International Journal of Advanced and Applied Sciences, 10(4) 2023, Pages: 76-87

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Time series forecasting of solid waste generation in selected states in Malaysia

Noryanti Nasir [1, 2, *], S. Sarifah Radiah Shariff [1, 2, 3], Siti Sarah Januri [3], Faridah Zulkipli [4], Zaitul Anna Melisa Md Yasin [5]

[1]School of Mathematical Sciences, College of Computing, Informatics, and Media, Universiti Teknologi MARA, Shah Alam, 40450, Selangor, Malaysia
[2]Logistic Modelling Research Group, College of Computing, Informatics, and Media Sciences, Universiti Teknologi MARA, Shah Alam, 40450, Selangor, Malaysia
[3]Malaysia Institute of Transport (MITRANS), Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[4]Mathematical Sciences Studies, College of Computing, Informatics, and Media, Universiti Teknologi MARA, Negeri Sembilan Branch, Seremban Campus, 70300 Seremban, Negeri Sembilan, Malaysia
[5]Mathematical Sciences Studies, College of Computing, Informatics, and Media Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, 35400, Tapah Road, Perak, Malaysia

A R T I C L E I N F O

Article history:
Received 9 September 2022
Received in revised form
30 December 2022
Accepted 6 January 2023

Keywords:
Solid waste management
Landfills
Forecasting
ARIMA

A B S T R A C T

This study aims to forecast Malaysian solid waste generation by identifying the state's landfill capacity to facilitate solid waste generated in the next two years. The solid waste management system depends extremely on landfill capacity. Due to the increased amount of solid waste generation, the authority is required to manage landfill utilization appropriately in selected regions, where landfill capacity was fully utilized. An accurate prediction of solid waste generation is required for the authority plan for landfill management. This paper provides the forecasting values for the seven states in Malaysia. The ARMA and ARIMA models are used to determine the best model for forecasting solid waste generation values. The results show that the ARIMA (2, 1, 1) model works best in Johor, Negeri Sembilan, and Wilayah Persekutuan Kuala Lumpur, while the ARIMA (1, 1, 2) model works best in Kedah and Perlis. Furthermore, the ARMA (1, 1) model is best for Pahang, and the ARMA (2, 1) model is best for Melaka. The ARIMA (3, 1, 1) model is the best for forecasting solid waste generation across all states. The findings are consistent with previous literature, which stated that solid waste generation would increase in one of Malaysia's districts over the next two years. They did not, however, consider the landfill's capacity to handle solid waste generation. These findings shed light on the potential volume of solid waste generated in the coming years, allowing authorized agencies to plan landfill capacity in Malaysia for environmental sustainability.

1. Introduction

An overload of solid waste is a never-ending concern not only in Malaysia but all around the world. Solid waste comes from human activities and falls into several categories which are household, commercial, construction, and demolition, industry and institutions are part of them. Solid waste management is a necessary part of the business to protect the environment and to ensure sustainability and quality of community life (Zulkipli et al., 2020). The main waste composition is composed of organic or food waste, paper, plastic, metal, glass, and others. Apart from the rapid increase in waste management system expenditure, the way people manage these wastes may also bring adverse effects on the environment and public health (Ferronato and Torretta, 2019). The most common practice of waste disposal is landfilling (Siddiqua et al., 2022). It is acknowledged as a significant alternative either now or in the near future, especially in low-income and middle-income countries, as it is the easiest and cheapest available technology (Sharifah and Latifah, 2013).

Malaysians produced about 38,000 tons of waste daily and 47 percent of the solid waste was food

* Corresponding Author.
Email Address: noryantinasir@uitm.edu.my (N. Nasir)
https://doi.org/10.21833/ijaas.2023.04.009
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-1475-0286

waste. The solid waste management system in Malaysia is highly dependent on landfills. The National Solid Waste Management Department reported in 2015, there are a total of 296 landfills in Malaysia where 165 are operational landfills while the other 131 are closed due to waste generated having reached maximum landfill capacity. Many landfills are reaching their design capacity and constructing new landfills becomes more challenging as population growth has led to land scarcity. In short, it is not surprising that Malaysia's solid waste management system has become tedious to handle. With major growth in population, land scarcity has become an obstacle in constructing new landfills for waste disposal. Landfill space was exhausted earlier than scheduled and was no longer sustainable in terms of security of disposal. The amount of waste generation increases parallel with development and population growth (Wu et al., 2020). Low awareness and participation in recycling among Malaysians should be improved to minimize the further impact of uncontrolled solid waste generation on the environment, society, and economy.

Landfill played a bigger role in waste management and sanitary landfill is one of the popular waste disposal methods. Here, the landfill needs to be poisoned and isolated from the environment until it is safe. Landfill could be environmentally harmful if not controlled. Every state in Malaysia has its own landfill however; it is limited due to environmental safety. Each landfill has a lifespan and maximum capacity that can handle the waste. As the number of waste generated increases few of the landfills had to be closed due to it reaching maximum capacity. Besides Malaysia, Nigeria had experienced the same issue of lack of landfill capacity, which is the residual life span is approximately 8 years (Emetere and Iroham 2021). Solid Waste Corporation (SWCorp) is one of the national agencies established under Act 672 and is in charge of solid waste management in Malaysia. This agency is in charge of seven Malaysian states including Johor, Kedah, Melaka, Negeri Sembilan, Pahang, Perlis, and Wilayah Persekutuan Kuala Lumpur. The objective of this study is to forecast solid waste generation in the selected states in Malaysia, thereby identifying whether the states' landfill capacity able to accommodate the solid waste generated in the next two years ahead.

Forecasting can be the formal procedure of making expectations by using the economic concept, mathematics, statistics, and econometric analysis (Asadullah et al., 2021). If the accurate prediction of exchange rates is critical for investment and business purposes, then the accurate prediction of solid waste generation is also critical for a better and healthier environment in the future. This is because individuals and local authorities consider forecasting when they make economic decisions. These decisions then affect the course in which the economy will continue.

Generally, there are numerous methods of forecasting. The renowned method Auto-Regressive Integrated Moving Average (ARIMA) has been widely used to forecast, introduced by Box & Jenkins in the 1970s. ARIMA is claimed to be the most popular and frequently used stochastic time series model that returns the highest forecasting accuracy. ARIMA model is reported able to predict the future values of a time series using a linear combination of past values and a series of errors (Zafra et al., 2017), and suitable for all kinds of data such as level, trend, seasonality, cyclicality and many more (Ceylan et al., 2020). This method performs well whether the data is stationary or non-stationary (Sriploy and Lertpocasombut, 2020). The ARIMA model forecasts were better than Winters' methods which are the additive and multiplicative methods in forecasting (Ayakeme et al., 2021). While the Support Vector Regression (SVR)-ARIMA model is the best model for forecasting solid waste generation (Chen and Dai, 2020).

The combination of the SVR and ARIMA models is effective to determine the linear and nonlinear trends in influencing factors and independent variables. Moreover, Ceylan et al. (2020) forecasted medical waste generation in Istanbul using Support Vector Regression (SVR), GM (1,1), ARIMA Model, and Linear Regression (LR). The study reveals that the ARIMA (0,1,2) model performed better than other methods based on maximum absolute error. The result indeed shows that the ARIMA model is the best model to make forecasting. Other than that, Mohamad et al. (2022) used Artificial Neural Network (ANN) to forecast the municipal solid waste generation in Klang, Selangor. However, the forecasting on municipal solid waste generation done for one district in Selangor State and moreover, considering the landfill capacity is omitted.

