Zagros J. Appl. Math. & Data. Anal.

Available online at https://zjamda.abru.ac.ir/
ISSN 0000-0000
Vol. 1, No. 1, 2025, 32–59
Research Article

Forecasting the Total Energy Consumption in the United States: A Semiparametric Markov Switching Approach

Younes Nademi*
Department of Economics, Faculty of Humanities, Ayatollah Boroujerdi University, Boroujerd, Iran

Arash Nademi
Department of statistics, Ilam branch, Islamic Azad University, Ilam, Iran

Abstract

Abstract: We applied a semiparametric Markov switching AR-ARCH (SMSARCH) model to forecast the total U.S. energy consumption in the residential, commercial, industrial, transportation and electric power sectors. For this purpose, we compared several SMSARCH models containing different core functions with the models such as ARIMA, GARCH, EGARCH, Markov switching in mean and GARCH based on their abilities to forecast the total energy consumption. The time period from January 2000 to December 2015 was used for the in-sample estimation, while the period for the out-of-sample forecasting was from January 2016 to December 2016. The root mean square error (RMSE) criterion for both in-sample and out-of-sample periods indicates that the forecasting abilities of the SMSARCH models in all the U.S. energy sectors are better than those of the other studied parametric models. Furthermore, the results of Diebold and Mariano test showed that there is a significant difference between the values of RMSE for all models.

Keywords: Forecasting, Total Energy, Semiparametric Markov Switching Models, Total Energy.

1 Introduction

Energy is one of the most important sources for the development of economy [17]. Therefore, energy supply is crucial for energy policy makers. Also, the fossil energy consumption plays a main role in the environmental pollution and threatens the public health [16]. The United States has the highest

*Corresponding author: younesnademi@abru.ac.ir

energy consumption and the biggest environmental polluters of the entire world. In December 2015, the Paris climate change agreement dealing with the mitigation of greenhouse gas emissions was adopted by 196 countries. The agreement has been signed by 195 countries in 2016. However, with the withdrawal of the United States from the agreement in 2017, many countries condemned this decision made by the U.S. government. Therefore, the future of the U.S. energy consumption is so important for the climate change and global warming problem. Hence, a precise forecasting of energy consumption has an important role on making energy and environmental policies in the U.S. and world economy. It is very difficult to forecast energy because many internal and external factors affect the energy consumption [28,30]. Researchers have used two general forecasting methods including time series and multifactor-influenced methods to achieve a precise forecast [30]. Although there are many studies on the energy consumption forecasting in developing countries especially in China, these studies in the developed countries especially in the United States are rare. For instance, Saab et. al [25], Hamzacebi [12], Adams and Shachmurove [2], Lee and Tong [17], Pao and Tsai [23], Sen et. al [26], Barak and Sadegh [3], Xu et. al [29], Sutthichaimethee, Ariyasajjakorn [27] and Kaboli et. al [15] and Wu. Et. al [32] forecast the energy consumption in developing countries specially in China. Craig et. al [6] reviewed the long-term energy forecast in the United States. They concluded that there are many limitations in energy forecast as a result of the exogenous shocks such as oil crisis in the 1970s. They suggest making an assumption about technologies and social systems for modeling energy consumption to achieve a precise energy forecast. Lee and Tong [16] forecasted the energy consumption in the United States and China by a hybrid dynamic model which combines a dynamic Grey model with genetic programming. They conclude that the hybrid dynamic model is more accurate and reliable than other competitive models including ARIMA, Grey Model (GM), Dynamic Grey Model (DGM) and Genetic Programming (GP) Model. Wang et.al [31] has predicted the residential solar energy consumption of the United States by a new grey model based on data grouping and buffer operator. Their results show that their proposed model has more efficiently identify the seasonal fluctuation and structural mutation of the residential solar energy consumption data.

There is no study on forecasting energy consumption in all different sections of the U.S. economy. Also, there is no study using a semiparametric Markov switching model to forecast the energy consumption. Therefore, because of the lack of investigations on forecasting energy consumption in the United States and the importance of the U.S. energy consumption in the world economy we were encouraged to write this paper.

In this paper, we used a semiparametric Markov switching AR-ARCH -(SMSARCH)- model introduced by Nademi and Farnoosh [19] to forecast the total U.S. energy consumption in residential, commercial, industrial, transportation and electric power sectors.

2 Methodology and Data

Consider $Z_1, \dots, Z_N$ as part of a stationary time series that are made by the following Markov switching model where the Switching between the states is governed by a hidden Markov chain F_t

with values in $S = \{1, \ldots, M\}$.

$$Z_t = \sum_{k=1}^{M} \lambda_{tk}(m(Z_{t-1}, Z_{t-2}; \rho_k) + \sigma(Z_{t-1}, Z_{t-2}; \omega_k, \alpha_k, \beta_k)\epsilon_{t,k}), \tag{1}$$

such that

$$m(Z_{t-1}, Z_{t-2}; \rho_k) = f_k(Z_{t-1}) + \rho_k(Z_{t-1} - f_k(Z_{t-2})), \tag{2}$$

and

$$\sigma^2(Z_{t-1}, Z_{t-2}; \omega_k, \alpha_k, \beta_k) = \omega_k + \alpha_k Z_{t-1}^2 + \beta_k Z_{t-2}^2,$$

and

$$\lambda_{tk} = \begin{cases} 1 & F_t = k, \\ 0 & \text{otherwise}, \end{cases}$$

where $\lambda_t = (\lambda_{t1}, \lambda_{t2}, \ldots, \lambda_{tM})^T$ are random with values of the unit vectors and the residuals $\epsilon_{t,k}, t = 1, \ldots, N, k = 1, \ldots, M$ are i.i.d. random variables with mean 0 and variance 1. F_t is an aperiodic, irreducible stochastic process and according to stationary processes the distribution of the hidden state process is given by the $M \times M$ transition probability matrix Δ, i.e, $\Delta_{jk} = pr(F_t = k|F_{t-1} = j)$. We can obtain the stationary distribution by solving the equation $\pi\Delta = \pi$ where $\pi = (\pi_1, \ldots, \pi_M)$ and $\pi_k = pr(F_t = k)$.
In the Markov switching models, in addition to the estimation of the parameters, the main goal is approximating the unknown function $f_k(x)$ in semiparametric framework in which $f_k(x) \in \{g(x, \theta_k)\xi_k(x); \theta_k \in V\}$ where $\xi_k(x)$ is a nonparametric adjustment factor, and $g(x, \theta_k)$ is a known function of x and θ_k and $V \subseteq R^p$ is the parametric space. Therefore, we can rewrite mean function (2) in the following form:

$$m(Z_{t-1}, Z_{t-2}; \theta_k, \rho_k) = g(Z_{t-1}, \theta_k)\xi_k(Z_{t-1}) + \rho_k(Z_{t-1} - g(Z_{t-2}, \theta_k)\xi_k(Z_{t-2})).$$

We consider $g(x, \theta_k)$ as the core function of semiparametric estimation and our goal is to find the proper core functions of SMSARCH models for forecasting the energy consumption in different sectors of the U.S. economy.

2.1 Estimation algorithm

The estimation processes introduced by Nademi and Farnoosh [19] is based on EM algorithm. For the estimation of the parameters, they followed the methodology of Franke et al. [10] in applying data log likelihood based on $(Z^{(N)}, \lambda^{(N)})$, rather than the usual log likelihood based on only $Z^{(N)}$. Accordingly, a complete log likelihood function was obtained as:

$$\begin{aligned} l_c(v, \Delta|Z^{(N)}, \lambda^{(N)}) &= \log \pi_{F_1} + \sum_{t=2}^{N} \log \Delta_{F_{t-1},F_t} \\ &+ \sum_{t=2}^{N}\sum_{k=1}^{M} \lambda_{tk} \log \frac{1}{\sigma(Z_{t-1}, Z_{t-2}; \omega_k, \alpha_k, \beta_k)} \varphi(\frac{Z_t - m(Z_{t-1}, Z_{t-2}; \theta_k, \rho_k)}{\sigma(Z_{t-1}, Z_{t-2}; \omega_k, \alpha_k, \beta_k)}), \end{aligned} \tag{3}$$

where $v = (\theta_1, \ldots, \theta_M, \rho_1, \ldots, \rho_M, \omega_1, \ldots, \omega_M, \alpha_1, \ldots, \alpha_M, \beta_1, \ldots, \beta_M)^T \in V$, $Z^{(N)} = (Z_1, \ldots, Z_N)$, $\lambda^{(N)} = (\lambda_1, \ldots, \lambda_N)$ and $\varphi(.)$ denote the normal density with mean $m(.)$ and standard deviation $\sigma(.)$. Therefore, we can maximize the relation (3) to estimate the parameters. However, maximizing (3) is not possible, because of the unobserving variables $\lambda^{(N)}$, without the use of the popular algorithm EM. The EM algorithm repeats between two expectation and maximization steps. In the expectation step (E-step), the approximation of the variables λ_{tk} are calculated by their conditional expectations given the observed data $Z^{(N)}$. In the maximization step (M-step), we maximize $l_c(v, \Delta|Z^{(N)}, \lambda^{(N)})$ to reach an estimate of Δ and v. These two steps are repeated until the satisfaction of some stopping criteria. The EM algorithm can be summarized in the following steps.

