Medium and Long Term Power Load Forecasting Based on Stacked-GRU

Zheng Yang[1], Jing Cui[1], Qiangjian Zhang[2], Chunlin Yin[1], Li Yang[1], Pengfeng Qiu[1], Kai Hu[3] and Junwen Yang[1,*]

[1] *Electric Power Research Institute of Yunnan Power Grid, Kunming 650217, China*
[2] *Key Laboratory in Software Engineering of Yunnan Province, School of Software, Yunnan University, Kunming 650091, China*
[3] *Yunnan Power Grid Co., Ltd, Kunming 650011, China*
E-mail: junwen05@163.com
[*] *Corresponding Author*

Received 09 May 2022; Accepted 06 June 2022;
Publication 24 September 2022

Abstract

Power load forecasting plays a critical role in energy economy development and distribution of power systems. Predicting medium and long term power loads have facilitated the development of power grids. In this paper, a stacked-gated recurrent unit (Stacked-GRU) is applied to establish a power load forecasting model by integrating economic factors. Meanwhile, it also conducts medium and long term power load (MLTPL) forecasting based on the power load data of Yunnan Province from 2009 to 2020. By comparing different optimizers, it is found that the Adam optimizer works the best on the Stacked-GRU architecture. In the experiment of medium and long term power load forecasting for Yunnan Province, the values of MAPE, RMSE, and MAE of the model are 9.76%, 1.412, and 1.14, respectively, all of which outperform other deep learning comparison algorithms.

Keywords: Medium and long term power load forecasting, time series forecasting, stacked gated RNN.

Strategic Planning for Energy and the Environment, Vol. 41_4, 363–378.
doi: 10.13052/spee1048-5236.4141

1 Introduction

Yunnan Province is rich in power resources. By the end of 2020, the province's installed capacity of power generation was about 103.4028 million kilowatts. Among them, the clean electricity accounts for more than 80%, and the installed capacity of thermal power accounts for about 15% [1].

Electric power load data has the characteristics of periodicity, which can be divided into daily periodicity, weekly periodicity and annual periodicity. Among them, the annual periodicity of load data is closely related to seasonal factors [2]. According to the time-division of power load forecasting, it can be divided into super short-term, short-term, medium-term, and long-term. The medium and long term power load (MLTPL) forecasting are instructive in guiding power grid planning department and helping analyze the status quo of power grid and predict saturated load [3]. Accurate MLTPL forecasting can not only support economic activities actitivities by providing a guide for the safe operation of power grid, but also improve the quality of power grid planning [4].

In recent years, scholars have proposed various solutions to obtain high-precision results of MLTPL forecasting. Some are listed as below:

(1) Traditional forecasting methods, such as linear regression [5, 6], least squares method [7], exponential smoothing method [8, 9], etc. These methods are based on statistical correlation principles to build predictive models. All the methods are quick and direct in forecast. However, these methods are less effective on nonlinear data [10].
(2) Machine learning methods, such as long short-term memory unit (LSTM) [11], random forest algorithm [12], and temporal convolutional network [13] are more powerful and have been used to predict power load.
(3) Methods of model combination, such as power load forecasting model integrated LSTM and eXtreme Gradient Boosting [14], chaotic sparrow search algorithm (CSSA), optimization and firefly algorithm (FA), improved extreme learning machine (ELM) [15], convolutional neural network (CNN) combined with LSTM [16], GRU-based CNN and CNN Hybrid Neural Network Model [17].

Power load is comprehensively affected by various factors (economy, weather, population, traffic, etc.). Therefore, in the process of designing the power load forecasting model, we should fully consider these influencing factors [18]. Economic and social factors are studied by many researchers

[19–21]. Liu D et al. [21] proposed a random forest algorithm which integrates integrats economic factors, and uses the electricity load data of China. Compared to using electricity data alone, the model was 15% more accurate. Ghanbari A et al. [22] proposed a model integrating two economic factors, GDP and population, and is applied to the annual electricity load data of Iran.

