Applied Energy 169 (2016) 738–747

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

A hybrid model for explaining the short-term dynamics of energy efficiency of China's thermal power plants

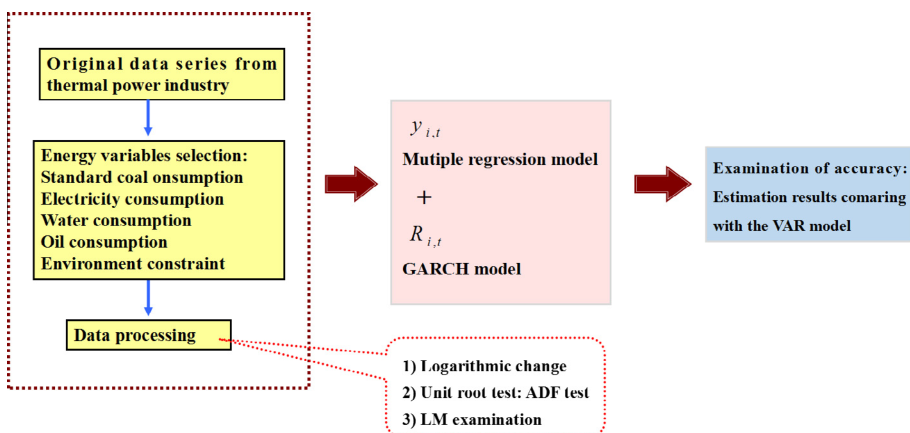
Ming-Jia Li, Chen-Xi Song, Wen-Quan Tao *

Key Laboratory of Thermo-Fluid Science and Engineering of Ministry of Education, School of Energy & Power Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

H I G H L I G H T S

- The new hybrid method consists of a multiple regression model and the GARCH model.
- The hybrid model is an improved benchmarking methodology.
- The method analyses the energy efficiency index of thermal power plants without artificial intervention.
- The volatility degree between determinants and energy efficiency index can be evaluated.
- Environment constraints of thermal power plants are considered.

G R A P H I C A L A B S T R A C T



A R T I C L E I N F O

Article history:
Received 16 July 2015
Received in revised form 12 February 2016
Accepted 14 February 2016
Available online 27 February 2016

Keywords:
Energy efficiency index
Thermal power plant
GARCH model
VAR model

A B S T R A C T

A new hybrid methodology is introduced which is a combination of multiple regression model and generalised autoregressive conditional heteroskedasticity (GARCH) model. Comparing the new approach and the vector auto-regression (VAR) model, this paper analyses the short-term dynamics of the energy efficiency index (EEI) in response to change in the five indicator variables for thermal power plants in China. The result indicates that: (i) The new hybrid model can directly calculate the EEIs of thermal power plants without artificial intervention. (ii) It can eliminate the disturbance of residual superposition. (iii) The new method will offer more direct information on the degree of volatility among determinants and operating inefficiency.

1. Introduction

1.1. Background

In this decade, there is an increasing number of studies paid more attention to evaluating, analysing energy efficiency around the world. The global warming is one of the world's most significant problems and global warming is primarily attributed to the emission of carbon dioxide (CO_2). China has become the second-largest economy entity in the world since the implementation of China's economic reform in 1978. However, this achievement has led to the inefficient natural resource utilisation, rising of emission of carbon dioxide and large energy consummation as well. Today, China is the world's second largest installed electricity generation nation. China's CO_2 emission is close to 10 billion tons in 2013, and

* Corresponding author.
E-mail address: wqtao@mail.xjtu.edu.cn (W.-Q. Tao).

http://dx.doi.org/10.1016/j.apenergy.2016.02.082
0306-2619/

Nomenclature

Δy_t	constant variance
$y_{i,t}$	energy efficiency indicators of item i at the end of period t
$x_{i,t}$	independent variables
R_t	energy efficiency
M_i	impact variable of the event i
h_t	conditional variance of R_t
υ_t	result of VAR model
a	constant
b	constant

Greek symbols

α	parameter
β	parameter
γ	coefficient of correlation
Γ	matrix valued polynomial
ε_t	residual error
υ	data series
μ_t	squared distribution

Subscripts

t	time index
0–1	state point
i	time-varying changes of impact variables
n	number of variables
q	the number of lags
c	point
k	point

coal-fired power generations contribute more than 80%. Thermal power generation is a procedure in which thermal energy obtained by the consumption of coal, gas and other fuels in transformed into electricity. Therefore, thermal energy is the main source of energy consumption and pollution in electricity generation industry.

To improve energy utilisation efficiency, protect the environment and implement sustainable development, various series of energy policies have been published by the Chinese government in order to establish a sufficient regulatory of energy use in different industries. A target has also been presented that the CO_2 emissions per unit gross domestic production (GDP) should be reduced by 40–45% by 2020 compared with the level in 2005. Government highly focuses on energy efficiency evaluation of high energy-consuming industries, especially thermal power industry. Therefore, evaluating and measuring the energy efficiency of thermal power industry with the consideration of environment constraint indicator is very vital for plants to reduce energy consumption and monitor environment pollution. The paper will adopt five variables of thermal power plants. They are namely: (i) standard coal consumption per unit product of power (hereinafter to be referred as SCC); (ii) the rate of electricity consumption of power plant (hereinafter to be referred as EC); (iii) the total water consumption per unit product of power generation (hereinafter to be referred as WC); (iv) the total oil consumption per year (hereinafter to be referred as OC); (v) the investment rate of desulfurization system (hereinafter to be referred as RDS) which is treated as an environmental constraint indicator of each thermal power plant.

1.2. Previous studies for evaluating utility performance

The importance of energy efficiency evaluating on one hand, and its complexity on the other hand, has motivated many studies in this area. There are several popular methodologies for energy efficiency evaluation. The first methods are linear approach and stochastic frontier analysis, which are the primal studies in evaluating utility efficiency. Farrell [1] contributed the linear regression approaches to measuring utility production efficiency based upon the foundation work of Knight [2] and Debreu [3]. The new linear regression approaches provided a distance function to evaluate efficiency in a primal system. The improvement work by Shephard [4] developed a mutual relationship between costs, production and benefits. Filippini and Hunt [5] examined a stochastic frontier approach to calculate the difference of energy efficiency among OECD countries. They adopted energy price to be an independent variable, which is in order to measure how well do the actions of consumers respond to the energy efficiency with the policy variables change. Zhou et al. [6] employed a parametric frontier approach to evaluate economic wide energy efficiency indicators. The methodology is the basis of index decomposition analysis, which can be applied to indicate a change in energy consumption. However, linear approaches can only indicate a single relation from input variables to outputs. It does not demonstrate the interconnection between variables. The behaviour of the changes of energy efficiency may not be completely captured by the linear techniques. Moreover, this method does not take into account random errors. Estimated results will be influenced by the residual superposition.