E-step: We suppose the parameter vector $\widehat{v}$ is given. Therefore, the conditional expectations of the variables λ_{tk} with given observations are estimated by

$$C_{tk} = \frac{\eta_k^t \tau_k^t}{\sum_{i=1}^M \eta_i^t \tau_i^t}, k = 1, \ldots, M, \ t = 1, \ldots, N,$$

where η_i^t and τ_i^t are calculated from

$$\eta_j^{t+1} = \varphi(Z_{t+1}; m(Z_t, Z_{t-1}; \theta_j, \rho_j), \sigma(Z_t, Z_{t-1}; \omega_j, \alpha_j, \beta_j)) \sum_{k=1}^M \Delta_{kj} \eta_k^t, \tag{4}$$

and

$$\tau_j^t = \sum_{k=1}^M \tau_k^{t+1} \varphi(Z_{t+1}; m(Z_t, Z_{t-1}; \theta_k, \rho_k), \sigma(Z_t, Z_{t-1}; \omega_k, \alpha_k, \beta_k)) \Delta_{jk}. \tag{5}$$

where $\varphi(Z_{t+1}; m(Z_t, Z_{t-1}; \theta_k, \rho_k), \sigma(Z_t, Z_{t-1}; \omega_k, \alpha_k, \beta_k))$ is the normal density with mean $m(.)$ and standard deviation $\sigma(.)$.

M-step: With estimation of random variables C_{tk}, the transition probabilities are estimated by

$$\widehat{\Delta}_{ij} = \frac{\sum_{t=1}^N \gamma_{ij}^{t,t+1}}{\sum_{t=1}^N C_{ti}},$$

where $\gamma_{ij}^{t,t+1}$ is calculated as

$$\gamma_{ij}^{t,t+1} = \frac{\tau_j^{t+1} \varphi(Z_{t+1}; m(Z_t, Z_{t-1}; \theta_j, \rho_j), \sigma(Z_t, Z_{t-1}; \omega_j, \alpha_j, \beta_j)) \Delta_{ij} \eta_i^t}{\sum_{k=1}^M \eta_k^t \tau_k^t}. \tag{6}$$

The probabilities $\pi_1, \ldots, \pi_M$ are estimated by

$$\widehat{\pi}_k = \frac{1}{N} \sum_{t=1}^N C_{tk}, \ k = 1, \ldots, M.$$

The M autoregression functions $\widehat{m}(x, z; \widehat{\theta}_k, \widehat{\rho}_k)$ are estimated by

$$\widehat{m}(x, z; \widehat{\theta}_k, \widehat{\rho}_k) = g(x, \widehat{\theta}_k)\widehat{\xi}_k(x) + \widehat{\rho}_k(x - g(z, \widehat{\theta}_k)\widehat{\xi}_k(z)), \tag{7}$$

such that $(\widehat{\rho}_k, \widehat{\theta}_k)$ get from $(\widehat{\rho}_k, \widehat{\theta}_k) = arg\min Q_n(\theta_k, \rho_k), \theta_k, \rho_k \in V, |\rho_k| < 1$ for $k = 1, \ldots, M$ where $Q_n(\theta_k, \rho_k)$ is

$$Q_n(\theta_k, \rho_k) = \sum_{t=2}^{N} C_{tk}(Z_t - g(Z_{t-1}, \theta_k) - \rho_k(Z_{t-1} - g(Z_{t-2}, \theta_k)))^2,$$

and

$$\widehat{\xi}_k(x) = \frac{\sum_{t=2}^{N} C_{tk}[k(\frac{Z_{t-1}-x}{h_k})g(Z_{t-1}, \widehat{\theta}_k)Z_t + k(\frac{Z_{t-2}-x}{h_k})g(Z_{t-2}, \widehat{\theta}_k)Z_{t-1}]}{\sum_{t=2}^{N} C_{tk}[k(\frac{Z_{t-1}-x}{h_k})g^2(Z_{t-1}, \widehat{\theta}_k) + k(\frac{Z_{t-2}-x}{h_k})g^2(Z_{t-2}, \widehat{\theta}_k)]},$$

where $k(.)$ is a kernel function and $(\widehat{\omega}_k, \widehat{\alpha}_k, \widehat{\beta}_k)$ are estimated by

$$(\widehat{\omega}_k, \widehat{\alpha}_k, \widehat{\beta}_k) = \arg\max \sum_{t=2}^{N} C_{tk} \log \frac{1}{\sigma(Z_{t-1}, Z_{t-2}; \omega_k, \alpha_k, \beta_k)} \varphi(\frac{Z_t - \widehat{m}(Z_{t-1}, Z_{t-2}; \widehat{\theta}_k, \widehat{\rho}_k)}{\sigma(Z_{t-1}, Z_{t-2}; \omega_k, \alpha_k, \beta_k)}),$$

for $k = 1, \ldots, M$. The optimal selection of the bandwidths h_k are calculated by

$$\widehat{h}_k = arg\min_{h_k} \sum_{t=2}^{N} C_{tk}[Z_t - g(Z_{t-1}, \widehat{\theta}_k)\widehat{\xi}_k(Z_{t-1}) - \widehat{\rho}_k(Z_{t-1} - g(Z_{t-2}, \widehat{\theta}_k)\widehat{\xi}_k(Z_{t-2}))]^2.$$

Finally, the estimation of the parameters is obtained by iterating these two steps until convergence (see [19,20]).

2.2 Competitive models

To evaluate the validity of the SMSARCH models, we compare the forecasting ability of this model containing different core functions with different models including ARIMA [5,11], GARCH [4,7,8], EGARCH [21], parametric Markov switching in mean (PMSM)[13,14,18] and parametric Markov switching GARCH (PMSGARCH) models [1]. Table 1 introduces these competitive models.

Table 1: The Competitive models for considering accuracy forecasting of the SMSARCH models

Model	Mean Equation	Variance Equation	Reference of the Models
ARIMA Models	AR(1), MA(1) , ARMA(1,1)	There is no variance equation	Box et. al (2015) (2015)
GARCH Models with Normal Distribution	Simple Mean, AR(1), MA(1), ARMA(1,1)	GARCH(1,1)	Bollerslev (1987), Diebold and Lopez (1995), Engle (2001)
EGARCH Models with Normal Distribution	Simple Mean, AR(1), MA(1), ARMA(1,1)	EGARCH(1,1)	Nelson (1991), Francq and Zakoian (2011)
Parametric Markov Switching Models with Normal Distribution (with switching Parameters between Two Regimes in Mean Equation)	Simple Mean, AR(1), MA(1), ARMA(1,1)	There is no variance equation	Hamilton(1989, 2010)
Parametric Markov Switching GARCH Models with Normal Distribution (with switching Parameters between Two Regimes in both of Mean and Variance Equations)	Simple Mean, AR(1), MA(1), ARMA(1,1)	GARCH(1,1)	Marcucci (2005), Abounoori, Elmi and Nademi (2016)

2.3 Data

We investigated the total U.S. energy consumption in five residential, commercial, industrial, transportation and electric power sectors. The data was taken from the U.S. energy information administration [1] from January 2000 to December 2016. Table 2 reports the descriptive statistics for 204 observations in each studied sector. We chose the time periods from January 2000 to December 2015 and from January 2016 to December 2016 as in-sample and out-of-sample data, respectively. Also, four different core functions were selected as:

$$g_1(x,\theta) = \exp(\theta x), g_2(x,\theta) = \exp(\theta\sqrt{x}), g_3(x,\theta) = \frac{\exp(\theta x)}{1+\exp(\theta x)}, g_4(x,\theta) = \theta\sin(x).$$

At first, we tested the unit root in five sectors by the Augmented Dickey-Fuller (ADF) test. Table 3 lists the results of ADF test based on in-sample data. Accordingly, the existence of unit root in

[1] http:// www.eia.gov

all five sectors is confirmed, and therefore, the total energy consumption is nonstationary.
Then, the first difference of the five series was tested by the ADF test. The results of the ADF test, reported in Table 4, indicate a stationary process in the first difference. Therefore, we should apply the first difference of total energy consumption for modeling and forecasting.

