Analysis of the impact of epidemic situation on total electricity consumption of the United States under the prediction scenario

Minglu Ma and Qiang Wang¹
School of Economics and Management, China University of Petroleum (East China), Qingdao 266580, People’s Republic China

¹E-mail: wangqiang7@upc.edu.cn

Abstract. The outbreak of COVID-19 epidemic has led to a decline in electricity consumption in the United States, the world's second largest consumer of electricity. Based on the monthly data of the United States from August 2015 to July 2020, this paper uses the ARIMA model and the ARIMA-BP model to forecast the power consumption of the United States in the next 17 months. The Mean Absolute Percentage Error (MAPE) is 5.6% and 4.5% respectively. It shows that the prediction results have high reliability. The results of prediction research will provide a scientific basis for the normal adjustment of electricity supply and demand in the United States, and the method used in this study can be used as a reference for the study of electricity consumption in other countries.

Nomenclature/Abbreviations

| Abbreviation | Description |
|--------------|-------------|
| EIA          | U.S. Energy Information Administration |
| ARIMA        | Autoregressive Integrated Moving Average |
| BP           | Backpropagation Artificial Neural Network Model |
| ARIMA-BP     | Autoregressive Integrated Moving Average- Backpropagation Artificial Neural Network Model |
| ADF          | Augmented Dickey-Fuller |
| MAPE         | Mean absolute percent error |

1. Introduction

The outbreak of COVID-19 epidemic endangers the health and safety all over the world. The adoption of a series of blockade measures has changed the social production mode and indirectly impacted the global energy industry. According to Short-term Energy Outlook released by EIA in August, the power consumption in the United States in 2020 decreased by 3.4% compared with that in 2019 [1]. This figure is based on the statistics of inter-annual comparison but cannot reflect the actual impact of the COVID-19 epidemic on the consumption of electrical energy in the United States. Therefore, it is necessary to calculate the actual impact on the energy consumption in the United States in the context of the epidemic by measuring the energy consumption mode under ideal scenarios.

Based on this research idea, this study adopted the nonlinear BP model to modify the linear ARIMA model and constituted the ARIMA-BP model, and uses the monthly data from August 2015 to December 2019 to forecast the monthly electricity consumption of the United States in the next 17
months. The monthly data obtained is regarded as total electricity consumption of the United States without the impact of COVID-19 epidemic. By comparing it with the actual electricity consumption in the United States in 2020, the impact of COVID-19 epidemic on the electricity consumption in the United States can be calculated. The research results not only have a clearer understanding of the energy development trend of the United States during the COVID-19 epidemic period, but also play a certain reference role for countries around the world to adjust their economic policies during the outbreak of COVID-19 epidemic.

The organizational structure of this paper is as follows: Part 2 is a review of the literature on social electricity consumption and forecasting methods; Part 3 is a detailed introduction of the combination method used in this paper; Part 4 is the presentation and discussion of the research results; Part 5 is a summary of the full paper.

2. Literature review
Total electricity consumption, as a weather vane for observing economic changes, has been widely concerned by experts and scholars from all over the world. After the COVID-19 epidemic, scholars' research on energy industry mainly focused on energy price, energy security and renewable energy development [2-4]. For example, the impact of climate change on energy development has been confirmed [5, 6]. Numerous analyses have pointed out the mechanism of action of meteorological parameters on ambient temperature [7, 8]. Menzi et al. [9] employed asymmetric multifractal detrended fluctuation analysis (a-fm-dfa) to analyze oil and gold prices and found that the efficiency of oil and gold markets was very sensitive during COVID-19 epidemic period. Adekoya [10] found that gold can effectively hedge the market risks related to the global oil and stock markets during the outbreak of COVID-19 epidemic, and the hedging potential will be greater when the oil price and stock price still remain higher. Jefferson [11] predicted that the crude oil price would rise before the third quarter of 2020 and would rise slightly to 40-50 USD per barrel in the first half of 2021. However, Wang et al. [12] made a detailed study on energy security, and found that to improve China's energy security coefficient, attention should be paid to energy supply and renewable energy technologies. Chen et al. [13] conducted studies on China’s oil import system and gave three suggestions to prevent external shocks. Hosseini [14] believed that although the COVID-19 epidemic had a great impact on the energy industry, renewable energy would not shrink significantly in the short term. In addition, Norouzi [15] and Edomah [16] respectively analyzed the electrical power of China and Nigeria.

