Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 7, S. 428-434, 4, 2023

Araştırma Makalesi

IJANSER

https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 7, pp. 428-434, 4, 2023

Research Article

Impact of the Size of Data on the Reliability of Short-Term Load Forecasting Using LSTM and GRU

Abdullah Ashraf* and Shaikh Saaqib Haroon[2]

[1]*Department of Electrical Engineering, University of Engineering and Technology Taxila, Pakistan*
[2] *Department of Electrical Engineering, University of Engineering and Technology Taxila, Pakistan*

**(abdullah.ashraf@students.uettaxila.edu.pk)*

(Received: 5 May 2023, Accepted: 27 May 2023)

(DOI: 10.59287/ijanser.773)

(1st International Conference on Contemporary Academic Research ICCAR 2023, May 17-19, 2023)

ATIF/REFERENCE: Ashraf, A.. & Haroon, S. S. (2023). Impact of the Size of Data on the Reliability of Short-Term Load Forecasting Using LSTM and GRU**.** *International Journal of Advanced Natural Sciences and Engineering Researches,* 7(4), 428- 434.

Abstract – Load forecasting has been an important aspect of power system operations and with the increase in integration of renewable energy resources in the main grid, the procedure is now more vital than ever. The methods developed to forecast the load of an area have also been improved with the use of artificial intelligence. This study proposes a forecasting training method using Gated Recurrent Units and compares it with the most widely used long-short term memory. The test systems are made of the historical load data from publicly available load data through PJM data miner 2 without the inclusion of weather parameters which reduces the training time of the models along with the reduction in data acquisition cost. The study also considers the impact of predicting future load without access to weather data.

Keywords – GRU, LSTM, Short-Term Load Forecasting, Machine Learning

I. Introduction

Load forecasting is an important aspect in many of the parts of the power system including planning, unit commitment, economic dispatch, storage management, power plant maintenance, power system supply design, and electricity market [1]. When the power system is mainly based on renewable energy resources (RERs), the forecasting of load on an hourly basis becomes much more significant due to the expensive and difficult solution of storage systems.

Forecasting has been classified into four types based on the duration of the forecast, this includes very short-term load forecasting which ranges from a few seconds to a few minutes, short term load forecasting ranging from hours to a few weeks, medium-term load forecasting which ranges from weeks to a few months and long term load forecasting which is the forecast in years or longer [2]. This study focuses on short-term load forecasting (STLF). Modern power system has smart metering in place which has made short-term forecasting of residential loads possible and implemented optimized load shaving along with the intelligent demand response [3]. The STLF is achieved using two different categories of methods namely traditional and intelligent, traditional methods involve statistical methods including regression and auto smoothing [4]. The intelligent methods include metaheuristics optimizations [5], artificial neural networks [6], and hybrid methods

[7][8]. These artificially intelligent methods enable forecasting with multiple input and output variables [9]. Load data is always a time series dataset whose specific statistical properties can be exploited as in [10] to transform the dataset into a form which exploited the advantages of the Convolutional Neural Networks (CNN).

Hybrid techniques comprise a combination of multiple machine learning techniques which integrate the merits of different models just as GRU-CNN [8] which showed improved results compared to both GRU and CNN separately. A similar case was observed in [7] where a hybrid of LSTM and CNN was proposed which showed improved better average prediction accuracy. Recently, deep learning, which is also part of the intelligent computational method, has become an active technology in load forecasting studies. It refers to stacking multiple layers of neural networks comprising nonlinear network layers which enable the nonlinear mapping and complication feature abstraction of the data and relies on stochastic optimization to achieve the computation.

Powerful machine learning techniques like recurrent neural networks (RNNs) have been proved to be much more effective in long term forecasting but LSTM and GRU which are types of RNNs have been employed increasingly in short-term forecasting scenarios in recent years because of the increase in computational capabilities of the available technologies [2][11].

A typical load forecast involves the following process:

- Obtain the historical load data and the weather data of a certain area.
- Get the weather prediction of the area and use a forecasting algorithm to forecast a load of that area depending on the weather [11].

Data preprocessing is also an important factor in getting accurate predictions from a model [12]. There are many techniques used to achieve the optimal data for the model to work on including longitudinal filling, regression prediction filling and normal interval filling all of which require the gap between the available data and missing data to be minimum. The data obtained in this study are precleaned and does not require any filling methods to be applied before passing it to the model for training.