Inspired by these methods, we put forward a power load forecasting model that integrates economic factors. In this paper, a stacked-gated recurrent neural network (RNN) is applied to model the MLTPL data in Yunnan Province. The two indicators of the year-end total population of Yunnan Province and the GDP of Yunnan Province are regarded as external factors in the power load forecasting model.

2 The Theory of Deep Learning

2.1 Gated RNN

Gated Recurrent Units (GRU) is an optimized model derived from classical neural recurrent networks [23–25]. GRU can alleviate some of the problems in traditional neural networks, such as exploding or vanishing gradients. GRU's architecture contributes to its efficient computing performance [26]. The gated recurrent neural unit in GRU contains a reset gate r and an update gate z and LSTM-like output gates are removed. The update gate decides how much past information will be fed into future, and the larger the value, the greater grater the previous state information is pereserved. The reset gate mainly determines how much past information needs to be forgotten. The smaller the value, the more the previous state is forgotten. In addition, gated recurrent neural units merge the LSTM's memory cell states and hidden layer node outputs. The structure diagram of the GRU is shown below.

The complete update process of the gated recurrent neural unit is shown in formulas (1)–(4):

(1) Reset gate

$$R_t = \sigma(W_r \cdot [H_{t-1}, x_t] + b_r) \tag{1}$$

(2) Update gate

$$Z_t = \sigma(W_z \cdot [H_{t-1}, x_t] + b_z) \tag{2}$$

(3) Candidate hidden layer states

$$\tilde{H}_t = \tanh(W_{\tilde{h}} \cdot [R_t \odot H_{t-1}, x_t] + b_h) \tag{3}$$

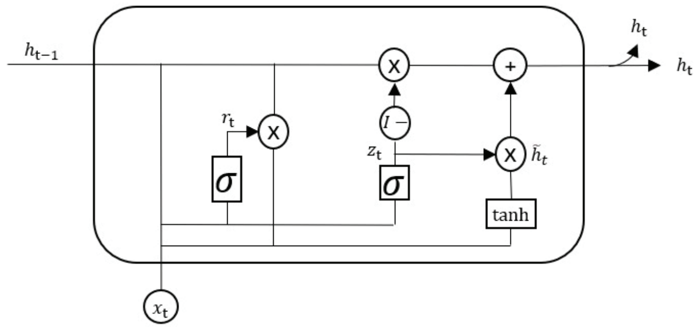

Figure 1 The structure diagram of GRU.

(4) The state of the hidden layer at time t

$$H_t = (1 - Z_t) \odot \tilde{H}_t + Z_t \odot H_{t-1} \tag{4}$$

Among them, x_t stands for the value of the model input at time t. H_{t-1} represents the state of the model at the last moment. H_t represents the state of the model at the current moment. R_t and Z_t represent the state of the reset gate and update gate at the current moment. A higher Z_t value, which ranges from 0 to 1, means a higher percentage of state information will be remembered and vicevise versa. $\tilde{H}_t$ represents the state of the candidate layer at the current moment. $W_r, W_z, W_{\tilde{h}}, b_r, b_z$ and b_h are weights and biases.

2.2 Stacked Gated RNNs

The overall architecture of long-term grid load forecasting model is shown in the figure below. After the power load data is normalized, the power load data, the year-end total population data of Yunnan Province, and the GDP of Yunnan Province in the same period are input into Stacked-GRU model, where network parameters *W* and *b* are optimized.

GRU is a shallow model with relatively weak feature extraction power. We stack multiple layers of GRUs to form a deeper network (as shown in Figure 3) to enhance model's feature extraction capability.

