DEVELOPING A FORECASTING MODEL TO ENHANCE THE EFFICIENCY OF SUSTAINABLE DEVELOPMENT POLICY: ENRICHING THE LS-ARIMAX MODEL

Pruethsan Sutthichaimethee[1], Thawatchai Pannarungsee[2], Surapol Suyaporm[3]
Boonton Dockthaisong[4], Naiyana Koetwichai[5]

Abstract

This study aims to develop a forecasting model by adapting the LS-ARIMAX model in order to enhance the efficiency of sustainable development policy. This unique LS-ARIMAX model is actually developed based on a great forecasting model known as ARIMA Model (p,d,q). AS the LS-ARIMAX model suggests, it comes in full name as Long Term and Short Term-Autocorrelation Integrated Moving Average with Exogeneous and Error Correction Mechanism (LS-ARIMAX model). As of many relevant studies are out in review, it reflects that LS-ARIMAX model is a newly-developed model designed to support in a policy formulation and planning of Thailand and other countries. This model of LS-ARIMAX is different compared to other existing models and comes with unique key features; deploying a stationary data, integrating with co-integration analysis, considering exogenous variables, implementing ECT to optimize a long-term forecasting capacity and eliminating an issue of Heteroskedasticity, Multicollinearity, and Autocorrelation. Therefore, the above model becomes an ideal and potential forecasting model to be utilized in the process of national policy formulation and planning as to achieve a sustainability.

Keywords: LS-ARIMAX model, Sustainable Development, Long term, Short term, Exogenous Variables

Introduction

A national development policy and a Sustainable Development Policy (SD policy) in particularly is a policy that aims to grow all aspects in economy, social and environment simultaneously (Office of the National Economic and Social Development Board (NESDB), 2018). Most countries around the world have emphasized and widely adapted such S.D. policy into their specific contexts (National

[1,3,4,5] Mahachulalongkornrajavidyalaya University
[2] Phitsanulok University

Statistic Office Ministry of Information and Communication Technology, 2018) in order to establish a sustainable future for the nation.

As for Thailand, the government has started taking a serious action to develop the economy by promoting an economic development in certain aspect to grow, including the promotion of foreign investment in large projects with a reduction of taxes, a financial aid for certain businesses and an extension of rental contract into long-term duration. All those policies are very much positively impactful as they could lift up a Gross Domestic Product (GDP) at a better position (NESDB, 2018). Moreover, there is a potential of growth in both population and social aspect, and they will continue growing. This can be seen that the economy has a direct relationship with a social move, while it is contradicted with an environment (Department of Alternative Energy Development and Efficiency: DADE, 2018). Hence, it can be further explained in a way that the current environment is deteriorating due to the economic and social development, and it has become a big issue for Thailand (Thailand greenhouse gas management organization, 2018).

Therefore, a key tool for an effective and efficient policy management is a long-term forecasting model to support in decision making of national planning and the nation moving forward in the right direction and sustainable path. As of this study, we choose to study many theories and studies as to produce an effective long-term forecasting model, which later is used as a tool to administer the country.

However, the best forecasting model on energy consumption must also be able to support sustainable development policy planning. From the various relevant studies that have been reviewed, there are different models and forecasting techniques optimized for different forecasting timelines, be it short-term or long-term. Therefore, it is necessary to examine what has been done in this area to increase the quality of the proposed model. In fact, there have been few stream studies exploring total energy consumption. For instance, Zhao, Zhao, and Guo (2016) started to estimate the electricity consumption of Inner Mongolia by deploying GM (1,1) optimized by moth-flame optimization (MFO) with a rolling mechanism from 2010 to 2014. Their study indicated which model could improve the forecasting performance of annual electricity consumption significantly. Li and Li (2017) also initiated a comparative study by using the ARIMA model, GM model, and

ARIMA–GM model to forecast energy consumption in Shandong, China from 2016 to 2020. Upon their analysis in the study, the prediction results showed that the energy demand of Shandong Province between those years will increase at an average annual rate of 3.9 percent. In the same stream, Xiong, Dang, Yao and Wang (2014) proposed a novel GM (gray model) (1,1) model based on optimizing the initial condition in accordance with the new information priority principle to predict China's energy consumption and production from 2013 to 2017. The study produced findings indicating that China's energy consumption and production will keep increasing, as will the gap between them. Furthermore, Panklib, Prakasvudhisarn, and Khummongkol (2015) attempted to forecast electricity consumption in Thailand by using an artificial neural network and multiple linear regression for the years 2010, 2015, and 2020. Based on their estimation from the study, it revealed that the electricity consumption of Thailand in 2010, 2015, and 2020, retrieved from the regression, will reach 160,136, 188,552, and 216,986 GWh, respectively, whereas 155,917, 174,394, and 188,137 GWh were the results obtained from the ANN model. Additionally, an ANN integrated with genetic algorithm was also presented by Azadeh, Ghaderi, Tarverdian, and Saberi (2007) to estimate the electricity consumption in the Iranian agriculture sector in 2008. Based on their results, they observed that the integrated genetic algorithm (GA) and Artificial Neural Network model (ANN) dominated the time series approach from the point of yielding less MAPE (mean absolute percentage error).

By incorporating values of socio-economic indicators and climatic conditions, Günay (2016) modelled artificial neural networks with the use of predicted values of socio-economic indicators and climatic conditions to predict the annual gross electricity demand of Turkey in 2028, which produced a result where the demand would become doubled, accounting for 460 TW in 2018 when compared to the years 2007 to 2013. Dai, Niu and Li (2018) explored energy consumption forecasting in China from 2018 until 2022 by adopting a model of ensemble empirical mode decomposition and least squares support vector machine with the technology of the improved shuffled frog leaping algorithm. Their results showed China's energy consumption to have a significant growth trend. Based on Wang and Li (2017), they tried to find whether China's coal consumption during 2016 to 2020 would be higher or lower than the level of 2014. Here, they optimized a time series model with a comprehensive analysis of data reliability.d According to the analysis, it indicated that the annual Chinese coal consumption during 2016–2020 would be lower than the level of 2014 given that the annual average GDP growth rate was less than 8.2 percent per year. Suganthi and Samuel (2016) developed econometric models to study the influence of the socioeconomic variables on energy consumption in India from 2030 to 2031 and found that the electricity demand depended on the GNP and electricity price, and the total energy requirement was found to be 22.944×10^{15} kJ.