2. Methodology

2.1. Description of data

A time-series data of solid waste generated from January 2016 to August 2020 under the solid waste management and Public Cleansing Corporation (SWCorp) is used to forecast the future solid waste generation in Johor, Kedah, Melaka, Negeri Sembilan, Pahang, Perlis and Wilayah Persekutuan Kuala Lumpur by using the ARIMA model. This data provides the total amount of solid waste (tons) generated in each state. Table 1 shows the total solid waste generated from 2016-2020 for Johor, Kedah, Melaka, Negeri Sembilan, Pahang, Perlis, and Wilayah Persekutuan Kuala Lumpur. The data are obtained from the official website of SWCorp. The state of Johor has the highest amount of solid waste generated, followed by Kuala Lumpur and Kedah due to its high population rate.

Table 1: Solid waste generated in seven states from 2016-2020

States	Total solid waste generated (tons)				
	2016	2017	2018	2019	2020
Johor	1,028,157.26	932,493.89	912,112.00	924,201.24	878,221.00
Kedah	563,945.89	456,181.96	500,843.36	502,426.20	479,660.95
Melaka	228,390.63	233,548.30	247,948.84	243,357.03	240,297.88
Negeri Sembilan	284,998.84	278,696.78	291,166.30	299,149.59	281,432.29
Pahang	310,977.97	302,306.39	289,311.85	278,071.45	284,065.99
Perlis	56,494.00	41,893.65	42,179.55	43,289.57	45,715.06
WPKL	823,830.89	773,684.17	759,900.86	765,146.42	677,090.92

2.2. Box Jenkins Method

Box Jenkins method is a systematic method of identifying, fitting, checking, and using ARIMA time series models. ARIMA model is a model that can represent stationary and non-stationary time series and produce an accurate forecast based on a description of historical data of a single variable. There are four steps in ARIMA models (Fattah et al., 2018), shown in Fig. 1.

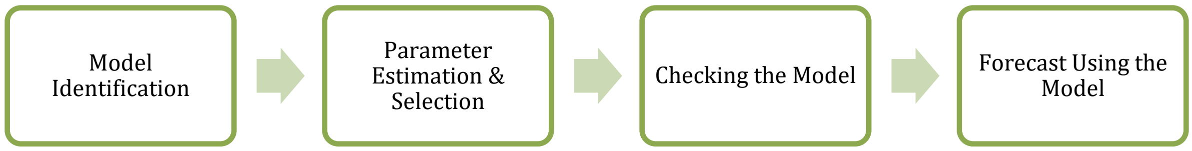


Fig. 1: The ARIMA procedures

Box-Jenkin's methodology assumed that a data series is stationary. For a series that did not fulfill the stationary condition, it would be called non-stationary series. The method of transforming non-non-stationary data series into a stationary data series is called differencing. By differencing, the mean of a time series was stabilized by eliminating changes in the level of a time series thus consequently dispensing trend and seasonality. When a data series appeared non-stationary, the first difference was performed. The first difference was defined as, $\Delta y_t = (y_t - y_{t-1})$ where y_t is the current value and y_{t-1} is the previous value. If the data remained non-stationary, second-order differencing was executed to obtain a stationary condition. After the stationary condition has been accomplished, the parameters of the ARIMA model must be identified. ARIMA model has three parameters which are autoregressive (p), differencing (d), and moving average (q) (Chintalapudi et al., 2020). The autocorrelation function (ACF) and partial correlation function (PACF) need to be plotted to identify autoregressive (p) and moving average (q) parameters. If the ACF decayed exponentially and PACF spiked, the process was an Autoregressive (AR) model. It was then identified as AR (p) where p is the number of spikes in the PACF.

The Moving Average (MA) was best used when PACF decayed and the ACF spiked. The value of parameter q is equal to the number of significant spikes in ACF. The parameter d refers to the order of differencing required by the time series to get stationery. In general terms, the model obtained was defined as ARIMA (p,d,q). It was important to differentiate the data series to achieve stationary in where the symbol 'd' defined the number of times the y variable must be different to reach stationary. A simple model case ARIMA (1,1,1) is as shown below,

$$w_t = \mu + \emptyset_1 w_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t \tag{1}$$

where, $w_t = y_t - y$, serve the first difference of the series and is considered to be stationary. In this scenario, the values of p=1, d=1, and q=1.

2.3. Unit root test

In order to determine the trend of the data in time series, a unit root test is used. It is to determine whether the data need to be differencing to achieve a stationary state. Unit root test is often the first step in the procedure of forecasting. Testing the stationary of data is important in time series analysis. Stationary is a statistical property that generates time series and does not change over time. Many statistical tests, analytical tools, and models rely on stationary data. The basic methods to determine the stationary properties are by plotting and visualizing the data. In this study, stationary is determined by plotting the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). The plot of ACF and PACF is called Correlogram. The data is stationary if the pattern of the Correlogram is decreasing to zero quickly while for non-stationary data the decrease is slow (Arzo et al., 2021). Additionally, the Augmented Dickey-Fuller (ADF) is used to analyze the stationary of the data series. Here the p-value obtained should be less than the significance level inferring that the series is stationary (Mohamed, 2008).

2.4. Validation of forecasting model

Specific models can generate good forecast values in a particular situation because each model type has unique characteristics which define its suitability to be fitted to the given data series. Evaluating the model is part of a significant process in most forecasting activities. Forecasters can identify how well the performance of the selected model forecasts. The large forecast errors by the wrong selection of forecasting method would give the effect of substantial financial loss to the organization.

Models can be evaluated to determine the one that may generate the best forecast values. Error measure is to differentiate between a poor forecast model and a good forecast model thus providing how close forecasts are to actual values. The standard measurement to determine a good forecast is when the error measure has the smallest value. The error can be computed from Mean Square Error (MSE) in Eq. 2, The Akaike information criterion (AIC) in Eq. 3, and Bayesian Information Criteria (BIC) in Eq. 4. The MSE equation is given as follows

$$MSE = \frac{\sum e_t^{\,2}}{n} \tag{2}$$

where, $e_t = y_t - \hat{y}_t$ where y_t is the actual observed value at time t and $\hat{y}_t$ is the fitted value at time t. The AIC is a mathematical method for assessing how efficiently a model fits the data from which it was produced. AIC was implemented to compare distinct models and discover which one is the best fit for the data (Niu et al., 2021). The AIC equation is given as

$$AIC = e^{\frac{2k}{T}} \sum_{t=1}^{T} \frac{e_t^2}{T} \tag{3}$$

where, $k=p+q+P+Q$ depict the number of parameters estimated in the model, for p and q the usual respective terms of the AR and MA parts the P and Q the seasonality part of the ARIMA model and T is the total number of observations in the data series. The test aimed to minimize the value of AIC by choosing the right p, q, P, and Q. For instance, a model was deemed to be having a better fit than other models if the value of the AIC was the lowest. The BIC was developed by Schwarz (1978), where Bayesian arguments for adopting it. The BIC aimed to choose a model that achieves the most accurate out-of-sample forecast by stabilizing between the models' complexity and goodness of fit (Hyndman, 2015). The BIC was calculated as

$$BIC = T^{\frac{k}{T}} \sum_{t=1}^{T} \frac{e_t^2}{T} \tag{4}$$

where, k is the number of parameters in the estimated model including the constant and T is the number of observations. The BIC was linked to AIC and one of the similarities was that the lower the value of BIC, the model was said to be the best ARIMA model. The Durbin Watson (DW) statistic indicates the presence of serial correlation of the residuals. Although serial correlation does not affect the consistency of the estimated regression coefficients, it does affect the ability to conduct valid statistical tests. The best DW value is around 2 which indicates no autocorrelation (Kim, 2022; Kumar and Kumar, 2021).