3 Power Load Forecasting Based on Stacked-GRU

3.1 Data Description

The GDP and population data of Yunnan Province involved in this article are obtained from "Yunnan province statistical yearbook" [27] and monthly

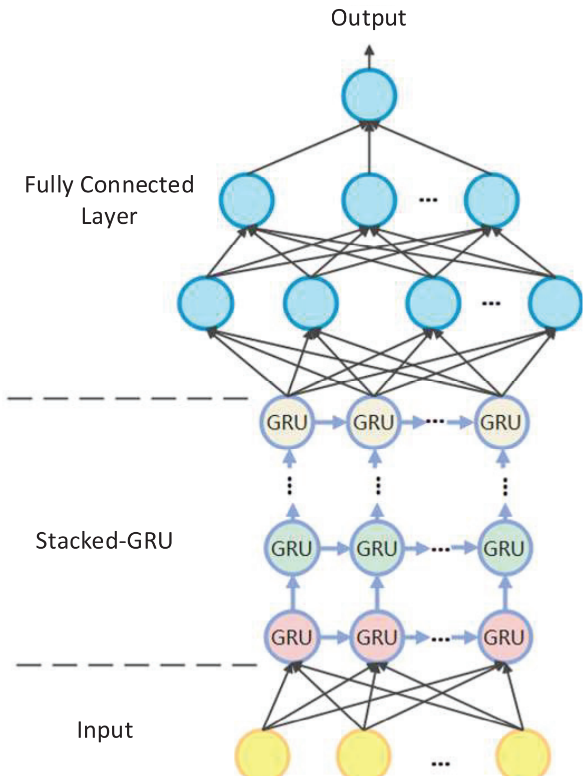

Figure 2 Stacked-GRU model.

statistical report data [28] on the official website of Yunnan Provincial Bureau of Statistics. The power load data of Yunnan Province ranges from January of 2009 to December of 2020 and consists of two parts: (1) monthly social electricity consumption of Yunnan Province, and (2) monthly electricity consumption data of cities and autonomous prefectures in Yunnan Province.

80% of the data set is used as the training set, and the remaining 20% is used as the test set. Data processing flow is presented in Figure 4.

3.2 Data Preprocessing

A linear normalization method is adopted, as shown in Equation (5).

$$X_{nom} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{5}$$

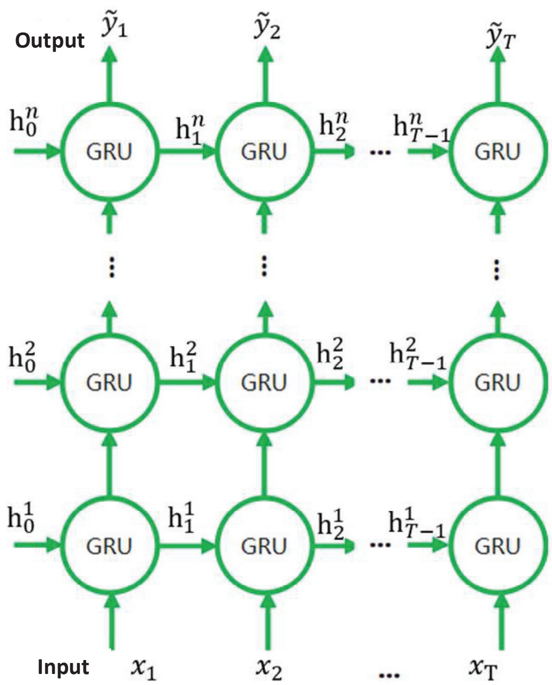

Figure 3 Detailed structure of Stacked-GRU model.

Where X_{nom} is the normalized power load data or economic data. X stands for the original data. X_{min} represents the minimum value of the data and X_{max} refers to the maximum value.

3.3 Evaluation Method for Forecast

This paper uses three indicators to evaluate forecast algorithmalrithm: Mean Absolute Error (MAE) [29], Mean Absolute Percentage Error (MAPE) [30], and Root Mean Squared Error (RMSE) [31]. The calculations are presented as follows.