In addition, Xu et al. (2012) analyzed the change of energy consumption and CO_2 emissions in China's cement industry and its driving factors over the period between 1990 to 2009 by applying a log-mean Divisia index (LMDI) method. With such analysis, the study reveals that, by applying the best available technology, an additional energy saving potential of 26% and a CO_2 mitigation potential of 33% can be gained when compared with 2009. Kishita, Yamaguchi, and Umeda (2016) tried to analyze electricity consumption in the telecommunications industry in 2030 by deploying an electricity demand model for the telecommunications industry (EDMoTI). The prediction results pointed out that electricity consumption in 2030 ranged between 0.7–1.6 times larger than the level of 2012 (10.7 TWh per year). To show a lot shorter time of prediction, Zhao, Wang and Lu (2014) conducted a study to forecast the monthly electricity consumption in China by proposing a time-varying-weight combining method; the High-order Markov chain based time-varying weighted average (HM–TWA) method. Apparently, their forecasting performance evaluation showed that the HM–TWA produced a better outcome for the component models and traditional combining methods.

Nonetheless, there are several studies that have examined the total energy demands and its consumption for a longer term of forecasting. For instance, Hamzacebi and Es (2014) implemented optimized grey modeling to forecast the total electric energy demand of Turkey from 2013 to 2025. From their prediction, it reflected that the direct forecasting approach resulted in better predictions than the iterative forecasting approach in estimating the electricity consumption in Turkey. An Improved Gray Forecast Model was also drawn by Mu, Dong, Wang, Ning, and Zhou (2002) to predict CO_2 emissions, energy consumption, and economic growth in China from 2011 and 2020 by using an improved grey model. Based on their prediction results, China's compound annual emissions, energy consumption, and real GDP growth for the predicted years was found to be 4.47%–0.06% and 6.67%, respectively. Furthermore, Zeng, Zhou, and Zhang (2017) proposed a Homologous Grey Prediction Model to predict the energy consumption of China's manufacturing from 2018 to 2024 where their study revealed that the total energy consumption in China's manufacturing was slowing down, however, the amount was still too large. Additionally, Jiang, Yang and Li (2018) adapted a metabolic grey model (MGM), autoregressive integrated moving average (ARIMA), MGM–RIMA, and back propagation neural network (BP) to forecast energy demand from 2017 to 2030. From their estimation, it showed that India's energy consumption would increase by 4.75% a year in the next 14 years at a 5 percent growth rate. By using the same forecasting model, but improved, Ediger and Akar (2007) analyzed the primary energy demand by fuel in Turkey from 2005 to 2020 using the autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) methods to estimate the above demand, and showed that the average annual growth rates of individual energy sources and total primary energy would decrease in all cases, except wood, and the animal–plant went negative.

Furthermore, Ekonomou (2010) developed an artificial neural network (ANN) to estimate the Greek long-term energy consumption from 2005 until 2008, 2010, 2012, and 2015. Overall, the study has constituted an accurate yet better tool for the forecasting problem in Greek long-term energy consumption. In addition, Ardakani and Ardehali (2014) utilized an IPSO (improved particle swarm optimization)–ANN model to forecast EEC (electrical energy consumption) for Iran and the U.S. from 2010 to 2030. Here, it resulted in the mean absolute percentage error of 1.94 and 1.51% for Iran and the U.S., respectively. In the context of Thailand, a study of characteristics and factors towards energy consumption was conducted by Supasa et al. (2017), who explored five household group energy consumption characteristics and seven driving forces of growth in residential energy consumption from 2000 to 2010 by applying the energy input–output method. Their calculations indicated that about 70 percent of total residential energy consumption was an indirect energy consumption form from consuming products and services. Another investigation of the impacts of urban land use on energy consumption in China from 2000 to 2010 was undertaken by Zhao, Thinh, and Li (2017). They used a panel data analysis with nighttime light (NTL) data estimation. With their study on sight, it has shown that an increase in the irregularity of urban land forms and the expansion of urban land will accelerate energy consumption, which indicates the relationship between urban growth and energy consumption. Along the same lines, Tian, Xiong, and Ma (2017) evaluated the potential impacts of China's industrial structure on energy consumption by deploying a fuzzy multi-objective optimization model based on the input–output model from 2015 until 2020. From their analysis, they concluded that the industrial structure adjustment had great potential in energy conservation, and such an adjustment could save energy by 19% (1129.17 Mtce) at the average annual growth rate of 7% GDP. Ayvaz and Kusakci (2017) employed a nonhomogeneous discrete grey model (NDGM) to forecast electricity consumption from 2014 to 2030. In their findings, they proved that the grey model (GM) proposed produced a better forecasting performance.

With previous studies in review, there are no other researches discussing about a forecasting in a longer-term (1-20 years) at better efficient. While available forecasting models are said to be general and common, including Multiple Linear Regression model (MLR model), the Artificial Neural Network model (ANN model), the Back Propagation neural network model (BP Model), the Gray model (GM(1,1)), the Autoregressive and Moving Average model (ARMA), and the Autoregressive Integrated and Moving Average model (ARIMA). To add on, there is a lack of consideration in term of specific context as well as an attempt of elimination of Heteroscedasticity, Multicollinearity, and Autocorrelation. Therefore, we realize to develop a forecasting model called LS-ARIMAX model in filling the above gap. Simultaneously, the model can be used to support in the national policy formulation as to achieve a sustainability. Whereas a discussion of Materials and Methods is placed in the following section.

