Prediction of China's Total Energy Consumption Based on Bayesian ARIMA-Nonlinear Regression Model

Zhang Kuna, Feng Weibingb

School of Science, Xi'an University of Science and Technology, Xi'an, 710600, China
aemail: E_ZhangKun@163.com, bemail: f_w001@163.com, f_w001@xust.edu.cn

Abstract: Energy management is the key to China's economy prosperity and environment protection in the future. This paper makes predictions on whether China's energy 13th Five-Year plan can be achieved and the trend of energy consumption in the next stage. Aiming at the shortcoming of insufficient extraction of nonlinear features in the energy system from existing energy consumption models, and low prediction accuracy due to single models. Based on the derivation of the joint posterior distribution density function of the stationary ARMA model parameters under the Normal-gamma prior distribution, this paper proposes a Bayesian ARIMA-nonlinear regression energy consumption prediction model. Bayesian parameter estimation makes full use of prior information and Markov Chain Monte Carlo (MCMC) technique gets rid of the tedious calculation of high-dimensional integrals. Compared with the existing model, the improved model has higher prediction accuracy. Its Root Mean Square Error and Mean Absolute Percentage Error are 3452.0294 and 0.4727%. In line with the development goals set out in the 13th Five-Year plan, an empirical study of the fluctuation law of China's total energy consumption was carried out using data on China's energy consumption from 1990 to 2019 and its factors. The results show that affected by the Corona Virus Disease 2019 epidemic, the growth rate of China's energy demand will fluctuate slightly in 2020. By the end of the 13th Five-Year plan, the total energy consumption can reach the expected goal. After the epidemic is under control, energy demand for the whole society will usher in a rebound growth. In the next five years (2020~2024), China's energy consumption will continue to maintain a steady upward trend of slowing growth. In view of these situations, this paper suggests that, on the one hand, we should attach equal importance to energy conservation and technology development to improve energy efficiency; on the other hand, we should adjust industrial structure and establish a unified energy management system to promote sustainable development of the energy industry.

1. Introduction
Energy is an important cornerstone of human survival, economic growth and social progress. Statistical Review of World Energy-2019[1] pointed out that the global primary energy consumption increased by 2.9% in 2018, which is close to twice the average growth rate (1.5%) of the past ten years. Statistical Review of World Energy -2020[2] pointed out that with the slowing of economic growth and the partial relaxation of some one-off factors that drove energy demand in 2018, the growth of the energy market in the United States, Russia and India has slowed down significantly. But China is an exception. In 2019, China's energy consumption has accelerated and accounted for more than
three-quarters of the global net growth. It has dominated the expansion of the global energy market and has become the largest energy driving force so far. According to the research of the International Energy Agency, there will be a significant change in energy supply and demand in the future. China is the world's largest energy consumer. The rapid growth of energy consumption has led to increasingly tight energy supply. Energy management is the key to future economic prosperity and environmental security.

The energy strategic planning of each country is the top priority of the government's work, and it is specific to China's recent energy strategy. Since the release of the 13th Five-Year energy plan in 2016, China has completed more than half of the 13th Five-Year period. Whether the energy planning goals proposed in the 13th Five-Year plan can be completed by the end of 2020 is a short-term issue that needs to be studied at this stage, and the trend of China's energy consumption in the next stage is a medium and long-term issue that needs to be studied at this stage. Accurately assessing the development trend of China's energy consumption in the future can provide a reference for relevant government management departments to grasp future energy consumption fluctuation trends and formulate long-term energy strategies. At the same time, it is beneficial to the sustainable development of the national economy. Therefore, it is very important to construct an accurate energy consumption prediction model.

Energy consumption is a hot topic in academic research. Many scholars at home and abroad have conducted research on related influencing factors. J L Harris and L M Liu[3] studied the dynamic structural relationship between electric energy consumption and several potential related variables based on the ARIMA model and transfer function model. B C O’Neill and M Desai[4] evaluated the accuracy of the US energy consumption forecast based on the IFFS model and the NEMS model of the Energy Information Administration, and found that the forecast result based on the GDP index was too high, and the forecast result based on the energy intensity index was too low. M Tunc, Ü Camdali, and C Parmaksizoğlu[5] used linear regression analysis to predict Turkey's electricity consumption rate from 2010 to 2020, and finally constructed a linear mathematical optimization model to predict the future distribution of electricity supply investment in Turkey. J Geng and S C Lu[6] predicted China's energy consumption from 2011 to 2013 by constructing a multiple linear model of energy consumption. A Al-Ghandoor, I Al-Hinti, J O Jaber and S A Sawalha[7] constructed an empirical regression model based on multivariate variables to predict the electrical energy demand of Jordan's industrial sector. N M Xie and S F Liu[8] established a gray GM(1,1) model based on energy production and consumption, and predicted the total energy consumption of Jiangsu Province in China. However, the scope of application of a single model has certain limitations. It is often unable to fully extract the useful information from the data, and the research method needs to be excessive to the fusion of multiple information. P Zhou, B W Ang and K L Poh[9] combined the traditional gray GM (1,1) model with the triangular residual correction technique to analyze China's electricity demand. Y S Lee and L I Tong[10] combined residual modification with genetic programming symbol estimation, and proposed a GPGM(1,1) combined prediction model, whose prediction accuracy and stability are better than a single gray model.

