# Energy Demand Forecast in Yunnan Province Based on Seq2Seq Model

Anrui Li[2], Shi Su[1], Tong Han[1], Chunlin Yin[1], Jie Li[1], Linjie Chen[2], Xinkai Wang[2], Shichen Wang[2] and Siyang Liu[1, *]

[1] Electric Power Research Institute, Yunnan Power Grid Co. Ltd, Kunming, China
[2] Key Laboratory in Software Engineering of Yunnan Province, School of Software, Yunnan University, China

**Abstract.** Energy demand forecast has an important practical significance for the sustainable development of the national economy, the reasonable allocation of resources, and the construction of modernization goals. This study is based on the analysis of coal, electricity, natural gas, and other energy data in Yunnan Province from 2011 to 2018 and uses long short-term memory, sequence to sequence, deep learning, and ridge regression coupling methods to construct an energy demand forecast model in Yunnan Province. Forecast results show the following. The total energy consumption of Yunnan Province from 2021 to 2025 will continue to increase. Moreover, the coal consumption of Yunnan Province will continue to decline from 2021 to 2025. Furthermore, the electricity consumption of Yunnan will increase by about 8.02% year-on-year from 2021 to 2025. The experiment proves that the forecasting effect of the energy demand forecast model proposed in this study is excellent.

## 1 Introduction

The impact of the construction of the Energy Internet on the traditional energy industry has revolutionary challenges and huge opportunities. The traditional energy industry should combine its own characteristics and rely on Internet technology and Internet concepts to make a reasonable change in its positioning to adapt to the development trend of the Energy Internet. Therefore, predicting the relevant demand of the energy industry is particularly important to determine the direction of the development of traditional industries.

In recent years, many scholars have conducted considerable research on energy demand forecasting. Common forecasting methods can be divided into trend extrapolation and scenario analysis. The trend extrapolation method is mainly based on past and current development trends to infer possible future states, such as time series method [1], energy elasticity coefficient method [2], and input–output method [3-5]. The scenario analysis method mainly identifies the external factors affecting the development of the main body of the research through the study of the environment and simulates various cross scenarios that may occur in the external factors to analyze and predict various possible prospects [6-9].

On the basis of the situation of Yunnan Province, this study selects and obtains comprehensive data, including energy, electric power, coal, wind power, natural gas, and non-fossil energy the from National Bureau of Statistics, power companies, and other data source units through national and local statistical channels, energy companies, and industry associations. On the basis of the energy data of Yunnan Province, the energy demand forecast model of Yunnan Province is established through a comprehensive analysis of the energy supply and demand relationship and energy structure. By using this model, the energy demand in Yunnan Province from 2021 to 2025 is predicted, and the energy development trend is also analyzed.

## 2 Models and forecasting methods

Long short-term memory (LSTM), sequence to sequence (Seq2Seq), deep learning, and ridge regression coupling prediction model are used to predict the total energy, coal, electricity, and natural gas consumption in Yunnan Province. LSTM performs well on unstable time series with fixed components, and deep learning performs accurately in model prediction.

### 2.1 Energy demand forecasting model based on LSTM

LSTM is a special RNN that selects long-term information through the "gate" design, thereby avoiding the problem of long-term dependence. The gates of LSTM at each sequence index position $t$ generally include forget, input, and output gates. The forget gate controls whether information is forgotten, the input gate processes the input of the current sequence position and determines what new information can be updated in the cell state, and the output gate is responsible for processing the output information.

The calculations corresponding to the forget, input, and output gates are as follows:

---

* Corresponding author: lsyfd27@sina.com

$$f_t = \sigma\left(W_f \cdot [h_{t\text{-}1}, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_i \cdot [h_{t\text{-}1}, x_t] + b_i\right) \tag{2}$$

$$C_t = f_t * C_{t\text{-}1} + i_t * tanh\left(W_C \cdot [h_{t\text{-}1}, x_t] + b_C\right) \tag{3}$$

$$h_t = \sigma(W_o[h_{t\text{-}1}, x_t]) * tanh(C_t) \tag{4}$$

When we use 7-day data to predict the energy demand for the next day, the input of the model is $[x_0, x_1, ..., x_6]$. For example, when calculating the expression of $x_1$, the parameter learns through the forget gate $W_f$. The expressions of $x_0$ (i.e., $h_0$) and $x_1$ are used to calculate the proportion that needs to be forgotten $f_1$. The input proportion $i_1$ is calculated by the parameter learned through the input gate $W_i$, the expression of $x_0$ (i.e., $h_0$), and $x_1$. The state of $x_0 C_0$ at the previous moment is controlled through the calculation results of the forget and input gates. Through the forgotten part of information and the passing ratio of the input $tanh(W_C[h_{t\text{-}1}, x_t] + b_C)$, the state of $x_1 C_1$ can be obtained. Finally, the output gate parameter $W_o$ can obtain the following.