Table 2: Descriptive statistics for Total Energy Consumption in five sectors (Trillion Btu) (Source: Own Calculation)

Sectors	N	Mean	Standard deviation	Minimum
Residential	204	1747.39	366.42	1282.63
Commercial	204	1484.23	120.92	1281.20
Industrial	204	2649.38	136.35	2247.87
Transportation	204	2272.01	112.08	1951.70
Electric Power	204	3221.26	323.40	2682.49
Sectors	Maximum	Skewness	Kurtosis	
Residential	2774.22	0.84	-0.28	
Commercial	1865.48	0.76	0.06	
Industrial	3037.25	-0.26	0.33	
Transportation	2541.26	-0.34	0.28	
Electric Power	4084.46	0.67	-0.37	

Table 3: ADF test for Total Energy Consumption by different Sectors baesd on in-sample data (Source: Own Calculation)

	ADF-Statistic	P-value	Result
Residential Sector	-1.34	0.17	Non-Stationary
Commercial Sector	-0.59	0.43	Non-Stationary
Industrial Sector	-0.50	0.47	Non-Stationary
Transportation Sector	-0.23	0.57	Non-Stationary
Electric Power Sector	-0.70	0.39	Non-Stationary

Table 4: ADF test for first difference of Total Energy Consumption baesd on in-sample data by different Sectors (Source: Own Calculation)

	ADF-Statistic	P-value	Result
Residential Sector	-9.13	0.001	Stationary
Commercial Sector	-12.67	0.001	Stationary
Industrial Sector	-28.24	0.001	Stationary
Transportation Sector	-24.86	0.001	Stationary
Electric Power Sector	-11.22	0.001	Stationary

One of the important problems in switching time series is to recognize the number of regimes on the behavior of time series. This problem refers to the selection of the regimes based on the numerical methods. However, choosing the proper primary point can help getting results in a faster way. In order to simplify showing the regimes in graphs, we applied the step functions to show the regimes based on some cut points and behavior functions for the samples. According to Figs. 1, 2 and 3 (step functions), we can suppose that the samples have two regimes by high and low consumption functions if we consider the points 2495, 1700 and 3776 as cut points for the residential, commercial and electric power sectors, respectively. We can consider two regimes based on the increasing and decreasing functions for the industrial sector. Furthermore, for the transportation sector, three regimes based on homogenous, increasing and decreasing functions can be considered (see Figs. 4 and 5). Therefore, we examine four SMSARCH models (called MS-SEMI-$G(i)$) based on g_1, g_2, g_3 and g_4, and evaluate the efficiency of these four models by the root-mean-square error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(Z_t - \widehat{Z}_t)^2}.$$

To compare the ability of the models in classifying the observations, we applied the index $max(C_{tk}),\ k = 1, \dots, M$. Such that Z_t belongs to the regime k if and only if $C_{tk} = \max(C_{ts}), s = 1, \dots, M$.

3 Empirical Results

Tables 5-9 report the estimated parameters of total energy consumption for all the studied sectors using the semi-parametric Markov switching models with different specifications. The RMSE criterion indicates that the MS-SEMI-G(1) model shows a more accurate fitting than other semi-parametric Markov switching models for the residential, commercial and transportation sectors. However, for the industrial and electric power sectors the MS-SEMI-G(4) model shows more efficiency than other switching models. The proper core function for the residential, commercial and transportation sectors is $g_1(x, \theta) = \exp(\theta x)$ and for the industrial and electric power sectors is $g_4(x, \theta) = \theta \sin(x)$. Figures 6-10 show the sample path for the selected models based on the best core function. Figures 11-15 show $max(C_{tk}),\ k = 1, \dots, M$ based on M for the four studied models. All the M values are greater than 0.8 for the MS-SEMI-G(1) and MS-SEMI-G(4) models showing that these two models are more powerful for the classification of the observations in comparison to the other models. On the other hand, the minimum RMSE index of in-sample data can be seen for the SMSARCH models (see table 10).

Table 5: The estimated parameters for Total Energy Consumption by the Residential Sector based on in-sample data (Source: Own Calculation)

	$g_1(x,\theta)=\exp(\theta x)$	$g_2(x,\theta)=\exp(\theta\sqrt{x})$	$g_3(x,\theta)=\frac{\exp(\theta x)}{1+\exp(\theta x)}$	$g_4(x,\theta)=\theta\sin(x)$
$\widehat{\Delta}_{12},\widehat{\Delta}_{21}$	0.2463,0.6523	0.2951,0.6610	0.3102,0.5834	0.2357,0.6348
	(0.104,0.214)	(0.116,0.205)	(0.103,0.354)	(0.184,0.199)
$\widehat{\omega}_1,\widehat{\omega}_2$	0.0114,0.0143	0.0101,0.0421	0.0214,0.0357	0.0214,0.0468
	(0.114,0.121)	(0.348,0.679)	(0.205,0.615)	(0.316,0.721)
$\widehat{\alpha}_1,\widehat{\beta}_1$	0.0015,0.0021	0.0611,0.0549	0.0313,0.0824	0.0164,0.0354
	(0.645,0.112)	(0.349,0.018)	(0.659,0.378)	(0.842,0.638)
$\widehat{\alpha}_2,\widehat{\beta}_2$	0.0012,0.0061	0.0049,0.0076	0.0046,0.0010	0.0084,0.0021
	(0.102,0.237)	(0.345,0.391)	(0.618,0.708)	(0.130,0.121)
$\widehat{\theta}_1,\widehat{\theta}_2$	0.1605,0.3112	0.2041,0.4870	0.3490,0.4672	0.1106,0.4872
	(0.126,0.103)	(0.264,0.518)	(0.326,0.124)	(0.457,0.885)
$\widehat{\rho}_1,\widehat{\rho}_2$	0.6150,0.5284	0.8472,0.6259	0.5104,0.7134	0.6015,0.4327
	(0.204,0.349)	(0.047,0.489)	(0.346,0.709)	(0.257,0.239)
$\widehat{\pi}_1,\widehat{\pi}_2$	0.7259,0.2741	0.6914,0.3086	0.6529,0.3471	0.7292,0.2708
	(0.015,0.038)	(0.164,0.205)	(0.109,0.037)	(0.323,0.107)
$\widehat{h}_1,\widehat{h}_2$	0.0421,0.0441	0.0205,0.0319	0.0249,0.0450	0.0410,0.0397
	(0.184,0.216)	(0.319,0.481)	(0.164,0.519)	(0.546,0.227)
$RMSE$	10.3128	12.1648	15.2214	11.1294

Table 6: The estimated parameters for Total Energy Consumption by the Commercial Sector based on in-sample data (Source: Own Calculation)

	$g_1(x,\theta)=\exp(\theta x)$	$g_2(x,\theta)=\exp(\theta\sqrt{x})$	$g_3(x,\theta)=\frac{\exp(\theta x)}{1+\exp(\theta x)}$	$g_4(x,\theta)=\theta\sin(x)$
$\widehat{\Delta}_{12},\widehat{\Delta}_{21}$	0.6145,0.4318	0.5986,0.3487	0.7842,0.4956	0.6480,0.3491
	(0.118,0.425)	(0.648,0.492)	(0.145,0.294)	(0.759,0.984)
$\widehat{\omega}_1,\widehat{\omega}_2$	0.0001,0.0040	0.0011,0.0002	0.0027,0.0011	0.0064,0.0019
	(0.362,0.184)	(0.110,0.203)	(0.418,0.648)	(0.267,0.164)
$\widehat{\alpha}_1,\widehat{\beta}_1$	0.0127,0.0026	0.0151,0.0094	0.0048,0.0375	0.0018,0.0101
	(0.208,0.846)	(0.785,0.629)	(0.598,0.213)	(0.754,0.252)
$\widehat{\alpha}_2,\widehat{\beta}_2$	0.0021,0.0038	0.0112,0.0209	0.0033,0.0128	0.0061,0.0650
	(0.326,0.489)	(0.128,0.162)	(0.325,0.951)	(0.618,0.493)
$\widehat{\theta}_1,\widehat{\theta}_2$	0.3610,0.4871	0.9521,0.4672	1.5268,1.2136	0.5942,0.1681
	(0.021,0.064)	(0.208,0.355)	(0.647,0.591)	(0.137,0.381)
$\widehat{\rho}_1,\widehat{\rho}_2$	0.6628,0.5138	0.5610,0.4346	0.4941,0.6892	0.7298,0.5032
	(0.284,0.952)	(0.464,0.108)	(0.627,0.493)	(0.424,0.648)
$\widehat{\pi}_1,\widehat{\pi}_2$	0.4127,0.5873	0.3681,0.6319	0.3872,0.6128	0.3501,0.6499
	(0.021,0.032)	(0.018,0.034)	(0.015,0.304)	(0.029,0.015)
$\widehat{h}_1,\widehat{h}_2$	0.0042,0.0032	0.0156,0.0208	0.0360,0.0034	0.0108,0.0054
	(0.327,0.643)	(0.938,0.821)	(0.032,0.284)	(0.349,0.184)
$RMSE$	13.3145	15.1846	18.2840	14.1246

Table 7: The estimated parameters for Total Energy Consumption by the Industrial Sector based on in-sample data (Source: Own Calculation)