However, there are still some problems in China's energy consumption forecasting: first, the model is single, and most existing studies are based on regression models, time series models, gray models and other single models to predict energy consumption; second, the nonlinear characteristics of the energy consumption system are not sufficiently extracted. The energy consumption system has both linear and nonlinear characteristics, and most models are only limited to the linear relationship in the energy consumption system; third, the prediction accuracy and robustness are not high. Energy consumption is a complex system that is affected by multiple factors and external environment. The prediction results obtained by using only a single indicator of energy consumption value are highly susceptible to epidemic diseases and other factors, so it is necessary to integrate the influences of multiple factors such as GDP, share of secondary industry and urbanization rate.

This paper focuses on whether the energy 13th Five-Year plan can be achieved and the prediction of energy consumption trends in the next stage, takes the total energy consumption from 1990 to 2019
as the research object, establishes the nonlinear regression model, ARIMA model and Bayesian ARIMA model to enrich the energy consumption prediction system. Due to the many factors that restrict energy consumption, there are both linear and nonlinear constraints[11-12]. Firstly, the time series curve is fitted. Considering that energy consumption is not an independent system, a non-linear regression model is constructed by integrating various economic and social factors to extract the non-linear characteristics of the energy consumption system. Secondly, ARIMA model is constructed to extract the linear relationship between current consumption and past consumption of energy, and Bayesian inference method is adopted to optimize the parameter estimation of ARIMA model, so as to make up for the inadequacy of single time series model in information utilization. Finally, ARIMA model and nonlinear regression model are combined optimally, and the Bayesian ARIMA-nonlinear regression model is proposed to integrate the advantages of the two models. Comparing the prediction accuracy of the models shows that the Bayesian ARIMA-non-linear combination model has a better prediction effect than a single model, and the model is used to predict China's total energy consumption in the next five years.

2. Methods and prediction framework

2.1. Bayesian inference of ARIMA(p,d,q) model

Suppose that the random variable sequence \( \{X_t\} \) satisfies the ARIMA(p,d,q) model, and the stationary sequence \( \{Y_t\} \) is obtained after d-order difference, which satisfies the ARMA(p,q) model:

\[
Y_t = \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \epsilon_t - \phi_1 \epsilon_{t-1} - \cdots - \phi_q \epsilon_{t-q}, t = 1, 2, \ldots, n
\]

\[
\theta_p \neq 0, \phi_q \neq 0
\]

\[
E(\epsilon_t) = 0, Var(\epsilon_t) = \sigma_\epsilon^2, E(\epsilon_t \epsilon_s) = 0, s \neq t
\]

\[
E(Y_t \epsilon_t) = 0, \forall s < t
\]

ARIMA(p,d,q) is essentially a combination of difference and ARMA(p,q). Any non-stationary time series can be analyzed and studied by ARMA(p,q) after proper difference reaches a stable state[13].

Given the initial value \( Y_{t-1} = y_{t-1}, Y_{t-2} = y_{t-2}, \ldots, Y_{t-p} = y_{t-p} \), then:

\[
(Y_t | Y_{t-1}, \ldots, Y_{t-p}; \theta, \phi, \sigma_\epsilon) \sim N(\theta_1 y_{t-1} + \cdots + \theta_p y_{t-p}, (1+\phi_1^2+\cdots+\phi_q^2)\sigma_\epsilon^2)
\]

The density function of the conditional distribution is:

\[
f_{Y_t | Y_{t-1}, \ldots, Y_{t-p}; \theta, \phi, \sigma_\epsilon} = \frac{1}{\sqrt{2\pi\sigma_\epsilon}} \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \left[ y_t - \left( \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} \right) \right]^2 \right\}
\]

The conditional likelihood function of the model is:

\[
L(\theta, \phi, \sigma_\epsilon) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma_\epsilon}} \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \left[ y_t - \left( \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} \right) \right]^2 \right\}
\]

\[
= \left( \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \right)^n \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \left[ \sum_{t=1}^{n} y_t^2 + \sum_{t=1}^{n} \left( \sum_{i=1}^{p} \theta_i y_{t-i} \right)^2 - 2\sum_{t=1}^{n} \sum_{i=1}^{p} \theta_i y_t y_{t-i} \right] \right\}
\]

In order to facilitate the calculation, the unknown parameters are transformed:

\[
\tau = \frac{1}{\sigma_\epsilon^2}
\]
The conditional likelihood function of the model is transformed into:

$$L(\theta, \phi, \tau) = \left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp \left\{-\frac{\tau}{2} \left[ \sum_{i=1}^{n} y_i^2 + \sum_{i=1}^{n} \left( \sum_{i=1}^{n} \theta_i y_{t-i} \right)^2 - 2\sum_{i=1}^{n} \sum_{i=1}^{n} \theta_i y_i y_{t-i} \right] \right\}$$

(6)

The conditional likelihood function of the model has the form of a Normal-gamma distribution density function kernel, so the prior distribution of model parameters selects the classic Normal-gamma distribution family. That is, given $\tau$, the prior distribution of $\theta$ and $\phi$ is the Normal distribution, and the prior distribution of $\tau$ is the Gamma distribution. Then the joint prior distribution density function of the model parameters is:

$$\pi(\theta, \phi, \tau) = \pi(\theta | \tau) \pi(\phi | \tau) \pi(\tau)$$

(7)