To solve this problem, the second approaches have been further developed. The multiple regression models are extended to explain the influences of many variables. Denholm et al. [7] adopted the model to assess the technological and environmental performance of wind power plants. They found that the energy efficiency of wind power plants will be at least five times greater than that of fossil combustion technology by improving capacity efficiency. Bernard and Cote [8] used principal component analysis (PCA), one type of multiple regression models, to calculate the energy efficiency of manufacturing. They treated environmental factor as an important determinant in system simulation. Besides that, they concluded that only regression approaches are unable to demonstrate the particular patterns of energy efficiency. It cannot provide a fair benchmarking of energy efficiency performance among different objectives. Based upon above mentioned, some studies further extended the data envelopment analysis (DEA) approach to involve more determinants. The DEA model is a nonparametric "black-box" multiple regression model. There are huge amounts of scholars applied the DEA model to evaluate the overall energy efficiency index (EEI) through involving different types of inputs. Examples of such studies include Song et al. [9], Khoshnevisan et al. [10], Bianchi et al. [11] and Mousavi-Avval et al. [12]. The drawback is that most models are invariant with respect to the decision making units (DMUs), and these models mainly focus on less input or higher output for better overall energy efficiency. This methodology simply includes all DMUs in one analysis while does not provides a mechanism for incorporating useful information such as volatility of interrelation into the analysis.

Instead of solving these methodological issues in an ordinary least squares regression, the generalised autoregressive conditional heteroskedasticity (GARCH) models were established. The earliest fundamental work of ARCH models was pioneered by Engle in 1982 [13], and the GARCH methodology was introduced by

Table 1
Variables description.

Variable	Indicator	Abbreviation	Unit	Definition
No. 1	The standard coal consumption per unit product of power units	SCC	gce/kW h	Coal consumption by boiler units
No. 2	The rate of electricity consumption of power plant	EC	%	The ratio of power consumption to generate capacity
No. 3	The total water consumption per unit product of power units	WC	kg/kW h	Fresh water consumption within the period of power generation (not containing repeated use of water)
No. 4	The total oil consumption per year	OC	tons/age	Oil consumption within the period of power generation
No. 5	The rate of desulfurization system	RDS	%	The efficiency rate of desulfurization system

Bollerslev in 1986 [14]. This approach is based on econometric models which was originally developed and widely adopted for stock markets. The GARCH models can solve the deficiencies mentioned above. The strength of GARCH models is not only to provide a stochastic volatility measurement between different variables but to allow the data to decide their own best parameters. It is most well-known for the dealing with mutual relationships. Nowadays, these models are widely adopted in many fields. Andersen and Bollerslev [15] published an affirmative result through constructing the continuous-time data series in the GARCH (1,1) model. It proved that the estimations of volatility are correlated highly with the different variables. The GARCH (1,1) model did provide good volatility. Day and Lewis [16] examined the relative accuracy of volatility using time series data from the crude oil future market. They made a comparison of near-term volatility and distant-term volatility for proving that the implied volatilities from oil market have significant within-sample explanatory power. Poon and Granger [17] further made a detailed review about how to forecast the volatility. They evaluated different types of GARCH model to demonstrate the best method. Most recently, Alberola et al. [18] and Byun and Cho [19] further applied the GARCH models to test the relationship between carbon price, energy price and the change in market structure in the European Union. Chevallier [20] used the GARCH models to examine the volatility degree of factors affecting EU carbon futures. The other GARCH studies, such as Goto and Karolyi [21], Sadorsky [22] and Elder and Serletis [23], explored the relationship between electricity market and natural gas, and proved the relation between oil price and macroeconomic factors. Li et al. [24] adopted the GARCH model to evaluate and predict the energy consumption in China.

According to the above studies [1–12], linear approaches and multiple regression models have been widely adopted to evaluate EEIs in many areas, but they cannot demonstrate the volatility degrees and interconnections among determinants. The weights used in the analysis are also defined artificially. Few papers considered environment constraints of thermal power units [6–12]. This paper focuses on estimating energy efficiency of thermal power plants without manual intervention, and analysing the volatility between all variables. Therefore, the multiple regression model and the GARCH model are combined for energy efficiency evaluating and volatility analysing.

The contributions and novelties of this study are as follows. First, this paper provides a new econometric methodology to measure the cross-sectional efficiency of thermal power industries. The new hybrid model combines the GARCH model and the multiple regression model. Because the EEI systematically varies over time, general regression methods cannot fully explain the mutual influence among various determinants. Second, normal regression methods usual lead the residual superposition, and estimate of efficiency is biased by construction. This hybrid method can overcome the above insufficiencies. It will capture the complex features and high volatility of data when data change over time, and the heteroskedastic regression will weaken the effect of errors. Third, this study treats thermal power plants as the whole system. It examines the changes in energy efficiency, and investigates the volatility degree among influential factors. The study identifies important factors that affect the performance of thermal power plants in China.

The reminder of the paper is organised as follows. Section 2 introduces the data adopted. Section 3 describes the methodology used in this paper. Section 4 demonstrates the analysis empirical result. Section 5 gives conclusions.

2. Data selection and descriptive statistics

According to wide practical researches, this study adopts five physical indicators as impact variables. Policy makers usually treat the SCC and the EC as the classic indicators compared with the international standard level. However, because of the increasing inefficient natural resource, it should be figured out that the WC and the OC are other two kinds of energy carriers. The RDS could be considered to measure the environment constraint level of every power unit. The five indicators and their descriptions are listed in Table 1.

The data set consists of the monthly data series of 600 MW thermal power plants from National Database of China Electricity Council [25]. The data set covers from January 01, 2010 to December 31, 2012, a time period lasts three years. Following issues need to be noted during data selection. First, this paper focuses on the coal-burning units. Therefore the data of coal-fired units are primarily selected from the database. Second, 67 thermal power plants are further considered. Every thermal power plant normally installed 3 or 4 units. A total number of sample observations for each variable will be around 200. The features of systems are different. It should be differentiated that some plants use steam turbine generators and others adopt gas turbine power systems. Different power generators lead different electricity used. Third, because of the physical indicators of equipment plant-levels, low-frequency data is more preferred compared with daily data. Finally, the paper will make data process before adopting them to the further analysis.

The descriptive statistics of the volatility indicators (Panel A) and their logarithmic change (Panel B) are presented in Table 2.

The result illustrates that the SCC and the EC are stable basically. The OC is the most volatile variable, and it ranges from −10.602 to 9.680. The WC volatiles a little. It varies from −6.215 to 6.141. The mean of the RDS is 87.7197. It demonstrates that most thermal power plants have good performances of desulfurization systems. All variables have long left tail because all kurtosis are greater than 3, and every variable has a leptokurtic distribution with asymmetric tails. It is to be noted that kurtosis is a measure of the shape of a probability distribution of random variables, while leptokurtic is a statistical distribution when the kurtosis value is larger than a normal distribution.

Table 2
Descriptive statistics.

