*inventions*

MDPI

*Article*

# Multi-Objective Optimization Algorithm Based Bidirectional Long Short Term Memory Network Model for Optimum Sizing of Distributed Generators and Shunt Capacitors for Distribution Systems

**Amarendra Alluri [1], Srinivasa Rao Gampa [1,*], Balaji Gutta [1], Mahesh Babu Basam [1], Kiran Jasthi [2], Nibir Baran Roy [3] and Debapriya Das [3]**

[1] Department of Electrical & Electronics Engineering, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Vijayawada 521356, Andhra Pradesh, India; amarendrabackup@gmail.com (A.A.); guttabalaji.eee@gmail.com (B.G.); bannu219@gmail.com (M.B.B.)
[2] Department of Electrical & Electronics Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad 500090, Telangana, India; jasthikiran88@gmail.com
[3] Department of Electrical Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, West Bengal, India; nbroy95@iitkgp.ac.in (N.B.R.); ddas@ee.iitkgp.ac.in (D.D.)
* Correspondence: gsr_gsrinu@yahoo.co.in

**Abstract:** In this paper, a multi-objective grey wolf optimization (GWO) algorithm based Bidirectional Long Short Term Memory (BiLSTM) network machine learning (ML) model is proposed for finding the optimum sizing of distributed generators (DGs) and shunt capacitors (SHCs) to enhance the performance of distribution systems at any desired load factor. The stochastic traits of evolutionary computing methods necessitate running the algorithm repeatedly to confirm the global optimum. In order to save utility engineers time and effort, this study introduces a BiLSTM network-based machine learning model to directly estimate the optimal values of DGs and SHCs, rather than relying on load flow estimates. At first, a multi-objective grey wolf optimizer determines the most suitable locations and capacities of DGs and SHCs at the unity load factor and the same locations are used to obtain optimum sizing of DGs and SHCs at other load factors also. The base case data sets consisting of substation apparent power, real power load, reactive power load, real power loss, reactive power loss and minimum node voltage at various load factors in per unit values are taken as input training data for the machine learning model. The optimal sizes of the DGs and SHCs for the corresponding load factors obtained using GWO algorithm are taken as target data sets in per unit values for the machine learning model. An adaptive moment estimation (adam) optimization approach is employed to train the BiLSTM ML model for identifying the ideal values of distributed generations and shunt capacitors at different load factors. The efficacy of the proposed ML-based sizing algorithm is demonstrated via simulation studies.

**Keywords:** distributed generators; shunt capacitors; machine learning; grey wolf optimization; BiLSTM model; multi-objective optimization



**Citation:** Alluri, A.; Gampa, S.R.; Gutta, B.; Basam, M.B.; Jasthi, K.; Roy, N.B.; Das, D. Multi-Objective Optimization Algorithm Based Bidirectional Long Short Term Memory Network Model for Optimum Sizing of Distributed Generators and Shunt Capacitors for Distribution Systems. *Inventions* **2024**, *9*, 114. https://doi.org/10.3390/inventions9060114

Academic Editor: Om P. Malik

Received: 2 October 2024
Revised: 31 October 2024
Accepted: 7 November 2024
Published: 12 November 2024



## 1. Introduction

Modern power distribution systems are integrated with distributed energy resources and shunt capacitors to meet the performance requirements and energy demands of consumers. It is important to choose appropriate sizes and locations of distributed generators (DGs) and shunt capacitors (SHCs) to gain the advantages of loss reduction and meet the voltage constraints. Many researchers have proposed conventional and evolutionary computing-based techniques for the optimal installation of DGs and SHCs with distribution systems, ensuring the utilities perform satisfactorily at different load factors.

In [1], a genetic algorithm-based multi-objective approach obtained the optimum sizing of DGs required at various load factors, thereby improving the distribution system voltage profile and reducing power losses. This approach employed a sensitivity index-based methodology to determine the optimal locations for DGs. A two-stage fuzzy multi-objective approach has been developed in [2] using a grasshopper optimization algorithm for optimal placement of DGs, shunt capacitors, and electric vehicle charging stations, and the energy savings gained by the DGs and shunt capacitor integration have been utilized for the placement of EV charging stations. The study in [3] employed a classical branch and bound methodology to optimally place distributed energy resources while proposing a second-order cone programming for their sizing. The analysis has taken variations in load and photovoltaic generation into account. In [4], an improved artificial hummingbird algorithm has been proposed for integrating biomass-based distributed generators with radial distribution systems, and performance analysis was carried out with DGs operating at different power factors. The authors in [5] have developed a whale optimization-based multi-objective approach to enhance the annual economic savings of the distribution system through the optimal placement of DGs. In [6], the authors proposed a genetic algorithm-based multi-objective approach with the objectives of minimizing real power loss, node voltage improvement, and distribution system stability enhancement. They developed a Thevenin impedance-based stability index, which limits the DG power injection to a safe value. The dwarf mongoose optimization (DMO) method put forward in [7] achieved the ideal placement of DGs and DSTATCOM simultaneously, enhancing the distribution system's performance and minimizing its operating costs. The loss sensitivity factors are used to identify the appropriate locations for DGs and DSTATCOM units. In [8], the synchronous compensators were also used along with static var compensators and DGs to minimize both real and reactive power losses of the distribution system. In [9], the optimal scheduling of battery energy storage systems and solar and wind DG units plans the distribution system operation during different seasons, taking into account the uncertainties of load and DG outputs. Network reconfiguration further reduces losses and improves power quality. Network reconfiguration further reduces losses and improves power quality. The authors in [10] have employed a multi-objective grey wolf optimizer to determine the best number and placement of distributed generators, capacitor banks, and fault current limiters (FCLs). They have allocated DGs, CBs, and FCLs simultaneously to reduce losses, harmonic distortion, and short-circuit current. Following the optimization processes, the distribution network operator determined the best operating configuration for the feeder using the Pareto optimal front. In the realm of distribution network planning, the authors in [11] have introduced a gravitational search algorithm-based multi-objective methodology for the optimal capacity and allocation of distributed generators and storage components necessary for distribution networks integrated with rooftop photovoltaic systems. This research aims to minimize losses, enhance voltage profiles, and reduce carbon emissions. In [12], the authors have proposed a modified gradient-based optimization (MGBO) technique that combines the gradient search and local escape mechanism with a binomial crossover operator for distributed generator and shunt capacitor integration in distribution feeders. The MGBO approach was utilized to enhance the network performance by minimizing technical power losses at peak load intervals. However, a planning problem, in reality, should be considered as a multi-objective optimization problem. The authors in [13] suggested a genetic algorithm-based strategy for optimizing the location and capacity of distributed generators in medium-voltage and low-voltage networks while also assigning capacitor banks on the low-voltage side of medium-voltage/low-voltage transformers. This method decreases power losses and improves voltage profiles in low-voltage buses without adding to investment expenses. In this article, the assessment of a GA chromosome only used voltage levels and transformer core impedance to compute losses, neglecting the network's architecture and distribution line losses. In [14], the authors presented an augmented subtraction-average-based technique (ASABT) to reduce energy-dissipated losses in electrical power supply. This study examined two distinct sce-

narios. To reduce power losses, the recommended ASABT was applied to shunt capacitors exclusively in the first scenario. Scenario two involved placing and sizing both PV units and capacitors simultaneously, significantly reducing the substation emissions. The authors in [15] have presented a new multi-objective allocation method using an improved golden jackal optimization (IGJO) algorithm for numerous capacitor banks and multi-type DGs in a distribution network. The suggested technique improves the accuracy and speeds up the convergence of the golden jackal optimization by integrating memory-based models and a random walk strategy. To decrease the search space of the optimization process, the created IGJO used a reactive loss sensitivity index to select potential nodes for DGs and CBs installation. The formulated planning problem aimed to improve voltage profiles and stability and reduce overall active power loss. In [16], the authors have put forward a combination of meta-heuristic mixed-discrete modified Jaya optimization and analytical distributed Q-PQV bus pair method to optimally position the properly sized distributed generators and shunt capacitor banks. Moreover, the intermittencies intrinsic to load demand and solar and wind generation are modelled using a worst-case realization-based robust approach. The authors in [17] proposed a grey wolf optimization-based approach for the simultaneous placement of renewable energy sources and DSTATCOM in distribution systems, with the aim of enhancing voltage stability and reducing losses under load fluctuations.

Machine learning-based approaches have become more popular in recent times for forecasting load patterns and solar and wind power predictions and playing an important role in power distribution system planning. The authors in [18] compared the performance of three machine learning models based on linear regression, logarithmic linear regression, and nonlinear autoregression in predicting the energy demands of residential and commercial sectors and found autoregressive models are more accurate. They have considered the role of renewable energy sources, population, natural gas price, and electricity price in energy consumption. A distributed neural network model was proposed in [19] for optimally distributing the total power demand among dispatchable distributed energy resources in an isolated microgrid. In this approach, the authors used a hierarchical decentralized optimization architecture algorithm to generate the training data. Non-dispatchable energy sources, such as solar and wind DGs, are considered as negative loads. The authors in [20] used a hybrid machine learning model combining LSTM and RNN networks for forecasting solar irradiance. They used the K-means clustering algorithm to divide the data sets into sunny and cloudy days and developed separate forecasting models. In [21], the authors proposed deep learning-based models for predicting solar and wind power curtailments and utilizing excess sources to store energy in battery and hydrogen systems. Their analysis showed that the gated recurrent unit networks can predict better than other machine learning models. A deep learning technique combining artificial neural networks and a rough neuron water cycle algorithm was proposed in [22] for predicting wind speed, solar irradiance, and load demand in the first stage. In the second stage, they used a cooperative game approach to estimate the microgrid's daily dispatch and energy transactions. They concluded that deep learning-based estimations can reduce the operating cost compared to conventional stochastic processes. In [23], the authors proposed a variance correction technique-based machine learning model to predict power generation from renewable energy sources and load profiles. Their analysis suggests that nuclear energy support is necessary to counterbalance the intermittent nature of renewable energy sources and achieve the higher goals of carbon-free energy consumption. The authors in [24] developed long-short-term memory-based machine learning models for predicting the day-ahead electricity demand of commercial buildings and to identify the charging and discharging behaviour of electric vehicles to minimize the peak electricity demand. The authors in [25] used long-short-term memory-based deep learning network modeling to forecast load, solar irradiance, and wind speed and use the results to optimize the sizing of renewable energy sources for active distribution networks. The authors in [26] proposed an LSTM machine learning network-based probabilistic forecasting approach for optimal scheduling of renewable energy microgrids. The authors found that they can accurately predict

solar power over longer time horizons than wind power, necessitating non-uniform time scheduling for the planning of distribution systems with renewable energy sources. The authors in [27] reviewed various machine learning approaches available in the literature and identified them as highly effective for predicting electricity production from renewable energy sources. Different machine learning-based modeling techniques were discussed in [28] that can be applied for load forecasting and predicting the power supply availability from renewable energy sources. Their analysis revealed that it is important to overcome the problems of data unavailability and smart meters shortage in order to effectively use machine learning models. Statistical machine learning models were proposed in [29] for simulating weather-based sensitive loads such as heaters and air conditioners and also for intermittent renewable energy sources. The probabilistic load flow model, improved with the nearest neighbourhood approach, is used in conjunction with stochastic models for optimum planning of shunt capacitors for distribution systems. The authors in [30] have proposed an LSTM-based deep learning model for providing a very short-term power generation forecast of the Ninh Thuan solar power plant in Vietnam. Furthermore, alternative deep learning architectures, including recurrent neural networks (RNN) and convolutional neural networks (CNN), were compared with the LSTM networks to illustrate the superiority of LSTM over these models.

