CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

Forecasting energy demand and consumption is essential to energy management systems and policy. They help energy suppliers and policymakers make proper plans for when and how much energy would be needed and what energy would need to be produced to avoid incurring losses due to excess energy supply or energy scarcity. It provides information valuable for maintaining economic stability, efficient allocation of resources, and the transition to efficient and ecological energy types. As an essential tool, it is widely used in nations experiencing high economic growth and development rates, where energy demand dynamics can be pretty complex. It may depend on economic activity, population growth, and technological change (Alam et al., 2013).

There are several reasons why energy consumption has to be forecasted. First, it ensures that the energy supply meets the users' expectations, reducing wastage and increasing the reliability of the energy systems (Hahn et al. , 2009). Secondly, it plays an essential role in informing infrastructure developments such as the requirement for new power plants, adding to an electricity grid, and other investments in the sector worldwide (Debnath & Mourshed, 2018). Third is forecast accuracy, which means that energy delivery or consumption risks can be identified promptly while budgets can be planned. Furthermore, energy consumption predictions play a vital role in environmental management, cutting greenhouse gasses and incorporating renewable energy systems into the grid (Suganthi & Samuel, 2012).

Several factors make energy consumption forecasts significant. First, it helps to avoid energy waste and increase the reliability of systems that store energy through the timely matching demand and supply. Second, it is helpful for infrastructure forecasting as it helps understand whether new power plants, new grids, and other investments in the energy sector are required (Debnath & Mourshed, 2018). Thirdly, the budget is essential in managing risks; therefore, energy companies require accurate forecasts to have proper budgets. Besides, accurate energy consumption predictions are crucial for environmental management with the agenda of minimizing carbon footprint and incorporating more renewable supplies into the power system (Suganthi & Samuel, 2012).

Energy demand in the context of Malaysia has been on the rise primarily because of issues such as industrialization, urbanization, and population growth. The Malaysian energy industry is under pressure to satisfy rising energy demand while trying to adopt more sustainable forms of energy. Energy consumption forecasting is essential to accurately meet Malaysia's challenges since it offers a framework for tackling them through strategic planning and policies. They help provide energy security, economic optimality, and ecological harmony (Lau et al., 2018). Besides, Malaysia has a clear vision and policies on reducing greenhouse gas emissions and raising the utilization of renewable energy; hence, accurate energy forecasts are vital for ensuring energy policies and environmental targets are synchronized (Chong et al., 2015).

2.1 Definition and Concepts of Time Series Forecasting

Time series forecasting is a technique that models a series' responses for a given period based on the previous time-ordered record. These models are crucial in many professions, such as economists, financiers, climatologists, and energy engineers. The scope of time series forecasting determines a series' dependence on its time element to make the forecasts. One must understand trends, seasonality, and noise when working with time series.

- **Trend** refers to the long-term movement in the data, indicating an upward or downward trajectory over time.
- **Seasonality** refers to regular patterns or behaviour cycles over a specific period, such as daily, monthly, or yearly.
- **Noise** represents random fluctuations that cannot be attributed to trend or seasonality and are typically considered random errors.

These components must be determined and considered to improve the efficiency of time series forecasting. Other approaches to forecasting include basic approaches such as moving averages and exponential smoothing and advanced techniques such as ARIMA models (Box et al., 2015).

2.2 Overview of Regression Techniques and Their Application in Time Series Analysis

Regression techniques are one of the most influential models available to analyze the impact of one or more factors on a variable of interest. Regression methods are used to build and analyze the forecasts using the previous data to set the values in time series forecasting. There are several types of regression techniques commonly used in time series analysis

- **Linear Regression**: Using linear equations, this technique arranges the data to describe the dependency between the dependent variable and one or more independent variables. It is most appropriate for datasets with linear relationships between predictors, though not necessarily strongly positive linear relationships (Montgomery et al., 2012).
- **Polynomial Regression**: This technique estimates the dependent variable using another independent variable to test various factors simultaneously. It is useful when the effect of the independent variable on the dependent variable needs to be understood by

simultaneously controlling other variables that may affect the dependent variable (Montgomery et al., 2012).

