Article

Energy Management through Cost Forecasting for Residential Buildings in New Zealand

Linlin Zhao [1], Zhansheng Liu [1,*], and Jasper Mbachu [2]

[1] College of Architecture and Civil Engineering, Beijing University of Technology, Beijing 100124, China; llzhao@bjut.edu.cn
[2] Faculty of Society & Design, Bond University, Gold Coast 4226, Australia; jmbachu@bond.edu.au
* Correspondence: liuzhansheng@bjut.edu.cn

Abstract: Over the last two decades, residential buildings have accounted for nearly 50 percent of total energy use in New Zealand. In order to reduce household energy use, the factors that influence energy use should be continuously monitored and managed. Building researchers and professionals have made efforts to investigate the factors that affect energy use. However, few have concentrated on the association between household energy use and the cost of residential buildings. This study examined the correlation between household energy use and residential building cost. Analysis of the correlation between energy use data and residential building cost indicated that residential building cost in the construction phase and energy use in the operation stage were significantly correlated. These findings suggest that correct monitoring of building costs can help to identify trends in energy use. Therefore, this study proposes a time series model for forecasting residential building costs of five categories of residential building (one-story house, two-story house, townhouse, apartment, retirement village) in New Zealand. The primary contribution of this paper is the identification of the close correlation between household energy use and residential building costs and provide a new area for optimize energy management.

Keywords: residential energy use; energy management; residential building costs; exponential smoothing method; ARIMA model

1. Introduction

Energy consumption in the residential sector is increasing due to rapid urbanization and social development. The residential sector consumed 45% of the nation's energy. The expansion of residential areas contributes significantly to the increase in energy consumption and CO_2 emission, resulting in global warming and air pollution. The existing literature indicates that the efficiency of end-use energy can significantly reduce total global energy consumption [1, 2]. Proper management of household energy use plays a key role in the management of energy demand. Knowledge of the factors that affect energy consumption is essential for proper energy management. To improve energy efficiency in the residential sector, decision-makers must continuously monitor and manage the factors that influence energy use in the residential sector. Several previous studies focused on factors that influence energy efficiency, including climate [3], time periods [4], environmental and economic impacts [5], building features [6], procurement process [7], building envelopes [8], heating, ventilation, and air conditioning (HVAC) [9], indoor temperatures [10], smart lighting and appliances [11] and occupant behavior [12]. However, only a limited number of studies examined the relationship between energy use and building costs. This study investigates the association between residential building cost and energy use in residential buildings.

Electricity, gas, and petrol are types of energy that are integral parts of human life. The residential sector is the greatest energy consumer. A significant amount of energy is used to provide comfortable indoor environments. Since the demands of all users should be satisfied, understanding to future trends of energy use has become important. It is important for energy demand side management to know the future trends of electricity and gas consumption in the residential sector. Prediction of the factors that influence household energy use is essential for continuous monitoring and management of energy consumption. Residential building costs are estimated in this study to properly manage household energy use since the study hypothesizes that residential building cost is significantly correlated with household energy use.

Given the utmost importance of cost forecasting and the important role of predicting it accurately, this study tends to use time series modelling techniques to forecast the building cost of five categories of residential building including one-story house (AR1), two-story house(AR2), townhouse (AR3), apartment (AR4), retirement village (AR5) in New Zealand. Building costs time series usually exhibit strong trends presenting challenges in developing useful models. How to effectively model building cost series and how to enhance forecasting performance are still outstanding questions. There are two most widely used time series forecasting methods: exponential smoothing and Autoregressive Integrated Moving Average (ARIMA). The performance of the two forecasting techniques was evaluated in terms of error measures.

Exponential smoothing method was originally introduced by [13, 14]; for short-term sales forecasting in support of supply chain management and production planning. The widespread usage of this method is mainly due to the fact that it is a relatively simple forecasting method requiring a small size sample and having a comprehensible statistical framework and model parameters. Exponential smoothing models developed are based on the trend and seasonality in time series, while ARIMA models are supposed to describe the autocorrelations in the time series. A framework for exponential smoothing methods was developed based on state-space models [15].

Autoregressive Integrated Moving Average (ARIMA) approach is a renowned and widely used linear method [16]. It carries more flexibility by representing various components of time series including autoregressive (AR), moving average (MA), and combined AR and MA. It is the most efficient approach for short-term forecasting with rapid changes. ARIMA models can also predict the future based on modelling the behavior of the serial correlation between the observations of the time series. The future predictions based on ARIMA models can be explained by previous or lagged values and the terms of the stochastic errors [17].

Time series forecasting techniques such as exponential smoothing models and the ARIMA models have not yet been examined for residential building cost forecasting in New Zealand. Therefore, the study as original contribution to the existing literature is for the first time to evaluate the forecasting performance of these models for residential building cost in New Zealand. Moreover, based on the comparison of the forecasting techniques, industry practitioners can derive a general understanding of forecasting techniques for building cost, and thereby improve forecasting accuracy.

The rest of this study is organized as follows. Section 2 presents the previous studies about factors influencing residential energy use and cost forecasting methods. Section 3 introduces correlation analysis, exponential smoothing method, and ARIMA models. Correlation analysis results, exponential smoothening and ARIMA models for the cost series are shown in Section 4. In Section 5, the comparison of cost models based on error measures is described. Results discussion is presented in Section 6. In the final section, conclusion is presented.

2. Literature Review

2.1. Influencing Factors of Residential Energy Use

Increased use of energy has already raised concern about exceeding supply capacities and severe environmental impacts, including global warming and climate change [18]. The residential sector significantly contributes to energy consumption, such as electricity, gas, and petrol. Many factors

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

influence household energy use, including weather, building design, energy systems, and socio-economic factors. Several studies have been conducted that focused on the investigation of the factors that influence energy efficiency.

A previous study explored the relationship between energy use and indoor thermal comfort [19], indicated that weather and climate have significant effects on household energy use. For example, in cold regions, people need more electricity or gas to warm up their houses. Recently, due to the increased demand for sustainable development more studies focused on how to improve energy efficiency. For example, some study stated that building features such as orientation can impact energy use by the proper use of natural resources, including sun light and natural ventilation [20]. Energy efficiency can be improved by using public-private partnership (PPP) practices [21, 22]. Certain financial and technological hints were used to manage energy that improved energy efficiency. Another study stated that certain innovative technologies can be incorporated into the project delivery process for reducing energy consumption [23, 24]. In addition, other studies explored the use of an evolved HVAC system and an advanced building envelope can improve energy efficiency in the operation stage [25-27]. For example, the smart lighting system can improve energy performance in a building [28]. Moreover, effective implementation of home appliances improve energy efficiency [29].

2.2. Cost Forecasting Methods

The main issue is forecasting building cost simply and accurately. For example, [30] developed a simple model to predict total project cost taking into account the effect of economic inflation. [31] illustrated a suitable method for examining cost overruns by involving political influences, and delays and economic inflation, during the project process. [32] introduced a mathematical model for investigating the accuracy of early cost estimates by using principal component analysis and regression analysis. [33] presented a probabilistic model based on the Poisson process for estimating project cost contingencies. [34] proposed an integrated regression model by including the strengths of probabilistic and parametric techniques to estimate conceptual cost. [35] categorized the leading cost drivers to evaluate future total project cost.

Although these methods are effective in identifying the leading cost drivers and appropriate estimation at the inception of the project, they are difficult to deal with as time-varying variables and reflect the time lag effects. Since much time-related data are dependent or have an autocorrelation [36], time-related techniques can be adopted to overcome these limitations.

2.3. Time Series Models for Cost Forecasting

In an attempt to solve time-related problems in the methods, time series techniques, which estimate future values of a certain variable according to past values of itself and random shock factors, have been adapted to cost forecasting in construction projects. For example, [37] used time-series models to provide reliable forecasts of building costs, tender prices, and the impacts of economic inflation on building projects. Moreover, [38] explained a way of applying neutral networks to forecast changes in the construction cost index. [39] suggested a time-series approach to identify the leading factors causing escalated construction cost. [40] introduced an integrated regression analysis and ARIMA techniques to predict a tender price index for Hong Kong building projects. [41] developed a Box-Jenkins model to estimate the labor market of the Hong Kong construction industry. [42] addressed a dynamic regression model to examine the relationship between the economic conditions in the market and construction cost. [43] illustrated a time series method that estimates future values according to past values and corresponding random errors and produces a reliable prediction of construction cost. These time series techniques provide systematic and time-related models to forecast future values. According to [41], it is possible to make accurate predications based on historical patterns.

