

Research Article

A Driving Force Analysis and Forecast for Gas Consumption Demand in China

Qing Zhu,^{1,2} Quan-Ying Lu,² Xiao-Yang Zhou,² and Kin Keung Lai^{2,3}

¹ School of Finance and Economics, Xi'an JiaoTong University, Xi'an 710061, China

² International Business School, Shaanxi Normal University, Xi'an 710062, China

³ Department of Management Sciences, City University of Hong Kong, Hong Kong

Correspondence should be addressed to Quan-Ying Lu; luquanying0705@163.com

Received 28 December 2013; Revised 15 March 2014; Accepted 10 April 2014; Published 7 May 2014

Academic Editor: Jianping Li

Copyright © 2014 Qing Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, gas has begun to be widely used in power generation as well as in manufacturing because of the environmental advantages, lower cost, and the relative safety and reliability. Accurate prediction of future gas consumption is of great importance for energy security. In this paper, we first use path analysis to select the core factors that impact gas consumption, then use the RBF-QRNN model and BVAR model separately to predict future gas consumption, and finally discuss and compare the results from the two models. The results show that if there is a positive growth in GDP, urbanization, and population, then gas consumption is expected to increase over a comparatively long time. In detail, gas consumption is expected to remain at about 25% of all energy consumption, GDP's contribution to gas consumption is expected to be about 53%, and the urbanization rate contribution would be about 15%. Gas consumption in China is predicted to reach around 178649.23 million cubic meters by the end of 2015 and about 264698.86 million cubic meters by 2020.

1. Introduction

In recent years, gas has begun to be widely used in power generation as well as in manufacturing because of the environmental advantages, lower cost, and the relative safety and reliability. The BP World Energy Statistics Yearbook (2013) reported that in 2012 the proportion of oil in total global energy consumption declined as oil grew only by 0.9% while the growth rate for gas was 2.2%, accounting for 23.9% of all basic energy consumption. According to "The 2012 domestic and international oil and gas industry development report," in 2012 gas consumption in China was 147.1 billion cubic meters, 13% more than in 2011, with imports making up 42.5 billion cubic meters, an increase of 31.1%, denoting 29% external dependency. The increasing gas consumption in China indicates the need for infrastructure development in terms of gas pipelines construction and maintenance. At the same time, China has specifically announced that the primary task is to address key technological aspects to lay the foundation for

large-scale development of shale gas. Under such a policy background, accurate prediction of future gas consumption is of great importance for pipeline construction planning, production management, and gas distribution scheduling. Further, such forecasts can assist in reducing gas distribution costs and thereby better serve the production and societal needs. With the rapid development and integration of artificial intelligence technology, forecasting methods such as the traditional time series, smoothing methods, regression models, development trend extrapolation methods and correlation analysis, fuzzy sets, and rough sets support vector machines (SVM), and genetic algorithms have been introduced to enhance energy demand forecasting. Modern artificial intelligence technology offers advantages in energy demand prediction, and traditional forecasting methods can provide good economic interpretations. Therefore, in this paper, we employ an RBF neural network quantile regression model and a BVAR model independently, to predict China's gas consumption.

2. Literature Review

Extant research on energy demand forecasting is sizeable, as are the gas consumption analyses and predictions. Because of the need to understand future energy demand, significant research has been conducted in the past and the effort continues. This paper discusses gas consumption from multiple perspectives and at different levels, with the main focus being on the following two aspects.

2.1. Gas Consumption Driving Forces. Many researchers have suggested that gas consumption is primarily connected with economic growth, population size, and economic structure. Das et al. (2013) [1] conducted the Granger causality test to examine gas consumption and the actual GDP in Bangladesh from 1980 to 2010 and discovered that GDP either had a notable effect on increasing gas consumption or had no apparent influence. Apergis and Payne (2010) [2] conducted further research on the connection between gas consumption and GDP and established a multiple error correction model for gas consumption in 67 nations from 1992 to 2005. A heterogeneous panel cointegration test was conducted which led to the conclusion that there was a long-run equilibrium relationship between gas consumption, real GDP, real asset investment, and the labor force, with a bidirectional causality existing between gas consumption and economic growth. The elasticity of gas demand was estimated to be 0.652, a little lower than that of renewable energy (0.76) but much higher than that of coal, which was between minus 0.142 to minus 0.251. These figures illustrate the strong relationship between economic development and gas consumption. Kankal et al. (2011) [3] predicted Turkey's gas consumption by building an artificial neural network and conducting a regression analysis. Taking the social economy and demography as the influential factors, using four test variables (GDP, population, import-export volume, and employment) and using stepwise regression, they came to the conclusion that these four factors impact gas consumption in this region with the final result indicating that gas consumption in Turkey in 2014 was expected to be between 117.0 and 175.4 million tons of oil equivalent. After analyzing the status of gas consumption in Italy, Bianco et al. (2014) [4] maintained that gas price, gross domestic product (GDP) population, and temperature data are the main factors influencing gas consumption and analyzed each factor separately. Based on an integrated analysis of such factors as population, gross national product, and import and export figure, Canyurt and Ozturk (2008) [5] used genetic algorithm demand estimation models (GA-DEM) to estimate the future coal, oil, and natural gas demand values in Turkey. The results suggested that coal, oil, and natural gas consumption values were estimated to increase almost 2.82, 1.73, and 4.83 times between 2000 and 2020.