For some areas, getting the historical weather can be an expensive step, and predicting the weather before the load forecast can create an extra computational cost. When achieving a short-term load forecast, sometimes only the daily weather of the area is available which is insufficient for the forecast to be accurate. Some studies achieve the forecast without access to the weather of the area [13]. Due to these problems, in this study, only the historical load of the areas is used which does not increase the cost of the forecasting while keeping the forecast accuracy reasonable.

II. RNN Approaches

In this section, the RNN methods used in this study are introduced. The models used here are LSTM and GRU, and are presented in sections III-A and section III-B respectively both of which are extensions of RNN architecture.

A. Long Short-Term Memory

This In applications, traditional RNNs were found to perform worst when the time intervals were long because of their inability to memorize previous information well because of the gradient vanishing problem [2]. To solve this problem, LSTM was introduced which combines the short-term memory of the traditional RNN with long-term memory through gate control. The internal structure of LSTM is presented in Fig. 1.

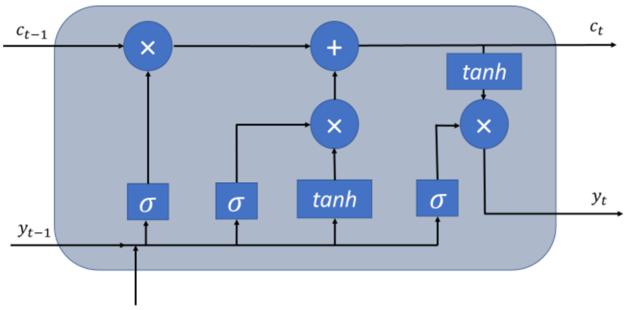


Fig. 1.LSTM internal structure

LSTM cells have a memory cell which is a special neuron structure with the capability to store information over an arbitrary time. There are three gate controls in the cell referred to as input, output, and forget gate.

Input gate i_t and forget gate f_t are expressed as follows:

$$i_t=\sigma\ (w_i\times[y_{t-1,}\ x_t]+b_i) \qquad (1)$$

$$f_t=\sigma\,(w_f\times[y_{t-1},\ x_t]+b_f) \tag{2}$$

where σ, the sigmoid activation function, is the variance whose values range between 0 and 1, w_i and w_f are the weight matrices, x_t is input while y_t is the output of the previous cell and b_i and b_f are the bias vectors. The cell state c_t is updated in the next step and is computed as follows:

$$c_t=f_t\times c_{t-1}+i_t\times(\tanh(w_c\times[y_{t-1},\,x_t]+b_c)) \tag{3}$$

where b_c is the bias vector of previous cell.

The output gate and final output y_t are expressed as:

$$o_t=\sigma\,(w_o\times[h_{t-1},\ x_t]+b_c) \tag{4}$$

$$y_t=\sigma_t\times\tanh(c_t) \tag{5}$$

B. Gated Recurrent Unit

The GRU is considered a variant of LSTM first introduced in [14]. The computation of GRU is simpler compared to LSTM which in turn creates less computational time while training [15]. The internal structure of GRU is shown in Fig. 2. There are two gate controls in GRU namely reset gate r_t and update gate u_t.

The mathematical equations which represent the gates in GRU are as follows:

$$r_t=\sigma\,(w_r.[y_{t-1}, x_t]) \tag{6}$$

$$u_t=\sigma\,(w_u.[y_{t-1}, x_t]) \tag{7}$$

The output gate y_t and ŷ which is the candidate state to determine the amount of information after the reset gate is presented as:

$$y_t=(1-u_t)\otimes y_{t-1}+y_t\otimes\hat{y}_t \tag{8}$$

$$\hat{y}_t=\tanh(w_y.[r_t\otimes y_{t-1}, x_t]) \tag{9}$$

where w_r, w_u and w_y are the weight parameters while σ and tanh are activation functions.

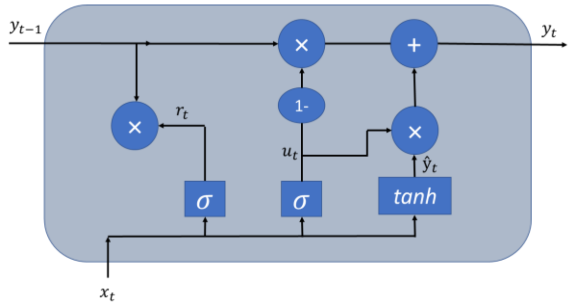


Fig. 2. GRU internal structure

III. Experimental Setup

A. Data description

In this study, the dataset is the historical load of 2016-2021 obtained from Pennsylvania-New Jersey-Maryland (PJM) [16] public API which includes the data of substations in the area. The dataset consists of the time of the reading, the load areas, and the power recorded in MW. Among the data for years 2016-2021 the selected substations are the ones which data available, only 6 substations are which are present in all the datasets to be tested which are given in Table 1. The data are divided into training, and testing data with 95% to 5% respectively.

