Prediction of energy consumption: Variable regression or time series? A case in China

Yan Li

College of Economics and Management, Taiyuan University of Technology, Taiyuan, China

Correspondence
Yan Li, College of Economics and Management, Taiyuan University of Technology, Taiyuan 030024, China. Email: liyan@tyut.edu.cn

Funding information
The national social science foundation of China, Grant/Award Number: 16CJY074

Abstract
Energy consumption is closely related to industrial structure, economic prosperity, and population. Because of the Granger causal relationship between GDP and energy consumption, many researchers consider the regression method to predict energy consumption using economic variables. However, some other researchers regard the time series method (forecast) to estimate for energy consumption. To address the advantages and disadvantages of two methods in energy consumption prediction, we performed the deep learning (DL) based on the case of China. The experiment results show that the accuracy of time series forecast is much higher than regression methods. In addition, we also explored the relationship between industry, economy, population, and energy consumption. We found that GDP, population, secondary sector, and tertiary sector were significant to energy consumption. A tertiary sector can help to reduce energy consumption to some extent. This could give suggestions about changing industry structure. More and more attention has been paid to energy intensity. The strategic approach to sustainable development requires reducing energy intensity while developing the economy. Since China is committed to achieving carbon emission peak by 2030, this paper predicts the energy intensity of China in the next decade based on a DL algorithm. The prediction results show that energy intensity would slightly decrease in the next decade. Different types of energy represent the industrial structure behind economic growth. Therefore, this paper was conducted using ridge regression (RR) to explore the relationship between different types of energy and energy intensity. In addition, the effect (positive or negative) of different energy sources in reducing energy intensity has also been explored. We found that the best way of developing the economy and reducing energy consumption is to use natural gas and hydroelectric instead of coal. This can guide policy development about changing energy source structure.

KEYWORDS
China, deep learning, energy consumption, long short-term memory, ridge regression, time series
INTRODUCTION

The policy of “Reform and Opening-up” has greatly promoted economic growth in China. China’s economy has been growing rapidly since 1978. However, huge energy cost is behind the rapidly growing economies. Many researchers have found that the economic growth has a strong relationship with energy consumption. Wang et al. have found that Granger causal relationship has existed between economic growth and energy consumption. In addition, they consider the relationship would be a bidirectional causal relationship. Therefore, energy consumption has become a double-edged sword to a country’s development. On the one hand, energy consumption can bring economic growth. On the other hand, increasing energy demand puts a heavy burden on the environment. Energy consumption has been generally recognized to be the main impact factor to cause CO₂ emissions.

Taking China as an example, although it has not yet achieved the tasks of industrialization and urbanization with the rapid economic growth, it has become the largest energy consumer and the world’s greatest CO₂ emitter. The contradiction between eliminating poverty, improving citizens’ lives, and environmental pollution makes China’s economic development still face an arduous journey. For sustainable development and environmental responsibility, one of China’s commitments under the Paris Accord is to achieve the peak of carbon dioxide emissions by 2030. Therefore, we take China as a case to study the relationship between energy consumption and other factors and find a reasonable way to predict energy consumption.

Energy consumption prediction has always been a hot issue. Many kind of literature focused on the relationship between energy consumption and economic growth by Granger Causality. They tried to find out some relative factors to energy consumption. However, with the development of artificial intelligence, more and more scholars use computer methods to find factors related to energy consumption and predict energy consumption. Artificial neural network (ANN), support vector machine (SVM), and many other algorithms are the most common way to do the prediction. It is clear that these methods’ performance is better than the linear models’ when dealing with the huge amount of non-linear data. (1) Some researchers use the regression method to predict energy consumption. Pao found out the effect of the national income (NI), population (POP), GDP, and consumer price index (CPI) on electricity consumption using linear regression and ANN. Szoplik predicted the natural gas demand in Szczecin (Poland) by multilayer perceptron model (MLP) using temperature and other factors. Günay used population, GDP per capita, inflation, and average summer temperature to regress the annual gross electricity demand of Turkey for the years 2014-2028 by ANN. (2) However, some researchers have used time series to predict energy consumption. Jebaraj et al. considered coal consumption in India as time series and predicted it by ANN. Wang et al. used support vector machine (SVM) to predict the annual electricity consumption of Beijing.

