Energy 35 (2010) 1709–1716

Contents lists available at ScienceDirect

Energy

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

Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India

Ujjwal Kumar [a,*], V.K. Jain [b]

[a] Flemish Institute for Technological Research (VITO), Environmental Modelling Unit, Boeretang 200, 2400 Mol, Belgium
[b] School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India

ARTICLE INFO

Article history:
Received 11 January 2009
Received in revised form
26 November 2009
Accepted 21 December 2009
Available online 27 January 2010

Keywords:
Energy consumption
GM (1, 1)
Grey-Markov
SSA

ABSTRACT

The present study applies three time series models, namely, Grey-Markov model, Grey-Model with rolling mechanism, and singular spectrum analysis (SSA) to forecast the consumption of conventional energy in India. Grey-Markov model has been employed to forecast crude-petroleum consumption while Grey-Model with rolling mechanism to forecast coal, electricity (in utilities) consumption and SSA to predict natural gas consumption. The models for each time series has been selected by carefully examining the structure of the individual time series. The mean absolute percentage errors (MAPE) for two out of sample forecasts have been obtained as follows: 1.6% for crude-petroleum, 3.5% for coal, 3.4% for electricity and 3.4% for natural gas consumption. For two out of sample forecasts, the prediction accuracy for coal consumption was 97.9%, 95.4% while for electricity consumption the prediction accuracy was 96.9%, 95.1%. Similarly, the prediction accuracy for crude-petroleum consumption was found to be 99.2%, 97.6% while for natural gas consumption these values were 98.6%, 94.5%. The results obtained have also been compared with those of Planning Commission of India's projection. The comparison clearly points to the enormous potential that these time series models possess in energy consumption forecasting and can be considered as a viable alternative.

1. Introduction

Forecasting energy demand constitutes a vital part of energy policy of a country, especially for a developing country like India whose energy demand is growing very rapidly. The energy consumption analysis also plays a crucial role in monitoring environmental issues such as green house gases emissions, air pollutants release in urban areas and the industrial region. With the rapid growth of economy, India has witnessed manifold increase in the energy consumption in the last three decades [1]. Since 1970–1971 to 2006–2007, the coal consumption has increased from 71.2 million tonnes (MT) to 462.7 MT, crude-petroleum consumption gone up from 18.4 MT to 146.5 MT, natural gas consumption rose from 0.64 Giga cubic meters (GCM) to 31.36 GCM while electricity consumption surged from 43.7 Tera Watt hour (TWh) to 443.1 TWh [1]. This mammoth energy consumption is also a great burden on Indian exchequer as almost 77% of crude-oil is being imported and petroleum and oil products constitute 33.8% of India's total import [2] (reference year is 2006–2007). Further, approximately 70% of electricity generation in India comes from thermal power plants which are mostly coal-based thus escalating the environmental problems with ever growing demand for electricity. In view of this, several studies have been conducted for energy demand analysis and forecasts. Parikh [3] presents a comprehensive report on modelling long term energy demand and energy policy implications for India. Sengupta [4] analyzes long term demand of India's commercial energy consumption and their long run supply by using several variants of regression models. Kulshreshtha and Parikh [5] used vector autoregressive (VAR) models with cointegrated variables to predict the demand for coal in four main sectors in India. Sharma et al. [6] formulated an econometric model using regression method to determine the demand for commercial energy namely, coal, petroleum products and electricity in different sectors in Kerala. Iniyan et al. [7] presents three models, namely, Modified Econometric Mathematical (MEM) model, Mathematical Programming Energy–Economy–Environment (MPEEE) model, and Optimal Renewable Energy Mathematical (OREM) model to project coal, oil and electricity consumption in India. Jebaraj et al. [8] compared various statistical models for energy consumption forecasting and concluded that Artificial neural network (ANN) is the best performing approach.

* Corresponding author. School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India. Tel.: +32 493479853,
E-mail addresses: ujjwalkumarin@yahoo.co.in, ujjwal.kumar@vito.be (U. Kumar).

0360-5442/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.energy.2009.12.021

Several rigorous studies have been conducted by various researchers for energy demand forecasting analysis for different countries. Bodger and Tay [9] used the logistic and energy substitution forecasting models to predict the electricity consumption in New Zealand using past consumption growth factor. Messner and Strubegger [10] presented a framework to analyze the consumption of energy (FACE), by considering growth factor, economics and technology as variables. Kaboudan [11] developed a non-linear dynamic econometric forecasting model to predict the electricity consumption in Zimbabwe through the year 2010 using 20 years of data. An exponential forecasting model had been developed to predict the Jordan's energy consumption by Tamimi and Kodah [12]. Al-Garni et al. [13] applied a forecasting regression model for the electrical energy consumption in Eastern Saudi Arabia, as a function of weather data, global solar radiation and population. Mackay and Probert [14] presented a modified logit-function demand forecasting model for predicting national crude-oil and natural gas consumptions based on saturation curve extrapolations for the appropriate energy intensity. Michalik et al. [15] used linguistic variables and fuzzy logic approach for the development of mathematical model to predict the energy demand in the residential sector. Chavez et al. [16] implemented univariate Box–Jenkins time series analyses (ARIMA) models to predict the energy production and consumption in Asturias, Northern Spain. Saab et al. [17] had investigated different univariate-modelling methodologies to forecast monthly electric energy consumption in Lebanon. Three univariate models were used namely, autoregressive, autoregressive integrated moving average (ARIMA) and a novel configuration combining an AR (1) with a high pass filter. Mackay and Probert [18] have developed a bottom–up technique-forecasting model to predict the supplies and demands of fluid fossil fuels for United Kingdom. Chow [19] discussed the sectoral energy consumption in Hong Kong for the period 1984–1997 with special emphasis on the household sector. Ediger and Tatlldil [20] used semi-statistical technique to formulate the forecasting model to predict the primary energy demand in Turkey and analysis of cyclic patterns. Aydinalp et al. [21] developed a neural network based energy consumption model for the Canadian residential sector. Huang et al. [22] developed a Grey-Markov forecasting model to forecast and analyze the electric-power supply and demand in China. The results showed that the forecast precisions were very high ranging from 97.56% to 99.42%. Akay and Atak [23] proposed an approach called Grey prediction with rolling mechanism (GPRM) to predict the Turkey's total and industrial electricity consumption. They found that the prediction by GPRM approach was better than the official Model of Analysis of the Energy Demand (MAED) by the Turkish Ministry of Energy and Natural Resources (MENR).

