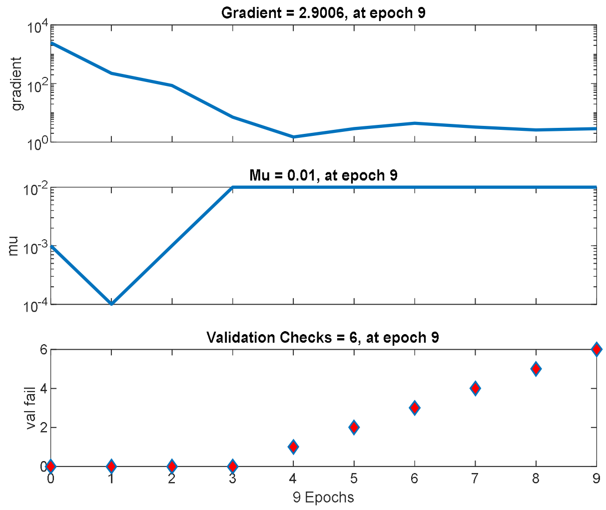


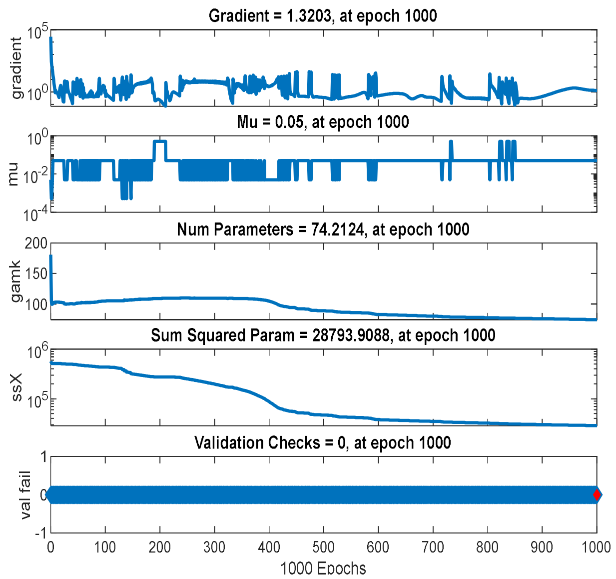
Fig. 3. Training state results of (a) Conventional MLP model and (b) Adaptive hybrid system predictor using the LM training algorithm for LOS location-1.

Figs. 3(a)−5(a) are training results of the conventional MLP model using the LM algorithm for LOS location-1 and NLOS location-2, respectively. For the LOS location-

1, it reached the local minima at a gradient of 1.4106 at epoch 10, and for NLOS location-2, it reached local minima at a gradient of 10.6312 at epoch 9. Training the same conventional MLP, using the BR training algorithm for LOS location-1 and NLOS location-2 in Figs. 4(a) and 6(a) respectively, the neural network model reached local minima at a gradient of 2.9006 at epoch 9 and at a gradient of 7.728, and at epoch 8, respectively. Training the designed adaptive hybrid system predictor with the LM training algorithm, as shown in Figs. 3(b) and 5(b), tends to reach the local minima at 0.19642 at epoch 1000 and 6.1116 at epoch 1000 for LOS location −1 and NLOS location-2, respectively. While training using the BR training algorithm for the same designed adaptive hybrid system predictor tends to reach local minima at 1.3203, and epoch 1000, and 0.41963 at epoch 1000 for LOS location-1 and NLOS location-2, as shown in Figs. 4(b)–6(b) respectively.



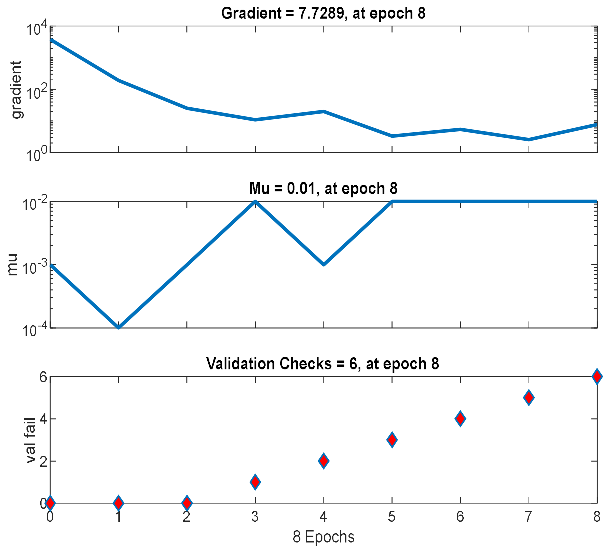
(a)