- **Multiple Linear Regression**: This technique involves multiple independent variables to predict the dependent variable, allowing for simultaneously analyzing the impact of various factors. It is useful when the dependent variable is influenced by several factors (Montgomery et al., 2012).
- **Regularization Techniques**: Common techniques like Ridge Regression and Lasso Regression are employed to solve the multicollinearity problem and the problem of overfitting. Another method of shrinkage is the Ridge Regression, which incorporates a penalty term into the equation, reducing the coefficients of the less significant variables. However, Lasso Regression imposes a penalty that can make some coefficients equal to zero, making selecting variables more efficient (Hastie et al., 2009).

Other techniques used in conjunction with time series are regression techniques in timeliness forecasts of a given series. For example, one of the techniques often used is including lagged values of the dependent variable into the set of predictors into a regression model to analyze the temporal patterns in the data (Box et al., 2015). The other typical socio-economic, demographic, and climatic factors should also be incorporated to improve the model's capacity.

Regression techniques are flexible and conservative methods for modelling and forecasting time series data. By using these methods, researchers and practitioners can better understand the temporal features and characteristics of time series data and, in turn, make proper predictions.

2.3 Historical Energy Consumption Trends in Malaysia

Over the past few decades, Malaysia has experienced significant growth in energy consumption driven by rapid economic development, industrialization, and urbanization.

2.3.1 Review of past studies on energy consumption trends in Malaysia

Energy consumption in Malaysia in recent decades has been increasing rapidly due to economic development, industrialization, and urbanization. These have been undertaken in order to differentiate energy demand trends and explore the determinants of energy consumption in the future.

Ang (2008) also explored available literature to establish that the relationship between energy consumption and economic growth in Malaysia is positively related. The study showed that energy consumption has risen enormously due to industrialization and economic growth. In the same context, Chandran et al. (2010) examined the causality between EC, growth, and carbon emissions while stressing the importance of industrial and economic activities in driving EC rates.

Furthermore, Tan et al. (2017) performed a literature review of Malaysia's energy usage and factors influencing it, like population and urbanization. The work stated that population growth in the urban area had necessitated high use of energy on-premises, while industrial growth triggered high commercial and industrial energy usage. These works highlight the interdependence of Malaysia’s economic processes on the one hand and energy consumption on the other.

2.3.2 Factors influencing historical energy consumption patterns (economic growth, population growth, industrialization, urbanization)

Malaysia’s economic growth remains one of the country's most important determinants of energy usage today. Overall, economic growth leads to an increase in energy that must be met across different industrial, commercial, and domestic uses. Over the previous years, the rise in GDP is directly proportional to the increase in energy use. Similarly, Ang (2008) and Chandran et al. (2010) have also confirmed that economic development results in enhanced energy demand due to industrialization and transport and commercial requirements.

It must be noted that population growth directly leads to changes in society's energy demands. Population rise leads to the demand for energy required for residential spaces through activities such as heating, cooling, lighting, and more. Also, population growth leads to special economic activities financially, which in turn increases the rate of energy consumption. In this way, Tan et al. (2017) pointed out that the increase in population and improving living standards are the major factors that have led to higher energy consumption in Malaysia.

Industrialization has been a significant factor in Malaysia's energy use. Industrialization and manufacturing industries consume considerable energy for their production processes and machinery, as well as transport. As Malaysia is witnessing a gradual shift towards industrialization, the demand for energy from these segments has increased. This study by Ang (2008) posited that industries influence energy consumption in the country.

This is another critical factor and is linked with energy utilization in urban areas. This is because more people are moving to urban centers to reside, work, and conduct businesses and activities, increasing demand for energy in these areas. A consequence is that more energy is consumed per head of population in urban areas than in rural areas because the accessibility and usage of such energy-intensive facilities are higher. Similarly, Tan & Kaur (2017) establish that urbanization has played a role in the energy consumption profile in Malaysia in that it identified the energy consumption of urban households as higher than that of their rural counterparts.

2.4 Methods of Energy Consumption Forecasting

Traditional forecasting methods, such as moving averages, exponential smoothing, and ARIMA models, have been widely used for predicting energy consumption due to their simplicity and effectiveness in handling time series data. Each technique has its strengths and is suited to different types of data and forecasting requirements.