## 2.2 Energy demand forecasting model based on Seq2Seq

When the Seq2seq model maps a sequence as an input to another output sequence, it mainly completes the two basic links of the encoder input and the decoder output. The encoding end encodes the input sequence data and outputs the encoding status on the network end. The decoder receives the final state of the encoder as the initial state, and the output value of each step after decoding is the input value of the next step.

On the basis of the energy demand data of different time scales, a chronological data sequence is obtained, such that each piece of data contains at least two parts, namely, the time sequence and the corresponding energy consumption. The time sequence is used as a feature to input into Seq2Seq, and the corresponding energy consumption as forecast content. For the energy demand forecasting task of the time scale of day, week, and month, weather characteristics can be added. The weather forecast data in the future can be used to form the comprehensive features with the time series as the input of the Seq2Seq model.

## 2.3 Deep learning and ridge regression coupling forecast model

In comparison with traditional machine learning methods, the deep learning model eliminates the feature selection process and can learn the feature selection strategy by itself. Ridge regression is a biased estimation regression method for multicollinearity data analysis. It is an improvement on the basis of the least squares method. It does not consider the unbiased of the least square method and loses part of the hidden information of the characteristic factors. Thus, it establishes a realistic regression model. The regression coefficients of the ridge

regression model can reflect the overall trend and avoid the overfitting and underfitting problems of the least squares linear regression, making the model practical. Coupling weather forecast data based on the ridge regression method and building an energy demand model and making predictions can stabilize forecast results and improve forecast accuracy.

An energy demand forecast model with a forecast period of 1 month is established. First, the monthly weather forecast data in this area of Yunnan Province is obtained based on the inverse distance weighting method, and the energy demand forecast model based on the LSTM network is obtained by using data training, with the forecast data obtained. Then, through monthly energy consumption rolling forecasts, on the basis of the measured data, the sliding load is carried forward time by time to revise the forecasts of energy demand for subsequent months. Finally, a ridge regression model is established, with the independent variables being the forecast data of weather factors and the energy consumption of next month predicted by the LSTM network model.

When collecting weather data, weather forecast data are collected for the next month to construct a mixed feature $x$ = [LSTM-out, temperature, humidity, weather...], normalize $x$, and send it to the ridge regression model, which obtains the hidden layer output $h_i = W x_i$, and finally obtains the prediction result through the output layer. The ridge regression model uses $J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2 + \lambda \sum_{j=1}^{n} \theta_j^2$ as the loss function. During training, the mean square error between the predicted value $h$ and the true value $y$ is calculated, and a regular term is added to alleviate the problem of model overfitting and underfitting.

# 3 Model building

On the basis of the different years and types of data collected, LSTM, Seq2Seq, deep learning, and ridge regression coupling prediction models are used to construct prediction models for total energy and each energy component. Particularly, the total energy consumption is subject to macro-control; hence, the Seq2Seq model is used. For the construction of a single energy demand model, the LSTM model and the coupling prediction model based on deep learning and ridge regression are used to build the prediction model.

## 3.1 Forecast model of total energy consumption in Yunnan Province

Experimental data: Energy consumption in Yunnan Province, as shown in Figure. 1.
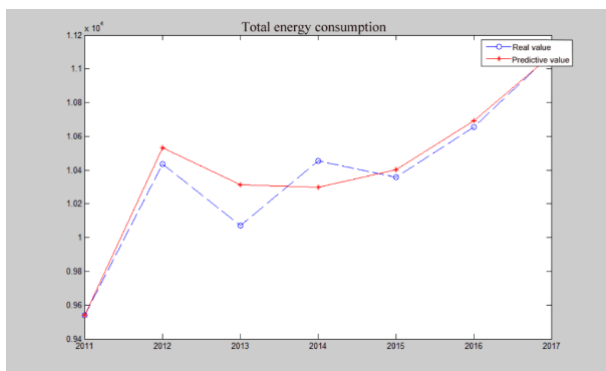
Application model: Energy demand forecasting model based on Seq2Seq