Some of the researchers also developed machine learning models for the optimal allocation of renewable energy sources, as well as capacitor banks, to improve the power quality of distribution networks. The authors in [31] developed a deep reinforcement learning-based volt-var optimization methodology for generating the optimum reactive power from DG smart inverters, capacitor banks, and voltage regulators for distribution systems. They have used a separate machine learning model at each node of the distribution system, considering the node voltage as the input for the model to control the volt-var devices. In [32], the authors proposed a decision tree-based classification model for the optimal placement of photovoltaic DGs, considering three performance metrics: node voltage risk index, corresponding branch power loss, and DG hosting capacity of the distribution system nodes. The methodology is mainly useful for allocating DGs of known capacity. Monte Carlo tree search-based reinforcement learning was proposed in [33] for the optimum sizing of energy storage systems, Static VAR Compensators, and capacity banks for active distribution systems. Their goal is to meet the desired performance of the distribution system with minimum load curtailment. In [34], the authors have developed deep neural network-based models for optimal planning of distributed energy resources and optimum network configuration using non-dominated sorting genetic algorithm II, considering resilience-based multi-objective functions. The methodology utilized a well-trained deep neural network to select an optimal solution, which includes the optimal sizing of renewable energy sources and switching combinations for various contingency conditions.

The methodologies proposed from [1–17] are mainly based on analytical and evolutionary computing techniques. The analytical techniques are not appropriate for multi-objective approaches that aim to place DG units and shunt capacitors simultaneously. The evolutionary computing techniques are based on stochastic parameters and require running the algorithm several times to confirm the global optimum value. The researchers from [18–30] mainly used machine learning algorithms for forecasting the energy demand and power supply available from renewable energy sources. The forecasted values only provide information about the maximum possible DG sizes, not the optimum sizing values. The authors in [31] used a machine learning model to find the best size for capacitor banks but didn't consider the distributed generations. Similarly, the decision tree-based method suggested in [32] can only find the best places for DGs of fixed sizes and not the best places for all DGs. The Monte Carlo tree search-based reinforcement learning methodology proposed in [33] requires repetition of the algorithm with load variations to find the optimum sizing of energy storage systems and capacitor banks. In order to validate the global optima of renewable energy sources under the necessary contingency circumstances, it is necessary

to identify performance indices with all potential values using the deep learning neural network-based methods suggested in [34].

Since the load profile varies with various factors like atmospheric temperature, customer demand, and availability of renewable energy sources, it is required to estimate the values of DGs and SHCs with the load factor variations. However, the machine learning-based methodologies proposed so far are inadequate, and evolutionary computing techniques require several attempts to find the global optimum at every desired load factor, which is difficult for a utility engineer. Therefore, the current work proposes a bidirectional long-short-term memory network (BiLSTM)-based machine learning model to assist utility engineers in directly determining the optimum sizing values of DGs and SHCs at any desired load factor without the need for load flow algorithms. The utility engineers should be provided with a software module developed based on the trained BiLSTM machine learning model for using this application for identifying the optimum sizing of DGs and SHCs at a desired load factor. The necessary data sets for training the BiLSTM model are acquired using the grey wolf optimization strategy at various samples.

The following are the major contributions of the proposed methodology.

- A GWO algorithm based multi-objective approach is developed for obtaining the optimum sizing of DGs and SHCs at different load factors.
- A BiLSTM network model is trained for direct determination of optimum sizing of DGs and SHCs without using load flows utilizing the data sets generated by the GWO algorithm.
- The performance of the BiLSTM network is tested with three different training algorithms known as root mean square propagation (rmsprop), stochastic gradient descent with momentum (sgdm) and adaptive moment estimation (adam).

The organization of the paper is as follows: Initially, the formulation of the multi-objective planning problem is discussed in Section 2. The grey wolf optimization algorithm for generating the input and target data for training the machine learning models is discussed in Section 3. The mathematical modeling of the machine learning networks and training algorithms are discussed in Section 4. The simulation studies are discussed in Section 5, and finally, the conclusions and future scope of this study are addressed in Section 6.

## 2. Formulation of the Multi-Objective Planning Problem

This section formulates the multi-objective planning framework for distributed generators and shunt capacitors using the grey wolf optimization (GWO) algorithm. The optimal sites and sizes of DGs and SHCs seek to reduce the apparent, real, and reactive power drawn from the substation while reducing real and reactive power losses, enhancing voltage profiles, and adhering to DG and reactive power penetration limits. The first three objective indices of substation apparent, real, and reactive power are evaluated to alleviate the burden on the power supply, while the subsequent three indices, real and reactive power loss and voltage profile, are assessed to enhance the steady-state performance of the distribution system. The last two objective indices, DG and reactive power penetrations, are considered to capture the optimal DG and SHC sizes at their respective load factors for real and reactive power, respectively. The objectives are normalized with their respective base values, and the weighted multi-objective function is formulated as illustrated below.

(i) Substation apparent power index

The substation apparent power index ($SSAI_{LF}$) is the ratio of substation apparent power considering DGs and SHCs to the ratio of S/S apparent power without considering the DGs and SHCs. The objective of the substation apparent power index is to minimize the overall apparent power delivered by the S/S with the optimum allocation of DGs and SHCs.

$$SSAI_{LF} = \frac{SSAP_{LF}^{DGQC}}{SSAP_{LF}^{Base}} \quad (1)$$

where $SSAI_{LF}$ is the S/S apparent power index at load factor LF. $SSAP_{LF}^{DGQC}$ and $SSAP_{LF}^{Base}$ are, respectively, the apparent power delivered by the S/S with the integration of DGs and SHCs and with base case values at load factor LF.

(ii) Substation real power index

The substation real power index ($SSPI_{LF}$) is defined as the ratio of real power delivered by the substation considering both DGs and SHCs to the power delivered without considering them.

$$SSPI_{LF} = \frac{SSPL_{LF}^{DGQC}}{SSPL_{LF}^{Base}} \tag{2}$$

$$SSPL_{LF}^{DGQC} = \sum_{i=2}^{NN} PL_{i,LF} + Ploss_{LF}^{DGQC} - \sum_{i=1}^{DGN} PDG_{i,LF} \tag{3}$$

$$SSPL_{LF}^{Base} = \sum_{i=2}^{NN} PL_{i,LF} + Ploss_{LF}^{Base} \tag{4}$$

Here, $SSPL_{LF}^{DGQC}$ and $SSPL_{LF}^{Base}$ are, respectively, the real power delivered by the S/S with integration of DGs and SHCs and with base case values at load factor LF. In the above equations, NN is the number of nodes of the distribution system, and DGN is the number of distributed generations.

(iii) Substation reactive power index

Substation reactive power index ($SSQI_{LF}$) is defined as the ratio of reactive power delivered by the substation considering DGs and SHCs to without considering DGs and SHCs.

$$SSQI_{LF} = \frac{SSQL_{LF}^{DGQC}}{SSQL_{LF}^{Base}} \tag{5}$$

$$SSQL_{LF}^{DGQC} = \sum_{i=2}^{NN} QL_{i,LF} + Qloss_{LF}^{DGQC} - \sum_{i=1}^{DGN} QDG_{i,LF} - \sum_{i=1}^{SCN} QSC_{i,LF} \tag{6}$$

$$SSQL_{LF}^{Base} = \sum_{i=2}^{NN} QL_{i,LF} + Qloss_{LF}^{Base} \tag{7}$$

where $SSQL_{LF}^{DGQC}$ and $SSQL_{LF}^{Base}$ are, respectively, the reactive power delivered by the S/S with the integration of DGs and SHCs and with base case values at load factor LF. In Equation (6), SCN is the number of shunt capacitors.

(iv) Real power loss index

Real power loss index ($PLoss_{LF}$) is defined as the ratio of the real power loss of the distribution system with DGs and SHCs to the real power loss with base case valueswhere $PLoss_{LF}^{DGQC}$ and $PLoss_{LF}^{Base}$ are the real power loss with the impact of DGs and SHCs and with base case values at load factor LF.

$$PLoss_{LF} = \frac{PLoss_{LF}^{DGQC}}{PLoss_{LF}^{Base}} \tag{8}$$

(v) Reactive power loss index

Reactive power loss index ($QLoss_{LF}$) is defined as the ratio of reactive power loss of the distribution system with DGs and SHCs to the reactive power loss with base case values. $QLoss_{LF}^{DGQC}$ and $QLoss_{LF}^{Base}$ are the reactive power loss with the impact of DGs and SHCs and with base case values at load factor LF.

$$QLoss_{LF} = \frac{QLoss_{LF}^{DGQC}}{QLoss_{LF}^{Base}} \tag{9}$$

(vi) Voltage profile index

The voltage profile index ($VDI_{LF}$) is defined as the ratio of the square root of the sum of the squares of the distribution system voltage deviations from unity with the integration of DGs and SHCs to the square root of the sum of the voltage deviations corresponding to the base case.

$$VDI_{LF} = \sqrt{\frac{\sum_{i=1}^{NN}\left(1 - V_{i,LF}^{DGQC}\right)^2}{\sum_{i=1}^{NN}\left(1 - V_{i,LF}^{Base}\right)^2}} \tag{10}$$

$V_{i,LF}^{DGQC}$ and $V_{i,LF}^{Base}$ are the voltage of the ith node of the distribution system considering the impact of DGs and SHCs and with base case values.

(vii) DG penetration index ($DGI_{LF}$)

The DG penetration index at a load factor is defined as the ratio of total DG sizing to the total real power load of the distribution system at that particular load factor.

$$DGI_{LF} = \frac{\sum_{k=1}^{DGN} DG_{k,LF}}{\sum_{i=2}^{NN} PL_{i,LF}} \tag{11}$$

where $DG_{k,LF}$ and $DGN$ are, respectively, the optimum sizing of kth DG at load factor $LF$ and the total number of DGs installed.

(viii) Reactive power penetration index ($QCI_{LF}$)

The reactive power index is considered as the ratio of total reactive power injected to the total reactive power load.

$$QCI_{LF} = \frac{\sum_{i=1}^{DGN} QDG_{i,LF} + \sum_{j=1}^{SCN} QSC_{j,LF}}{\sum_{i=2}^{NN} QL_{i,LF}} \tag{12}$$

The numerator of the index comprises total reactive power delivered by the DGs and total reactive power injected by the shunt capacitors. The denominator indicates the total reactive power load at that particular load factor. The optimal sizes of the DGs and SHCs explicitly provide economically optimal values of the DGs and SHCs since the costs of these components are directly related to their sizing values. Utility engineers have the flexibility to choose from a variety of DGs according to their availability.

The multi-objective function can be formulated using the indices developed by the following equation.

$$J_{LF} = w_1\, SSAI_{LF} + w_2\, SSPI_{LF} + w_3 SSQI_{LF} + w_4 PLoss_{LF} + w_5 QLoss_{LF} + w_6 VDI_{LF} \quad + w_7 DGI_{LF} \quad + w_8 QCI_{LF} \tag{13}$$

Depending on the needs at the substation, utility engineers can choose different weights in the weighted sum multi-objective function, such as how much weight to give to loss reduction, voltage profile improvement, DG sizing, or shunt capacitor sizing. Here, in the multi-objective function (Equation (13)), all the weights are set to identical values since all the goals are deemed equally important.