2.4.1 Moving Averages

Moving averages are essential and among the most popular working methods with time series data. This method involves assessing a certain number of previous values and assuming this as the forecast for the next period. There are two standard moving averages, SMA and WMA, wherein SMA sums up the values of previous periods and divides them by the number of periods. At the same time, WMA attributes more weight to current data than SMA.

1) **Simple Moving Average (SMA)**: SMA computes the average of the last ‘n’ values without weighting given to the most recent values. It eliminates fine trends in short-term movements and reveals long-term secular trends or cycles. However, it treats all observations as identical, which might not be best if the data has trends or seasonal structure (Hyndman & Athanasopoulos, 2018).

2) **Weighted Moving Average (WMA)**: WMA gives more importance to recent observations; therefore, different weights are assigned to different data points. This can apply more to data with trends so that the forecaster could affect small changes in subsequent terms (Makridakis et al., 1998).

2.4.2 Exponential Smoothing

Exponential smoothing is one of the most important and eclectic approaches in forecasting, as it assigns decreasing weights to previous observations. This is more effective when applied to data that have trends and seasonality inherent in them. Single exponential smoothing, double exponential smoothing, and triple exponential smoothing: These three are the most basic forms of exponential smoothing techniques

- **Single Exponential Smoothing**: It is applicable when the data contains no cycles or trends. It employs a smoothing factor to weigh more past observations, prioritizing the recent observations with exponentially decreasing weights to the older observations (Hyndman & Athanasopoulos, 2018).
- **Double Exponential Smoothing (Holt's Linear Trend Model)**: Linear trend decomposition enhances single exponential smoothing and incorporates additional factors to model linear trends in the data. It uses two smoothing constants, the level component and a trend component, which makes it feasible to extrapolate data with trends, as Holt (2004) recommended.
- **Triple Exponential Smoothing (Holt-Winters Method)**: TES is an extension of double exponential smoothing, where a new component is introduced to accommodate seasonality.

2.4.3 ARIMA Model

ARIMA (AutoRegressive Integrated Moving Average) models are statistical models for analyzing and forecasting time series data. They combine three components: autoregression (AR), differencing (I for integration), and moving average (MA).

- **Autoregression (AR)**: This component tries to capture the link between an observation and some previous observations, lagged in a certain way (Box et al., 2015).
- **Integration (I)**: This process involves differencing the data so that the variable becomes stationary and, therefore, free from trends or seasonality (Box et al., 2015).

- **Moving Average (MA)**: This component describes the association between observation and random error concerning a moving average equation of lagged variables (Box et al., 2015).

The general form of an ARIMA model is ARIMA (p, d, q), where:

- p is the number of lag observations the model includes (autoregressive order of differenced series).
- d represents the work done on the raw observations (degree of differencing).
- q is the size of the moving average window, which is the characteristics of the moving average (moving average order).

ARIMA models are robust and adaptable and can, therefore, be used for almost any type of time series forecasting. They can express complicated data features by transforming both plain and differenced series with autoregressive and moving average terms (Hyndman & Athanasopoulos, 2018).

2.5 Regression Techniques in Time Series Forecasting

Regression techniques are an essential statistical method that is employed to describe and predict dependence between a variable that is dependent on or the dependent variable and another one or more variables on which the dependent variable depends or the independent variables. In time series forecasting, simple regression models forecast future values using past data and other factors that may influence the particular variable. Most of these methods can handle linear and non-linear relationships and are more applicable in multivariate time series data analysis.

2.5.1 Linear Regression

Linear regression can be described as the former type of regression analysis that defines Y's functioning dependent on X by using a linear function. The equation of a simple linear regression model is:

Y=β0+β1X+ϵY = \beta_0 + \beta_1X + \epsilonY=β0+β1X+ϵ

Where ‘β0\beta_0β0’ is the intercept, ‘β1\beta_1β1’ is the slope of the line, and ‘ϵ\epsilonϵ’ represents the error term. Linear regression assumes that the relationship between the variables is linear and is particularly effective for datasets where this assumption holds (Montgomery et al., 2012).