By sorting out the corresponding literature, we found that, after the outbreak of COVID-19 epidemic, scholars paid more attention to the change of energy price and its impact evaluation on the safety index of oil and methods for improving the energy security, or the study on the development and change of renewable energy after the COVID-19 epidemic. However, there are few researches on energy forecast after the COVID-19 epidemic, especially on total electricity consumption of the United States. In addition, existing research uses inter-annual comparisons when assessing changes in electricity consumption under the epidemic. This measurement method is not scientific enough, because it is assumed that the electricity consumption in 2020 without an epidemic is the same as in 2019. In summary, this research will improve the above two aspects and then carry out research.

The innovation points are as follows. First, this study designed a new research framework for measuring the impact of the epidemic on the electricity consumption in the United States. In this framework, we simulated the U.S. electricity consumption in 2020 when there was no epidemic based on the historical change trajectory. This is also called the data under the ideal scenario. By comparing
the value under the ideal scenario with the value under the actual scenario, the difference between the two is regarded as the decrease in electricity consumption in the United States caused by the epidemic. Second, when simulating the electricity consumption under the condition of no epidemic situation, this study adopted the principle of "error correction + secondary modeling" to construct an ARIMA-BP combined model. The monthly data from August 2015 to December 2019 was used in the modeling process of the ARIMA-BP model. Through calculation, the monthly data from January 2020 to December 2021 are predicted. This part of the forecast data is also regarded as the electricity consumption in the state of no epidemic situation. Comparing this with the actual electricity consumption, the difference is regarded as the decrease in electricity consumption caused by the epidemic.

On the one hand, we can accurately understand the total electricity consumption of the United States, which can be used as a weather vane to observe the economic changes in the coming year and provide data support for the adjustment of energy and economic policies in the United States and other countries. On the other hand, it provides a reference method for other countries to study electrical power and energy after the outbreak of COVID-19 epidemic.

3. Methodology

3.1. Forecast principle of ARIMA model
ARIMA model, with the full name of Autoregressive Integrated Moving Average model, also known as Integrated Moving Average model, is one of the time series forecast analysis methods [21]. The evolved model can accurately fit and measure non-stationary time series. Specifically, firstly, the non-stationary time series can be smoothed by the difference tool, and secondly, the regression parameters of the model are determined by judging the correlation coefficient. ARIMA model has the advantages of simple structure, fast modeling speed and high forecast accuracy in calculation, and is widely used in different industries because it can solve the problem of non-stationary time series.

ARIMA (p, d, q) model is divided into three steps in the forecast process:

Step 1: the process of stabilization. Using the difference tool, the non-stationary time series is changed into stationary time series [22]. After d-order difference, the stationary time series can be expressed as:

\[ X_t^* = (1 - B)^d X_t \]

Step 2: Autoregressive (AR) process. The interpreted variables are regarded as regression functions of previous data and current data [23]:

\[ X_t^* = a_0 + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_KX_{t-K} + \varepsilon_t \]

In which, \( a_i \) is the coefficient; \( X_t^* \) is the interpreted variable; \( X_{t-K} \) is preliminary data; \( \varepsilon_t \) is the error term.

Step 3: Moving Average (MA) process. In this process, the interpreted variable is regarded as the regression function of current error and previous error terms [24], and the corresponding mathematical model is as below:

\[ X_t^* = u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \cdots + \beta_q u_{t-q} \]

In which, \( \beta_i \) is the coefficient, \( u_{t-q} \) is error term, \( X_t^* \) is interpreted variable.

By regression of previous data, current data and error terms, we can fit the monthly data from 2015 to 2020. Then through comparison with the original data, we can obtain the average relative error, get the accuracy of prediction, and then apply it to the prediction.

3.2. Forecast principle of BP model
Backpropagation artificial neural network model (BP) is the abbreviation of "Error Backpropagation", which is a common method used to train artificial neural network in combination with optimization methods (such as gradient descent method). Firstly, the gradient of loss function is calculated for all weights in the network, and then the gradient is fed back to the optimization method, and the weights are updated continuously to minimize the loss function. The back propagation artificial neural network
model calculates the gradient of loss function according to the desired output, so it is generally considered as a self-learning method.