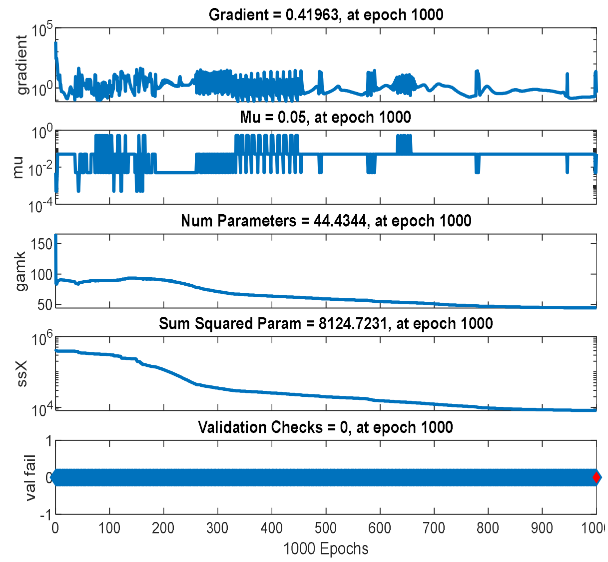
(b)

Fig. 4. Training state results of (a) Conventional MLP model and (b) Adaptive hybrid system predictor using the BR training algorithm for LOS location-1.

The trend from the training results demonstrates performance function reduction at each iteration, using the LM training algorithm with the LM training algorithm showing very high training speed. Thus, the iterative results were obtained with very few training epochs in comparison to training using the BR training algorithm, which trains the entire 1000 epochs, resulting in low training speed. High training speed with few training epochs is one of the pros of the LM training algorithm over the BR training algorithm. However, the LM training algorithm has the problem of memory and computational overhead resulting from gradient and Hessian matrix computation; thus, the BR training algorithm effectively reduces squared error and weights and determines an appropriate combination, resulting in a better-trained and generalized network with improved training results.



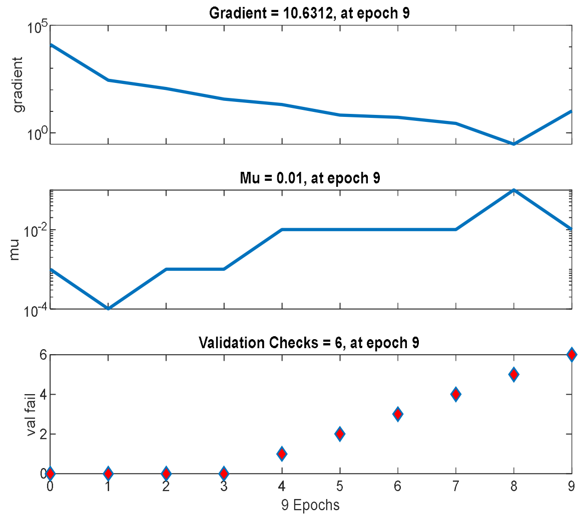
(a)



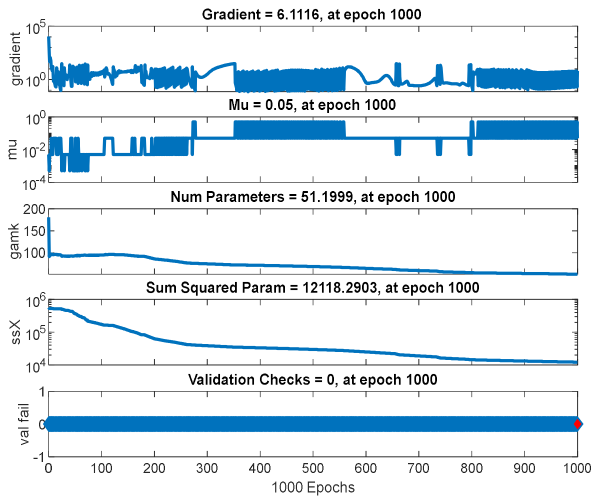
(b)

Fig. 5. Training state results of (a) Conventional MLP model and (b) Adaptive hybrid system predictor using the LM training algorithm for NLOS location-2.