Materials and Methods

1. Autoregressive Model and Moving Average Model

The autoregressive model and moving average model or Box-Jenkins are two models that emphasize only the stationary data (Dickey and Fuller,1981) described as follows.

In the case of the random seasonal process, it is an uncertain or specific seasonal process. For instance, a country encounters a political conflict for the past many years. At the same time, a demonstration occurs at the second quarter. This situation causes a sales drop. However, the political conflicts may seem stable this year in the same quarter. Thus, the sale is consistent. Here, the seasonal that took place last year at the second quarter temporarily affects this year's second quarter, or can be called the stationary seasonal process. In this case, it is not necessary to drop off the season, but it can be incorporated into the model. This can be called the seasonal autoregressive moving average or seasonal ARMA in short. The model is explained as below (MacKinnon, 1991; Enders, 2010):

X_t is a quarter time series and falls under a stationary seasonal process. This time series X_t can be written as:

$$X_t = A_1 X_{t-4} + v_t \qquad |A_1| < 1 \qquad (1)$$

Where v_t is a random error variable, which is a white noise. The above equation is a AR (4) model where the coefficient of X_{t-1}, X_{t-2} and X_{t-3} is 0, and $|A_1|<1$ is the condition indicating the stationary seasonal process in time series X_t. If this X_t is brought to find an average value, a variance of the Theoretical Autocorrelation Function (TAC) and Theoretical Partial Autocorrelation Function (TPAC) is computed through the following equation:

Let $\mu = 0$, Variance $\gamma_0 = \dfrac{\sigma^2}{1-A_1^2}$, the TAC is pointed in Equation (2) and TPAC is drawn in Equation (3).

$$\rho_k = \begin{cases} (A_1)^{\frac{k}{4}}, & k = 0,4,8,... \\ 0, & \text{when it is other case} \end{cases} \qquad (2)$$

$$\phi_{kk} = \begin{cases} \rho_4 & k = 4 \\ 0, & \text{when it is other case} \end{cases} \qquad (3)$$

Since $|A_1|<1$, and when considering Equation (2), it can be concluded that if $0 < A_1 < 1$, TAC will exponentially reduce at time 4, 8, 12, ..., and if $-1 < A_1 < 0$ as time slowly passes, TAC will be exponentially up and down at time 4, 8, 12, ... If $|A_1|$ is closely approaching 1, a seasonal pattern will be clearer and last long. In contrast, if $|A_1|$ is close to 0, the pattern will disappear. Equation (3) shows that the TPAC is not equivalent to 0.

Based on Equation (1), it reflects only on the impact of season in AR, but the time series can be as the ARMA in practice, and this can be written as below:

$$A(L^s)X_t = B(L^s)v_t \tag{4}$$

where s is the time duration of season.

$$A(L^s) = 1 - A_1 L^s - A_2 L^{2s} - \dots A_p L^{ps} \tag{5}$$

$$B(L^s) = 1 - B_1 L^s - B_2 L^{2s} - \dots B_Q L^{Qs} \tag{6}$$

We consider Equation (4) as the pure seasonal $ARMA$ model at $(P,Q)_s$. In practice, it is possible that X_t is in AR(1), together influencing the season as the equation below:

$$X_t = A_1 X_{t-1} + A_1 X_{t-4} + v_t, \quad |\alpha_1| < 1 \text{ and } |A_1| < 1 \tag{7}$$

Equation (7) indicates the influence of seasonal process when $s = 4$, and time series X_t in quarter 2 is related to quarter 1, while quarter 2 of this year is related to quarter 2 last year.

In the meantime, Equation (4) shows X_t in some seasons of the $ARMA$ model at $(P,Q)_s$ or $ARMA\,(p,q)$ together, and v_t in $ARMA\,(p,q)$ is as follows:

$$v_t = \frac{\beta(L)}{\alpha(L)}\varepsilon_t \tag{8}$$

where ε_t is the random error variable with white noise $\alpha(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots \alpha_p L^p$ and $\beta(L) = 1 - \beta_1 L - \beta_2 L^2 - \dots \beta_q L^q$. Therefore, Equation (4) can be drawn as below:

$$A(L^s)\alpha(L)X_t = B(L^s)\beta(L)\varepsilon_t \tag{9}$$

Equation (9) will be called the multiplicative seasonal $ARMA$ model at $(p,q)\times(P,Q)_s$, and can be denoted as $ARMA\,(p,q)(P,Q)_s$ or $ARMA\,(p,q)\times(P,Q)_s$.

In order to make a better understanding about $ARMA\,(p,q)(P,Q)_s$, $ARMA\,(0,1)(0,1)_s$ is given to explain further by the following details.

According to $ARMA\,(0,1)(0,1)_s$, it can tell that $p = 0$ and $q = 1$. This means $\alpha(L) = 1$ and $\beta(L) = 1 - \beta_1 L$, respectively. Additionally, it explains that $P = 0$ and $Q = 1$ while the general time range is s, and that retrieves $A(L^s) = 1$ and $B(L) = 1 - B_1 L^s$, respectively. Hence, $ARMA\,(0,1)(0,1)_s$ model is written as an equation below:

$$A(L^s)\alpha(L)X_t = B(L^s)\beta(L)\varepsilon_t$$

$$(1)(1)X_t = (1 - B_1L^s)(1 - \beta_1L)\varepsilon_t$$

$$X_t = (1 - \beta_1L - B_1L^s + \beta_1B_1L^{s+1})\varepsilon_t$$

$$X_t = \varepsilon_t - \beta_1\varepsilon_{t-1} - B_1\varepsilon_{t-s} + \beta_1B_1\varepsilon_{t-s-1} \qquad (10)$$

Equation (10) constitutes like the moving average model. If X_t in the $ARMA\ (0,1)(0,1)_s$ pattern is taken to compute the average, variance, and TAC, it can be obtained from the equation below.