	Observation	Standard deviation	Skewness	Kurtosis	Mean	Maximum	Minimum
Panel A: Levels							
SCC	646	11.795	0.681	3.075	312.775	354.030	288.200
EC	650	0.279	0.064	3.573	0.002	0.761	−0.827
WC	670	7.694	24.841	625.942	1.271	195.000	0.000
OC	672	26.736	−3.152	17.159	166.181	2083.970	0.000
RDS	642	21.3626	−2.382	8.308	87.720	100.000	0.120
Panel B: First logarithmic change							
SCC	646	0.050	−0.102	3.162	0.001***	−0.138	0.135
EC	650	0.279	−0.059	3.550	0.002	−0.827	0.761
WC	670	1.383	−0.036	3.362	0.015***	−6.215	6.141
OC	672	1.002	13.97	14.796	0.072	−10.602	9.680
RDS	642	0.691	−6.735	58.520	4.380	4.605	0.120

Note: full sample of variables are collected from every thermal power plant yearly closing data from January 01, 2010 to December 31, 2012.
*** Significance at the 1% level.

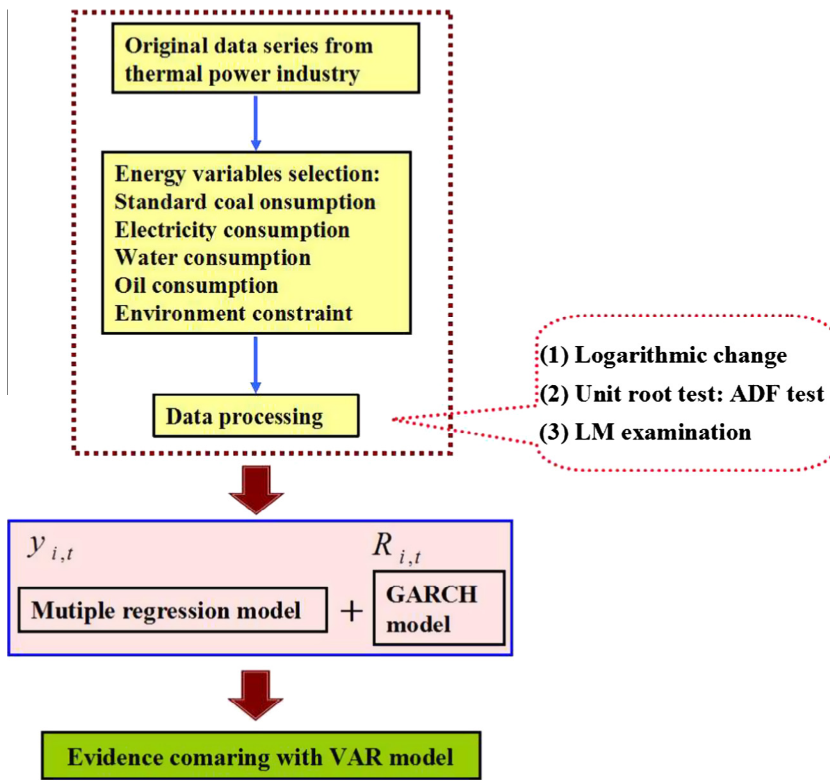


Fig. 1. Flowchart of the proposed method.

3. Econometric methods

This section adopts econometric methods to investigate short-term relationships among the volatility of energy efficiency, the SCC, the EC, the WC, the OC and the RDS every year. The flowchart of the proposed method is sketched in Fig. 1. The framework of the hybrid model is listed as follows:

(1) The augmented Dickey Fuller (ADF) unit root test [4] is adopted to analyse whether data are stationary for further calculation. In order to measure the effectiveness of the result, the Phillips Perrson (PP) test [4] is employed to make a comparison.
(2) The Lagrange multiplier (LM) test [13] is further used to examine whether time series data have heteroscedasticity.
(3) After above econometric tests, the hybrid method is further established to calculate energy efficiency indicators and to analyse the influence of different variables on the changes of energy efficiency.
(4) Vector auto-regression (VAR) model [14], a well-known approach of the AR-based models, is employed to measure the effectiveness of the proposed method. The vital advantage of VAR model is to provide a simultaneous estimation of the interconnection under a mature structure, while taking into account the time series data. It is an appropriate option to examine the short-term fluctuation for proving the accuracy of estimation results of the hybrid model.

3.1. Unit root test: ADF test and PP test

As mentioned above, energy efficiency can be highly volatile. An appropriate GARCH model should be adopted because this model considers every time series moment as variants. The GARCH (1,1) model is widely employed since it efficiently explains systematic variations of assets volatility [12]. The GARCH (1,1) model requires the covariance matrix and the data to be systematically changed over time. Therefore, the test of time series is necessary before the GARCH (1,1) model is adopted. In order to describe the unbiased estimators for the classical regression model, the data should be stationary because that the existence of correlations between variables is a good evidence of a causal relationship [12]. This paper applies two methodologies of unit root tests: the ADF test and the PP test. Both include extra lagged terms of the dependent variables for capturing the auto-correlation. The ADF test equations are shown from Eqs. (1) to (3) and results of the PP test will be adopted to make a fair comparison. Equations of the ADF test are as follows:

$$\Delta y_t = \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \mu_t \tag{1}$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \mu_t \tag{2}$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + a_2 t + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \mu_t \tag{3}$$

where α is a constant, $a_2 t$ represents trend of equations. γ and β denote the short-run coefficients. Eq. (1) presents that the ADF test has no constant and no trend in the series. Eq. (2) represents the test as having a constant but no trend. Eq. (3) shows that the test not only has a constant but a trend.

3.2. LM test

The LM test is the second step after examining the stationary of data series. According to previous studies, the classical regression issue of the joint hypotheses has to be considered. Because the study would employ the GARCH (1,1) model in the following section, this part would choose the lag number at 1 to test whether the ARCH effect exists. The LM test [7] applies the null hypothesis that there is no serial correlation up to lag order p, where the lag is equal to 1 in this test. It tests for first order serial correlation.

$$\mu_t^2 = \gamma_0 + \gamma_1 \upsilon_{t-1}^2 + \cdots + \gamma_q \upsilon_{t-q}^2 + \mu_t \tag{4}$$

where μ_t is squared distribution of Δy_t, and it is obtained from Eq. (3). υ is data series. q is the number of lags, and γ represents the coefficient.

3.3. Multiple regression method

After the previous examination, this study will adopt the new hybrid model combining the multiple regression method and the GARCH model. This section evaluates the EEI primarily.

$$y_{i,t} = \beta_0 + \beta_i x_{i,t} + u_t \tag{5}$$

where $y_{i,t}$ is the EEI of item i at the end of period t. $x_{i,t}$ is independent variables of model. β_i can be estimated by the correlation matrix. The important choice for this method relates to the EEI to be calculated.

