Gated recurrent unit based demand response for preventing voltage collapse in a distribution system

Venkateswarlu GUNDU[*], **Sishaj pulikottil SIMON**, **Kinattingal SUNDARESWARAN**,
Srinivasa rao nayak PANUGOTHU
Department of Electrical and Electronics Engineering, National Institute of Technology,
Tiruchirappalli, Tamil Nadu, India

Abstract: This paper presents the application of deep learning algorithms towards demand response management. Demand limit violation and voltage stability are the major problems associated with a secondary distribution system. These problems are solved using demand response models by day ahead scheduling loads at every 15 min interval through linear integer programming and based on short term forecasting of load (kW). A new architecture for short term load forecasting is presented namely gated recurrent unit in which statistical analysis is carried out to get the optimal architecture of the neural network model. Reliability indices such as loss of load probability (LOLP) is evaluated to handle uncertainties that may occur in forecasting due to overestimation and underestimation. Here a novel dynamic power flow is carried out to check limit violation of the voltage. Also, scheduling is performed for two types of loads namely deferrable with interruptible and deferrable with uninterruptable, when either of maximum demand or voltage limit violation occurs. Finally, the suggested model is validated on a modified 12 bus radial distribution system. The result analysis shows that the suggested gated recurrent unit minimizes the forecast error and demand response program schedules household appliances without a demand limit violation and ensures the prevention of voltage collapse.

Key words: Long short-term memory, load forecasting, gated recurrent unit, linear integer programming, voltage stability analysis, deep learning

1. Introduction

Recent developments in smart grid technologies have enabled efficient demand response (DR) programs that operates power system advantageously to both consumers and utilities. DR has effectively handled maximum demand violation, reduction in electricity pricing for consumers, outage management etc. [1–5]. However, voltage stability in distribution system operations is still considered to be a major challenge for utilities when system is stressed at peak load intervals as shown in Figure 1. Hence the success of demand response programs depends on the stability of the distribution system by maintaining the voltage within the specified limit and simultaneously flattening the peak load through efficient load scheduling. Intensive research has been carried out to manage demand response effectively. Some of these techniques are discussed here [6–9]. In [6], a demand response model is developed for flattens the load curve, clips the peak load and decreases the customer's payment cost. In [7], renewable energy sources are effectively used in DR programs. In [8], an incentive-based demand response optimization model is presented to schedule the household appliances for minimum usage during peak hours. In [9], a cloud computing-based demand response (DR) program is discussed, in which cloud computing is leveraged to gather the data generated in the internet of the energy network and performs analytics to manage

---

[*]Correspondence: psv2482109@gmail.com

DR. Based on the available literature, in demand response management, researchers have concentrated more on demand limit violation than voltage stability in the distribution system. Hence this paper presents a DR scheme which simultaneously mitigates demand limit violation and voltage collapse through advanced deep neural network models that can be used for load forecasting.
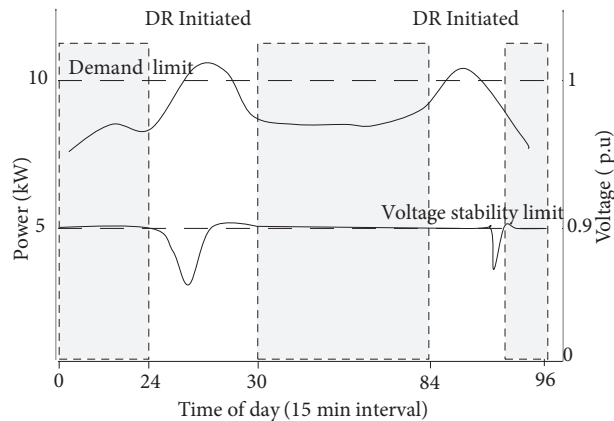

Figure 1: Typical load and voltage curves.

Forecasting and time sequence analysis have been intensively studied for almost four decades with statistical and intelligent models. Statistical models, such as artificial neural network models, linear regression, logistic regression, polynomial regression and autoregressive integrated moving average etc. [10–14], and intelligent models, such as fuzzy logic based neural networks, support vector machine etc., has already shown its dominance as a successful forecasting tools in several fields [15, 16]. However, accurate forecasting is still a challenge for lot of researchers. The recent developments in deep learning algorithms for its efficiency in processing time series data is found to be encouraging [17]. Therefore, these models can be attempted as forecasting tools in demand response programs. Recurrent neural networks are one of the deep learning models that has already established a reputation for coping with time sequence data through recurring neural connections. The effect of hidden layers in recurrent neural network causes the neural network output to either decrease or blow up exponentially which is found to be a limitation [18]. This limitation is addressed in LSTM (long short-term memory) and GRU (gated recurrent unit) based neural networks by altering the framework of the hidden neurons in traditional recurrent neural network [19]. LSTM and GRU models had already shown better results in various applications involving forecasting of time series data [20–22]. Hence this research work attempts to implement GRU and LSTM for forecasting load (kW) for the development of efficient demand response programs.

The suggested DR programs is validated on a modified 12 bus/node radial distribution system. The sensitive node is identified using forward and backward sweep load flow algorithm. Linear integer programming (LIP) is used to schedule the connected loads of the sensitive node, thereby voltage stability is achieved. It should be noted that scheduling of loads is carried out without the violation of maximum demand (MD) and ensuring prevention of voltage collapse.

2. LSTM/GRU architecture

Long term memory-based networks, usually referred to as "LSTMs" are a superior class of recurrent neural networks (RNN) adequate for long term dependency learning. A component called the memory block is the

key component that improves the ability of RNNs to model long-term dependencies [19]. As shown in Figure 2, the memory block (LSTM/GRU block) is a repetitively connected subnet with functional segments called the memory cell and gates. The memory cell is responsible for remembering the neural network's temporal status and multiplicative unit gates are accountable for monitoring the information flow pattern. Conferring to the relevant practical features these gates are categorized as input gate, forget gate and output gate. The input gate (i_t) regulates the amount of information that enters the memory cell, while the forget gate (f_t) directs the memory cell about the amount of information thus still remains in the present memory cell through recurring connection and the output gate(o_t) determines amount of data used to calculate the memory cell's output activation and further flows to the rest of the neural network. The following equations represents activations in an LSTM.