## 3. Optimum Placement of DGs and SHCs Using Grey Wolf Optimization Algorithm

The optimum sizing of DGs and SHCs is obtained using the Grey Wolf Optimization algorithm using the multi-objective function delineated in Equation (13).

### *3.1. Grey Wolf Optimization Algorithm*

The grey wolf optimization technique is a meta-heuristic algorithm [35] inspired by the leadership hierarchy and hunting mechanism of grey wolves. The mathematical modelling of the GWO technique is designed based on the hunting and social hierarchy

procedure followed by the grey wolves, and they are divided into four subsections. The four subsections are known as social hierarchy: tracking, encircling, and attacking prey. In the present section, we discuss the development of the GWO algorithm based on these four steps for optimal placement of DGs and SHCs.

In the social hierarchy, the first three best solutions are considered to guide the rest of the population. The first three best solutions are named as alpha ($\alpha$), beta ($\beta$) and delta ($\delta$). The first three best candidates $\alpha$, $\beta$ and $\delta$ guide the remaining set of candidates in hunting their prey.

The grey wolves encircling strategy of their prey and hunting techniques can be modeled mathematically using the following equations.

The equations from Equation (14) to Equation (17) represent the formulas for obtaining the first best solution.

$$S_\alpha = |R_1 \, X_\alpha(t) - X(t)| \tag{14}$$

$$X_1(t+1) = X_\alpha(t) - P_1 \quad S_\alpha \tag{15}$$

$$P_1 = 2 \, b \, r_{11} - b \tag{16}$$

$$R_1 = 2 \, r_{12} \tag{17}$$

The second best solution can be given by the following equations from Equation (18) to Equation (21).

$$S_\beta = \left|R_2 \, X_\beta(t) - X(t)\right| \tag{18}$$

$$X_2(t+1) = X_\beta(t) - P_2 \quad S_\beta \tag{19}$$

$$P_2 = 2 \, b \, r_{21} - b \tag{20}$$

$$R_2 = 2 \, r_{22} \tag{21}$$

The third best solution can be given by the following equations from Equation (22) to Equation (25).

$$S_\delta = |R_3 \, X_\delta(t) - X(t)| \tag{22}$$

$$X_3(t+1) = X_\delta(t) - P_3 \quad S_\delta \tag{23}$$

$$P_3 = 2 \, b \, r_{31} - b \tag{24}$$

$$R_3 = 2 \, r_{32} \tag{25}$$

where $b$ is a linearly decreasing function from 2 to zero as the number of iterations increases and $r_{11}$, $r_{12}$, $r_{21}$, $r_{22}$, $r_{31}$ and $r_{32}$ are the random numbers whose values are randomly selected between 0 and 1. The sets $X_1$, $X_2$ and $X_3$ are updated values from the first three best solutions using Grey Wolf Optimizer.

The final best solution can be taken as the average of the updated sets of $X_1$, $X_2$ and $X_3$ and can be expressed by the following Equation (26).

$$X(t+1) = \frac{X_1(t+1) + X_2(t+1) + X_3(t+1)}{3} \tag{26}$$

The algorithm steps are repeated using equations from Equation (14) to Equation (26) with the updated population until convergence to a global minimum is achieved.

### *3.2. Optimum Sizing of DGs and SHCs Using GWO*

Most evolutionary techniques require the proper selection of many algorithm parameters to achieve the global optimum. The grey wolf optimization algorithm requires only the selection of two parameters: population size and no of iterations. Hence, the authors selected this algorithm as a tool for generating the target data set for training the machine learning models.

The optimum sizing values are obtained considering the weighted sum multi-objective function described by Equation (13) as the fitness function. The optimum sizing data of

DGs and SHCs required for training the machine learning models at selected load factors is obtained using GWO. The flow chart describing the algorithm is shown in the following Figure 1.

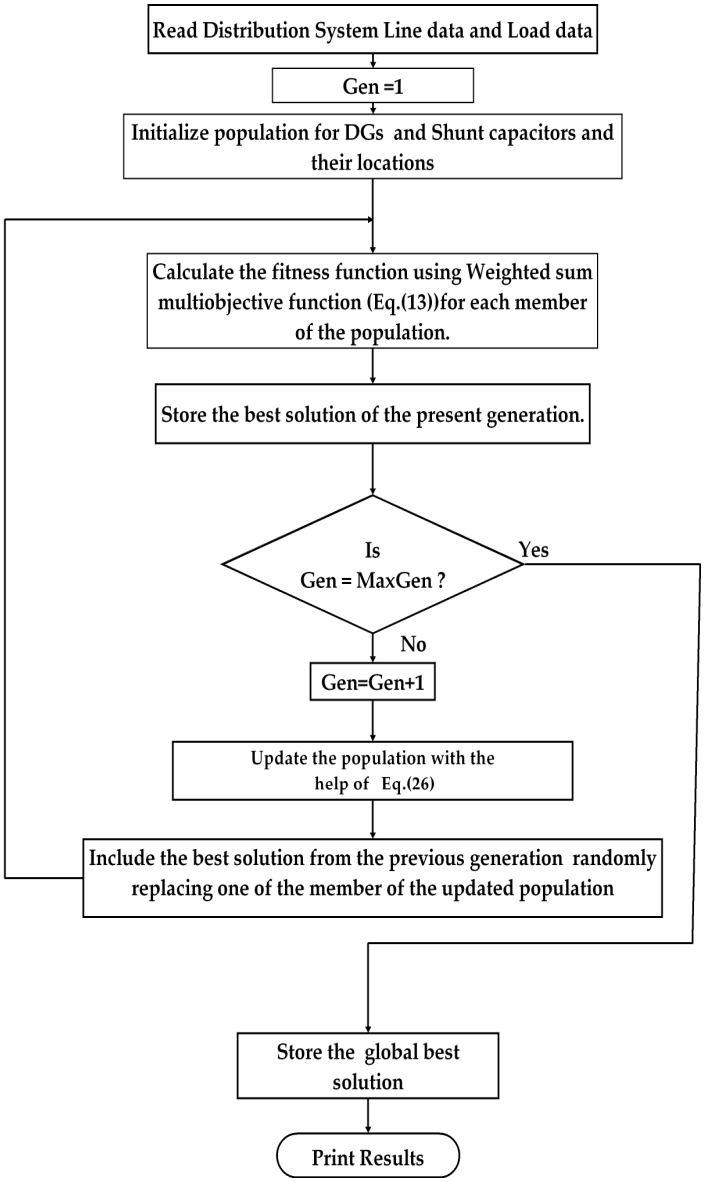


**Figure 1.** GWO algorithm for optimum sizing of DGs and SHCs.

## 4. Mathematical Modeling of Machine Learning Networks

Machine learning models are advanced versions of neural networks and are more efficient in processing sequential data and time series data for applications ranging from load forecasting to image classification. Many machine learning models have been developed recently, and among them, long short term memory networks (LSTM) have become more popular, and Bidirectional Long Short Term Memory (BiLSTM) networks are advanced versions of LSTM networks [36].

LSTM networks are enhanced versions of recurrent neural networks to avoid the problem of vanishing gradients and to improve the accuracy in predicting long-range dependent sequences.

### *4.1. Long Short-Term Memory Networks*

Long short-term memory networks can be represented by the following architecture and mathematical equations. The LSTM block will have a cell state and hidden state computed using three types of gates known as forget gate, input gate and output gate, as shown in Figure 2. The gate values are computed from the previous hidden and cell states and input sequence. The forget gate eliminates unnecessary information from the cell state, and the input gate regulates the addition of useful information. Finally, the hidden state is computed from the output gate.

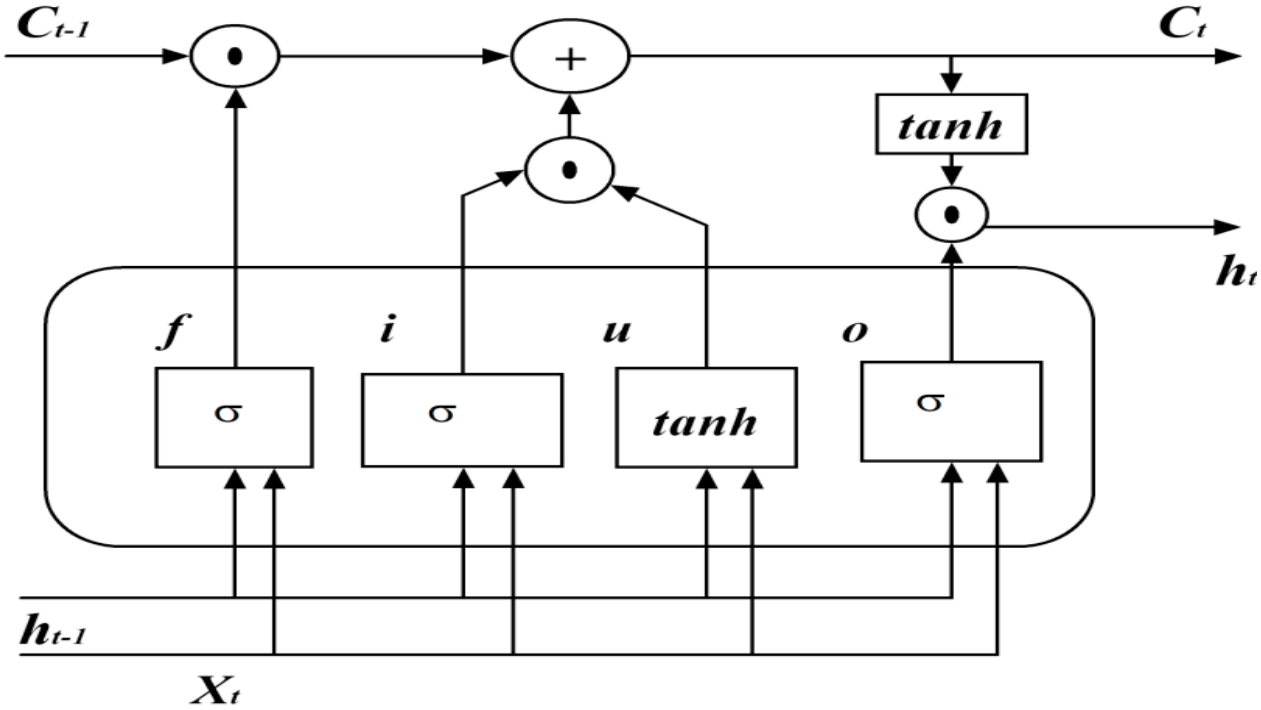


**Figure 2.** LSTM network block diagram.

The hidden states of the LSTM block can be computed by the following equations

$$f_t = \sigma\left(IW_f X_t + RW_f h_{t-1} + b_f\right) \tag{27}$$

$$i_t = \sigma(IW_i X_t + RW_i h_{t-1} + b_i) \tag{28}$$

$$u_t = tanh(IW_u X_t + RW_u h_{t-1} + b_u) \tag{29}$$

$$o_t = \sigma(IW_o X_t + RW_o h_{t-1} + b_o) \tag{30}$$

where $\sigma$ and $tanh$ are sigmoid and hyperbolic tangent functions. The $f_t$, $i_t$, $u_t$ and $o_t$ of the above equations are the forget gate, input gate, cell state and output gate, respectively. Input weights, recurrent weight and bias values are represented by $IW$, $RW$ and $b$.