3.4. GARCH (1,1) model

In order to provide a comprehensive analysis regarding the stochastic activity of all time series data, the volatility model is further applied. The volatility is described as conditional standard deviations of time series variables. The model is generally known as the GARCH model, and it has been employed in wide fields of study. The GARCH (1,1) model [13] formulas are as follows:

$$R_{i,t} = c + \sum_{k\geqslant 1} a_k R_{t-k} + \sum_i b_i M_{i,t} + \varepsilon_t \tag{6}$$

$$\varepsilon_t \sim (0, h_t) \tag{7}$$

$$h_{i,t} = \alpha_0 + \beta h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \sum_i \gamma_i M_{i,t} \tag{8}$$

where $R_{i,t}$ is the EEI which is estimated by the ordinary least square approach in Eq. (5). $h_{i,t}$ explains the conditional variance of R_t, and it is depended upon $M_{i,t}$. $M_{i,t}$ is the impact variable of the event i and the event of i represents the time-varying changes of impact variable. The coefficients of Eq. (6) need to satisfy $a_k \geqslant 0, b_i \geqslant 0$, and $c > 0$, to ensure that the conditional variance is positive. α_0, β and α_1 have to be estimated in the variance to determine the time-varying conditional variance. γ_i stands for the asymmetric effect.

In order to specifically analyse the five impact variables of energy efficiency of thermal power plants, the model is extended as follows:

$$R_t = c + \sum_{k\geqslant 1} a_k R_{t-k} + b_{SCC} M_{SCC,t} + b_{EC} M_{EC,t} + b_{WC} M_{WC,t} + b_{OC} M_{OC,t} + b_{RDS} M_{RDS,t} + \varepsilon_t \tag{9}$$

$$\varepsilon_t \to (0, h_t) \tag{10}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma_{SCC} M_{SCC,t} + \gamma_{EC} M_{EC,t} + \gamma_{WC} M_{WC,t} + \gamma_{OC} M_{OC,t} + \gamma_{RDS} M_{RDS,t} \tag{11}$$

where $M_{SCC,t}$ stands for the standard coal consumption per unit product of power, $M_{EC,t}$ is the rate of electricity consumption of power plant, $M_{WC,t}$ represents the total water consumption per unit product of power generation and $M_{OC,t}$ is the total oil consumption per year. $M_{RDS,t}$ means the investment rate of desulfurization system.

3.5. VAR model

The VAR model is one classical well-known model of AR-based models. It is employed to prove the effectiveness through providing a comparison between its estimation and the empirical results of the hybrid model. Formulas of VAR model are as follows:

$$\Gamma(Y_{EEI}) = \Gamma_1 x_{i,t} + \Gamma_2 x_{i,t} + \cdots + \varepsilon_t \tag{12}$$

$$\upsilon_t = \Gamma_0^{-1} \varepsilon_t \tag{13}$$

where $\Gamma(Y_{EEI})$ is a matrix valued polynomial in positive powers of the variable EEI, n is the number of variables in the system, ε_t represents the residual error of x_t, and υ_t is the result from the VAR model. In accordance with the standard likelihood ratio tests, the selected optimal lag length is 1.

Considering an identification that the Y_{EEI} can be explained by a set of variables. Variables are namely: (i) SCC; (ii) EC; (iii) WC; (iv) OC; and (v) RDS.

4. Computational results

4.1. Unit root test

Before detailed analysis is presented, it is important to look at the unit root test in order to glean preliminary information about time series data of variables. The results of unit root test for the SCC, the EC, the WC, the OC and the RDS are reported in Table 3.

Table 3
Unit root test of five variables.

Variables	Automatic lag length	ADF test	PP test	1% level of critical value	Inference
lng (SCC)	1	−47.319***	−71.231***	−3.497	$Inf \sim I(1)$
lng (EC)	1	−42.003***	−34.428***	−3.336	$Inf \sim I(1)$
lng (WC)	1	−35.374***	−38.032***	−3.284	$Inf \sim I(1)$
lng (OC)	1	−30.782***	−36.146***	−3.405	$Inf \sim I(1)$
lng (RDS)	1	−20.607***	−27.381***	−3.411	$Inf \sim I(1)$

*** Significance at the 1% level.

Applying the unit root test is a pre-condition of the GARCH (1,1) model, and the usefulness is to confirm whether variables are stationary to establish the time series model. The null hypothesis is that there is an unit root in variables. The number of total observations is 3280. The following test demonstrates that the *t*-statistics value generated lower than the critical value, and it means the series data do not have a unit root problem. All the first order differenced series are stationary. The data can be adopted to do further calculation. Expressing the time series data in natural logs is customary and useful because the coefficients of the estimated regression are elastic.

4.2. LM test

After the stationary test, the LM test is further employed. The Obs*R-squared shown in Table 4 provides the Breusch-Godfrey LM test statistics [8]. In order to achieve whether the null hypothesis is rejected, the LM test statistics value (125.392) should be compared with the critical value of Chi-Squared (1) value. The critical value of Chi-Squared (1) is selected as 3.62 from the Statistical Table. As 125.392 exceeds the critical value 3.84, thus no doubt the null hypothesis in these two groups can be rejected. Therefore, a significant serial correlation exists between variables.

For a better understanding the trend of variables' fluctuation from 2010 to 2012, the patterns of the five variables are analysed. Fig. 2 presents the time series pattern of the volatility of SCC each year. The total trend is active, and it generally fluctuates within around 304 gce/kW h to 320 gce/kW h in three years. Fig. 3 shows time series pattern of the volatility of EC from 2010 to 2012, and the means are increased from 4.89% (2010) to 5.15% (2011). Afterwards it reaches 5.28% (2012). Fig. 4 is the trend of the volatility of WC, and the means are decreased from 1.05 kg/kW h (2010) to 0.96 kg/kW h (2011), and then it finally achieves 0.95 kg/kW h (2012). Fig. 5 shows the pattern of the volatility of OC; the means are declined from 219.45 tons/age (2010) to 158.49 tons/age (2011), and further decreased to 155.75 tons/age (2012). Fig. 6 suggests that volatilities of RDS normally range from 89% to 100%, especially the performances of RDS are in better position in 2012 than the level in 2011. These results primarily describe the uncertainty innovations of different determinants will, to some extent, affect expected volatility changes of energy efficiency.

4.3. Correlation matrix

One function of GARCH model is to allow the data to decide their own best weights. Table 5 presents contemporaneous correlations among five variables as well as their logarithmic series. The interval of the correlation coefficient could range between −1.0 and 1.0, and a correlation of 1.0 means two variables are perfect positively corresponded to each other. Generally, the correlation is quite high when it is over 0.8 [9]. The correlations between the SCC and other variables are positive. It indicates that the expected volatilities of these indicators seem to change in the same direction within the sample period. Moreover, the SCC and the EC are highly correlated at the level of 0.853, presenting that the EC plays an important role in evaluating the SCC of thermal power plants. The WC has a weak relationship with the SCC at the level of 0.331, implying that the WC does not impact the SCC well. The OC is positively correlated with the SCC, the EC and the WC, which means that the OC may cause small fluctuation of these three variables.

Table 4
LM test for ARCH of the variables.

Lags (p)	Prob. Chi-square	Obs*R-squared
1	0.0003	125.392

Note: The Prob. Chi-square is probability Chi-squared distribution, it is a continuous probability distribution. The Obs*R-squared is observation R-squared distribution.