	$g_1(x,\theta)=\exp(\theta x)$	$g_2(x,\theta)=\exp(\theta\sqrt{x})$	$g_3(x,\theta)=\frac{\exp(\theta x)}{1+\exp(\theta x)}$	$g_4(x,\theta)=\theta\sin(x)$
$\widehat{\Delta}_{12},\widehat{\Delta}_{21}$	0.4578,0.5162	0.6465,0.3258	0.5941,0.2654	0.6549,0.4512
	(0.152,0.952)	(0.349,0.645)	(0.651,0.321)	(0.120,0.215)
$\widehat{\omega}_1,\widehat{\omega}_2$	0.0002,0.0011	0.0021,0.0016	0.0031,0.0085	0.0010,0.0004
	(0.025,0.059)	(0.132,0.902)	(0.112,0.348)	(0.258,0.346)
$\widehat{\alpha}_1,\widehat{\beta}_1$	0.0127,0.0325	0.0101,0.0032	0.0150,0.0658	0.0412,0.0016
	(0.264,0.679)	(0.845,0.628)	(0.201,0.349)	(0.802,0.329)
$\widehat{\alpha}_2,\widehat{\beta}_2$	0.0048,0.0056	0.0012,0.0031	0.0023,0.0315	0.0058,0.0032
	(0.205,0.349)	(0.508,0.267)	(0.401,0.308)	(0.294,0.349)
$\widehat{\theta}_1,\widehat{\theta}_2$	0.6580,0.3218	0.3190,0.2648	1.2648,2.001	0.2655,0.9842
	(0.330,0.842)	(0.895,0.649)	(0.799,0.235)	(0.294,0.264)
$\widehat{\rho}_1,\widehat{\rho}_2$	0.5628,0.6218	0.4697,0.5102	0.6987,0.5684	0.5512,0.5694
	(0.528,0.167)	(0.305,0.194)	(0.302,0.841)	(0.637,0.264)
$\widehat{\pi}_1,\widehat{\pi}_2$	0.5300,0.4700	0.3351,0.6649	0.3088,0.6912	0.4079,0.5921
	(0.021,0.013)	(0.024,0.010)	(0.011,0.023)	(0.014,0.031)
$\widehat{h}_1,\widehat{h}_2$	0.0215,0.0152	0.0024,0.0034	0.0012,0.0026	0.0320,0.0057
	(0.209,0.306)	(0.419,0.578)	(0.608,0.112)	(0.246,0.637)
$RMSE$	11.6478	14.2648	14.1644	10.1015

Table 8: The estimated parameters for Total Energy Consumption by the Transportation Sector based on in-sample data (Source: Own Calculation)

	$g_1(x,\theta) = \exp(\theta x)$	$g_2(x,\theta) = \exp(\theta\sqrt{x})$	$g_3(x,\theta) = \frac{\exp(\theta x)}{1+\exp(\theta x)}$	$g_4(x,\theta) = \theta\sin(x)$
$\widehat{\Delta}_{12}, \widehat{\Delta}_{13}$	0.2584,0.2245	0.2103,0.2115	0.2391,0.2433	0.2689,0.2165
	(0.466,0.645)	(0.310,0.562)	(0.652,0.054)	(0.048,0.108)
$\widehat{\Delta}_{21}, \widehat{\Delta}_{23}$	0.3127,0.1014	0.3201,0.1403	0.3651,0.1659	0.3368,0.1103
	(0.516,0.348)	(0.125,0.013)	(0.483,0.761)	(0.159,0.617)
$\widehat{\Delta}_{31}, \widehat{\Delta}_{32}$	0.4194,0.4029	0.4523,0.4410	0.4025,0.4002	0.4381,0.4219
	(0.358,0.428)	(0.139,0.750)	(0.763,0.813)	(0.617,0.048)
$\widehat{\omega}_1, \widehat{\omega}_2, \widehat{\omega}_3$	0.0001,0.0003,0.0035	0.0018,0.0021,0.0010	0.0007,0.0024,0.0061	0.0032,0.0046,0.0021
	(0.056,0.042,0.079)	(0.023,0.150,0.047)	(0.349,0.491,0.617)	(0.110,0.320,0.594)
$\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{\alpha}_3$	0.0013,0.0058,0.0010	0.0064,0.0076,0.0001	0.0017,0.0026,0.0005	0.0059,0.0095,0.0011
	(0.016,0.301,0.112)	(0.455,0.341,0.210)	(0.306,0.140,0.023)	(0.346,0.250,0.109)
$\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3$	0.0002,0.0040,0.0015	0.0002,0.0013,0.0024	0.0004,0.0017,0.0002	0.0021,0.0011,0.0005
	(0.560,0.543,0.548)	(0.500,0.208,0.349)	(0.113,0.389,0.011)	(0.187,0.162,0.219)
$\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3$	0.5984,0.3348,0.0251	0.2594,0.3591,0.1020	0.3541,1.659,1.5208	0.4670,0.8452,0.2691
	(0.064,0.091,0.108)	(0.023,0.206,0.130)	(0.016,0.049,0.209)	(0.112,0.041,0.052)
$\widehat{\rho}_1, \widehat{\rho}_2, , \widehat{\rho}_3$	0.4419,0.5234,0.3261	0.2490,0.6617,0.1152	0.3641,0.6260,0.3461	0.5921,0.6456,0.2240
	(0.021,0.048,0.441)	(0.114,0.109,0.629)	(0.058,0.067,0.437)	(0.116,0.227,0.211)
$\widehat{\pi}_1, \widehat{\pi}_2, \widehat{\pi}_3$	0.4152,0.4197,0.1651	0.4613,0.3712,0.1675	0.4399,0.3537,0.2064	0.4294,0.4099,0.1607
	(0.003,0.008,0.042)	(0.004,0.021,0.042)	(0.009,0.054,0.042)	(0.006,0.004,0.042)
$\widehat{h}_1, \widehat{h}_2, \widehat{h}_3$	0.0025,0.0051,0.0021	0.0079,0.0023,0.0040	0.0203,0.0041,0.0031	0.0264,0.0015,0.0019
	(0.084,0.031,0.518)	(0.403,0.059,0.012)	(0.049,0.019,0.110)	(0.107,0.039,0.126)
$RMSE$	12.0548	16.2401	15.1131	12.3451

Table 9: The estimated parameters for Total Energy Consumption by the Electric Power Sector based on in-sample data (Source: Own Calculation)

	$g_1(x,\theta)=\exp(\theta x)$	$g_2(x,\theta)=\exp(\theta\sqrt{x})$	$g_3(x,\theta)=\frac{\exp(\theta x)}{1+\exp(\theta x)}$	$g_4(x,\theta)=\theta\sin(x)$
$\widehat{\Delta}_{12},\widehat{\Delta}_{21}$	0.6659,0.4126	0.5426,0.5528	0.5984,0.5023	0.4620,0.5438
	(0.361,0.321)	(0.980,0.320)	(0.120,0.132)	(0.118,0.103)
$\widehat{\omega}_1,\widehat{\omega}_2$	0.0001,0.0001	0.0002,0.0004	0.0001,0.0002	0.0001,0.0013
	(0.156,0.116)	(0.204,0.296)	(0.346,0.551)	(0.325,0.842)
$\widehat{\alpha}_1,\widehat{\beta}_1$	0.0002,0.0005	0.0003,0.0010	0.0004,0.0001	0.0003,0.0002
	(0.359,0.119)	(0.428,0.205)	(0.313,0.056)	(0.084,0.037)
$\widehat{\alpha}_2,\widehat{\beta}_2$	0.0010,0.0015	0.0042,0.0036	0.0042,0.0072	0.0003,0.0005
	(0.059,0.048)	(0.019,0.134)	(0.492,0.783)	(0.264,0.084)
$\widehat{\theta}_1,\widehat{\theta}_2$	0.6658,0.1572	0.3326,0.6294	1.5594,2.3648	0.3249,0.5977
	(0.204,0.349)	(0.059,0.037)	(0.101,0.237)	(0.309,0.024)
$\widehat{\rho}_1,\widehat{\rho}_2$	0.6658,0.3958	0.4678,0.5625	0.4892,0.6058	0.5528,0.5394
	(0.008,0.027)	(0.023,0.160)	(0.033,0.049)	(0.215,0.119)
$\widehat{\pi}_1,\widehat{\pi}_2$	0.3826,0.6174	0.5047,0.4953	0.4563,0.5437	0.5407,0.4593
	(0.006,0.013)	(0.014,0.003)	(0.002,0.003)	(0.005,0.011)
$\widehat{h}_1,\widehat{h}_2$	0.0251,0.0153	0.0120,0.0452	0.0351,0.0356	0.0038,0.0026
	(0.012,0.076)	(0.108,0.346)	(0.149,0.307)	(0.451,0.007)
$RMSE$	14.1081	17.6627	14.9845	13.0051

Table 10: The $RMSE$ Criteria for In-sample Estimation for different models in the U.S energy sectors