The momentum parameter (Mu) cogitates an update at each iteration and determines the linear combination of the previous gradient and update. It is a variation of stochastic gradient descent used for the swift convergence of the loss function. The training results show that both the designed adaptive system predictor and the conventional MLP-ANN training with the BR training algorithm yielded high training results of 0.05Mu over training with the LM algorithm which gave 0.01Mu for all the training instances for both LOS location-1 and NLOS location-2 as shown in Figs. 3(a)−6(b) accordingly. The huge Mu of the BR training algorithm is compensated by the use of a low training learning rate of 0.1 during the neural network models training to limit it from running past optimal with vast steps.



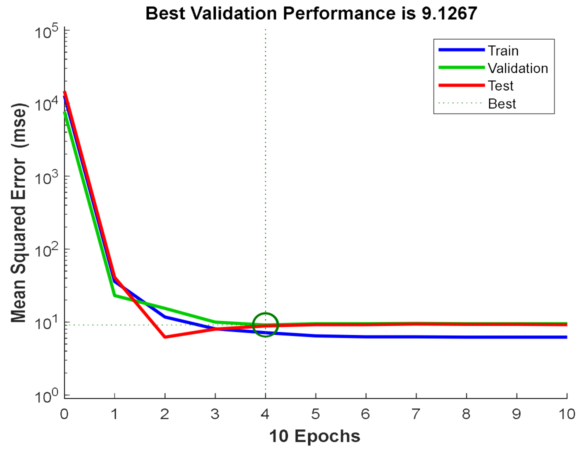
(a)



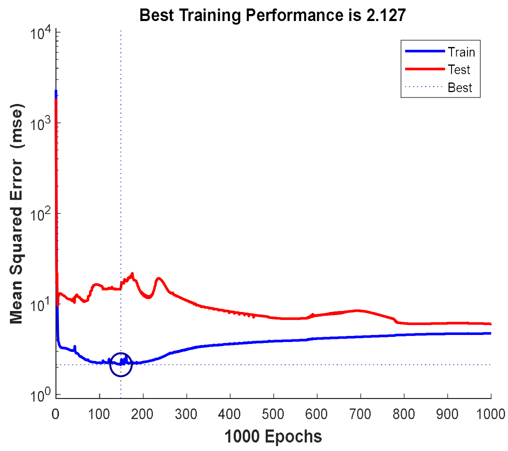
(b)

Fig. 6. Training state results of (a) Conventional MLP model and (b) Adaptive hybrid system predictor using the BR training algorithm for NLOS location-2.

The MSE measures the closeness of the regression line to a set of data points. It is the average of the squared differences between the measured values and the predicted values, i.e., the risk function corresponding to the expected values of the squared error loss. This is calculated during the neural network training by taking the mean average of errors squared from the dataset as it relates to a function, and the lower the value, the better the result. An MSE closer to zero tends to be a perfect model. The performance MSE results analysis is shown in Figs. 7(a)−(b) for the training of the conventional MLP-ANN model and the designed adaptive hybrid system predictor, respectively, using the LM training algorithm. Figs. 8 (a)−(b) show the training result analysis for the conventional MLP-ANN model and the designed adaptive hybrid system predictor, respectively, using the BR training algorithm, both for LOS location-1.



(a)



(b)

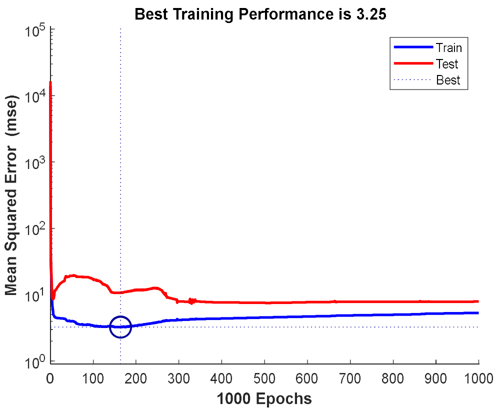
Fig. 7. Training performance MSE results of (a) Conventional MLP model, (b) Designed adaptive hybrid predictor, using the LM training algorithm for LOS location-1.