Table 1. Load Areas in Test Cases

No.	Load Area	Abbreviation
1	Allegheny Power Systems	AP
2	Dayton Power and Light Co.	DAY
3	Duke Energy Ohio and Kentucky Corp.	DEOK
4	East Kentucky Power Cooperative	EKPC
5	Potomac Electric Power Co.	PEPCO
6	Rockland Electric Co.	RECO

B. Evaluation metrics

The results obtained by training and testing the algorithms are evaluated by their root-mean-square error (RMSE), and mean absolute percentage error (MAPE).

RMSE is defined as a measure of the distance between the prediction of a point with its true value. To remove the case where the prediction of two points is equal and opposite to their true values, a square is placed so that the sum may not add up to zero. RMSE is represented as:

$$\mathrm{RMSE}\left(y, y_{\mathrm{predicted}}\right) = \sqrt{\sum_{i=0}^{n} \frac{\left(y_{\mathrm{predicted}_i} - y_i\right.}{n}} \quad (10)$$

MAPE is the evaluation of regression loss in prediction problems and is represented as:

$$\mathrm{MAPE}\left(y, y_{\mathrm{predicted}}\right) = \frac{1}{n} \sum_{i=0}^{n-1} \frac{\left|y_i - y_{\mathrm{predicted}_i}\right|}{\max(\epsilon, \left|y_i\right|)} \quad (11)$$

where y is the test data to be validated by $y_{\mathrm{predicted}_i}$, n is number of observations (8760 for each load in a single year).

C. Training of LSTM model

LSTM is dependent on the data being clear in making the patterns because of its fundamental nature of learning from long-term and short-term data. This property can cause prediction losses when the data is incomplete or is not headed towards an incorrect path. For example, training the model on summer load data and testing it to predict load in winter. Because of that, the model for the yearly forecast is trained on 95% of the year's data.

Following are the parameters on which the model is trained:

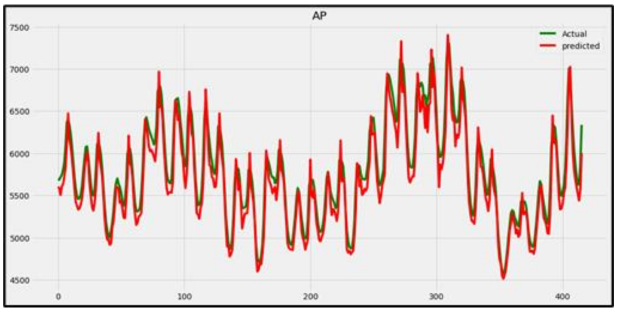
- Number of units = 256
- Optimizer = adam
- Number of dense layers = 2
- Number of epochs = 2

D. Training of GRU model

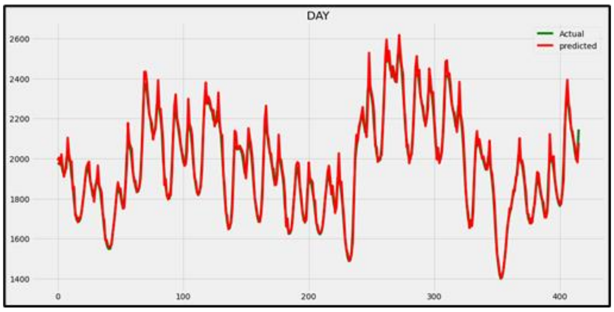
GRU is built to have the advantages of LSTM while removing the disadvantages of higher computational time and complexity while also providing the benefit of higher performance in the case of sequential data. The problem with GRU is its inability to perform when dealing with multiple dimensions but in this study, the data is of a single dimension.