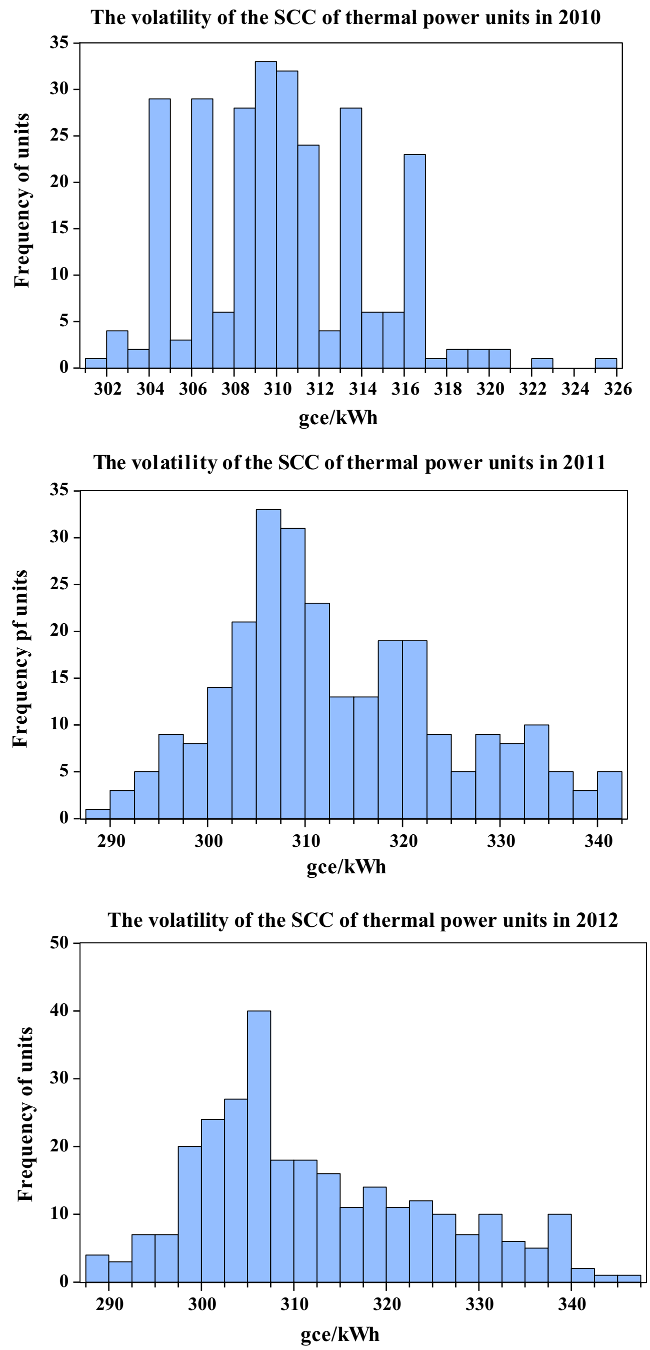


Fig. 2. The volatility of the SCC of the 67 thermal power plants from 2010 to 2012.

4.4. Application of hybrid model

How to use the proposed hybrid model for the implementation is a key issue. Due to the purpose of energy conservation in high consuming industries, government needs to make decisions on whether new regulations should be approved. This hybrid model can be adopted to evaluate the EEI conveniently.

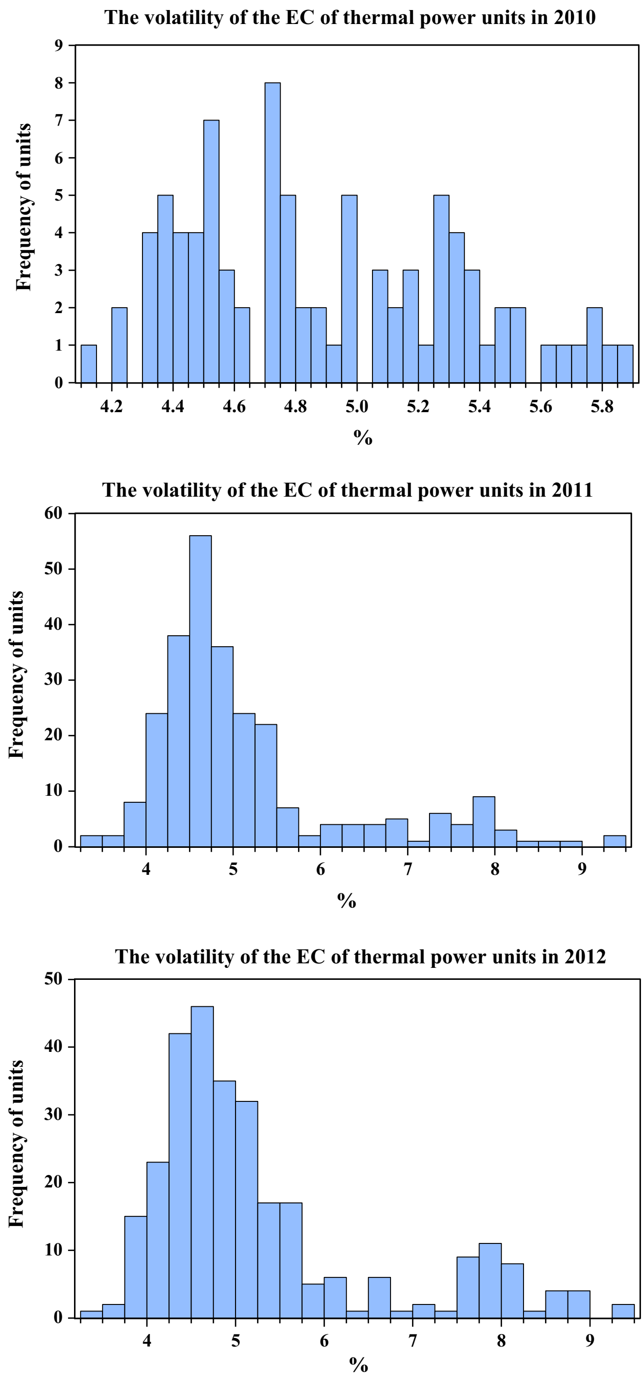


Fig. 3. The volatility of the EC of the 67 thermal power plants from 2010 to 2012.