3. Research Methods

3.1. Data

The building cost index is useful for construction professionals to quantify cost variations [44]. The index can provide information of cost changes caused by a combination of changes in material, labor, and equipment. Hence, the cost index has been used widely in the industry for cost estimation [43]. The cost index provided by QV cost-builder have been accepted in the Architecture, Engineering and Construction (AEC) industry in New Zealand. QV cost-builder carried out various surveys on construction economics including construction material, labor, and equipment costs to provide comprehensive statistical information. As for many other industries and sectors, QV cost-builder compiles historical data to guide construction organizations and industry professionals and to identify cost fluctuations in the construction industry. The cost forecasting application undertaken in this study is based on the average quarterly building cost data for the five categories of residential building in New Zealand. The variables of residential energy use include household electricity use, household gas use, and household petrol use. The energy variables are obtained from Statistics New Zealand which compiles historical data of household energy use.

The available dataset consists of quarterly data over a period of 18 years (72 observations) for three types of household energy use and five categories of residential building costs in New Zealand. The cost data are separated into two sections: in-sample data and out-of-sample data. The in-sample data are used for model fitting and the out-of-sample data aims to evaluate the forecasting performance of the model [45]. The data (72 observations) was split into two parts: the training part for model fitting and the testing part for evaluating forecasting performance by comparing forecasts with observations [46]. There is no clear rule for this dividing; in this study, about 72% of the data (2001:Q1-2013:Q4) were used for model fitting and the remaining 28% (2014:Q1-2018:Q4) were used for out-of-sample forecasts evaluation. The quarterly average building cost for the five categories of residential building in New Zealand from 2001:Q1-2018:Q4 are depicted in Figure 1.

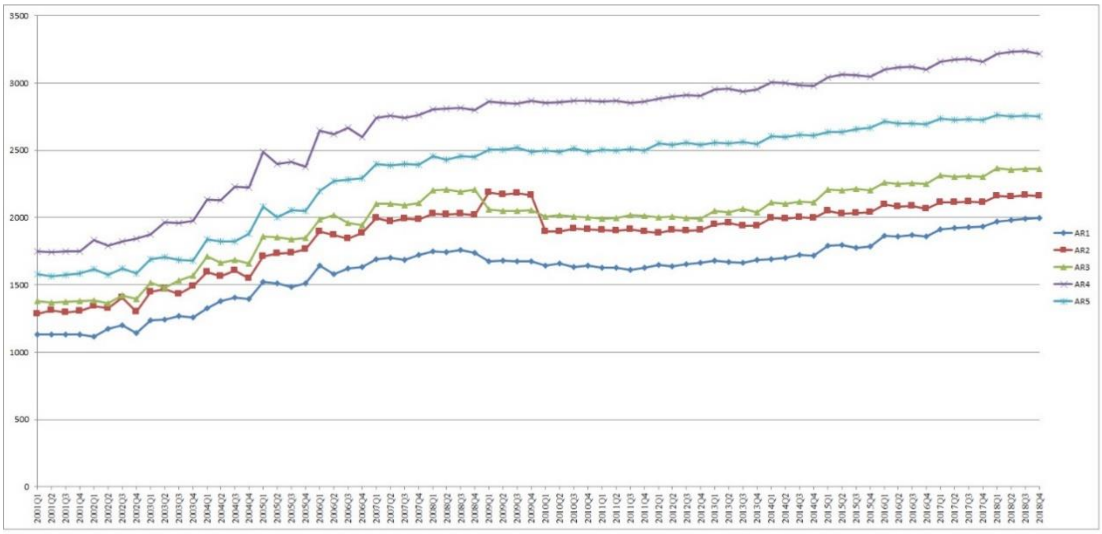

Figure 1. Building cost time series for five categories of residential building in New Zealand

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

3.2. Correlation Analysis

Correlation analysis is a statistical method used to evaluate the significance of correlation between two variables [47]. A significant correlation indicates that two variables have a significant relationship, while a weak correlation indicates that the variables are weakly related. In other words, correlation analysis can be used to examine the significance of the relationship between two variables. Pearson's correlation coefficient, also called linear correlation coefficient, assesses the linear relationship between two variables [48]. Let i and j be two random variables of the same sample n. To calculate Pearson's correlation coefficient r, use Equation (1) as follows:

$$r = \frac{\sum_{k=1}^{n}(i_k-\hat{\imath})(j_k-\hat{\jmath})}{\sqrt{\sum_{k=1}^{n}(i_k-\hat{\imath})^2}\sqrt{\sum_{k=1}^{n}(j_k-\hat{\jmath})^2}}, \tag{1}$$

Where

$$\hat{\imath} = \frac{1}{n}\sum_{k=1}^{n} i_k \; and \; \hat{\jmath} = \frac{1}{n}\sum_{k=1}^{n} j_k,$$

are the means of the variables i and j, respectively. The correlation coefficient r ranges between -1 and +1. If the linear correlation between i and j is positive (the increase of one variable is related to the increase of the other), then the correlation coefficient r >0, whereas if the linear correlation between i and j is negative (the increase of one variable is related to the decrease of the other), then the correlation result r <0. The value r = 0 indicates absence of any association between i and j. The sign of the correlation coefficient indicates the direction of the relationship. The magnitude of r indicates the significance of the correlation. For example, if r is close to 1, then the two variables are positively associated at a significant level, which also indicates that the increase of one variable is related to the increase of the other. If r is close to -1, then the two variables are negatively associated with each other at a significant level, which indicates that the increase of one variable is related to the decrease of the other. If r = 0, this usually indicates the two variables are unrelated.

3.3. Exponential Smoothing

Exponential smoothing is one of the most effective forecasting methods when a time series has a trend that has changed over time, for example, since the 1950s [49]. It unequally weights the observed time series values. More recently observed values are weighted more heavily than more remote observations. The weights for the observed time series values decrease exponentially as one moves further into the remote. A smoothing constant can determine the rate at which the weights of older observed values decrease. Exponential smoothing techniques include simple exponential smoothing, linear trend corrected exponential smoothing, Holt-Winters methods, and damped trend exponential smoothing [49].

According to [50], exponential smoothing models have been widely used in many research fields and industry practices due to their relative simplicity and good overall forecasting performance as well as considering trends, seasonality and other features of the data. A large number of existing research and studies also indicated their extensive industrial applications [51, 52]. In this study, Holt-Winters exponential smoothing method was adopted.

3.3.1. Holt-Winters Method

Holt-Winters method can be applied to time series data displaying trend and seasonality; it has level and trend smoothing parameters (α and β) in addition to a seasonal parameter (γ). Although there is no strong evidence for seasonality in the time series of the residential building costs in New

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

Zealand, Holt-Winters method is used to evaluate whether the involvement of a seasonal parameter can improve the model.

Holt-Winters methods are designed for time series that exhibit linear trend and seasonal variation, which include additive Holt-Winters method and multiplicative Holt-Winters methods [49]. An advantage of these methods is that they can model data seasonality directly instead of stationary transforming for the data. If a time series has a linear trend and additive seasonal pattern, the additive Holt-Winters method is appropriate. Then the time series can be described in Equation (2).

$$Y_t = (\beta_0 + \beta_1 t) + S_t + \epsilon_t, \tag{2}$$

where β_1 is growth rate; S_t is a seasonal pattern; ϵ_t is error term.

For such time series, the mean, the growth rate, and the seasonal variation may be changing over time. A state space model for these changing components can be found in Equation (3-6).

$$l_t = l_{t-1} + b_{t-1} + \alpha[Y_t - (l_{t-1} + b_{t-1} + S_{t-L})], \tag{3}$$

$$b_t = b_{t-1} + b_{t-1} + \alpha\gamma[Y_t - (l_{t-1} + b_{t-1} + S_{t-L})], \tag{4}$$

$$S_t = S_{t-L} + (1-\alpha)\delta[Y_t - (l_{t-1} + b_{t-1} + S_{t-L})], \tag{5}$$

$$\hat{Y}_t = l_{t-1} + b_{t-1} + S_{t-L}, \tag{6}$$

To begin the estimation the initial values for level, growth rate and seasonal variation should be estimated. Hence, first, a least squares regression model should be generated based on available data. The regression model can be expressed in Equation (7). The initial values l_0, $b_0$were also obtained from the model.

$$\hat{Y}_t = l_0 + b_0 t, \tag{7}$$

Obtain estimated values for each time period based on the above regression model. The initial seasonal factor in each of L seasons can be calculated in Equation (8-11).

$$S_{L1} = \frac{(y_1 - \hat{y}_1) + (y_{1+L} - \hat{y}_{1+L}) + (y_{1+2L} - \hat{y}_{1+2L}) + \cdots + (y_{1+nL} - \hat{y}_{1+nL})}{L}, \tag{8}$$

$$S_{L2} = \frac{(y_2 - \hat{y}_2) + (y_{2+L} - \hat{y}_{2+L}) + (y_{2+2L} - \hat{y}_{2+2L}) + \cdots + (y_{2+nL} - \hat{y}_{2+nL})}{L}, \tag{9}$$

$$S_{L3} = \frac{(y_3 - \hat{y}_3) + (y_{3+L} - \hat{y}_{3+L}) + (y_{3+2L} - \hat{y}_{3+2L}) + \cdots + (y_{3+nL} - \hat{y}_{3+nL})}{L}, \tag{10}$$

$$S_{LL} = \frac{(y_L - \hat{y}_L) + (y_{L+L} - \hat{y}_{L+L}) + (y_{L+2L} - \hat{y}_{L+2L}) + \cdots + (y_{L+nL} - \hat{y}_{L+nL})}{L}, \tag{11}$$

Where $S_{L1}, S_{L2}, S_{L3}, \cdots, S_{LL}$ are seasonal factors; L is the number of seasons in a year.