3. Results and discussions

This section presents the analysis in determining the trend analysis and forecasting values of solid waste generation for Johor, Kedah, Melaka, Negeri Sembilan, Pahang, Perlis, and Wilayah Persekutuan Kuala Lumpur. In addition, combinations of all states are also analyzed to give an overview of Malaysia's solid waste generation. The results are determined according to the objectives of the studies. In this section, the findings are presented and explained comprehensively based on the objectives.

3.1. Trend analysis on solid waste generation

The data of the solid waste generated in Johor, Kedah, Melaka, Negeri Sembilan, Pahang, Perlis, and Wilayah Persekutuan Kuala Lumpur (WPKL) are analyzed to identify whether landfill capacity can hold the solid waste generation in the future or not. Besides, the data of all states also are summed up to be analyzed from January 2016 to December 2020.

Fig. 2a until 2g show the trend analysis on solid waste generation over time in Johor, Kedah, Melaka, Negeri Sembilan, Pahang, Perlis, and Wilayah Persekutuan Kuala Lumpur, respectively while Fig. 2h illustrates the trend analysis of all states combined. In Fig. 2a, Johor state shows a decreasing pattern from 2016 to 2017, and then an increasing pattern until the middle of 2018. Following that, there is no obvious upward or downward trend until the first quarter of 2020. A short-term memory random shock occurred in April 2020, possibly because citizens are obligated to stay at home in accordance with the first Movement Control Order. Trend analysis on solid waste generation in Fig. 2b illustrates an irregularity pattern of a turning point in Kedah. The amount of waste generated in 2016 stayed above 40,000 tons and peaked in December 2016. Then, in 2017 the pattern started to fall and never rise above 48,000 tons up until the end of 2020.

Fig. 2c of Melaka's solid waste generation shows a major increase in May 2018. The increment may be due to the start of Ramadan for Muslims in Malaysia. The streets would be buzzing with bazaars a few hours before Iftar. Even non-Muslims enjoy going to the bazaars since there will be a smorgasbord of local delicacies. Consequently, people tend to waste food during this month. Towards the end of Ramadan, people will start preparing for Raya where the shopping mall and the night market will be swamped with people. More waste is generated during this phase since the restaurants or eateries are packed with people either to break their fast or to have supper. On the other hand, there exist irregularities in the data series. The occurrence of random shock (short-term memory) which affects the trend and later returns to the normal level can be seen from March 2020 to May 2020. For the short-term memory effect, the level returns to normal within a short period of time. Based on the graph, the sudden decrease in solid waste generated during that period may be due to the Restricted Movement Control Order (RMCO) implemented to curb the global virus outbreak, Covid-19. During that phase, people are obligated to stay at home. All outdoor activities including shopping, going on vacation, and eating at the eatery were not allowed. These contribute to less solid waste generated during that period.

Fig. 2d shows that in Negeri Sembilan from July 2018 to July 2019, the volume of the solid waste generated can be seen increasing compared to the previous month and then reduced in early 2020. The decrements may be due to people staying at home during the Restricted Movement Control Order (RMCO). However, the solid waste generated increased tremendously starting April 2020 onward. This is because the RMCO was lifted during that period and implementation of the Conditional Movement Control Order (CMCO) begins where people are allowed to go out as long as they follow the Standard Operating Procedure (SOP).

Next, the graph in Fig. 2e illustrates a gradually decreasing from March 2016 to March 2020 for Pahang State. It started to increase from April 2020 after the lowest volume of solid waste was generated from the year 2016 to the year 2020. Fig. 2f shows the trend analysis in Perlis. The generated graph remained the same from January 16 to December 16. The highest volume of solid waste generated is on June 18. It started to drop dramatically on July 18 and increased gently towards the end of 2020. Moreover, the January 2016 to December 2020 graph of waste generation in Kuala Lumpur in Fig. 2g shows a gradual downward trend, with a random shock depicted during early 2020. This may happen due to the same reason as RMCO implementation. When combining all states involved, solid waste generated by Malaysians is more than 260,000 tons for the year 2016 and continues to drop for the following year. The ton of waste fluctuates at an average of 260,000 tons. During early 2020, when COVID-19 hits and Movement Control Order was implemented, it is acceptable to say that Malaysians followed the restriction as the solid waste generated experienced a sudden drop which can be classified as a random shock (short-term memory). It can be concluded that when people stay at home, less solid waste is generated.

3.2. Stationary

To ensure a better relevant stationary situation, the data is analyzed by plotting the autocorrelation function (ACF) and partial autocorrelation function (PACF). Another way to determine the stationary of data is by conducting a unit root test. The correlograms of respective states are shown in Fig. 3a until Fig. 3h. Fig. 3a shows the correlogram of the Johor state after the first differencing has been made. From Fig. 3a, it is obvious to ARIMA (1,1,1) as the most noticeable spike can be seen at lag 1 for both ACF and PACF. However, model identification must be done to carefully select the best model. Based on Augmented Dickey-Fuller (ADF) performance, the unit root test resulted in all three critical values being higher than the t-statistic of-13.756. The same rules are applied to Kedah as the data is not stationary. At the first level of differentiation, Fig. 3b shows no decaying pattern, and no values surpass the significance limit. Besides, the value of t-statistics in Kedah is-8.582 with a probability value of 0.000, which is substantially lower than all three critical values. It proved that the Kedah sequence is stationary after the first level of differentiation.

In Fig. 3c, Melaka state reveals that there is no decaying trend for both ACF and PACF correlogram. Therefore, it can be concluded that the series is stationary, and no differentiation is needed. To be precise, a unit root test for Melaka also proved the stationery with a t-statistic value of-5.853 and a probability value of 0.000 is lower than all critical values obtained. While in Fig. 3d Negeri Sembilan, the values of ACF and PACF plots indicate no decaying pattern after the first level of difference. The data is stationary as the value of the t-statistic is-14.287 which is lower than other critical values with a 0.000 probability value. Fig. 3e shows no values surpassing the maximum significance. It is concluded that the series is stationary and no differentiation from the initial data is needed. In addition, the t-statistic of-4.858 with a probability value of 0.000 is significantly lower than all three critical values given.

For Perlis, the values of ACF and PACF plots in Fig. 3f have a decaying pattern at the level. Hence, differentiation is needed as the data is not stationary. After performing differentiation, Fig. 3f reveals no decaying trend for both ACF and PACF. Besides, the value t-statistic of Perlis is-8.931 which is lower than other critical values with a probability value of 0.000. Similar to Fig. 3a of the Johor state, Fig. 3g also depicts the obvious model to be selected which can be seen at lag 1. ACF and PACF value at lag 1 shows the most noticeable spike, thus ARIMA (1,1,1) is selected. The t-statistic value of-13.032 is lower than all critical values. No decaying pattern after the first differencing indicates that the series in Fig. 3h is stationary. Also, the t-statistic of-11.996 with a probability value of 0.000 is lower than all critical values.