$$
\begin{aligned}
i_t &= \sigma(w_{ih}h_{t-1} + w_{ix}x_t + b_i) \\
o_t &= \sigma(w_{oh}h_{t-1} + w_{ox}x_t + b_o) \\
f_t &= \sigma(w_{fh}h_{t-1} + w_{fx}x_t + b_f) \\
g_t &= tanh(w_{gh}h_{t-1} + w_{gx}x_t + b_g) \\
c_t &= (f_t c_{t-1} + i_t g_t) \\
h_t &= o_t(tanh(c_t))
\end{aligned}
\tag{1}
$$

where c_t=Cell state, h_t=hidden state, w and b=weights and bias of each component

From Figure 2 LSTM cell highlights the functioning mechanism of LSTM block. The information processing between gates and the memory cell enables effective learning and retrieving activity in the recurrent structures. However, several improvements had taken place from its original architecture. GRU is one of the modified structures of LSTM. Gated recurrent unit (GRU) is a gating mechanism in recurrent neural networks, introduced in 2014 by Cho et al. [14] . The GRU is like a long short-term memory (LSTM) with forget gate [23] but has fewer parameters than LSTM, as it lacks an output gate shown in Figure 2. GRU's performance on certain tasks of polyphonic music modeling and speech signal modeling is similar to that of an LSTM [24]. Even though GRUs performance is similar to that of an LSTM, in some applications its performance enhanced with smaller datasets [25]. The following equations represents activations in a GRU.

$$
\begin{aligned}
z_t &= \sigma(w_{zh}h_{t-1} + w_{zx}x_t + b_z) \\
r_t &= \sigma(w_{rh}h_{t-1} + w_{rx}x_t + b_r) \\
H_t &= tanh(r_t h_{t-1} w_{Hh} + w_{Hh}x_t + b_H) \\
h_t &= (1 - z_t)h_{t-1} + z_t H_t
\end{aligned}
\tag{2}
$$

Like LSTM cell, GRU remembers the temporal status of the neural networks. Here the gates formed by multiplicative units are responsible for controlling the information flow pattern. According to the corresponding practical features these gates are classified as the reset gate(r_t) and update gate(z_t). r_t and z_t control the amount of past memory to be retrieved in the network.

2.1. Optimal architecture

The architecture of the LSTM/GRU network models is given in Figure 2. This network model comprises of input nodes and output nodes with hidden units in the LSTM/GRU layer. Each hidden node consists of their
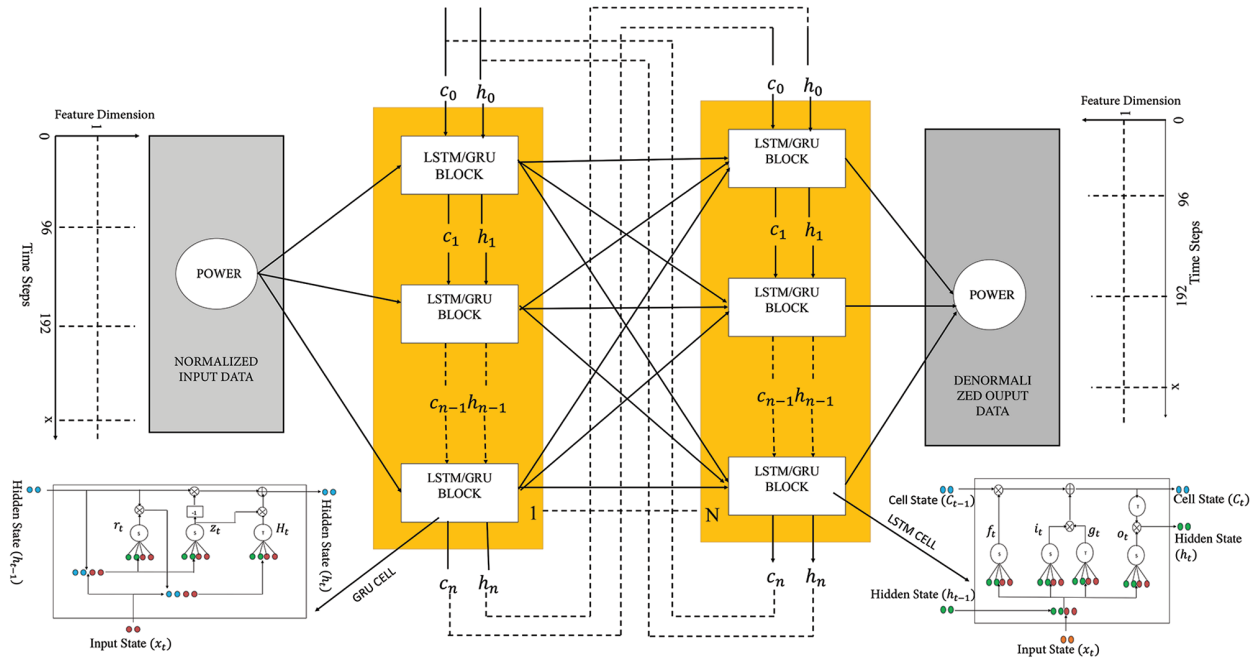
Figure 2: LSTM/GRU layer architecture.

(LSTM/GRU) respective gates as discussed in Section 2. The input data (I) is smoothened by normalization to avoid unnecessary dominance of certain variables using Equation (3).

$$N(I)_x \quad = \quad \frac{I_x - \bar{I}}{S} \tag{3}$$

where I_x = Input data set $(I_1, I_2 \dots I_x)$, x = Number of samples,
$\bar{I}$ = Mean of the samples, S = Standard deviation of the samples.
The past historical data of power is given as inputs to obtain forecasted day ahead power shown in Figure 2. The number of layers and its respective hidden nodes is selected based on statistical analysis and calculating mean absolute percentage error (MAPE) as given in Equation (4).

$$MAPE \quad = \quad \frac{1}{x} \sum_{I=1}^{x} |\frac{actual - forecast}{actual}| \tag{4}$$

Initially number of layers and hidden nodes are selected based on trial and error so that the networks (LSTM/GRU) converge. Once the layers and nodes are fixed, the number of layers (i) are changed from 1 to k and i is fixed for which MAPE is minimum. Then the number of nodes (j) is varied from 1 to m for i = 1 and j is fixed for which MAPE is minimum. The same procedure is carried out for k-1 layers. Hence the architecture $(L_{ibest}(X_{jbest}))$ is selected based on the generalized equation given in Equation (5).