Some literature showed the superiority of ANN when it is used in the prediction of energy consumption. However, with the rapid development of ANN, deep learning (DL) methods have become a more powerful way to solve all the problems which ANN could do. ANN is the shallow neural network but DL is the deep neural network with stronger learning ability, which means it could find the more precise regular pattern behind the huge nonlinear data.

Therefore, this paper used a DL algorithm to compare the difference between the regression method and time series prediction. On the one hand, some parameters related to energy consumption, such as GDP, population, were mapped to energy consumption through DL to achieve prediction. On the other hand, we constructed a deep learning model to predict by considering energy consumption as a time series. Long short-term memory (LSTM) which is a kind of DL method is chosen in this paper, since it is widely used in presenting many aspects of the economy and energy.

Selecting variables before performing regression is also an important problem. Some traditional methods such as principal components analysis (PCA), Kernel principal components analysis (KPCA), or Bayes are widely used in the field of energy economics. However, these methods could not give the statistical significance between energy consumption and variables. Therefore, to overcome the multicollinearity between variables, ridge regression (RR) was used in this paper. RR not only selects significant variables for regression but also gives us P-value of each variable. It is widely used for variable selection in various fields.

In this paper, we performed RR to selected variables and figured out the relationship between variables and energy consumption. Then, we used LSTM to predict energy consumption in two ways: variables regression and time series prediction. Then, we compared the results of the two methods and selected a better method to predict the energy intensity. Finally, we found the relationship between energy sources and GDP and energy intensity using RR. This result would provide a way to change industry structure and energy consumption structure so that China can promote economy and reduce energy consumption.

METHODS

2.1 Work frame

Our work consists of three parts. (1). The first part is the model comparison. In this part, we compared regression
algorithm with time series prediction algorithm and tried to figure out the reason that causes the difference. In addition, we also identified the relationship between variables and energy consumption. (2). Prediction of energy intensity. (3). Relationship between energy type and energy intensity and relationship between energy type and GDP were found.

To explain our work, detailed information of methods we used are introduced by the sections below.

2.2 | Algorithms used in part 1

The workflow of part 1 is shown in Figure 1. Firstly, we chose GDP, population, primary sector of the economy, secondary sector of the economy, and tertiary sector of the economy as the potential factors that affect energy consumption. Then, we used RR to select variables. For variables regression, we totally built three models, we used all variables for the first model, the selected variables involved in the second model, and the last one included only GDP.

For time series prediction, we only used energy consumption data to build a model. Finally, we compared the results of these four models.

2.2.1 | Time series method

Feature selection

Since we processed the energy consumption data as a time series, we should reshape the series to a format that LSTM could process. We tried to use 5-year data to predict for the 6th year. Therefore, the length of each sample was 5. The relation between samples and target variables is shown in Formula (1):

\[
\begin{bmatrix}
data_1 & \cdots & data_m \\
data_2 & \cdots & data_{m+1} \\
\vdots & \vdots & \vdots \\
data_{n-m} & \cdots & data_{n-1}
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
data_{m+1} \\
data_{m+2} \\
\vdots \\
data_n
\end{bmatrix}
\]

(1)

In Formula (1), the left side is samples and the right side is the corresponding target variables. In this case, m is 5 and n is the whole size of our data.

In order to ensure that the prediction will not break off, after using all the raw data, the predicted \(data_{n+1}\) will be used to generate a new sample, like \([data_{n-1}data_{n-2}data_{n-1}data_n\,data_{n+1}]\).
Then, we can obtain \((n + 2)\)th data value \(\text{data}_{n+2}\). Theoretically, based on continuous circulation, the energy consumption of any moment can be obtained.

In fact, due to the effects of various factors, the prediction data are unstable. Therefore, it is difficult to do long-term prediction. So we only predict the energy consumption from 2018 to 2028.