After finding the values for the seasonal factors, the state space models are employed to obtain model parameters that minimize the sum of the squared errors. Future values of the time series are predicted by the state space model in Equation (6).

3.3.2. Multiplicative Holt-Winters Method

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

If a time series has a linear trend with multiplicative seasonal variations, the multiplicative Holt-Winters is appropriate to be used. The state space models for this method can be described in Equations (12-15).

$$l_t = l_{t-1} + b_{t-1} + \alpha \frac{[Y_t - (l_{t-1} + b_{t-1})S_{t-L}]}{S_{t-L}}, \quad (12)$$

$$b_t = b_{t-1} + \alpha\gamma \frac{[Y_t - (l_{t-1} + b_{t-1})S_{t-L}]}{S_{t-L}}, \quad (13)$$

$$S_t = S_{t-L} + (1-\alpha)\delta \frac{[Y_t - (l_{t-1} + b_{t-1})S_{t-L}]}{l_t}, \quad (14)$$

$$Y_t = (l_{t-1} + b_{t-1})S_{t-L}, \quad (15)$$

And the seasonal factors can be computed in the following Equations (16-19).

$$S_{L1} = \frac{(y_1/\hat{y}_1) + (y_{1+L}/\hat{y}_{1+L}) + (y_{1+2L}/\hat{y}_{1+2L}) + \cdots + (y_{1+nL}/\hat{y}_{1+nL})}{L}, \quad (16)$$

$$S_{L2} = \frac{(y_2/\hat{y}_2) + (y_{2+L}/\hat{y}_{2+L}) + (y_{2+2L}/\hat{y}_{2+2L}) + \cdots + (y_{2+nL}/\hat{y}_{2+nL})}{L}, \quad (17)$$

$$S_{L3} = \frac{(y_3/\hat{y}_3) + (y_{3+L}/\hat{y}_{3+L}) + (y_{3+2L}/\hat{y}_{3+2L}) + \cdots + (y_{3+nL}/\hat{y}_{3+nL})}{L}, \quad (18)$$

$$S_{LL} = \frac{(y_1/\hat{y}_1) + (y_{1+L}/\hat{y}_{1+L}) + (y_{1+2L}/\hat{y}_{1+2L}) + \cdots + (y_{1+nL}/\hat{y}_{1+nL})}{L}, \quad (19)$$

3.4. Autoregressive Integrated Moving Average (ARIMA)

There are four steps to select an appropriate model for the time series data in the Box-Jenkins approach [53]. The development process of an ARIMA model is shown in Figure 2. ARIMA models are flexible and adaptive since they can forecast data values of a time series by a linear combination of its past values, past errors (in terms of univariate analysis) and past and present values of other time series (in terms of multivariate analysis). Besides, in univariate time series analysis, the development processes of ARIMA models provide a comprehensive set of tools for model identification, parameters estimation, diagnosis checking, and forecasting. Taking into account the seasonality of the time series, a seasonal ARIMA model denoted as ARIMA $(p,d,q)(P,D,Q)_L$ is introduced, where P represents seasonal autoregressive orders, D indicates seasonal differencing orders, Q represents seasonal moving average orders, and L indicates the number of seasons.

Seasonality implies that a pattern repeats itself over a fixed time interval [54]. In this study, the quarterly data present a seasonal period of four quarters. The auto-correlation function (ACF) and partial auto-correlation function (PACF) were employed to determine the stationarity in the dataset. The seasonal difference was used to transform the non-stationary seasonal data into stationary by taking the difference between the current observation and the corresponding observation from the previous year. A seasonal ARIMA model can be shown in Equation (20).

$$\emptyset_p(B)\varphi_P(B^L)\nabla_L^D\nabla^d y_t = \delta + \theta_q(B)\vartheta_Q(B^L)a_t, \quad (20)$$

Where

$$\emptyset_p(B) = \left(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p\right),$$

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

$$\varphi_P(B^L) = \left(1 - \varphi_{1,L}B^L - \varphi_{2,L}B^{2L} - \cdots - \varphi_{P,L}B^{PL}\right),$$

$$\nabla_L^D \nabla^d y_t = (1 - B^L)^D (1 - B)^d y_t z,$$

$$\theta_q(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right),$$

$$\vartheta_Q(B^L) = \left(1 - \vartheta_{1,L}B^L - \vartheta_{2,L}B^{2L} - \cdots - \vartheta_{Q,L}B^{PL}\right),$$

where B is backshift operator; L is the number of seasons in a year (**L**=4 for quarterly data and **L**=12 for monthly data); δ is a constant term; $a_t, a_{t-1}, \cdots$ are random shocks; $\emptyset_1, \emptyset_2, \cdots, \emptyset_p$ are non-seasonal autoregressive parameters; $\varphi_{1,L}, \varphi_{2,L}, \cdots, \varphi_{P,L}$ are seasonal autoregressive parameters; $\theta_1, \theta_2, \cdots, \theta_q$ are non-seasonal moving average parameters, $\vartheta_{1,L}, \vartheta_{2,L}, \cdots, \vartheta_{Q,L}$ are seasonal moving average parameters.

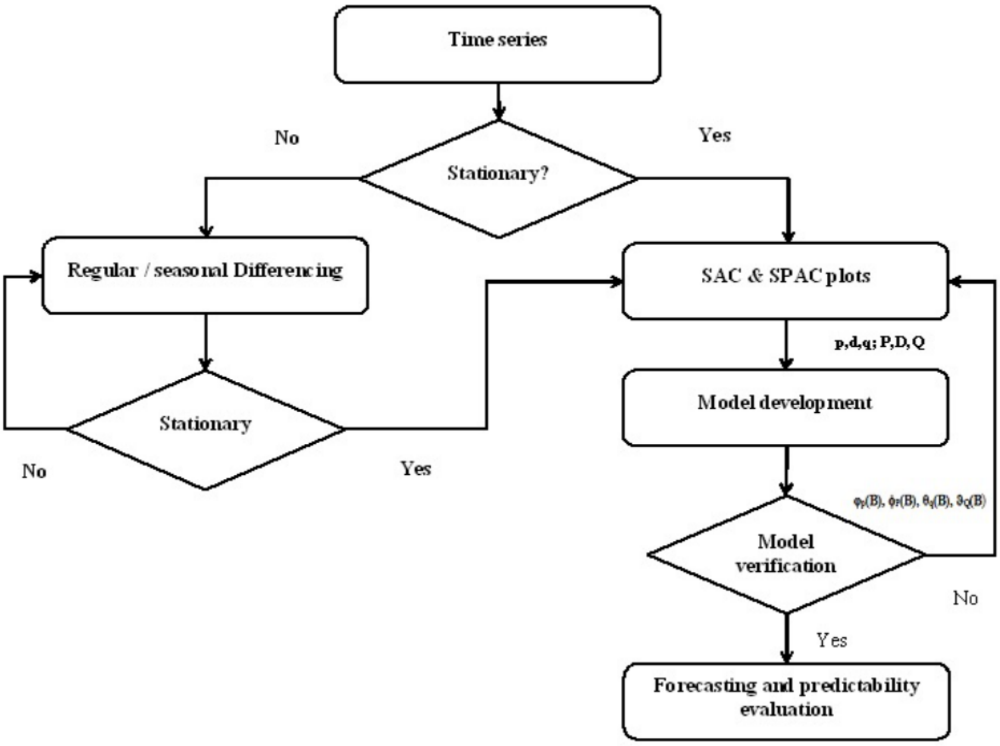

Figure 2. ARIMA model development process

3.4.1. Stationary Checking

Classical ARIMA models are usually used to describe stationary time series. Thus, in order to identify an appropriate ARIMA model, the stationary of the times series should be determined at first. If the time series is not stationary, the transformation of the time series to stationary should be undertaken. A stationary time series can be described as the statistical properties (e.g. the mean and the variance) of it are essentially constant over time [49]. The first differences of the non-stationary time series are usually employed to transform the non-stationary time series values into stationary time series values. The first differences of the time series values are shown in Equation (21).

$$z_t = y_t - y_{t-1}, \tag{21}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

where $\boldsymbol{z_t}$ indicates the first differences time series values; $\boldsymbol{y_t}$ indicates non-stationary times series values.