Let $\mu = 0$, $\gamma_0 = (1+\beta_1^2)(1+B_1^2)\sigma^2$

$$\rho_k = \begin{cases} \dfrac{-\beta_1}{(1+\beta_1^2)} & k = 1 \\ \dfrac{-B_1}{(1+\beta_1^2)} & k = s \\ \dfrac{\beta_1B_1}{(1+\beta_1^2)(1+B_1^2)} & k = s-1, s+1 \\ 0 & k \neq 0,1,s-1,s,s+1 \end{cases} \qquad (11)$$

The pattern of TAC can be considered from Equation (11). If $s = 4$, TAC will not be 0 at time 1, 2, 4, 5, and is equal to 0 at other times. If $s = 12$, TAC will not be 0 at time 1, 11, 12, 13, and it becomes 0 at other times.

2. ARMA (p,q) Model

Consider the ARMA(p,q) (Cryer, 2008) model is written as

$$X_t = \alpha_0 + \alpha_1X_{t-1} + \alpha_2X_{t-2} + \ldots + \alpha_pX_{t-p} + \varepsilon_t - \beta_1\varepsilon_{t-1} - \beta_2\varepsilon_{t-2} - \ldots - \beta_q\varepsilon_{t-q}$$

where $t = 1,2,\ldots,T$. If we consider at time T, the $ARMA\ (p,q)$ model becomes

$$X_T = \alpha_0 + \alpha_1X_{T-1} + \alpha_2X_{T-2} + \ldots + \alpha_pX_{T-p} + \varepsilon_T - \beta_1\varepsilon_{T-1} - \beta_2\varepsilon_{T-2} - \ldots - \beta_q\varepsilon_{T-q} \qquad (12)$$

or it can be written in another form as

$$\alpha(L)X_T = \alpha_0 + \beta(L)\varepsilon_T \qquad (13)$$

where $\alpha(L) = 1 - \alpha_1L - \alpha_2L - \ldots - \alpha_pL^p$ and $\beta(L) = 1 - \beta_1L - \beta_2L - \ldots - \beta_qL^q$ while the information at time T can be replaced by $I_T = \{X_1 \ldots, X_{T,}\varepsilon_1 \ldots, \varepsilon_T\}$. Equation (12) produces X_{T+1} and X_{T+2} as the below equation.

$$X_{T+1} = \alpha_0 + \alpha_1 X_T + \alpha_2 X_{T-1} + \ldots + \alpha_p X_{T+(1-p)}$$
$$+\varepsilon_{T+1} - \beta_1 \varepsilon_T - \beta_2 \varepsilon_{T-1} - \ldots - \beta_q \varepsilon_{T+(1-q)}$$

$$X_{T+2} = \alpha_0 + \alpha_1 X_{T+1} + \alpha_2 X_T + \ldots + \alpha_p X_{T+(2-p)}$$
$$+\varepsilon_{T+2} - \beta_1 \varepsilon_{T+1} - \beta_2 \varepsilon_T - \ldots - \beta_q \varepsilon_{T+(2-q)}$$

A forecasting of time series 1 and 2 from $ARMA(p,q)$ can be drawn below.

$$\hat{X}_T(1) = \mathrm{E}(X_{T+1} | I_T)$$

$$= \alpha_0 + \alpha_1 X_T + \alpha_2 X_{T-1} + \ldots + \alpha_p X_{T+(1-p)} - \beta_1 \varepsilon_T - \ldots - \beta_q \varepsilon_{T+(1-q)}$$

$$\hat{X}_T(2) = \mathrm{E}(X_{T+2} | I_T)$$

$$= \alpha_0 + \alpha_1 \hat{X}_T(1) + \alpha_2 X_T + \ldots + \alpha_p X_{T-(p-2)} - \beta_2 \varepsilon_T - \ldots - \beta_q \varepsilon_{T+(2-q)}$$

While we can formulate X_{T+j} in general form as:

$$X_{T+j} = \alpha_0 + \alpha_1 X_{T+(j-1)} + \ldots + \alpha_p X_{T+(j-p)} + \varepsilon_{T+J} - \beta_1 \varepsilon_{T+(J-1)} - \ldots - \beta_q \varepsilon_{T+(J-q)} \quad (14)$$

While we can formulate $\hat{X}_T(j)$ in general form as:

$$\hat{X}_T(j) = \mathrm{E}(X_{T+j} | I_T)$$

$$\hat{X}_T(j) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \hat{X}_T(j-i) - \sum_{i=1}^{q} \beta_i \varepsilon_T(j-i) \quad (15)$$

Where

$$\hat{X}_T(j-1) = X_{T+(j-i)} \text{ when } j-i \le 0$$

$$\varepsilon_T(j-i) = \begin{cases} \varepsilon_{T+(j-i)} & ,if \quad j-i \le 0 \\ 0 & ,if \quad j-i > 0 \end{cases}$$

Besides, we can also check $ARMA\,(1,1)$ as the above explanation when $j \to \infty$, and the forecasting can be executed from:

$$\hat{X}_T(j) = \frac{\alpha_0}{1 - \alpha_1 - \ldots - \alpha_p} \quad (16)$$

Equation (16) tells us when to forecast further where the forecasting result will approach $\frac{\alpha_0}{1 - \alpha_1 - \ldots - \alpha_p} = \mathrm{E}(X_t)$, and this is the average of time series X_t in the $ARMA(p,q)$ model.

In addition, j-step ahead forecast error and its variance can be easily executed when altering the $ARMA(p,q)$ model into $MA(\infty)$ as explained below.

Since the time series X_t is stationary, Equation (13) can be rewritten as:

$$X_T = \frac{\alpha_0}{\alpha(L)} + \frac{\beta_0}{\alpha(L)}\varepsilon_T \quad (17)$$

When considering $\frac{\alpha_0}{\alpha(L)} = \frac{\alpha_0}{1-\alpha_1-\ldots-\alpha_p} = \mathrm{E}(X_t)$, which is the average. When $\frac{\beta(L)}{\alpha(L)}\varepsilon_T$ is considered, it shows a relativeness to ε_T, and that $\frac{\beta(L)}{\alpha(L)}\varepsilon_T = \frac{1-\beta_1 L-\cdots-\beta_q L}{1-\alpha_1 L-\cdots-\alpha_p L}\varepsilon_T$ with inconstancy in value.