Associated data: Yunnan's GDP, Yunnan's population, and output value of main industrial products
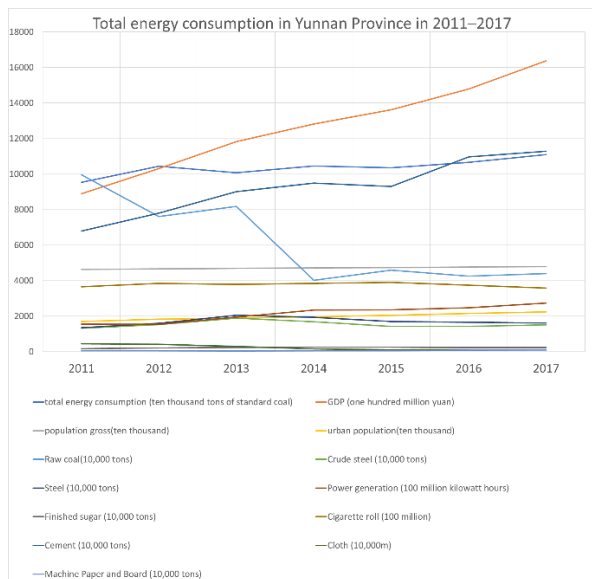
**Figure 1.** Total energy consumption in Yunnan Province in 2011–2017 (Data source: Yunnan Statistical Yearbook)

The fitting result is shown in Figure. 2.



**Figure 2.** Energy consumption forecasting model of Yunnan Province based on Seq2Seq model

### 3.2 Thermal power production forecast model in Yunnan Province

Experimental data: Thermal power production in Yunnan Province, as shown in Table 1, and hydropower production in Yunnan Province

Application model: Energy demand forecasting model based on LSTM

Associated data: Hydroelectric power generation and precipitation in Yunnan Province

**Table1.** Thermal power production in Yunnan Province

| Year | Thermal power generation (hundred million kWh) | Annual precipitation (mm) |
|---|---|---|
| 2013 | 479.3 | 835.271172 |
| 2014 | 391.41 | 980.9 |
| 2015 | 266.47 | 1106.60238 |
| 2016 | 238.04 | 1154.6 |

The fitting result is shown in Figure. 3.
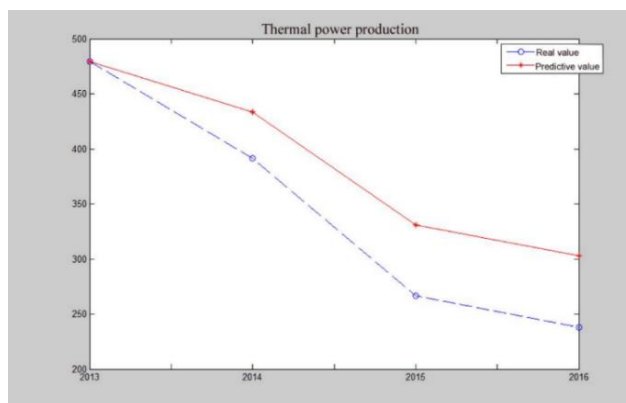


**Figure 3.** Thermal power production forecast model in Yunnan Province based on LSTM model

### 3.3 Simulation of wind power production in Yunnan Province

Data: Wind power production in Yunnan Province, as shown in Table 2., and temperature in Yunnan Province (Yunnan Statistical Yearbook)
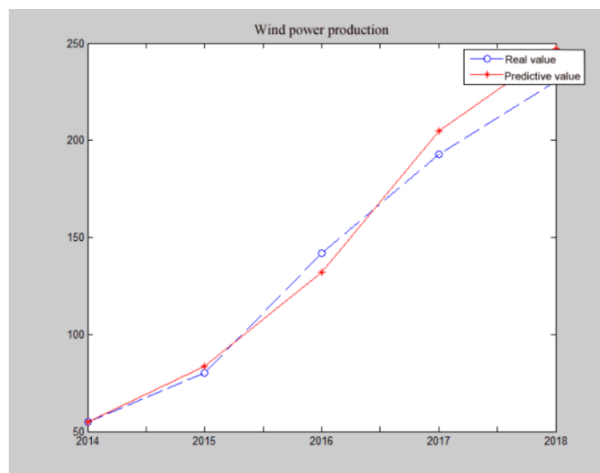
Application model: Coupling prediction model based on deep learning and ridge regression

Associated data: Weather in Yunnan Province

**Table2.** Wind power generation in Yunnan Province from 2014 to 2018 (hundred million kWh)

| | Wind power generation (hundred million kWh) |
|---|---|
| 2018 | 230.7 |
| 2017 | 192.9 |
| 2016 | 141.7 |
| 2015 | 80.3 |
| 2014 | 55 |

The fitting result is shown in Figure. 4.