If a seasonal time series was analyzed, then a seasonal transformation is used to produce a stationary time series. The first seasonal differencing can be described in Equation (22).

$$z_t = y_t - y_{t-L}, \tag{22}$$

where L indicate the numbers of seasons in a year ($\boldsymbol{L}$=4 for quarterly data and $\boldsymbol{L}$=12 for monthly data).

Moreover, the first regular differencing and seasonal differencing was usually adopted to transform a seasonal non-stationary series to a stationary time series. The transformation is expressed in Equation (23).

$$z_t = (y_t - y_{t-1}) - (y_{t-L} - y_{t-L-1}) = y_t - y_{t-1} - y_{t-L} + y_{t-L-1}, \tag{23}$$

Further, the sample auto-correlation function (SAC) also can be used to determine whether the time series is stationary. For example, if the SAC of a time series values either dies down quickly or cuts off quickly at both seasonal lags and non-seasonal lags, then the time series can be considered as stationary. If the SAC of the time series values dies down extremely slowly either at seasonal lags or non-seasonal lags, it is reasonable to decide the time series is non-stationary.

3.4.2. Tentative Identification

The identification of the ARIMA models is usually dependent on the sample auto-correlation function (SAC) and the sample partial auto-correlation function (SPAC) for the values of a stationary time series. The behaviour of the SAC and SPAC at the seasonal level to be that at lags L, 2L, 3L, and 4L. The sample autocorrelation at lag k can be calculated through Equation (24).

$$r_k = \frac{\sum_{t=1}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2}, \tag{24}$$

Where z_t indicates the stationary time series values; $\bar{z}$ is the mean of the time series.

The formula for the sample partial autocorrelation at lag k is presented in Equation (25).

$$r_{kk} = \begin{bmatrix} r_1 & if\ k = 1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & if\ k = 2, 3, \ldots \end{bmatrix}, \tag{25}$$

Where

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j},$$

3.4.3. Parameter Estimation

After identifying a tentative model for a time series, the parameters of the model should be estimated. The parameters are usually estimated by the least square method. The least square estimation method means that the model parameters minimize the sum of the squared errors. The sum of the squared errors can be computed by using Equation (26).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at *Energies* **2019**, *12*, 2888; doi:10.3390/en12152888

$$SSE = \sum_{t=1}^{n} (y_t - \hat{y}_t)^2, \tag{26}$$

Where y_t is the real value of the time series; $\hat{y}_t$ is the value estimated by the tentative model.

3.4.4. Diagnostic Checking

The obtained models should be checked for whether the ARIMA assumptions are satisfied. As a more accurate test, the Ljung-Box test is usually undertaken to examine whether the autocorrelation of the residuals are statistically different from an expected white noise process. If the *p*-value is greater than 0.05, indicating no significant autocorrelation in residuals, in turn, the model is adequate [55].

3.4.5. Forecasting Error Measure

Although a model may well fit the historical data, it is not valid to determine that the model has good forecasting performance. The forecasting performance of a model can only be determined by the accuracy of the out-of-sample forecasts [56]. The accuracy of the forecasts was evaluated by mean absolute percentage error (MAPE) between the actual and predicted values of building cost. The lower the values are, the better the forecasting performance of the proposed model.

The most widely used criterion for forecasting models is accuracy, which has many forms, including root mean square error (RMSE) [57], mean absolute error (MAE) [58], and mean absolute percentage error (MAPE) [59]. This study evaluated the accuracy of the forecasts by mean absolute percentage error (MAPE) between the actual and predicted values of building cost. The lower the values are, the better the forecasting performance of the proposed model. Denote the real observations for the time series by (y_i) and the forecasting values for the same series by $(\hat{y}_i)$. Mean absolute percentage error (MAPE) can be computed in Equation (27).

$$MAPE = \frac{\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \times 100\%, \tag{27}$$

3.5. *t-Test*

The *t*-test is a statistical method that is also referred to as the Student's *t*-test. The *t*-test includes one-sample *t*-test, two independent samples *t*-test, and paired sample *t*-test (2). Unlike some statistical methods that heavily rely on sample size for their effectiveness, the *t*-test can be used with small sample sizes (such as n <30) [60]. The *t*-test can be used to test the difference between one sample and a set mean value (one sample *t*-test). The *t*-test can also be used to compare the difference between the mean of two samples (two sample *t*-test). Also, it can be used to test the mean difference between paired samples before and after the experiment (paired sample *t*-test). In this study, the paired sample *t*-test was used. To calculate *t*-value, Equation (28) was used, as follows.

$$t = \frac{(\sum_{i=1}^{n} D)/N}{\sqrt{\frac{\sum_{i=1}^{n} D^2 - (\sum_{i=1}^{n} D)^2/n}{(N-1)N}}}, \tag{28}$$

Where D is the difference between sample x and y and n is the sample size.

Then, the *t*-value can be calculated. Next, the *t*-critical value is calculated based on the sample size n. If the *t*-value is greater than the *t* critical value, there is no significant difference between the

 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at *Energies* **2019**, *12*, 2888; doi:10.3390/en12152888

two samples. If the *t*-value is less than the *t*-critical value, there is a significant difference between the two samples.

4. Data Analysis

4.1. Correlation Analysis Results

Correlation analysis was performed to examine whether there is a significant correlation between household energy use and residential building costs, which indicated the reliability of using residential building costs as an indicator of the future trend of household energy use. Based on the correlation coefficient, the results of correlation analysis show a significant correlation between the two variables. The significance level of the variables is validated by two-tailed significant correlation values at the 0.05 level.

Correlation analysis was conducted using the statistical software program SPSS (Statistical Packages for the Social Sciences, versions 23). Results of the analysis indicated that residential building costs positively correlate with household energy use. This significant correlation supports the research hypothesis. Therefore, there is a significant correlation between the energy use and residential building costs. The researcher predicts the correlation between residential building costs and household energy use. The results of correlation analysis are essential for understanding the effectiveness of residential building costs to serve as an indicator of future trends of household energy use.

The results of correlation analysis between household energy use variables and residential building cost are shown in Table 1, which shows that all household energy use variables are positively correlated with residential building costs. The highest correlation coefficient was obtained from the correlation between household gas use and residential building cost of two-story house (AR2), with a significant value of r = 0.994. A weakest correlation was observed between household petrol use and residential building cost of retirement village (AR5), with a coefficient of r = 0.640. Despite having a relative weak correlation, all the household energy use variables correlate with residential building cost at a significant level.

Table 1. Correlation analysis results.

	Residential Building Cost				
Energies	AR1	AR2	AR3	AR4	AR5
Electricity	0.974**	0.977**	0.984**	0.782**	0.833**
Gas	0.976**	0.994**	0.968**	0.966**	0.898**
Petrol	0.919**	0.916**	0.937**	0.884**	0.640**

** indicate significance at 0.05 level.

4.2. Exponential Models for Building Cost

Both additive Holt-Winters and multiplicative Holt-Winters models were applied to the five cost series. Following the methods outlined in [50], the model parameters were estimated. The results of the exponential smoothing models for the cost of the five categories of the residential building are displayed in Table 2. The *p*-value of the model parameters indicate that they are effective. Moreover, the model fit R-square and error measures including root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) were also generated. In addition, the model parsimony measure Bayesian Information Criterion (BIC) was also obtained. The results are shown in Table 4. They all indicate Holt-Winters models can fit the cost series fairly well because the models can identify the trend and seasonal variation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at *Energies* **2019**, *12*, 2888; doi:10.3390/en12152888

Table 2. Estimated parameter values with significant test for exponential smoothing models

Series	Exponential Smoothing model	Parameter	Estimate	SE	*p*-value
AR1	ES(AHW)	α	0.370	0.116	0.002
		β	0.634	0.287	0.032**
		γ	0	0.112	0.993
	ES(MHW)	α	0.379	0.112	0.001**
		β	0.537	0.251	0.037**
		γ	0.528	0.171	0.003**
AR2	ES(AHW)	α	0.899	0.150	***
		β	0	0.047	1
		γ	0	0.696	1
	ES(MHW)	α	0.846	0.145	***
		β	0.001	0.045	0.983
		γ	0.028	0.298	0.925
AR3	ES(AHW)	α	0.683	0.135	***
		β	0.218	0.113	0.059*
		γ	0.001	0.171	0.995
	ES(MHW)	α	0.578	0.128	***
		β	0.269	0.128	0.042**
		γ	0.020	0.091	0.830
AR4	ES(AHW)	α	0.200	0.079	0.014**
		β	1.000	0.467	0.037**
		γ	0	0.091	1
	ES(MHW)	α	0.198	0.083	0.021**
		β	1.000	0.503	0.052*
		γ	0.040	0.062	0.524
AR5	ES(AHW)	α	0.469	0.118	***
		β	0.441	0.194	0.027**
		γ	0.014	0.096	0.883
	ES(MHW)	α	0.361	0.096	***
		β	0.511	0.225	0.028**
		γ	0.479	0.151	0.003**