2.5.2 Polynomial Regression

Polynomial regression is an extension of linear regression where a polynomial is fitted to the data you have. This technique can be used when the relationship between the dependent and the independent variables is curved. The general form of a polynomial regression model is:

Y=β0+β1X+β2X2+⋯+βnXn+ϵY = \beta_0 + \beta_1X + \beta_2X^2 + \cdots + \beta_nX^n + \epsilonY=β0+β1X+β2X2+⋯+βnXn+ϵ

Where ‘nnn’ represents the degree of the polynomial. By increasing the degree, polynomial regression can capture even more complex relationships in the data. However, higher-degree polynomials can lead to overfitting, where the model fits the training data too closely and performs poorly on new data (Seber & Lee, 2012).

2.5.3 Multiple Linear Regression

Multiple linear regression models the relationship between a dependent variable and multiple independent variables. The equation for multiple linear regression is:

Y=β0+β1X1+β2X2+⋯+βpXp+ϵY = \beta_0 + \beta_1X_1 + \beta_2X_2 + \cdots + \beta_pX_p + \epsilonY=β0+β1X1+β2X2+⋯+βpXp+ϵ

where ‘X1,X2,…,XpX_1, X_2, \ldots’, ‘X_pX1,X2,…,Xp’ are the independent variables. This technique allows for the inclusion of several predictors, enabling a more accurate analysis of the factors influencing the dependent variable. Multiple linear regression is widely used in time series forecasting to incorporate various predictors such as economic indicators, demographic factors, and climatic variables (Montgomery et al., 2012).

2.5.4 Regularization Techniques

Regularization methods affect the excess fitting of regression models by adding loss condition. This assists in reducing the value of the model coefficients so that it is easier for the model to generalize new information. Two popular methods are Ridge Regression and Lasso Regression.

- **Ridge Regression**: Ridge regression, also known as Tikhonov regularization, adds a penalty term proportional to the square of the magnitude of the coefficients to the loss function. The ridge regression objective function is:

  Minimize(∑i=1n(yi−β0−∑j=1pβjxij)2+λ∑j=1pβj2)\text{Minimize} \left(\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p \beta_j^2 \right)Minimize(∑i=1n(yi−β0−∑j=1pβjxij)2+λ∑j=1pβj2)

  Where ‘λ\lambdaλ’ is the regularization parameter that controls the strength of the penalty. Ridge regression effectively handles multicollinearity and prevents overfitting by shrinking the coefficients of less important variables toward zero but not exactly to zero (Hastie et al., 2009).

- **Lasso Regression**: Lasso (Least Absolute Shrinkage and Selection Operator) regression adds a penalty term proportional to the absolute value of the coefficients to the loss function. The lasso regression objective function is:

  Minimize(∑i=1n(yi−β0−∑j=1pβjxij)2+λ∑j=1p∣βj∣)\text{Minimize} \left(\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| \right)Minimize(∑i=1n(yi−β0−∑j=1pβjxij)2+λ∑j=1p∣βj∣)

  Lasso regression can lower some coefficients to zero, effectively performing variable selection. This makes it useful for models where feature selection is necessary to identify the most significant predictors.

2.5 Key Influencing Factors on Energy Consumption

Energy consumption can be influenced by various factors broadly categorized as economic, demographic, climatic, and technological. Understanding these factors is crucial for accurate energy consumption forecasting and effective energy management.

2.5.1 Economic Factors

- **Gross Domestic Product (GDP)**: GDP is commonly considered one of the most essential measures that characterize the economic growth of a country. It shows that GDP has a positive association with energy consumption, where any growth in GDP affects energy due to higher industrial activities, door commercial service, and improved standard of living, as Ang (2008) highlighted. Research has established that economic development increases energy consumption in industry, transport, and end-user sections.

- **Industrial Output**: Industrial output is another pivotal element for determining energy consumption. As one moves from the pre-production process towards production, more energy-demanding processes, equipment, and plants are called for. Any country that witnesses rapid industrialization, such as Malaysia, has witnessed a sharp rise in energy demand in proportion to industrial development (Chandran et al., 2010). Another factor that influences energy demand is the type of industries in a country.