The following are required equations for computing cell state and hidden state.

$$C_t = f_t \odot C_{t-1} + i_t \odot u_t \tag{31}$$

$$h_t = o_t \odot tanh(C_t) \tag{32}$$

The symbol $\odot$ is the Hadamard product used for the element-wise multiplication of vectors.

### *4.2. Bidirectional Long Short-Term Memory Networks*

The Bidirectional LSTM (BiLSTM) network is developed from the LSTM network by combining the sequences in the forward and backward directions to improve the convergence of the model. The Bi-LSTM network architecture is shown in Figure 3.

$$y_t = \sigma\left(U_f \overrightarrow{h}_t + U_b \overleftarrow{h}_t + b_y\right) \tag{33}$$

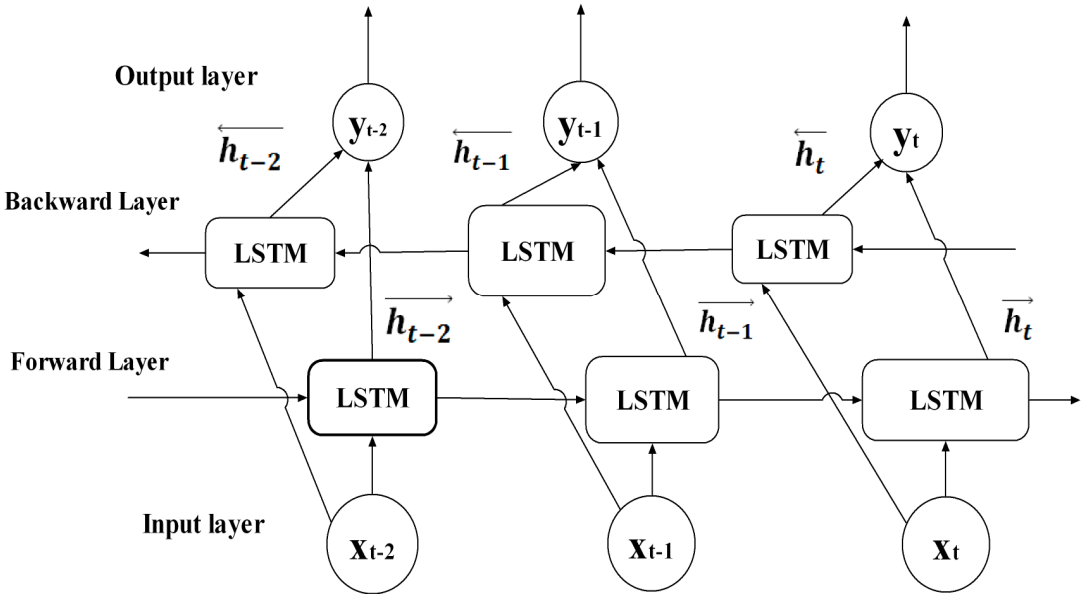


**Figure 3.** BiLSTM network architecture.

The output of the BiLSTM is calculated from the above equation, where $\overrightarrow{h}_t$ and $\overleftarrow{h}_t$ are the forward and backward outputs and $U_f$, $U_b$ are the weights of the outputs. $b_y$ is the bias value of the output.

### *4.3. Adaptive Moment Estimation Optimizer (Adam)*

The machine learning network models are trained using the Adam moment estimation optimization algorithm [37], also known as the Adam optimizer. The Adam optimizer method updates the parameters of the network model from the moments of stochastic objective function gradients.

$$gm_t = \delta_1 gm_{t-1} + (1-\delta_1)\nabla E(\phi_t) \tag{34}$$

$$gv_t = \delta_2 gv_{t-1} + (1-\delta_2)[\nabla \mathrm{E}(\phi_t)]^2 \tag{35}$$

The $gm_t$ and $gv_t$ are the moments and $\delta_1$ and $\delta_2$ are the exponential decay factors. $E(\phi_t)$ is the objective function and $\phi_t$ is the weight parameter of the network to be updated.

$$g\hat{m}_t = \frac{gm_t}{(1-\delta_1)} \tag{36}$$

$$g\hat{v}_t = \frac{gv_t}{(1-\delta_2)} \tag{37}$$

$$\phi_t = \phi_{t-1} - \frac{\propto g\hat{m}_t}{\sqrt{g\hat{v}_t} + \epsilon} \tag{38}$$

The first moment and second moments, $gm_t$ and $gv_t$, are calculated based on the gradient and gradient square values of the objective function selected. These moving averages are then used to update the weight parameters. The parameters are finally updated using equations from Equation (36) to Equation (38).

### *4.4. Performance Metric of the Machine Learning Model*

The performance of the machine learning model networks can be evaluated using the root mean square error (RMSE) function. The better the convergence of these two functions indicates, the better the accuracy of the trained machine-learning models will be.

The root mean square error function (RMSE) mainly gives a measure of the deviation of the output determined by the trained network and the actual value of the output.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left(y_{ij} - t_{ij}\right)^2}{N}} \tag{39}$$

The above equation describes the RMSE function, where $t_{ij}$ represents the target value of the ith sample for the jth output member and $y_{ij}$ is the value of the ith sample for the jth output member determined by the trained network. M in the above equation is the total number of elements in a target data set of the model, and N is the number of samples. The convergence of the RMSE value to the best possible lowest value is the measure of the performance of the trained network for providing accurate predicted values.

### *4.5. GWO Algorithm-Based Machine Learning Model for Optimum Sizing of DGs and SHCs*

Machine learning models are proposed in this work for computing the optimum sizing of DGs and SHCs. In the present work, initially, the optimum locations and sizing of the DGs and SHCs are identified using the grey wolf optimization algorithm considering the weighted sum multi-objective function described by Equation (13) as a fitness function at the unity load factor. The optimum sizing of the DGs and SHCs are obtained at the same locations at the other load factors considering the same multi-objective function. The input data for the machine learning model is determined using a forward/backward sweep load flow algorithm for the distribution systems. The corresponding optimum values of the DGs and Shunt capacitors are obtained using the grey wolf optimization algorithm and used as the target data for the machine learning model. The machine learning model is trained considering the vectors X and Y as the input and target data to the machine learning network model. The input vector X consists of substation apparent power, real power load, reactive power load, real power loss, reactive power loss and minimum voltage. The DG sizing values and shunt capacitor sizing values are considered as the target vector Y.

The machine learning model will be trained with the input data (X) and target data (Y) using the Adam optimization algorithm described in Section 4.3. The root mean square error (RMSE) function described in Section 4.4 is considered to be the performance metric for evaluating the performance of the machine learning model. The trained machine learning model can be used for obtaining optimum sizing of DGs and shunt capacitors directly without using the grey wolf optimization algorithm and load flow solutions at any desired load factor. In this work, the machine learning models such as LSTM and BiLSTM are trained and compared. The overall idea of the application of the proposed methodology is shown in Figure 4.

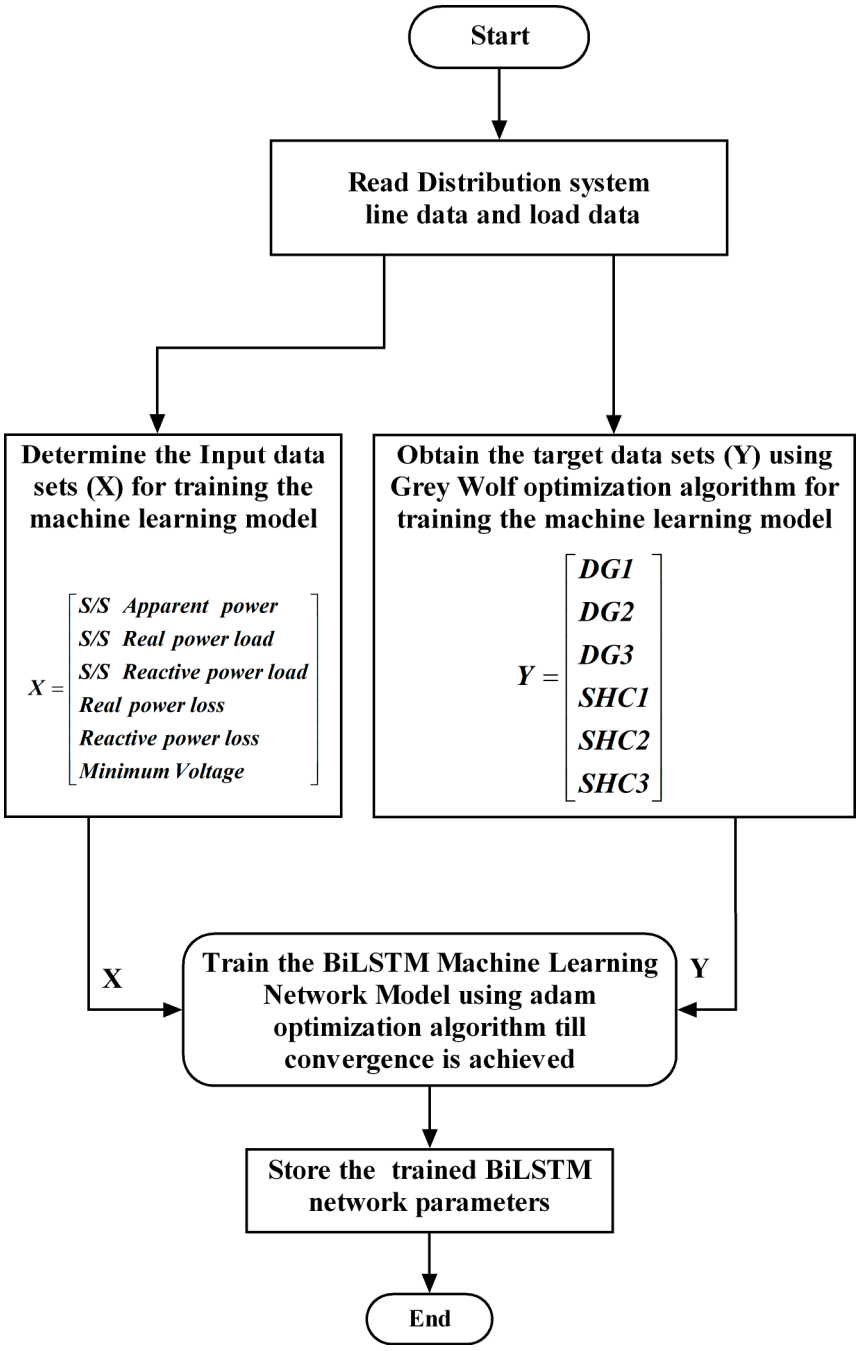


**Figure 4.** Machine learning model for DGs and SHCs computation using GWO.

## 5. Results and Discussions

The proposed methodology is tested on 11 kV, 51 bus [1] and 12.66 kV, 69 bus [38] distribution systems and the analysis is presented in this section. The network diagrams of the 51-bus and 69-bus test systems are shown in Figures 5 and 6, respectively.

The Grey wolf optimization algorithm-based multi-objective approach considering Equation (13) as a fitness function is used at first for obtaining optimum locations and sizing of DGs and SHCs at unity load factor. The optimum nodes identified for the 51-bus system for DG placement are 9, 42 and 5, and for shunt capacitors placement, they are 14, 25 and 36. The optimum nodes for DG placement in the 69 bus system are 61, 5, and 16, and for the shunt capacitors placement, they are 11, 62, and 65. Tables 1 and 2 show the base case values of 51 and 69 bus test systems, the input data required for the machine learning models.