E. Performance Mean Squared Error Result Analysis

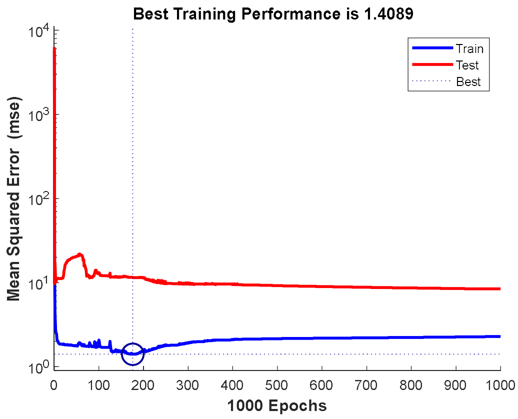
From the performance MSE training results from Figs. 7(a)−(b), on training the conventional MLP-ANN model, and the designed adaptive hybrid system predictor with the LM training algorithm, the performance MSE for the conventional MLP model is 9.1267 and 2.127 for the designed adaptive hybrid system predictor. Training the conventional MLP-ANN model with the BR training algorithm gave performance MSE of 3.25 and 1.4089 for the designed adaptive hybrid system predictor. The results clearly show the superiority of the designed adaptive hybrid system predictor, which has low MSE in all instances in comparison to the performance MSE training results of the conventional MLP model. Also, the training using the BR training algorithm in Figs. 8(a)−(b) gave lower performance MSE in comparison to the training using the LM training algorithm in Figs. 7(a)−(b).

F. Analysis of Regression Results

Regression is a measure of the strength of the relationship between the measured data and the predicted values. The regression results analysis from training the conventional MLP-ANN model and the designed adaptive hybrid system predictor using the LM and the BR training algorithms for LOS location-1 and NLOS location-2 are shown in Figs. 9 (a)−9(b) accordingly.



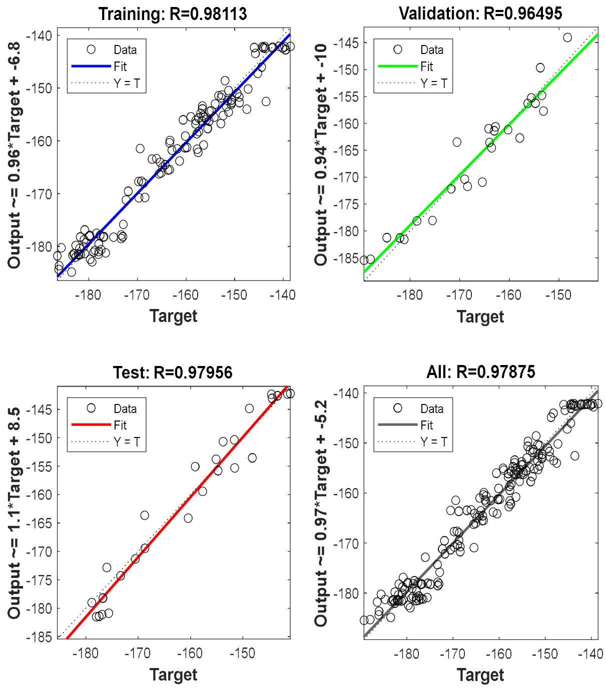
(a)



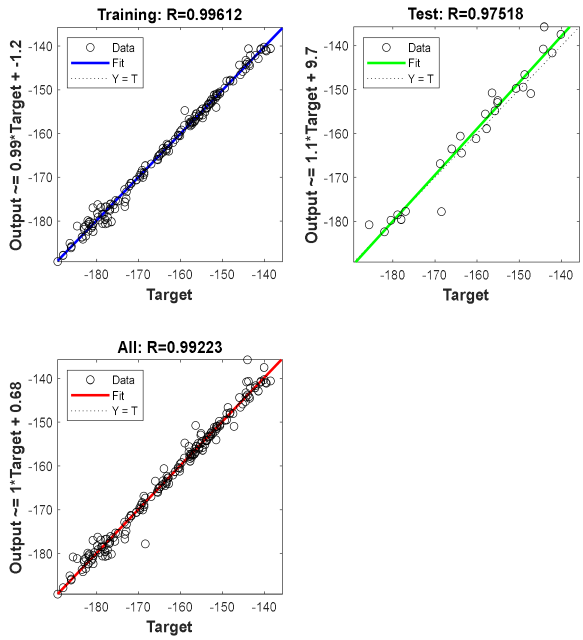
(b)

Fig. 8. Training performance MSE results of (a) Conventional MLP model, (b) Designed adaptive hybrid predictor, using the BR training algorithm for LOS location-1.