Let $\frac{\alpha_0}{\alpha(L)} = \mu$

$$\frac{\beta(L)}{\alpha(L)} = \varphi(L) = 1+\varphi_1 L+\varphi_1^2 L^2+\ldots \quad (18)$$

Thus, Equation (17) with $ARMA(p,q)$ can be formulated into the $MA(\infty)$ form as:

$$X_T = \mu+\varepsilon_T+\varphi_1\varepsilon_{T-1}+\varphi_1\varepsilon_{T-2}+\ldots, \text{ or } X_T = \mu+\varphi(L)\varepsilon_T \quad (19)$$

We call this $\varphi_i\,(i=1,2,\ldots)$ as the impulse response function of the $ARMA$ model. When the time series X_T is stationary, $\varphi_1,\varphi_2,\varphi_3,\ldots$ will rapidly decrease exponentially. However, Equation (19) can be used to compute the j-step ahead forecast error and its variance through the following description.

From Equation (19), the time series X_{T+1}, X_{T+2}, and X_{T+3} can be written as follows:

$$X_{T+1} = \mu+\varepsilon_{T+1}+\varphi_1\varepsilon_T+\varphi_2\varepsilon_{T-1}+\ldots$$

$$X_{T+2} = \mu+\varepsilon_{T+2}+\varphi_1\varepsilon_{T+1}+\varphi_2\varepsilon_T+\varphi_3\varepsilon_{T-1}+\ldots$$

$$X_{T+3} = \mu+\varepsilon_{T+3}+\varphi_1\varepsilon_{T+2}+\varphi_2\varepsilon_{T+1}+\varphi_3\varepsilon_T+\varphi_4\varepsilon_{T-1}+\ldots$$

Then, the forecasting value of 1, 2, and 3 ahead is derived from the following:

$$\hat{X}_T(1) = \mathrm{E}(X_{T+1}|I_T) = \mu+\varphi_1\varepsilon_T+\varphi_2\varepsilon_{T-1}+\ldots$$

$$\hat{X}_T(2) = \mathrm{E}(X_{T+2}|I_T) = \mu+\varphi_2\varepsilon_T+\varphi_3\varepsilon_{T-1}+\ldots$$

$$\hat{X}_T(3) = \mathrm{E}(X_{T+3}|I_T) = \mu+\varphi_3\varepsilon_T+\varphi_4\varepsilon_{T-1}+\ldots$$

While its error at 1, 2, and 3 ahead is as follows:

$$e_T(1) = X_{T+1} - \hat{X}_T(1) = \varepsilon_{T+1}$$

$$e_T(2) = X_{T+2} - \hat{X}_T(2) = \varepsilon_{T+2} + \varphi_1 \varepsilon_{T+1}$$

$$e_T(3) = X_{T+3} - \hat{X}_T(3) = \varepsilon_{T+3} + \varphi_1 \varepsilon_{T+2} + \varphi_2 \varepsilon_{T+1}$$

Moreover, its variance at 1, 2, and 3 ahead is as below:

$$Var(e_T(1)) = \sigma^2$$

$$Var(e_T(2)) = (1 + \varphi_1^2)\sigma^2$$

$$Var(e_T(3)) = (1 + \varphi_1^2 + \varphi_2^2)\sigma^2$$

However, the j-step ahead forecast error and its variance can be drawn in an equation as follows:

$$e_T(j) = \varepsilon_{T+j} + \varphi_1 \varepsilon_{T+(j-1)} + \varphi_2 \varepsilon_{T+(j-2)} + \ldots + \varphi_{j-1} \varepsilon_{T+1} \quad (20)$$

$$Var(e_T(j)) = (1 + \varphi_1^2 + \varphi_2^2 + \ldots \varphi_{j-1}^2)\sigma^2 \quad (21)$$

3. A Forecasting Model with LS-ARIMAX model

In the construction of the LS-ARIMAX model for forecasting, the autoregressive model (AR) and moving average model (MA) were basically integrated to first structure an ARIMA model. Once the ARIMA model was obtained, it was then applied to generate the LS-ARIMAX model together with a co-integration test at the same level for every variable in the equations. In addition, there was also an adaptation of an error correction mechanism test into this particular model as discussed below.

It is a notion that differentiating at d with a particular time series will make non-stationary a stationary. With such differentiation applied in the Box–Jenkins model, it can become known as $ARIMA(p,d,q)$ (Harvey, 1989).

To create a better understanding, X_t is denoted as the non-stationary time series, where $Z_t = \Delta X_t = X_t - X_{t-1}$ is the stationary time series. Here, a proper model for this time series X_t is $ARIMA(1,1,0)$ and it can be written as:

$$\Delta X_t = \alpha_0 + \alpha_1 \Delta X_{t-1} + \varepsilon_t \qquad \text{where } t = 1,2,\ldots,T \quad (22)$$

or $Z_t = \alpha_0 + \alpha_1 Z_{t-1} + \varepsilon_t$ where $t = 1,2,\ldots,T$ (23)

If time at T is taken into account, the $ARIMA(1,1,0)$ becomes:

$$Z_T = \alpha_0 + \alpha_1 Z_{T-1} + \varepsilon_T \tag{24}$$

and $X_1, X_2, \ldots X_T$ (or denoted as I_T) is now known for their value.