$$L_{ibest}(X_{jbest}) \quad = \quad minMAPE(L_i(X_j)) \tag{5}$$

$\forall \; i = 1, 2, \dots k$ and $j = 1, 2, \dots m$

2.2. Sensitive node analysis

Sensitive node analysis is carried out to find the node which is vulnerable to voltage stability issues. Here 12 bus radial basis system is used to validate the suggested methodology [26] . Even though there are several voltage stability indices (such as line index, line stability factor, voltage collapse proximity indicator, voltage instability proximity index and integral steady state margin etc.) [27], it is not a function of voltage. It should be noted that, if any stability issues occur in the distribution system, then it first affects the sensitive node of the distribution system, thereby spreading to all other nodes. Therefore, the calculation of voltage stability index requires the knowledge of node voltage which comes from the load flow solution. Since the method we used to find the voltage stability index (VSI) is a function of voltage, it directly relates the node which is vulnerable to voltage stability issues. Hence, Equation (6) is used which is a function of voltage. The sensitive node of the test system is found through forward and backward sweep power flow using the load and line data given in Tables 1a and 1b [28]. Here, a DR program is implemented at the sensitive node of the 12 bus system.

$$VSI = v(m_2)^4 - [4P(m_2)X_{jj} - Q(m_2)R_{jj}] - [4P(m_2)R_{jj} - Q(m_2)X_{jj}]V(m_2)^2 \qquad (6)$$

Table 1. Twelve bus system data.

(a) Load data

Bus no.	P (kW)	Q (kVAr)	Bus no.	P (kW)	Q (kVAr)
1	0	0	7	55	55
2	60	60	8	45	45
3	40	30	9	40	40
4	55	55	10	35	30
5	30	30	11	40	30
6	20	15	12	15	15

(b) Line data

Branch no.	Sending node	Receiving node	R (Ohms)	X (Ohms)
1	1	2	1.093	0.455
2	2	3	1.184	0.494
3	3	4	2.095	0.873
4	4	5	3.188	1.329
5	5	6	1.093	0.455
6	6	7	1.002	0.417
7	7	8	4.403	1.215
8	8	9	5.642	1.597
9	9	10	2.89	0.818
10	10	11	1.514	0.428
11	11	12	1.238	0.351

2.3. Load forecast uncertainity

The uncertainty in the forecasted curve is incorporated by calculating reliability indices loss of load probability (LOLP) [29]. The forecasted uncertainty is assumed to be normally distributed. The probability distribution

of load can be described by four step model $(0, \pm 1\sigma, \pm 2\sigma, \pm 3\sigma)$. Where the standard deviation (σ) is usually 3 percentage of forecasted load [30]. The forecast uncertainty is calculated using the following equation:

$$LOLP_k = \sum_{l=1}^{4} LOLP_k(l) \times PL(l), k\epsilon[1, 96] \tag{7}$$

where PL(l) is the probability of load step l. $LOLP_k(l)$ is the LOLP for the load step l of hour k. The voltage profile at the sensitive node is calculated from the forecasted load with uncertainty using modified forward and backward sweep load flow given in Section 3.3, which will be used in scheduling of deferrable loads to check the violation of voltage limits.

3. Methodology

System design and the mathematical formulations are explained in Sections 3.1 and 3.2.

3.1. System design

The system consists of residential loads which is assumed to be lumped at the sensitive node of the 12 bus radial system is as shown in Figure 3. Here the constant load at 12th node is modified as a dynamic load which is obtained from the GEFCom (Global Energy Forecasting Competition) data for the residential buildings[1]. The residential loads comprise of houses with various electrical loads such as base loads. Base loads such as TV, lights and fans etc., are the loads which cannot shift their working time to any other time slot. Interruptible nondeferrable loads (INDL) such as air conditioner, geyser are the loads that cannot shift from one time period to other time period, but these loads can be interruptible. Deferrable loads are the loads that can be shift from one-time period to other time period to maintain load balance such as washing machine, electric vehicle, grinder and vacuum cleaner etc. These loads can be turned ON and OFF intermittently without degrading its performance. Hence in this work deferrable loads are scheduled as per the requirement of the DR program. Let the total scheduling period of time is finite, i.e. one day. The time period is split into T subintervals (e.g., 96 each of subintervals has 15 min duration).
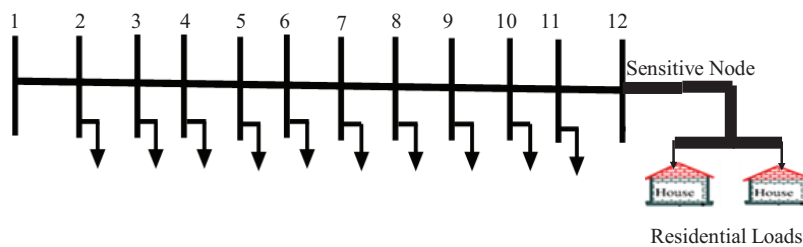

Figure 3: Twelve bus test system.

The total load of the system is considered as sum of the connected loads of the individual houses. The total energy usage by p_t^h , where p_t^h is a sum of base (L_{ht}^B), deferrable (L_{ht}^D), interruptible nondeferrable loads (L_{ht}^{INDL}).

---

[1]GEFCom (2012).Global Energy Forecasting Competition [online]. Website https://www.kaggle.com/data63177 [accessed 10 July 2019]

$$p_t^h \quad = \quad L_{ht}^B + L_{ht}^D + L_h t^{INDL} \tag{8}$$

where h=$(1,2,...H)$ and t=$(1,2,..T)$

3.2. Problem formulation

The problem is formulated to curtail the load within the demand limit by optimally scheduling deferrable loads and simultaneously maintaining the voltage within the limits. The objective function based on the problem formulation is given in Equation (9) as follows:

$$(f(A_{ta}^h, X_{ta}^h) = min \sum_{h=1}^{H} (L_{ht}^B + [L_{ht}^D + A_{ta}^h \times X_{ta}^h] + L_{ht}^{INDL}) \tag{9}$$

Demand limit constraints:

When the forecasted demand (kW) or voltage (pu) exceeds the specified limits, linear integer program (LIP) is initiated as given in Equation (10) which generates a feasible solution of X_{ta}^h for each of the deferrable appliances 'a' which may be in ON(A_{ta}^h=1) or OFF (A_{ta}^h=0) condition at time t. The total load at each sub interval should be less than or equal to the maximum demand, B_t [31] and is given as Equation (10).

$$\sum_{h=1}^{H} L_{ht}^B + L_{ht}^D + [A_{1a}^h \quad A_{2a}^h...A_{ta}^h] \begin{pmatrix} X_{11}^h & X_{12}^h & . & . & . & X_{1a}^h \\ X_{21}^h & X_{22}^h & . & . & . & X_{2a}^h \\ . & & . & . & . & . \\ . & & . & . & . & . \\ X_{t1}^h & X_{t2}^h & . & . & . & X_{ta}^h \end{pmatrix} + L_{ht}^{INDL} \leq B_t \tag{10}$$