** indicate significance at 0.05 level, *** indicate significance at 0.01 level.

4.3. Seasonal ARIMA Models

4.3.1. Model Selection

[16] suggested that to properly implement the ARIMA method a time series with at least 30 observations is required. In this study, for each cost series, a total of 52 observations from 2001:Q1 to 2013:Q4 were used to obtain the proposed models. For the stationary analysis of the five cost series autocorrelation function (ACF) and partial autocorrelation function (PACF) were used; results are shown in Figure 3. Investigate the graphs of ACF and PACF for the five building cost series; it can be observed that the ACFs decay very slowly at both non-seasonal and seasonal lags. For each cost

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

series, the appropriate number of differencing should be determined. Hence, it is reasonable to transform to a stationary series by taking four quarter differencing of data to remove seasonality and regular differencing to remove trends for the four cost series, except the cost series for the two-story house in New Zealand. The cost series has only made a regular differencing to transform the data into stationary. After the differencing, the results of ACFs and PACFs for the five cost series are shown in Figure 4. The seasonal ARIMA models for the five cost series are shown in Table 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at *Energies* **2019**, *12*, 2888; doi:10.3390/en12152888

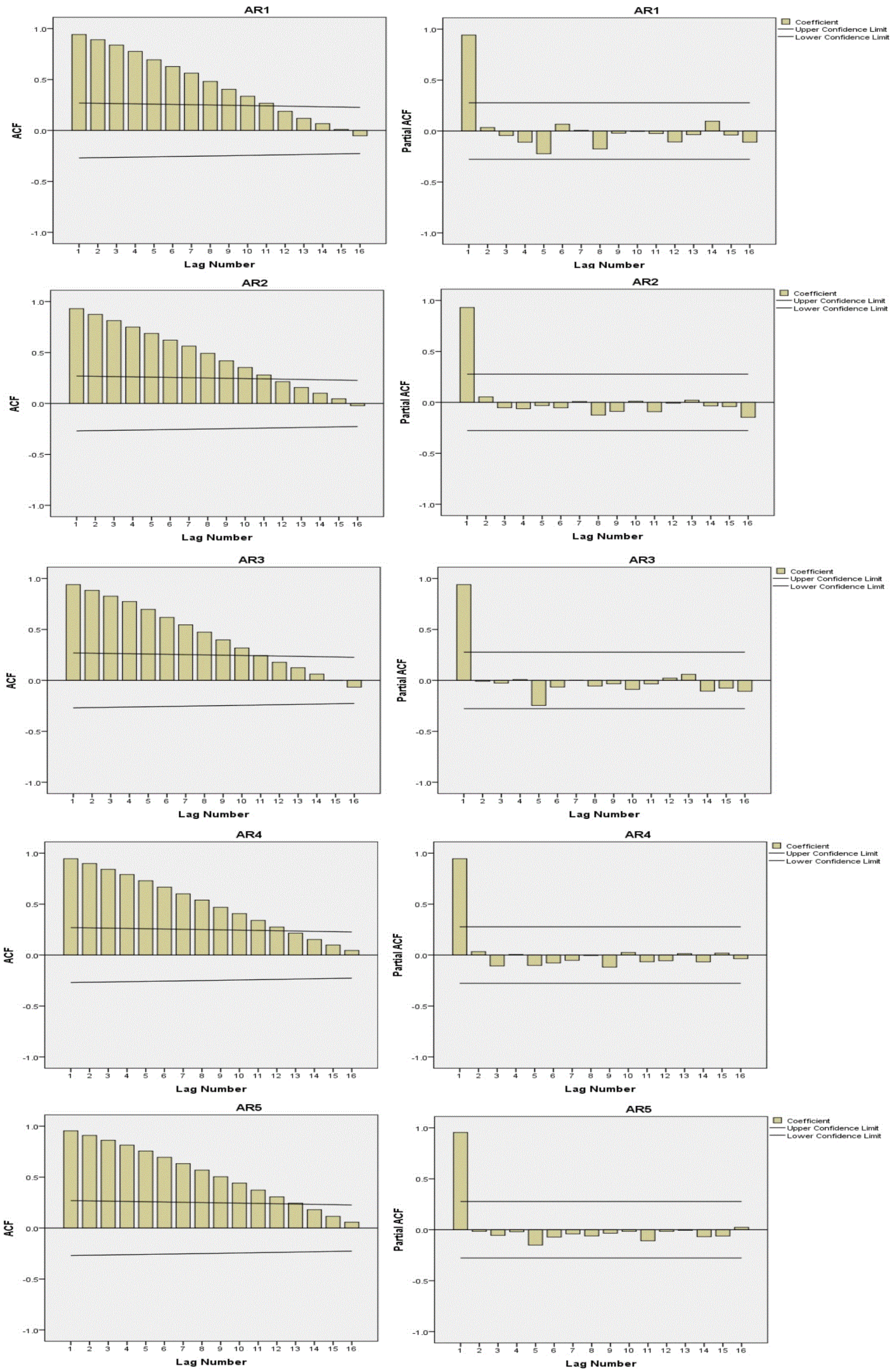

Figure 3. Sample ACF (left panels) and sample PA

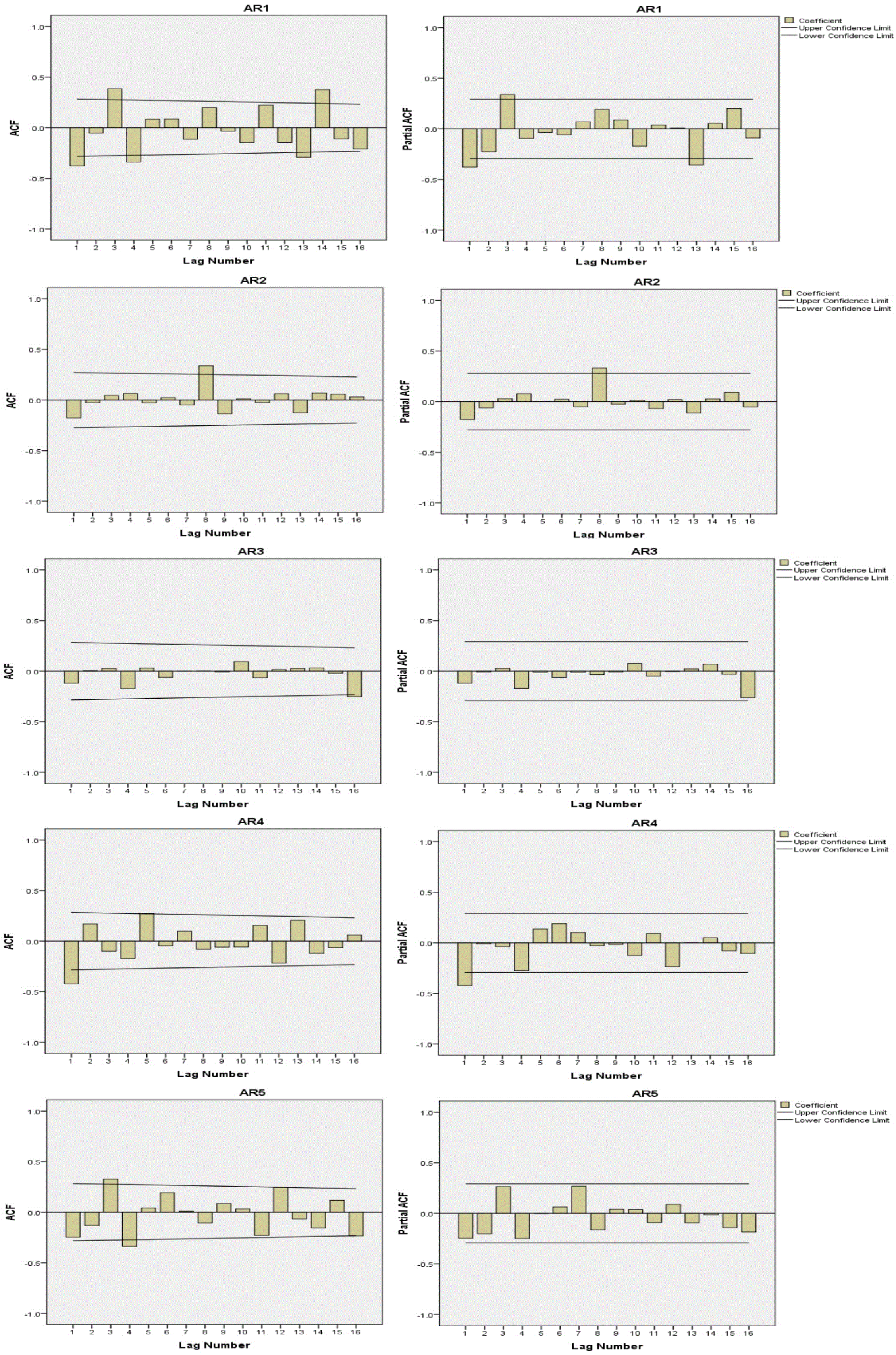

Figure 4. ACFs (left panels) and PACFs (right panels) of the differenced data series