MAE is a linear score that infer the mean of absolute error between output value and observed value, where all individual differences are equally weighted.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{6}$$

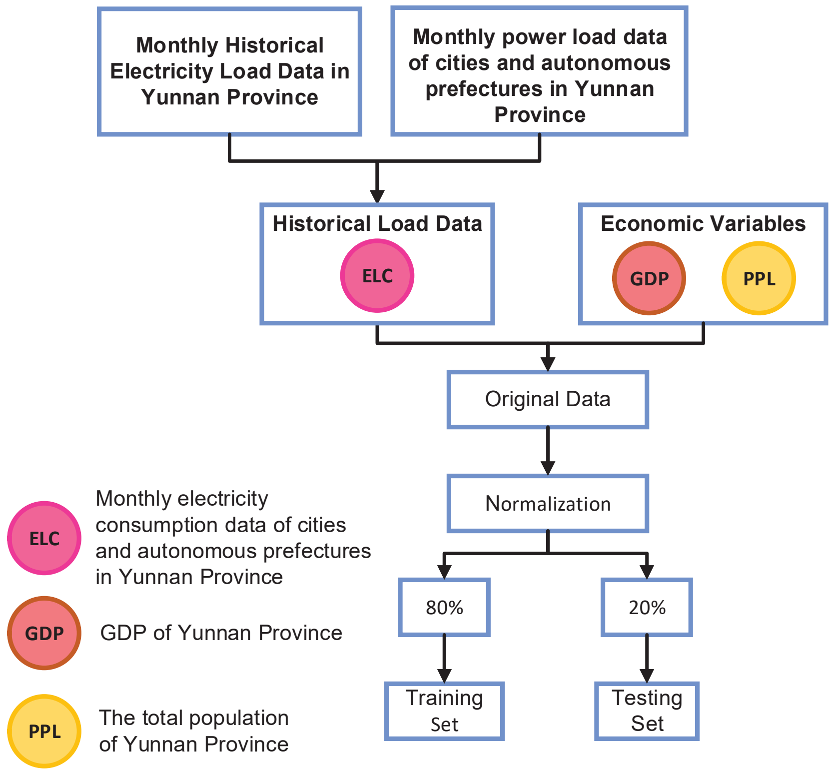

Figure 4 Data processing flowchart.

MAPE refers to the ratio of the difference between observed value and model output value to the absolute values of all observations.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{7}$$

RMSE is the square root of the ratio of the square of the deviation between the model output value and the observed value to the number of observations, which is sensitive to outliers in a data set.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{8}$$

Table 1 Comparison of Adam and other optimizers

Optimizer Name	The Best MAPE (%)
Adam	1.032
Adagrad	1.58
RMSPorp	3.207
SGD	26.78

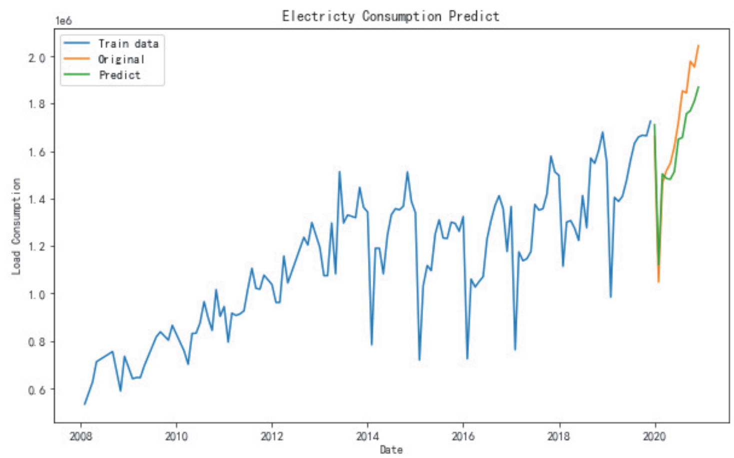

Figure 5 Consumption curve predicted using Stacked-GRU.