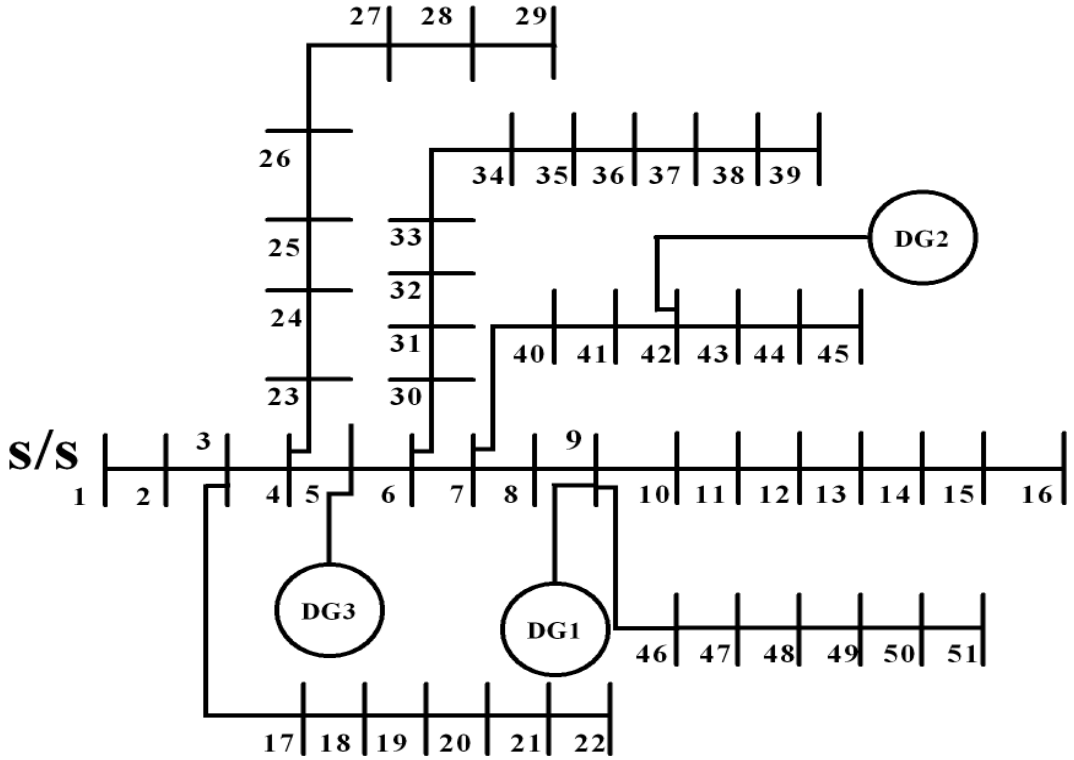


**Figure 5.** 51-bus radial distribution system.

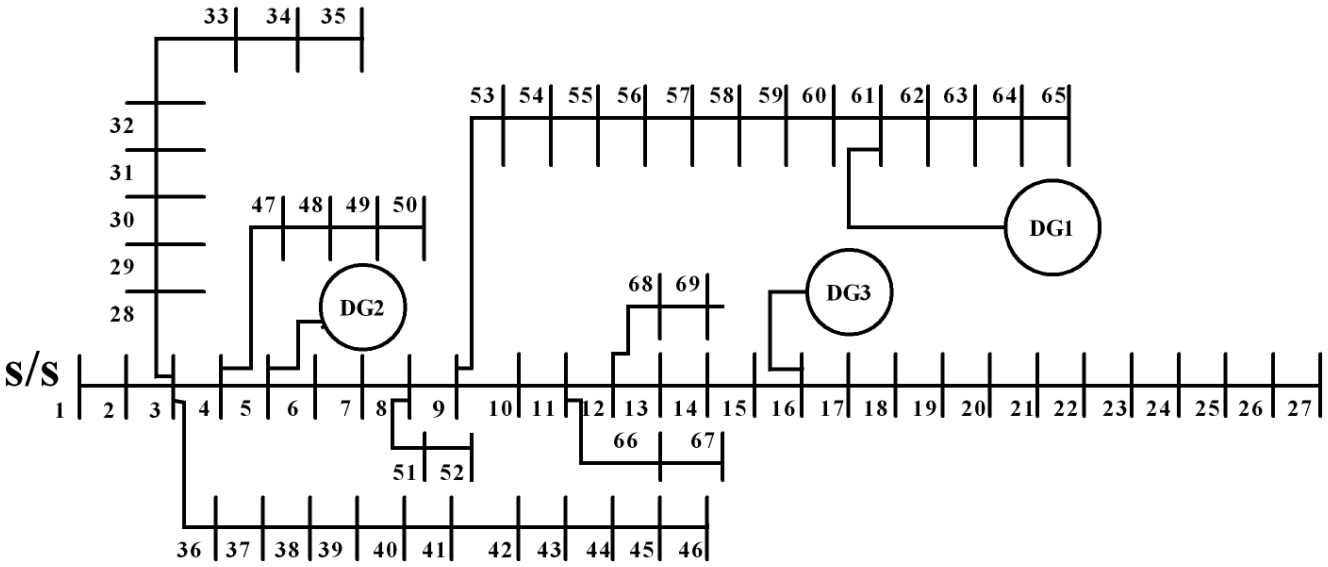


**Figure 6.** 69-bus radial distribution system.

The optimum sizing values of DGs and shunt capacitors at other load factors are identified, considering the same optimum locations and the same fitness function using the GWO algorithm. In this work, it is assumed that the DGs are operating at 0.95 lagging power factor. The optimum values of DGs and SHCs obtained at various load factors are shown in Tables 3 and 4 for the 51-bus system and 69-bus system, respectively.

**Table 1.** 51-bus system without DGs and SHCs.

| Load Factor | S/S Apparent Power (kVA) | Real Power Load (kW) | Reactive Power Load (kVAr) | Real Power Loss (kW) | Reactive Power Loss (kVAr) | Minimum Node Voltage (kV) |
|---|---|---|---|---|---|---|
| 0.40 | 1193.11 | 985.20 | 627.60 | 19.10 | 16.51 | 10.61 |
| 0.45 | 1345.96 | 1108.35 | 706.05 | 24.33 | 21.02 | 10.56 |
| 0.50 | 1499.69 | 1231.50 | 784.50 | 30.23 | 26.12 | 10.51 |
| 0.55 | 1654.32 | 1354.65 | 862.95 | 36.82 | 31.80 | 10.46 |
| 0.60 | 1809.87 | 1477.80 | 941.40 | 44.11 | 38.09 | 10.41 |
| 0.65 | 1966.35 | 1600.95 | 1019.85 | 52.12 | 45.00 | 10.36 |
| 0.70 | 2123.79 | 1724.10 | 1098.30 | 60.85 | 52.53 | 10.31 |
| 0.75 | 2282.20 | 1847.25 | 1176.75 | 70.34 | 60.70 | 10.26 |
| 0.80 | 2441.60 | 1970.40 | 1255.20 | 80.58 | 69.53 | 10.20 |
| 0.85 | 2602.03 | 2093.55 | 1333.65 | 91.61 | 79.02 | 10.15 |
| 0.90 | 2763.50 | 2216.70 | 1412.10 | 103.43 | 89.20 | 10.10 |
| 0.95 | 2926.03 | 2339.85 | 1490.55 | 116.07 | 100.08 | 10.04 |
| 1.00 | 3089.67 | 2463.00 | 1569.00 | 129.56 | 111.68 | 9.99 |
| 1.05 | 3254.41 | 2586.15 | 1647.45 | 143.89 | 124.01 | 9.93 |
| 1.10 | 3420.30 | 2709.30 | 1725.90 | 159.10 | 137.09 | 9.88 |
| 1.15 | 3587.36 | 2832.45 | 1804.35 | 175.21 | 150.93 | 9.82 |
| 1.20 | 3755.63 | 2955.60 | 1882.80 | 192.25 | 165.56 | 9.77 |
| 1.25 | 3925.13 | 3078.75 | 1961.25 | 210.23 | 181.00 | 9.71 |
| 1.30 | 4095.89 | 3201.90 | 2039.70 | 229.18 | 197.27 | 9.65 |

**Table 2.** 69-bus system without DGs and SHCs.

| Load Factor | S/S Apparent Power (kVA) | Real Power Load (kW) | Reactive Power Load (kVAr) | Real Power Loss (kW) | Reactive Power Loss (kVAr) | Minimum Node Voltage (kV) |
|---|---|---|---|---|---|---|
| 0.40 | 1899.21 | 1520.88 | 1077.84 | 32.51 | 14.85 | 12.23 |
| 0.45 | 2141.89 | 1710.99 | 1212.57 | 41.47 | 18.93 | 12.17 |
| 0.50 | 2385.85 | 1901.10 | 1347.30 | 51.61 | 23.55 | 12.11 |
| 0.55 | 2631.11 | 2091.20 | 1482.03 | 62.95 | 28.71 | 12.05 |
| 0.60 | 2877.70 | 2281.31 | 1616.76 | 75.53 | 34.44 | 12.00 |
| 0.65 | 3125.66 | 2471.42 | 1751.49 | 89.38 | 40.73 | 11.94 |
| 0.70 | 3375.03 | 2661.53 | 1886.22 | 104.54 | 47.61 | 11.88 |
| 0.75 | 3625.85 | 2851.64 | 2020.95 | 121.03 | 55.10 | 11.82 |
| 0.80 | 3878.15 | 3041.75 | 2155.68 | 138.90 | 63.20 | 11.76 |
| 0.85 | 4131.97 | 3231.86 | 2290.41 | 158.20 | 71.94 | 11.70 |
| 0.90 | 4387.37 | 3421.97 | 2425.14 | 178.95 | 81.34 | 11.64 |
| 0.95 | 4644.39 | 3612.08 | 2559.87 | 201.20 | 91.41 | 11.57 |
| 1.00 | 4903.08 | 3802.19 | 2694.60 | 225.00 | 102.17 | 11.51 |
| 1.05 | 5163.49 | 3992.30 | 2829.33 | 250.40 | 113.64 | 11.45 |
| 1.10 | 5425.67 | 4182.41 | 2964.06 | 277.45 | 125.85 | 11.38 |

**Table 2.** *Cont.*

| Load Factor | S/S Apparent Power (kVA) | Real Power Load (kW) | Reactive Power Load (kVAr) | Real Power Loss (kW) | Reactive Power Loss (kVAr) | Minimum Node Voltage (kV) |
|---|---|---|---|---|---|---|
| 1.15 | 5689.69 | 4372.52 | 3098.79 | 306.21 | 138.81 | 11.32 |
| 1.20 | 5955.60 | 4562.63 | 3233.52 | 336.72 | 152.56 | 11.25 |
| 1.25 | 6223.48 | 4752.74 | 3368.25 | 369.06 | 167.12 | 11.18 |
| 1.30 | 6493.39 | 4942.85 | 3502.98 | 403.29 | 182.52 | 11.12 |