Four steps of the operation of BP model are as follows:

Step 1: Preprocessing data. In this step, the model standardizes the training data and test data so that they can be used for the next operation. The commonly used inverse normalization method is an effective tool in such preprocessing.

Step 2: Selecting the number of neurons in hidden layer. Follow the following formula, constantly change the range of "n", and compare the models one by one to find the most accurate value.

\[ n = \sqrt{a + b + c}, \quad n = \sqrt{ab}, \quad n = \log_2 a \]

In the formula, "a" is the number of neurons in the input layer, "n" is the number of neurons in the hidden layer, "b" is the number of neurons in the output layer, and "c" is a value ranging from 1 to 10.

Step 3: Setting parameters. In the running process of the model, we usually need to set standardized parameters in the training. In Matlab software, we set the minimum training error as \(1 \times 10^{-7}\), the training times as 1000 times, and the learning rate as 0.01.

Step 4: Model forecasting and testing. In this step, the error is returned to the setting process of the model, and the model is repeatedly tested by correcting the error, so as to obtain the optimal model through calculation.

3.3. Forecast principle of ARIMA-BP model

On the basis of the above two single models, we used the combination principle of "error correction + secondary modeling" to conduct combined modeling for ARIMA and BP and thereby created ARIMA-BP model. Specifically, ARIMA-BP model is based on ARIMA model, and the initial prediction results are obtained through the prediction of the data. After that, the initial prediction results are subtracted from the original data to obtain the initial error sequence. Then, the BP neural network model is used to carry out secondary modeling and prediction for the initial error sequence, so as to achieve the effect of error correction. Finally, the initial prediction result is combined with the corrected error value, and then the final unique prediction value can be thereby obtained. The specific combination process is as shown in Figure 1.

3.4. Augmented Dickey-Fuller unit root test

ADF test, also known as Augmented Dickey-Fuller test, is mainly used to detect the stationarity of sequence. If the sequence is stationary, there is no unit root, otherwise there will be. The null hypothesis of ADF is that the sequence has unit root, that is, it is not stationary. For a stationary time-series data, it is required to be significant at a given confidence level, so the null hypothesis is rejected. In the calculation process, we need to obtain the ADF statistics of the sequence through mathematical analysis. If the obtained statistics are significantly less than the critical statistics of the three confidence levels (1%, 5%, 10%), it is indicated that the null hypothesis is rejected.
4. Empirical analysis

4.1. Data sources
In this paper, the monthly electricity consumption in the United States from August 2015 to December 2019 is selected as the basic prediction data. Firstly, ARIMA model is used to carry out modeling and prediction for this group of data, and the initial prediction value is calculated thereby. After that, the initial prediction value is subtracted from the real value to get the initial error. Then, BP model is employed to conduct secondary modeling for the initial error, and the corrected error is thus obtained. Finally, the initial prediction value is combined with the corrected error value, so that the fitted value of previous data and the prediction value of future data based on ARIMA-BP model could be obtained.

4.2. Initial prediction based on ARIMA model
The prediction of ARIMA model is divided into the following steps:

Step 1: Check the unit root of the original data. In this test, the prediction data is stabilized by means of a difference tool. The difference order corresponds to the value of d in ARIMA (p, d, q) model parameters. Table 1 shows the results of unit root test. As shown in table 1, the value of t is smaller than the value within the Third-Order standard deviation, and the value of P is 0 < 0.1. This proves that the data has passed the unit root test, which shows that d=0.

| Augmented Dickey-Fuller test statistic | t-Statistic | Prob.* |
|----------------------------------------|------------|--------|
| Test critical values                   |            |        |
| 1% level                               | -4.148465  | 0.0000 |
| 5% level                               | -3.500495  |        |
| 10% level                              | -3.179617  |        |

Step 2: After d=0 is determined, the values of p and q can be obtained through the autocorrelation coefficient diagram. Figure 2 is a diagram of correlation coefficient obtained after 0-order difference. It can be seen from this diagram that the partial autocorrelation coefficient graph tends to be stable after the second-order difference, while the autocorrelation coefficient diagram tends to be stationary after the first-order difference, so it can be concluded that p=2 and q=1.

![Figure 2. Correlation coefficient diagram of stationary sequence.](image-url)
After ARIMA (2, 0, 1) is determined, SPSS software is used for the next operation of this model, and the parameter values obtained from the operation are as shown in Table 2. It can be seen from Table 2 that the stationary R value is 0.528, which meets the goodness test standard.