Voltage limit constraints:

Once the deferrable loads are scheduled, the voltage profile at each sub interval should be within the limits as given in Equation (11), where $V_{min} = 0.9$ and $V_{max} = 1.1$.

$$V_{min} \leq V(\sum_{h=1}^{H} L_{ht}^B + L_{ht}^D + [A_{1a}^h \quad A_{2a}^h...A_{ta}^h] \begin{pmatrix} X_{11}^h & X_{12}^h & . & . & . & X_{1a}^h \\ X_{21}^h & X_{22}^h & . & . & . & X_{2a}^h \\ . & & . & . & . & . \\ . & & . & . & . & . \\ X_{t1}^h & X_{t2}^h & . & . & . & X_{ta}^h \end{pmatrix} + L_{ht}^{INDL}) \leq V_{max} \tag{11}$$

Deferrable load constraints:

Deferrable loads are of two types such as deferrable with interruptible (DI) and deferrable with uninterruptable (DUI). The first type belongs to the loads which can be operated intermittently without degrading their performance. Appliances such as electric vehicles and vacuum cleaners are few examples. The user has the provision to program the operating schedule of each appliances well in advance and is represented as a constraint given in Equation (12) as follows:

$$X_{ta}^h \quad = \quad \{x_{ta}^h | \begin{pmatrix} E_{aDINL} \leq X_{taDI}^h \leq E_{aDIFL} \\ \forall A_{taDI}^h = 1 \\ \forall t \in T \end{pmatrix} |\} \tag{12}$$

where

E_{aDINL} = No load capacity of DI appliances

E_{aDIFL} = Full load capacity of DI appliances

Here, to complete the job of an appliance in its operating schedule, the available power for that appliance should be within its maximum capacity (E_{aDIFL}) mentioned. The second type belongs to the loads with uninterruptable operation for a total time period (T_i) appliances such as Iron, clothes dryer, dishwasher, well pump, washing machine are few examples. The user has the provision to program the operating schedule of each appliances well in advance is represented as a constraint given in equation (13) as follows:

$$X_{ta}^h = \left\{ x_{ta}^h \middle| \begin{pmatrix} E_{aDUINL} \leq X_{taDUI}^h \leq E_{aDUIFL} \\ \forall A_{taDUI}^h = 1 \\ \forall t \in T_i \end{pmatrix} \middle| \right\} \tag{13}$$

where E_{aDUINL} =No load capacity of DUI appliances, E_{aDUIFL} =Full load capacity of DUI appliances

T_i = Total time period for which DUI appliance is in ON state. Here, to complete the job of an appliance in its operating schedule, the available power for that appliances should be within its maximum capacity (E_{aDUIFL}) mentioned.

3.3. Demand response program

The DR algorithm is initiated when maximum demand (B_t) or voltage limit violation occurs which is given in the flowchart (Figure 4). In the secondary distribution system, the area of the conductor is kept small when compared to the transmission system. Therefore, the resistance of secondary distribution system is high, thereby R/X ratio is dominant. Hence, in the suggested method voltage stability analysis is carried out with real power [32] and, modified forward and backward sweep load flow is used as given below.

Algorithm: Modified forward and backward sweep load flow (dynamic power flow).

Initialization

Step 1. Read bus data and calculate pu quantities of bus data

Step 2. Read line data and calculate pu quantities of line data

Step 3. Calculate the impedance matrix

Step 4. Calculate VSI

Step 5. If node is equal to sensitive node, then update load data with 15 min time block else go to step 1

Step 6. for t=1:96

iter=1

while ($\triangle V$)< epsilon

Compute load current and branch current of sensitive node

Calculate voltage drop (V_d)

Calculate sensitive node voltage

Compute the absolute change in voltage ($\triangle V$)

iter=iter+1

End

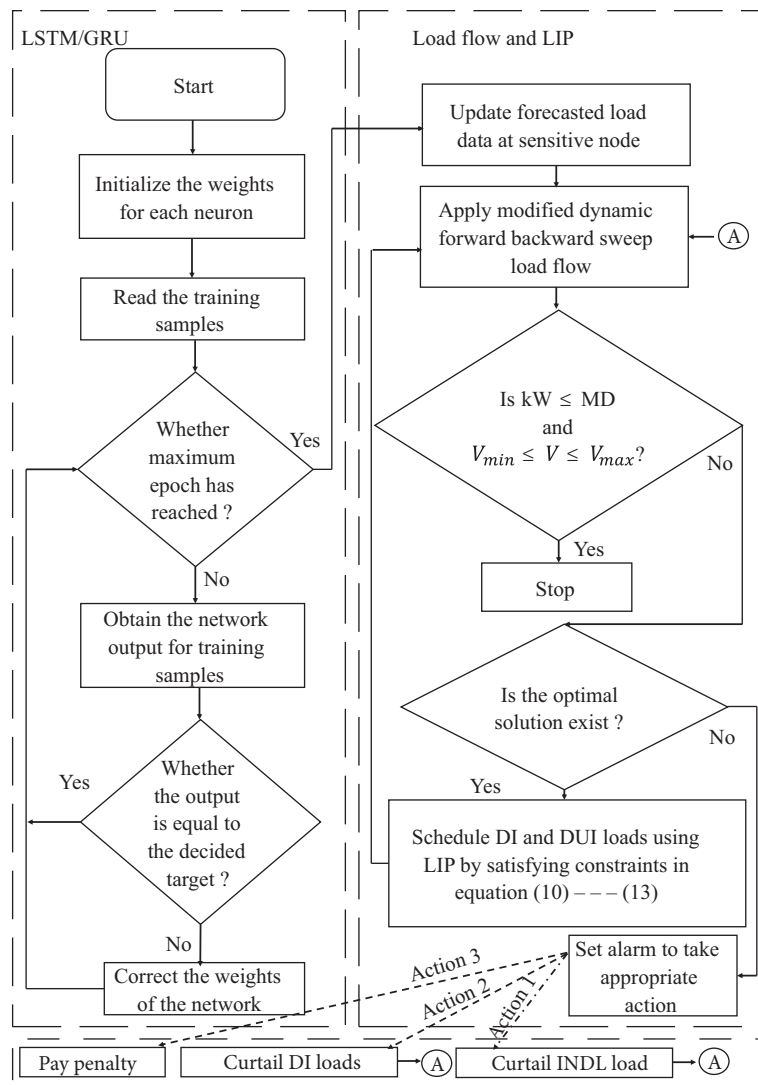Update voltage profile for 15 min interval

End

End

Figure 4: Flowchart of DR.