3.4 Model Optimizer Selection

There are many optimizers for deep learning. It takes trial and error to determine which optimizer to choose. In this study, we choose four common optimizers, namely Adam, Adagrad, RMSPorp, and SGD, to run on the Stacked-GRU model to conduct comparative experiments on the MLTPL data in Yunnan Province.

We use MAPE as metric for the model optimizer, given that MAPE is a relative error measure and is suitable for the comparison task on the accuracy of different time series based forecasting models.

Experiments show that the model with Adam optimizer performs the best among all four optimizers and the results can be seen in Table 1.

3.5 Performance on Test Set

We run the trained Stacked-GRU model on test set to verify its effectiveness. The test set used in this paper is the power load data of Yunnan Province from 2009 to 2020. As shown in Figure 5, the power consumption data after

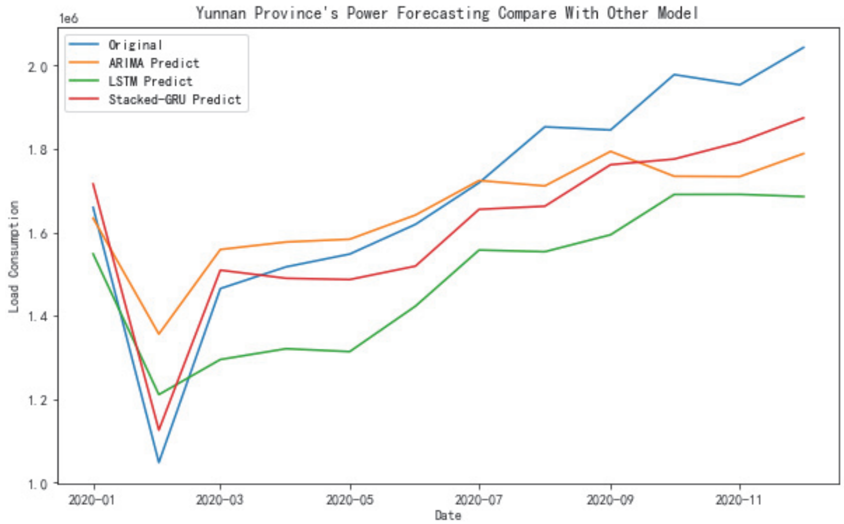

Figure 6 Yunnan province's power forecasting compare with other model.

year 2020 was satisfactorily predicted, including the sudden steepening of the power consumption curve at the beginning of year 2020.

To verify the effectiveness of the model, comparative experiments are carried out by using ARIMA, LSTM and GRU methods. The experimental results show that our model has the best effect on data set.

4 Conclusion

Based on the power load data of Yunnan Province from 2009 to 2020, a medium and long term power load forecasting model is designed and tested. The main innovations of this work includes: (1) a neural network architecture based on Stacked-GRU is proposed for medium and long term load forecasting; (2) forecasting model's input is designed by incorporating social and economic factors of regional GDP and population. In addition, we have proved through comparative experiments that Adam optimizer works the best on the Stacked-GRU model than other optimizers do. Taking MAPE, RMSE and MAE as evaluation metrics, the experimental results show that the proposed model outperforms other deep learning methods on the task of power load forecasting in Yunnan Province.

Since we can only download annual or monthly power load data from the public websites of Yunnan Provincial Bureau of Statistics and the National

Bureau of Statistics, there are limitations in data acquisition and data completeness. At the same time, some unquantifiable factors, such as government policies and emergencies, will also disturb power load forecast. Therefore, future work needs to be done to extend data acquisition channel and to consider more social and economic ecomonic factors, so as to improve prediction accuracy at data level.

If more fine-grained power consumption data at city level can be obtained, the spatial relationship of power consumption can then be mined. With the improvement of power consumption perception based on spatial relationship, one can expect a more accurate prediction of power consumption on a smaller region in Yunnan Province.