When using Equation (24), we will forecast $\hat{Z}_{T+1}, \hat{Z}_{T+2}, \hat{Z}_{T+3}$ from the following equation.

$$\left.\begin{array}{lcccc} \hat{Z}_{T+1} & = & \alpha_0 & + & \alpha_1 \Delta X_T \\ \hat{Z}_{T+2} & = & \alpha_0 & + & \alpha_1 \Delta \hat{X}_{T+1} \\ \hat{Z}_{T+3} & = & \alpha_0 & + & \alpha_1 \Delta \hat{X}_{T+2} \\ & & \vdots & & \\ \hat{Z}_{T+j} & = & \alpha_0 & + & \alpha_1 \Delta \hat{X}_{T+(j-1)} \end{array}\right\} \tag{25}$$

At this point, it can be noticed that the result of $ARIMA(1,1,0)$ based on Equation (24) will have a difference at level 1 ($\hat{Z}_{T+j} = \Delta \hat{X}_{T+j}$), but the forecasting of time series $\hat{X}_{T+j}$. From Equation (25), it can be seen that $\hat{X}_T$ may not need forecasting. This is because the true information is known, which is X_T, hence, the forecasting result of $\hat{X}_{T+1}$ can be computed from.

$$\hat{X}_{T+1} = X_T + \hat{Z}_{T+1} \tag{26}$$

While the forecasting result of $\hat{X}_{T+2}, \hat{X}_{T+3}, \ldots, \hat{X}_{T+j}$ can be calculated as follows:

$$\hat{X}_{T+2} = X_T + \hat{Z}_{T+1} + \hat{Z}_{T+2} \tag{27}$$

$$\hat{X}_3 = X_T + \hat{Z}_{T+1} + \hat{Z}_{T+2} + \hat{Z}_{T+3}$$

$$\vdots$$

$$\hat{X}_{T+j} = X_T + \sum_{k=1}^{j} \hat{Z}_{T+k} \tag{28}$$

Another method to compute the forecasting result of $\hat{X}_{T+1}, \hat{X}_{T+2}, \ldots$ from $ARIMA(1,1,0)$ is to reform the equation as below:

$$X_t - X_{t-1} = \alpha_0 + \alpha_1 (X_{t-1} - X_{t-2}) + \varepsilon_t$$

$$X_t = \alpha_0 + (\alpha_1 + 1)X_{t-1} - \alpha_1 X_{t-2} + \varepsilon_t \text{ where } t = 1,2,\ldots,T \quad (29)$$

$$\text{or } X_t = \phi_0 + \phi_1 X_{t-1} - \varphi_2 X_{t-2} + \varepsilon_t \text{ where } t = 1,2,\ldots,T \quad (30)$$

where $\varphi_0 = \alpha_{0,} \varphi_1 = \alpha_1 + 1, \varphi_2 = -\alpha_1$. We can use Equation (30) to forecast $\hat{X}_{T+j}$. The forecast error and its variance can be found by applying the mentioned concept discussed earlier. The difference made here is that the variance of forecast error will be higher yet inconstant because the time series X_t is non-stationary.

As for forecasting with $ARIMA(p,1,q)$, it can be applied by Equations (28) or (30), but the equation transformation is complicated.

Assuming the $ARIMA(p,1,q)$ model is written as below:

$$\Delta X_t = \alpha_0 + \alpha_1 \Delta X_{t-1} + \alpha_2 \Delta X_{t-2} + \ldots + \alpha_p \Delta X_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots - \beta_1 \varepsilon_{t-q}$$

$$\begin{aligned} X_t - X_{t-1} &= \alpha_0 + \alpha_1 (X_{t-1} - X_{t-2}) + \alpha_2 (X_{t-2} - X_{t-3}) \\ &+ \ldots + \alpha_p (X_{t-p} - X_{t-p-1}) + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} \\ &- \ldots - \beta_1 \varepsilon_{t-q} \end{aligned}$$

$$\begin{aligned} X_t &= \alpha_0 + (\alpha_1 + 1)X_{t-1} + (\alpha_2 - \alpha_1)X_{t-2} + (\alpha_3 - \alpha_2)X_{t-3} \\ &+ \ldots + (\alpha_p - \alpha_{p-1})X_{t-p} - \alpha_p X_{t-p-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1} \\ &- \beta_2 \varepsilon_{t-2} - \ldots - \beta_1 \varepsilon_{t-q} \end{aligned}$$

Or rewritten as:

$$\begin{aligned} X_t &= \varphi_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varphi_3 X_{t-3} + \ldots + \varphi_p X_{t-p} \\ &+ \varphi_{p+1} X_{t-p-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots - \beta_1 \varepsilon_{t-q} \end{aligned} \quad (31)$$

Where $\varphi_0 = \alpha_{0,} \varphi_1 = \alpha_1 + 1, \varphi_j = \alpha_j - \alpha_{j-1}$ and $\varphi_{p+1} = -\alpha_p$

When the ARIMA model is obtained, it can then be used to construct the LS-ARIMAX model. The construction is explained below.

(1) For the Long Term-ARIMAXS Model (LS-ARIMAX model), we have adapted the concept from the basic models including autoregressive (AR), integrate (I), and moving average (MA) models [32]. This LS-ARIMAX model was examined for the unit

root test and variable selection for stationary into this model formation. We have determined the Level (I(0)) or first difference (I(1)) in order to analyze co-integration (Johansen, S.; Juselius,1990). This point of analysis must reflect the relationship at the same level. However, this LS-ARIMAX model must consist of co-integration and an error correction mechanism test (ECT) to increase efficiency and the zero error in the model. In addition, the LS-ARIMAX model comes with the suitability of future application in different areas in line with the policy of a particular country. This is due to the difference of the LS-ARIMAX model with other models so that the ARIMA model focuses on the variables of Autoregressive (AR), Integrated (I) , and Moving Average (MA) only at time t-i especially in past data. In this paper, the LS-ARIMAX model differs from other old models due to the emphasis of Exogeneous Variables (Exogeneous Variables: $\sum_{i=1}^{p} Y_{t-i}$), which is believed to be an important yet appropriate variable in the study. As for the reason, it is the influential variable that can affect the dependent variable. Additionally, the LS-ARIMAX model uses the variable of Autoregressive (AR), Integrated (I), and Moving Average (MA) during time t-i in the study's model. The LS-ARIMAX model utilizes the co-integration and error correction mechanism test from the theory of Johansen Juselius (1990) in order to increase the effectiveness of the model. The co-integration model and error correction mechanism model can be explained below.