The joint posterior density of model parameters is:

$$\pi(\theta, \phi, \tau | y) = \frac{L(\theta, \phi, \tau) \pi(\theta, \phi, \tau)}{\int_{\tau \in (0, +\infty)} \int_{\theta \in \mathbb{R}} \int_{\phi \in \mathbb{R}} L(\theta, \phi, \tau) \pi(\theta, \phi, \tau) d\theta d\phi d\tau}$$

$$= \frac{\left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp \left\{-\frac{\tau}{2} \left[ \sum_{i=1}^{n} y_i^2 + \sum_{i=1}^{n} \left( \sum_{i=1}^{n} \theta_i y_{t-i} \right)^2 - 2\sum_{i=1}^{n} \sum_{i=1}^{n} \theta_i y_i y_{t-i} \right] \right\} \pi(\theta, \phi, \tau)}{\int_{\tau \in (0, +\infty)} \int_{\theta \in \mathbb{R}} \int_{\phi \in \mathbb{R}} \left(\frac{\tau}{2\pi}\right)^{\frac{n}{2}} \exp \left\{-\frac{\tau}{2} \left[ \sum_{i=1}^{n} y_i^2 + \sum_{i=1}^{n} \left( \sum_{i=1}^{n} \theta_i y_{t-i} \right)^2 - 2\sum_{i=1}^{n} \sum_{i=1}^{n} \theta_i y_i y_{t-i} \right] \right\} \pi(\theta, \phi, \tau) d\theta d\phi d\tau}$$

(8)

From $\pi(\theta, \phi, \tau | y)$, integrate $\tau$ on $(0, +\infty)$, and integrate $\phi$ and $\theta$ on $\mathbb{R}$ respectively. Get the edge posterior density of the autoregressive coefficient $\theta$ and moving average coefficient $\phi$ of the ARMA(p,q) model. In the same way, integrate $\theta$ on $\mathbb{R}$ and integrate $\phi$ on $\mathbb{R}$ to obtain the edge posterior density of parameter $\tau$.

2.2. MCMC and Gibbs sampling

The key of Bayesian estimation lies in the calculation of posterior distribution. However, in practice, the difficulty of posterior distribution in high-dimensional computation leads to the limitation of Bayesian inference in practical application. With the rapid development of software technology, MCMC is proposed to solve the problem of high dimensional integration involved in Bayesian analysis.

MCMC is a computer stochastic simulation method that effectively handles complex high-dimensional integrals[14]. The basic idea is: Through a probability transfer matrix, a Markov chain with a stationary distribution of $\pi(\theta)$ is obtained, and $\pi(\theta)$ is the posterior distribution of the parameter to be estimated. The posterior distribution samples are generated through this Markov chain, and Monte Carlo integration is performed based on the extracted samples.

WinBUGS is a professional software developed by the British Imperial College and the MRC (Medical Research Council) Biostatistics Association to use MCMC algorithm for Bayesian inference[15]. The basic principle is Gibbs sampling and Metropolis-Hastings algorithm. Gibbs sampling was first proposed by Geman S and Geman D when discussing image restoration in 1984. It is a special case of Metropolis-Hastings algorithm. It can be applied to situations where the target distribution is multi-dimensional, and it is one of the most widely used methods in the MCMC sampling algorithm[16]. The basic idea is:

- Give initial vector of parameters $\theta = (\theta_1^{(0)}, \theta_2^{(0)}, \cdots, \theta_p^{(0)})$;
- Perform the following iterative updates
Draw sample $\theta_1^{(1)}$ from distribution $\pi(\theta_1 | \theta_2^{(0)}, \theta_3^{(0)}, \cdots, \theta_p^{(0)})$;

Draw sample $\theta_2^{(1)}$ from distribution $\pi(\theta_2 | \theta_1^{(1)}, \theta_3^{(0)}, \cdots, \theta_p^{(0)})$;

......

Draw sample $\theta_p^{(1)}$ from distribution $\pi(\theta_p | \theta_1^{(1)}, \theta_2^{(1)}, \cdots, \theta_{p-1}^{(1)})$;

- After $n$ iterations, the posterior sample $\theta_1, \theta_2, \cdots, \theta_p$ is obtained, and a Markov chain is formed.

Where $\theta^i = (\theta_1^{(i)}, \theta_2^{(i)}, \cdots, \theta_p^{(i)})$, $i = 1, 2, \cdots, n$;

- According to the posterior sample to calculate each order moment, and then perform statistical inference.

2.3. Bayesian ARIMA-nonlinear regression prediction model

The energy consumption system is a complex system with linear and nonlinear characteristics, but it is difficult to fully extract the useful information of the system from a single model. In this paper, a Bayesian ARIMA-nonlinear regression prediction model is proposed, the combined model integrates the advantages of different models, Bayesian parameter estimation makes full use of prior information, and MCMC technology gets rid of the cumbersome high-dimensional integral calculation, so as to improve the accuracy of energy consumption prediction. The basic idea of the model is as follows: firstly, a nonlinear regression model is established by integrating various influencing factors of energy consumption to extract the nonlinear characteristics of the energy consumption system. In addition, as energy consumption is affected by the trend of short-term energy consumption in the past, ARIMA model is adopted to extract the linear relationship between current consumption and past consumption of energy, and Bayesian method is adopted to improve the estimation accuracy of unknown parameters of the model. Finally, the weights are determined according to the model fitting accuracy, and the predictive values of the two models are weighted to determine the predictive values of the Bayesian ARIMA-nonlinear regression model. The detailed process is as follows (Figure 1).

Set $\hat{y}_i$ as the energy consumption forecast value of Bayesian ARIMA-nonlinear regression model.
in \( r \)th year, \( \hat{y}_1 \) as the energy consumption forecast value of Bayesian ARIMA model in \( r \)th year, \( \hat{y}_2 \) as the energy consumption forecast value of nonlinear regression model in \( r \)th year, the formula for Bayesian ARIMA-nonlinear regression model is:

\[
\hat{y}_r = a\hat{y}_1 + b\hat{y}_2,
\]

\[
\begin{align*}
    a &= \frac{\alpha}{\alpha + \beta} \\
    \beta &= \frac{\beta}{\alpha + \beta}
\end{align*}
\]

Where, \( a + b = 1 \) and \( a, b \in [0,1] \), \( \alpha \) is the Mean Absolute Percentage Error of the nonlinear regression model \( (MAPE) \), and \( \beta \) is the Mean Absolute Percentage Error of the Bayesian ARIMA model.