	MS-SEMI-G(1)	MS-SEMI-G(2)	MS-SEMI-G(3)
Residential Sector	10.31	12.17	15.22
Commercial Sector	13.32	15.19	18.28
Industrial Sector	11.65	14.27	14.16
Transportation Sector	12.06	16.24	15.11
Electric Power Sector	14.11	17.66	14.99
	MS-SEMI-G(4)	AR(1)	MA(1)
Residential Sector	11.13	390.35	558.61
Commercial Sector	14.13	131.36	162.93
Industrial Sector	10.10	120.84	151.07
Transportation Sector	12.35	159.68	171.53
Electric Power Sector	13.01	629.36	354.54
	ARMA(1,1)	GARCH (1,1)	AR(1)-GARCH(1,1)
Residential Sector	573.30	1014.73	434.67
Commercial Sector	159.55	2031.96	124.43
Industrial Sector	122.22	611.92	125.19
Transportation Sector	178.88	1510.13	154.46
Electric Power Sector	333.43	1146.72	411.79
	MA(1)-GARCH(1,1)	ARMA (1,1)-GARCH(1,1)	EGARCH (1,1)
Residential Sector	542.46	1705.22	10221.36
Commercial Sector	163.86	185.35	154.68
Industrial Sector	152.59	126.44	1751.03
Transportation Sector	151.25	176.11	3980.8
Electric Power Sector	345.19	688.00	1829.68

Table 10: The $RMSE$ Criteria for In-sample Estimation for different models in the U.S energy sectors (continue)

	AR(1)-EGARCH(1,1)	MA(1)-)EGARCH(1,1)	ARMA (1,1)-EGARCH(1,1)
Residential Sector	1731.40	638.69	11064.49
Commercial Sector	458.86	604.43	343.88
Industrial Sector	443.08	127.76	129.35
Transportation Sector	1954.29	645.86	587.94
Electric Power Sector	1303.13	991.52	1148.25
	Parametric MS with Simple Mean	Parametric MS with AR(1) process in Mean	Parametric MS with ARMA (1,1) process in Mean
Residential Sector	433.53	442.73	1062.60
Commercial Sector	182.47	122.17	235.51
Industrial Sector	185.36	191.18	237.34
Transportation Sector	319.77	203.17	488.77
Electric Power Sector	310.01	911.29	772.61
	Parametric MS GARCH with Simple Mean	Parametric MS GARCH with AR(1) process in Mean	Parametric MS GARCH with ARMA (1,1) process in Mean
Residential Sector	754.68	900.83	1030.39
Commercial Sector	188.37	121.44	306.53
Industrial Sector	266.29	205.14	340.44
Transportation Sector	518.50	448.77	503.11
Electric Power Sector	361.18	1656.70	1401.10

We also compared the forecasting ability of out-of-sample data for all the models during the time period from January 2016 to December 2016. Accordingly, the RMSE values for five studied sectors are reported in Table 11. The results indicate that the SMSARCH models show a more accurate fitting than other competing parametric models. For the residential, commercial and transportation sectors, the MS-SEMI-G(1) model shows a more accurate forecasting ability than other semiparametric Markov switching models, while for the industrial and electric power sectors, the MS-SEMI-G(4) model forecasts better. The difference between the RMSE values of parametric and semiparametric models is very high. Also, the Diebold and Mariano test in Table 12 confirms that the difference between the RMSE values is significant, and therefore, the forecasting ability between parametric and semiparametric models are different.

Table 11: The $RMSE$ Criteria for out of sample (2016m01-2016m12) for different models in the U.S energy sectors

	MS-SEMI-G(1)	MS-SEMI-G(2)	MS-SEMI-G(3)
Residential Sector	12.26	15.30	19.12
Commercial Sector	14.16	16.54	21.85
Industrial Sector	12.10	16.41	16.08
Transportation Sector	12.48	19.62	18.53
Electric Power Sector	15.22	17.19	16.73
	MS-SEMI-G(4)	AR(1)	MA(1)
Residential Sector	13.94	560.52	487.16
Commercial Sector	16.01	135.49	135.79
Industrial Sector	10.97	78.83	73.76
Transportation Sector	13.01	107.81	96.86
Electric Power Sector	13.24	391.95	391.16
	ARMA(1,1)	GARCH (1,1)	AR(1)-GARCH(1,1)
Residential Sector	344.23	366.68	404.56
Commercial Sector	125.76	206.71	134.23
Industrial Sector	76.98	91.83	79.75
Transportation Sector	98.86	158.31	94.20
Electric Power Sector	375.82	440.84	403.07
	MA(1)-GARCH(1,1)	ARMA (1,1)-GARCH(1,1)	EGARCH (1,1)
Residential Sector	398.86	545.25	612.02
Commercial Sector	132.91	133.01	131.13
Industrial Sector	78.84	76.59	131.13
Transportation Sector	94.30	92.44	308.9
Electric Power Sector	408.53	373.97	467.98

Table 11: The *RMSE* Criteria for out of sample (2016m01-2016m12) for different models in the U.S energy sectors(continue)

	AR(1)-EGARCH(1,1)	MA(1)-) EGARCH(1,1)	ARMA (1,1)-EGARCH(1,1)
Residential Sector	484.59	485.63	1096.91
Commercial Sector	147.63	148.19	140.15
Industrial Sector	77.58	74.41	79.87
Transportation Sector	113.58	86.65	86.35
Electric Power Sector	419.9	421.06	417.11
	Parametric MS with Simple Mean	Parametric MS with AR(1) process in Mean	Parametric MS with ARMA (1,1) process in Mean
Residential Sector	107.64	1050.91	871.61
Commercial Sector	164.30	161.02	268.78
Industrial Sector	263.57	454.93	292.91
Transportation Sector	257.92	294.52	407.06
Electric Power Sector	187.70	225.72	729.25
	Parametric MS GARCH with Simple Mean	Parametric MS GARCH with AR(1) process in Mean	Parametric MS GARCH with ARMA (1,1) process in Mean
Residential Sector	457.58	7717.67	779.21
Commercial Sector	166.03	294.17	197.21
Industrial Sector	563.6	745.26	322.97
Transportation Sector	4778.64	581.17	366.09
Electric Power Sector	211.54	1842.26	2588.81

Table 12: The Diebold-Mariano Test (MS-SEMI-G(1)) based on RMSE Statistic (P-Value)

	MS-SEMI-G(2)	MS-SEMI-G(3)	MS-SEMI-G(4)
Residential Sector	0.00	0.00	0.02
Commercial Sector	0.00	0.00	0.00
Industrial Sector	0.00	0.00	0.00
Transportation Sector	0.00	0.00	0.01
Electric Power Sector	0.00	0.01	0.00
	AR(1)	MA(1)	ARMA(1,1)
Residential Sector	0.00	0.00	0.00
Commercial Sector	0.00	0.00	0.00
Industrial Sector	0.00	0.00	0.00
Transportation Sector	0.00	0.00	0.00
Electric Power Sector	0.00	0.00	0.00
	GARCH (1,1)	AR(1)-GARCH(1,1)	MA(1)-GARCH(1,1)
Residential Sector	0.00	0.00	0.00
Commercial Sector	0.00	0.00	0.00
Industrial Sector	0.00	0.00	0.00
Transportation Sector	0.00	0.00	0.00
Electric Power Sector	0.00	0.00	0.00
	ARMA (1,1)-GARCH(1,1)	EGARCH (1,1)	
Residential Sector	0.00	0.00	
Commercial Sector	0.00	0.00	
Industrial Sector	0.00	0.00	
Transportation Sector	0.00	0.00	
Electric Power Sector	0.00	0.00	

Table 12: The Diebold-Mariano Test (MS-SEMI-G(1)) based on RMSE Statistic (P-Value)(continue)

	AR(1)-EGARCH(1,1)	MA(1)-)EGARCH(1,1)	ARMA (1,1)-EGARCH(1,1)
Residential Sector	0.00	0.00	0.00
Commercial Sector	0.00	0.00	0.00
Industrial Sector	0.00	0.00	0.00
Transportation Sector	0.00	0.00	0.00
Electric Power Sector	0.00	0.00	0.00
	Parametric MS with Simple Mean	Parametric MS with AR(1) process in Mean	Parametric MS with ARMA (1,1) process in Mean
Residential Sector	0.00	0.00	0.00
Commercial Sector	0.00	0.00	0.00
Industrial Sector	0.00	0.00	0.00
Transportation Sector	0.00	0.00	0.00
Electric Power Sector	0.00	0.00	0.00
	Parametric MS GARCH with Simple Mean	Parametric MS GARCH with AR(1) process in Mean	Parametric MS GARCH with ARMA (1,1) process in Mean
Residential Sector	0.00	0.00	0.00
Commercial Sector	0.00	0.00	0.00
Industrial Sector	0.00	0.00	0.00
Transportation Sector	0.00	0.00	0.00
Electric Power Sector	0.00	0.00	0.00

4 Conclusion

We forecast the total U.S. energy consumption in five residential, commercial, industrial, transportation and electric power sectors. We applied the SMSARCH model to forecast the total energy consumption in the United States during the time period from January 2000 to December 2015. Because of nonstationary of energy consumption in level, at first, we used the first difference of energy consumption for modeling and forecasting the energy consumption. Then, we estimated the SMSARCH models with different core functions. Also, we used different parametric models including ARIMA, GARCH, EGARCH, PMSM and PMSGARCH models. The RMSE criterion in both in-sample and out-of-sample forecasting indicates that the SMSARCH models are more accurate than other parametric models for all the studied sectors. Furthermore, the empirical results indicate that the MS-SEMI-G(1) and MS-SEMI-G(4) models are more powerful in the classification of the observations than the other models.