(2) If the variables are stationary at the first difference, cointegration analysis, which is the method proposed by Johansen and Juselius (1990), will be used to examine the long-term relationship between the variables. The essence of cointegration is that the linear combination of variables is stationary. Cointegration tests also require that all variables are integrated in the same order. If series X_t and series Y_t have a long-term equilibrium relationship, we can use the following formula to conduct the cointegration test.

$$X_{t-i} = \alpha_0 + \alpha_1 Y_{t-i} + \varepsilon_t \quad (32)$$

Where ε_t is the residual term and t denotes time. If ε_t is a stable series, we can conclude that the series X_{t-i} and Y_{t-i} are co-integrated. Otherwise, a co-integration relationship does not exist between series X_{t-i} and Y_{t-i}.

However, when taking stationary data for co-integration analysis, the Error correction mechanism test becomes necessary. We can find that the change of X_t not only depends on the change of Y_t but also depends on the change of the last period Y_{t-1} and X_{t-1}. Considering the non-stationarity, the OLS test cannot be used to perform the regression. Therefore, Equation (32) can deform to the equation below.

$$\Delta X_t = \beta_1 \Delta Y_t - (1-\delta)(X_{t-1} - \frac{\beta_0}{1-\delta} - \frac{\beta_1 + \beta_2}{1-\delta} Y_{t-1}) + \varepsilon_t \quad (33)$$

The LS-ARIMAX model is written below.

$$\begin{aligned} X_t &= \varphi_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varphi_3 X_{t-3} + \ldots + \varphi_p X_{t-p} \\ &+ \varphi_{p+1} X_{t-p-1} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots - \beta_1 \varepsilon_{t-q} \\ &+ \sum_{i=1}^{p} Y_{t-i} + \sum_{i=1}^{p} ECT_{t-i} \end{aligned} \quad (34)$$

Let $\varphi_0 = \alpha_0, \varphi_1 = \alpha_1 + 1, \varphi_j = \alpha_{j-1}, \varphi_{p+1} = -\alpha_p$, $\sum_{i=1}^{p} Y_{t-i}$ = exogenous variables, which are stationary at the level and $\sum_{i=1}^{p} ECT_{t-i}$ = the error correction mechanism test.

Equation (34) indicates the components of the LS-ARIMAX model comprised of (1) Autoregressive variables (AR), (2) Moving Average (MA), (3) exogenous variables ($\sum_{i=1}^{p} Y_{t-i}$), and (4) error correction mechanism $\sum_{i=1}^{p} ECT_{t-i}$. The LS-ARIMAX model is built and developed with the assurance of which Heteroscedasticity, Multicollinearity, and Autocorrelation are free. There is also an analysis of period identification with the Q-statistics test as to ensure that the model is not spurious while it becomes efficient in the forecasting with fewer errors. The model is then able to be applied in a different context and management policy.

However, as of the LS-ARIMAX model, it is a newly developed method completed by adapting various concepts from the general models including autoregressive (AR), integrate (I), and moving average (MA). For the variable selection criterion, the variables must only be causal factors or stationary at the same level. The stationary level I(0) or first difference I(1) in order to test the unit root test. Upon the right variables, we fulfill the criterion. Those variables are put forth for a co-integration test. When they are found to be co-integrated, they are then used in structuring the LS-ARIMAX model (p, d, q, Xi, $ECT_{(t-1)}$) with an appropriateness check of period (t-i) through the implementation of a white noise process by the Q test statistic method. In this paper, the LS-ARIMAX model (p, d, q, Xi, $ECT_{(t-1)}$) must not be free from heteroskedasticity, multicollinearity, and autocorrelation. Testing the LS-ARIMAX model (p, d, q, Xi, $ECT_{(t-1)}$) can be done based on the mean absolute percentage error (MAPE) and the root mean square error (RMSE) and compare MAPE and RMSE with existing models. Once the model is obtained, forecasting the future is the next essential step. We have combined the data set by using Microsoft Office Excel. In addition, the model is deployed an EViews 9.5 software and it flows like below.

1. Place the stationary variables at the same level in the analysis of the long-term relationship based on the Johansen Juselius concept.

2. Create a forecasting model by adapting the advance statistics of the so-called LS-ARIMAX model with full consideration of the relationship of all causal variables both in terms of the error correction mechanism test and the co-integration test.

3. Examine the goodness of fit in two aspects: (1) Appropriateness check of period (t-i) through the implementation of a white noise process by the Q test statistic method, (2) performance test for the LS-ARIMAX model based on the mean absolute percentage error (MAPE) and the root mean square error (RMSE) , and we compare those two values derived from the LS-ARIMAX model with the existing model including the Multiple Linear Regression model (MLR model), the Artificial Neural Network model (ANN model), the Back Propagation neural network model (BP Model), the Gray model (GM(1,1)), the Autoregressive and Moving Average model (ARMA), and the Autoregressive Integrated and Moving Average model (ARIMA).

4. Forecast final energy consumption by using the LS-ARIMAX model.

Conclusion

Establishing a long-term forecasting model (1-20 years) is a difficult task due to several factors with a possibility of failure, and that may allow the model become spurious. If an improper forecasting model is used for the national planning and policy formulation in a long run, it will get the planning failed, cost a massive damage to a nation, and create a difficulty to tackle the damage. Therefore, in choosing a forecasting model, the model must be efficient and functional in a context of a country or situation. However, such a model establishment requires a serious, thorough attention and deep understanding in theories and modeling as so that model can be developed and functional. Also, such model can be a useful tool for various stages of national planning, while achieving a sustainable future. Not to mention, it can become a good guidance and foundation for future research.