The regression results of the trained conventional MLP-ANN model and the designed adaptive hybrid system predictor for the LOS location-1 yielded R of 0.97875 with the conventional MLP model and R of 0.99223 with the designed adaptive hybrid system predictor, on training with the LM training algorithm. On training with the BR training algorithm, R of 0.97922 for the conventional MLP model and 0.99378 for the designed adaptive hybrid system predictor were recorded for the LOS location-1. For NLOS location-2, R yielded 0.92848 and 0.96119 for the conventional MLP model and the designed adaptive hybrid system predictor, respectively, on training with the LM training algorithm. On training the neural network with the BR training algorithm for the same NLOS location-2, R gave 0.96184 for conventional MLP and 0.96336 for the designed adaptive hybrid system predictor. The training, testing, and validation values are all shown in Figs. 11 (a)−12 (b) accordingly. Training using the BR training algorithm does not require a validation dataset, as seen from the figures trained with the BR training algorithm. This is because the essence of checking for validation during neural network training is to monitor the error state and stop the neural network training if it gets worse; however, the BR training algorithm error is not just based on how well the model performs on the dataset, but also on the size of the weights. Thus, the BR training algorithm does not let the network explore larger weights during the training process, even if the validation step is on.

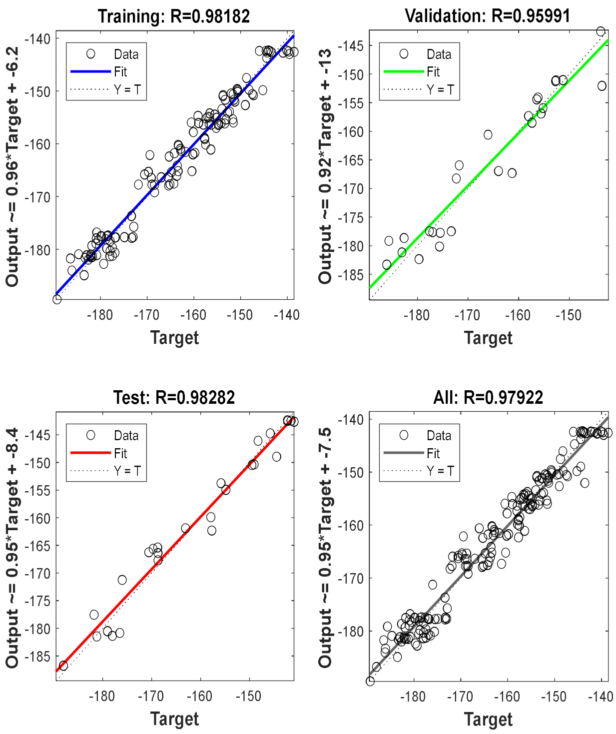


(a)

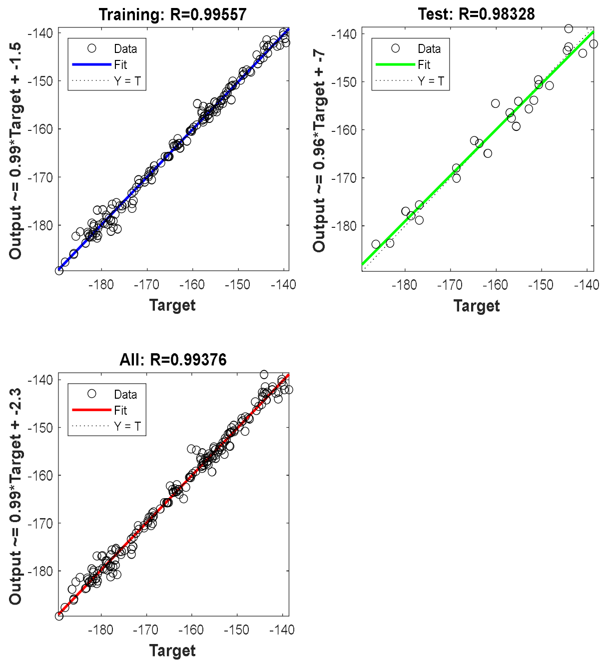


(b)

Fig. 9. Training regression results analysis of (a) Conventional MLP model and (b) Designed adaptive hybrid system predictor for LOS location-1 using the LM training algorithm.