**Figure 4.** Forecast model of wind power production in Yunnan Province based on the coupling model of deep learning and ridge regression

## 4 Model accuracy analysis

The comparison between the predicted value of the model simulation part and the true value curve indicates that the prediction simulation results are quite impressive. Thus, the total energy, raw coal, electricity, and natural gas in Yunnan Province of China are predicted by this prediction model. Some of the data in 2019 in the trained model are compared with the corresponding data in 2019 in the 2020 Statistical Yearbook of Yunnan Province to analyze the accuracy of the model.

**Table3.** Part of the energy demand forecast tables in Yunnan Province

| Category | Forecast of 2019 (ten thousand tons of standard coal) | Official data for 2019(ten thousand tons of standard coal) |
|---|---|---|
| Total energy consumption | 12663.08 | 12157.65 |
| Raw coal consumption | 5921.596 | 6759.46 |
| Power consumption | 1256.4 | 1112.40 |

According to the analysis of the forecast table, the forecast of total energy consumption and electricity consumption is accurate, but the forecast value of raw coal consumption is slightly lower than the data published in the Yunnan Statistical Yearbook. The main reason for the difference of raw coal consumption may be that in 2019, the government increases the input of raw coal consumption to support the corresponding industrial development, but the prediction results of other energy consumption structure are basically consistent with the actual situation, and the accuracy can meet the prediction demand.

## 5 Model prediction

### 5.1 Prediction of total energy consumption

On the basis of the National Statistical Yearbook of China, the regional GDP of Yunnan Province has maintained a rapid growth, which is higher than the national level. From 2011 to 2017, Yunnan's GDP rose from 0.88 trillion yuan to 1.8 trillion yuan.

The simulation results of the 2011–2017 data of Yunnan Province based on the forecast model indicate that the total energy consumption of Yunnan Province from 2021 to 2025 will still show an upward trend, as shown in Table 4.

**Table4.** Forecast of total energy consumption in Yunnan Province

| Year | Forecast value of total energy consumption (ten thousand tons of standard coal) |
|---|---|
| 2021 | 13211.21 |
| 2022 | 13771.29 |
| 2023 | 14353.93 |
| 2024 | 14958.37 |
| 2025 | 15554.95 |

### 5.2 Raw coal consumption forecast

On the basis of the gray prediction model, the output of raw coal in Yunnan Province in 2021 can reach 50 million tons. According to policy forecasts, from 2021 to 2025, Yunnan's raw coal output can reach 55 million tons, which basically meets the coal consumption demand in the province.

The simulation results of the 2012–2018 data of Yunnan Province based on the forecast model indicate that the coal consumption of Yunnan Province from 2021 to 2025 will show a downward trend.

However, combined with the advancement of hydropower, aluminum, hydropower, and silicon integration projects in Yunnan Province, the power supply and demand in Yunnan Province will change from oversupply to a basic balance between supply and demand. During the dry period, power shortage will be experienced, the role of thermal power will be played out, and the decline trend of coal consumption tends to be flat. Therefore, the forecast of coal consumption in Yunnan Province from 2021 to 2025 is shown in Table 5.

**Table5.** Forecast of raw coal consumption in Yunnan Province

| Year | Raw coal (unit: hundred million tons) |
|---|---|
| 2021 | 5832.5 |
| 2022 | 5744.744 |
| 2023 | 5658.308 |
| 2024 | 5573.173 |
| 2025 | 5489.319 |

### 5.3 Electricity load forecasting

On the basis of model prediction, the electricity consumption in Yunnan Province will increase by about 8.02% year-on-year from 2021 to 2025. However, as the new large industrial load in Yunnan Province is expected to be 3.4 million kW in 2020, under the condition that production is put into operation on time, it is estimated that the new electricity will be about 15 billion kWh, and the largest electricity load includes Matang Hydropower Aluminum Phase I (820000 kW), Funing Shenhuo Group Hydropower Aluminum Project Phase I (750000 kW), Qiya Electrolytic Aluminum (540000 kW), Heqing Yixin Aluminum Co., Ltd. Phase II (350000 kW), Lianghe County Yunju Technology Co., Ltd. (200000 kW), Julida Iron and Steel Co., Ltd. of Yunnan Yonggang Iron and Steel Group (100000 kW), Dehonghebao Technology Co., Ltd. (70000 kW), Lijiang Longji Silicon Material Company Phase II (65,000 kW), and Baoshan Longji Silicon Material Company Phase III (62000 kW). Through a comprehensive consideration, on the basis of the abovementioned production situation, the total electricity consumption in the province in 2020 is about 192.9 billion kWh, with a year-on-year increase of 12.3%.

Without the construction of megaprojects, the electricity consumption in Yunnan Province is predicted to increase by about 8.02% year-on-year from 2021 to 2025.