References

[1] Abounoori, E., Elmi, Z. M., & Nademi, Y. (2016). Forecasting Tehran stock exchange volatility; Markov switching GARCH approach. *Physica A: Statistical Mechanics and its Applications*,

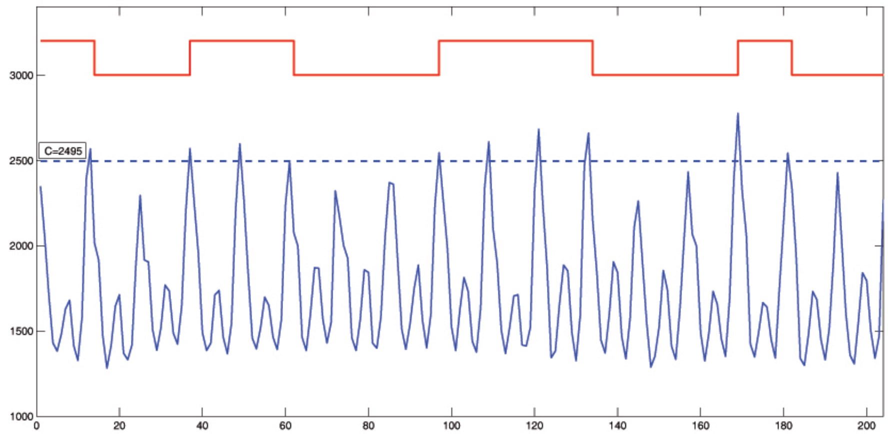

Figure 1: The step plot of Total Energy Consumed by the Residential Sector.

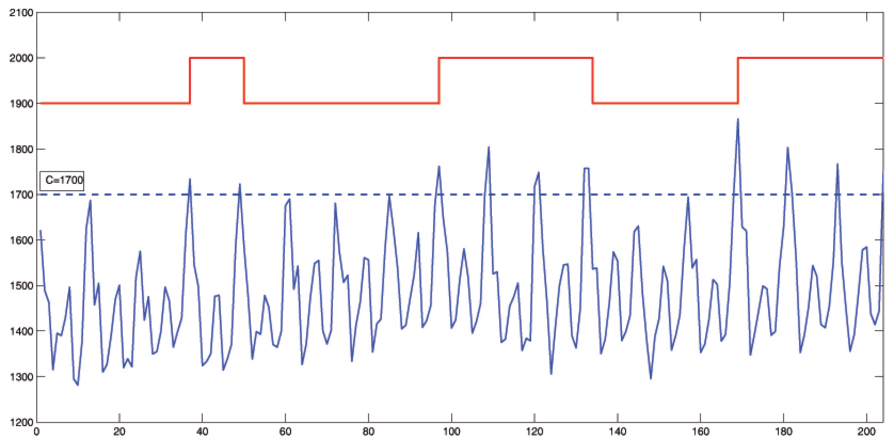

Figure 2: The step plot of Total Energy Consumed by the Commercial Sector.

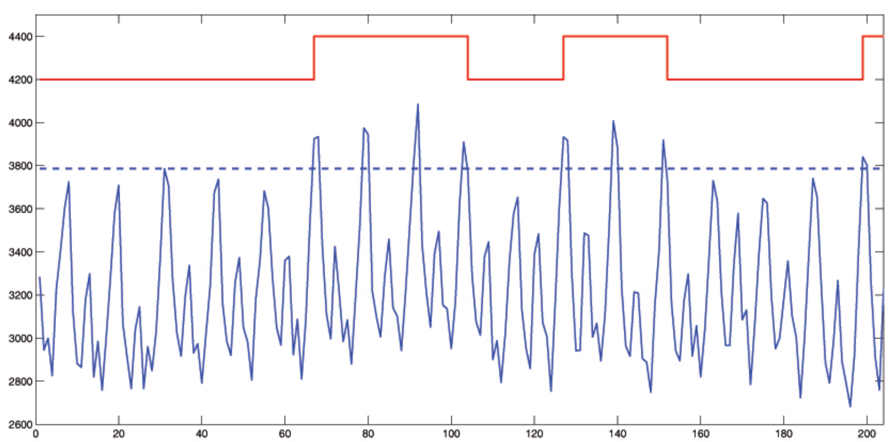

Figure 3: The step plot of Total Energy Consumed by the Electric Power Sector.

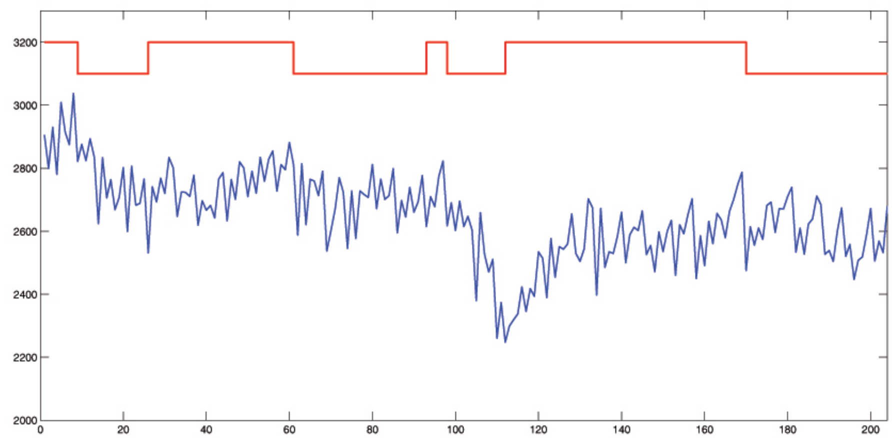

Figure 4: The step plot of Total Energy Consumed by the Industrial Sector.

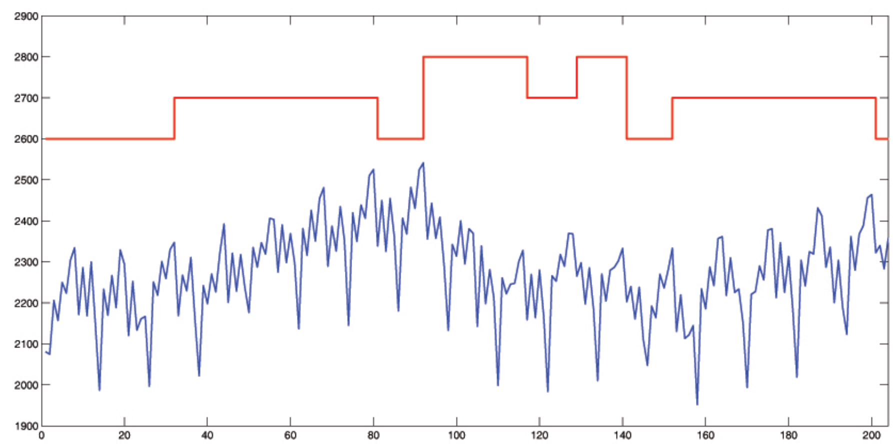

Figure 5: The step plot of Total Energy Consumed by the Transportation Sector.

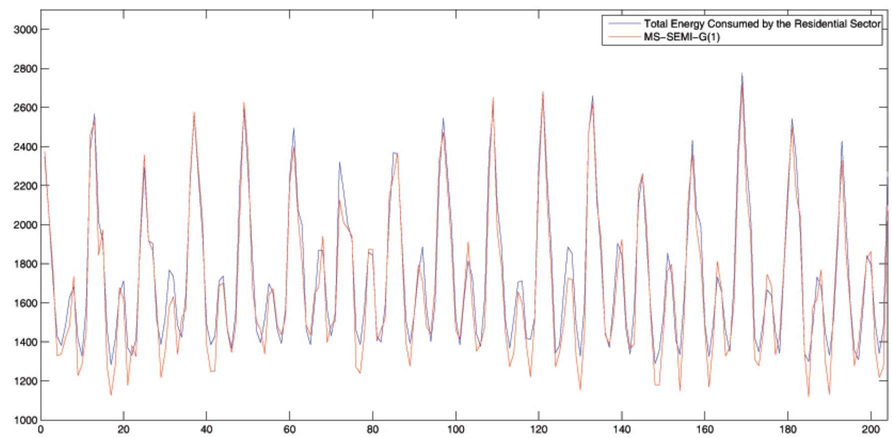

Figure 6: The sample path and the approximated model for Total Energy Consumed by the Residential Sector.