**Basic principles of LSTM**

Long short-term memory is a specific form of recurrent neural network (RNN). Generally, RNN includes three features:

(A) RNN can produce an output at each time node, and the connection between hidden elements is cyclic;

(B) The output at one time is only connected with the hidden unit of the next time node.

(C) RNN contains hidden units with cyclic connections and can process sequential data and output a single prediction.

Unlike RNN, LSTM has three “gates.” They are forget gate, input gate, and output gate. The forget gate makes LSTM better than traditional RNN. Because LSTM adds forget gates, the original neuron adds a state, \(C\), to "reasonably" preserve its long-term state. The added state \(C\) is called “cell state” and it replaces the traditional hidden neuron nodes so it is also called “memory block.” It is responsible for transferring the memory information from the initial position of the sequence to the end of the sequence.

As we can see in Figure 2, there are three inputs in the current neuron: output from the previous moment, input from this moment, and cell state from the previous moment.

The forget gate is crucial for our work since data from few decades ago would not provide much useful information for predicting the future. But we could not know which year the amount of information begins to decrease. So we need to use LSTM to select some data from ages to be forgotten.

\(C_t\) is the cell state (where the loop occurs), \(H_t\) denotes the output of the current cell. \(\sigma\) denotes sigmoid function and \(\tanh\) denotes hyperbolic tangent function. \(X_t\) denotes the input vector.

\[
\begin{align*}
    f_t & = \sigma(W_f[h_{t-1},x_t] + b_f) \quad (2) \\
    i_t & = \sigma(W_i[h_{t-1},x_t] + b_i) \quad (3) \\
    \tilde{C}_t & = \tanh(W_C[h_{t-1},x_t] + b_C) \quad (4) \\
    C_t & = f_t \ast C_{t-1} + i_t \ast \tilde{C}_t \quad (5) \\
    o_t & = \sigma(W_o[h_{t-1},x_t] + b_o) \quad (6) \\
    h_t & = o_t \ast \tanh(C_t) \quad (7)
\end{align*}
\]

\(f_t\) denotes forget gate. \(i_t\) denotes input gate. \(\tilde{C}_t\) denotes the state in the previous moment. \(C_t\) denotes cell state in the current moment. \(O_t\) denotes output gate. \(H_t\) denotes output. All the \(W\) denotes the weight of different gates, and \(b\) denotes the bias of different gate.

**Building model by LSTM**

We build a three layers network. Each layer is the structure of LSTM. The input is a time series vector with dimension \((1,5)\), and output is the prediction result. The parameters are set as Table 1.

**2.2.2 Regression method**

**Feature selection**

We find five variables related to energy consumption from three aspects: industry, economy, and population. They are GDP, population, the primary sector of the economy, the...
secondary sector of the economy, and tertiary sector of the economy. Because the values of these variables are not in an order of magnitude, we use Z-score normalization to transform each variable into normal distribution data.

\[ x' = \frac{x - \text{mean}(x)}{\text{std}(x)} \]  

(8)

\[ y' = wx \]  

(9)

\[ L(w) = \|y' - y\|^2_2 \]  

(10)

Since RR considers the complexity of the model, so the objective function could be denoted as:

\[ L(w) = \|y' - y\|^2_2 + \lambda \|w\|^2_2 \]  

(11)

\( \lambda \) denotes the regularization term. To solve (11), we derive Equation 11.

\[ \frac{\partial L(w)}{\partial w} = 2x^T x w - x^T y - x^T y + 2\lambda w = 0 \]  

(12)

Finally, we could obtain the value of \( w \), then we could know which variables are significantly related to energy consumption.

\[ w = (x^T x + \lambda E)^{-1} x^T y \]  

(13)

**Building model by LSTM**

After selecting variables, we could put the significant variables into LSTM to build a model. Then, the input is significant variables and output is energy consumption.