The present study attempts to forecast the consumption of conventional energy sources namely, coal, crude-petroleum, natural gas and electricity in India by investigating the time series pattern of each energy sources. Each of the energy consumption patterns exhibits their own unique characteristics and no single model can be uniformly applied to all time series. Therefore, in this study, based on the individual time series characteristics we have selected Grey prediction with rolling mechanism (GPRM) approach to forecast coal and electricity consumption in India while Grey-Markov model has been applied for crude-petroleum consumption and singular spectrum analysis (SSA) for natural gas consumption. Singular Spectrum Analysis (SSA) is relatively new in energy forecasting and has been applied for the first time to forecast energy consumption in this study. Section 2 briefly describes these modelling methodologies. Section 3 presents the results and discussion and Section 4 concludes the present study.

2. Modelling methodologies

Energy Consumption data usually presents interesting cases for time series modelling. If a time series energy consumption pattern is under consideration, this time series contains all the effects of all the factors that, in any way, affect the energy consumption pattern. Because if it is not so, that factor is irrelevant for energy consumption as it doesn't change the energy consumption pattern at all. It means, any change in energy consumption due to any factor should be seen in time series data/pattern and the idea of time series model is to capture such patterns. Most of the changes in the society are gradual, be it technological or behavioural, if these are changing and affecting the energy consumption, that must also be evident in time series energy consumption pattern. Time series models do attempt to capture such patterns to the greatest possible extent. The purpose of univariate time series analysis is "forecasting" based on the modelled time series pattern if we don't expect all of a sudden change in the scenario which is very much real. These forecasts can very well supplement the analysis dealing with energy policy issues. The further references exclusively on the significance of time series analysis can be found in various available excellent literatures [e.g., Priestley [24], Brockwell and Davis [25], Stoffer and Shumway [26]]. The present study identifies appropriate time series models based on individual time series pattern and subsequently forecasts have been made. In this section, we discuss the theories of three time series modelling approaches that have been identified and applied to forecast India's energy consumption.

2.1. Grey prediction with rolling mechanism (GPRM)

GPRM is a variant of GM (1, 1) Grey-forecasting model [27] that adopts the essential part of Grey system theory [28]. GM (1, 1) Grey forecasting can be used in circumstances with relatively little data ($n \geq 4$) [29]. The GM (1, 1) uses a first order differential equation to characterize an unknown system. A GM (1, 1) modelling algorithm is described below [23,29]:

Step I:
Assume that the original raw data series $y^{(0)}$ with n samples is expressed as:

$$y^{(0)} = \left[y^{(0)}(1), y^{(0)}(2), \ldots, y^{(0)}(n)\right], \quad n \geq 4$$

where the subscript (0) represents the original series. The original data are assumed to be positive. Negative values are prohibited in grey-modelling.

Step II:
A new series $y^{(1)}$ is generated by accumulated generating operation (AGO)

$$y^{(1)} = \left[y^{(1)}(1), y^{(1)}(2), \ldots, y^{(1)}(n)\right],$$

where

$$y^{(1)}(k) = \sum_{i=1}^{k} y^{(0)}(i), k = 1, 2, \ldots, n \tag{2.1.1}$$

Step III:
Establish a first order differential equation,

$$\frac{\mathrm{d}y^{(1)}}{\mathrm{d}t} + a_g y^{(1)} = u_g \tag{2.1.2}$$

where the parameters a_g and u_g are called the development coefficient and grey input, respectively. This grey model is referred as

GM (1, 1), in which the first number in the brackets denotes the order of differential equation and second number indicates the number of variables.

We define $z^{(1)}(k)$ as a sequence obtained by applying the MEAN operation to $y^{(1)}$

$$z^{(1)}(k) = \text{MEAN}\cdot y^{(1)} = \frac{1}{2}\left[y^{(1)}(k)+y^{(1)}(k-1)\right],\ k = 2,3,\ldots n$$

Generally the mean operation may be expressed as [30]:

$$z^{(1)}(k) = a\cdot y^{(1)}(k)+(1-a)\cdot y^{(1)}(k-1), \quad a\in[0,1] \tag{2.1.3}$$

Since the sampling time is 1, we have:

$$\frac{dy^{(1)}}{dt} = y^{(1)}(k)-y^{(1)}(k-1) \overset{\text{AGO}}{=} y^{(0)}(k) \tag{2.1.4}$$

By substituting equations (2.1.1, 2.1.3, 2.1.4) into equation (2.1.2), one has the grey differential equation.

$$y^{(0)}(k)+a_g z^{(1)}(k) = u_g \tag{2.1.5}$$

Equation (2.1.5) is solved by using least square method to obtain

$$\hat{y}^{(1)}(n+p) = \left(y^{(0)}(1)-\frac{u_g}{a_g}\right)\cdot e^{-a_g\cdot(n+p-1)}+\frac{u_g}{a_g} \quad n\geq 4$$

p is the step-size and $k + 1 = n + p$ is the time instant of the prediction.

Step IV:
Taking the inverse AGO on sequence $\hat{y}^{(1)}(k)$, we have:

$$\hat{y}^{(0)}(n+p) = \hat{y}^{(1)}(n+p)-\hat{y}^{(1)}(n+p-1)$$

$$\hat{y}^{(0)}(n+p) = \left(\hat{y}^{(0)}(1)-\frac{u_g}{a_g}\right)\cdot e^{-a_g\cdot(n+p-1)}\cdot(1-e^{a}) \tag{2.1.6}$$

this gives the predicted values.

In GM (1, 1), the whole data-set is used for prediction. It is, however, recommended to use only recent data to increase forecasting accuracy in future prediction [23]. GPRM is such a technique used in GM (1, 1) for ensuring this approach. In GPRM, $y^{(0)}(k+1)$ is predicted by applying GM (1, 1) to $y^{(0)} = (y^{(0)}(1), y^{(0)}(2), \ldots, y^{(0)}(k))$ where $k < n$. After the result is found, procedure is repeated, but this time the newly predicted entry $y^{(0)}(k+1)$ is added at the end of series, and the oldest data $y^{(0)}(1)$ is removed from series. Next, $y^{(0)} = (y^{(0)}(2), y^{(0)}(3), \ldots, y^{(0)}(k+1))$ is used to predict $y^{(0)}(k+2)$ [23].