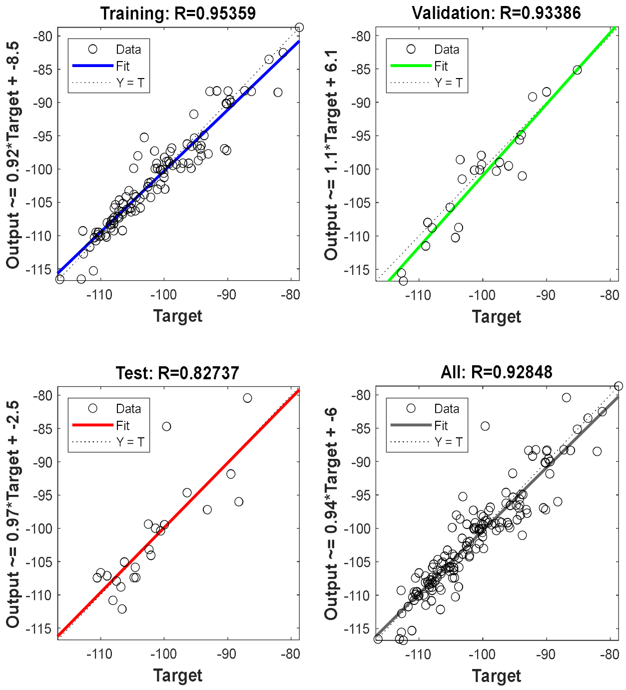


(a)

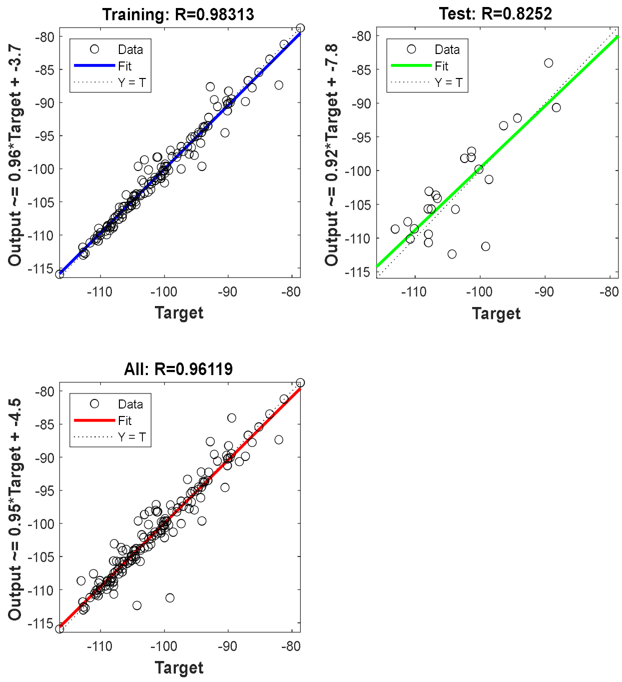


(b)

Fig. 10. Training regression results analysis of (a) conventional MLP-ANN model and (b) designed adaptive hybrid system predictor for LOS location-1 using the BR training algorithm.

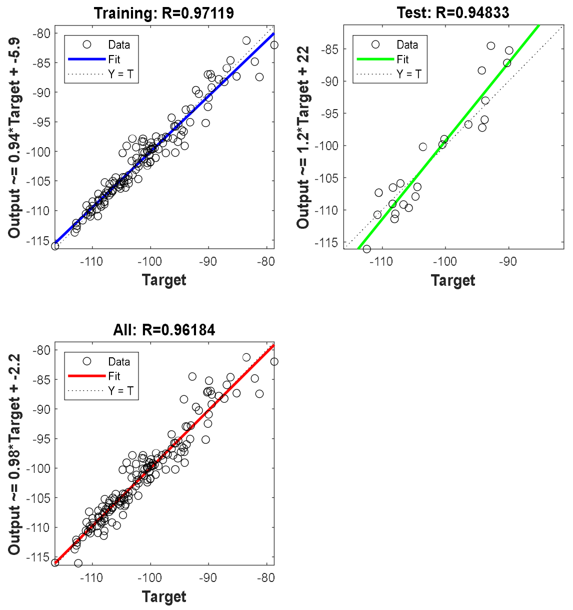


(a)