3. Prediction of China's total energy consumption based on the combination model

3.1. Data source

Energy is the material basis of social production and an important guarantee for stable economic development. Whether the future development of the energy industry can maintain the sustainable development of the economy is a key issue of concern to the whole society. China is the first energy supplier and consumer in the world. In the industrial field, China's energy consumption has accounted for more than 1/3 of the world's energy consumption. China plays a pivotal role in global energy consumption, so this paper selects China's total energy consumption (10,000 tons of standard coal) from 1990 to 2019 as the research object. Since the energy consumption system is a complex system affected by multiple factors and external environments, this paper comprehensively considers the impact of GDP (RMB 100 million), the share of the secondary industry (%), and the urbanization rate (%) on energy consumption. All data are from the National Bureau of Statistics of the People’s Republic of China China Statistical Yearbook-2019[17].

Table 1  Data on China's total energy consumption and influencing factors from 1990 to 2019.

| Year | Total energy consumption | gross domestic product (GDP) | Share of the secondary industry | Urbanization rate |
|------|--------------------------|------------------------------|---------------------------------|-------------------|
| 1990 | 98703.0                  | 18872.9                      | 41.0329                         | 26.4097           |
| 1991 | 103783.0                 | 22005.6                      | 41.4876                         | 26.9402           |
| 1992 | 109170.0                 | 27194.5                      | 43.1153                         | 27.4599           |
| 1993 | 115993.0                 | 35673.2                      | 46.1767                         | 27.9901           |
| 1994 | 122737.0                 | 48637.5                      | 46.1629                         | 28.5098           |
| 1995 | 131176.0                 | 61399.9                      | 46.7505                         | 29.0404           |
| 1996 | 135192.0                 | 71813.6                      | 47.1043                         | 30.4799           |
| 1997 | 135909.0                 | 79715.0                      | 47.0990                         | 31.9100           |
| 1998 | 136184.0                 | 85195.5                      | 45.7976                         | 33.3502           |
| 1999 | 140569.0                 | 90564.4                      | 45.3599                         | 34.7797           |
| 2000 | 146964.0                 | 100280.1                     | 45.5362                         | 36.2198           |
| 2001 | 155547.0                 | 110863.1                     | 44.7934                         | 37.6597           |
| 2002 | 169577.0                 | 121717.4                     | 44.4506                         | 39.0898           |
| 2003 | 197083.0                 | 137422.0                     | 45.6228                         | 40.5302           |
| 2004 | 230281.0                 | 161840.2                     | 45.9002                         | 41.7600           |
| 2005 | 261369.0                 | 187318.9                     | 47.0226                         | 42.9900           |
2006 286467.0 219438.5 47.5574 44.3430  
2007 311442.0 270092.3 46.8842 45.8892  
2008 320611.0 319244.6 46.9712 46.9895  
2009 336126.0 348517.7 45.9571 48.3417  
2010 360648.0 412119.3 46.4978 49.9497  
2011 387043.0 487940.2 46.5293 51.2703  
2012 402138.0 538580.0 45.4230 52.5701  
2013 416913.0 592963.2 44.1767 53.7296  
2014 425806.0 643563.1 43.0856 54.7701  
2015 429905.0 688858.2 40.8413 56.0999  
2016 435819.0 746395.1 39.5806 57.3497  
2017 448529.1 832035.9 39.8517 58.5197  
2018 464000.0 919281.1 39.6870 59.5802  
2019 486000.0 990865.1 38.9725 60.6000  

Data source: National Bureau of Statistics of the People’s Republic of China

3.2. Nonlinear regression model

From the energy consumption scatter diagram (Figure 2), it can be seen that China's total energy consumption from 1990 to 2019 did not have obvious cyclical changes, and showed a non-linear growth trajectory as a whole. In the first 10 years, China's energy consumption has been steadily rising. Since 2000, China's energy consumption has started to rise rapidly with economic development. Affected by the international financial crisis, the rate of increase in China's energy consumption has slowed from 2008 to 2009. By 2014, there was a slight fluctuation again, and then it began to accelerate gradually.

With the help of SPSS software, six time series curve regression models were tried for the total energy consumption sequence of China from 1990 to 2019:

\[
\begin{align*}
\text{Logarithm: } y_t &= a + b \log(t), \quad t = 1, 2, \ldots, n \\
\text{Quadratic: } y_t &= a + bt + ct^2, \quad t = 1, 2, \ldots, n \\
\text{Cubic: } y_t &= a + bt + ct^2 + dt^3, \quad t = 1, 2, \ldots, n \\
\text{Power: } y_t &= at^b, \quad t = 1, 2, \ldots, n \\
\text{Index: } y_t &= ae^{bt}, \quad t = 1, 2, \ldots, n \\
\text{Logistic: } y_t &= \frac{1}{a + be^{-ct}}, \quad t = 1, 2, \ldots, n
\end{align*}
\]

(11)