Table 3. Estimated parameter values for seasonal ARIMA models

Series	Model	AR	MA	SAR	SMA
AR1	ARIMA(0,1,3)(0,1,1)l		MA(l)=0.341 MA(2)=-0.10l MA(3)=-0.295		SMA(l)=0.447
	ARIMA(0,1,1)(0,1,1)l		MA(l)=0.3l7		SMA(l)=0.290
AR2	ARIMA(0,1,0)(2,0,0)l			SAR(1)=0.038 SAR(2)=0.348	
	ARIMA(0,1,0)(0,0,2)l				SMA(l)=-0.005 SMA(2)=-0.368
AR3	ARIMA(0,1,0)(1,0,0)l			SAR(l)=0.562	
	ARIMA(0,1,0)(0,1,0)l				
AR4	ARIMA(1,1,0)(0,1,0)l	AR(l)=-0.4l9			
	ARIMA(0,1,1)(0,1,0)l		MA(l)=0.404		
AR5	ARIMA(0,1,0)(0,1,1)l				SMA(l)=0.554
	ARIMA(0,1,0)(0,1,0)l				

4.3.2. Proposed Seasonal ARIMA Models

Based on the approach provided by [53], the model parameters, model fit, and error measures were estimated for the five cost series. In order to select proper seasonal ARIMA models, different models with various combinations of regular orders (p and q) and seasonal orders (P and Q) were evaluated. The model parameters of the five cost series are presented in Table 3. Furthermore, the model fit and error measures of the ARIMA models are provided in Table 4. As seen from Table 4, ARIMA models fit the cost data fairly well.

4.3.3. Model Validation

The Ljung-Box Q test was employed to examine the autocorrelation of model residuals. If the p-value is greater than the value of 0.05, the null hypothesis that the data are not correlated should be accepted [16]. To examine the normality of the residuals the analysis applied the Shapiro-Wilk test. If the p-value of the test is greater than the value of 0.05, it indicates that there is no evidence to reject the null hypothesis that the data follow a normal distribution [61]. As seen from Table 4, the residuals of all the models pass the tests, indicating the proposed models are adequate. According to the estimation results, the model fit measures and error measures are acceptable. This suggests that the proposed models fit the data fairly well.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

Table 4. Model fit statistics and residual statistics

Series	Model	R-Square	RMSE	MAPE	MAE	BIC	Ljung-Box	Shapiro-Wilk
AR1	ARIMA(0,1,3)(0,1,1)$_4$	0.959	37.234	1.716	25.866	7.644	0.873	0.461
	ARIMA(0,1,1)(0,1,1)$_4$	0.953	38.665	1.856	27.895	7.556	0.519	0.158
	ES(AHW)	0.976	33.205	1.655	24.52	7.233	0.238	0.203
	ES(MHW)	0.974	34.311	1.581	23.722	7.299	0.805	0.102
AR2	ARIMA(0,1,0)(2,0,0)$_4$	0.942	63.899	2.221	38.057	8.546	0.877	0.184
	ARIMA(0,1,0)(0,0,2)$_4$	0.942	63.865	2.256	38.766	8.545	0.898	0.136
	ES(AHW)	0.947	62.506	2.066	35.292	8.498	0.891	0.153
	ES(MHW)	0.943	64.989	2.159	37.134	8.576	0.904	0.133
AR3	ARIMA(0,1,0)(0,1,0)$_4$	0.944	54.277	1.913	35.836	8.07	0.855	0.153
	ARIMA(0,1,0)(4,1,0)$_4$	0.95	53.423	1.823	33.945	8.366	0.956	0.122
	ES(AHW)	0.964	52.405	1.953	36.27	8.146	0.13	0.103
	ES(MHW)	0.959	55.473	2.054	38.568	8.26	0.106	0.173
AR4	ARIMA(1,1,0)(0,1,0)$_4$	0.981	51.468	1.53	37.383	8.046	0.657	0.391
	ARIMA(0,1,1)(0,1,0)$_4$	0.981	51.889	1.505	36.868	8.062	0.628	0.24
	ES(AHW)	0.989	45.577	1.32	31.825	7.867	0.593	0.127
	ES(MHW)	0.988	47.836	1.385	34.24	7.964	0.333	0.104
AR5	ARIMA(0,1,0)(0,1,1)$_4$	0.986	40.162	1.325	27.306	7.55	0.489	0.127
	ARIMA(0,1,0)(0,1,0)$_4$	0.983	43.299	1.52	31.335	7.618	0.141	0.103
	ES(AHW)	0.991	36.083	1.312	27.824	7.4	0.477	0.393
	ES(MHW)	0.989	39.374	1.356	28.715	7.574	0.415	0.647

5. Model Selection

5.1. Comparisons of the Models

Although a model can fit the data fairly well, it does not indicate that the model can produce better forecasts [62]. The forecasting accuracy of a model is affected by many factors, such as the number of observations in the time series, the number of forecast time origins examined, and the number of forecast lead times regarded [63]. In general, exponential smoothing models outperform ARIMA models in predicting in sample period. Despite superior performance over the in-sample period, the good performance of the exponential smoothing models does not translate into out-of-sample forecasts. However, seasonal ARIMA models are able to produce consistent forecasts based on error measures.

The forecasting performance of the univariate methods was evaluated by MAPE statistics. The results for residential building cost of one-story house (AR1) presented in Table 5 suggest that the exponential smoothing models generate better results in comparison to seasonal ARIMA models. In particular, the additive Holt-Winters model produces better forecasts based on MAPE measurement. When analyzing the results for cost of two-story house (AR2), the additive Holt-Winters model outperforms other models based on MAPE measurement. The results regarding cost of townhouse (AR3), the ARIMA model has the best forecasting performance among the proposed four models. For results for cost of apartment (AR4) suggest that ARIMA (0,1,1) (0,1,0)$_4$ is the best forecasting model. For results considering cost of retirement village (AR5), ARIMA (0,1,0) (0,1,1)$_4$ produced the best forecasts among the proposed models. Therefore, these results suggest that the ARIMA approach is better than the exponential smoothing method for building cost of the town house (AR3), apartment (AR4) and retirement village (AR5) in New Zealand. Although the seasonal ARIMA models did not outperform exponential smooth models in the model training process, their error measures are slightly larger than that of exponential models. The results show that seasonal ARIMA models perform better for predicting building cost for town house, apartments and retirement village in New Zealand, while exponential smooth models are superior in cost forecasting for both the one-story house and the two-story house in New Zealand. This outcome may be due to the relative stability of the cost series for the one-story house and the two-story house.

From the above results, it can be seen that both the exponential smoothing method and the ARIMA approach can produce good forecasts for residential building cost in New Zealand. Which method is better, depends on the characteristics of the data. For example, the ARIMA approach can produce better forecasts for building costs of town house, apartment and retirement village, which indicates that these costs have a random walk characteristic.

The MAPE of the proposed models for all the five cost series are presented in Table 5. Bold type is utilized in these tables to identify the lowest values of MAPE for each proposed model. As the results show, no single forecasting method is better for all data series. This confirms the generally acceptable idea that no individual forecasting approach can describe all the situations [64].

Table 5. Forecast values for building cost of one-story house in New Zealand

Series	Model	MAPE	t-value
AR1	ARIMA(0,1,3)(0,1,1)L	1.813	
	ARIMA(0,1,1)(0,1,1)L	1.651	
	ES(AHW)	**1.260**	**3.287**
	ES(MHW)	1.556	
AR2	ARIMA(0,1,0)(2,0,0)L	0.922	
	ARIMA(0,1,0)(0,0,2)L	0.794	
	ES(AHW)	**0.395**	**4.961**
	ES(MHW)	0.650	
AR3	ARIMA(0,1,0)(0,1,0)L	**1.020**	**3.589**
	ARIMA(0,1,0)(4,1,0)L	2.318	
	ES(AHW)	1.211	
	ES(MHW)	1.130	
AR4	ARIMA(1,1,0)(0,1,0)L	0.501	
	ARIMA(0,1,1)(0,1,0)L	**0.446**	**5.268**
	ES(AHW)	0.748	
	ES(MHW)	0.809	
AR5	ARIMA(0,1,0)(0,1,1)L	**0.853**	**4.891**
	ARIMA(0,1,0)(0,1,0)L	1.213	
	ES(AHW)	0.917	
	ES(MHW)	0.949	

5.2. t-Test results

In this study, the paired sample *t*-test was used to compare the mean difference between the estimated residential building cost with the real residential building cost. The *t*-test was performed using SPSS. The *t*-value was greater than the *t*-critical value (*t*-critical=1.666, n = 72), which means there is no significant difference between the real residential building costs and the estimated residential building costs. The *t*-value for the best-fit models are shown in Table 5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2019 doi:10.20944/preprints201906.0099.v1

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

6. Results Discussion

Correlation analysis was performed to test the correlation between residential energy use and residential building costs. These results validate that a correlation still exists between household energy use (electricity, gas, petrol) and estimated residential building cost at a significant level.

