

Using the Chaos SRU Neural Network Method to Predict Short-term Load of the Power Grid

Feifei He¹

¹Changjiang River Scientific Research Institute, Changjiang Water Resources Commission of the Ministry of Water Resources of China, Wuhan, Hubei, China, 430010.

Corresponding Author: Feifei He

Date of Submission: 28-12-2023

Date of Acceptance: 30-12-2023

ABSTRACT: The power grid needs to predict future daily load values for the daily planning of cascade hydropower stations. To achieve this, short-term load forecasting is done. In this article, a power grid load forecasting method based on a chaotic Simple Recurrent Unit (SRU) neural network is proposed. The method involves using the C-C method to calculate the time delay t and embedding dimension m of the load time series, then reconstructing the phase space. After that, it combines with the SRU neural network to predict future load values. The model was simulated and verified using the power grid load data of the Hubei Power Grid in 2015. The results indicate that the proposed model has a good fitting effect.

KEYWORDS: Load forecasting, Chaos, Simple Recurrent Unit; C-C method

I. INTRODUCTION

Short-term load forecasting is crucial for the safe and stable operation of the power system. Its main objective is to predict the electricity demand for the near future. This prediction has a direct impact on the operation of the electricity market, generation scheduling, grid stability, and economy.

The methods for short-term load forecasting mainly include statistical models, time series analysis, and machine learning. Statistical modeling methods use the statistical characteristics of historical load data for prediction. The moving average method, exponential smoothing method, etc. are simple and easy to use but have poor adaptability to complex load patterns[1].

The time series analysis method considers the time dependence of load data, such as the autoregressive moving average model (ARMA), autoregressive integral moving average model (ARIMA)[2], etc. These models can capture the dynamic changes in load data but have a weaker ability to handle nonlinear relationships[3]. In recent years, prediction methods based on artificial intelligence and machine learning have been widely

applied. Machine learning methods include artificial neural networks (ANN)[4], support vector regression (SVR)[5], Kalman filtering method[6], extreme learning machine (ELM)[7], etc. These methods can handle complex nonlinear relationships and high-dimensional data, thereby improving prediction accuracy.[8,9]

Although significant progress has been made in the research of short-term load forecasting, there are still some challenges. One of the most significant challenges is the widespread chaotic characteristics of power grid loads. Therefore, it is necessary to explore more accurate and robust prediction methods, such as deep learning methods that consider the chaotic characteristics of power grid loads to improve prediction performance.

II. MATERIALS AND METHODS

2.1 Phase space reconstruction

The method of chaos prediction involves finding a nonlinear model that can approximate the dynamic characteristics of the system and predict its behavior for a certain period of time. This process does not require a specific model to be set, nor does it need to consider factors that affect load or establish a holiday load model or different models for different user types. Instead, it relies on the self-similarity of chaos itself at different levels and uses phase space reconstruction to restore the multi-dimensional chaotic system of the original power system in actual load data[10].

According to the Takens embedding theorem, chaotic time series can be reconstructed into multiple vectors with an m -dimensional delay time t , and its main conditional parameters are the embedding dimension m and delay time t . By extending the original system to a high-dimensional phase space, the embedding dimensions present the hidden information features of nonlinear systems in another intuitive way.

To initiate chaos prediction, the chaotic time series (load time series) must first be recorded. Next, the embedding dimension m and delay time t

can be selected, and the phase space can be shown using formula 1.

$$X = \begin{bmatrix} x_1 & x_{1+\tau} & \dots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \dots & x_{2+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n-(m-1)\tau} & x_{n-(m-1)\tau+1} & \dots & x_{n-1} \end{bmatrix}, Y = \begin{bmatrix} x_{2+(m-1)\tau} \\ x_{3+(m-1)\tau} \\ \vdots \\ x_n \end{bmatrix} \quad (1)$$

Where $k = 1, 2, \dots, n - (m-1)\tau$.