Following are the parameters on which the model is trained:

- Number of units = 256
- Optimizer = adam
- Number of dense layers = 2
- Number of epochs = 2



(a)



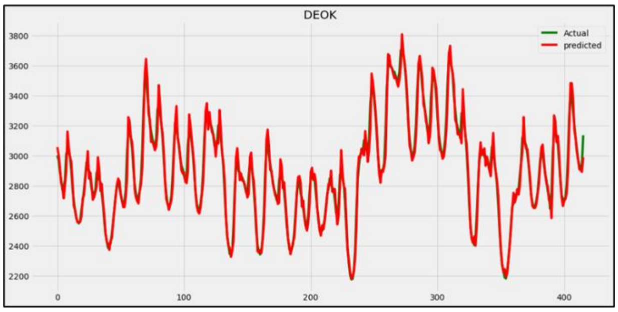
(b)

Fig. 3. Forecast using LSTM for 1-year data
(x-axis: hours, y-axis: MW)
(a) AP (b) DAY

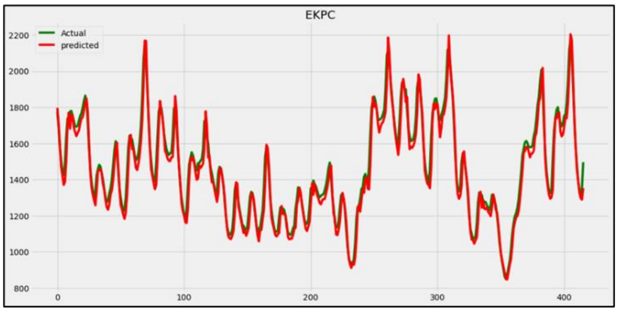
Table 2. Detailed results of LSTM and GRU with 1-year and 5-year data

Load Areas	**1-year data**				**5-year data**			
	LSTM		**GRU**		**LSTM**		**GRU**	
	RMSE	***MAPE***	***RMSE***	***MAPE***	***RMSE***	***MAPE***	***RMSE***	***MAPE***
AP	173.17	0.068	160.06	0.062	103.91	0.044	103.28	0.046
DAY	39.53	0.037	39.80	0.037	36.37	0.032	35.00	0.032
DEOK	56.69	0.044	52.27	0.044	65.14	0.034	64.86	0.038

Load Areas	1-year data				5-year data			
	LSTM		GRU		LSTM		GRU	
	RMSE	*MAPE*	*RMSE*	*MAPE*	*RMSE*	*MAPE*	*RMSE*	*MAPE*
EKPC	54.42	0.071	55.78	0.072	52.75	0.065	49.06	0.063
PEPCO	76.09	0.021	78.21	0.022	75.95	0.018	80.48	0.019
RECO	3.76	0.061	3.77	0.062	5.67	0.034	5.75	0.033

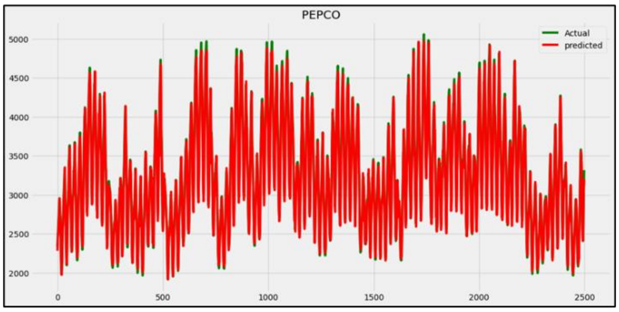


(a)

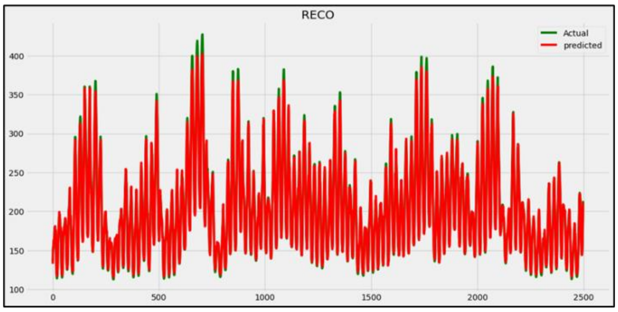


(b)

Fig. 4. Forecast using GRU for 1-year data
(x-axis: hours, y-axis: MW)
(a) DEOK (b) EKPC



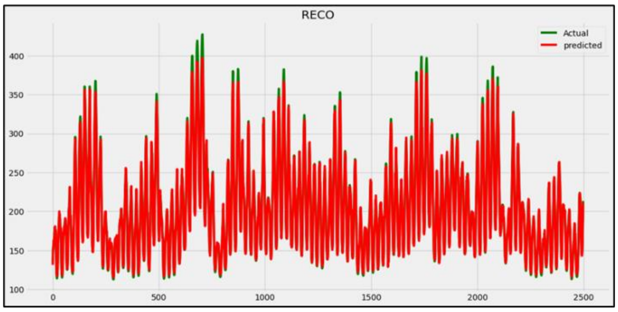
(a)