Bibliography

Office of the National Economic and Social Development Board (NESDB). Available online: http://www.nesdb.go.th/nesdb_en/more_news.php?cid=154&filename=index (accessed on 1 August 2018).

National Statistic Office Ministry of Information and Communication Technology. Available online: http://web.nso.go.th/index.htm (accessed on 1 August 2018).

Department of Alternative Energy Development and Efficiency. Available online: http://www.dede.go.th/ewtadmin/ewt/dede_web/ewt_news.php?nid=47140 (accessed on 1 August 2018).

Achawangkul, Y. Thailand's Alternative Energy Development Plan, Available online: http://www.unescap.org/sites/default/files/MoE%20_%20AE%20policies.pdf (accessed on 1 August 2018).

Thailand greenhouse gas management organization (public organization). Available online: http://www.tgo.or.th/2015/thai/content.php?s1=7&s2=16&sub3=sub3 (accessed on 1 August 2018).

Zhao, H.R.; Zhao, H.R.; Guo, S. (2016). Using GM (1,1) Optimized by MFO with Rolling Mechanism to Forecast the Electricity Consumption of Inner Mongolia. *Appl. Sci*., 6, 20.

Li, S.; Li, R. (2017). Comparison of forecasting energy consumption in Shandong, China Using the ARIMA model, GM model, and ARIMA-GM model. *Sustainability*, 9: 1181.

Xiong, P.P.; Dang, Y.G.; Yao, T.X.; Wang, Z.X. (2014). Optimal modeling and forecasting of the energy consumption and production in China. *Energy*, 77: 623–634.

Panklib, K.; Prakasvudhisarn, C.; Khummongkol, D. (2015). Electricity Consumption Forecasting in Thailand Using an Artificial Neural Network and Multiple Linear Regression. *Energy Sources Part B*, 10: 427–434.

Azadeh, A.; Ghaderi, S.F.; Tarverdian, S.; Saberi, M. (2007). Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. *Appl. Math. Comput.*, 186: 1731–1741.

Günay, M.E. (2016). Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. *Energy Policy*, 90: 92–101.

Dai, S.; Niu, D.; Li, Y. (2018). Forecasting of Energy Consumption in China Based on Ensemble Empirical Mode Decomposition and Least Squares Support Vector Machine Optimized by Improved Shuffled Frog Leaping Algorithm. *Appl. Sci.*, 8: 678.

Wang, Q.; Li, R. (2017). Decline in China's coal consumption: An evidence of peak coal or a temporary blip. *Energy Policy*, 108: 696–701.

Suganthi, L.; Samuel, A.A. (2016). Modelling and forecasting energy consumption in INDIA: Influence of socioeconomic variables. *Energy Sources Part B Econ. Plan. Policy*, 11: 404–411.

Xu, J.; Fleiter, T.; Eichhammer, W.; Fan, Y. (2012). Energy consumption and CO2 emissions in China's cement industry: A perspective from LMDI decomposition analysis. *Energy Policy*, 50: 821–832.

Kishita, Y.; Yamaguchi, Y.; Umeda, Y. (2016). Describing Long-Term Electricity Demand Scenarios in the Telecommunications Industry: A Case Study of Japan. *Sustainability*, 8: 52.

Zhao, W.; Wang, J.; Lu, H. (2014). Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model. *Omega*, 45: 80–91.

Hamzacebi, C.; Es, H.A. (2014). Forecasting the annual electricity consumption of Turkey using an optimized grey model. *Energy*, 70: 165–171.

Mu, H.; Dong, X.; Wang, W.; Ning, Y.; Zhou, W. (2002). Improved Gray Forecast Models for China's Energy Consumption and CO, Emission. *J. Desert Res*., 22: 142–149.

Zeng, B.; Zhou, M.; Zhang, J. (2017). Forecasting the Energy Consumption of China's Manufacturing Using a Homologous Grey Prediction Model. *Sustainability*, 9: 1975.

Jiang, F.; Yang, X.; Li, S. (2018). Comparison of Forecasting India's Energy Demand Using an MGM, ARIMA Model, MGM-ARIMA Model, and BP Neural Network Model. *Sustainability*, 10, 7: 2225.

Ediger, V.S.; Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, 35: 1701–1708.

Ekonomou, L. (201). Greek long-term energy consumption prediction using artificial neural networks. *Energy*, 35: 512–517.

Ardakani, F.J.; Ardehali, M.M. (2014). Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types. *Energy*, 65: 452–461.

Supasa, T.; Hsiau, S. S.; Lin, S, M.; Wongsapai, W.; Wu, J. C. (2017). Household Energy Consumption Behaviour for Different Demographic Regions in Thailand from 2000 to 2010. *Sustainability*, 9: 2328.

Zhao, J.; Thinh, N. X.; Li, C. (2017). Investigation of the Impacts of Urban Land Use Patterns on Energy Consumption in China: A Case Study of 20 Provincial Capital Cities. *Sustainability*, 9: 1383.

Tian, Y.; Xiong, S.; Ma, X. (2017). Analysis of the Potential Impacts on China's Industrial Structure in Energy Consumption. *Sustainability*, 9: 2284.

Ayvaz, B.; Kusakci, A.O. (2017). Electricity consumption forecasting for Turkey with nonhomogeneous discrete grey model. Energy Sources Part B Econ. Plan. *Policy*, 12: 260–267.

Dickey, D.A.; Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*: 1057–1072.

Enders, W. (2010). *Applied Econometrics Time Series.* Wiley Series in Probability and Statistics. University of Alabama: Tuscaloosa, AL, USA.

MacKinnon, J. (1991). *Critical Values for Cointegration Test. In Long-Run Economic Relationships; Engle, R., Granger, C., Eds.* Oxford University Press: Oxford, UK.

Cryer, J. D.; Chan, K. (2008). *Time Series Analysis with Applications in R*. 2nd edition. Springer-Verlag: New York.

Johansen, S.; Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxford Bull. Econ. Statist.*, 52: 169–210.