The parameters are set as Table 2.

| Parameter       | Value |
|-----------------|-------|
| Layers          | 4     |
| Loss            | Mean squared error |
| Optimizer       | adam  |
| Epochs          | 100   |
| Batch size      | 10    |
| Activation      | Relu  |

---

**3 | RESULTS**

**3.1 | Data description**

We obtained the energy consumption, GDP, population, primary sector of the economy, the secondary sector of the economy, and tertiary sector of the economic data of China from 1965 to 2017. There are 6 kinds of different energy consumption data in the dataset. They are coal, oil, hydroelectric, nuclear, natural gas, and other (other renewables).
We drew a figure of GDP and energy consumption as Figure 4. The blue line is energy consumption. The unit of it is billion-ton oil equivalent. The green line is GDP. The unit of it is billion RMB. As we can see in Figure 4, they are in the same trend, which means they are highly relevant.

3.2 Results of model comparison

Firstly, we should use RR to select the most relevant variables. As we can see in Table 3, all the P-values are shown quite low except primary sector of the economy. It makes sense since the primary sector mainly includes farming, forestry, mining, and fishing. They do not cost too much energy. Another important point is that the coefficient of the tertiary sector is a negative number, which means the development of the tertiary industry will relatively reduce energy consumption.

After screening variables, we used 70% data to build a model and the rest 30% to test for validation, so we used the data from 1965 to 2001 to train and data from 2002 to 2017 to judge the accuracy of the model that we built. We finally built four models. (1) since GDP is generally considered highly correlated with energy consumption, we only used GDP to predict energy consumption. (2) we used all variables to predict energy consumption. (3) we used significant variables (GDP, population, secondary sector, and tertiary sector) to predict energy consumption. (4) we regarded energy consumption data as time series to predict.

The results of the prediction are shown in Figure 5.

As we can see in Figure 5, the precision of these four methods was almost the same. To figure out the best one, we calculated the error of four results.

As we can see in Table 4, there are three evaluating indicators. MAE means mean absolute error. MSE means mean square error. MaxError means the max error. It is easy to figure out that time series prediction is the best. Prediction by GDP is the worst. If we only use the most relevant variables to regress, the precision would increase.

Through these experiments, we could know that the precision of time series prediction is higher than regression method (at least in China’s data) since we used the same method (LSTM) to do the experiments in two different ways. The superiority of time series prediction lies not only in high accuracy but also inconvenient calculation. Firstly, we only need energy consumption data to predict future development, while regression methods need more variables. In addition, if we use regression methods to predict energy consumption in 2020, we should at least know the GDP and other factors in that year, however, if we use time series prediction, we only need the energy consumption data from past years (The more data, the higher the accuracy).

Due to the superiority of time series prediction, we used it to predict energy intensity.
3.3 Prediction of energy intensity

Firstly, we converted units of energy consumption to ton oil equivalent and convert units of GDP to billion dollars. Then, we calculated energy intensity as energy consumption/GDP. Finally, we used LSTM to predict the energy intensity of China in the next decade.

To verify the accuracy of our method, we firstly used 80% data to build a model and used 20% data to test the precision of our model. The result is shown in Figure 6.

As we can see in Figure 6, the blue (test result) almost coincides with a red line (true energy intensity), which means that our prediction result is very accurate. The green line is the training result. The MAE of this experiment is 0.059, which proves that our method could be used to predict the energy intensity of China in the next decade (2018-2028).

Figure 7 is the result of a prediction. We could see that the energy intensity would decrease in the next decade, but the degree of decline is not very large.

The decline in energy intensity illustrates China’s raising awareness of environmental protection and the changes in energy consumption structure. China has gradually reduced its dependence on coal and has developed and applied clean energy. Correspondingly, the proportion of the secondary industry is gradually decreasing, while the proportion of the tertiary industry is rising very fast. According to this trend, China is very likely to reach a peak of carbon emissions before 2030.

3.4 Energy type and energy intensity & GDP

We also wanted to know which kind of energy contributes to the development of GDP most. Therefore, we used RR to calculate the relationship between Energy types and GDP.

As we can see in Table 5, to our surprise, coal and oil are not considered as the main factor to affect GDP. The reason should be that they are always the major energy source and there is no significant change in these two energy sources.