2.2. Predicting Gas Consumption. Sánchez-Úbeda and Berzosa (2007) [6] presented a statistical decomposition model for daily data of the Spanish industrial gas time series and forecast the short-term demand for gas. Gutiérrez et al. (2005) [7] used annual gas consumption data from Spain from 1973 to 1997 as the sample and used the random

Gompertz innovation diffusion (stochastic Gompertz innovative diffusion) model to predict the total gas consumption for the period 1998–2000. The predictions from this model, in terms of goodness-of-fit of the real values, were significantly more accurate than any other random diffusion growth models such as the logic (diffusion-innovation) models or the lognormal (diffusion-non-innovative) models. Combination forecasting models have also been applied gradually. Khotanzad et al. (2000) [8] proposed a two-stage artificial neural networks method using multilayer feed forward neural networks and functional connectivity neural networks to predict the volumes of six gas utilities in the US. Xu and Wang (2010) [9] combined a 2nd order polynomial curve model with a moving average model to model the PCMACP to forecast China's total demand for gas in 2007–2008. The difference between the predicted and the real value was 3.82%, which was significantly lower than the conventional second-order models, BP neural network models, or gray prediction models. Soldo (2012) [10] in a review of gas forecasting research summarized the status of gas forecasting in terms of gas research and development process, the study area, the predicted angle (hours/days/months/years/combo), data types, and forecasting tools and noted that the Hubert curve model and gray model were the main tools used while an optimized combination of tools and classical regression tools would likely be the main tools for future research in this area. Gorucu (2004) [11] used artificial neural network (ANN) model to forecast gas consumption in the capital city of Ankara, Turkey. Lee and Tong (2011) [12] examined China's energy consumption data from 1990 to 2007 to construct a grey model improved by incorporating genetic programming GPGM (1, 1) using the gray system theory, an uncertainty analysis method. Azadeh et al. (2011) [13] presented an adaptive network-based fuzzy inference system-stochastic frontier analysis (ANFIS-SFA) approach for long-term gas consumption prediction. The results indicated that the ANFIS-SFA approach is capable of dealing with complexity, uncertainty, and randomness as well as several other unique features.

Summarizing the findings of previous studies, most scholars agree that the ARMA method based on historical data and the exponential smoothing method are most suitable for short-term forecasts. The gray prediction method, however, has been found to be more useful for predicting the broader short- and long-term demand but offers lower accuracy. The artificial neural network methods with adaptive, self-organizing, and real-time learning were found to be suited to dynamic forecasting of gas consumption with a higher predictive accuracy. Although a simple neural network algorithm is able to digitize all information under conditions of unconstrained nonlinearity with high accuracy, it is not effective in reflecting the uncertain relationships among economic variables. In order to find new ways to avoid forecast error uncertainty, some scholars have used probabilistic prediction research methods, such as the quantile regression method, combined with a neural network model, giving the probability of a certain confidence level for longer range predictions. Amongst the nonlinear regression forecasting models, neural network models have been shown to have

high accuracy, while quantile regression models are able to highlight local impacts, so combining these two models can yield better prediction results. He et al. (2013) [14] used an RBF neural network and a quantile regression model to build a combination of the two methods and also used time-series data to determine a complete power load probability distribution function for every moment on a given day. This RBF neural network quantile regression (RBF-QRNN) model has obvious advantages in terms of prediction accuracy but lacks the ability to offer economic explanations, unlike the traditional forecast methods.

Traditional forecasting methods, which describe the structural relationships between variables, are based on the economic theory. These methods have more advantages in terms of being able to incorporate the economic theory. However, there are some problems when we establish a model based only on the traditional methods as it is necessary to distinguish the endogenous and exogenous variables and each endogenous variable in the equation is associated with a disturbance, which makes the estimation of the model parameters very complex, resulting in large errors. In a word, it is unable to accurately reflect the dynamic relationship between variables, especially when the selected sample data is limited. Because of this freedom restraint, it cannot effectively predict the model. The VAR model has been shown to be an effective predictive model with a time-series variable interconnected system, so it has been frequently used to analyze the influence of different types of random error terms on the dynamics of the system variables. However, when the number of selected samples is insufficient, the VAR model causes multicollinearity and the degrees of freedom fall, so parameter estimation errors occur. The BVAR model has been shown to remedy the defects in the VAR model as it can sufficiently handle small samples. In recent years, the BVAR method has been widely employed in energy prediction. Crompton and Wu [15] used the Bayesian method of vector autoregression (BVAR) to predict China's energy consumption and discuss the potential impacts.

Since modern artificial intelligence technology has advantages for energy prediction and the traditional forecasting methods have a better ability to determine economic changes; here, we employ an RBF neural network quantile regression model and a BVAR model independently to predict China's gas consumption. First, we use path analysis to select the core factors that affect China's gas consumption. Then, we build an RBF neural network quantile regression model to predict China's gas consumption based on the theory of probability density and then use a standard gradient optimization algorithm to estimate the model's parameter vectors. To forecast gas consumption from 2012 to 2020, we have to allocate gas consumption to different subdigit levels, compare the relative errors under these different levels, and select the lowest error median. In the following, we use a cross-validation method to determine the center, width, and punishment values for the kernel function in the RBF neural network model. Finally, we build a BVAR model with a Minnesota prior (normal-flat prior) distribution for analysis and forecasting.

3. Core Effect Factors

Most previous studies have used multivariate regression analysis but these methods do not eliminate multicollinearity. To solve this problem, scholars have proposed many methods, such as stepwise regression, ridge regression, partial least-squares regression, and path analysis, but none of these has proven to be perfect. Here, we use path analysis to screen out the core factors driving China's gas consumption. In 1921, the population geneticist Sewall Wright proposed the path analysis method, which is an extension of simple correlation analysis. Using multiple regression, a correlation coefficient is resolved and the direct effect of some variables on the dependent variable, as well as the indirect and combined effects of other variables on the dependent variable, can be expressed through a direct path, an indirect path, and a total path coefficient. Chai et al. (2011) [16] used this method to analyze factors influencing oil prices.

From the literature review, we know the factors that affect gas consumption but determining the mutual relationships between the factors is even more complex. Here, we select gas production, population, urbanization rate, GDP, urban per capita consumption, industrial fuel ex-factory price, and the electricity price (X_1, X_2, \dots, X_8) as the factors that impact gas consumption (Y), based on a literature review of papers which examined supply and demand, alternative energies, cost factors, and national policy. We selected annual data from 1990 to 2011 and used path analysis to screen the core factors affecting China's gas consumption (data from Yearbook of China Statistics, WIND database, and CEIC database).