3.3. Model identification

This study conducted the model identification through the ACF and PACF pattern from the stationary series. ACF and PACF are included in the process of determining the appropriate models to be fitted to the data series. However, the model based on the number of spikes in ACF and PACF is not simple to determine precisely.

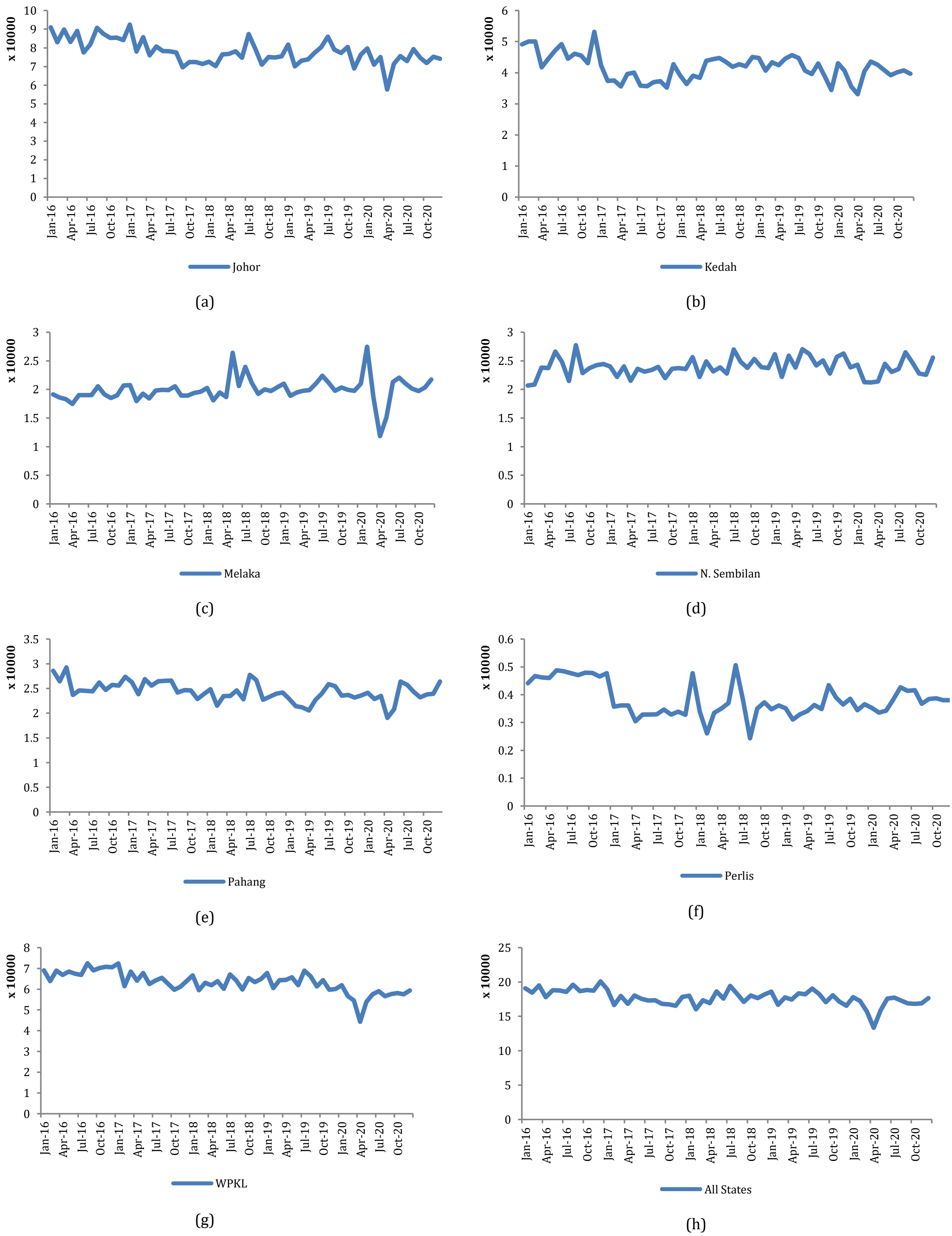
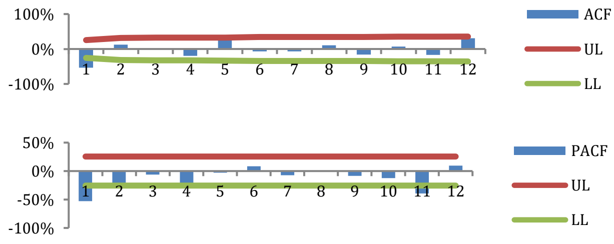
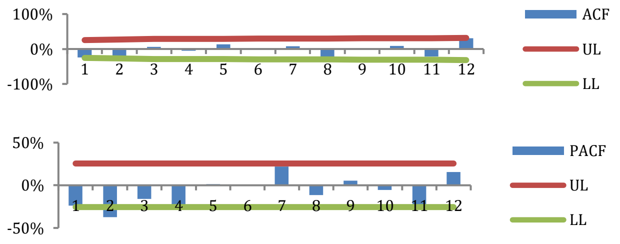


Fig. 2: Trend analysis on solid waste generation for each state (in tons)

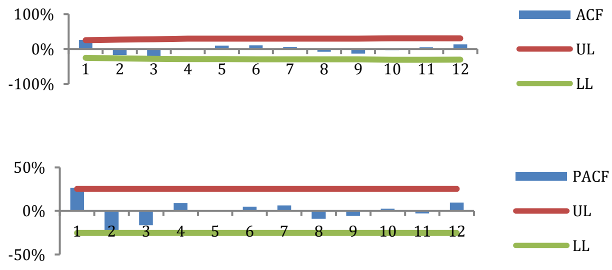
To determine the number of lags needed, a careful judgment of the position and size of spikes is necessary. The ACF and PACF graph in Fig. 3 represents the MA and the AR, where both AR and MA will be denoted as p and q, respectively. Since some of the data is not stationary, the value of differencing, d is detonated as 1. Based on the correlogram in Fig. 3a, the most noticeable spike in ACF and PACF for Johor is at lag 1 with the estimated model of ARIMA (1,1,1). The model that can be evaluated for Kedah is ARIMA (1,1,2). For Melaka, ARMA (2,1) is used to evaluate the model as the noticeable spike for ACF is at lag 2 and PACF at lag 1. The model that can be evaluated for Negeri Sembilan is ARIMA (2,1,1) as the most noticeable for ACF is seen at lag 1 and PACF at lag 2.