**Table 3.** Optimum sizes of DGs and SHCs for 51-bus system.

| Load Factor | DG1 (kW) | DG2 (kW) | DG3 (kW) | SHC1 (kVAr) | SHC2 (kVAr) | SHC3 (kVAr) |
|---|---|---|---|---|---|---|
| 0.40 | 394.08 | 188.03 | 394.08 | 111.49 | 89.28 | 106.60 |
| 0.45 | 443.34 | 211.61 | 443.34 | 125.17 | 101.01 | 119.76 |
| 0.50 | 492.60 | 235.19 | 492.60 | 139.27 | 112.07 | 132.98 |
| 0.55 | 541.86 | 258.87 | 541.86 | 153.34 | 122.86 | 146.74 |
| 0.60 | 591.12 | 282.67 | 591.12 | 167.12 | 133.33 | 161.03 |
| 0.65 | 640.38 | 306.56 | 640.38 | 181.38 | 145.57 | 173.30 |
| 0.70 | 689.64 | 330.13 | 689.64 | 195.42 | 156.52 | 186.59 |
| 0.75 | 738.90 | 354.65 | 738.90 | 209.02 | 167.73 | 200.37 |
| 0.80 | 788.16 | 378.43 | 788.16 | 223.13 | 178.95 | 213.62 |
| 0.85 | 837.42 | 402.50 | 837.42 | 237.32 | 189.07 | 227.58 |
| 0.90 | 886.68 | 426.17 | 886.68 | 251.54 | 200.11 | 241.11 |
| 0.95 | 935.94 | 450.63 | 935.94 | 265.08 | 211.83 | 254.45 |
| 1.00 | 985.2 | 474.46 | 985.2 | 279.38 | 222.10 | 268.36 |
| 1.05 | 1034.46 | 499.06 | 1034.46 | 293.22 | 233.54 | 281.15 |
| 1.10 | 1083.72 | 522.64 | 1083.72 | 308.22 | 244.43 | 294.06 |
| 1.15 | 1132.98 | 547.55 | 1132.98 | 322.09 | 255.18 | 307.81 |
| 1.20 | 1182.24 | 571.80 | 1182.24 | 336.03 | 266.02 | 321.39 |
| 1.25 | 1231.5 | 595.79 | 1231.5 | 350.41 | 276.88 | 334.65 |
| 1.30 | 1280.76 | 619.36 | 1280.76 | 363.85 | 287.58 | 349.25 |

**Table 4.** Optimum sizes of DGs and SHCs for 69-bus system.

| Load Factor | DG1 (kW) | DG2 (kW) | DG3 (kW) | SHC1 (kVAr) | SHC2 (kVAr) | SHC3 (kVAr) |
|---|---|---|---|---|---|---|
| 0.40 | 608.35 | 608.35 | 257.00 | 238.38 | 284.47 | 56.07 |
| 0.45 | 684.39 | 684.39 | 289.02 | 268.62 | 317.82 | 66.23 |
| 0.50 | 760.44 | 760.44 | 321.55 | 297.62 | 356.27 | 69.65 |
| 0.55 | 836.48 | 836.48 | 353.63 | 327.43 | 388.92 | 79.68 |
| 0.60 | 912.53 | 912.53 | 385.53 | 358.08 | 421.24 | 89.43 |
| 0.65 | 988.57 | 988.57 | 418.12 | 386.56 | 464.04 | 90.21 |
| 0.70 | 1064.61 | 1064.61 | 450.26 | 415.88 | 492.66 | 104.57 |
| 0.75 | 1140.66 | 1140.66 | 482.52 | 445.62 | 530.86 | 108.69 |
| 0.80 | 1216.70 | 1216.70 | 514.58 | 476.14 | 570.43 | 111.27 |
| 0.85 | 1292.74 | 1292.74 | 546.29 | 512.47 | 604.44 | 120.93 |

**Table 4.** *Cont.*

| Load Factor | DG1 (kW) | DG2 (kW) | DG3 (kW) | SHC1 (kVAr) | SHC2 (kVAr) | SHC3 (kVAr) |
|---|---|---|---|---|---|---|
| 0.90 | 1368.79 | 1368.79 | 578.94 | 534.91 | 636.77 | 130.94 |
| 0.95 | 1444.83 | 1444.83 | 611.45 | 565.27 | 670.44 | 139.51 |
| 1.00 | 1520.88 | 1520.88 | 643.62 | 596.27 | 706.80 | 144.29 |
| 1.05 | 1596.92 | 1596.92 | 675.98 | 623.99 | 739.16 | 155.90 |
| 1.10 | 1672.96 | 1672.96 | 707.64 | 655.35 | 778.54 | 157.85 |
| 1.15 | 1749.01 | 1749.01 | 740.30 | 683.50 | 812.53 | 168.12 |
| 1.20 | 1825.05 | 1825.05 | 772.56 | 712.13 | 850.19 | 173.11 |
| 1.25 | 1901.10 | 1901.10 | 804.69 | 744.06 | 888.73 | 175.07 |
| 1.30 | 1977.14 | 1977.14 | 836.90 | 772.67 | 930.79 | 176.64 |

Tables 5 and 6 illustrate the performance of the distribution system for select samples before and after the installation of distributed generators and shunt capacitors, therefore affirming the improvement in distribution system performance resulting from the integration of DGs and SHCs. Figure 7 displays the comparison of the minimum voltage profiles across various systems, revealing a significant rise in voltage levels with the use of DGs and shunt capacitors at all load factors.

**Table 5.** 51-bus system with DGs and SHCs.

| Load Factor | Real Power Loss (kW) | | Reactive Power Loss (kVAr) | |
|---|---|---|---|---|
| | Without DGs and SHCs | With DGs and SHCs | Without DGs and SHCs | With DGs and SHCs |
| 0.40 | 19.10 | 4.58 | 16.51 | 1.78 |
| 0.60 | 44.11 | 10.34 | 38.09 | 4.02 |
| 0.80 | 80.58 | 18.46 | 69.53 | 7.16 |
| 1.00 | 129.56 | 28.96 | 111.68 | 11.21 |
| 1.20 | 192.25 | 41.90 | 165.56 | 16.18 |

**Table 6.** 69-bus system with DGs and SHCs.

| Load Factor | Real Power Loss (kW) | | Reactive Power Loss (kVAr) | |
|---|---|---|---|---|
| | Without DGs and SHCs | With DGs and SHCs | Without DGs and SHCs | With DGs and SHCs |
| 0.40 | 32.51 | 1.59 | 14.85 | 1.46 |
| 0.60 | 75.53 | 3.59 | 34.44 | 3.29 |
| 0.80 | 138.90 | 6.39 | 63.20 | 5.87 |
| 1.00 | 225.00 | 9.99 | 102.17 | 9.19 |
| 1.20 | 336.72 | 14.41 | 152.56 | 13.26 |

The data shown in Tables 1 and 3 are the input and target data sets for the 51-bus system for training the machine learning models. Similarly, the data shown in Tables 2 and 4 are input and target data sets for the 69-bus system for training the machine learning models. The per unit values of the base case values of substation apparent power, real power load, reactive power load, real power loss, reactive power loss and minimum voltage are taken as the input vector elements for training the LSTM and BiLSTM network models. The substation's apparent power at the unity load factor and the base voltage

values are considered as base values for generating the per unit values at all the load factors. The base values of apparent power and voltage are 3089.67 kVA, 11 kV for the 51-bus system, and 4903.08 kVA and 12.66 kV for the 69-bus system, respectively. The models are trained with the three different training algorithms known as root mean square propagation (rmsprop), stochastic gradient descent with momentum (sgdm) and adaptive moment estimation (adam). The LSTM and BiLSTM models are trained using functions of the deep learning toolbox of MATLAB R2021a software. The number of hidden layers used for both the models is two, and for each hidden layer, 30 neurons are considered.

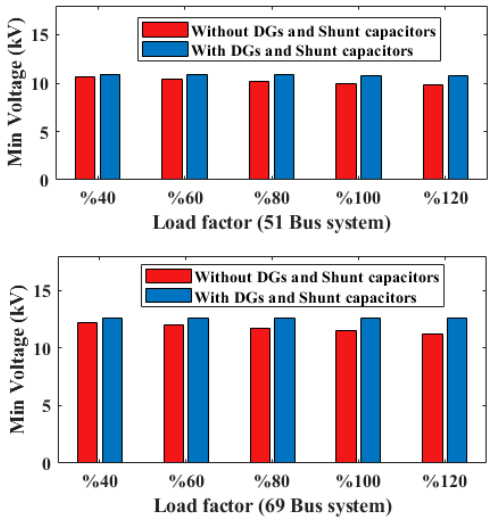


**Figure 7.** Voltage improvement comparison.

From Table 7 it can be realized that adam optimizer algorithm is more efficient in achieving least root mean square error value (RMSE) compared to other two algorithms rmsprop and sgdm. It is also observed that the BiLSTM network is slightly better than LSTM network and has lesser RMSE value. The convergence characteristics of the different optimization algorithms are shown in Figures 8 and 9 respectively for 51 and 69 bus systems for BiLSTM network model.

**Table 7.** Comparison of RMSE between LSTM and BiLSTM networks.

| Name of the Solver | 51 Bus System | | 69 Bus System | |
|---|---|---|---|---|
| | LSTM | BiLSTM | LSTM | BiLSTM |
| rmsprop | 0.01528 | 0.01708 | 0.01362 | 0.01845 |
| sgdm | 0.00519 | 0.00113 | 0.00294 | 0.00306 |
| adam | 0.00158 | 0.00107 | 0.00220 | 0.00129 |

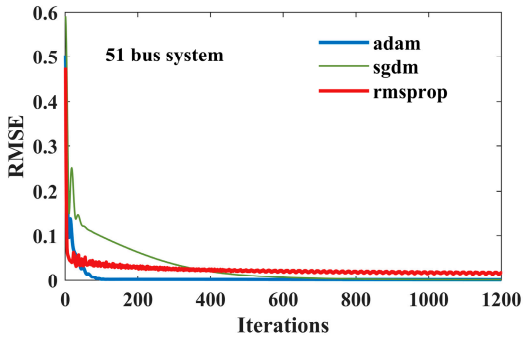


**Figure 8.** Comparison of convergence characteristics of RMSE for 51-bus system.

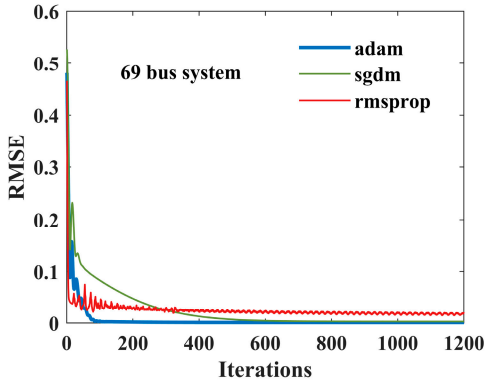


**Figure 9.** Comparison of convergence characteristics of RMSE for 69-bus.

From the convergence characteristics shown in Figures 8 and 9 for the BiLSTM network model for 51 and 69 bus systems, it can be said that the Adam optimizer algorithm is much better in terms of fast convergence and lesser RMSE value. The lesser the RMSE value of the model, the greater its efficiency. The total DG capacity of three DG sources calculated from a multi-objective based GWO algorithm and BiLSTM-trained network are compared and shown in Figures 10 and 11.