We tested the normality of the dependent variable (y) using the Shapiro-Wilk method before path analysis and then determined the statistic for y to be 0.780, with a skewness coefficient of 0, which implied that y followed a normal distribution. Hence, we could use path analysis to screen out the core driving factors. Three independent variables, total population, urbanization rate, and GDP were chosen. Magnitude of the direct effect (in absolute value) was represented as $X_4 > X_2 > X_3$ and the combined effects (in absolute value of the correlation coefficient) are denominated as $X_4 > X_3 > X_2$. From this we found that GDP had the most significant effect on China's gas consumption, among direct and combined effects. This can be explained by the fact that China is still in a development stage, with economic growth still being led by industrial production growth which leads to higher energy consumption. The increase in population and the urbanization rate, as well as the commensurate transportation and infrastructure expansion, are eventually reflected in GDP growth. The decision coefficient for the effect factors can be calculated as $R(2)^2 = 1.429608$, $R(3)^2 = 1.645803$, and $R(4)^2 = -0.015015$. Therefore, population growth rates and urbanization growth rates are the main push factors for gas consumption with the effect of urbanization growth rate being greater than that of the population growth rate function. GDP is a limiting factor for gas consumption because the relationship between GDP and energy consumption appears to be very complex. Because gas reserves are insufficient in China, gas could be seen to be a high-end

TABLE 1: Results of the normality test.

	Shapiro-Wilk		
	Statistics	df	Sig.
Y	0.780	22	0.000

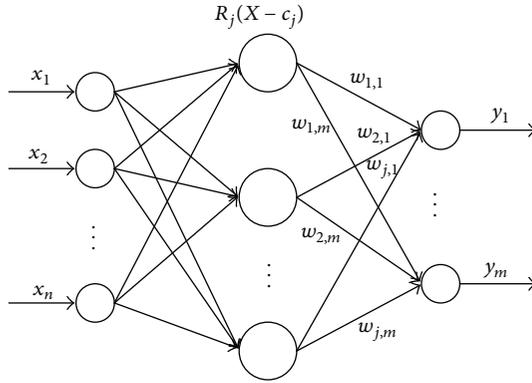


FIGURE 1: Structure of RBF neural network.

energy relative to coal and oil. The determination coefficient $R^2 = \sum_i^3 p_{iy} r_{iy} = 0.998935$ shows that the interpretation capacity of the selected factors for the dependent variable reaches 99.89 percent, which indicates that the path analysis efficiently covered the primary effect factors. Detailed results are shown in Tables 1 and 2.

4. Modeling Approach and Data Analysis

4.1. RBF Neural Network Structure. Faced with a reliance on convergence to the initial value and a local convergence in a traditional BP neural network, Moody and Darken (1989) [17] proposed a three-layer (input layer, hidden layer, and output layer) forward neural RBF network (see Figure 1) where the radial basis function was the hidden network neuron activation function. This offered a faster computation, a stronger nonlinear mapping ability, and a better forecasting performance.

RBF neural network structures are achieved by mapping between two layers, that is, a nonlinear transformation from the input layer to the hidden layer and a linear transformation from the hidden layer to the output layer. To establish the RBF neural network structure, we first need to determine the center functions c_j and variances σ_j of the hidden layer and weights $w_{j,i}$ from the hidden layer to the output layer to complete the mapping from the input to the output.

4.2. The Quantile Regression. In 1978, Koenker and Bassett (1978) [18] proposed a quantile regression theory which characterized the changing trends of Y under different quantiles when the independent variable X is given. It measured the impact of the independent variable in the distribution's center, upper tail, and under tail and highlighted the correlation among the local effects. Assuming Y is affected by

X_1, X_2, \dots, X_k , then we can express the quantile regression as in the following equation:

$$Q_Y(\tau | X) = \beta_0(\tau) + \beta_1(\tau) X_1 + \beta_2(\tau) X_2 + \dots + \beta_k(\tau) X_k \equiv X' \beta(\tau), \quad (1)$$

where $X = [X_1, X_2, \dots, X_k]'$ is a set of explanatory variables, $\tau \in (0, 1)$ is the quantile, and $\beta(\tau) = [\beta_0(\tau), \beta_1(\tau), \beta_2(\tau), \dots, \beta_k(\tau)]'$ is the regression coefficient vector.

4.3. The RBF Neural Network Quantile Regression. Taylor (2000) [19] proposed a single-layer neural network-based quantile regression model:

$$Q_Y(\tau | X) = f[X, W(\tau), V(\tau)], \quad (2)$$

where $W(\tau) = \{w_{jk}(\tau)\}_{j=1,2,3,\dots,j; k=1,2,3,\dots,K}$ is the connection weight vector of the input layer to the hidden layer; $V(\tau) = \{v_j\}_{j=1,2,\dots,J}$ is the connection weight vector between the hidden layer and the output layer; K is the hidden layer node; and f is a nonlinear function. The equation can be expressed as follows:

$$f[X, W(\tau), V(\tau)] = g_2 \left\{ \sum_{j=1}^J v_j(\tau) g_1 \left[\sum_{k=1}^K w_{jk}(\tau) x_{kj} \right] \right\}. \quad (3)$$

This equation shows the explanatory variables X to Y 's response to the variable nonlinear structure. In this paper, as the annual gas consumption data is continuous, the RBF neural network was deemed suitable for fitting the continuously changing data into many types of neural networks.

The parameter vectors $\widehat{W}(\tau), \widehat{V}(\tau)$ of RBF neural network quantile regression model (2) can be estimated as follows:

$$\begin{aligned} \min_{W, V} & \left\{ \sum_{i=1}^N \rho_\tau [Y_i - f(X_i, W, V)] + \lambda_1 \sum_{j,i} w_{ji}^2 + \lambda_2 \sum_i v_i^2 \right\} \\ & = \min_{W, V} \left(\sum_{i|Y_i \geq f(X_i, W, V)} \tau |Y_i - f(X_i, W, V)| \right. \\ & \quad + \sum_{i|Y_i < f(X_i, W, V)} (1 - \tau) |Y_i - f(X_i, W, V)| \\ & \quad \left. + \lambda_1 \sum_{j,i} w_{ji}^2 + \lambda_2 \sum_i v_i^2 \right), \end{aligned} \quad (4)$$

where λ_1, λ_2 are the penalty parameters to avoid an overfit. The optimal value for penalty parameters λ_1, λ_2 and the number of hidden layer's node K can also be determined using cross-validation methods. Parameter vectors $W(\tau), V(\tau)$ can be estimated using a standard gradient optimization algorithm. Therefore, we determined $\widehat{W}(\tau), \widehat{V}(\tau)$ and put them into (2). As a result, the conditional quantile for Y can be estimated as in the following equation:

$$\widehat{Q}_Y(\tau | X) = f(X, \widehat{W}(\tau), \widehat{V}(\tau)). \quad (5)$$

TABLE 2: Path Analysis for effect of the total population, proportion of urban residents and GDP.