2.2. Grey-Markov model

The forecasting precision of GM (1, 1) Grey-forecasting model for data sequence with large random fluctuations is low. Markov-chain forecasting model can be used to forecast a system with randomly varying time series. Therefore, Markov-chain can improve the GM (1, 1) model especially when fluctuation in data-set is large [27]. The Grey-Markov model has been adapted in this study after Huang et al. [22] and described below.

2.2.1. Partition of states by Markov-chain forecasting model

From step IV of previous subsection, a trend curve equation can be built

$$\hat{X}(k) = \hat{y}(k+1)\cdot ge^{-a_g k}$$

where $g = (y^{(0)}(1)-u_g/a_g)\cdot(1-e^{a_g}),\ k = 0,1,\ldots,(n-1).\hat{X}(k)$ denotes the Grey-model forecasted data.

The values of $y^{(0)}(k+1)$ are distributed in the region of the trend curve $\hat{X}(k)$ which may be divided into a convenient number of contiguous intervals. When $y^{(0)}(k+1)$ falls in interval i, one of S such intervals, it may be regarded as corresponding to a state E_i in an m order Markov unstable sequence, E_i can be signified as follows:

$$E_i = [E_{1i}, E_{2i}],$$

where $i = 1,2,\ldots,S$. S is the amount of states.

$$E_{1i} = \hat{X}(k)+A_i,$$

$$E_{2i} = \hat{X}(k)+B_i,$$

$\hat{X}(k)$ is a time function, so E_{1i} and E_{2i} will vary with the time series.

2.2.2. Calculating the transition probability P

For Markov-chain series, the transition probability from state E_i to E_j can be established using an equation as follows:

$$P_{ij}(m) = \frac{M_{ij}(m)}{M_i}, \quad (i,j = 1,2,\ldots,S).$$

where $P_{ij}(m)$ is the transition probability of state E_j transferred from state E_i for m steps. $M_{ij}(m)$ is the number of original data of state E_j transferred from state E_i for m-steps, M_i is the number of original data points in E_i.

These $P_{ij}(m)$ values can be presented as a transition probability matrix $R(m)$,

$$R_{ij}(m) = \begin{bmatrix} p_{11}(m) & p_{12}(m) & \ldots & p_{1j}(m) \\ p_{21}(m) & p_{22}(m) & \ldots & p_{2j}(m) \\ \ldots & \ldots & \ldots & \ldots \\ p_{i1}(m) & p_{i2}(m) & \cdot & p_{ij}(m) \end{bmatrix} \quad (i,j = 1,2,\ldots,S).$$

The state transition probability $p_{ij}(m)$ reflects the statistical law of each state transition in a system, which is foundation of Markov probability matrix forecast. The future development of the system can be forecasted by studying the state transition probability matrix $R(m)$. Generally, it is necessary to observe one-step transition matrix $R(1)$. Suppose the object to be forecasted is in state $E_Q(1\leq Q\leq S)$, row Q in matrix $R(1)$ should be considered. If max $P_{Qj}(1) = P_{QL}(1)(j = 1, 2,\ldots,S, (1\leq Q\leq S))$, then what will most probably happen in the next moment is the transition from state E_Q to state E_L.

2.2.3. Calculating the forecast data

After the determination of the future state transition of a system, i.e., the determination of Grey-elements E_{1j}, E_{2j} the changing interval of the forecast value is between E_{1j} and E_{2j}. The most probable forecast value $\hat{X}(k+1)$, is considered to be the middle value of the determined state interval, that is

$$\hat{X}(k+1) = 1/2(E_{1j}+E_{2j}) = \hat{X}(k)+1/2(A_i+B_i)$$

2.3. Singular spectrum analysis (SSA)

Five basic steps can be identified in SSA-forecasting: (I) Embedding the sampled time series in a vector space of dimension M; (II) Computing the $M\times M$ lag-covariance matrix $\mathbf{C_D}$ of the data; (III) Diagonalizing $\mathbf{C_D}$; (IV) Recovering the time series; and (V) forecasting. The procedure described in this section has been adapted after Elsner and Tsonis [31]

(I) Embedding the sampled time series in a vector space of dimension M;

One-dimensional time series $(x_1, x_2, \ldots, x_N)$ is represented as multidimensional time series as follows:

$$X = (x_{ij})_{i,j=1}^{k,M} = \begin{pmatrix} x_1 & x_2 & x_3 & \ldots & x_M \\ x_2 & x_3 & x_4 & \ldots & x_{M+1} \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ x_k & x_{k+1} & x_{k+2} & \ldots & x_N \end{pmatrix}$$

its dimension is known as *window length* (M). The matrix X is called *trajectory matrix*.

(II) Computing the $M \times M$ lag-covariance matrix $\mathbf{C_D}$ of the data;

Broomhead and King [32] proposed

$$(\mathbf{C_D})_{ij}^{(1)} = \frac{1}{N-M+1}\sum_{t=1}^{N-m+1} x_{i+t-1}x_{j+t-1}$$

Alternatively, Vautard and Ghil [33] suggested

$$(\mathbf{C_D})_{ij}^{(2)} = \frac{1}{N-|i-j|}\sum_{t=1}^{N-|i-j|} x_{|i-j|+t}x_t$$

$(\mathbf{C_D})^{(2)}$ is a Toeplitz matrix, meaning that all elements along each of diagonals are the same.

(III) *Diagonalizing* $\mathbf{C_D}$

Since $\mathbf{C_D}$ is real and symmetric, there exists a diagonalizing matrix E whose columns are orthonormal and a diagonal matrix $\mathbf{\Lambda}$ [34] such that

$$\mathbf{C_D} = \mathbf{E\Lambda E^T}$$

or

$$\mathbf{C_D E} = \mathbf{E\Lambda}$$

which is called the spectral decomposition of $\mathbf{C_D}$. Also $\mathbf{L}$ is a diagonal matrix whose nonnegative entries are the eigenvalues of $\mathbf{C_D}$. Using the definition of $\mathbf{C_D}$ we have

$$\mathbf{X^T X E} = \mathbf{E\Lambda},$$

$$\mathbf{E^T X^T X E} = \mathbf{\Lambda},$$

or

$$(\mathbf{XE})^{\mathbf{T}}(\mathbf{XE}) = \mathbf{\Lambda}.$$

The matrix $\mathbf{XE}$ is the trajectory matrix projected onto the basis $\mathbf{E}$. The components of $\mathbf{X}$ aligned along the basis $\mathbf{E}$ are uncorrelated since $\mathbf{E}$ is composed of orthogonal vectors $\mathbf{E_k}$ called the singular vectors of $\mathbf{X}$. The diagonal matrix $\mathbf{\Lambda}$ consists of ordered values $0 \le \lambda_1 \le \lambda_2 \le \ldots \le \lambda_M$ whose square roots are called the singular values that are referred to collectively as the *singular spectrum*. These terms give SSA its name.