The sample data is created by selecting the first $n - 1 - (m-1)\tau$ data points from the trajectory matrix. Each row of the trajectory matrix is used as input for the prediction model, and the data immediately following that row is used as the output. These inputs and outputs make up the samples that form the prediction model.

2.2 Simple Recurrent Unit (SRU)

The Simple Recurrent Unit (SRU) is a variant of Recurrent Neural Networks (RNN) that has a simpler architecture designed to speed up the training process[11]. Unlike other RNN variants such as LSTM and GRU, SRU has a faster training speed due to its unique structure. The basic architecture of SRU consists of two components, "light recurrence" and "highway network," as shown in Figure 1. The light recurrence component reads the input vector x_t and calculates the state sequence c_t , which captures the sequential information. The light recurrent procedure can be summarized using the following equations:

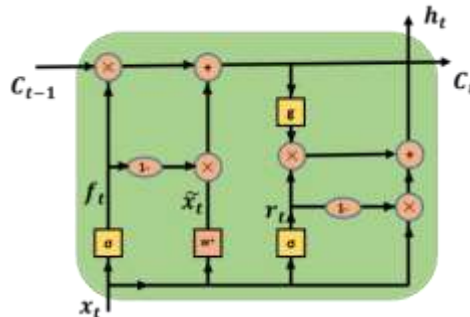


Fig. 1. The structure of the Simple Recurrent Unit (SRU)

$$\tilde{x}_t = Wx_t \quad (2)$$

$$f_t = \sigma(W_f x_t + b_f) \quad (3)$$

$$c_t = f_t \otimes c_{t-1} + (1 - f_t) \otimes (Wx_t) \quad (4)$$

Where, W_f , W , and b_f are the parameter matrices, which will be determined through the training process. $\sigma()$ is a sigmoid function and the f_t is the output of this function ranging from 0 to 1. \otimes is the Kronecker product operation.

The SRU consists of two parts, one of which is the highway network. Its primary objective is to facilitate the gradient-based training of deep networks. To accomplish this, a reset gate r_t is used to combine the input vector x_t and the current state c_t from the light recurrence. The equations for the highway network can be formulated as follows:

$$r_t = \sigma(W_r x_t + b_r) \quad (5)$$

$$h_t = r_t \otimes g(c_t) + (1 - r_t) \otimes x_t \quad (6)$$

Where W_r and b_r are also the parameters learned by the training procedure. Especially, $(1 - r_t) \otimes x_t$ is a skip connection allows the gradient to directly propagate to the previous layer, which has been shown to improve the scalability.

2.3 Calculation of Time Delay and Embedding Dimension

The selection of embedding dimension and delay parameters is critical in reconstructing phase space and requires careful consideration. The embedding theorem involves projecting the phase space of a system onto the embedding space. If the embedding dimension is too small, the phase space trajectory will be projected onto a lower dimensional space, leading to many erroneous intersections. An increase in embedding dimension will decrease the number of inaccurate crossings. However, if the embedding dimension is too large, noise in the observation data will occupy most of the embedding space, while the system attractor will only occupy a small part, masking important information with noise. Moreover, a larger embedding dimension also leads to greater computational complexity. Currently, there is no unified algorithm for selecting these two parameters. Two main views exist on the selection of embedding dimension and delay parameters. One view is that the two are unrelated, while the other suggests that the selection of one parameter is dependent on the other. Several methods have been developed to select embedding dimensions and delay parameters for both perspectives. The primary methods for selecting embedding dimensions include the correlation dimension method, the Cao method, and

the false nearest neighbor method. The primary methods for selecting delay parameters include the autocorrelation function method, the C-C method, and the mutual information method. In this paper, the C-C method is chosen to select the embedding dimension and delay parameters for sequence phase space reconstruction.

III. SIMULATION

In this section, we will be using the power grid load in Hubei Province as an example for our

research. Hubei Province is situated in central China and is located north of Dongting Lake in the middle reaches of the Yangtze River. It has a total area of 185900 square kilometers and a permanent population of nearly 60 million.