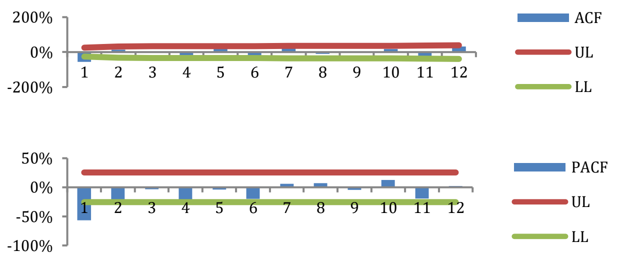
(a). For Johor at first difference



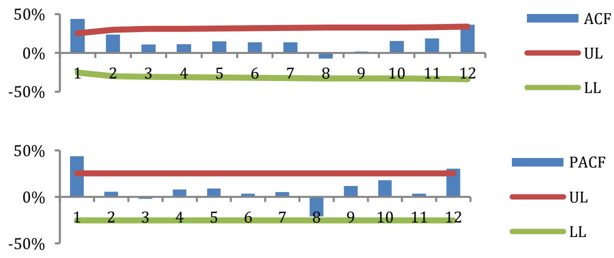
(b). For Kedah at first difference



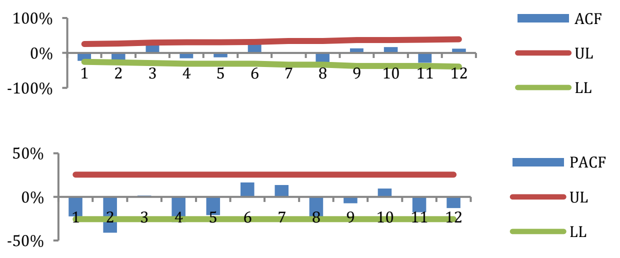
(c). For Melaka at level



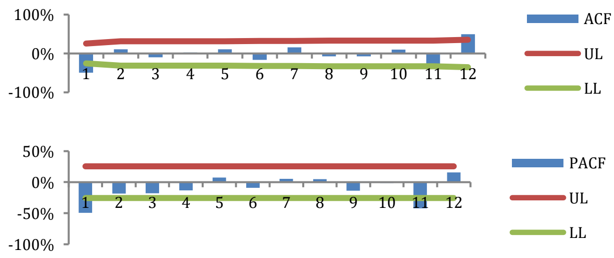
(d). For Negeri Sembilan at first difference



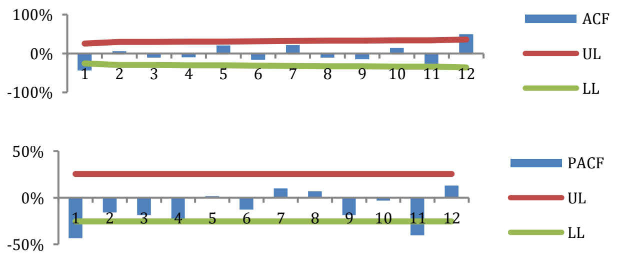
(e). For Pahang at level



(f). For Perlis at first difference



(g). For WPKL at first difference



(h). For all states at first difference

Fig. 3: Correlogram and ADF

The model that can be estimated for Pahang is ARMA (1,1). From the correlogram for Perlis in Fig. 3f, based on the most noticeable spike, the model that can be evaluated for Perlis is ARIMA (1,1,2). The model that can be evaluated for WPKL is ARIMA (1,1,1) using the parsimony principle. After summing up all the states, the model that can be evaluated is ARIMA (3,1,1) as the most noticeable spike in ACF and PACF can be noticed at lag 1 and 3.

3.4. Performance evaluation

From the model identification, five models of ARIMA are assumed and examined to obtain more definite evidence. In this stage, all five models are evaluated by choosing the lowest Normalized Bayesian Information Criterion (BIC), Akaike Info Criterion (AIC), and Durbin Watson (DW) value. Table 2a shows the statistics value for Johor. According to Table 2a, the results of DW for ARIMA (3,1,1) is around 2, however ARIMA (2,1,1) has the lowest results of AIC and BIC. Therefore ARIMA (2,1,1) is chosen as the best model to analyze the solid waste generation for the Johor state. Table 2b shows the statistics value for Kedah. ARIMA (1,1,2) is chosen as the best model to analyze solid waste generation as it has the lowest AIC and BIC with a value of DW around 2. Even though ARIMA (2,1,1) has a relatively similar value to ARIMA (1,1,2) yet based on the parsimony concept, ARIMA (1,1,2) fulfills more criteria to be the best model for the Kedah state. Table 2c shows the statistics value for Melaka. Model parsimony is also considered for Melaka, with the lowest value of AIC and BIC for ARMA (2,1). The DW for ARMA (2,1) is also near 2 with a value of 1.9373. While Table 2d shows the statistics value for Negeri Sembilan. After first differencing, five ARIMA models for Negeri Sembilan are analyzed to perform model validation. Based on Table 2d, ARIMA (2,1,1) is the best to analyze solid waste generation in Negeri Sembilan with the lowest AIC and BIC value. Table 2e shows the statistics value for Pahang. As Pahang does not need differencing, thus ARMA model is suitable for this state. However, both ARMA (1,1) and ARMA (1,2) give quite similar results. Thus, based on the parsimony concept, ARMA (1,1) is better to be selected as the best model. Table 2f shows the statistics value for Perlis. Since the data is not stationary, thus differencing is needed in model identification. Based on Table 2f, ARIMA (1,1,2) is selected as the best model to analyze solid waste generation for this state. This is because ARIMA (1,1,2) obtained the lowest value of AIC and BIC, with a DW value nearest to 2. Table 2g shows the statistics value for WPKL. Since WPKL is not stationary when performing unit root test at level, differencing needs to be done. After the first differencing, five ARIMA models are assumed and examined to perform validation and diagnostic tests. Based on Table 2g, ARIMA (2,1,1) turned out to be the best-fit model by applying the model parsimony concept. Lastly, Table 3 shows the statistic value for all states combined. ARIMA (3,1,1) is chosen as the best model to analyze the solid waste generation of all states combined, with the lowest AIC and BIC value, and DW nearest to 2.

Table 2: Statistics value in the selected state in Malaysia

Statistics	Model				
	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (3,1,1)	ARIMA (3,1,2)	ARIMA (2,1,2)
AIC	20.0986	20.0869	20.1028	20.4869	20.4784
BIC	20.2042	20.1925	20.2085	20.5926	20.584
DW	1.9607	1.931	2.0636	2.8564	2.874

(a). Johor

Statistics	Model				
	ARIMA (1,1,1)	ARIMA (1,1,2)	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (3,1,2)
AIC	19.1867	19.1317	19.1424	19.2589	19.2709
BIC	19.2924	19.2373	19.2481	19.3646	19.3765
DW	1.928	2.0196	1.973	2.6343	2.6975

(b). Kedah

Statistics	Model				
	ARMA (1,1)	ARMA (1,2)	ARMA (2,1)	ARMA (2,2)	ARMA (3,2)
AIC	18.1751	18.1657	18.1639	18.2436	18.1903
BIC	18.3148	18.3053	18.3035	18.3833	18.323
DW	1.9785	1.9365	1.9373	1.435	1.4247

(c). Melaka

Statistics	Model				
	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (3,1,1)	ARIMA (3,1,2)
AIC	17.8114	17.7871	18.3373	17.8097	18.3348
BIC	17.917	17.8927	18.4429	17.9154	18.4405
DW	1.9423	1.8748	2.9217	1.8896	2.8733