(IV) *Recovering the time series*

The Principal Components (PCs) of original time series can be calculated as follows:

$$a_i^k = \sum_{j=1}^{m} x_{i+j-1}e_j^k$$

for i = 1, 2,….,N and k = 1, 2,….., M where represents the jth component of the kth eigenvector.

To get back to the original time series we convolve these principal components with their associated eigenvectors, which amount to

$$x_{i+j-1} = \sum_{k=1}^{m} a_i^k e_j^k$$

where i = 1, 2, 3,……., N and j = 1, 2, 3,…., M

The original time series can be filtered this way by using only *a selection of the possible principal components.*

(V) *Iterative approach to forecast*

This strategy was introduced by Zhang et al. [35] and Zhang and Pan [36]. To obtain a prediction of the next future value for the time series we add to the observed record an initial guess $\hat{x}_0$ of the prediction – for example, the mean. Above SSA procedure is repeated from the augmented time series having $N_t + 1$ values. An updated prediction of the future value is made by letting

$$\hat{x}_{(1)} = x'_{N_t+1} = \sum_{k=1}^{d_s} a_N^k e_m^k,$$

where a_N^k is the Nth value (in this case $N = N_t - m + 2$) of the kth principal component and e_M^k is the mth component of the kth eigenvector of the augmented trajectory matrix.

The updated prediction $\hat{x}_{(1)}$ is then used to replace $\hat{x}_{(0)}$ in the augmented record and the process is iterated until

$$|\hat{x}_{(q)} - \hat{x}_{(q-1)}| < \varepsilon \quad \text{for } q = 1, 2, \ldots.$$

Upon convergence $\hat{x}_{(q)}$ is accepted as the predicted value for $\hat{x}_{N_t+1}$.

3. Results and discussion

All the computations required for the present study have been carried out on MATLAB® 7.1 workstation by preparing appropriate MATLAB routines. The data used in the present study have been procured from Energy-Statistics (2007) [1] and Energy-Statistics (2008) [37] maintained by Ministry of Statistics and Programme Implementation, Govt. of India. In case of discrepancy between Energy-Statistics (2007) [1] and Energy Statistics (2008) [37] data, newer data, i.e., Energy Statistics (2008) [37] has been preferred.

Figs. 1–4 show the annual consumption of coal, electricity (utilities), crude-petroleum and natural gas, respectively, in India since 1970–1971 to 2005–2006. An examination of Figs. 1–4 clearly reveals that there have been tremendous rise in energy consumption for each energy sources but the temporal consumption pattern for each of the energy sources vary. Accordingly, the models applied on each of the time series also vary. Section 3.1 deals with the coal and electricity (utilities) consumption. Section 3.2 discusses crude-petroleum consumption while Section 3.3 elaborates upon natural gas consumption.

3.1. Forecasting coal and electricity (utilities) consumption

An examination of Figs. 1 and 2 suggest that there is nearly a steady exponential kind of growth pattern in these time series. Such patterns are very much suitable for grey GM (1, 1) model. We have applied Grey-model with rolling mechanism in both the cases.

Recent 6 years data (x_k, x_{k-1}, x_{k-2},…, x_{k-5}) has been used to predict (k + 1)th data in case of coal. The calculated $(k+1)$th value has now been made part of new series of 6 data that consists of calculated (k + 1)th value and (x_{k+1}, x_k, x_k,…, x_{k-4})th values (it is

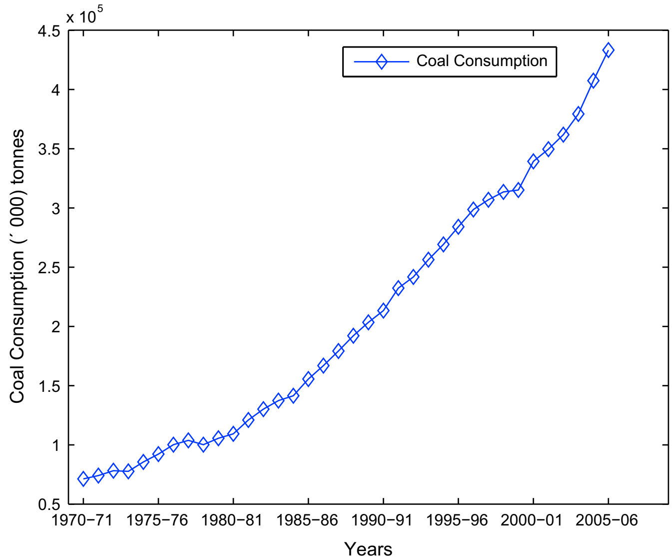

Fig. 1. Time series of Coal Consumption ('000 tonnes) in India since 1970–1971 to 2005–2006.

notable that x_{k-5} has now been excluded) to predict (k + 2)th value. This way two out of sample of forecasts have been calculated using each possible time series of 6 lagged data. The absolute percentage error for each out of sample forecasts has been calculated. Thus, taking mean of all out of sample forecasts gives us mean absolute percentage error (MAPE) for all out of sample forecasts which turned out to be 3.5% in case of coal consumption. The MAPE of in-sample values or data-fitting was 0.98%. The prediction accuracy ([22,23]) for the last two out of sample data (i.e., for 2004–05 and 2005–06) were found to be 97.9% and 95.4% which are reasonably good. The coal consumption forecasts for next 10 years till 2015–2016 have been shown in Fig. 5. Fig. 5 also shows the last two out of sample forecasts in red colour.

In case of Electricity power consumption, recent five years data ($x_k, x_{k-1}, \ldots, x_{k-4}$) have been used to predict $(k+1)$th data. As done in case of coal consumption, two out of sample forecasts have been calculated using each possible time series of 5 lagged data in case of electricity consumption. The mean absolute percentage error (MAPE) turned out to be 3.4% in case of electricity consumption. The MAPE of in-sample values or data-fitting was 0.99%. The forecast precisions for the last two out of sample data (i.e., for 2004–2005 and 2005–2006) were found to be 96.9% and 95.1%. The electricity (in utilities) consumption forecasts for the next 10 years (till 2015–2016) have been depicted in Fig. 6 along with last two out of sample forecasted data in red.