4. Results and analysis

The suggested methodology is implemented using GEFCom (Global Energy Forecasting Competition) data. Based on the analysis which is discussed in Section 2.2 the VSI is found to be minimum at the 12th node which can be inferred from Figure 5. A typical load curve (24 March 2007) at sensitive node with DI and DUI loads (refer Table 2) is shown in Figure 6a. The required data for training and validation is taken from 01 February 2007 to 15 March 2007 shown in Figure 6b. Here 70% of load samples are used for training and 30% of load samples are used for validation.

Both, day ahead and week ahead forecasting is carried out. In day ahead forecasting, nth day load is mapped to $(n+1)th$ day load profile [33]. However, in week ahead forecasting n^{th} day load is mapped with $(n+7)^{th}$ day load profile for training [34]. The optimal architecture of LSTM/GRU is necessary for effective demand response management. Hence, the optimal architecture for LSTM/GRU is found using statistical
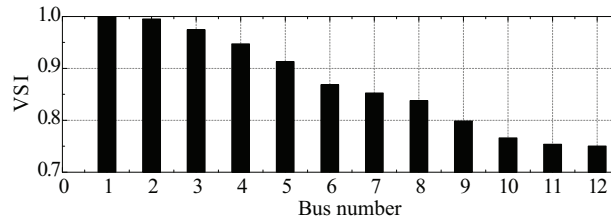
Figure 5: VSI of distribution system.

Table 2: Deferrable loads at sensitive node.

Set	Appliance	Quantity	Rated power (kW)
DI	Hybrid electric vehicle	1.0	2.5
DUI	Well pump	2.0	2.0
DUI	Water boiler	2.0	1.5
DUI	Dish washer	2.0	1.0
DUI	Washing machine	2.0	1.5
DI	Vacuum cleaner	2.0	0.8
DUI	Grinder	2.0	0.5

analysis as discussed in Section 2.1. The required data for testing LSTM/GRU architecture is taken from 23 March 2007 to 29 March 2007 (test week). The MAPE for day ahead and week ahead forecasting along with the structural parameters is given in Tables 3a and 3b. Statistical analysis for optimal layer selection of the LSTM/GRU layers is performed with three layers and a combination of 10,15 and 20 hidden units. The various combinations of architectures for both day ahead and week ahead is given in Table 3a. The optimal number of layers is the configuration which is able to give minimum average MAPE for the chosen test week.

From Table 3a, it has been observed that the architecture with two layers gives the minimum average MAPE for the given test week when compared to other combination. Hence the number of layers for LSTM/GRU is fixed as 2 for further analysis. Then the best combination of nodes is statistically carried out through trial and error for the 2 layers by calculating the MAPE which is tabulated in Table 3b. Finally, the optimal number of nodes is the one which gives the minimum average MAPE. Based on the above analysis, GRU day ahead with two layers (layer1 = 10 hidden nodes and layer2 = 20 hidden nodes) is selected as the best architecture for load forecasting. Once the forecasting results are obtained, uncertainty is incorporated as discussed in Section 2.3. The resulting training and validation plots for forecasted data with GRU is shown in Figure 6c. In this analysis, the performance of the GRU based DR program is verified by comparing with respect to the available literature results which is given in Tables 4a and 4b. Also, forecasting load curve with uncertainty is shown in Figure 7a. Here, the forecasted load curve with uncertainty (MAPE = 0.0489) is much closer with the actual load curve when compared with the forecasted load curve without uncertainty (MAPE = 0.0579). Also, it can be observed that the load violates its set values 10 kW at most of the intervals shown in Figure 7a. Once the forecasted load with uncertainty is determined the voltage profile is calculated using modified forward and backward sweep load flow at all subintervals as discussed in Section 3.3 (refer Figure 7b). It can be inferred that the voltage violates the set value of 0.9 pu at most of the intervals. Hence, scheduling of DI and DUI using linear integer programming is carried out for the objective function given in Equation (9). The LIP schedules the loads
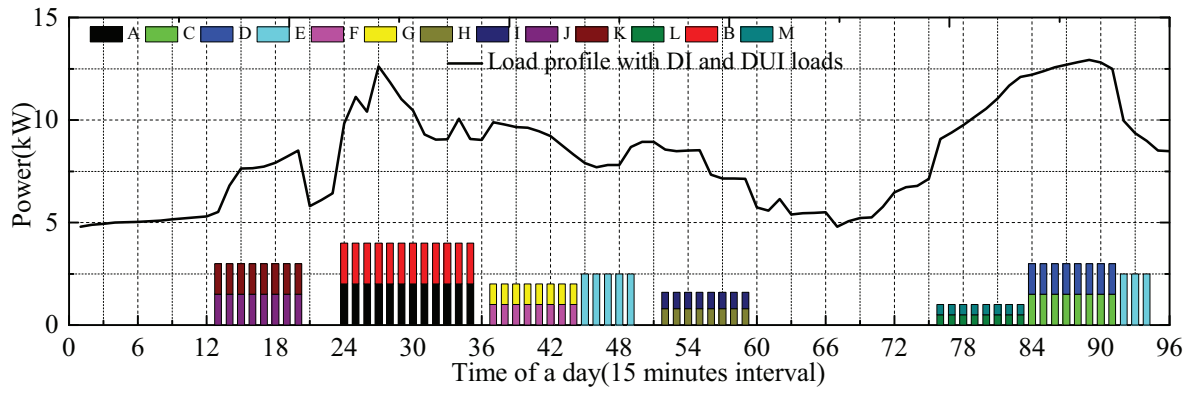
(a)
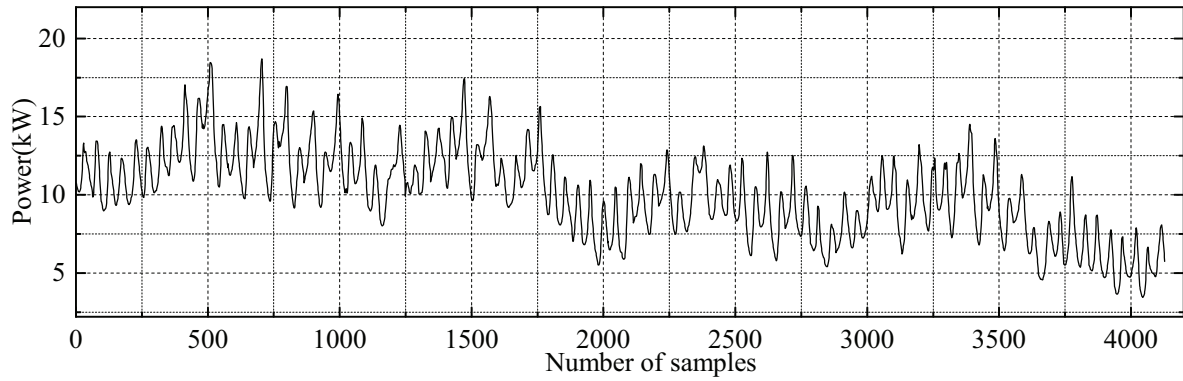
(b)
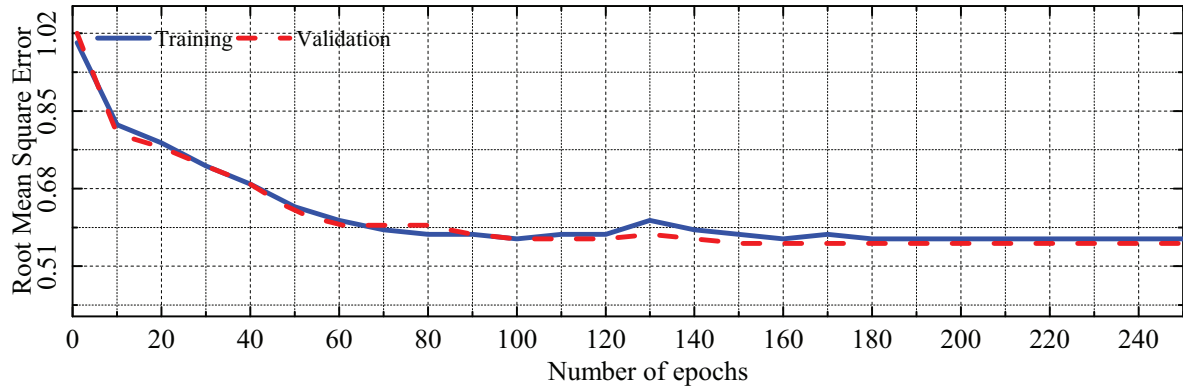
(c)