2.5.2 Demographic Factors

- **Population Growth**: As the population rises, energy demand increases because many people require energy to heat, cool, light, and power their homes. Also, as the increased population leads to higher consumption of products within the commercial and industrial segments, it boosts energy utilization (Tan et al., 2017). Population growth is a critical driver of energy demand, and as the population in the cities grows, this is anticipated to rise exponentially, particularly given improved physical access to energy-consuming devices and amenities.
- **Urbanization**: Urbanization refers to the increasing concentration of populations in urban areas. Residents in urban areas typically have higher per capita energy consumption than their rural counterparts due to better access to

electricity, higher living standards, and the widespread use of energy-intensive appliances and technologies. Urbanization also leads to greater demand for infrastructure development, transportation, and commercial activities, all of which contribute to higher energy consumption (Tan et al., 2017).

2.5.3 Climatic Factors

- **Temperature**: Climate is another element that describes the external conditions and variables that strongly influence energy consumption for heating and cooling. In warmer climates, changes in trade temperatures increase the utility of air conditioning, while in colder climates, changes in trade temperatures increase the utility of heating. Potential predictors such as temperature changes throughout the year also cause significant changes in energy usage levels (Al-Ghandoor et al., 2013). The climate in Malaysia is tropical; therefore, the temperatures are high throughout the year, so the requirement for cooling power is reasonably constant.

- **Humidity**: Humidity impacts energy usage profiles due to the changes in energy demand when it is hot or humid outside. Relative humidity is one factor that raises discomfort in a room since it leads to increased usage of AC and dehumidifiers. Pérez-Lombard et al., 2008 have documented that the energy used for indoor air conditioning and achieving comfort, especially in buildings for living and working, is colossal in humid regions.

References

1. Alam, M. J., Begum, I. A., Buysse, J., Rahman, S., & Van Huylenbroeck, G. (2013). Energy consumption, CO2 emissions and economic growth nexus in Bangladesh: Cointegration and dynamic causality analysis. *Energy Policy*, 61, 47-54.
2. Chong, C. S., Lau, K. Y., Tan, C. W., & Yatim, A. H. M. (2015). Statistical analysis of Malaysia wind speed data using R software: A case study. *Renewable Energy*, 83, 29-38.
3. Debnath, K. B., & Mourshed, M. (2018). Forecasting methods in energy planning models. *Renewable and Sustainable Energy Reviews*, 88, 297-325.
4. Hahn, H., Meyer-Nieberg, S., & Pickl, S. (2009). Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research*, 199(3), 902-907.
5. Lau, K. Y., Ching, C. S., Lim, Y. S., & Yatim, A. H. M. (2018). Assessing the impact of weather on energy demand in Malaysia: An analysis of residential electricity consumption. *Energy Procedia*, 143, 564-569.
6. Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, 16(2), 1223-1240.
7. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.
8. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
9. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons.
10. Seber, G. A. F., & Lee, A. J. (2012). *Linear Regression Analysis*. John Wiley & Sons.
11. Ang, J. B. (2008). Economic development, pollutant emissions and energy consumption in Malaysia. *Journal of Policy Modeling*, 30(2), 271-278.
12. Chandran, V. G. R., Sharma, S., & Madhavan, K. (2010). Electricity consumption–growth nexus: The case of Malaysia. *Energy Policy*, 38(1), 606-612.
13. Tan, F., Lau, L., & Lee, K. (2017). Energy consumption, economic growth and environmental degradation in Malaysia. *Energy Policy*, 107, 53-61.
14. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

15. Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5-10.
16. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
17. Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications*. John Wiley & Sons.
18. Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324-342.
19. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
20. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons.
21. Seber, G. A. F., & Lee, A. J. (2012). *Linear Regression Analysis*. John Wiley & Sons.
22. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
23. Al-Ghandoor, A., Jaber, J. O., Al-Hinti, I., & Sawalha, S. A. (2013). Residential past and future energy consumption: Potential savings and environmental impact. *Renewable and Sustainable Energy Reviews*, 24, 103-109.
24. Ang, J. B. (2008). Economic development, pollutant emissions and energy consumption in Malaysia. *Journal of Policy Modeling*, 30(2), 271-278.
25. Chandran, V. G. R., Sharma, S., & Madhavan, K. (2010). Electricity consumption–growth nexus: The case of Malaysia. *Energy Policy*, 38(1), 606-612.
26. Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings*, 40(3), 394-398.
27. Tan, F., Lau, L., & Lee, K. (2017). Energy consumption, economic growth and environmental degradation in Malaysia. *Energy Policy*, 107, 53-61.