Independent variables	Combined action	Direct action	Indirect action	Decision coefficient	t-test
X_2	0.801821	-0.379401	1.181222	1.429608	-9.622777
X_3	0.898468	0.223623	0.674845	1.645803	3.9144691
X_4	0.992009	1.111109	-0.119100	-0.150105	41.083457

TABLE 3: Forecasts of independent variables.

Forecasting Result	Dependent variable Independent variables	Gas consumption (millions of cubic meters)		
		GDP (One hundred million yuan)	Total population (one million)	Urbanization rate (%)
	2012	506221.28	1357.05	52.295400
	2013	541656.77	1366.82	53.341308
	2014	579572.74	1376.66	54.408134
	2015	620142.84	1386.57	55.496297
	2016	663552.83	1396.56	56.606223
	2017	710001.53	1406.61	57.738347
	2018	759701.64	1416.74	58.893114
	2019	812880.75	1426.94	60.070976
	2020	873033.93	1437.22	61.272396

4.4. Gas Forecast Consumption

4.4.1. Data Selection. We selected data for the period 1990 to 2011, with China’s total gas consumption as the output variables, and used path analysis to filter out the three core factors (GDP, total population, and urbanization rate) as the input variables to construct the RBF neural network quantile regression model. This was conducted in two phases. In the first stage, we selected data from 1990 to 2006 as the training samples to determine value of each parameter for the model and then selected data from 2007 to 2011 as the test samples to predict China’s future gas consumption. We compared the forecast values with the actual gas consumption values from 2007 to 2011 and derived the prediction errors. In the second stage, we selected data from 1990 to 2011 as the training samples and used the above parameter values to forecast gas consumption from 2012 to 2020. After training, the results showed that the maximum number of iterations was 1000 and the structure of the neural network was 3-1-1, which indicated that the number of nodes in the input layer, hidden layer, and output layer was 3, 1, and 1, respectively. Both penalty parameters for the RBF neural network quantile regression model (λ_1, λ_2) were 0.00001 and the width of the kernel function (σ) was 10. All data was normalized using

$$X_s = \frac{X - \bar{X}}{S}, \tag{6}$$

where X_s is the sample vector, X_t is the vector of the original sample, \bar{X} is the average of a column vector, and S is the standard deviation of the column vectors.

The scenario analysis for the core factors was carried out for 2012 to 2015. According to China’s “Twelfth Five Year Plan,” the annual population growth rate was to be controlled within 7.2%. Further, China’s gross domestic product (GDP) in 2015 was estimated to be 55.8 trillion yuan and the average

annual growth was predicted to be 7%. From this information we calculated the GDP with a 7% annual average growth. The urbanization rate is expected to reach more than 54% by 2015 having already reached 51.27% in 2011. After the 18th CPC National Congress, the Chinese government was devoted to reducing the gap between the rich and the poor to ensure sustainable economic development. The most important measures for the achievement of these two goals were stated to be urbanization, narrowing the gap between the rich and the poor, infrastructure development, industrial transfer, and a structural adjustment in the urbanization process to encourage investment and employment. Most population experts estimated that, after the 18th CPC National Congress, China’s urbanization would reach more than 60% by 2020. According to these forecasts, annual growth of urbanization is expected to reach about 2%, so here we assumed a growth rate of 2% for China’s urbanization from 2011 to 2020. The results are shown in Table 3.

4.4.2. Determine the τ Quantile. In a RBF neural network quantile regression, there can be differences between the results predicted with different quantile values (Table 4). When the quantile $\tau = 0.1$, the prediction error is 3.02%, which is the minimum prediction error. Figure 2 also shows that the corresponding curve ($\tau = 0.1$) is the closest to the real value, so the most appropriate quantile (τ) for the RBF neural network model is 0.1. By training the model and comparing the gas consumption from 1990 to 2011 with real data, the average prediction error was determined to be only 0.08%, with fitting precision of the model being 99.92%. From these results it is concluded that the RBF neural network quantile regression is superior for a complex nonlinear prediction. The prediction results for China’s gas consumption from 2012 to 2020 using the RBF neural network quantile regression are shown in Table 8.

TABLE 4: Comparison of gas consumption prediction under different quantile levels (million cubic meters).

τ value	2007	2008	2009	2010	2011	Average absolute relative deviation
0.1	70560.15	84601.49	94529.42	112460.70	130227.26	3.02%
0.2	70703.16	84874.13	94891.75	113142.60	131391.15	3.42%
0.3	70797.26	84976.99	94999.72	113301.53	131586.66	3.56%
0.4	70650.43	84185.64	93788.88	110529.23	125615.36	3.13%
0.5	71499.94	87316.19	98612.99	120456.55	144340.02	8.43%
0.6	71734.13	88306.52	100207.47	123738.03	150610.36	10.67%
0.7	71820.29	88731.32	100952.97	125331.19	153760.68	11.75%
0.8	71953.44	89158.33	101625.68	126708.88	156386.20	12.70%
0.9	72502.79	91450.19	105348.59	134785.30	172277.66	18.20%
Real value	70523.14	81293.43	89519.84	106940.00	130530.00	0