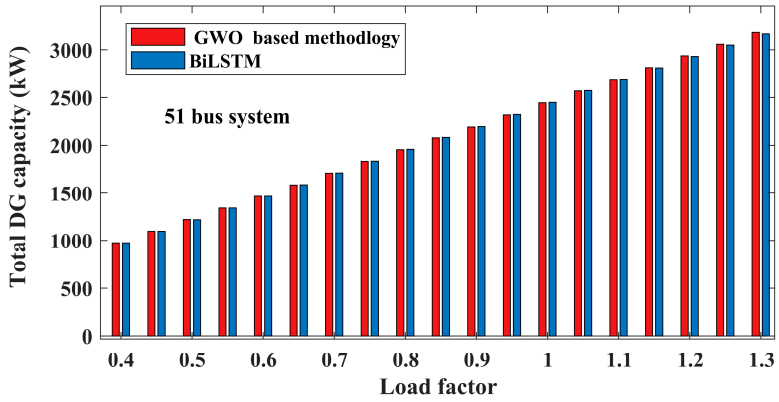


**Figure 10.** Comparison of total DG capacity for 51 bus system.

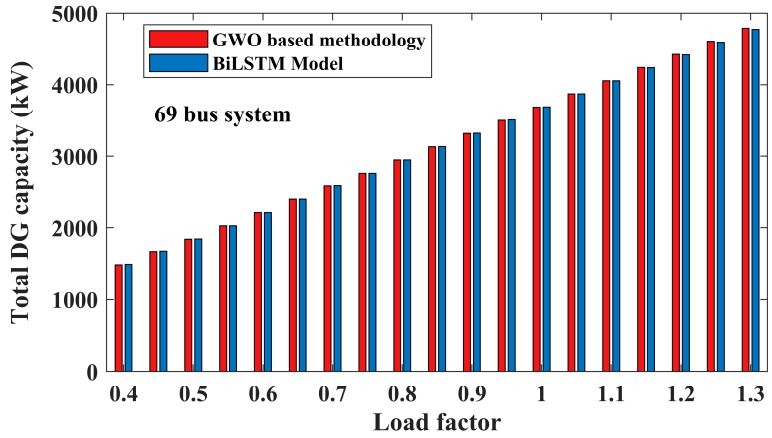


**Figure 11.** Comparison of total DG capacity for 69 bus system.

The DGs calculated from the BiLSTM model are almost equal to the optimum sizing values determined from the GWO-based methodology, as shown in Figures 10 and 11. Tables 8 and 9 display the input data required for validating the trained machine learning models for the 51 and 69 bus systems, respectively. The load factor data used for validation is distinct from that employed in training the models.

**Table 8.** Validation of Input data for 51 bus system.

| Load Factor | S/S Apparent Power (kVA) | Real Power Load (kW) | Reactive Power Load (kVAr) | Real Power Loss (kW) | Reactive Power Loss (kVAr) | Minimum Node Voltage (kV) |
|---|---|---|---|---|---|---|
| 0.475 | 1422.71 | 1169.93 | 745.28 | 27.20 | 23.50 | 10.54 |
| 0.675 | 2044.95 | 1662.53 | 1059.08 | 56.39 | 48.68 | 10.34 |
| 0.875 | 2682.63 | 2155.13 | 1372.88 | 97.42 | 84.03 | 10.12 |
| 1.125 | 3503.68 | 2770.88 | 1765.13 | 167.04 | 143.91 | 9.85 |

**Table 9.** Validation of Input data for 69 bus system.

| Load Factor | S/S Apparent Power (kVA) | Real Power Load (kW) | Reactive Power Load (kVAr) | Real Power Loss (kW) | Reactive Power Loss (kVAr) | Minimum Node Voltage (kV) |
|---|---|---|---|---|---|---|
| 0.475 | 2263.71 | 1806.04 | 1279.94 | 46.39 | 21.18 | 12.14 |
| 0.675 | 3250.17 | 2566.48 | 1818.86 | 96.79 | 44.10 | 11.91 |
| 0.875 | 4259.47 | 3326.92 | 2357.78 | 168.39 | 76.56 | 11.67 |
| 1.125 | 5557.45 | 4277.46 | 3031.43 | 291.61 | 132.23 | 11.35 |

The trained distribution network BiLSTM model for optimum sizing of DGs and SHCs is tested with data at load factors other than the data used for training, and the comparison results with the multi-objective GWO approach are shown in Tables 10 and 11, respectively.

**Table 10.** Validation of results for BiLSTM model of 51 bus system.

| | LF = 0.475 | | LF = 0.675 | | LF = 0.875 | | LF = 1.125 | |
|---|---|---|---|---|---|---|---|---|
| | GWO | BiLSTM | GWO | BiLSTM | GWO | BiLSTM | GWO | BiLSTM |
| DG1 (kW) | 467.97 | 467.65 | 665.01 | 665.34 | 862.05 | 863.74 | 1108.35 | 1108.34 |
| DG2 (kW) | 223.37 | 222.91 | 318.77 | 319.05 | 414.54 | 415.80 | 535.32 | 534.95 |
| DG3 (kW) | 467.97 | 467.21 | 665.01 | 664.93 | 862.05 | 863.74 | 1108.35 | 1108.75 |
| SHC1 (kVAr) | 132.01 | 132.28 | 188.94 | 188.81 | 244.47 | 245.34 | 314.58 | 314.49 |
| SHC2 (kVAr) | 106.30 | 106.91 | 150.04 | 151.07 | 194.82 | 195.20 | 249.90 | 249.32 |
| SHC3 (kVAr) | 126.92 | 127.29 | 180.19 | 180.31 | 234.05 | 234.10 | 301.39 | 301.17 |

From Tables 10 and 11, it can be observed that for both distribution systems the DGs and SHCs sizing values are almost close to the values obtained with GWO methodology.

The efficacy of the distribution system for real power loss and enhancement of minimum node voltage is compared between trained BiLSTM and GWO methodologies, and the results are shown in Tables 12 and 13 for the 51 and 69 bus systems, respectively. Tables 12 and 13 indicate that the performance of the BiLSTM model is almost identical to that of the GWO-based method. The optimal sizes of DGs and SHCs are rounded to the closest multiples of five for practical considerations, and it was observed that the performance achieved by the distribution systems closely approximates the actual values calculated from the machine learning models. Any type of available DGs, either renewable energy sources like wind generators and solar photovoltaic units or diesel generators, battery energy storage systems and fuel cells, can be used based on the availability at the substation

areas. The variations in renewable energy sources have to be compensated with energy storage systems or diesel generators to meet the required demand. The implementation of load curtailment is necessary if there is any deficiency in the power supply at the substation due to grid failures.

**Table 11.** Validation of results for BiLSTM model of 69 bus system.

| | **LF = 0.475** | | **LF = 0.675** | | **LF = 0.875** | | **LF = 1.125** | |
|---|---|---|---|---|---|---|---|---|
| | **GWO** | **BiLSTM** | **GWO** | **BiLSTM** | **GWO** | **BiLSTM** | **GWO** | **BiLSTM** |
| DG1 (kW) | 722.42 | 723.00 | 1026.59 | 1026.68 | 1332.37 | 1330.77 | 1710.99 | 1710.91 |
| DG2 (kW) | 722.42 | 723.20 | 1026.59 | 1026.51 | 1332.16 | 1330.77 | 1710.99 | 1711.03 |
| DG3 (kW) | 305.09 | 306.40 | 434.45 | 434.28 | 563.28 | 562.75 | 724.09 | 723.61 |
| SHC1 (kVAr) | 282.66 | 283.51 | 401.33 | 402.68 | 522.17 | 520.84 | 669.33 | 669.06 |
| SHC2 (kVAr) | 336.57 | 336.31 | 479.87 | 478.01 | 620.89 | 616.84 | 796.79 | 797.80 |
| SHC3 (kVAr) | 68.54 | 68.55 | 95.81 | 97.80 | 126.47 | 128.91 | 160.65 | 160.75 |

**Table 12.** Performance comparison of BiLSTM model for 51-bus system.

| **Load Factor** | **Real Power Loss (kW)** | | **Minimum Node Voltage (kV)** | |
|---|---|---|---|---|
| | **GWO** | **BiLSTM** | **GWO** | **BiLSTM** |
| 0.475 | 6.464 | 6.460 | 10.898 | 10.898 |
| 0.675 | 13.115 | 13.125 | 10.855 | 10.855 |
| 0.875 | 22.123 | 22.177 | 10.811 | 10.811 |
| 1.125 | 36.763 | 36.748 | 10.756 | 10.756 |

**Table 13.** Performance comparison of BiLSTM model for 69-bus system.

| **Load Factor** | **Real Power Loss (kW)** | | **Minimum Node Voltage (kV)** | |
|---|---|---|---|---|
| | **GWO** | **BiLSTM** | **GWO** | **BiLSTM** |
| 0.475 | 2.249 | 2.253 | 12.626 | 12.626 |
| 0.675 | 4.549 | 4.556 | 12.612 | 12.612 |
| 0.875 | 7.650 | 7.668 | 12.597 | 12.597 |
| 1.125 | 12.640 | 12.652 | 12.579 | 12.579 |

## 6. Conclusions

In this work, a multi-objective grey wolf optimization algorithm-based machine learning model has been developed to determine the optimum sizing of distributed generators and shunt capacitors for improving the performance of the distribution systems at the required load factor. A set of base case data consisting of substation apparent power, real power load, reactive power load, real power loss, reactive power loss, and minimum node voltage at various load factors are used as input data, and the corresponding optimum sizing values of DGs and SHCs obtained with the GWO-based approach have been used as the target data for training the LSTM and BiLSTM models. The per-unit values of the input and target data are considered for training the network model, and the adam optimization algorithm is used for updating the network parameters. The simulation results revealed that the trained BiLSTM model had better convergence properties and a slightly lower RMSE than the LSTM model, while the adam optimization algorithm outperformed the sgdm and rmsprop optimization techniques. The trained BiLSTM model was tested with

the validation data considered at load factors other than data used for training and found the results are satisfactory and almost close to the values obtained with the multi-objective GWO approach. The evolutionary computing techniques for optimum sizing of DGs and SHCs at various load factors necessitate running the algorithm multiple times to confirm the global optimum. However, a well-trained BiLSTM machine learning network model can compute the optimum values of DGs and SHCs directly at any desired load factor without using evolutionary computing techniques and load flow solutions, which will be more convenient for a utility engineer. Future work could involve managing the total DG capacity requirements at various seasons, using different types of available DGs at utilities to meet consumer load demands.