To begin with, we will be using the C-C method to determine the delay parameters of the time series. We will be selecting different sequence lengths and calculating the time delay, as shown in the figure.

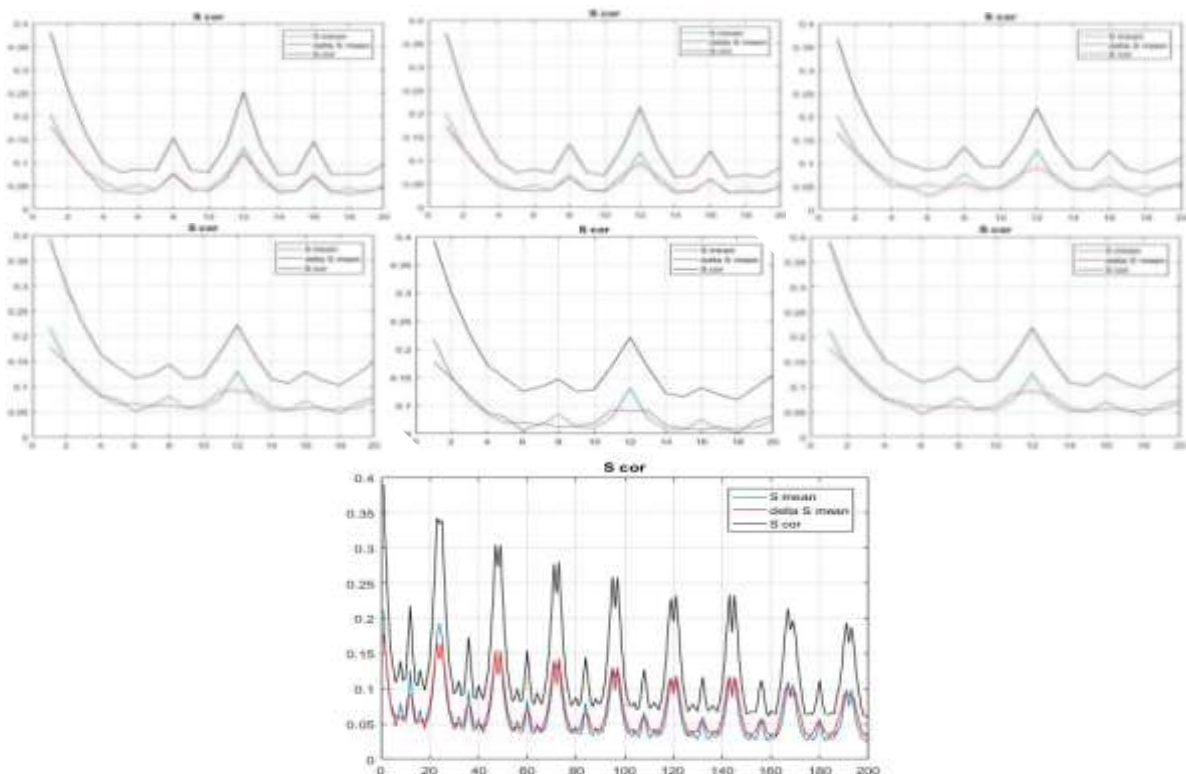


Fig. 2. Using the c-c method to determine the optimal embedding window

Table 1. Calculation results of time delay corresponding to different time series lengths

Sequence length (t)	1416	2880	4344	5832	7296	8760
Time delay (τ)	6	6	6	6	6	6

After analyzing the time delays corresponding to different sequence lengths, it was discovered that the length of the sequence has little effect on the time delay. The relationship between them is shown in Table 1. Therefore, the time delay $t=6$. From Figure 2, it can be observed that the embedding window and the embedding dimension

$m=5$ are obtained using the formula.