(d). Negeri Sembilan

Statistics	Model				
	ARMA (1,1)	ARMA (1,2)	ARMA (2,1)	ARMA (2,2)	ARMA (3,2)
AIC	17.8687	17.8708	17.8705	18.0198	18.0518
BIC	18.0083	18.0105	18.0102	18.1594	18.1914
DW	1.9503	1.9707	1.9649	1.2161	1.1857

(e). Pahang

Statistics	Model				
	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (2,1,3)	ARIMA (1,1,2)	ARIMA (1,1,1)
AIC	15.2316	15.3048	15.3375	15.1902	15.2618
BIC	15.3372	15.4105	15.4432	15.2959	15.3674
DW	1.927	2.6253	2.517	2.024	1.923

(f). Perlis

Statistics	Model				
	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (1,1,2)	ARIMA (1,1,3)	ARIMA (2,1,2)
AIC	19.2735	19.2715	19.2716	19.326	19.6031
BIC	19.3792	19.3771	19.3772	19.4316	19.7088
DW	1.9785	1.9801	1.9898	2.1869	2.8579

(g). WPKL

Table 3: Statistics value for all states combined

Statistics	Model				
	ARIMA (1,1,1)	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (3,1,2)	ARIMA (3,1,1)
AIC	22.0657	22.0735	22.3633	22.3544	22.0519
BIC	22.1714	22.1792	22.469	22.46	22.1576
DW	2.0213	1.8613	2.7809	2.8305	1.9994

3.5. Forecasting using ARMA and ARIMA models

The forecasted graph using ARMA and ARIMA models of respective states are shown in Fig. 4a until Fig. 4h. Referring to Fig. 4a for Johor State, after first differencing, five models were suggested to be analyzed and ARIMA (2,1,1) is selected as the best-fit model. With a maximum landfill capacity of 42638 tons per month, Johor landfills are expected to be enough to accommodate the solid waste that will be generated. However, as more than half of the landfills are occupied, thus all of the landfills in Johor must be observed from time to time. Based on Fig. 4b, the best model to forecast the solid waste generated is ARIMA (1,1,2) as it needs differentiation. The maximum landfill capacity for Kedah is 24,104.4 tons per month. This shows good signs as the forecast values are decreasing gently from January 2021 to December 2022. However, the solid waste generated in Kedah has totally overflowed the landfill capacity. Overflowed waste that keeps piling up is dangerous and may lead to serious climate change.

Next, with no differencing, ARMA (2,1) is chosen to be the best-fit model for Melaka and is used for model application as shown in Fig. 4c. There is only one landfill in Melaka which has a maximum landfill capacity of 851.77 tons per day, which is approximately 25553 tons per month. According to Fig. 4c, the solid waste generated overflowed the landfill capacity sometime in February 2020. The estimated waste is roughly measured at around 19000 to 20000 tons per month, which is still below the maximum capacity of the landfill in Melaka. ARIMA (2,1,1) is chosen to forecast solid waste generation in Negeri Sembilan, Fig. 4d shows the actual and predicted values from January 2019 to December 2022. Observing the actual values and predicted values of solid waste generated, clearly shows the actual values have already exceeded the maximum landfill capacity. Therefore, the landfill capacity might be inadequate to hold the volume of solid waste generated in the future as the forecast values are increasing towards the end of 2022. The best model to forecast solid waste generated for Pahang is ARMA (1,1) as the model does not need differentiation. The graph shown in Fig. 4e presents the predicted values from January 2021 to December 2022. It is seen that the estimated values of solid waste generated exceeded the maximum landfill capacity with an amount of 22,065 tons per month. In the future, the landfill capacity no longer can hold the solid waste generated. The authorities should find other areas to hold solid waste generation as it already overflowed the landfill capacity in Pahang.

While on Fig. 4f for Perlis illustrates the predicted values for model ARIMA (1,1,2) as it is the best model to forecast solid waste generated. The solid waste generated approaches the maximum capacity which is 4,441.8 tons per month. The forecast values from January 2021 to December 2022 show the decreasing volume of solid waste generated as it is assumed from the effect of the pandemic COVID-19. However, the maximum landfill capacity might not be able to keep the solid waste generated in the future as it already comprised three-quarters of landfill capacity.

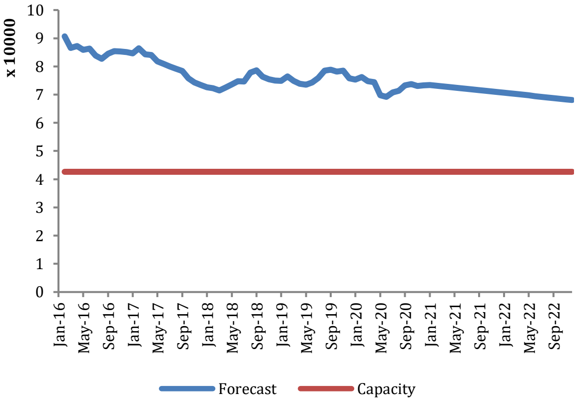
Fig. 4g shows the actual solid waste generated in KL and the forecasted value using the best-fit model, ARIMA (2,1,1). Kuala Lumpur's transfer station of 5.25 hectares has a maximum landfill capacity of approximately 2350.54 tons per day. This is equivalent to 70516 tons per month. The actual and predicted solid waste generated does not overflow the landfill capacity, however, strict monitoring must be done since there is not much room left until the maximum landfill capacity is reached. Combining all states in Fig. 4h, ARIMA (3,1,1) is selected based on the lowest MSE as the best model. It is worrying that the actual versus forecasted solid waste generated line graph of all states shows that the maximum landfill capacity of 214003 tons per month has been reached and is overflowing. Analysis of forecasting values for all stated states using ARIMA and ARMA models generates different results. The good signs for forecast values can be shown in Kuala Lumpur, Melaka, and Perlis. For Negeri Sembilan, the forecast values keep increasing towards the end of 2022 and it will eventually reach the landfill capacity. Meanwhile, solid waste generation for Johor, Kedah, and Pahang already comprised all the landfill capacity. Table 4 shows the forecasted value versus landfill capacity of solid waste generation for all states. It can be concluded that most landfills in Malaysia are not enough to cater to all the solid waste produced by the citizens. Hence, immediate actions must be done. Comparing Table 4 with the result from Mohamad et al. (2022) shows that the results are parallel, and the value of solid waste keeps on increasing. However, Mohamad et al. (2022) only forecast an area and did not investigate the importance of landfills capacity that accommodates solid waste generation in the future.

4. Conclusions

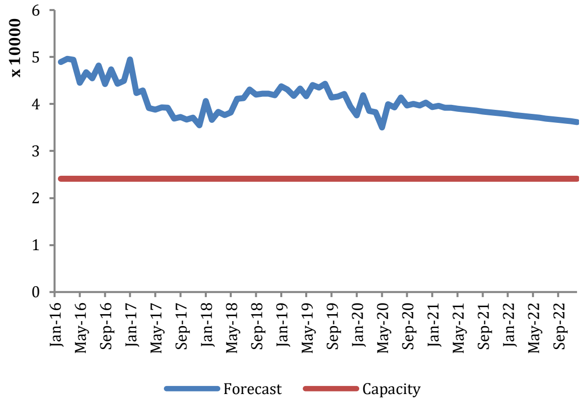
Due to the rapid growth of the population and economy, solid waste generation keeps on increasing and requires special attention. This study aims to predict the total solid waste generated in the future years and provide insight information among the authorized bodies as well as the society. The results of this study show that the solid waste generated in some of the states exceeds the maximum landfill capacity. However, in some of the studied states, WPKL, Melaka, and Perlis, show promising signs of stabilizing in terms of forecast value.