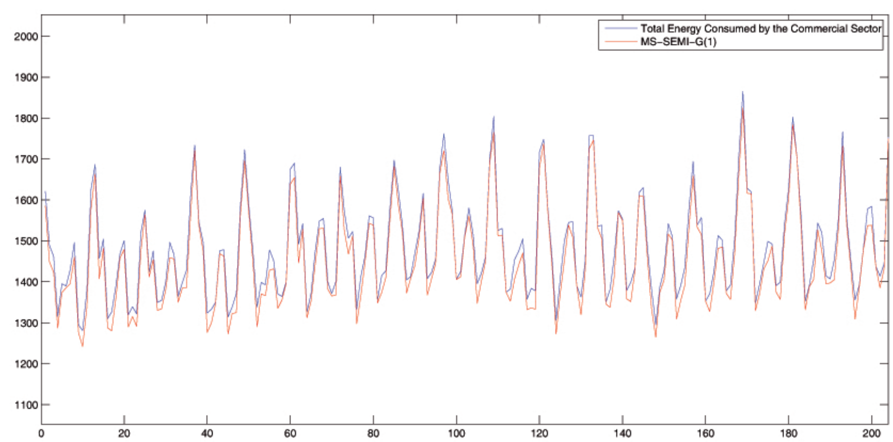

Figure 7: The sample path and the approximated model for Total Energy Consumed by the Commercial Sector.

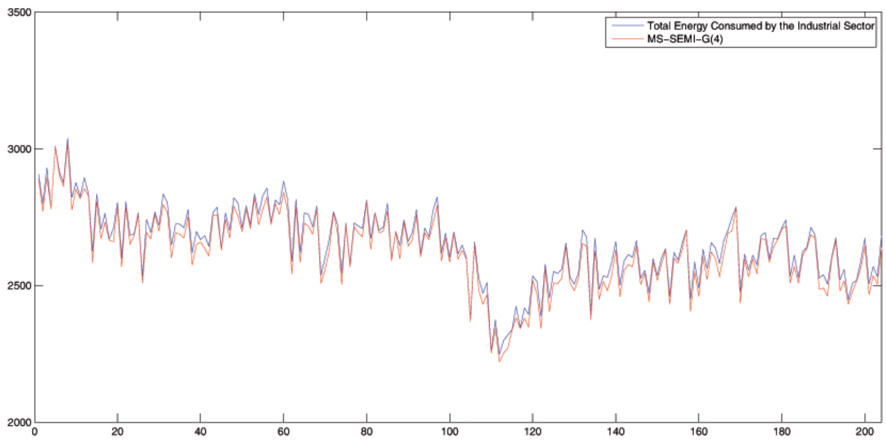

Figure 8: The sample path and the approximated model for Total Energy Consumed by the Industrial Sector.

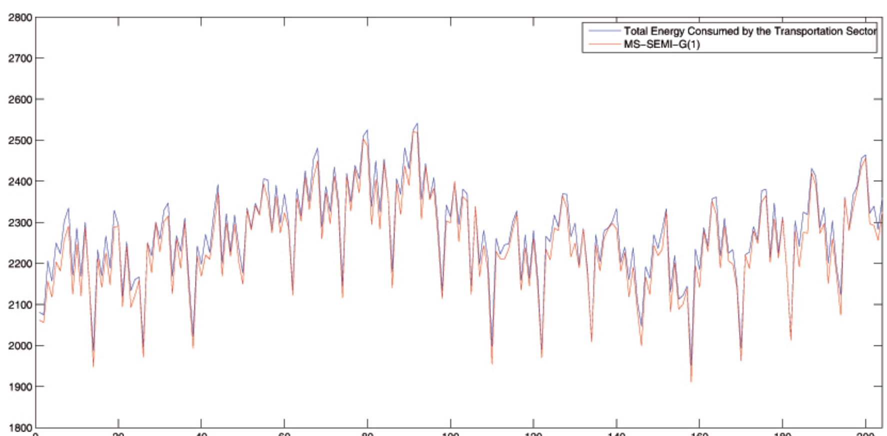

Figure 9: The sample path and the approximated model for Total Energy Consumed by the Transportation Sector.

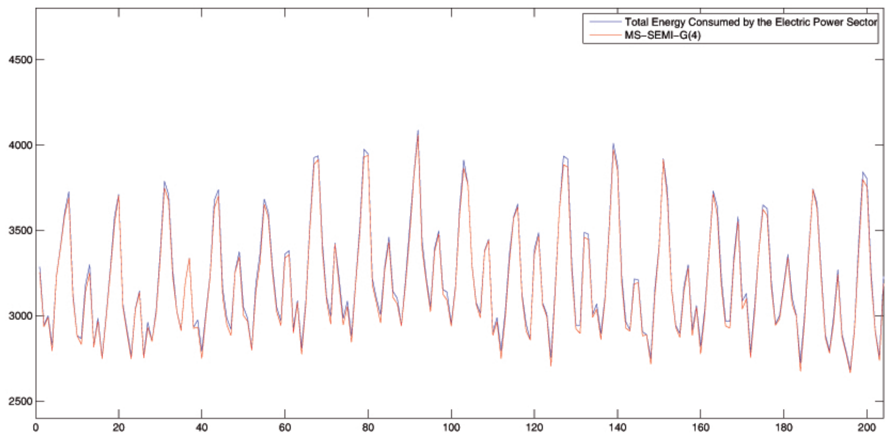

Figure 10: The sample path and the approximated model for Total Energy Consumed by the Electric Power Sector.

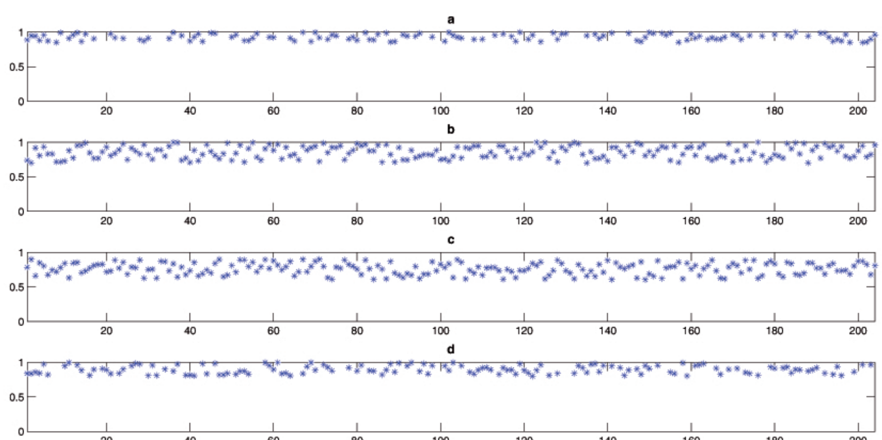

Figure 11: The $max(C_{t1}; C_{t2})$ of Total Energy Consumed by the Commercial Sector according to a. MS-SEMI-$G(1)$, b. MS-SEMI-$G(2)$, c. MS-SEMI-$G(3)$, d. MS-SEMI-$G(4)$.

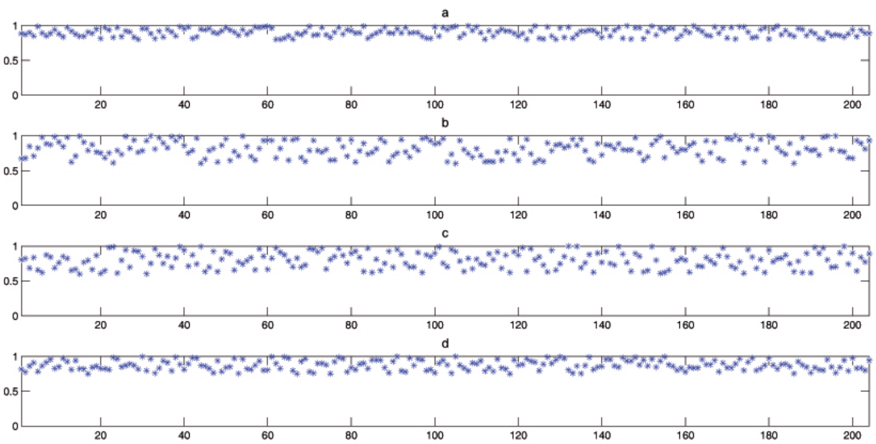

Figure 12: The $max(C_{t1}; C_{t2})$ of Total Energy Consumed by the Residential Sector according to a. MS-SEMI-$G(1)$, b. MS-SEMI-$G(2)$, c. MS-SEMI-$G(3)$, d. MS-SEMI-$G(4)$.