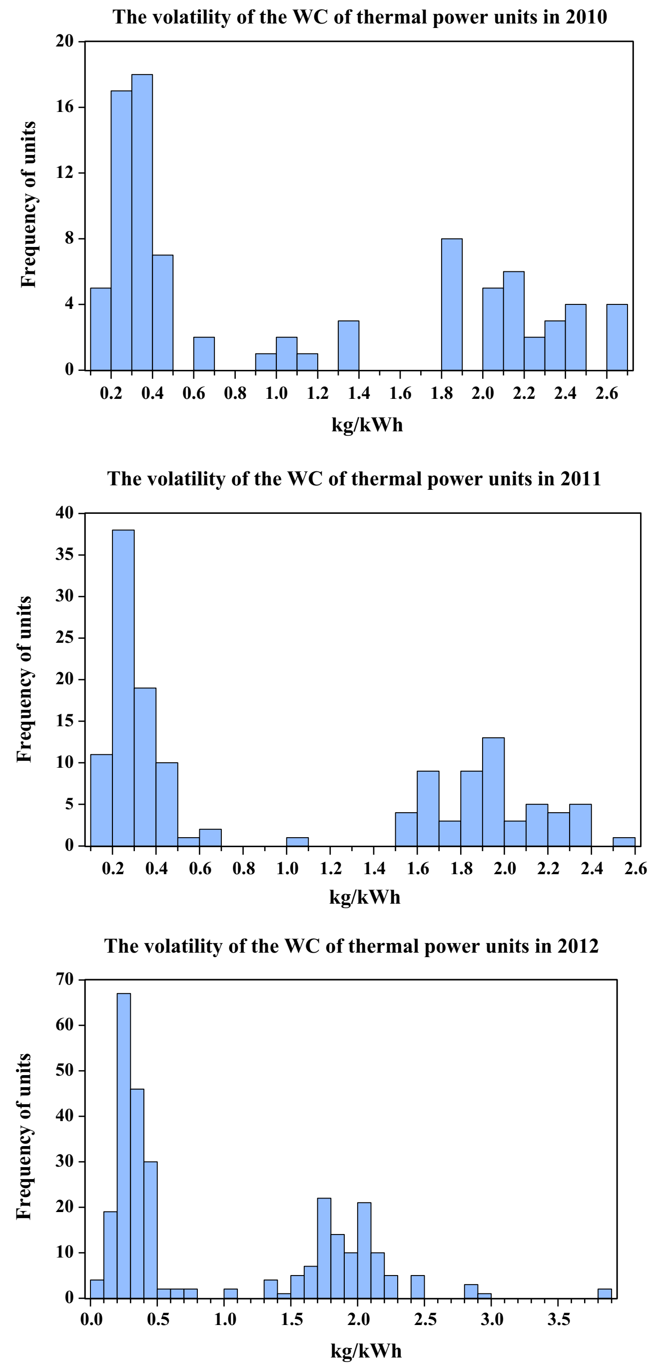


Fig. 4. The volatility of the WC of the 67 thermal power plants from 2010 to 2012.

4.4.1. Multiple regression model

To apply all series data of variables and correlation coefficients in Eq. (5), Table 6 shows EEI changes of thermal power plants during the sample period. It can be found out that the EEI of No. 12 and No. 66 thermal power plants have fluctuations. After a detailed comparison of these plants, it reveals that plants have higher efficiency indicators because they installed steam turbine power plants. These plants have potential for improving energy efficiency by new equipment with high operational quality.

The EEI by fuel is obviously important. It is vital to measure energy supply, energy use and efficient level of different types of equipment. While relatively high EEI for fuel can be treated as a sufficient level to access, it is significant for achieve governmental requirement in a secure and reliable energy supply.

It should be noticed that the meanings of empirical data of this paper are very different from the normal outputs. First, almost no existing paper considers the influence of environmental constraint indicators on the EEI. They merely calculate the operational efficiency of power plants [6–8,10–12]. Second, to the authors knowledge there is no papers can straight examine the volatility degree among the EEI and indicator variables without artificial intervention. Third, nowadays most studies calculate the EEI of industries through "black-box" calculation procedure, eg. DEA model [9] and PCA model [8]. These EEIs are directly obtained without detail

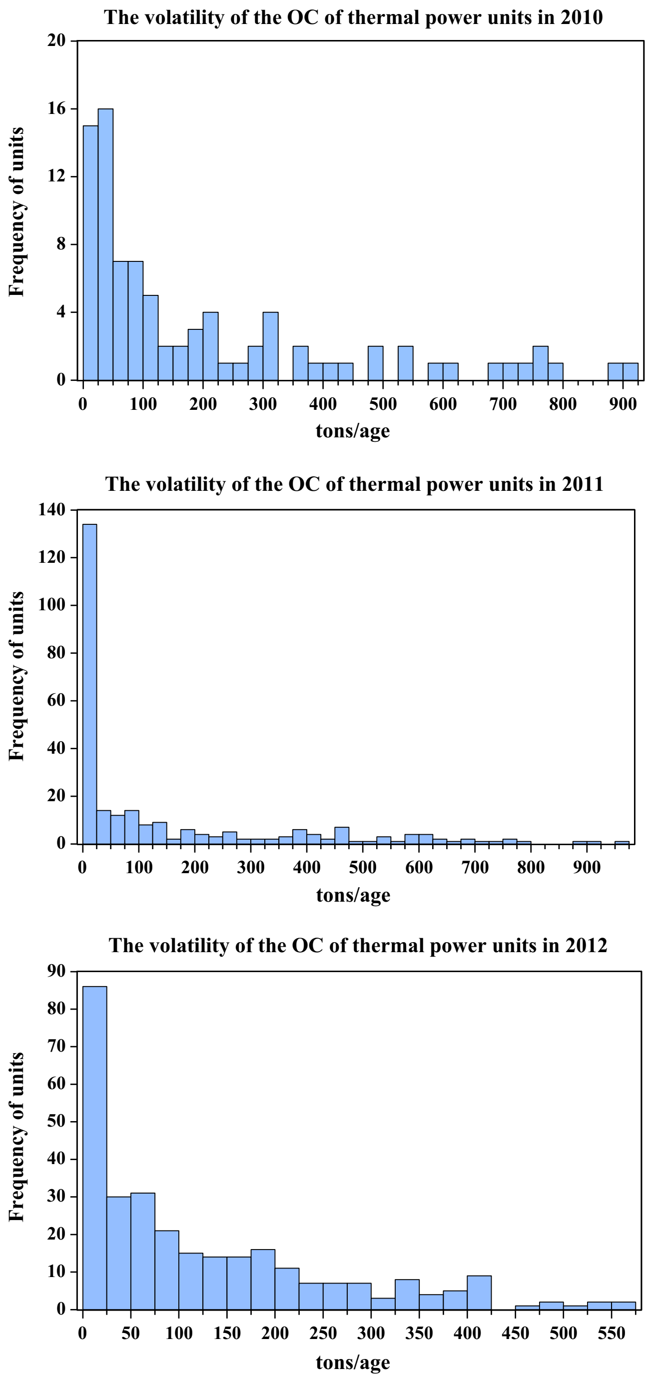


Fig. 5. The volatility of the OC of the 67 thermal power plants from 2010 to 2012.