Exponential smoothing (ES) approach and ARIMA technique are both effective time series forecasting methods as they both can fairly well describe trend movement in the time series, but they have both strengths and weaknesses. For example, the ARIMA approach is more readily expanded to model interventions, outliers, variations and variance changes in time series; but it is a relatively sophisticated technique. Due to different data patterns and limited sample size, it is unjust to attempt to determine whether one time series forecasting method is better than the other. Therefore, either the exponential smoothing method or the ARIMA approach should be given a chance to demonstrate its maximum potential in any empirical case study.

While ES method is based on describing the trend and seasonality in the time series, ARIMA approach is focused on a description of the autocorrelation in the data. There is an idea that ARIMA approach is more advanced than ES method since the former has fewer parameters to be estimated. Although ARIMA models are more general, ES models can provide framework that is sufficient to capture the dynamics in the data series. ARIMA models are excellent for short-term forecasting. When they are used for long-term forecasting, the models need to be remodeled based on updated data incorporated into the model training process. ES method can be very competitive by automatically incorporating updated information into the model, and then producing better forecasts for long-term forecasting. An advantage of ARIMA technique is that only several parameters need to be estimated for generating good forecasting results. However, extreme values in the dataset are difficult to be estimated by ARIMA models due to the univariate nature of the model and the lack of a specific ability to simulate unexpected events.

In fact, for all the five cost series of residential building in New Zealand to which the statistical approaches were applied, the exponential smoothing models showed an excellent performance at the model training phase. However, seasonal ARIMA models produced more accurate forecasts for cost series of the town house, apartment and retirement village in the forecasting process. In the case of the cost for town house, apartment, and retirement village in New Zealand, the cost data must be examined in greater detail to identify previous short-term variations. Any distinct changes in the cost series may result in ARIMA models being wholly unsuitable.

7. Conclusions

Several methods have been used to predict future household energy use, but most of them are based on building factors such as thermal envelope or HVAC systems. However, in this study, residential building cost was used as an indicator of future trends of household energy use. Correlation analysis showed a significant positive correlation between household energy use and residential building cost. This result not only provides a clear guide for the future trend of household energy use, but also forms the relationship between the house building costs at the construction stage and the energy used at the operational stage. It is not necessary to evaluate every component of the building; future trends of energy use can be obtained from the residential building cost. In addition, results of correlation analysis will help developers or investors to make informed decisions by connecting current investment (construction cost) and future expenses (energy used). Accurate prediction of residential building cost provides a clear indication of the future trend of household energy use.

In this study, quarterly building costs data (over an 18-year range 2001:Q1-2018:Q4) for five categories of residential building (one-story house, two-story house, town house, apartment, and

Peer-reviewed version available at Energies 2019, 12, 2888; doi:10.3390/en12152888

retirement village) in New Zealand, were analyzed. It was found that time series data of residential building costs are non-stationary and autocorrelated and do not display a very strong seasonal pattern. Based on the identified characteristics, two time series forecasting techniques, exponential smoothing method and ARIMA approach, were adopted to take into account variations of residential building costs in predicting their future trends. It was concluded that both methods can produce proper forecasts. The analysis of model residuals explored that the underlying modelling assumptions hold true.

The primary contributions of this study to the existing body of knowledge are twofold: (1) explore the relationship between residential energy use and residential building costs; and (2) develop univariate forecasting models to predict the future trend of residential building cost with reasonable accuracy. The findings of this study can help industry professionals prepare energy use plan, help in decision making, and provide a new clue for energy management. Although this study used the QV's residential building cost index and energy variables of Statistics New Zealand, the proposed methods can be used for similar data sets in other cities as well as globally. Although these forecasting techniques are to predict future values of building cost, they can be further applied to other modelling purposes.

Abbreviations

ACF	Auto-Correlation Function
AR	Autoregressive
AR1	One-story House
AR2	Two-story House
AR3	Townhouse
AR4	Apartment
AR5	Retirement Village
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
ES	Exponential Smoothing Method
HVAC	Heating, Ventilation and Air Conditioning
HW	Holt-Winter Method
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MHW	Multiplicative Holt-Winter Method
PACF	Partial Auto-Correlation Function
PPP	Public-Private Partnership
RMSE	Root Mean Square Error
SAC	Sample Auto-Correlation Function
SAR	Seasonal Autoregressive
SE	Standard Error
SMA	Seasonal Moving Average
SPAC	Sample Partial Auto-Correlation Function

Author Contributions: Conceptualization, L.Z. and J.M.; methodology, L.Z.; software, Z.L.; validation, L.Z., Z.L.; writing—original draft preparation, L.Z.; writing—review and editing, J.M.; supervision, J.M.; project administration, J.M.; funding acquisition, L.Z

Funding: This research was funded by China Scholarship Council, grant number 201206130069, Massey University, grant number 09166424, and Beijing University of Technology, the grant number is 004000514119067. The APC was funded by the three grants.

Acknowledgments: The authors would like to thank the China Scholarship Council (CSC) for its support through the research project and also the Massey University. The authors would like to thank Reserve Bank of New Zealand and Ministry of Business, Innovation, and Employment for providing data to conduct this research. In addition, I would like to thank all practitioners who contributed to this project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reinisch, C.; Kofler, M.; Iglesias, F.; Kastner, W., Thinkhome energy efficiency in future smart homes. *EURASIP Journal on Embedded Systems* **2011**, 104617.
2. Lee, Y. M.; An, L.; Liu, F.; Horesh, R.; Chae, Y. T.; Zhang, R., Applying science and mathematics to big data for smarter buildings. *Annals of the New York Academy of Sciences* **2013**, *1295* (1), 18-25.
3. Li, D. H. W.; Yang, L.; Lam, J. C., Impact of climate changes on energy use in the built environment in different climate zones-A review. *Energy* **2012**, *42* (1), 103-112.
4. Murphy, K. D.; McCartney, J. S., Seasonal response of energy foundations during building operation. *Geotechnical and Geological Engineering* **2014**, *33* (2), 343-356.
5. Buildings Department *Comprehensive environmental performance assessment scheme for buildings: CEPAS application guidelines*; HKSAR Government: HK, 2006.
6. Muringathuparambila, R. J.; Musangoa, J. K.; Brentb, A. C.; Curriea, P., Developing building typologies to examine energy efficiency in representative low cost buildings in Cape Town townships. *Sustainable Cities and Society* **2017**, *33*, 1-17.
7. Copiello, S., Achieving affordable housing through energy efficiency strategy. *Energy Policy* **2015**, *85*, 288-298.
8. Konstantinidou, C. A.; Lang, W.; Papadopoulos, A. M., Multiobjective optimization of a building envelope with the use of phase change materials (PCMs) in Mediterranean climates. *International Journal of Energy Research 1*, 1-18.
9. Abdallah, M.; El-Rayes, K.; Liu, L., Economic and GHG emission analysis of implementing sustainable measures in existing public buildings. *Journal of Performance of Constructed Facilities* **2016**, *30* (6), 1-14.
10. Groissböck, M.; Heydari, S.; Mera, A.; Perea, E.; Siddiqui, A. S.; Stadler, M., Optimizing building energy operations via dynamic zonal temperature settings. *Journal of Energy Engineering* **2014**, *140* (1), 1-10.
11. Byun, J.; Shin, T., Design and Implementation of an energy-saving lighting control system considering user satisfaction. *IEEE Transactions on Consumer Electronics* **2018**, *64* (1), 61-68.

12. Menassa, C. C.; Kamat, V. R.; Lee, S. H.; Azar, E.; Feng, C.; Anderson, K., Conceptual framework to optimize building energy consumption by coupling distributed energy simulation and occupancy models. *Journal of Computing in Civil Engineering* **2014**, *28* (1), 50-62.

13. Holt, C. C., Forecasting trends and seasonals by exponentially weighted moving averages. *O.N.R. Memorandum* **1957**, *52* (52).

14. Brown, R. G., *Statistical forecasting for inventory control*. McGraw-Hill.: New York, 1959.

15. Hyndman, R. J.; Koehler, A. B.; Ord, J. K.; Snyder, R. D., *Forecasting with exponential smoothing: The state space approaches*. Springer-Verlag: Berlin, 2008.

16. Box, G. E. P.; Jenkins, G. M.; Reinsel, G. C., *Time Series Analysis: Forecasting and Control*. Holden-Day.: Oakland, California, 2011.