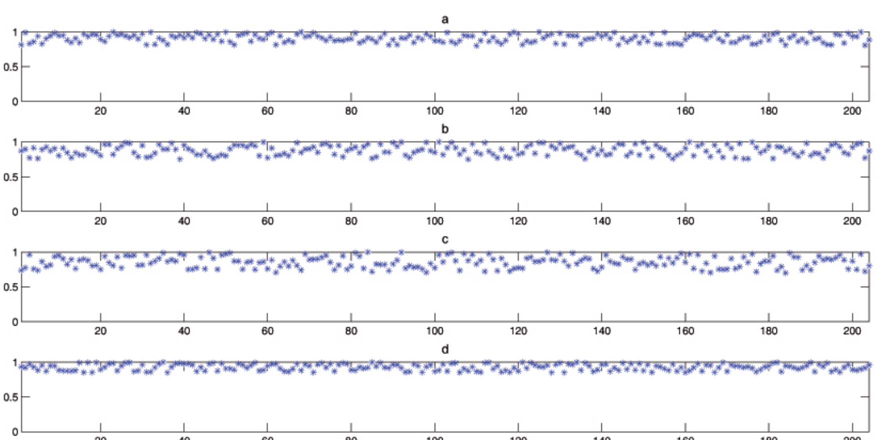

Figure 13: The $max(C_{t1}; C_{t2})$ of Total Energy Consumed by the Industrial Sector according to a. MS-SEMI-$G(1)$, b. MS-SEMI-$G(2)$, c. MS-SEMI-$G(3)$, d. MS-SEMI-$G(4)$.

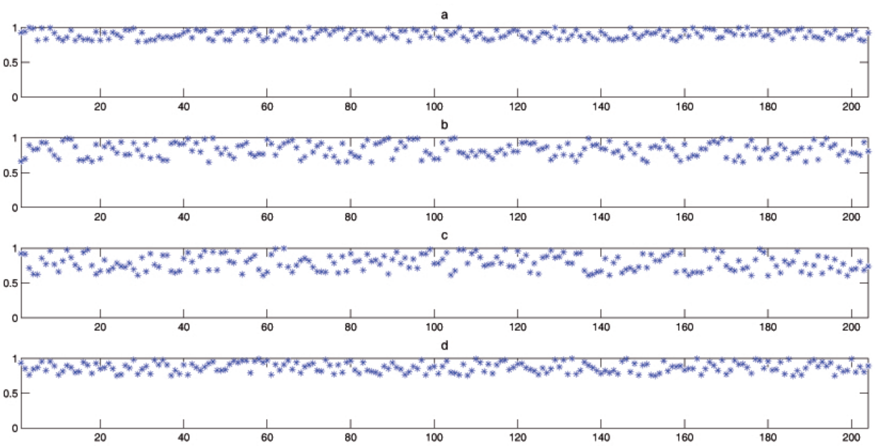

Figure 14: The $max(C_{t1}; C_{t2}; C_{t3})$ of Total Energy Consumed by the Transportation Sector according to a. MS-SEMI-$G(1)$, b. MS-SEMI-$G(2)$, c. MS-SEMI-$G(3)$, d. MS-SEMI-$G(4)$.

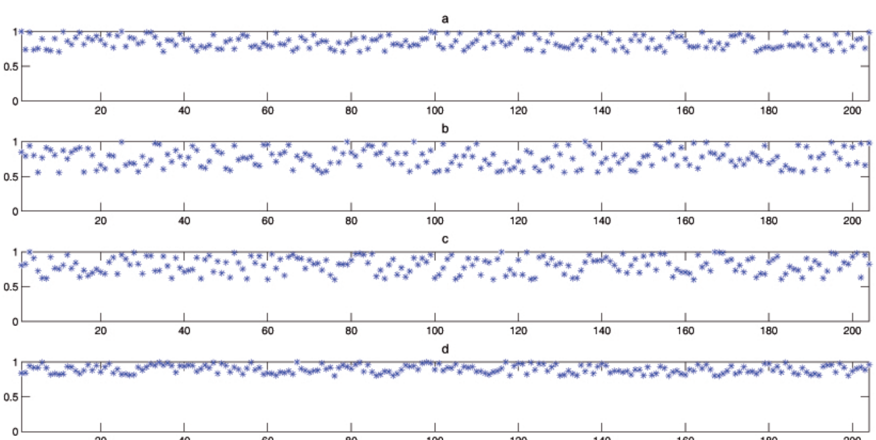

Figure 15: The $max(C_{t1}; C_{t2})$ of Total Energy Consumed by the Electric Power Sector according to a. MS-SEMI-$G(1)$, b. MS-SEMI-$G(2)$, c. MS-SEMI-$G(3)$, d. MS-SEMI-$G(4)$.

445, 264–282.

[2] Adams, F. G., & Shachmurove, Y. (2008). Modeling and forecasting energy consumption in China: Implications for Chinese energy demand and imports in 2020. *Energy Economics*, 30(3), 1263–1278.

[3] Barak, S., & Sadegh, S. S. (2016). Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm. *International Journal of Electrical Power and Energy Systems*, 82, 92–104.

[4] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 542–547.

[5] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

[6] Craig, P. P., Gadgil, A., & Koomey, J. G. (2002). What can history teach us? A retrospective examination of long-term energy forecasts for the United States. *Annual Review of Energy and the Environment*, 27(1), 83–118.

[7] Diebold, F. X., & Lopez, J. A. (1995). Modeling volatility dynamics. *Macroeconometrics: Developments, Tensions and Prospects*, 427–472.

[8] Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 20(1), 134–144.

[9] Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *The Journal of Economic Perspectives*, 15(4), 157–168.

[10] Franke, J., Stockis, J. P., Kamgaing, J. T., & Li, W. K. (2011). Mixtures of nonparametric autoregressions. *Journal of Nonparametric Statistics*, 23, 287–303.

[11] Francq, C., & Zakoian, J. M. (2011). *GARCH Models: Structure, Statistical Inference and Financial Applications*. John Wiley & Sons.

[12] Hamzaçebi, C. (2007). Forecasting of Turkey's net electricity energy consumption on sectoral bases. *Energy Policy*, 35(3), 2009–2016.

[13] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 357–384.

[14] Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and Time Series Analysis* (pp. 202–209). Palgrave Macmillan, UK.

[15] Kaboli, S. H. A., Fallahpour, A., Selvaraj, J., & Rahim, N. A. (2017). Long-term electrical energy consumption formulating and forecasting via optimized gene expression programming. *Energy*, 126, 144–164.

[16] Lee, Y. S., & Tong, L. I. (2011). Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Conversion and Management*, 52(1), 147–152.

[17] Lee, Y.-S., & Tong, L.-I. (2011). Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Conversion and Management*, 52, 147–152.

[18] Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, 9(4).

[19] Nademi, A., & Farnoosh, R. (2014). Mixtures of autoregressive–autoregressive conditionally heteroscedastic models: Semi-parametric approach. *Journal of Applied Statistics*, 41(2), 275–293.

[20] Nademi, A., & Nademi, Y. (2018). Forecasting crude oil prices by a semiparametric Markov switching model: OPEC, WTI, and BRENT cases. *Energy Economics*. doi:10.1016/j.eneco.2018.06.020

[21] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 347–370.

[22] Pao, H. T. (2009). Forecasting energy consumption in Taiwan using hybrid nonlinear models. *Energy*, 34(10), 1438–1446.

[23] Pao, H. T., & Tsai, C. M. (2011). Modeling and forecasting the CO_2 emissions, energy consumption, and economic growth in Brazil. *Energy*, 36(5), 2450–2458.

[24] Qu, W. H., Xu, L., Qu, G. H., Yan, Z. J., & Wang, J. X. (2017). The impact of energy consumption on the environment and public health in China. *Natural Hazards*, 87(2), 675–697.

[25] Saab, S., Badr, E., & Nasr, G. (2001). Univariate modeling and forecasting of energy consumption: The case of electricity in Lebanon. *Energy*, 26(1), 1–14.

[26] Sen, P., Roy, M., & Pal, P. (2016). Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization. *Energy*, 116, 1031–1038.

[27] Sutthichaimethee, P., & Ariyasajjakorn, D. (2017). Forcasting energy consumpyion in short-term and long-term period by using ARIMAx model in the construction and materials sectoe in Thailand. *Journal of Ecological Engineering*, 18(4), 52–59.

[28] Tang, L., Wang, S., He, K. J., & Wang, S. Y. (2014). A novel mode-characteristic-based decomposition ensemble model for nuclear energy consumption forecasting. *Annals of Operations Research*, 234(1), 111–132.

[29] Xu, N., Dang, Y., & Gong, Y. (2017). Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. *Energy*, 118, 473–480.

[30] Zeng, Y.-R., Zeng, Y., Choi, B., & Wang, L. (2017). Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network. *Energy*. doi:10.1016/j.energy.2017.03.094

[31] Wang, Z. X., He, L. Y., & Zheng, H. H. (2019). Forecasting the residential solar energy consumption of the United States. *Energy*, 178, 610–623.

[32] Wu, W., Ma, X., Zeng, B., Wang, Y., & Cai, W. (2019). Forecasting short-term renewable energy consumption of China using a novel fractional nonlinear grey Bernoulli model. *Renewable Energy*, 140, 70–87.