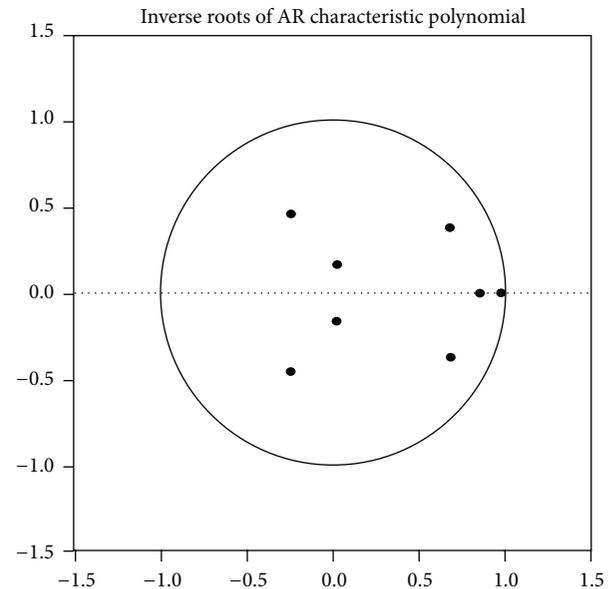
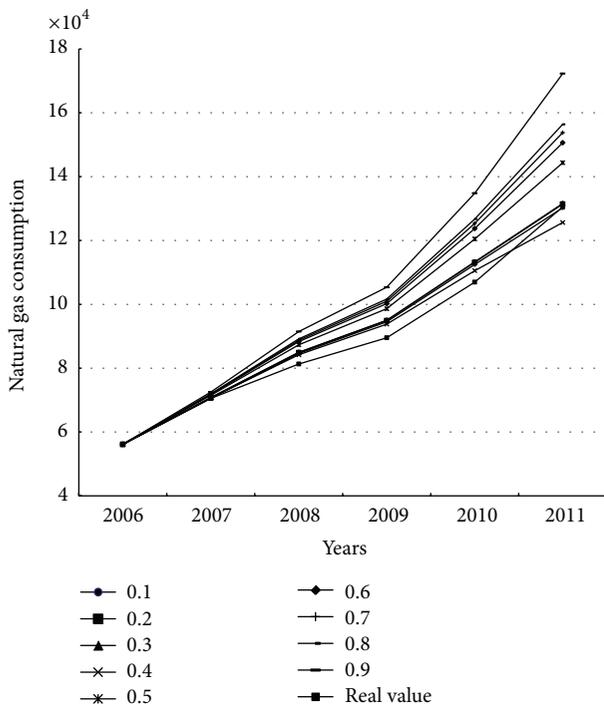


FIGURE 3: Unit root graph.

FIGURE 2: Trend of gas consumption under different quantile levels (million cubic meters).

4.5. BVAR Model

4.5.1. VAR Model Approach. The general VAR model can be expressed as in the following equation:

$$y_t = c + A_1 y_{t-1} + \dots + A_p y_{t-p} + \mu_t, \quad t = 1, 2, \dots, n, \quad (7)$$

where $y_t = (y_{1t} \dots y_{kt})$ is a $k \times 1$ random vector, A_1 to A_p are $m \times m$ coefficient matrixes, c is a m -dimensional vector, μ_t is a white noise sequence, $E(\mu_t) = 0$, $E(\mu_t \mu_t') = \Sigma$, Σ is an $m \times m$ positive definite matrix, and $E(\mu_t \mu_s') = 0$, ($t \neq s$).

Before establishing the VAR model, it is necessary to first test the multivariate cointegration. When using a nonstationary time series, the differential sequence not only can prevent

spurious regression but is also able to maintain the long-term relationships between the variables. In the following, we use the Johansen test, a maximum likelihood method first proposed by Johansen and Juselius, to test the multivariate cointegration. The maximum likelihood estimation for a cointegration system requires that all independent cointegration relationships be analyzed. The number of cointegration relationships is not assumed in advance and the coefficients do not need to be regulated. The unit root test of every variable is required in the Johansen test. After the unit root test, we established a VAR system consisting of different variables and tested the stability of the system. By comparing the VAR model system with the LNY, LNGDP, LNCZ, and LNP (logarithmic gas consumption, GDP, urbanization rate, and population), it was found to be the most stable as it satisfied the two-order integration. The results are shown in Tables 5, 6, and 7 and in Figure 3.

TABLE 5: The variables of unit root test results.

Variable	ADF statistics	5% critical value	Conclusion
LNY	0.048960	-3.040391	Not smooth
LNGDP	0.454627	-3.052169	Not smooth
LNCZ	-0.700623	-3.020686	Not smooth
LNP	-2.564774	-3.052169	Not smooth
D(LNY)	-0.652005	-3.040391	Not smooth
D(LNGDP)	-5.743269	-3.052169	Smooth
D(LNCZ)	-2.076184	-3.020686	Not smooth
D(LNP)	-2.740084	-3.020686	Not smooth
D(LNY, 2)	-6.069975	-3.040391	Not smooth
D(LNGDP, 2)	-3.206952	-3.081002	Smooth
D(LNCZ, 2)	-4.574096	-3.029970	Smooth
D(LNP, 2)	-4.192706	-3.029970	Smooth

TABLE 6: The estimation results of VAR model.

	LNY	LNGDP	LNCZ	LNP
R-squared	0.998013	0.988922	0.999815	0.999990
Adj. R-squared	0.996568	0.998138	0.999681	0.999982
Sum sq. resids	0.017876	0.013288	0.000150	360E - 07
S.E. equation	0.040312	0.034757	0.003699	0.000181
F-statistic	690.7156	1273.812	7442.299	133230.7
Log likelihood	41.82176	44.78723	89.59510	149.9582
Akaike AIC	-3.282176	-3.578723	-8.059510	-14.09582
Schwarz SC	-2.834096	-3.130643	-7.611430	-13.64774
Mean dependent	10.45492	11.72739	3.628654	7.147743
S.D. dependent	0.688150	0.805368	0.207067	0.042830
Determinant resid covariance (dof adj.)			150E - 19	
Determinant resid covariance			1.37E - 20	
Log likelihood			343.8536	
Akaike information criterion			-30.78536	
Schwarz criterion			-28.99304	

TABLE 7: JB test results.

Component	Skewness	Chi-sq	df	Prob.
1	-0.139030	0.097296	1	0.7551
2	-0.271715	0.367644	1	0.5443
3	0.068714	0.023836	1	0.8773
4	0.672789	2.104472	1	0.1469
Joint		2.593247	4	0.6280
Component	Kurtosis	Chi-sq	df	Prob.
1	2.609985	0.245432	1	0.6203
2	3.401941	2.504162	1	0.1135
3	2.281862	0.006698	1	0.9348
4	3.685034	1.250519	1	0.2635
Joint		4.006812	4	0.4051
Component	Jarque-Bera	df	Prob.	
1	0.342728	2	0.8425	
2	2.871806	2	0.2379	
3	0.030534	2	0.9848	
4	3.354991	2	0.1868	
Joint	6.600059	8	0.5803	

TABLE 8: Forecasts of gas consumption (million cubic meters).