In another way to explain this, coal and oil are the important support for national economic development, but the upper limit of economic development needs other clean and efficient energy sources. Another important information that the table shows is the development of other renewables causes the decrease in GDP (coefficient of “Other” is negative). The reason might be the contradiction between high-cost development technology and low utilization rate. The development of other renewable energy technologies in China is still immature, which undoubtedly has an impact on China’s economy.

The other important thing we found is the relevance between energy types and energy intensity.

As we can see in Table 6, oil and hydroelectric are significantly relevant to energy intensity and natural gas has less significant relevance. It is worth noting that natural gas and hydroelectric are significantly relevant to both energy intensity and GDP, which means that development of these two energy sources not only increase GDP but also reduce the cost of energy. In addition, the coefficient of natural gas is negative, which means that it can help reduce energy intensity. It shows China a road of low-carbon and high-speed development.

4 DISCUSSION

In this paper, we mainly focused on three problems. (1). We compared variables regression method with time series

![TABLE 5](image)
which means that China’s economic development will gradually get rid of the huge cost of energy consumption. A lower and lower price is achieved to convert energy into GDP.

(4). Hydroelectric, nuclear, and natural gas are significantly relevant to GDP and other renewables are less relevant to GDP. It means that the development of a new clean energy source could achieve economic prosperity too. In addition, due to the cost of development of other renewables, other renewables will lead to a decline in GDP in the short run.

(5). Oil and hydroelectric are significantly relevant to energy intensity. Natural gas is less relevant to energy intensity. Combining with (4), we could know that the development of natural gas and hydroelectric can promote economic development while reducing energy intensity.

5 | CONCLUSIONS

China is experiencing rapid economic growth with great energy consumption. Since China has promised to reach the carbon emission peak in 2030, changing in energy structure and industry structure is essential. In this paper, we give suggestions about changing the energy structure and industry structure by data analysis.

In this paper, we introduced LSTM which is one of the popular algorithms in deep learning to predict the energy consumption in two ways. One is variable regression and the other one is time series prediction. We compared the two methods and found that time series prediction is more accurate and easier to achieve. We also figured out that energy consumption is significantly relevant to GDP, population, secondary sector, and tertiary sector. The tertiary sector is the key to promote GDP and reduce energy consumption. Since time series prediction is better than variables regression, we used it to predict energy intensity. The results showed that the energy intensity of China will continue to decline then remain stable in a few years. Finally, we discovered that hydroelectric, nuclear, and natural gas are significantly relevant to GDP and other renewables are less relevant to GDP. Oil and hydroelectric are significantly relevant to energy intensity. Natural gas is less relevant to energy intensity. Hydroelectric and natural gas might be the key to promote GDP and reduce energy consumption.

Overall, the best way to develop Economy and reduce Energy Consumption is to develop tertiary industry and use natural gas and hydroelectric to replace coal.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

---

**TABLE 6**  Relevance between Energy types and Energy intensity

| Energy Type | Coefficient | T value | P-value  |
|-------------|-------------|---------|----------|
| Coal        | 2.7e-03     | 0.24    | 0.812    |
| Oil         | 0.28        | 6.14    | 7.82e-10*** |
| Hydroelectric | 0.39     | 3.45    | 0.0005*** |
| Nuclear     | 0.39        | 0.66    | 0.51     |
| Natural gas | -0.33       | 2.32    | 0.02*    |
| Other       | 0.30        | 1.08    | 0.28     |
DATA AVAILABILITY STATEMENT

Anyone interested in data and code can contact the author to ask for it. The data and code are in the supplemental file.