17. Gujarati, D. N., *Basic Econometrics*. Elsevier: Rio de Janeiro, 2006.

18. Pérez-Lombard, L.; Ortiz, J.; Pout, C., A review on buildings energy consumption information. *Energy and Buildings* **2008**, *40* (3), 394-398.

19. Costa, A.; Keane, M. M.; Raftery, P.; O'Donnell, J., Key factors methodology—A novel support to the decision making process of the building energy manager in defining optimal operation strategies. *Energy and Buildings* **2012**, *49*, 158-163.

20. Ballarini, I.; Corrado, V.; Madonna, F.; Paduos, S.; Ravasio, F., Energy refurbishment of the Italian residential building stock: energy and cost analysis through the application of the building typology. *Energy Policy* **2017**, *105*, 148-160.

21. Galik, C. S.; Abt, R. C.; Latta, G.; Meley, A.; Henderson, J. D., Meeting renewable energy and land use objectives through public-private biomass supply partnerships. *Applied Energy* **2016**, *172*, 264-274.

22. Carbonara, N.; Pellegrino, R., Public-private partnerships for energy efficiency projects: A win-win model to choose the energy performance contracting structure. *Journal of Cleaner Production* **2018**, *170*, 1064-1075.

23. Zhai, Z.; Abarr, M. L. L.; Al-Saadi, S. N. J.; Yate, P., Energy storage technologies for residential buildings. *Journal of Architectural Engineering* **2014**, *20* (4), 1-10.

24. Gökçe, H. U.; Gökçe, K. U., Integrated system platform for energy efficient building operations. *Journal of Computing in Civil Engineering* **2014**, *28* (6), 1-11.

25. Ascione, F.; Bianco, N.; Mauro, G. M.; Napolitano, D. F., Building envelope design: Multi-objective optimization to minimize energy consumption, global cost and thermal discomfort. Application to different Italian climatic zones. *Energy* **2019**, *174*, 359-374.

26. Mostavi, E.; Asadi, S.; Boussaa, D., Framework for energy-efficient building envelope design optimization tool. *Journal of Architectural Engineering* **2018**, *24* (2), 1-13.

27. Ham, Y.; Golparvar-Fard, M., Three-dimensional thermography-based method for cost-benefit analysis of energy efficiency building envelope retrofits. *Journal of Computing in Civil Engineering* **2015**, *29* (4), 1-11.

28. Aghemo, C.; Blaso, L.; Pellegrino, A., Building automation and control systems: A case study to evaluate the energy and environmental performances of a lighting control system in offices. *Automation in Construction* **2014**, *43*, 10-22.

29. Zhou, S.; Wu, Z.; Li, J.; Zhang, X., Real-time energy control approach for smart home energy management system. *Electric Power Components and Systems* **2014**, *42*, 315-326.

30. Ranasinghe, M., Total project cost: A simplified model for decision makers. *Construction Management and Economics* **1996**, *14* (6), 497-505.

31. Akpan, E. O.; Igwe, O., Methodology for determining price variation in project execution. *Journal of Construction Engineering and Management* **2001**, *127* (5), 367-373.

Peer-reviewed version available at *Energies* **2019**, *12*, 2888; doi:10.3390/en12152888

32. Trost, S.; Oberlender, G. D., Predicting accuracy of early cost estimates using factor analysis and multivariate regression. *Journal of Construction Engineering and Management* **2003**, *129* (2), 198204.
33. Touran, A., Probabilistic model for cost contingency. *Journal of Construction Engineering and Management* **2003**, *129* (3), 280-284.
34. Sonmez, R., Parametric range estimating of building costs using regression models and bootstrap. *Journal of Construction Engineering and Management* **2008**, *134* (12), 1011-1016.
35. Shane, J. S.; Molenaar, K. R.; Anderson, S.; Schexnayder, C., Construction project cost escalation factors. *Journal of Management Engineering* **2009**, *25* (4), 221-229.
36. Lu, T.; AbouRizk, S. M., Automated Box-Jenkins forecasting modelling. *Automation in Construction* **2009**, *18* (5), 547-558.
37. Fellows, R. F., Escalation management: Forecasting the effects of inflation on building projects. *Construction Management and Economics* **1991**, *9* (2), 187-204.
38. Williams, T. P., Predicting changes in construction cost indexes using neural networks. *Journal of Construction Engineering and Management* **1994**, *120* (2), 306-320.
39. Akintoye, A.; Bowen, P.; Hardcastle, C., Macro-economic leading indicators of construction contract prices. *Construction Management and Economics* **1998**, *16* (2), 159-175.
40. Ng, S. T.; Cheung, S. O.; Skitmore, M.; Wong, T. C., An integrated regression analysis and time series model for construction tender price index forecasting. *Construction Management and Economics* **2004**, *22* (5), 483-493.
41. Wong, J. M. W.; Chan, A.; Chiang, Y., Time series forecasts of construction labor market in Hong Kong: The Box-Jenkins approach. *Construction Management and Economics* **2005**, *23* (9), 979-991.
42. Hwang, S.; Liu, L. Y., Contemporaneous time series and forecasting methodologies for predicting short-term productivity. *Journal of Construction Engineering and Management* **2010**, *136* (9), 1047-1055.
43. Ashuri, B.; Lu, J., Time series analysis of ENR construction cost index. *ASCE Journal of Construction Engineering and Management* **2010**, *136* (11), 1227-1237.
44. Nam, H.; Han, S. H.; Kim, H., Time series analysis of construction cost index using wavelet transformation and a neutral network. In *24th International Symposium on Automation & Robotics in Construction* Madras., 2007.
45. Arlot, S.; Alain, C., A survey of cross-validation procedures for model selection. *Statistical Survey* **2010**, *4*, 40-79.
46. Chatfield, C., *Time series forecasting*. Chapman & Hall/CRC.: London, 2001.
47. Franzese, M.; Iuliano, A., Correlation analysis. *Encyclopedia of Bioinformatics and Computational Biology* **2019**, *1*, 706-721.
48. Pearson, K., Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London: 1895; Vol. 58, pp 240-242.
49. Bowerman, B. L.; O'Connell, R.; Koehler, A., *Forecasting, time series, and regression.* . Thomson Learning Inc.: U.S, 2005.
50. Hyndman, R. J.; Koehler, A. B.; Snyder, R. D.; Grose, S., A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* **2002**, *18* (3), 439-454.
51. Gardner Jr., E. S., Exponential smoothing: The state of the art. *International Journal of Forecasting* **2006**, *22* (4), 637-666.
52. Maia, A. L. S.; de Carvalho, F. A. T., Holtas exponential smoothing and neural network models for forecasting interval-valued time series. *International Journal of Forecasting* **2011**, *27* (3), 740-759.

Peer-reviewed version available at *Energies* **2019**, *12*, 2888; doi:10.3390/en12152888

53. Box-Jenkins, G. E. P.; Gwilym, M.; Reinsel, G. C., *Time Series Analysis: Forecasting and Control.* Prentice Hall: New Jersey, 1994.

54. Taneja, K.; Ahmad, S.; Ahmad, K.; Attri, S. D., Time series analysis of aerosol optical depth over New Delhi using Box-Jenkins ARIMA modeling approach. *Atmospheric Pollution Research* **2016**, *7*, 585-596.

55. Ljung, G. M.; Box, G. E., On a measure of a lack of fit in time series models. *Biometrika* **1978**, *65*, 297-303.

56. Hyndman, R. J.; Athanasopoulous, G., *Forecasting: principles and practice*. Wiley: Australia, 2012.

57. Krol, R., Evaluating state revenue forecasting under a flexible loss function. *International Journal of Forecasting* **2013**, *29* (2), 282-289.

58. Kankal, M.; Uzlu, E., Neural network approach with teaching-learning-based optimization for modelling and forecasting long-term electric energy demand in Turkey. *Neural Computing and Applications* **2016**, 1-11.

59. Musttaffa, Z.; Yusof, Y.; Kamaruddin, S. S., Enhanced artificial bee colony for training least squares support vector machines in commodity price forecasting. *Journal of Computer Science* **2014**, *5* (2), 196-205.

60. McNeish, D. M.; Stapleton, L. M., The effect of small sample size on two-level model estimates: A review and illustration *Educational Psychology Review* **2014**, *28* (2), 295-314.

61. Hanusz, Z.; Tarasinska, J.; Zielinski, W., Shapiro-Wilk test with known mean. *REVSTAT- Statistical Journal* **2016**, *14* (1), 89-100.

62. Makridakis, S.; Wheelwright, S.; Hyndman, R., *Forecasting: Methods and applications*. 3rd ed.; John Wiley & Sons: New York, 1998.

63. Pack, D. J., In defense of ARIMA modelling. *International Journal of Forecasting* **1990**, *6*, 211-218.

64. Chatfield, C., What is the best method of forecasting? *Journal of Applied Statistics* **1988**, *15* (1), 19-38.