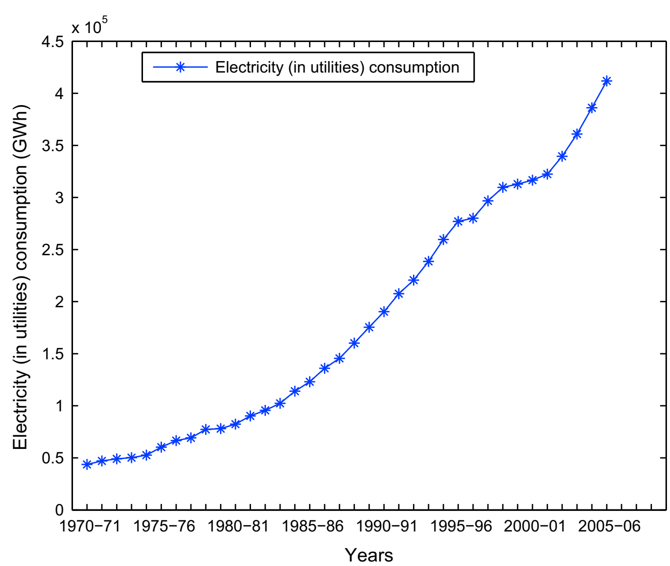

Fig. 2. Time series of Electricity (utilities) Consumption (GWh) in India since 1970–1971 to 2005–2006.

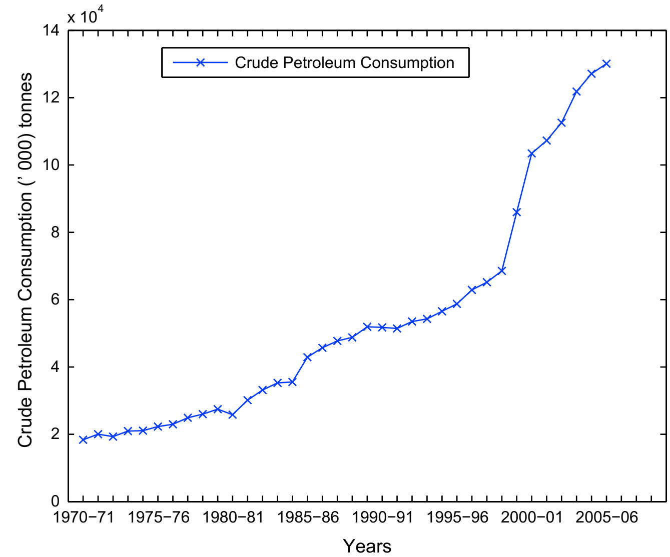

Fig. 3. Time series of crude-petroleum consumption ('000 tonnes) in India since 1970–1971 to 2005–2006.

3.2. *Forecasting crude-petroleum consumption*

A visual inspection of Fig. 3 reveals that although crude-petroleum consumption has a rising trend, there are great turns/fluctuations throughout the time series. These fluctuations in time series greatly reduce the forecasting ability of Grey GM (1, 1) model even if it is used by applying rolling mechanism. Since Markov-chain forecasting model can be used to forecast a system with randomly varying time series, we have applied Grey-Markov model for this time series.

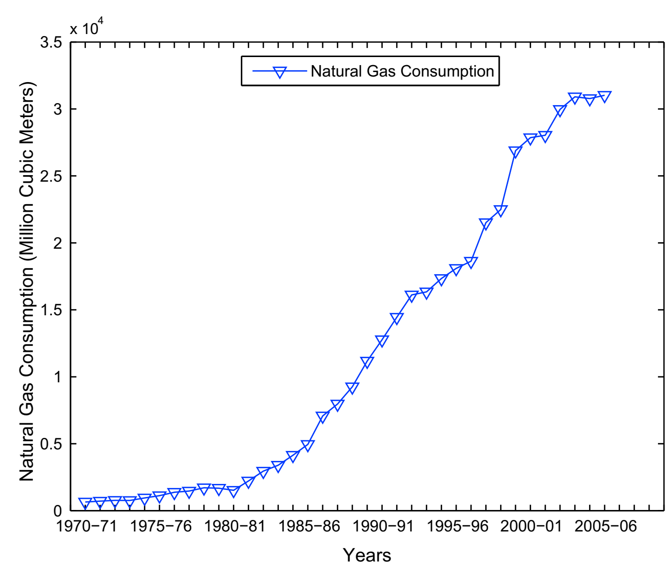

Fig. 4. Time series of natural gas consumption (million cubic meters) in India since 1970–1971 to 2005–2006.

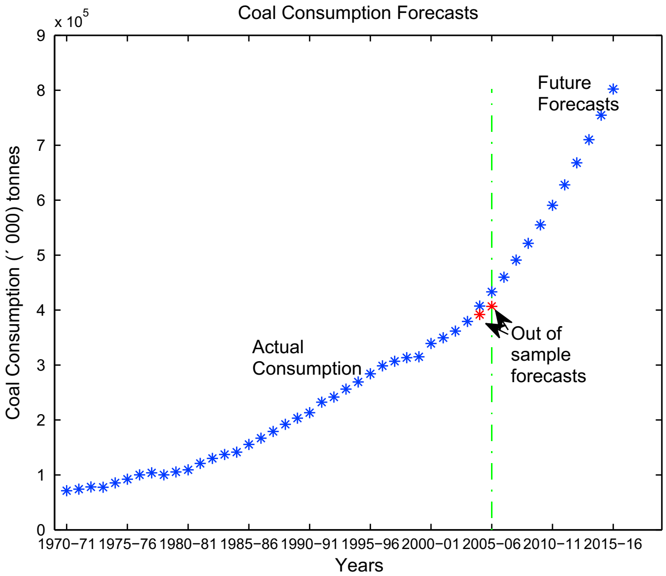

Fig. 5. Coal Consumption forecasts for next 10 years till 2015–2016. Two red star (*) represents the last two out of sample forecasts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