Next, the model is trained using the data from April 16, 2012, to April 16, 2015, as the training set. The test set consists of the load data of Hubei Province's power grid on April 17, 2015. The goal is to analyze the predictive performance of the model.



Fig. 3. Comparison chart and error distribution chart of load forecasting results

Table 2. Table of load forecasting results and relative error

Time	Real(MW)	Forecast(MW)	Relative Error	Time	Real(MW)	Forecast(MW)	Relative Error
0	14386	14299	0.006	12	16181	16780	0.037
1	14094	14235	0.010	13	16381	16794	0.025
2	13923	14058	0.010	14	16531	16761	0.014
3	13747	13910	0.012	15	16700	16873	0.010
4	13727	14090	0.026	16	17058	17068	0.001
5	14179	14657	0.034	17	17107	17162	0.003
6	15636	15946	0.020	18	16547	17297	0.045
7	16834	17010	0.010	19	17231	16911	0.019
8	17536	17578	0.002	20	17025	16260	0.045
9	17705	17675	0.002	21	16020	15695	0.020
10	17135	17331	0.011	22	15713	15219	0.031
11	16922	17109	0.011	23	14767	14591	0.012
average							0.017

From Figure 3 and Table 2, it can be seen that the model has good fitting performance, with an average relative error of 0.017, achieving good predictive performance. It is evident that traditional power load forecasting models typically use statistical methods to establish the mapping relationship directly between historical sequences and predicted values. These methods do not take into account the laws of short-term load changes in the power grid. However, after phase space reconstruction, the intrinsic regularity of power load data can be discovered. This regularity is a non-linear mapping that can be described, and the model proposed in this article can reflect this non-linear mapping relationship accurately, thus achieving high prediction accuracy.

IV. CONCLUSION

This article proposes a short-term load forecasting model based on a chaotic SRU neural network model that combines chaotic phase space reconstruction with an SRU deep neural network to achieve short-term load forecasting. The article uses Hubei Province as an example to analyze the load forecasting results of Hubei Province. In the case study of power grid load forecasting on April 17, 2015, the proposed model's relative error was 0.017, indicating that the model has a high fitting effect. The model proposed in this article has practical value in improving the short-term load forecasting of the power grid.

This work was supported by the Natural Science Foundation of Hubei Province (2021CFB151); the Central research institutes of

basic research and public service special operations (CKSF2021441/SZ); and special thanks are given to the anonymous reviewers and editors for their constructive comments.

REFERENCES

- [1]. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- [2]. Box GE, Pierce DA. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J Am Stat Assoc* 1970;65:1509–26
- [3]. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- [4]. Chan ZSH, Ngan HW, Rad AB, David AK, Kasabov N. Short-term ANN load forecasting from limited data using generalization learning strategies. *NEUROCOMPUTING*. 2006;70:409-419.
- [5]. Che J, Wang J. Short-term load forecasting using a kernel-based support vector regression combination model. *APPL ENERG*. 2014;132:602-609.
- [6]. Paliwal K, Basu A, "A speech enhancement method based on Kalman filtering," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'87.*, (IEEE, 1987), pp. 177-180.
- [7]. Li S, Goel L, Wang P. An ensemble approach for short-term load forecasting by extreme learning machine. *APPL ENERG*. 2016;170:22-29.
- [8]. Subbiah, Siva Sankari, and Jayakumar Chinnappan. "A review of short-term load forecasting using deep learning." *International Journal on Emerging Technologies* 11.2 (2020): 378-384.
- [9]. Kong, X., Li, C., Zheng, F., & Wang, C. (2019). Improved deep belief network for short-term load forecasting considering demand-side management. *IEEE transactions on power systems*, 35(2), 1531-1538.
- [10]. Zhang Yong-qiang. "Study on short term power load forecasting based on chaos theory" ShenYang Ligong University, 2013
- [11]. LEI, T. & DAI, H. (2017), "Simple Recurrent Units for Highly Parallelizable Recurrence".