The findings also conclude that the overflowed solid waste generated in Negeri Sembilan, Johor, Kedah, and Pahang is worrying as it can cause a negative impact on the environment as well as people's health. Therefore, it is proposed that new areas must be developed quickly to ensure that solid waste generation is not being scattered or dispersed all over the place. Because the landfill capacity will reach the maximum capacity due to the overflow of solid waste generation, all people including authorities and citizens must take this issue thoughtfully. The authorities need to ensure that solid waste generation is not overflowed as it will be detrimental to the environment.

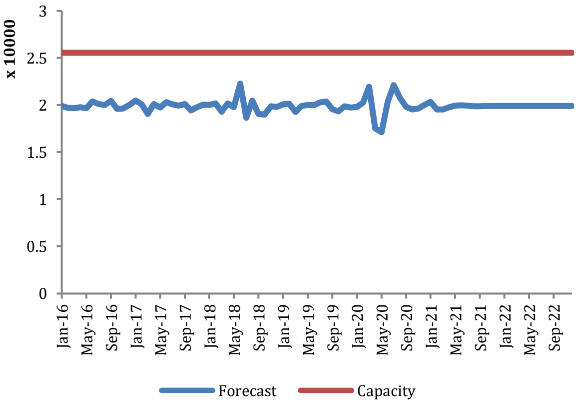
In Malaysia, the standard waste management hierarchy consists of five key steps: Reuse, reduce, recycle, treat, and dispose of. Malaysia's practice can be described as effective and convincing. Based on the Malaysian Investment Development Authority in 2020, the recycling rate increased from 5% in 2005 to 17% in 2018.



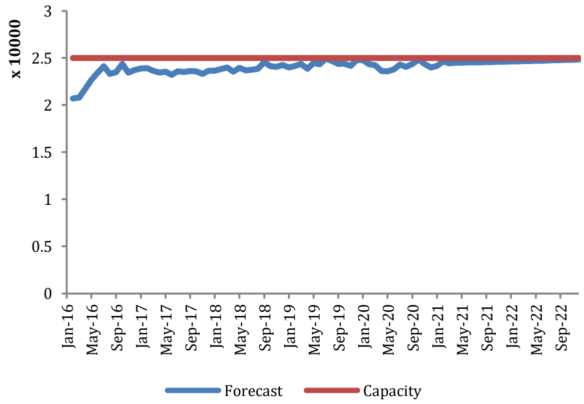
(a). Johor



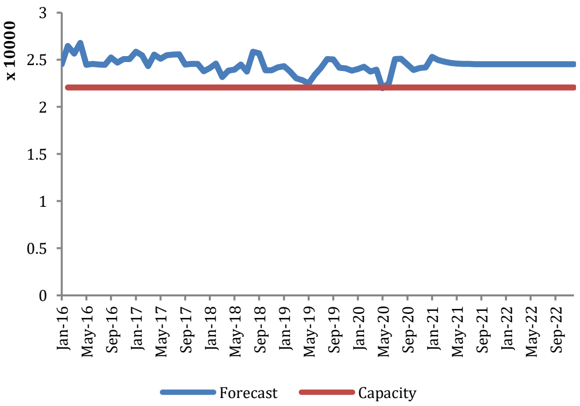
(b). Kedah



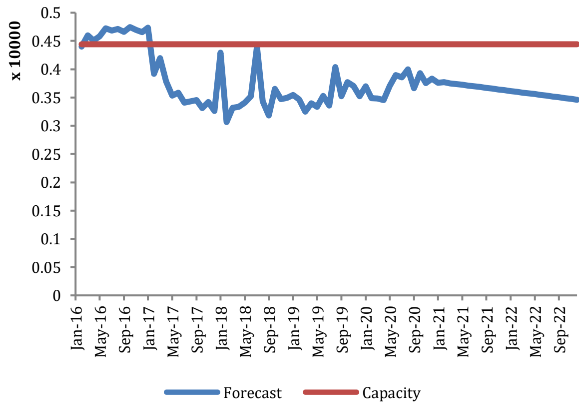
(c). Melaka



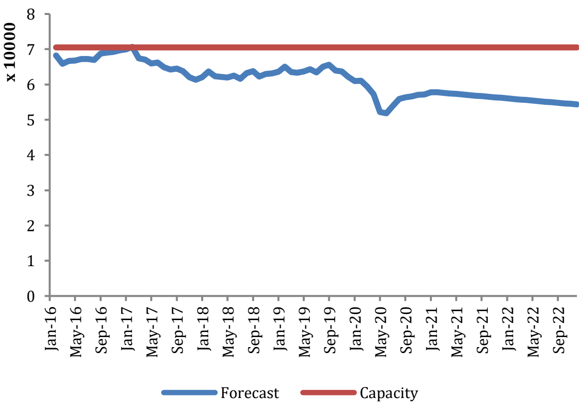
(d). Negeri Sembilan



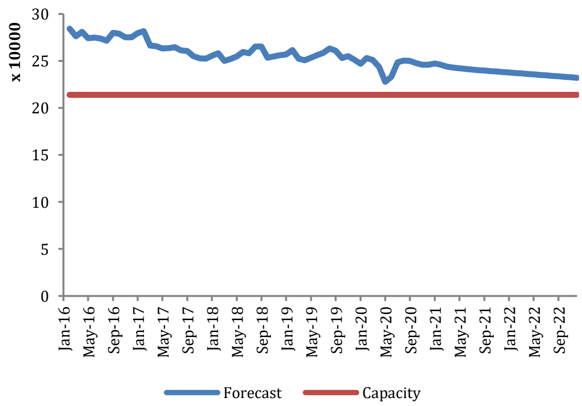
(e). Pahang



(f). Perlis



(g). WPKL



(h). All states combined

Fig. 4: Estimated solid waste generation in the selected state in Malaysia (in tons)

Table 4: Forecast value versus maximum landfill capacity of all states combined (in tons)