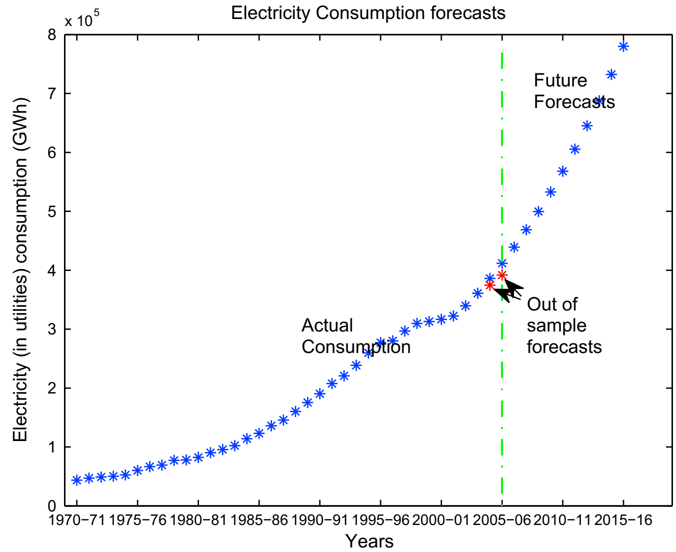

Fig. 6. Electricity(in utilities) Consumption forecasts for next 10 years till 2015–16. Two red star (*) represents the last two out of sample forecasts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

A trend curve equation is built by Grey-forecasting model GM (1, 1) for the time series of crude-petroleum consumption. This trend curve has been shown as green (middle) curve in Fig. 7. For Markov-chain forecasting model, partitions of states were done by establishing six contiguous intervals. An inspection of Fig. 7 reveals that the number of historical data in each interval are as follows (excluding the data of the years 2004–2005 and 2005–2006 which have been kept out of model building so as to check the validity of model for out of sample forecasts):

$$M_1 = 2; M_2 = 3; M_3 = 4; M_4 = 16; M_5 = 5; M_6 = 4;$$

Where M_i denotes the number of historical data in the interval i, and $i = 1, 2, 3, 4, 5, 6$. The numbers of one-step transiting to E_1, E_2, E_3, E_4, E_5, E_6 from E_1 are as follows:

$$M_{11}(1) = 1; M_{12}(1) = 0; M_{13}(1) = 0; M_{14}(1) = 1; M_{15}(1) = 0; M_{16}(1) = 0.$$

Next $M_{ij}(1)$, where $i = 2, 3, 4, 5, 6$ and $j = 1,2,3,4, 5, 6$ can be calculated in the same way. Then one-step transition probability for every state can be calculated by using the formula $M_{ij}(1)/M_i$. The calculated probability values have been presented in the one-step transition matrix as follows:

$$R(1) = \begin{bmatrix} 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 2/4 & 1/4 & 0 & 0 \\ 0 & 0 & 2/16 & 12/16 & 1/16 & 1/16 \\ 0 & 0 & 0 & 1/5 & 4/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3/4 \end{bmatrix}$$

To calculate the out of sample forecast, we observe the state of the last data. For example, to calculate the value of 2004–2005, we observe that 2003–2004 data lies in state 6, the sixth row of transition matrix suggests that most likely transition is E_{66}, i.e., the value of 2004–2005 lies between 123 380 and 127 940 thousand tonnes. The forecasted value for 2004–2005 would be 125 660 tonnes. Similarly, by obtaining two-step transition matrix, the forecast for 2005–2006 has been calculated as 133,240 thousand tonnes.

The mean absolute percentage error for both out of sample forecasts is found to be 1.6%.

The prediction accuracy obtained for both the years are 99.2% and 97.6%, respectively. The forecasts for next 10 years crude-petroleum consumption has been represented in Fig. 8 along with last two out of sample forecasts in red.

3.3. Forecasting natural gas consumption

The variation of annual consumption of natural gas (Fig. 4) is very peculiar among the four conventional energy sources consumption pattern. Although, looking superficially, natural gas consumption pattern can be considered as having overall a rising trend, a close inspection clearly reveals that many of the years the consumption of natural gas remained almost stagnant or even

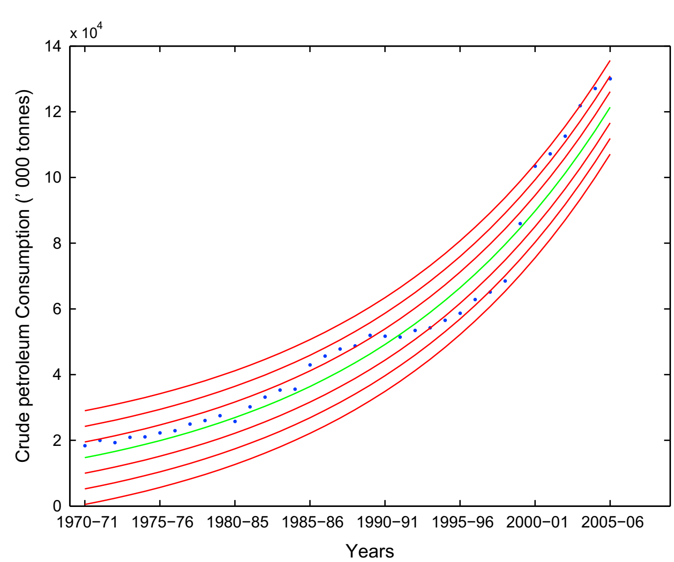

Fig. 7. Partition of states into six contiguous intervals for the time series of Crude-Petroleum Consumption. Dots represent the original data. The green (middle) curve shows the obtained trend curve by GM (1, 1) model.

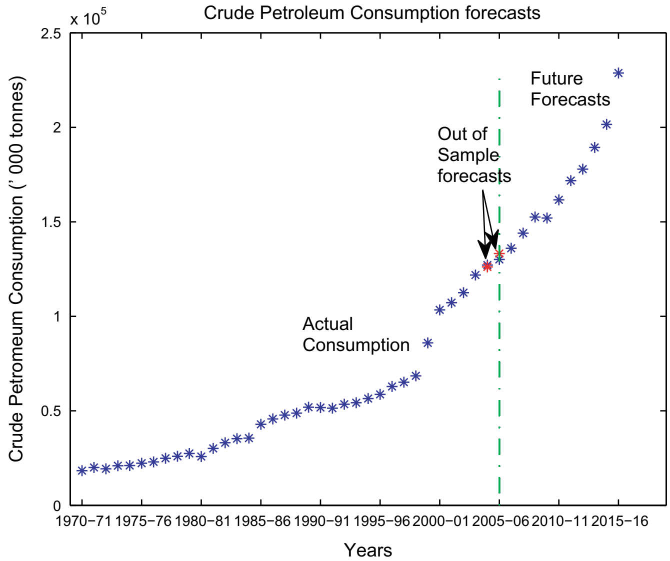

Fig. 8. Crude-petroleum Consumption forecasts for next 10 years till 2015–2016. Two red star (*) represents the last two out of sample forecasts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

decreased in some of the years. The GM (1, 1) Grey-forecasting model is prone to produce large amount of errors for such consumption pattern as grey-model supports only exponential kind of growth. Since fluctuation of growth pattern is not around GM (1, 1) trend curve, the application of Grey-Markov model is also not feasible because suitable partition of states is practically not viable. However, singular spectrum analysis (SSA) is a data adaptive technique (Elsner and Tsonis [31], 1996) and therefore it has potential to capture such variations and can prove extremely promising for short-term forecasts.