Figure 6. Deferrable loads and convergence plots (a) load curve with deferrable loads, (b) data samples (c) training and validation plots.

A and B = Well pump, C and D = Washing machine, E = Electric vehicle,F and G = Dish washer, H and I = Vacuum cleaner, J and K = Water boiler, L and M = Grinder.

ensuring the nonviolation of both voltage and maximum demand constraints. From Figures 7b and 7c, it can be inferred that the violation of voltage limits (between 24th and 30th intervals, between 81st and 93rd intervals)

Table 3. Optimal selection of architecture.

(a) Optimal selection of layers.

Type of forecast	Layers	Nodes	Average MAPE	Type of forecast	Layers	Nodes	Average MAPE
LSTM with day ahead	1	10	0.0679	GRU with day ahead	1	10	0.0674
	1	15	0.0614		1	15	0.0578
	1	20	0.0690		1	20	0.0684
	2	10	0.0538		2	10	0.0568
	2	15	0.0582		2	15	0.0562
	2	20	0.0555		2	20	0.0557
	3	10	0.0592		3	10	0.0895
	3	15	0.0730		3	15	0.0890
	3	20	0.0900		3	20	0.0710
LSTM with week ahead	1	10	0.0774	GRU with week ahead	1	10	0.0903
	1	15	0.0760		1	15	0.0717
	1	20	0.0765		1	20	0.0725
	2	10	0.0721		2	10	0.0711
	2	15	0.0729		2	15	0.0672
	2	20	0.0725		2	20	0.0658
	3	10	0.0757		3	10	0.0930
	3	15	0.1093		3	15	0.1196
	3	20	0.1127		3	20	0.1188

(b) Optimal selection of nodes.

Type of forecast	Layer1 nodes	Layer2 nodes	Average MAPE	Type of forecast	Layer1 nodes	Layer2 nodes	Average MAPE
LSTM with day ahead	10	10	0.0538	GRU with day ahead	10	10	0.0568
	10	15	0.0596		10	15	0.0555
	10	20	0.0617		10	20	0.0521
	15	10	0.0572		15	10	0.0664
	15	15	0.0582		15	15	0.0562
	15	20	0.0592		15	20	0.0624
	20	10	0.0640		20	10	0.0688
	20	15	0.0800		20	15	0.0713
	20	20	0.0555		20	20	0.0557
LSTM with week ahead	10	10	0.0721	GRU with week ahead	10	10	0.0711
	10	15	0.0760		10	15	0.0717
	10	20	0.1039		10	20	0.0647
	15	10	0.0749		15	10	0.0674
	15	15	0.0729		15	15	0.0672
	15	20	0.0796		15	20	0.0589
	20	10	0.0765		20	10	0.0670
	20	15	0.1315		20	15	0.0794
	20	20	0.0725		20	20	0.0658

and violation of maximum demand limits (between 24th and 30th intervals, between 78th and 93rd intervals) is mitigating after the execution of the DR algorithm given in Figure 4. If the DR algorithm does not have the options to schedule the loads satisfying all the constraints, then the algorithm either curtails the interruptible loads or asks the user to pay the penalty. This decision solely depends on mutual agreements between the utility and the consumer. Therefore, based on the simulation and performance analysis it is observed that GRU based DR program is suitable in preventing voltage collapse and maintaining demand limit in a distribution system.

**Table 4.** Performance comparison.

(a) MAPE comparison.

| Zone Id	Existing method(MAPE) [35]	GRU architecture (MAPE)
2	7.018	6.127
3	6.915	5.217
5	11.276	9.235
6	7.006	6.012
8	7.823	6.813
10	9.438	8.254

(b) Voltage profile (at sensitive node).

| Case study	DG at 9th bus	Without DR	With DR
Base case	0 kW	0.9304	0.9315
Murthy and Kumar [36]	235 kW at unity pf	0.9803	0.9824
Murthy and Kumar [36]	305 kVA at 0.9 pf lag	0.9904	0.9915
Yuvaraj, Ravi, Devabalaji [37]	235.5 kW at unity pf	0.9826	0.9836