Acknowledgement

SFYP Science Foundation of Yunnan Province No.202101AT070167; OFKLSEYPG Open Foundation of Key Laboratory in Software Engineering of Yunnan Province under Grant No.2020SE311;

YKLCPSAOF Yunnan Key Laboratory of Crop Production and Smart Agriculture 2022 Open Fund No.2022ZHNY10;

MPHREOS Major Project of High-Resolution Earth Observation System No.89-Y50G31-9001-22/23.

References

[1] Yunnan Provincial Bureau of Statistics. (2021, February 1). Report on Energy Production in Yunnan in 2020. Retrieved April 26, 2022, from http://stats.yn.gov.cn/tjsj/jjxx/202102/t20210201_1044039.html
[2] Pingfei Wang. (2021). Research on Power load Prediction based on Time Seriesconvolution Lstm. Master's Thesis, Sichuan University.
[3] Shuxin Luo, Minhua Ma, Lin Jiang, Bingjie Jin, Yong Lin, Xuhao Diao, Canbing Li and Bo Yang. (2020). Medium and Long-term Load Forecasting Method Considering Multi-time Scale Data. Proceedings of the CSEE, (S1), 11–19.
[4] Qing Zhong, Wen Sun, Nanhua Yu, Chunfang Liu, Fang Wang and Xin Zhang. (2014). Load and Power Forecasting in Active Distribution Network Planning. Proceedings of the CSEE, 34, (19), 3050–3056.
[5] Song, K. B., Baek, Y. S., Hong, D. H., and Jang, G. (2005). Short-term load forecasting for the holidays using fuzzy linear regression method. IEEE transactions on power systems, 20(1), 96–101.

[6] Bracale, A., Caramia, P., De Falco, P., and Hong, T. (2019). Multivariate quantile regression for short-term probabilistic load forecasting. IEEE Transactions on Power Systems, 35(1), 628–638.

[7] Weipeng Li. (2019). Application analysis of power load forecasting based on partial least square method. Practical Electronicstor, (12), 74–75+86.

[8] Göb, R., Lurz, K., and Pievatolo, A. (2013). Electrical load forecasting by exponential smoothing with covariates. Applied Stochastic Models in Business and Industry, 29(6), 629–645.

[9] Abderrezak, L., Mourad, M., and Djalel, D. (2014, December). Very short-term electricity demand forecasting using adaptive exponential smoothing methods. In 2014 15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA) (pp. 553–557). IEEE.

[10] Liao, G. C. (2021). Fusion of Improved Sparrow Search Algorithm and Long Short-Term Memory Neural Network Application in Load Forecasting. Energies, 15(1), 130.

[11] Chuanjun Pang, Bo Zhang and Jianming Yu. (2021). Short-term power load forecasting based on LSTM recurrent neural network. Electric Power Engineering Technology, (01), 175–180+194.

[12] Xiaoyu Wu, Jinghan He, Pei Zhang and Jun Hu. (2015). Power System Short-term Load Forecasting Based on Improved Random Forest with Grey Relation Projection. Automation of Electric Power Systems, 39(12), 50–55.

[13] Yang Zhao, Hanmo Wang, Li Kang, Zhaoyun Zhang. (2022). Temporal Convolution Network-Based Short-Term Electrical Load Forecasting. Transactions of China Electrotechnical Society, 37(05), 1242–1251.

[14] Li, C., Chen, Z., Liu, J., Li, D., Gao, X., Di, F., . . . and Ji, X. (2019, August). Power load forecasting based on the combined model of LSTM and XGBoost. In Proceedings of the 2019 the International Conference on Pattern Recognition and Artificial Intelligence (pp. 46–51).

[15] Wang, Z., Wang, X., Ma, C., and Song, Z. (2021). A Power Load Forecasting Model Based on FA-CSSA-ELM. Mathematical Problems in Engineering, 2021.

[16] Alhussein, M., Aurangzeb, K., and Haider, S. I. (2020). Hybrid CNN-LSTM model for short-term individual household load forecasting. IEEE Access, 8, 180544–180557.