Where \( y_t \) represents China's total energy consumption in the \( t \)th year, that is: when \( t = 1 \), China's total energy consumption in 1990 was \( y_1 \) million tons of standard coal.
Comparing the six time series regression curves, the cubic regression curve is in the highest agreement with the energy consumption series. The fitting results of the time series regression model output by SPSS software are as follows (Table 2):

| Model  | R2     | F      | df1 | df2 | Significance |
|--------|--------|--------|-----|-----|--------------|
| Logarithm | 0.707  | 67.660 | 1   | 28  | 0.000        |
| Quadratic | 0.970  | 441.434| 2   | 27  | 0.000        |
| Cubic   | 0.987  | 841.443| 3   | 26  | 0.000        |
| Power   | 0.806  | 116.211| 1   | 28  | 0.000        |
| Index   | 0.966  | 790.340| 1   | 28  | 0.000        |
| Logistic| 0.966  | 790.340| 1   | 28  | 0.000        |

Comparing the fitting results of six time series regression models, all model parameters have passed the significance test at a significance level of 0.05. Among them, the coefficient of determination of the cubic regression curve ($R^2 = 0.987$) is the largest, which means that 98.7% of China's total energy consumption can be explained by this nonlinear relationship. The model has the highest goodness of fit. Therefore, the preliminary prediction model of China's total energy consumption is as follows:

$$y_i = a + bt + ct^2 + dt^3, t = 1, 2, \cdots, n$$ \hspace{1cm} (12)$$

Where $y_i$ represents China's total energy consumption in $t^{th}$ year.

Energy consumption system is a complex system closely related to the level of economic development, industrial structure and population size. Fluctuations in energy consumption growth due to economic and social factors cannot be ignored. The time series regression model using a single indicator data is difficult to accurately predict energy consumption, so the nonlinear regression model is improved as follows:

$$y_i = a + bt + ct^2 + dt^3 + e_{xt} + f_{xt} + g_{xt}, t = 1, 2, \cdots, n$$ \hspace{1cm} (13)$$

Where $x_{it}$ indicates China's gross domestic product (billion yuan) in $t^{th}$ year, $x_{2it}$ indicates the share of China's secondary industry(%) in $t^{th}$ year, and $x_{3it}$ indicates China's urbanization rate (%).
in the year.

Table 3 Estimated results of nonlinear regression models.

| parameter | Estimate  | Standard error | 95% confidence interval |
|-----------|-----------|----------------|-------------------------|
| a         | -68204.409 | 145541.321     | -369279.570 - 232870.752 |
| b         | -17392.658 | 2742.818       | -23066.610 - -11718.706 |
| c         | 2579.107   | 290.482        | 1978.199 - 3180.015     |
| d         | -40.566    | 5.539          | -52.025 - -29.108       |
| e         | -0.083     | 0.079          | -0.246 - 0.081          |
| f         | 9284.917   | 1150.812       | 6904.281 - 11665.554    |
| g         | -7218.271  | 4483.722       | -16493.557 - 2057.015  |

Compared with the single-time series regression model, the nonlinear regression model fitting advantage value, which combines the three influence factors of GDP (RMB 100 million), secondary industry share (%) and urbanization rate (%) is increased to 0.999, which indicates that the variation of 99.9% of China's total energy consumption can be explained by this nonlinear regression relationship. Using SPSS software to estimate model parameters, the final improvement of the nonlinear regression model is:

\[ y_t = -68204.4 - 17392.7t + 2579.107t^2 - 40.566t^3 - 0.083x_{tb} + 9284.917x_{ts} - 7218.271x_{tg}, t = 1, 2, \ldots, n \]  (14)

3.3. Bayesian ARIMA model based on MCMC

3.3.1. Determination of the ARIMA (2,1,0) model.

First use the Box-Jenkin method to get the appropriate ARIMA model. Because sequence \( \{X_t\} \) is not stable, it can't be modeled directly with the ARIMA model, and it needs to be smoothed. In order to weaken the correlation between energy consumption sequences and ensure the trend stability of modeling sequences, the difference is made on sequence \( \{X_t\} \):

\[ Y_t = X_t - X_{t-1}, t = 1, 2, \ldots, n \]  (15)

Where \( X_t \) represents total energy consumption in the \( t \)th year, \( Y_t \) represents the difference between energy consumption in the \( t \)th year and the \( (t-1) \)th year.

Unit root testing of first-order differential sequence \( \{Y_t\} \) with the Eviews software:

Table 4 The results of the ADF test.

| t-Statistic | Prob  |
|-------------|-------|
| Augmented Dickey-Fuller test statistic | -6.4963 | 0.0000 |
| Test critical values | |
| 1% level | -3.7529 |
| 5% level | -2.9981 |
| 10% level | -2.6388 |

For the first-order differential sequence \( \{Y_t\} \), the t-test statistical value is -6.4963, and at the three significance levels of 1%, 5% and 10%, the existence of unit root is accepted with zero probability, indicating that the first-order differential sequence \( \{Y_t\} \) has good stability.
Figure 3  Autocorrelogram and partial autocorrelation diagram.

Observation the autocorrelogram (Figure 3.a), which shows obvious periodic sine fluctuations, and the continuous gradient of the process of attenuation to zero, the self-correlation coefficient is judged to be dragging its tail. Observe the partial autocorrelation diagram (Figure 3.b), the partial autocorrelation coefficient of the delayed first order is outside the twice standard deviation range and falls within the twice standard deviation range after the first order. And the process of reducing the positive and negative phases to the vicinity of the zero value to do small value fluctuations is very sudden. So the partial autocorrelation coefficient is judged to be a first order truncation.

According to the basic principle of ARIMA model ordering[18], an appropriate model is preliminarily formulated as ARIMA(1,1,0):

\[ Y_t = \theta_1 Y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \]  

Where \( \varepsilon_t \) is the residual, \( \sigma^2 \) is the variance of the residual.

In order to avoid the model error caused by subjective identification inaccuracies, the ARIMA (2,1,0) model is obtained by using R software for automatic ordering of the system:

\[ Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \]  

Where \( \varepsilon_t \) is the residual, \( \sigma^2 \) is the variance of the residual.