## 6 Discussion

On the basis of the prediction results of various future energy consumption indicators by our trained prediction model, our recommendations are as follows.

(1) Yunnan Province should develop clean coal power actively and promote deep peak shaving of thermal power units.

With the rapid growth of power consumption in and outside the province and the slowdown in power production in the province, the scale of surplus hydropower in Yunnan Province during the flood season has shown a trend of shrinking year by year, and the contradiction of power supply balance in the whole province has gradually changed from promoting the full consumption of clean energy to the transformation of power supply in the dry period. The supporting role of thermal power supply in dry period has become prominent. With the profound development of the Energy Internet, the power market compensation mechanism is gradually improving, and the thermal power in Yunnan Province can be fully positioned as a peak-shaving power source and a backup power source. In future energy planning, thermal power location, scale, and access system should be discussed according to the above positioning.

(2) Yunnan Province should further explore commercial innovations to promote the transformation of traditional energy companies into integrated energy service providers.

With the development of distributed power generation and regional energy networks, integrated energy companies will exist in the future, which can centrally supply water, electricity, gas, cooling, heat, and other resources and realize the "water and electricity centralized reading" of intelligent meter, intelligent water meter, and intelligent gas meter through power fiber to the home. The two-way interactive demand of users in the era of mobile Internet will prompt enterprises to continue to extend their service areas and tap the added value of energy products and services. Traditional energy enterprises in Yunnan Province should promote the construction of intelligent power service system actively, deploy intelligent power interactive terminals, improve residents' awareness of intelligent power consumption, build demonstration intelligent communities, improve the intelligent level of charging, support the innovation of charging business model, and transform from a single energy supplier to a comprehensive energy service provider.

(3) Yunnan Province should speed up the replacement of electric energy to build a green Yunnan.

Yunnan shall accelerate the replacement of electric energy, use secondary electric energy to replace traditional primary fossil energy consumption, such as electric vehicles/hydrogen energy vehicles instead of fuel vehicles, and promote the combination of the advantages of clean energy of hydropower, silicon and aluminum resources, and industrial basis, push forward the integrated development of hydropower aluminum and hydropower silicon industry, and accelerate the full consumption of clean energy in Yunnan Province.

## Acknowledgments

## References

1. T. Wang, H. Leung, J. Zhao and W. Wang, "Multiseries Featural LSTM for Partial Periodic Time-Series Prediction: A Case Study for Steel Industry," in *IEEE Trans. Instrum. Meas.*, **69**, no. 9, pp. 5994-6003 (2020)

2. H. Yang, J. Zhang, J. Qiu, S. Zhang, M. Lai and Z. Y. Dong, "A Practical Pricing Approach to Smart Grid Demand Response Based on Load Classification," in *IEEE Trans. Smart Grid*, **9**, no. 1, pp. 179-190 (2018)

3. J. Ke et al., "Comprehensive Economic Benefit Assessment Method and Example of Energy Storage Based on Power Grid," *iSPEC - IEEE Sustain. Power Energy Conf.: Grid Mod. Energy Revolut., Proc.*, pp. 2287-2292 (2019)

4. A. Enescu, H. Andrei, E. Diaconu and V. Ion, "Numerical method for modeling the input-output characteristic in a cogeneration power plant," *Proc. Int. Conf. Electron., Comput. Artif. Intell., ECAI*, pp. 1-6 (2019)

5. M. Yan, N. Zhang, X. Ai, M. Shahidehpour, C. Kang and J. Wen, "Robust Two-Stage Regional-District Scheduling of Multi-carrier Energy Systems With a Large Penetration of Wind Power," in *IEEE Trans. Sustainable Energy,* **10**, no. 3, pp. 1227-1239 (2019)

6. F. Zhang et al., "Understanding WeChat User Preferences and "Wow" Diffusion," in *IEEE Trans Knowl Data Eng* (2021)

7. Z. Zhao, H. Zhou, L. Qi, L. Chang and M. Zhou, "Inductive Representation Learning via CNN for Partially-Unseen Attributed Networks," in *IEEE Trans. Netw. Sci. Eng.*, **8**, no. 1, pp. 695-706 (2021)

8. Z. Zhao, H. Zhou, L. Qi, L. Chang and M. Zhou, "Inductive Representation Learning via CNN for Partially-Unseen Attributed Networks," in *IEEE Trans. Netw. Sci. Eng.*, **8**, no. 1, pp. 695-706 (2021)

9. W. Ahmed et al., "Machine Learning Based Energy Management Model for Smart Grid and Renewable

Energy Districts," in *IEEE Access*, **8**, pp. 185059-185078 (2020)