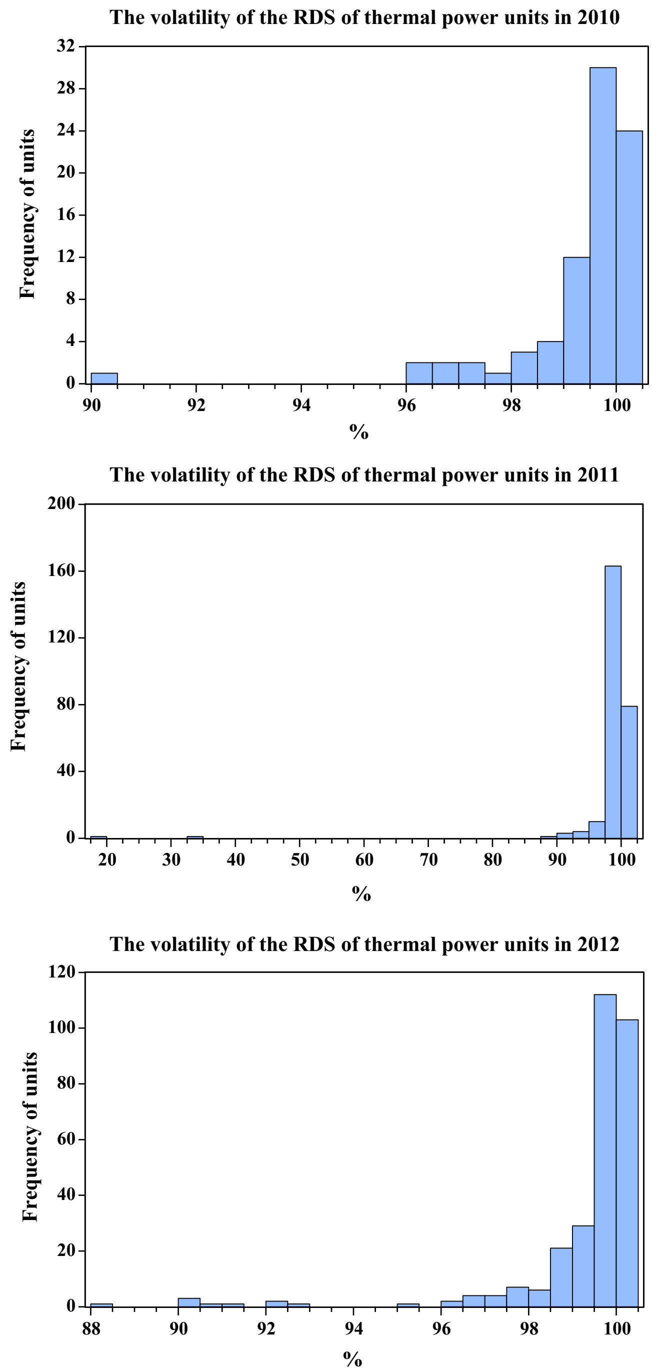


Fig. 6. The volatility of the RDS of the 67 thermal power plants from 2010 to 2012.

observation process. The new hybrid model in this paper provides "white-box" of each procedure.

4.4.2. Fixed effect regressive analysis estimated by the GARCH model, compared with VAR model

The purpose of this section is to investigate the volatility relation between the EEI and the five variables. Estimation results of GARCH model are summarised in Table 7. It demonstrates the evidence of a significant relation between the EEI and variables. The estimated coefficients from the mean equation of GARCH model have confirmed the result of the nonlinear regression model. As shown in Table 7, the most influential variables of energy efficiency in thermal power plants are the SCC and the EC. It presents that

Table 5
Correlation matrix of the variables.

	SCC	EC	WC	OC
Panel A: level				
EC	0.862***			
WC	0.721***	0.180***		
OC	0.314***	0.038***	0.361***	
RDS	0.032***	0.018***	0.019***	−0.017***
Panel B: first logarithmic change				
EC	0.853***			
WC	0.331***	0.366***		
OC	0.042***	0.024***	0.199***	
RDS	0.165***	0.433***	−0.007***	0.002***

*** Significance at the 1% level.

Table 6
Energy efficiency indicators change of China's thermal power plants 2010–2012.

Thermal power plants	2010	2011	2012	Cumulative
No. 1	1.263	0.972	0.938	1.058
No. 2	0.987	0.864	1.327	1.059
No. 3	1.245	0.709	0.737	0.897
No. 4	0.903	1.821	1.083	1.269
No. 5	0.966	0.579	1.241	0.929
No. 6	1.097	1.382	1.397	1.292
No. 7	1.083	1.327	0.725	1.045
No. 8	0.394	0.540	0.742	0.559
No. 9	0.672	0.693	0.651	0.672
No. 10	0.579	1.581	1.284	1.148
No. 11	0.392	1.028	0.865	0.762
No. 12	1.177	0.963	1.277	1.139
		....		
		....		
No. 65	0.890	0.886	0.794	0.857
No. 66	0.784	1.230	1.201	1.072
No. 67	1.380	1.371	2.103	1.618

Table 7
Estimation volatility result of GARCH model.

Variables	Volatility effects	Standard error	*Z*-Statistics
a	0.049	0.023	2.796
SCC	0.672***	0.004	0.801
EC	0.543***	0.056	−1.680
WC	0.291***	0.007	−1.392
OC	0.236***	0.001	0.457
RDS	0.483***	0.006	0.283

*** Significance at the 1% level.

higher EEI will cause a positive effect on the volatility of SCC, consequently leading to a high fluctuation of EC. The RDS is the second influential variable to the EEI. It can be found out that there is a substantial persistence in the RDS, as the coefficient associated with its own lag is significant. Both the WC and the OC are fairly correlated to the volatility of energy efficiency, with average slopes of 0.291 and 0.236. This finding indicates that a large amount of WC and OC may cause small fluctuation. The influence result of these variables is that they can reduce the uncertainty of volatility of energy efficiency in thermal power plants over the short-term.

The VAR model is further adopted to prove the accuracy of proposed hybrid model by making a fair comparison between their empirical estimations. It is a classical model which is often applied to examine the volatility and Granger causality among variables. Primary advantage of this model is that it can be used for aggregated level of integration. The estimation results of VAR model are summarised in Table 8, which is correspondent with the result of the proposed hybrid model. The estimation result of VAR model demonstrates that there is positive relation between the EC and the SCC. The obvious effect of the EC on the EEI is expected, and the result also presents that an increase in the RDS has a positive influence on the EC. Moreover, the WC and the OC have weak significant influence. Fig. 7 provides a test of root graph of VAR model. All the points are in a circular with a radius of 1. Based on the principle of VAR model, it proves that econometric analysis of the proposed model is perfectly stationary.

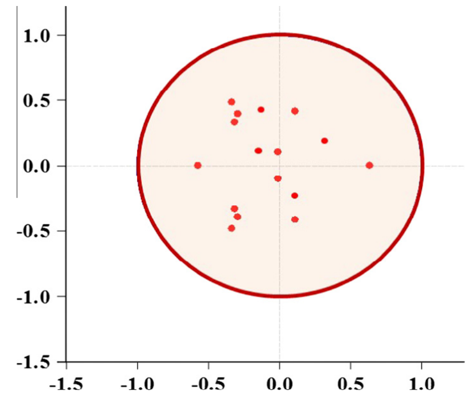


Fig. 7. Test of characteristic polynomial of VAR model.

Therefore, it can be summarised that the new econometric hybrid model is an appropriate model to evaluate short-term dynamics of thermal power units accurately. It contains more information on the degree of volatility among determinants and operating inefficiency. This has a surprise result that the variable RDS is positive and significant, which demonstrates that a wide use of desulfurization system will cause a rise of EEI.