**Author Contributions:** Conceptualization, A.A., S.R.G., B.G., M.B.B. and K.J.; methodology, S.R.G., K.J., N.B.R. and D.D.; software, S.R.G. and K.J.; validation, A.A., B.G., N.B.R. and D.D., formal analysis, S.R.G., M.B.B., K.J. and N.B.R.; writing—original draft preparation, S.R.G., N.B.R. and D.D.; writing—review and editing, A.A., S.R.G., N.B.R. and D.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| DGN | Number of distributed generations |
| NN | Number of nodes of the distribution system |
| $DGI_{LF}$ | DG penetration index |
| $PL_{i,LF}$ | Real power load at ith node load factor LF |
| $Ploss_{LF}^{DGQC}$ | Real power loss with DGs and shunt capacitors |
| $Ploss_{LF}^{Base}$ | Real power loss at the base case at load factor LF |
| $PDG_{i,LF}$ | Real power delivered by the ith DG at load factor LF |
| $QL_{i,LF}$ | Reactive power load at ith node at load factor LF |
| $Qloss_{LF}^{DGQC}$ | Reactive power loss with DGs and shunt capacitors |
| $QDG_{i,LF}$ | Reactive power delivered by the ith DG |
| $QSC_{i,LF}$ | Reactive power delivered by the ith shunt capacitor |
| $QL_{i,LF}$ | Reactive power load of ith node at load factor LF |
| $Qloss_{LF}^{Base}$ | Reactive power loss at bases case at load factor LF |
| $QCI_{LF}$ | Reactive power penetration index |
| $SSAI_{LF}$ | Substation apparent power index at load factor LF |
| $SSAP_{LF}^{DGQC}$ | Substation Apparent power with the integration of DGs and SHCs at LF |
| $SSAP_{LF}^{Base}$ | Substation Apparent power with base case |
| $SSPI_{LF}$ | Substation real power index |
| $SSPL_{LF}^{DGQC}$ | Real power delivered by the S/S with integration of DGs and SHCs |
| $SSPL_{LF}^{Base}$ | Real power delivered by the S/S with base case |
| $SSQI_{LF}$ | Substation reactive power index |
| $SSQL_{LF}^{DGQC}$ | Reactive power delivered by the S/S with integration of DGs and SHCs |
| $SSQL_{LF}^{Base}$ | Reactive power delivered by the S/S at base case |
| $VDI_{LF}$ | Voltage profile index at load factor LF |
| $V_{i,LF}^{DGQC}$ | Voltage of ith node with DGs and SHCs at LF |
| $V_{i,LF}^{Base}$ | Voltage of ith node with base case at LF |

## References


1. Gampa, S.R.; Das, D. Optimum placement and sizing of DGs considering average hourly variations of load. *Int. J. Electr. Power Energy Syst.* **2015**, *66*, 25–40. [CrossRef]
2. Gampa, S.R.; Jasthi, K.; Goli, P.; Das, D.; Bansal, R.C. Grasshopper optimization algorithm based two stage fuzzy multiobjective approach for optimum sizing and placement of distributed generations, shunt capacitors and electric vehicle charging stations. *J. Energy Storage* **2020**, *27*, 101117. [CrossRef]
3. Gil-González, W.; Garces, A.; Montoya, O.D.; Hernández, J.C. A Mixed-Integer Convex Model for the Optimal Placement and Sizing of Distributed Generators in Power Distribution Networks. *Appl. Sci.* **2021**, *11*, 627. [CrossRef]
4. Fathy, A. A novel artificial hummingbird algorithm for integrating renewable based biomass distributed generators in radial distribution systems. *Appl. Energy* **2022**, *323*, 119605. [CrossRef]
5. Hari Prasad, C.; Subbaramaiah, K.; Sujatha, P. Optimal DG unit placement in distribution networks by multi-objective whale optimization algorithm & its techno-economic analysis. *Electr. Power Syst. Res.* **2023**, *214 Pt A*, 108869.
6. Gümüş, T.E.; Emiroglu, S.; Yalcin, M.A. Optimal DG allocation and sizing in distribution systems with Thevenin based impedance stability index. *Int. J. Electr. Power Energy Syst.* **2023**, *144*, 108555. [CrossRef]
7. Raj, A.F.; Saravanan, A.G. An optimization approach for optimal location & size of DSTATCOM and DG. *Appl. Energy* **2023**, *336*, 120797.
8. Shukla, V.; Mukherjee, V.; Singh, B. Integration of distributed generations and static var compensators with static synchronous compensators to reduce power losses. *Eng. Appl. Artif. Intell.* **2023**, *126*, 107208. [CrossRef]
9. Ahlawat, A.; Das, D. Optimal sizing and scheduling of battery energy storage system with solar and wind DG under seasonal load variations considering uncertainties. *J. Energy Storage* **2023**, *74*, 109377. [CrossRef]
10. Azeredo, L.F.; Yahyaoui, I.; Fiorotti, R.; Fardin, J.F.; Garcia-Pereira, H.; Rocha, H.R. Study of reducing losses, short-circuit currents and harmonics by allocation of distributed generation, capacitor banks and fault current limiters in distribution grids. *Appl. Energy* **2023**, *350*, 121760. [CrossRef]
11. Roy, A.; Verma, V.; Gampa, S.R.; Bansal, R.C. Planning of distribution system considering residential roof top photovoltaic systems, distributed generations and shunt capacitors using gravitational search algorithm. *Comput. Electr. Eng.* **2023**, *111 Pt B*, 108960. [CrossRef]
12. El-Rifaie, A.M.; Shaheen, A.M.; Tolba, M.A.; Smaili, I.H.; Moustafa, G.; Ginidi, A.R.; Elshahed, M. Modified gradient-based algorithm for distributed generation and capacitors integration in radial distribution networks. *IEEE Access* **2023**, 120899–120917. [CrossRef]
13. Ćetković, D.; Komen, V. Optimal distributed generation and capacitor bank allocation and sizing at two voltage levels. *IEEE Syst. J.* **2023**, *17*, 5831–5841. [CrossRef]
14. Smaili, I.H.; Almalawi, D.R.; Shaheen, A.M.; Mansour, H.S.E. Optimizing PV Sources and Shunt Capacitors for Energy Efficiency Improvement in Distribution Systems Using Subtraction-Average Algorithm. *Mathematics* **2024**, *12*, 625. [CrossRef]
15. Elseify, M.A.; Hashim, F.A.; Hussien, A.G.; Kamel, S. Single and multi-objectives based on an improved golden jackal optimization algorithm for simultaneous integration of multiple capacitors and multi-type DGs in distribution systems. *Appl. Energy* **2024**, *353 Pt A*, 122054. [CrossRef]
16. Nareshkumar, K.; Roy, N.B.; Das, D. A novel distributed Q-PQV bus pair approach for optimal planning of distributed generators and capacitor banks in a distribution network. *Comput. Electr. Eng.* **2024**, *120 Pt A*, 109718. [CrossRef]
17. Mohammedi, R.D.; Kouzou, A.; Mosbah, M.; Souli, A.; Rodriguez, J.; Abdelrahem, M. Allocation and Sizing of DSTATCOM with Renewable Energy Systems and Load Uncertainty Using Enhanced Gray Wolf Optimization. *Appl. Sci.* **2024**, *14*, 556. [CrossRef]
18. Nabavi, S.A.; Aslani, A.; Zaidan, M.A.; Zandi, M.; Mohammadi, S.; Hossein Motlagh, N. Machine learning modeling for energy consumption of residential and commercial sectors. *Energies* **2020**, *13*, 5171. [CrossRef]
19. Guo, F.; Xu, B.; Zhang, W.A.; Wen, C.; Zhang, D.; Yu, L. Training deep neural network for optimal power allocation in islanded microgrid systems: A distributed learning-based approach. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 2057–2069. [CrossRef]
20. Zafar, R.; Vu, B.H.; Husein, M.; Chung, I.Y. Day-Ahead solar irradiance forecasting using hybrid recurrent neural network with weather classification for power system scheduling. *Appl. Sci.* **2021**, *11*, 6738. [CrossRef]
21. Shams, M.H.; Niaz, H.; Na, J.; Anvari-Moghaddam, A.; Liu, J.J. Machine learning-based utilization of renewable power curtailments under uncertainty by planning of hydrogen systems and battery storages. *J. Energy Storage* **2021**, *41*, 103010. [CrossRef]
22. Bidgoli, M.A.; Ahmadian, A. Multi-stage optimal scheduling of multi-microgrids using deep-learning artificial neural network and cooperative game approach. *Energy* **2022**, *239*, 122036. [CrossRef]
23. Mayer, M.J.; Biró, B.; Szücs, B.; Aszódi, A. Probabilistic modeling of future electricity systems with high renewable energy penetration using machine learning. *Appl. Energy* **2023**, *336*, 120801. [CrossRef]
24. Ghafoori, M.; Abdallah, M.; Kim, S. Electricity peak shaving for commercial buildings using machine learning and vehicle to building (V2B) system. *Appl. Energy* **2023**, *340*, 121052. [CrossRef]
25. Al-Ja'Afreh, M.A.; Amjad, B.; Rowe, K.; Mokryani, G.; Marquez, J.L. Optimal planning and forecasting of active distribution networks using a multi-stage deep learning based technique. *Energy Rep.* **2023**, *10*, 686–705. [CrossRef]
26. Aguilar, D.; Quinones, J.J.; Pineda, L.R.; Ostanek, J.; Castillo, L. Optimal scheduling of renewable energy microgrids: A robust multi-objective approach with machine learning-based probabilistic forecasting. *Appl. Energy* **2024**, *369*, 123548. [CrossRef]

27. Krechowicz, A.; Krechowicz, M.; Poczeta, K. Machine Learning Approaches to Predict Electricity Production from Renewable Energy Sources. *Energies* **2022**, *15*, 9146. [CrossRef]
28. Matijašević, T.; Antić, T.; Capuder, T. A systematic review of machine learning applications in the operation of smart distribution systems. *Energy Rep.* **2022**, *8*, 12379–12407. [CrossRef]
29. Fu, X.; Guo, Q.; Sun, H. Statistical machine learning model for stochastic optimal planning of distribution networks considering a dynamic correlation and dimension reduction. *IEEE Trans. Smart Grid* **2020**, *11*, 2904–2917. [CrossRef]
30. Nguyen, B.N.; Ogliari, E.; Pafumi, E.; Alberti, D.; Leva, S.; Duong, M.Q. Forecasting Generating Power of Sun Tracking PV Plant using Long-Short Term Memory Neural Network Model: A case study in Ninh Thuan-Vietnam. In Proceedings of the 2024 Tenth International Conference on Communications and Electronics, Danang, Vietnam, 31 July–2 August 2024; pp. 333–338.
31. Zhang, Y.; Wang, X.; Wang, J.; Zhang, Y. Deep reinforcement learning based volt-var optimization in smart distribution systems. *IEEE Trans. Smart Grid* **2020**, *12*, 361–371. [CrossRef]
32. Oladeji, I.; Zamora, R.; Lie, T.T. Security constrained optimal placement of renewable energy sources distributed generation for modern grid operations. *Sustain. Energy Grids Netw.* **2022**, *32*, 100897. [CrossRef]
33. Zhang, X.; Hua, W.; Liu, Y.; Duan, J.; Tang, Z.; Liu, J. Reinforcement learning for active distribution network planning based on Monte Carlo tree search. *Int. J. Electr. Power Energy Syst.* **2022**, *138*, 107885. [CrossRef]
34. Vilaisarn, Y.; Rodrigues, Y.R.; Abdelaziz, M.M.; Cros, J. A deep learning based multiobjective optimization for the planning of resilience oriented microgrids in active distribution system. *IEEE Access* **2022**, *10*, 84330–84364. [CrossRef]
35. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
36. Cui, Z.; Ke, R.; Pu, Z.; Wang, Y. Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values. *Transp. Res. Part C Emerg. Technol.* **2020**, *118*, 102674. [CrossRef]
37. Vanga, J.; Ranimekhala, D.P.; Jonnala, S.; Jamalapuram, J.; Gutta, B.; Gampa, S.R.; Alluri, A. Fault classification of three phase induction motors using Bi-LSTM networks. *J. Electr. Syst. Inf. Technol.* **2023**, *10*, 28. [CrossRef]
38. Chakraborty, M.; Das, D. Voltage stability analysis of radial distribution networks. *Int. J. Electr. Power Energy Syst.* **2001**, *23*, 129–135. [CrossRef]