[17] Wu, L., Kong, C., Hao, X., and Chen, W. (2020). A short-term load forecasting method based on GRU-CNN hybrid neural network model. Mathematical Problems in Engineering, 2020.
[18] Chongqing Kang, Qing Xia and Boming Zhang. (2004). Review and Development of Load Forecasting in Power System. Automation of Electric Power Systems, (17), 1–11.
[19] Han, Y., Sha, X., Grover-Silva, E., and Michiardi, P. (2014, October). On the impact of socio-economic factors on power load forecasting. In 2014 IEEE International Conference on Big Data (Big Data) (pp. 742–747). IEEE.
[20] Moral-Carcedo, J., and Pérez-García, J. (2017). Integrating long-term economic scenarios into peak load forecasting: An application to Spain. Energy, 140, 682–695.
[21] Liu, D., Sun, K., Huang, H., and Tang, P. (2018). Monthly load forecasting based on economic data by decomposition integration theory. Sustainability, 10(9), 3282.
[22] Ghanbari, A., Naghavi, A., Ghaderi, S. F., and Sabaghian, M. (2009, March). Artificial Neural Networks and regression approaches comparison for forecasting Iran's annual electricity load. In 2009 International conference on power engineering, energy and electrical drives (pp. 675–679). IEEE.
[23] Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., and Khudanpur, S. (2010, September). Recurrent neural network based language model. In Interspeech (Vol. 2, No. 3, pp. 1045–1048).
[24] Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2015, June). Gated feedback recurrent neural networks. In International conference on machine learning (pp. 2067–2075). PMLR.
[25] Zhewen Niu, Zeyuan Yu, Bo Li, Wenhu Tang. (2018). Short-term wind power forecasting model based on deep gated recurrent unit neural network. Electric Power Automation Equipment, (05), 36–42.
[26] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.
[27] Yunnan Provincial Bureau of Statistics. (2021, November 25). Statistical Yearbook . Retrieved April 26, 2022, from http://stats.yn.gov.cn/tjsj/tjnj/
[28] Yunnan Provincial Bureau of Statistics. (2021, November 25). Yunnan Monthly Statistics. Retrieved April 26, 2022, from http://stats.yn.gov.cn/tjsj/tjnj/

[29] Willmott, C. J., and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research, 30(1), 79–82.
[30] Goodwin, P., and Lawton, R. (1999). On the asymmetry of the symmetric MAPE. International journal of forecasting, 15(4), 405–408.
[31] Chai, T., and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. Geoscientific model development, 7(3), 1247–1250.

Biographies

Zheng Yang received his M. S. degree from Beijing Jiao Tong University in 2012. He is currently work for the Electric Power Research Institute of Yunnan Power Grid as an engineer. His main research interests are computer networking and security.

Jing Cui received his M. S. degree from Kunming University of Science and Technology in 2021. Her main research fields are computer analysis and deep learning.

Qiangjian Zhang received his B.S. degree in industrial engineering from Kunming University in 2019. He is currently a master's student of Yunnan University. His main research fields are Machine Learning and Computer Vision.

Chunlin Yin received his M. S. degree from Yun Nan University in 2017. He is currently work for the Electric Power Research Institute of Yunnan Power Grid as an engineer. His main research interests are transfer learning.

Li Yang received her MA.Eng degree from Kunming University of Science and Technology in 2011. She is currently work for the Electric Power Research Institute of Yunnan Power Grid as an senior engineer. Her main research interests are digital transformation.

Pengfeng Qiu received his M. S. degree from Chongqing University in 2017. He is currently work for the Electric Power Research Institute of Yunnan Power Grid as an engineer. His main research interests are High Voltage Engineering and Power Electronics.

Kai Hu is currently work for Yunnan Power Grid Co., Ltd. as an engineer. His main research interests are Power System Analysis and Planning.

Junwen Yang, master, senior engineer, working in China Southern Power Grid. His research interest is power system analysis and planning.