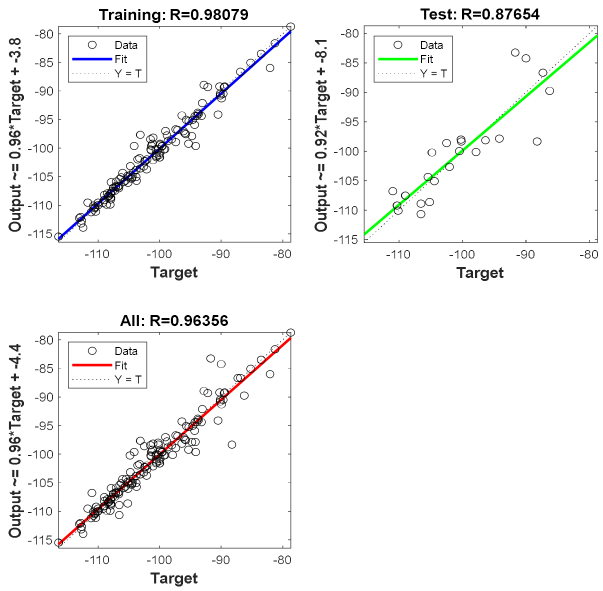


(b)

Fig. 11. Training regression results analysis of (a) Conventional MLP-ANN model and (b) Designed adaptive hybrid system predictor for NLOS location-2 using the LM training algorithm.



(a)



(b)

Fig. 12. Training regression results analysis of (a) Conventional MLP-ANN model and (b) Designed adaptive hybrid system predictor for NLOS location-2 using the BR training algorithm.

The training results shown in both tables and figures demonstrate the superiority of the designed adaptive hybrid system predictor over the conventional MLP-ANN model, as well as the superiority in prediction capability of the BR training algorithm over the LM training algorithm.

VIII. Conclusions and Future Work

This research work designed an adaptive hybrid system predictor with a well-structured architectural implementation empowered with the grid search-based hyperparameter tuning technique for optimal signal power loss prediction and approximation between the base station and the mobile station path lengths. The degree of the prediction accuracy of the developed adaptive hybrid system predictor over the conventional MLP-ANN is also clearly provided using the first-order statistics performance indices in the tables and graphs.

The trained dataset was from two LTE micro-cell environments, a LOS location-1, and an NLOS location-2.

The results clearly show superior performance of signal power loss prediction using the adaptive hybrid system predictor, which yielded minimal prediction error in comparison to prediction using the conventional MLP model. The BR training algorithm exhibits superior training performance compared to the LM training algorithm; however, this comes with certain tradeoffs, such as increased training time.

In conclusion, the proposed adaptive hybrid system predictor is a well-structured implementation of a network architecture empowered with the correct hyperparameter tuning algorithms, ensuring efficient prediction of signal power loss and easy adaptability in diverse environments. This contrasts with the conventional MLP-ANN, whose efficient architectural composition to suit the underlying problem is based on trial and error.

The developed adaptive hybrid system predictor, the selection of the adequate training algorithm, and the hyperparameters have a clear impact on the quality of the prediction competence. Explicitly in terms of the MAE, RMSE, R, and SD statistical values, the proposed adaptive hybrid system predictor yielded reasonable performance prediction accuracy improvement over the conventional MLP-ANN on the acquired LTE path loss datasets.

Future work will compare the developed model with other modern approaches recently applied to the same problem, such as CNNs etc. It will also test the model's scalability to larger datasets, different frequencies, and diverse environments.

Conflict of Interest

The authors declare no conflict of interest.

Authors Contributions

Virginia C. Ebhota and Thokozani Shongwe conducted this research; Virginia C. Ebhota trained and analysed the model with data and wrote the paper; Thokozani Shongwe verified the result with the designed model; both authors had approved the final version.

Funding

This work is supported and funded by the University of Johannesburg, South Africa.

Reference

[1] V. C. Ebhota and V. M. Srivastava, "Effect of architectural composition of MLP ANN in neural network learning for signal power loss prediction," *Journal of Communication*, vol. 16, pp. 20−30, 2021.

[2] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Investigation and comparison of the performance of multi-layer perceptron and radial Basis function artificial neural networks for signal power loss prediction," *International Journal on Communications Antenna and Propagation,* vol. 9, pp.43−54, 2019.

[3] V. C. Ebhota and T. Shongwe, "Neural network optimization and implications of high and low gradient results in training vanilla and hybrid adaptive neural network models for effective signal power loss prediction," *International Journal of Review on Modelling and Simulations.* vol. 16, pp. 274−285, 2023.