Table 5  Estimated results of the ARIMA (2,1,0) model.

| Coefficients(\( \theta_1 \)) | Coefficients(\( \theta_2 \)) | sigma^2 (\( \sigma^2 \)) | AIC        | Log likelihood |
|-----------------------------|-----------------------------|---------------------------|------------|----------------|
| 1.0635                      | -0.3325                     | 31591593                  | 589.16     | -290.58        |

The AIC value of the ARIMA (2,1,0) model is calculated to be 589.16, which is less than the AIC value of the ARIMA (1,1,0) model 591.79. Through continuous attempts, the ARIMA (2,1,0) model is finally established as the optimal model, and the estimate of unknown parameters is obtained by maximum likelihood estimation (Table 5):

\[ Y_t = 1.0635Y_{t-1} - 0.3325Y_{t-2} + \varepsilon_t, \varepsilon_t \sim N(0, 3.159 \times 10^7) \]  

3.3.2. MCMC simplifies high-dimensional integrals in bayesian statistical inferences

Then the MCMC[19] technique was used to simplify the problem of high-dimensional integrals in Bayesian statistical inferences:
\[
\begin{align*}
\pi(\theta_1 | y) &= \int_{\tau > 0} \pi(\theta_1, \theta_2, \tau | y) d\theta_2 d\tau \\
\pi(\theta_2 | y) &= \int_{\tau > 0} \pi(\theta_1, \theta_2, \tau | y) d\theta_2 d\tau
\end{align*}
\] (19)

From \(\pi(\theta_1, \theta_2, \tau | y)\), integrate \(\tau\) on \((0, +\infty)\), and integrate \(\theta_1\) and \(\theta_2\) on \(R\) respectively. The edge posterior density of the autoregressive coefficients \(\theta_1\) and \(\theta_2\) of the ARIMA (2,1,0) model is obtained.

\[
\pi(\tau | y) = \int_{\theta_1, \theta_2} \pi(\theta_1, \theta_2, \tau | y) d\theta_1 d\theta_2
\] (20)

Integrate \(\theta_1\) on \(R\) and integrate \(\theta_2\) on \(R\) to obtain the edge posterior density of parameter \(\tau\).

The priori distribution of unknown parameters is based on the Normal-gamma distribution:

\[
\begin{align*}
\theta_1 &\sim N(0,1,0.00001) \\
\theta_2 &\sim N(0,1,0.00001) \\
\tau &\sim \text{Gamma}(0.1,0.001)
\end{align*}
\] (21)

In order to ensure the convergence of unknown parameters, 1000st Gibbs pre-iterations were carried out first[20]. Excluding the results of the first 1000st initial iterations, the output results of WinBUGS[21] from 1001st to 20000st iterations are as follow:

| node  | tau (\(\tau\)) | Theta1(\(\theta_1\)) | Theta2(\(\theta_2\)) |
|-------|---------------|----------------------|----------------------|
| mean  | 2.597\times10^{-8} | 1.197 | -0.2578 |
| sd    | 7.553\times10^{-9} | 0.2117 | 0.2152 |
| MC error | 6.054\times10^{-11} | 0.001537 | 0.001602 |
| 2.50% | 1.338\times10^{-8} | 0.781 | -0.6855 |
| median | 2.518\times10^{-8} | 1.197 | -0.2586 |
| 97.50% | 4.285\times10^{-8} | 1.616 | 0.1693 |
| start | 1001          | 1001 | 1001 |
| sample | 20000         | 20000 | 20000 |

Table 6 shows the Bayesian estimation results of unknown parameters of Bayesian ARIMA (2,1,0) model. The Bayesian estimation of \(\tau\) is \(2.597\times10^{-8}\), and the confidence interval of 95% confidence level is \((1.338\times10^{-8}, 4.285\times10^{-8})\); The Bayesian estimation of \(\theta_1\) is 1.197, and the confidence interval of 95% confidence level is \((0.7810, 1.6160); \) The Bayesian estimate of \(\theta_2\) is -0.2578, and the confidence interval of 95% confidence level is \((-0.6855, 0.1693)\).

The Bayesian ARIMA(2,1,0) model is obtained as follows:

\[
Y_t = 1.197Y_{t-1} - 0.2578Y_{t-2} + \varepsilon_t, \varepsilon_t \sim N(0, 2.597\times10^{-8})
\] (22)

### 4. Prediction results of Bayesian ARIMA-nonlinear regression model

The nonlinear regression model, the ARIMA(2,1,0) model and the ARIMA(2,1,0) model are respectively calculated to predict the total energy consumption of China in the recent five years (2015-2019) (Table 7):

| Year | Nonlinear regression model | ARIMA(2,1,0) model | Bayesian ARIMA(2,1,0) model |
|------|---------------------------|--------------------|---------------------------|

Unit: 10,000 tons of standard coal
The accuracy of the above three models in the prediction of China's total energy consumption is measured by two indexes, Root Mean Square Error ($RMSE$) and Mean Absolute Percentage Error ($MAPE$):

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

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

(23)

Where, $n$ represents the predicted sample size, $y_i$ represents the actual value of China's total energy consumption in the $t$th year, and $\hat{y}_i$ represents the predicted value of China's total energy consumption in the $t$th year.