	Model	RBF-QRNN	BVAR Prior = 1	Combined forecast	The year-on-year growth rate
Forecasting results for gas consumption demand	2012	138996.8	141110.99	140053.89	/
	2013	151325.0	152619.47	151972.24	8.81%
	2014	164503.6	165139.69	164821.65	8.46%
	2015	178532.1	178766.36	178649.23	8.39%
	2016	193414.4	193604.36	193509.38	8.32%
	2017	209136.1	209768.12	209452.11	8.24%
	2018	225677.4	227384.34	226530.87	8.15%
	2019	243009.0	246591.89	244800.45	8.06%
	2020	261853.0	267544.72	264698.86	8.13%
	AAGR	8.24%	8.33%	8.28%	/

The established VAR model can be expressed as in the following equations:

$$\begin{aligned} \text{LNY} = & 0.9555\text{LNY}(-1) - 0.0140\text{LNY}(-2) \\ & + 0.0642\text{LNGDP}(-1) - 0.0108\text{LNGDP}(-2) \\ & + 0.6191\text{LNCZ}(-1) - 0.1634\text{LNCZ}(-2) \\ & - 41.6128\text{LNP}(-1) + 38.3431\text{LNP}(-2) + 22.0978, \end{aligned}$$

$$\begin{aligned} \text{LNGDP} = & 0.4688\text{LNY}(-1) + 0.4316\text{LNY}(-2) \\ & + 0.7172\text{LNGDP}(-1) - 0.7475\text{LNGDP}(-2) \\ & - 1.5692\text{LNCZ}(-1) - 2.4301\text{LNCZ}(-2) \\ & + 14.4091\text{LNP}(-1) + 10.3005\text{LNP}(-2) \\ & - 159.3559, \end{aligned}$$

$$\begin{aligned} \text{LNCZ} = & 0.0637\text{LNY}(-1) - 0.0938\text{LNY}(-2) \\ & - 0.0807\text{LNGDP}(-1) + 0.1726\text{LNGDP}(-2) \\ & + 0.5011\text{LNCZ}(-1) + 0.3619\text{LNCZ}(-2) \\ & - 12.2478\text{LNP}(-1) + 10.7965\text{LNP}(-2) \\ & + 10.2506, \end{aligned}$$

$$\begin{aligned} \text{LNP} = & -0.0010\text{LNY}(-1) + 0.0001\text{LNY}(-2) \\ & + 0.0048\text{LNGDP}(-1) - 0.0007\text{LNGDP}(-2) \\ & + 0.0133\text{LNCZ}(-1) - 0.0124\text{LNCZ}(-2) \\ & + 0.6274\text{LNP}(-1) + 0.2368\text{LNP}(-2) + 0.9366. \end{aligned}$$

(8)

The results of this model estimation show that $AIC = -30.78536$ is small enough and the log likelihood = 343.8536 is large enough, so the overall effect of the model is comparatively good. After establishing the model, we tested its effectiveness. According to values of the JB statistic, we were able to accept the null hypothesis, which indicated that the residuals were normally distributed. Stability test results show that all the unit roots were in the unit circle, which means the VAR model system is stable.

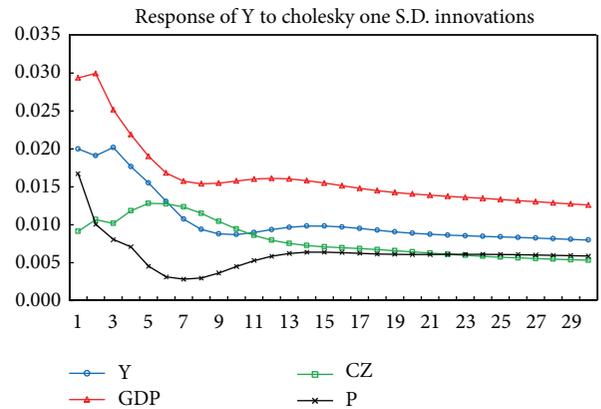


FIGURE 4: Impulse response.

Figure 4 describes the gas consumption impulse response function. The horizontal axis represents the lag periods for the impact of innovation (unit: years) and the vertical axis represents the response degree of the explanatory variables to the explained variable. The curves of the impulse response function represent the dynamic effects of each variable. The impulse function test indicates that GDP, urbanization rate, and population growth positively and persistently impact gas consumption. When we assumed a positive impact of GDP, gas consumption reaches a peak in the second period, then declines gradually, and then begins to converge in the 13th period, indicating that the GDP impact on gas consumption is significant from the second year onwards and lasts for a long period. This observation is consistent with China's development trajectory as GDP growth depends strongly on industrial growth. When a positive impact of urbanization growth rate is assumed, gas consumption reaches a peak in the 5th period after a period of fluctuation and then can be seen to be declining in the 6th period, with convergence beginning in the 11th period. This result indicates that the impact of urbanization growth rate on gas consumption is significant in the 5th year and continues to be so for a long period of time. When positive impact of population growth rate is assumed, gas consumption can be seen to be declining

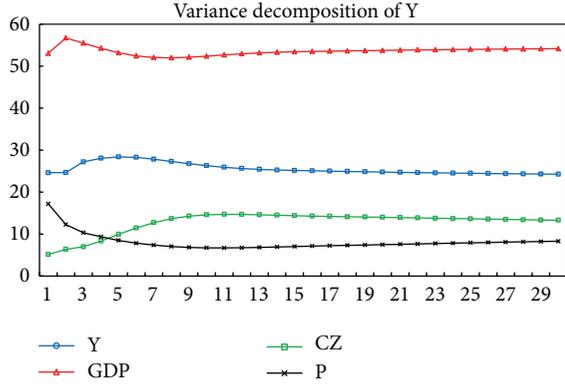


FIGURE 5: Variance decomposition.

gradually before the 9th period and then begins to increase steadily until convergence in the 11th period. This indicates that the impact of population growth on gas consumption is significant in the 8th year. However, a decline in population growth has a negative impact on gas consumption.