In this case also, last two data has been kept out of model-fitting to validate the out of sample forecasts. Window length has been taken as 4. The selection of window length is trivial. The main idea in selecting window length is that possible variations or periodicity is covered within this length of time series because SSA captures mainly the patterns in a time series that are contained within the window length. However, it is notable that variations of the window length about a sufficient large M only stretch or compress the spectrum of eigenvalues, leaving the relative magnitudes of the individual eigenvalues unchanged [31]. The eigenvalues of the present time series at window length 4 has been shown in Fig. 9. It is clear that there is one dominant eigenvalue and rest eigenvalues are comparatively very small.

In the process of recovering time series [step IV of Section 3.2.(iii)] we have selected three principal components. The convergence error ϵ has been set at 0.001. The iterative forecasts for the two out of sample values were found to be 30 348 and 29 135 million cubic meters while actual values were 30 775 and 31 025 million cubic meters. The mean absolute percentage error is 3.4% while prediction accuracies are 98.6% and 94.5%. The natural gas consumption forecasts for next 10 years (till 2015–2016) have been shown in Fig. 10 along with last two out of sample forecasts.

3.4. Discussion

The results obtained in this study can be compared with those of Planning Commission of India [Planning Commission (2002–2007) [38] and Planning Commission (2007–2011) [39]]. Planning Commission (2002–2007) [38] provides estimated energy demand for 2006–2007 and this has been compared with the model's forecasts presented in this study (Table 1). Table 1 indicates that the forecasts for Coal and Oil (Crude-petroleum) are comparable, however, natural gas consumption is grossly overestimated in Planning Commission (2002–2007) [38] study and hence forecast precision is quite low.

Planning Commission (2007–2011) [39] projects the coal consumption requirement to be increased by 35% till 2011–2012 from the year 2006–2007 while the present study puts this growth at 36.8%. In case of natural gas consumption, Planning Commission (2007–2011) [38] puts this growth at 38.7% while the present study puts this figure at 11.7%. Oil consumption is expected to grow by 40.1% till 2011–2012 by Planning Commission (2007–2011) while present study projects oil consumption requirement to grow by 30% till 2011–2012 from 2006 to 2007.

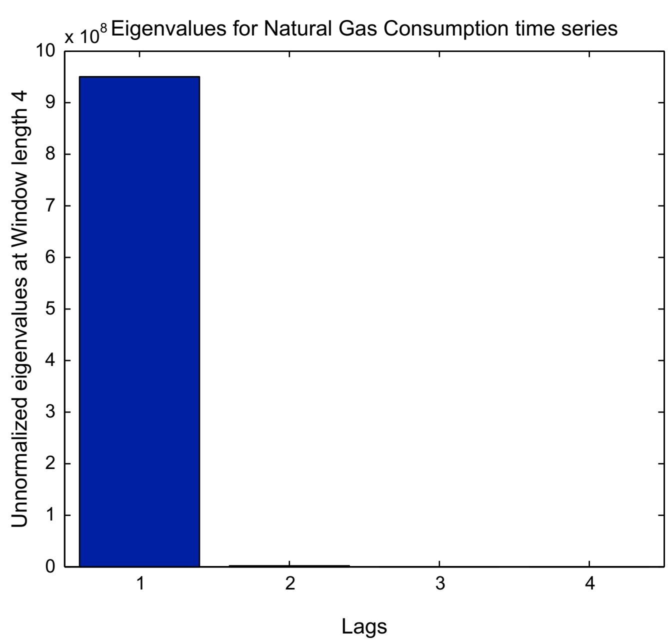

Fig. 9. Eigenvalues at window length 4 for the time series of Natural Gas Consumption.

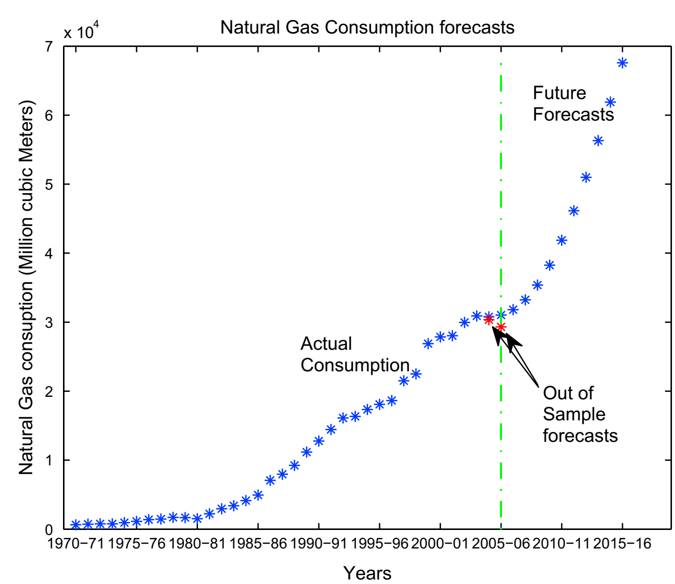

Fig. 10. Natural Gas Consumption forecasts for next 10 years till 2015–16. Two red star (*) represents the last two out of sample forecasts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 1
A Comparison of projected energy demand for 2006–2007 by the models of the present study and the Planning Commission (2006–2007).

Energy Sources	Energy consumption forecast precision for 2006–2007	
	Models in this study	Planning Commission (2002–2007)
Coal	99.5%	99.6%
Oil	92.8%	91.7%
Natural gas	98.6%	48.7%

4. Conclusion

The present study has explored the potential of three time series models in energy consumption forecasting for India. For short-term forecasting, the results clearly indicate that values forecasted by these models are very much close to the real values. A comparison of projected energy demand by the models presented in this study and the Planning Commission (2002–2007) [Table 1] shows that these time series models can be a viable alternative to project the future energy requirements. However, it should be noted that appropriate time series models should be applied to each of the time series. Normally, energy consumption pattern has rising trend for most of the countries. Therefore, these time series models have potential for wider application. They can be very helpful and supplementary to frame suitable energy policy.