Comparing the prediction errors of the three models mentioned above (Table 8), the Root Mean Square Error and Mean Absolute Percentage Error of ARIMA(2,1,0) model are the largest, followed by the nonlinear regression model, and the Bayesian ARIMA(2,1,0) model is the smallest. Compared with the ARIMA(2,1,0) model, Root Mean Square Error and Mean Absolute Percentage Error of Bayesian ARIMA(2,1,0) model are smaller, indicating that the Bayesian method optimizes the unknown parameter estimation of the model and is helpful to improve the accuracy of the energy consumption prediction model. Therefore, the weights $a$ and $b$ are determined according to Mean Absolute Percentage Error of the Bayesian ARIMA (2,1,0) model and the nonlinear regression model with small prediction error.

Considering the timeliness of prediction, $a = 0.489037, b = 0.510963$ is obtained according to the prediction error calculation in the past five years. The prediction values of the combined model are determined by weighting the predicted values of Bayesian ARIMA (2,1,0) model and nonlinear regression model, and the prediction error of Bayesian ARIMA-nonlinear regression model is calculated (Table 8).

| Error | Nonlinear regression model | ARIMA model | Bayesian ARIMA model | Bayesian ARIMA-nonlinear regression model |
|-------|----------------------------|-------------|----------------------|------------------------------------------|
| RMSE  | 5233.0337                  | 6196.9567   | 4731.8515            | 3452.0294                                |
| MAPE  | 0.9737%                    | 1.1460%     | 0.9319%              | 0.4727%                                  |

The Root Mean Square Error and Mean Absolute Percentage Error of Bayesian ARIMA-nonlinear regression model are respectively 3452.0294 and 0.4727%, which are the lowest values among the four models. It shows that the combined model can improve the prediction accuracy of the two single models, and the prediction accuracy of the China's energy consumption prediction is higher. In view of the complexity of energy consumption system, this model integrates various influencing factors of energy consumption to extract nonlinear characteristics, and gives full play to the advantages of ARIMA model for the high accuracy of linear time series prediction and Bayesian method for the high accuracy of unknown parameters estimation. It is a model that can improve the accuracy of China's energy consumption prediction.

Bayesian ARIMA (2,1,0) - nonlinear regression model was used to predict China's energy consumption in the next five years (2020–2024). Firstly, the ARIMA model is used to predict China's
GDP, share of the secondary industry and urbanization rate in the next five years, and the final energy consumption forecast value is obtained by weighting the Bayesian ARIMA (2,1,0) prediction results and nonlinear regression prediction results. The results are as follows (Table 9):

Table 9  Prediction results of Bayesian ARIMA-nonlinear regression model.

| Year   | Predictive value |
|--------|------------------|
| 2020   | 497616.0404      |
| 2021   | 511300.7172      |
| 2022   | 521194.8299      |
| 2023   | 529811.8371      |
| 2024   | 535910.1913      |

For the convenience of observation, combined with the actual value of China's total energy consumption from 1990 to 2019, the predicted results of nonlinear regression model, ARIMA model and Bayesian ARIMA-nonlinear regression model are plotted as follows (Figure 4):

Figure 4  The comparison of the model prediction curve with the actual curve.

Energy consumption is closely related to many factors such as economic development, national policies and social emergencies. Affected by the continuous impact of the international financial crisis, China's total energy consumption growth slowed down significantly from 2008 to 2009. With the introduction of subsequent economic stimulus policies, energy demand has increased. After a period of low-speed growth, China's energy consumption gradually rebounded from 2017. Affected by the Corona Virus Disease 2019 epidemic, it is expected that the growth rate of China's energy demand will fluctuate slightly during the epidemic control stage. After the epidemic is effectively controlled, the energy demand of the whole society will see a rebound growth. Based on the Bayesian ARIMA-nonlinear regression prediction model, it is estimated that China's total energy consumption in 2020 will reach 497,616,040,000 tons of standard coal. Referring to the total energy consumption target of China in 2020 formulated in the 13th Five-Year plan, it is within 500,000 tons of standard coal, which has reached the binding target, but is very close to the constraint conditions. In recent years, China has paid more and more attention to the control of total energy consumption, which has been reasonably controlled through the adjustment of industrial structure, rational planning and improvement of energy efficiency. It is expected that China's energy consumption in the next five years (2020~2024) will maintain an upward trend of slow growth and reach 5359.101913 tons of standard coal in 2024. This is a new challenge for the further control of energy consumption, and the development of the energy industry will be under great pressure. The next 20 years will be a key period of global energy industry adjustment, but also an opportunity for China's energy transformation. According to the prediction results of Bayesian ARIMA-nonlinear regression model, China's total
energy consumption will still maintain the growth momentum in the future, which requires the government, the energy industry association and relevant management departments should optimize the long-term energy supply strategy to adapt to the development of the new situation.

- Insist on energy conservation and development technology. Strengthen the awareness of energy conservation and emission reduction in various industries, and popularize advanced technologies for energy conservation and high efficiency. Because China's coal reserves are abundant and indispensable to all industries, improving the level of coal conversion and using efficiency technology is conducive to enterprises in reducing energy consumption.
- Diversified energy development. With green and low-carbon development as the main direction, vigorously research and development to water, electricity, wind and solar energy and other renewable new energy. Widespread use of biomass energy to promote the development of non-fossil energy.
- Adjust the industrial structure. Reduce the proportion of high energy consumption industries, vigorously develop low-carbon industries, and establish a sustainable economic development mode.
- Establish a unified energy management system. Establish energy management information system platform, coordinate the development and utilization of various energy resources, and implement macro energy management mode. Implement various energy policies and regulations to provide policy support for the orderly development of energy market.