Figure 5 shows the results of the variance decomposition; the horizontal axis represents the lag periods for the impact of innovation (unit: years); the vertical axis represents the rate of contribution of economic variables to gas consumption. Contribution of economic variables to gas consumption remained at around 25% with the contribution of GDP being the largest at about 53% and reaching a maximum 56.71% in the second period. Contribution of urbanization rate can be seen to be increasing with a contribution of 14.72% in the 11th period and a contribution rate of 15% after that. This suggests that China's urbanization rate has been increasing in recent years; as more people migrate to cities and towns, demand for gas increases, implying that urbanization has been a significant factor for the increase in gas consumption. The contribution rate of the population on gas consumption appears to be minimal and has an overall declining tendency. The lowest contribution of population can be seen to be 6.7% in the 11th period and it remains at around 7% after that. Because China's population growth has been controlled effectively by family planning measures, the birth rate in the last few years has fallen. At the same time, an increase in the proportion of the aging in the population has led to an overall decline in population growth. Further, technological progress has meant that overall per capita energy consumption has declined, leading to a decline in contribution of the population growth rate to increase in energy consumption.

4.5.2. BVAR Model Approach. As already discussed, the VAR model does not consider the economy and so is unable to give an accurate structural interpretation. Blanchard and Quah (1989) used a vector autoregression model to analyze macroeconomic fluctuations in the USA. Ray and Tsay (2002) [20] proved that the structural vector autoregression model is equivalent to the general vector autoregression model. The main characteristics of the VAR model are that quite a few parameters need to be estimated and a large sequence of data samples is required. When the sample size

is small, the estimated parameter error is significant. Bias theory, however, has significant advantages in the case of small samples and can provide a new method for estimating VAR model parameters. In 1986, Litterman forecast seven macroeconomic indicators (GNP, etc.) in Minnesota State using the BVAR model.

In the BVAR model, coefficients of all variables fluctuate around their respective mean values and a prior distribution function is given. There are many prior distributions, such as the conjugate prior distribution, the maximum entropy prior distribution, the ML-II distribution, and the hierarchical prior distribution. However, different prior distribution methods lead to different prediction results. Because the choice of a prior distribution method plays an important role in the BVAR model's prediction accuracy, many prior distributions can be (theoretically) selected but some of the prior distributions may not be estimated or the estimated error may be large because of limitations of the selected data. When examining the three prior distributions incorporated in the MSBVAR software package, the normal-flat prior appears to be more appropriate than the other two for the selected data. The normal-flat prior (Minnesota prior) has been applied widely in previous studies, so in this paper, we used the normal-flat prior distribution to establish the BVAR model. Results of the BVAR model are shown in Table 8 (0 = normal-Wishart prior, 1 = normal-flat prior, and 2 = flat-flat prior (i.e., akin to MLE)).

The matrix form of the VAR model can be expressed as in the following equation:

$$y_t = Z_t B' + \mu, \quad t = 1, 2, \dots, n. \quad (9)$$

If y_t , $Z_t B'$, and μ_t ($t = 1, 2, \dots, n$) are arranged according to the transpose, then each is a $n \times m$ matrix. The equations can be written as follows:

$$Y = ZB + U, \quad U \sim N_{nm}(0, \Sigma \otimes I_n), \quad (10)$$

where $Y = [y_1^T, y_2^T, \dots, y_n^T]_{n \times m}^T$, $Z = [z_1^T, z_2^T, \dots, z_n^T]_{n \times (mp+1)}^T$, $U = [u_1^T, u_2^T, \dots, u_n^T]_{n \times m}^T$, $B = [c^T, A_1^T, \dots, A_p^T]_{(mp+1) \times m}^T$, and $Z_t = [1, y_{t-1}, \dots, y_{t-p}]_{(mp+1) \times 1}^T$.

Then, a general multivariate normal distribution is introduced as the prior distribution coefficient:

$$P(B, \Sigma) \propto |\Sigma|^{-(m+1)/2} \exp \left[-\frac{1}{2} (B - \bar{B})' C^{-1} (B - \bar{B}) \right]. \quad (11)$$

Here, \bar{B} , C are the prior mean and covariance, respectively. By combining the likelihood function of B and Σ , a posterior distribution can be determined as follows:

$$\begin{aligned} P(B, \Sigma | Z, Y) &\propto \exp \left[-\frac{1}{2} (B - \bar{B})' \bar{S}^{-1} \otimes Z' Z (B - \bar{B}) \right]^{-n/2} \\ &\times \exp \left[-\frac{1}{2} (B - \bar{B})' C^{-1} (B - \bar{B}) \right] \\ &\propto \exp \left\{ -\frac{1}{2} (B - \bar{B})' F (B - \bar{B}) \right\}, \end{aligned} \quad (12)$$

where $\bar{B} = (C^{-1} + \bar{S}^{-1} \otimes Z'Z)^{-1} (C^{-1}\bar{B} + \bar{S}^{-1} \otimes Z'Z \hat{B})$, $F = C^{-1} + \bar{S}^{-1} \otimes Z'Z$, and the posterior distribution of the coefficient matrix can be considered a multivariate normal distribution. The mean (variance) is the weighted average value based on the prior and sample mean (variance).

As the RBF neural network quantile regression model and the BVAR model have different prediction advantages and disadvantages, we combined the results of the two models to determine the combinational forecast. The key to an accurate combinational forecast is to determine the weighting coefficient for each forecast model. Kurz-Kim (2008) [21] analyzed the weighting coefficients in a combinational forecast and selected the same weighting coefficient for each forecast model. Had different weighting coefficients been selected, the combinational forecast would have lost its meaning, so the equally weighted combinational forecast had good rationality and applicability. The results of the combinational forecast are shown in Table 8.

According to the combinational forecast China's gas consumption was to be 140053.89 billion cubic meters in 2012 and the actual figure was 132774.13 billion cubic meters, a relative error of 5.48%, which is within an acceptable range, as the average annual growth was 8.28%, but the year-on-year growth rate slowed. The reason for this could be the increasing use of clean energy sources such as solar energy, biomass energy and geothermal energy applications, the slowing of China's GDP growth, or a decline in the price of other alternative energies. Overall, however, it can be seen from the broad trends that gas consumption in China is continuing to rise at a stable rate.