**Table 2. Model statistics.**

| Model          | Number of prediction variables | Model fitting statistics | Number of Outliers |
|----------------|--------------------------------|--------------------------|--------------------|
| ARIMA (2,0,1)  | 1                              | 0.528                    | 0                  |

Finally, the fitting results from August 2015 to July 2020 calculated by ARIMA (2,0,1) are as shown in Figure 3, from which we can clearly see the similarity and error between the fitted values and the actual values.

![Figure 3. ARIMA model fitting effect diagram.](image)

Figure 3 clearly shows the changing trend between the original data and ARIMA prediction data. The fluctuation gap of electricity consumption in the first 60 months remains small, and the overall changing direction tends to be consistent. In addition, the error calculated by the fitting part is only 5.64%, and the small error indicates that ARIMA (2,0,1) is suitable for this study.

4.3. **Prediction of ARIMA-BP model**

Based on the initial results and initial errors of ARIMA model, we use BP model to further correct the initial errors. The comparison between errors before and after correction is as shown in the following Figure 4. As can be seen from Figure 4, the corrected error is flatter than the initial error. It is proved that BP model has obvious correction effect.

Combining the corrected error with the initial prediction result, the final prediction value based on ARIMA-BP model can be calculated. The results are as shown in Figure 5.
As shown in the figure, by comparing the original data of the total electricity consumption of the United States from 2015 to 2019 with the ARIMA-BP prediction data, the trend of the two groups of data presents has still higher similarity.
By using the index of average relative error \( \text{MAPE} = \frac{|Y_t^* - Y_t|}{Y_t} \), the accuracy of fitting is calculated (as shown in Figure 6). In the formula, \( Y_t^* \) is the predicted value, \( Y_t \) is the actual value and \( N \) the number of samples. The accuracy of ARIMA model and ARIMA-BP model is 94.4% and 95.4%, respectively. Therefore, the prediction accuracy of the combined models in this study will be still higher, and the results obtained have higher credibility.

![Figure 6. Prediction accuracy of ARIMA model and ARIMA-BP model.](image)

According to the statistics, from February 2020 to April 2020, the total electricity consumption in the United States decreased significantly, but it began to grow slowly from May 2020. The ARIMA-BP model predicts that total US electricity use will gradually recover to pre-February 2020 levels in 2021 and even exceed the average of the previous 54 months.

5. Conclusions
The impact of the epidemic on the energy sector, especially the power industry, is significant. When calculating the impact on electricity consumption, scholars have come to the results based on inter-year comparisons. This ignores the law of power changes in the epidemic year. In order to more accurately calculate the decline in power consumption due to the epidemic, this study set up a scenario where no epidemic occurred, and simulated the power consumption in this scenario using the ARIMA-BP model. The errors calculated by fitting the data of the first 60 months with these two models are 5.6% and 4.5% respectively. The accuracy of the prediction results of these two models is over 94%, indicating that the reliability of the prediction results is very high.
By analyzing the difference between the ideal scenario and the actual scenario in the United States, several conclusions have been drawn as follows.

(1) Due to the impact of the COVID-19 epidemic, the electricity consumption in February-April 2020 decreased significantly. However, since the full resumption of work in the United States in May, the electricity consumption of the whole society began to increase gradually.

(2) In the next 17 months, although the total electricity consumption in the United States will rise and fall slightly. The economic recovery is the main reason for this. In addition, a cold winter, the collapse of the Texas system and the government's new energy policies are other factors contributing to the rapid increase in electricity consumption.

(3) The total electricity consumption will gradually return to the state before the COVID-19 outbreak. In 2021, the total electricity consumption in the United States will increase by 0.85% compared with that in 2020, which is basically consistent with the prediction of EIA that the total electricity consumption in the United States will increase by 0.9% in 2021.

The prediction results of this paper also reflect that although the economy of the United States is affected by the COVID-19 epidemic in 2020, the US economy will gradually return to the state before the outbreak of COVID-19 epidemic in the coming year. A rapid rebound in energy consumption on the back of an economic recovery would threaten the balance between supply and demand in the energy market. Only by anticipating this situation in advance can energy economic policies and measures be adjusted in advance to ensure energy security. The U.S. phenomenon also provides a warning to other regions that a steady economic recovery and improved energy efficiency will prevent energy shortages in the later stages of the pandemic.