5. Conclusion

A comprehensive optimized GRU architecture with demand response program has been presented to optimally schedule the loads of residential houses at the sensitive to maintain the load and voltage profile below a predefined desired threshold. Voltage stability index of the 12 bus test system is calculated based on which 12th node is identified as the sensitive node of the test system. Once the sensitive node of the test system is identified, statistical analysis is carried out for optimal selection of deep neural network model. Statistical analysis gives the gated recurrent unit as the suitable network architecture for effective load forecasting at the sensitive node of the distribution system. Also, the incorporation of uncertainty in the forecasted load curve reduced the MAPE by 15.54 percentage with respect to the actual load curve. Based on the forecasted load, linear integer program schedules the deferrable loads of the residential houses without demand violation and ensuring voltage stability. Voltage profile of the test system for every subinterval is calculated using modified forward and backward sweep load flow technique. Also, the DR program is flexible enough to take decisions either to curtail loads or to pay the penalty during violations of contractual limits. The validation of results using 12 bus radial distribution system and practical load data from GEFCom shows the robustness of the GRU based demand response program for future implementation.
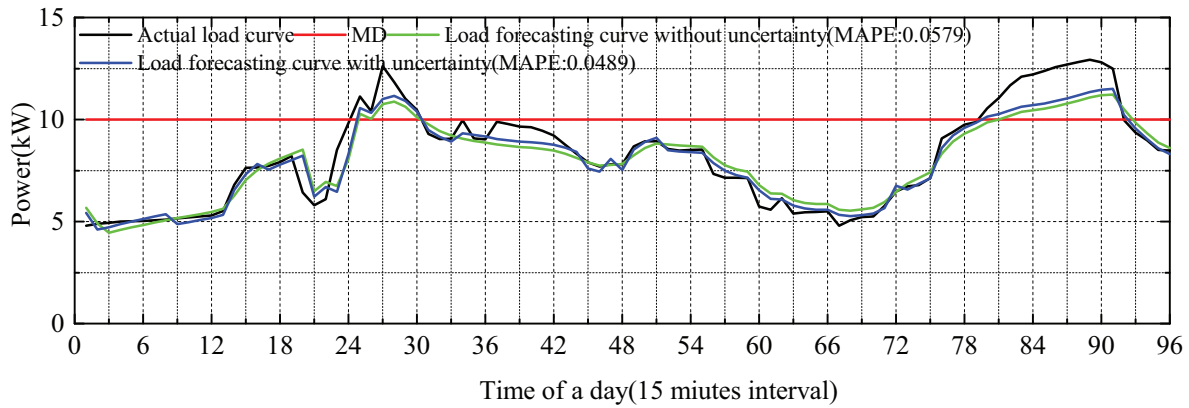
(a)
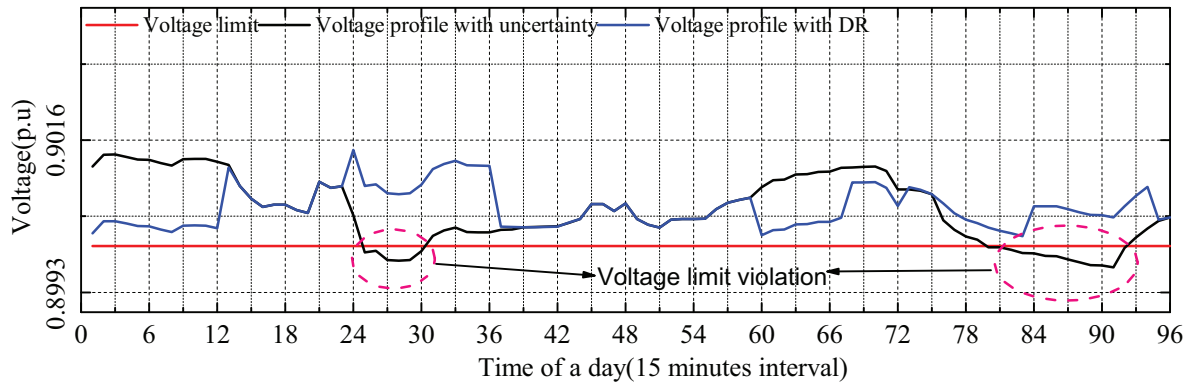
(b)
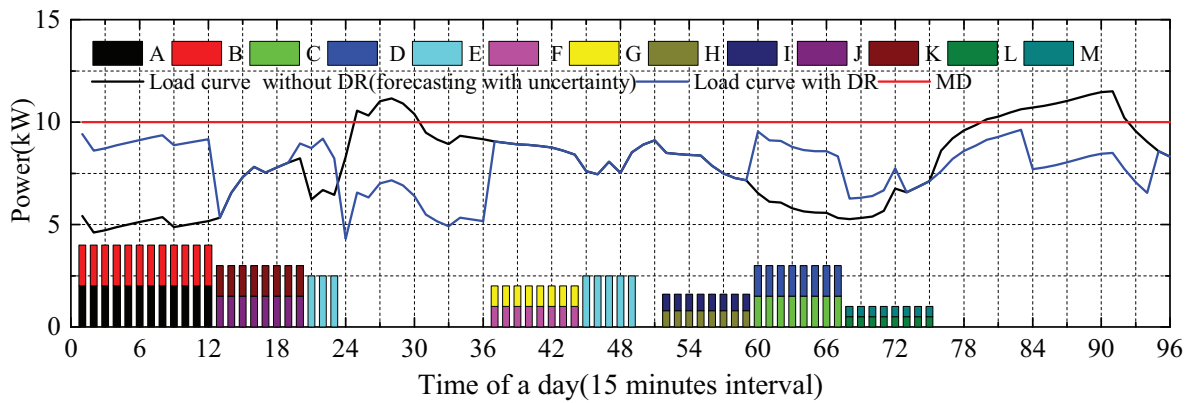
(c)

Figure 7. Load and voltage curves (a) load forecast curves, (b) voltage profile at sensitive node, (c) scheduling deferrable loads with DR.

References

[1] Mohsenian-Rad AH, Wong VW, Jatskevich J, Schober R, Leon-Garcia A. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. Institute of Electrical and

Electronics Engineers Transactions on Smart Grid 2010; 1 (3): 320-331.

[2] Albadi MH, El-Saadany EF. A summary of demand response in electricity markets. Electric Power Systems Research 2008; 78 (11): 1989-1996.

[3] Fraser H. The importance of an active demand side in the electricity industry. The Electricity Journal 2001; 14 (9): 52-73.

[4] Chen C, Wang J, Kishore S. A distributed direct load control approach for large-scale residential demand response. Institute of Electrical and Electronics Engineers Transactions on Power Systems 2014; 29 (5): 2219-2228.