5. Conclusion

In this paper, we examined China's gas consumption and used path analysis to screen out the core driving factors. We found that population growth and urbanization rates were the main factors influencing gas consumption and GDP was the main limiting factor. At the same time, results of the impulse function in the VAR model showed that GDP, urbanization rate, and the population growth rate positively and persistently impact gas consumption and that these three core factors have an evident long-term impact. Contribution of GDP to change in gas consumption was shown to be at around 25%. Contribution rate of GDP on gas consumption was the highest at around 53% with a maximum 56.71% achieved in the second period. Contribution of urbanization growth rate to gas consumption was shown to be increasing with 14.72% estimated in the 11th period and around 15% after that. Contribution of population growth rate to gas consumption was shown to be minimal and had a gradual declining effect, with the lowest contribution being 6.7% in the 11th period and around 7% after that. This study relates to the current structure of energy consumption where coal is the main source of energy. However, in recent years the pressure of environmental pollution problems has meant that China needs to change the composition of its energy consumption. Therefore, the use of gas as a cleaner energy source is shown to have a significant growth trend. The prediction results of

the model also show that, with a change in China's energy consumption structure, gas consumption is expected to reach around 178649.23 million cubic meters by the end of 2015 and about 264698.86 million cubic meters by 2020. The average gas consumption growth rate from 2012 to 2020 was estimated at 8.28 percent. This indicates that the relevant government departments need to take measures to accommodate this rise in gas consumption, such as developing reasonable allocation, pricing, and planning policies. Also, there needs to be a greater focus by the government on gas exploration and development within China, such as the promotion of coal seam gas, shale gas, and other unconventional oil and gas resources development and utilization; the need for increased gas reserves; improvements in the domestic oil and gas trunk pipeline network; importation of gas; development of liquefied gas receiving stations; cross-regional backbone gas transmission and distribution pipeline network construction; and research on effective extraction of coal-bed methane, as well as a coal gas supply coordinated development plan.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (NSFC) under Grant no. 71350007; the Postdoctoral Science Foundation of China, no. 2012M510580; and the Soft Science Research Program of Shaanxi Province, no. 2012KRM95. The authors would like to thank the anonymous referees as well as the editors.

References

- [1] A. Das, A. A. McFarlane, and M. Chowdhury, "The dynamics of natural gas consumption and GDP in Bangladesh," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 269–274, 2013.
- [2] N. Apergis and J. E. Payne, "Natural gas consumption and economic growth: a panel investigation of 67 countries," *Applied Energy*, vol. 87, no. 8, pp. 2759–2763, 2010.
- [3] M. Kankal, A. Akpınar, M. I. Kömürçü, and T. Ş. Özşahin, "Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables," *Applied Energy*, vol. 88, no. 5, pp. 1927–1939, 2011.
- [4] V. Bianco, F. Scarpa, and L. A. Tagliafico, "Scenario analysis of nonresidential natural gas consumption in Italy," *Applied Energy*, vol. 113, pp. 392–403, 2014.
- [5] O. E. Canyurt and H. K. Ozturk, "Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey," *Energy Policy*, vol. 36, no. 7, pp. 2562–2569, 2008.
- [6] E. F. Sánchez-Úbeda and A. Berzosa, "Modeling and forecasting industrial end-use natural gas consumption," *Energy Economics*, vol. 29, no. 4, pp. 710–742, 2007.
- [7] R. Gutiérrez, A. Nafidi, and R. G. Sánchez, "Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model," *Applied Energy*, vol. 80, no. 2, pp. 115–124, 2005.

- [8] A. Khotanzad, H. Elragal, and T.-L. Lu, "Combination of artificial neural-network forecasters for prediction of natural gas consumption," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 464–473, 2000.
- [9] G. Xu and W. Wang, "Forecasting China's natural gas consumption based on a combination model," *Journal of Natural Gas Chemistry*, vol. 19, no. 5, pp. 493–496, 2010.
- [10] B. Soldo, "Forecasting natural gas consumption," *Applied Energy*, vol. 92, pp. 26–37, 2012.
- [11] F. B. Gorucu, "Artificial neural network modeling for forecasting Gas consumption," *Energy Sources*, vol. 26, no. 3, pp. 299–307, 2004.
- [12] Y.-S. Lee and L.-I. Tong, "Forecasting energy consumption using a grey model improved by incorporating genetic programming," *Energy Conversion and Management*, vol. 52, no. 1, pp. 147–152, 2011.
- [13] A. Azadeh, S. M. Asadzadeh, M. Saberi, V. Nadimi, A. Tajvidi, and M. Sheikalishahi, "A neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting and behavior analysis: the cases of Bahrain, Saudi Arabia, Syria, and UAE," *Applied Energy*, vol. 88, no. 11, pp. 3850–3859, 2011.
- [14] Y.-Y. He, Q.-F. Xu, S.-l. Yang, and B.-G. Xu, "A power load probability density forecasting method based on RBF neural network quantile Regression," *Proceedings of the CSEE*, vol. 33, no. 1, pp. 93–98, 2013 (Chinese).
- [15] P. Crompton and Y. Wu, "Energy consumption in China: past trends and future directions," *Energy Economics*, vol. 27, no. 1, pp. 195–208, 2005.
- [16] J. Chai, J.-E. Guo, L. Meng, and S.-Y. Wang, "Exploring the core factors and its dynamic effects on oil price: an application on path analysis and BVAR-TVP model," *Energy Policy*, vol. 39, no. 12, pp. 8022–8036, 2011.
- [17] J. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Computation*, vol. 1, no. 2, pp. 281–294, 1989.
- [18] R. Koenker and G. Bassett Jr., "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.
- [19] J. W. Taylor, "A quantile regression neural network approach to estimating the conditional density of multiperiod returns," *Journal of Forecasting*, vol. 19, no. 4, pp. 299–311, 2000.
- [20] B. K. Ray and R. S. Tsay, "Long-range dependence in daily stock volatilities," *Journal of Business and Economic Statistics*, vol. 18, no. 2, pp. 254–262, 2000.
- [21] J.-R. Kurz-Kim, "Combining forecasts using optimal combination weight and generalized autoregression," *Journal of Forecasting*, vol. 27, no. 5, pp. 419–432, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