(b)

Fig. 5. Forecast using LSTM for 5-year data
(x-axis: hours, y-axis: MW)
(a) PEPCO (b) RECO

(a)



(b)

Fig. 6. Forecast using GRU for 5-year data
(x-axis: hours, y-axis: MW)
(a) PEPCO (b) RECO

IV. Results and Discussion

The systems are tested for two cases, one with one-year data while the second is trained for 5 years of data. From the results, we can see the prediction of both systems is very near to each other. The results from GRU surpass LSTM in 3 out of 6 load areas while in the case of 5-year data, GRU showed a significant improvement compared to LSTM. Both methods have presented similar values even with increasing the amount of data five folds. Fig. 3 and Fig. 4 show the forecast from LSTM and GRU respectively for different load areas for the 1-year test data. In

Fig. 5 and Fig. 6, the 5-year data from LSTM and GRU respectively is presented. The detailed results of all areas are shown in Table 2

V. CONCLUSION

This paper compares two state-of-the-art used intelligent forecasting techniques and presents the results on publicly available historical load data. The study showed a minimum difference in results even with the increase in the data 5 folds. The number of LSTM and GRU units being 256 caused the training time of the models to be higher. The study presented a method to reliably forecast a load of an area without the weather forecast being present which is to increase the units of the machine learning model.

References

[1] S. Muzaffar and A. Afshari, "Short-term load forecasts using LSTM networks," *Energy Procedia*, vol. 158, pp. 2922–2927, 2019, doi: 10.1016/j.egypro.2019.01.952.

[2] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network," *Energies*, vol. 11, no. 12, 2018, doi: 10.3390/en11123493.

[3] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019, doi: 10.1109/TSG.2017.2753802.

[4] W. Charytoniuk and M. S. Chen, "Nonparametric regression based short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 725–730, 1998, doi: 10.1109/59.708572.

[5] Z. Gao, J. Yu, A. Zhao, Q. Hu, and S. Yang, "A hybrid method of cooling load forecasting for large commercial building based on extreme learning machine," *Energy*, vol. 238, p. 122073, 2022, doi: 10.1016/j.energy.2021.122073.

[6] J. Lee and Y. Cho, "National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model?," *Energy*, vol. 239, p. 122366, 2022, doi: 10.1016/j.energy.2021.122366.

[7] B. Farsi, M. Amayri, N. Bouguila, and U. Eicker, "On short-term load forecasting using machine learning techniques and a novel parallel deep LSTM-CNN approach," *IEEE Access*, vol. 9, pp. 31191–31212, 2021, doi: 10.1109/ACCESS.2021.3060290.

[8] L. Wu, C. Kong, X. Hao, and W. Chen, "A Short-Term Load Forecasting Method Based on GRU-CNN Hybrid Neural Network Model," *Math. Probl. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/1428104.

[9] H. Zang *et al.*, "Residential load forecasting based on LSTM fusing self-attention mechanism with pooling," *Energy*, vol. 229, p. 120682, 2021, doi: 10.1016/j.energy.2021.120682.

[10] N. Andriopoulos *et al.*, "Short term electric load forecasting based on data transformation and statistical machine learning," *Appl. Sci.*, vol. 11, no. 1, pp. 1–22, 2021, doi: 10.3390/app11010158.

[11] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep Learning for solar power forecasting - An approach using AutoEncoder and LSTM Neural Networks," *2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc.*, pp. 2858–2865, 2017, doi: 10.1109/SMC.2016.7844673.

[12] W. Guo, L. Che, M. Shahidehpour, and X. Wan, "Machine-Learning based methods in short-term load forecasting," *Electr. J.*, vol. 34, no. 1, p. 106884, 2021, doi: 10.1016/j.tej.2020.106884.

[13] M. Ghomi, M. Goodarzi, and M. Goodarzi, "Peak Load Forecasting of electric Utilities for West province of IRAN by using Neural Network without Weather information," *12th Int. Conf. Comput. Model. Simul. Peak*, pp. 592–597, 2010, doi: 10.1109/UKSIM.2010.113.

[14] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," pp. 1–9, 2014, [Online]. Available: http://arxiv.org/abs/1412.3555.

[15] E. Yiğit, U. Özkaya, Ş. Öztürk, D. Singh & H. Gritli "Automatic detection of power quality disturbance using convolutional neural network structure with gated recurrent unit. Mobile" Information Systems, 2021, 1-11.

[16] PJM, "PJM Data Miner 2," 2022. https://dataminer2.pjm.com/list.