5. Conclusion

This paper provides a new hybrid methodology for measuring energy efficiency of thermal power plants and explaining its short-term dynamics among plant-level physical indicators and environment constraint.

Four conclusions can be obtained based upon the study. First, the SCC is highly correlated with the EC. The OC is considered to be a new driving force affecting fluctuates in energy efficiency. Second, the EC has a positive impact on the volatility of energy efficiency. The higher EEI will lead the lower EC. Third, rise in the EDS drives up the EEI and the EC. Higher consumption on EC presents that plants installing desulfurization systems have good performances. Fourth, the results clearly demonstrate that the hybrid methodology is an accurate model through a comparison with the VAR model. These findings are useful for a multiple of managers. They should manage the energy use of plant-levels and pollution emission through environmental management facilities.

Table 8
Estimation result of VAR model.

	EEI	ln SCC	ln EC	ln WC	ln OC	ln RDS
EEI (−1)	5.573***	−0.829	−1.371	0.265	2.439***	1.674
ln SCC (−1)	−0.671***	1.610	1.566***	0.428	0.762***	−0.476***
ln EC (−1)	−1.283***	1.833***	−1.724	−0.114***	0.538	1.383***
ln WC (−1)	−0.613	0.197	0.186***	−0.293	0.329***	0.817
ln OC (−1)	−0.309	0.387***	0.201	−0.621***	−1.183	0.235
ln RDS (−1)	0.663***	−1.426	1.492***	−0.479	−0.238***	2.264
c	3.412	−3.342	2.546	−0.742	−0.892	10.162
R-squared	0.726	0.638	0.417	0.378	0.542	0.573
Adj. *R*-squared	0.549	0.528	0.491	0.264	0.298	0.287
Log likelihood	−211.547	115.511	51.882	−126.719	−241.039	−330.324

*** Significance at the 1% level.

In addition, two more major implications can be drawn based on the comparison of the hybrid model and the VAR model. First, the paper presents a competing methodology which is combining the multiple regression model and the GARCH model. It straight calculates the EEI of thermal power industry without artificial intervention. Second, it can eliminate the disturbance of changing outputs' structures. In other words, the new method will offer more information on the degree of volatility among determinants and operating inefficiency. Understanding the performance of units is significant in determining the benefits associated with a shift towards competition. Because of the feature of the GARCH method, adopting more explainable variables not only enlarge EEIs but also obviously estimate the volatility of energy efficiency.

Based on this study, some suggestions are listed as follows. First, due to the energy shortage and tough environment pollution, the standard regulation of water consumption in thermal power industry should be promoted, and the investment rate of desulfurization systems should be considered during the further performance benchmarking in thermal power industry. Second, this hybrid method could be adopted in different high-consuming industries for measuring EEIs. Third, the empirical results have relevance to Chinese's energy policymakers who are focusing the energy efficiency performance and aiming to reduce energy consumption in order to decrease the emission.

It should be mentioned that, the hybrid model proposed in the present work focuses on the evaluation of the energy efficiency of China's thermal power plants. Because the data from China Electricity Council is the only official resource to collect information of thermal power plants in China, the data used in the present work is, in some sense, limited. Further research work is under way in order to apply this hybrid methodology on different cases and improve it by extended data sources.

Acknowledgment

This work is supported by the National Key Basic Research Program of China (973 Program) (2013CB228304).

The authors gratefully acknowledge the support from National Energy Conservation Centre and anonymous reviewers for their valuable comments on this paper.

References

[1] Farrell MJ. The measurement of productive efficiency. J Roy Stat Soc 1957;3:253–90.
[2] Knight F. The economic organization. New York: Harper and Row; 1993.
[3] Debreu G. The coefficient of resource utilization. Econometrica 1951;19 (3):273–92.
[4] Shephard R. Cost and production functions. Princetion: Princeton University Press; 1953.
[5] Filippini M, Hunt LC. Energy demand and energy efficiency in the OECD countries: a stochastic demand frontier approach. Energy 2011;32(2):59–80.
[6] Zhou P, Ang BW, Zhou DQ. Measuring economy-wide energy efficiency performance: a parametric frontier approach. Appl Energy 2012;90 (1):196–200.
[7] Denholm P, Kulcinski GL, Holloway T. Emissions and energy efficiency assessment of baseload wind energy systems. Environ Sci Technol 2005;39:1903–11.
[8] Bernard JT, Cote B. The measurement of the energy intensity of manufacturing industries: a principal components analysis. Energy Policy 2005;33:221–33.
[9] Song CX, Li MJ, Zhang F, et al. A data envelopment analysis for energy efficiency of coal-fired power units in China. Energy Convers Manage 2015;102:121–30.
[10] Khoshnevisan B, Rafiee S, Omid M, et al. Reduction of CO_2 emission by improving energy use efficiency of greenhouse cucumber production using DEA approach. Energy 2013;55:676–82.
[11] Bianchi M, Branchini L, Cesari S, et al. Repowering existing under-utilized WTE power plant with gas turbines. Appl Energy 2015;160:902–11.
[12] Mousavi-Avval SH, Rafiee S, Jafari A, et al. Optimization of energy consumption for soybean production using Data Envelopment Analysis (DEA) approach. Appl Energy 2011;88(11):3765–72.
[13] Engle RF. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 1982;50:987–1007.
[14] Bollerslev T. Generalized autoregressive conditional heteroskedasticity. J Econ 1986;31:307–27.
[15] Andersen TG, Bollerslev T. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int Econ Rev 1998;39(4):885–905.
[16] Day TE, Lewis CM. Stock-market volatility and the information-content of stock index options. J Econ 1992;52(1):267–87.
[17] Poon SH, Granger CWJ. Forecasting volatility in financial markets: a review. J Econ Lit 2003;41(2):478–539.
[18] Alberola E, Chevallier J, Chèze B. Price drivers and structural breaks in European carbon price 2005–2007. Energy Policy 2008;36(2):787–97.
[19] Byun SJ, Cho H. Forecasting carbon futures volatility using GARCH models with energy volatilities. Energy Econ 2013;40:207–21.
[20] Chevallier J. Carbon futures and macroeconomic risk factors: a view from the EU ETS. Energy Econ 2009;31(4):614–25.
[21] Goto M, Karolyi GA. Understanding electricity price volatility within and across markets. Working Paper Series, 2004, Dice center working paper No. 2004-12.
[22] Sadorsky P. Correlations and volatility spillovers between oil prices and the stock prices of clean energy and technology companies. Energy Econ 2012;34:248–55.
[23] Elder J, Serletis A. Oil price uncertainty. J Money Credit Bank 2010;42:1137–59.
[24] Li MJ, He YL, Tao WQ. Modeling a hybrid methodology for evaluating and forecasting regional energy efficiency in China. Appl Energy 2015. http://dx.doi.org/10.1016/j.apenergy.2015.11.082.
[25] China Electricity Council, 2015. The national benchmarking of energy efficiency indicators for power industry, Available at: http://www.cec.org.cn/qghdnxdb.html [online].