[4] V. C. Ebhota and V. M. Srivastava, "Performance comparison of non-linear median filter built on MLP-ANN and conventional MLP-ANN: Using improved dataset training in micro-cell environment," *Journal of Communication.* vol. 16, no 11, pp. 508−515, 2021.

[5] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Improved adaptive signal power loss prediction using combined vector statistics based smoothing and neural network approach," *International Journal on Progress in Electromagnetics Research C.*, vol. 82, pp. 155−169, 2018.

[6] C. Nguyen and A. A. Cheema, "A deep neural network-based multi-frequency path loss prediction model from 0.8 GHz to 70 GHz," *Journal of Sensors*, vol. 21, pp. 51−60, 2021.

[7] K. C. Igwe, O. D. Oyedum, A. M. Aibinu, M. O. Ajewole, and A. S. Moses, "Application of artificial neural network modeling techniques to signal strength computation," *Journal of Heliyon*, vol. 7, pp. 1−9, 2021.

[8] G. Bhumika and S. N. Shailendra, "Image denoising with linear and non-linear filters: A review," *International Journal of Computer Science*, vol. 10, pp. 1694−0784, 2013.

[9] J. Isabona and V. M. Srivastava, "A neural network-based model for signal coverage propagation loss prediction in urban radio communication environment," *International Journal of Applied Engineering Research,* vol. 11, pp. 11002−11008, 2016.

[10] J. Isabona and V. M. Srivastava, "Hybrid neural network approach for predicting signal propagation loss in urban microcells," in *Proc. 2016 IEEE Region 10 Humanitarian Technology Conference*, 2016.

[11] A. Ali, R. Ghazali, and M. M. Deris, "The wavelet multilayer perceptron for the prediction of earthquake time series data," in *Proc. 13th International Conference on Information Integration and Web-based Applications and Services*, 2011, pp. 138−143.

[12] H. Anysz, A. Zbiciak, and N. Ibadova, "The influence of input data standardization method on prediction accuracy of artificial neural networks," *Procedia Engineering*, vol. 153, pp. 66−70, 2016.

[13] N. M. Nawi, W. H. Atomi, and M. Z. Zehman, "The effect of data pre-processing on optimized training of artificial neural networks," *Procedia Technology*, vol. 11, pp. 32−39, 2013.

[14] K. A. Dotche, F. Sekyere, and W. Banuenulmah, "LPC for signal analysis in cellular network coverage," *Open Access Library Journal,* vol. 3, pp. 7−10, 2016.

[15] X. Jia, B. D. Brabandere, T. Tuytelaars, and L. V. Gool, "Dynamic filter networks," in *Proc. 30th Conference on Neural Information Processing Systems*, 2016, pp. 2046−2054.

[16] V Jothiprakash and A. S. Kote, "Improving the performance of data-driven techniques through data pre-processing for modeling daily reservoir inflow," *Hydrological Science,* vol. 56, pp. 168−186, 2011.

[17] B. E. Gschwendtner and F. M. Landstorfer, "Adaptive propagation modeling using a hybrid neural technique," *Electronics Letters,* vol. 32, pp. 162−164, 1996.

[18] H. Yonaba, F. Anctil, and V. Fortin, "Comparing sigmoid transfer functions for neural network multi-step ahead streamflow forecasting," *Journal of Hydrologic Engineering,* vol. 15, pp. 275−283, 2010.

[19] V. C. Ebhota, J. Isabona, and V. M. Srivastava, "Signal power loss prediction based on artificial neural networks in microcell environment," in *Proc. IEEE International Conference on Electro-Technology for National Development*, 2017, pp. 250−257.

[20] W. R. Loh, S. Y. Lim, I. F. M. Rafie, J. S. Ho, and K. S. Tze, "Intelligent base station placement in urban areas with machine learning," *IEEE Antennas Wireless. Propagation. Letter*, vol. 22, pp. 2220–2224, 2023.

[21] A. C. Deme, "Radio propagation modelling of a typical sudan savanna belt rural terrain using soft computing and empirical techniques," *Journal of Multidisciplinary Engineering Science and Technology*, vol. 6, pp. 11320−11325, 2019

[22] V. C. Ebhota and T. Shongwe, "Analysis of the strengths of various intermediate layer neurons of radial basis function and vanilla neural networks in the prediction of signal power loss using measurement data from micro-cell environment," *International Journal on Communications, Antenna and Propagation*, vol. 13, pp. 301−310, 2023.