References
[1] Short-Term Energy Outlook. 2020
[2] Kannan N and D Vakeesan 2016 Solar energy for future world: - A review Renewable and Sustainable Energy Reviews 62 1092-1105
[3] Agbo E P, et al. 2021 Solar energy: A panacea for the electricity generation crisis in Nigeria Heliyon 7(5) e07016
[4] Da Da B and O Ec 2017 Estimating Daily Solar Radiation from Monthly Values Over Selected Nigeria Stations for Solar Energy Utilization Journal of Fundamentals of Renewable Energy & Applications 7(6)
[5] Ajayi G O and L B Kolawole 1984 Centimeter and millimeter wave attenuation by atmospheric gases and rainfall at a tropical station International Journal of Infrared and Millimeter Waves 5(7) 919-935
[6] Agbo E P, E B Ettah and E E Eno 2021 The impacts of meteorological parameters on the seasonal, monthly, and annual variation of radio refractivity Indian Journal of Physics 95(2) 195-207
[7] Agbo E P, C M Ekpo and C O Edet 2021 Analysis of the effects of meteorological parameters on radio refractivity, equivalent potential temperature and field strength via Mann-Kendall test Theoretical and Applied Climatology
[8] Agbo E P and C M Ekpo 2021 Trend Analysis of the Variations of Ambient Temperature Using Mann- Kendall Test
[9] Me Ns I W, Sensoy A, Xuan V V, et al. 2020 Impact of COVID-19 outbreak on asymmetric multifractality of gold and oil prices[J]. Resources Policy 101829
[10] Adekoya O B, Oliyide J and Oduyemi G 2020 How COVID-19 upturns the hedging potentials of gold against oil and stock markets risks: Nonlinear evidences through threshold regression and markov-regime switching models[J]. Resources Policy 70(3)
[11] Jefferson M 2020 A crude future? COVID-19s challenges for oil demand, supply and prices[J]. Energy Research & Social Science 68 101669
[12] Wang D, Tian S, Fang L, et al. 2020 A functional index model for dynamically evaluating China's energy security[J]. Energy Policy 147
[13] S Chen, M Zhang, Y T Ding, et al. 2020 Resilience of China's oil import system under external shocks: A system dynamics simulation analysis[J]. Energy Policy 146
[14] Hosseini S E 2020 An Outlook on the Global Development of Renewable and Sustainable Energy at the Time of COVID-19[J]. Energy Research & Social Science
[15] Norouzi M R, et al. 2014 Short-term environmental/economic hydrothermal scheduling Electric Power Systems Research 116 117-127
[16] N Edomah and Ndulue G 2020 Energy transition in a lockdown: An analysis of the impact of COVID-19 on changes in electricity demand in Lagos Nigeria[J]. Global Transitions 2 127-137
[17] Guefano S, Tamba J G, Monkam L, et al. 2020 Forecast for the Cameroon's Residential Electricity Demand Based on the Multilinear Regression Model[J]. Energy and Power Engineering 12(5) 182-192
[18] Munkhammar J, Meer D and J Widén 2021 Very short-term load forecasting of residential electricity consumption using the Markov-chain mixture distribution (MCM) model[J]. Applied Energy 282
[19] Fan G F, X Wei, Li Y T, et al. 2020 Forecasting electricity consumption using a novel hybrid model[J]. Sustainable Cities and Society 61 102320
[20] Kaytez F 2020 A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption[J]. Energy 197
[21] Wang Q, Li S, Li R, et al. 2018 Forecasting U.S. shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model[J]. Energy 160 378-387
[22] Shuyu Li, et al. 2019 India's dependence on foreign oil will exceed 90% around 2025 - The forecasting results based on two hybridized NMGM-ARIMA and NMGM-BP models[J]. Journal of Cleaner Production
[23] Wang Qiang, S Li, and R Li 2018 China's Dependency on Foreign Oil Will Exceed 80% by 2030: Developing a Novel NMGM-ARIMA to Forecast China's Foreign Oil Dependence from Two Dimensions Energy 163 151-167
[24] Wang Q, S Li and R Li 2019 Will Trump's coal revival plan work? - Comparison of results based on the optimal combined forecasting technique and an extended IPAT forecasting technique Energy 169 762-775