ORCID

Yan Li https://orcid.org/0000-0003-4064-4532

REFERENCES

1. Wang S, Fang C, Wang Y, Huang Y, Ma H. Quantifying the relationship between urban development intensity and carbon dioxide emissions using a panel data analysis. Ecol Ind. 2015;49:121-131.
2. Wang S, Li Q, Fang C, Zhou C. The relationship between economic growth, energy consumption, and CO2 emissions: Empirical evidence from China. Sci Total Environ. 2016;542:360-371.
3. Nejat P, Jomehzadeh F, Taheri MM, Gohari M, Majid M. A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries). Renew Sustain Energy Rev. 2015;43:843-862.
4. Zhu B, Jiang M, He K, Chevallier J, Xie R. Allocating CO2 allowances to emitters in China: A multi-objective decision approach. Energy Pol. 2018;121:441-451.
5. Mi Z, Wei Y-M, Wang B, et al. Socioeconomic impact assessment of China's CO2 emissions peak prior to 2030. J Clean Prod. 2017;142:2227-2236.
6. Paul S, Bhattacharya RN. Causality between energy consumption and economic growth in India: a note on conflicting results. Energy Econ. 2004;26(6):977-983.
7. Chiu-Wei SZ, Chen C-F, Zhu Z. Economic growth and energy consumption revisited—evidence from linear and nonlinear Granger causality. Energy Econ. 2008;30(6):3063-3076.
8. Dreyfus G. 'Neural networks: methodology and applications', 2005, Springer Science & Business Media.
9. Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev. 2015;50:1352-1372.
10. Zhang X-P, Cheng X-M. Energy consumption, carbon emissions, and economic growth in China. Ecol Econ. 2009;68(10):2706-2712.
11. Pao H-T, Tsai C-M. Multivariate Granger causality between CO2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. Energy. 2011;36(1):685-693.
12. Szoplik J. Forecasting of natural gas consumption with artificial neural networks. Energy. 2015;85:208-220.
13. Güney ME. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. Energy Pol. 2016;90:92-101.
14. Jebaraj S, Iniyavan S, Goic R. Forecasting of coal consumption using an artificial neural network and comparison with various forecasting techniques. Energy Sources, Part A Recov Utilizat Environment Effects. 2011;33(14):1305-1316.
15. Wang J, Li L, Niu D, Tan Z. An annual load forecasting model based on support vector regression with differential evolution algorithm. Appl Energy. 2012;94:65-70.
16. Schmidhuber J. Deep learning in neural networks: An overview. Neural Net. 2015;61:85-117.
17. GravesA and SchmidhuberJ. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Net. 2005;18(5–6):602-610.
18. ChenK, ZhouY, and DaiF. A LSTM-based method for stock returns prediction: A case study of China stock market. Big Data (Big Data), 2015 IEEE International Conference on, 2015, IEEE, 2823-2824.
19. Gensler H, Henze BS, Raabe N. Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks'. Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on, 2016, IEEE, 002858-002865.
20. AmeyawB and LiY. Analyzing the impact of GDP on CO2 emissions and forecasting Africa's total CO2 emissions with non-assumption driven bidirectional long short-term memory. Sustainability. 2018;10(9):3110.
21. Sheng Y, Shi X. Energy market integration and equitable growth across countries. Appl Energy. 2013;104:319-325.
22. Qu Y. Evaluation of smart city development potential in major cities of China based on the KPCA method, 2016 International Conference on Management Science and Engineering (ICMSE), 2016, IEEE, 855–862.
23. Camarero M, Forte A, Garcia-Donato G, Mendoza Y, Ordoñez J. Variable selection in the analysis of energy consumption–growth nexus. Energy Econ. 2015;52:207-216.
24. Kubokawa T, Srivastava MS. Selection of variables in multivariate regression models for large dimensions. Communicat Stat-Theory Methods. 2012;41(13–14):2465-2489.
25. Wang P, Wu W, Zhu B, Wei Y. Examining the impact factors of energy-related CO2 emissions using the STIRPAT model in Guangdong Province, China. Appl Energy. 2013;106:65-71.
26. Deng S, Zhang J, Shen F, Guo H, Li Y-W, Xiao H. The relationship between industry structure, household-number and energy consumption in China. Energy Sources Part B. 2014;9(4):325-333.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Li Y. Prediction of energy consumption: Variable regression or time series? A case in China. Energy Sci Eng. 2019;7:2510–2518. https://doi.org/10.1002/ese3.439