5. Conclusions

Use the ARIMA model to extract the linear characteristics of energy consumption in the current period and the past, and integrate China’s GDP, the share of the secondary industry and the urbanization rate to fit a nonlinear regression model to extract the nonlinear characteristics of energy consumption. The predicted value of the two models is weighted according to Mean Absolute Percentage Error to determine the predicted value of the Bayesian ARIMA-nonlinear regression combined model. The results show that the Root Mean Square Error and Mean Absolute Percentage Error values of the Bayesian ARIMA-nonlinear regression model are lower than those of the other two single models. It is an effective and high-precision energy consumption prediction model, which has certain reference significance for the energy management department to accurately grasp the fluctuation trend of China's future energy consumption and formulate effective measures in time. Finally, the Bayesian ARIMA-nonlinear regression model was used to predict China's total energy consumption in the next five years. It is estimated that China's total energy consumption will reach 4,976,160,404 tons of standard coal in 2020. In the next five years, there will be an upward trend of slowing growth, and it is expected to reach 5,351,101,913 tons of standard coal in 2024.

The Bayesian inference method makes comprehensive use of the total information, sample information and prior information, and effectively integrates the data and historical information. It solves the problem of large error in the case of insufficient samples and poor quality of traditional inference methods, and it is an effective tool for parameter estimation of time series models. Bayesian inference improves the quality of the prediction model and makes the improved ARIMA model more suitable for solving practical problems. Among them, MCMC technology and WinBUGS software simplify the high-dimensional probability function integral of the posterior quantity in the process of Bayesian inference and promote the bayesian ARIMA model to be widely used in practice.

Combined with the prediction results of Bayesian ARIMA-nonlinear regression model and the actual energy consumption, China's energy consumption status is basically controlled within the target set in the 13th Five-Year energy plan, but is very close to the constraint conditions. The medium and long-term forecast results also put forward higher requirements for the development of China's energy consumption, and it is expected that the development of the energy industry will maintain the momentum of growth in the future. In view of the current situation of energy consumption in China, some suggestions are put forward, such as paying equal attention to developing technology and saving energy, developing energy diversification, adjusting industrial structure and establishing unified energy management system, so as to promote the sustainable development of energy industry.
Acknowledgments
This work was financially supported by the National Natural Science Foundation: Research on Theory and Application of Coal Resources Multi-stage Investment Management under Uncertain Environment (71473194).

References
[1]BP Statistical Review of World Energy 2019. (2019) China Statistics Press., Beijing.
[2]BP Statistical Review of World Energy 2020. (2020) China Statistics Press., Beijing.
[3]Harris, J.L., Liu, L.M. (1993). Dynamic Structural Analysis and Forecasting of Residential Electricity Consumption. International Journal of Forecasting., 9(4):437-455.
[4]O’Neill, B.C., Desai, M. (2005). The Accuracy of Past Projections of U.S. Energy Consumption. Energy Policy., 33(8):979-993.
[5]Tunc, M., Çamdali, Ü., Parmaksizoğlu, C. (2006). Comparison of Turkey’s Electrical Energy Consumption and Production with Some European Countries and Optimization of Future Electrical Power Supply Investments in Turkey. Energy Policy., 34(1):50-59.
[6]Geng, J., Lu, S.C. (2014). Research on my country's Energy Consumption Forecast Based on Multiple Regression Analysis. Resources Development and Market., 30(11):1317-1319.
[7]Al-Ghandoor, A., Al-Hinti, I., Jaber, J.O., Sawalha, S.A. (2008). Electricity Consumption and Associated GHG Emissions of the Jordanian Industrial Sector: Empirical Analysis and Future Projection. Energy Policy., 36(1):258-267.
[8]Xie, N.M., Liu, S.F. (2009). Forecast analysis of total energy consumption and structure: Taking Jiangsu Province as an example. Industrial Technology & Economy., 28(1):46-49.
[9]Zhou, P., Ang, B.W., Poh, K.L. (2006). A Trigonometric Grey Prediction Approach to Forecasting Electricity Demand. Energy., 31(14):2839-2847.
[10]Lee, Y.S., Tong, I.I. (2011). Forecasting Energy Consumption Using A Grey Model Improved by Incorporating Genetic Programming. Energy Conversion and Management., 52(1):147-152.
[11]Gui, M., Liu, L.H. (2016). Based on the ARIMA model to predict the total retail sales of consumer goods in Hainan Province. Mathematics in Practice and Knowledge., 47(3):25-30.
[12]Ye, T. (2017) Stock forecasting method based on wavelet analysis and ARIMA-SVR model. International Conference on Information Management. pp. 102-106.
[13]Wang, Y. (2015) Time series analysis: based on R. Renmin University of China Press., Beijing.
[14]Zhou, J. (2008). MCMC is applied to Bayesian estimation of parameters. MA thesis., Hubei University.
[15]Zhou, L.L., Ding, D.Y. (2011). Bayesian hierarchical credit risk assessment model based on MCMC simulation. Statistics and Information Forum., 26(12): 26-31.
[16]Zhu, X.L. (2009). Review of Markov Chain Monte Carlo Method Research. Statistics and Decision., (21): 151-153.
[17]National Bureau of Statistics of the People’s Republic of China. China Statistical Yearbook 2019. (2019) China Statistics Press., Beijing.
[18]Xie, X.J., Qiu, Y.L., Shi, L. (2019). Energy consumption forecast based on the combined model of ARIMA and BP neural network. Mathematics in Practice and Knowledge., 49(10):292-298.
[19]Mao, S.S., Tang, Y.C. (2019) Bayesian Statistics. 2rd ed. China Statistics Press., Beijing.
[20]Zhang, Z.C. (2008). MCMC method and application. MA thesis., Hubei University.
[21]Zheng, J.C., Zhu, H.M. (2005). Bayesian AR(P) model analysis based on MCMC method. Statistics and Decision., (20): 4-6.