References

[1] Energy Statistics, http://mospi.nic.in/es07_main.htm; 2007 [accessed Dec 2008].
[2] Economic Survey. Government of India, Ministry of Finance, Economic Division. India: Oxford University Press; 2007–08.
[3] Parikh JK. Modelling Energy Demand for Policy Analysis. New Delhi: Govt. of India, Planning Commission, http://planningcommission.gov.in/reports/publications/pub_enrgy1981.pdf; 1981 [accessed Oct 2009].
[4] Sengupta R. Energy Modelling for India. New Delhi: Planning Commission, Govt. of India, http://planningcommission.gov.in/reports/publications/pub_enrgy1993.pdf; 1993 [accessed Oct 2009].
[5] Kulshreshtha M, Parikh JK. Modelling demand for coal in India: vector autoregressive models with cointegrated variables. Energy 2000;25:149–68.
[6] Sharma PD, Chandramohanan Nair PS, Balasubramanian R. Demand for commercial energy in the state of Kerala, India: an econometric analysis with medium-range projections. Energy Policy 2002;30:781–91.
[7] Iniyan S, Suganthi L, Samuel AA. Energy models for commercial energy prediction and substitution of renewable energy sources. Energy Policy 2006;34:2640–53.
[8] Jebaraj S, Iniyan S, Kota H. Forecasting of commercial energy consumption in India using artificial neural network. Int J Global Energ Issues 2007;27:276–301.
[9] Bodger PS, Tay HS. Logistic and energy substitution models for electricity forecasting: a comparison using New Zealand consumption data. Technol Forecast Social Change 1987;31:27–48.
[10] Messner S, Strubegger M. A framework to analyse the consumption of energy. In: Proceedings of the workshop of economics, modelling, planning and management of energy, Trieste, Italy; 1987. p. 315–28.
[11] Kaboudan M. An econometric model for Zimbabwe's future electricity consumption. Energy 1989;1:75–85.
[12] Tamimi A, Kodah Z. Modelling of energy consumption and projection in Jordan. Energy Convers Manage 1993;34:427–32.
[13] Al-Garni AZ, Zubair SM, Nizami JS. A regression model for electric-energy consumption forecasting in Eastern Saudi Arabia. Energy 1994:1043–9.
[14] Mackay RM, Probert SD. Modified logit-function demand model for predicting natural crude-oil and natural-gas consumptions. Appl Energy 1994;49:75–90.
[15] Michalik G, Khan ME, Bonwick WJ, Mielczarski W. Structural modeling of energy demands in the residential sector: 1. Development of structural models. Energy 1997;22:937–47.
[16] Chavez GS, Xiberta Bernat J, Llaneza Coalla H. Forecasting of energy production and consumption in Asturias (Northern Spain). Energy 1999;24:183–98.
[17] Saab S, Badr E, Nasr G. Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. Energy 2001;26:1–14.
[18] Mackay RM, Probert SD. Forecasting the United Kingdom's supplies and demands for fluid fossil-fuels. Appl Energy 2001;69:161–89.
[19] Chow LC-h. A study of sectoral energy consumption in Hong Kong (1984–97) with special emphasis on the household sector. Energy Policy 2001;29:1099–110.
[20] Ediger VS, Tatlldil H. Forecasting the primary energy demand in Turkey and analysis of cyclic patterns. Energy Convers Manage 2002;43:473–87.
[21] Aydinalp M, Ismet Ugursal V, Fung AS. Modelling of the appliance, lighting and space-cooling energy consumptions in the residential sector using neural networks. Appl Energy 2002;71:87–110.
[22] Huang M, He Y, Cen H. Predictive analysis on electric power supply and demand in China. Renew Energy 2007;32:1165–74.
[23] Akay D, Atak M. Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. Energy 2007;32:1670–5.
[24] Priestley MB. Spectral analysis and time series. New York: Academic Press; 1981.
[25] Brockwell PJ, Davis RA. Introduction to Time Series and Forecasting. New York: Springer-Verlag; 2002.
[26] Shumway RH, Stoffer DS. Time series analysis and its applications with R examples. New York: Springer-Verlag; 2005.
[27] He Y, Bao YD. Grey-Markov forecasting model and its application. Syst Eng 1992;9(4):59–63 [Theory Practice].
[28] Deng JL. Control problems of grey system. Wuzhang University of Science and Technology Press; 1990.
[29] Dounis AI, Tiropanis P, Tseles D, Nikolaou G, Sycros GP. A comparison of grey model and fuzzy predictive model for time series. International Journal of Computational Intelligence 2005;2:176–81.
[30] Chang T.-C., Wen K.-L., Chang H.-T., You M.-L. Inverse approach to find an optimum for grey prediction model. In: IEEE International Conference on SMC, Oct 1999. Tokyo Japan, 1999: p. III (309–313).
[31] Elsner JB, Tsonis AA. Singular spectrum analysis – a new tool in time series analysis. New York and London: Plenum Press; 1996.
[32] Broomhead DS, King GP. Extracting qualitative dynamics from experimental data. Physica 1986;20D:217–36.
[33] Vautard R, Ghil M. Singular-spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. Physica D 1989;35:395–424.
[34] Press WH, Teukolsky SA, Vellerling WT, Flannery BP. Numerical recipes in C++: the art of scientific computing. Cambridge: Cambridge University Press; 2002.
[35] Zhang B, Lie J, Sun Z. A new multidimensional time series forecasting method based on the EOF iteration scheme. Adv Atmos Sci 1993;10:243–7.
[36] Zhang B, Pan J. Forecasts of tropical Pacific SST anomalies using a statistical (EOF) iteration model. Exp Long-Lead Forecast Bull 1995;4:38–40.
[37] Energy Statistics, http://mospi.nic.in/es08_main.htm; 2008 [accessed Nov 2009].
[38] Planning Commission. 10th Five year plan. Govt. of India, http://planningcommission.gov.in/plans/planrel/fiveyr/welcome.html; 2002–07 [accessed Nov 2009].
[39] Planning Commission. 11th Five year plan. Govt of India, http://planningcommission.gov.in/plans/planrel/fiveyr/welcome.html; 2007–12 [accessed Nov 2009].