[5] Bahrami S, Parniani M, Vafaeimehr A. A modified approach for residential load scheduling using smart meters. In: 2012 3rd Institute of Electrical and Electronics Engineers Power and Energy Society Innovative Smart Grid Technologies Europe; Berlin, Germany; 2012. pp.1-8.

[6] Nikoukar J. Unit commitment considering the emergency demand response programs and interruptible/curtailable loads. Turkish Journal of Electrical Engineering and Computer Sciences 2018; 26 (2): 1069-1080.

[7] Zehir MA, Batman A, Bagriyanik M. Review and comparison of demand response options for more effective use of renewable energy at consumer level. Renewable and Sustainable Energy Reviews 2016; 56: 631-642.

[8] Ruzbahani HM, Rahimnejad A, Karimipour H. Smart households demand response management with micro grid. In: 2019 Institute of Electrical and Electronics Engineers Power and Energy Society Innovative Smart Grid Technologies; Bucharest, Romania; 2019. pp 1-5.

[9] Jindal A, Kumar N, Singh M. Internet of energy-based demand response management scheme for smart homes and PHEVs using SVM. Future Generation Computer Systems 2020; 108: 1058-1068.

[10] Charytoniuk W, Chen MS. Very short-term load forecasting using artificial neural networks. Institute of Electrical and Electronics Engineers transactions on Power Systems 2000; 15 (1): 263-268.

[11] Taylor JW, Buizza R. Neural network load forecasting with weather ensemble predictions. Institute of Electrical and Electronics Engineers Transactions on Power systems 2002; 17 (3): 626-632.

[12] Eapen RR, Simon SP, Sundareswaran K, Nayak PS. User centric economic demand response management in a secondary distribution system in India. Institution of Engineering and Technology Renewable Power Generation 2019; 13 (6): 834-846.

[13] Peter SE, Raglend IJ, Simon SP. An architectural frame work of ANN based short term electricity price forecast engine for indian energy exchange using similar day approach. International Journal of Research in Engineering and Technology 2014; 2 (4): 111-122.

[14] Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing; Doha, Qatar; 2014. pp. 1724-1734.

[15] Sapankevych NI, Sankar R. Time series prediction using support vector machines: a survey. Institute of Electrical and Electronics Engineers Computational Intelligence Magazine 2009; 4 (2): 24-38.

[16] Hong WC. Application of seasonal SVR with chaotic immune algorithm in traffic flow forecasting. Neural Computing and Applications 2012; 21 (3): 583-593.

[17] Arı A, Hanbay D. Deep learning based brain tumor classification and detection system. Turkish Journal of Electrical Engineering and Computer Sciences 2018; 26 (5): 2275-2286.

[18] Hua Y, Zhao Z, Li R, Chen X, Liu Z et al. Deep learning with long short-term memory for time series prediction. Institute of Electrical and Electronics Engineers Communications Magazine 2019; 57 (6): 114-119.

[19] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation 1997; 9 (8): 1735-1780.

[20] Wang J, Tang J, Xu Z, Wang Y, Xue G et al. Spatiotemporal modeling and prediction in cellular networks: a big data enabled deep learning approach. In: Institute of Electrical and Electronics Engineers Conference on Computer Communications; Atlanta, GA, USA; 2017. pp. 1-9.

[21] Alahi A, Goel K, Ramanathan V, Robicquet A, Fei-Fei L et al. Social LSTM: human trajectory prediction in crowded spaces. In: 2016 Proceedings of the Institute of Electrical and Electronics Engineers Conference on Computer Vision and Pattern Recognition; Las Vegas, NV, USA; 2016. pp. 961-971.

[22] Ravanelli M, Brakel P, Omologo M, Bengio Y. Light gated recurrent units for speech recognition. Institute of Electrical and Electronics Engineers Transactions on Emerging Topics in Computational Intelligence 2018; 2 (2): 92-102.

[23] Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. Institution of Engineering and Technology 1999; 2: 850-855.

[24] De Campos LM. Time series prediction with direct and recurrent neural networks. Turkish Journal of Forecasting 2017; 1 (1): 7-15.

[25] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv 2014; arXiv:1412.3555 [cs.NE].

[26] Das D, Nagi HS, Kothari DP. Novel method for solving radial distribution networks. Institution of Electrical Engineers Proceedings-Generation. Transmission and Distribution 1994; 141 (4): 291-298.

[27] Danish MS, Senjyu T, Danish SM, Sabory NR, Mandal PA. A recap of voltage stability indices in the past three decades. Energies 2019; 12 (8): 1544.

[28] Chakravorty M, Das D. Voltage stability analysis of radial distribution networks. Electrical Power and Energy Systems 2001; 23 (2): 129-135.

[29] Chandrasekaran K, Simon SP, Padhy NP. SCUC problem for solar/thermal power system addressing smart grid issues using FF algorithm. International Journal of Electrical Power and Energy Systems 2014; 62: 450-460.

[30] Amjady N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. Institute of Electrical and Electronics Engineers Transactions on Power Systems 2001; 16 (3): 498-505.

[31] Arun SL, Selvan MP. Intelligent residential energy management system for dynamic demand response in smart buildings. Institute of Electrical and Electronics Engineers Systems Journal 2018; 12 (2): 1329-1340.

[32] Szultka A, Malkowski R, Czapp S, Szultka S. Impact of R/X ratio of distribution network on selection and control of energy storage units. In: 2017 International Conference on Information and Digital Technologies; Zilina, Slovakia; 2017. pp.359-364.

[33] Zhao T, Wang J, Zhang Y. Day-ahead hierarchical probabilistic load forecasting with linear quantile regression and empirical copulas. Institute of Electrical and Electronics Engineers Access 2019; 7: 80969-80979.

[34] Chen Y, Luh PB, Guan C, Zhao Y, Michel LD et al. Short-term load forecasting: similar day-based wavelet neural networks. Institute of Electrical and Electronics Engineers Transactions on Power Systems 2010; 25 (1): 322-330.

[35] Eapen RR, Simon SP. Performance analysis of combined similar day and day ahead short term electrical load forecasting using sequential hybrid neural networks. Institution of Electronics and Telecommunication Engineers Journal of Research 2019; 65 (2): 216-226.

[36] Murty VV, Kumar A. Optimal placement of DG in radial distribution systems based on new voltage stability index under load growth. International Journal of Electrical Power and Energy Systems 2015; 69: 246-256.

[37] Yuvaraj T, Ravi K, Devabalaji KR. Optimal allocation of DG and DSTATCOM in radial distribution system using cuckoo search optimization algorithm. Modelling and Simulation in Engineering 2017; 1: 1-20.