Month	Forecast value	Maximum landfill capacity	Overflowed solid waste	Percentage overflowed solid waste (%)
Jan-21	247,494.66	214,003.12	33,491.54	15.65
Feb-21	246,087.69	214,003.12	32,084.57	14.99
Mar-21	243,990.21	214,003.12	29,987.09	14.01
Apr-21	243,048.48	214,003.12	29,045.36	13.57
May-21	242,146.10	214,003.12	28,142.98	13.15
Jun-21	241,494.32	214,003.12	27,491.20	12.85
Jul-21	240,883.64	214,003.12	26,880.52	12.56
Aug-21	240,330.21	214,003.12	26,327.09	12.3
Sep-21	239,792.57	214,003.12	25,789.45	12.05
Oct-21	239,268.78	214,003.12	25,265.66	11.81
Nov-21	238,750.01	214,003.12	24,746.89	11.56
Dec-21	238,234.73	214,003.12	24,231.61	11.32
Jan-22	237,720.94	214,003.12	23,717.82	11.08
Feb-22	237,208.06	214,003.12	23,204.94	10.84
Mar-22	236,695.60	214,003.12	22,692.48	10.6
Apr-22	236,183.38	214,003.12	22,180.26	10.36
May-22	235,671.28	214,003.12	21,668.16	10.13
Jun-22	235,159.24	214,003.12	21,156.12	9.89
Jul-22	234,647.24	214,003.12	20,644.12	9.65
Aug-22	234,135.25	214,003.12	20,132.13	9.41
Sep-22	233,623.28	214,003.12	19,620.16	9.17
Oct-22	233,111.30	214,003.12	19,108.18	8.93
Nov-22	232,599.33	214,003.12	18,596.21	8.69
Dec-22	232,087.37	214,003.12	18,084.25	8.45
Jan-23	231,575.40	214,003.12	17,572.28	8.21
Feb-23	231,063.43	214,003.12	17,060.31	7.97
Mar-23	230,551.46	214,003.12	16,548.34	7.73
Apr-23	230,039.50	214,003.12	16,036.38	7.49
May-23	229,527.53	214,003.12	15,524.41	7.25
Jun-23	229,015.56	214,003.12	15,012.44	7.02
Jul-23	228,503.59	214,003.12	14,500.47	6.78
Aug-23	227,991.63	214,003.12	13,988.51	6.54
Sep-23	227,479.66	214,003.12	13,476.54	6.3
Oct-23	226,967.69	214,003.12	12,964.57	6.06
Nov-23	226,455.73	214,003.12	12,452.61	5.82
Dec-23	225,943.76	214,003.12	11,940.64	5.58

The Malaysian government, by encouraging people to implement reuse and recycle activities, can preserve the use of landfill capacity. One way to overcome this issue is by increasing recycling activities in which the authorities can prepare more recycling bins to encourage people to separate the type of solid waste generation meticulously. The increase in waste sorting can reduce the mixing of recycled items hence saving up the use of landfills.

Besides, campaigns or awareness on solid waste management must be performed constantly to ensure all citizens remember the importance of solid waste management and practice the right method to dispose of waste. Above all, people play a major part in controlling the amount of waste generation. These findings are very helpful in assisting authorities of respective states to prepare upcoming plans or programs in order to ensure landfill capacity in Malaysia is sufficient to hold upcoming solid waste generation. Besides, solid waste management must be handled efficiently to provide a clean and healthy environment for future generations. In order to handle the waste efficiently, a thorough and systematic observation of solid waste generation should be monitored frequently.

Acknowledgment

This research was supported by Universiti Teknologi MARA (UiTM) and funded under UiTM internal. Grant no. 600-RMC/GPM LPHD 5/3 (064/2021). This support is gratefully acknowledged. The authors would like to thank other researchers, lecturers, and friends for their ideas and discussion. A special appreciation to the College of Computing, Informatics, and Media, Universiti Teknologi MARA for supporting the publication of this paper.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Arzo A, Naznin S, and Moloy MDJ (2021). Modeling and forecasting of time series data using different techniques. Journal Multicultural Education, 7(11): 474-482. https://doi.org/10.5281/zenodo.5730512

Asadullah M, Bashir A, and Aleemi AR (2021). Forecasting exchange rates: An empirical application to Pakistani rupee. The Journal of Asian Finance, Economics and Business, 8(4): 339-347. https://doi.org/10.13106/jafeb.2021.vol8.no4.0339

Ayakeme TI, Biu OE, Enegesele D, and Wonu N (2021). Forecasting of Bayelsa state internally generated revenue using ARIMA model and winter methods. International Journal of Statistics and Applied Mathematics, 6(1): 107-116.

Ceylan Z, Bulkan S, and Elevli S (2020). Prediction of medical waste generation using SVR, GM (1, 1) and ARIMA models: A case study for megacity Istanbul. Journal of Environmental Health Science and Engineering, 18(2): 687-697. https://doi.org/10.1007/s40201-020-00495-8 **PMid:33312594 PMCid:PMC7721841**

Chen Y and Dai F (2020). Integrating SVR and ARIMA Approach to build the municipal solid waste generation prediction system. Journal of Computers, 31(3): 216-225.

Chintalapudi N, Battineni G, and Amenta F (2020). COVID-19 virus outbreak forecasting of registered and recovered cases after

sixty day lockdown in Italy: A data driven model approach. Journal of Microbiology, Immunology and Infection, 53(3): 396-403.
https://doi.org/10.1016/j.jmii.2020.04.004
PMid:32305271 PMCid:PMC7152918

Emetere ME and Iroham CO (2021). Computational forecast of municipal waste in Lagos: What may happen in 2025? In the IOP Conference Series: Materials Science and Engineering. IOP Publishing, Bristol, UK: 012014.
https://doi.org/10.1088/1757-899X/1036/1/012014

Fattah J, Ezzine L, Aman Z, El Moussami H, and Lachhab A (2018). Forecasting of demand using ARIMA model. International Journal of Engineering Business Management, 10: 1-9.
https://doi.org/10.1177/1847979018808673

Ferronato N and Torretta V (2019). Waste mismanagement in developing countries: A review of global issues. International Journal of Environmental Research and Public Health, 16(6): 1060.
https://doi.org/10.3390/ijerph16061060
PMid:30909625 PMCid:PMC6466021

Hyndman RJ (2015). Measuring forecast accuracy. In: Gilliland M, Tashman L, and Sglavo U (Eds.), Business forecasting: Practical problems and solutions: 177-184. John Wiley & Sons, Hoboken, USA.

Kim H (2022). A finite sample correction for the panel Durbin–Watson test. Applied Economics, 54(28): 3197-3205.
https://doi.org/10.1080/00036846.2020.1869172

Kumar S and Kumar R (2021). Forecasting of municipal solid waste generation using non-linear autoregressive (NAR) neural models. Waste Management, 121: 206-214.
https://doi.org/10.1016/j.wasman.2020.12.011
PMid:33360819

Mohamad NAJ, Yatim SRM, Abdullah S, Azmin MT, and Alwi N (2022). Forecasting municipal solid waste (MSW) generation in Klang, Selangor using artificial neural network (ANN). Malaysian Journal of Medicine and Health Sciences, 18(8): 151-158.

Mohamed IE (2008). Time series analysis using SAS-part I-the augmented Dickey-Fuller (ADF) test. In the SAS Conference Proceedings, Pittsburgh, USA.

Niu D, Wu F, Dai S, He S, and Wu B (2021). Detection of long-term effect in forecasting municipal solid waste using a long short-term memory neural network. Journal of Cleaner Production, 290: 125187. https://doi.org/10.1016/j.jclepro.2020.125187

Schwarz G (1978). Estimating the dimension of a model. The Annals of Statistics, 6(2): 461-464.
https://doi.org/10.1214/aos/1176344136

Sharifah NSI and Latifah AM (2013). The challenge of future landfill: A case study of Malaysia. Journal of Toxicology and Environmental Health Sciences, 5(6): 86-96.
https://doi.org/10.5897/JTEHS12.058

Siddiqua A, Hahladakis JN, and Al-Attiya WAK (2022). An overview of the environmental pollution and health effects associated with waste landfilling and open dumping. Environmental Science and Pollution Research, 29: 58514-58536.
https://doi.org/10.1007/s11356-022-21578-z
PMid:35778661 PMCid:PMC9399006

Sriploy S and Lertpocasombut K (2020). Industrial wastes to wastes disposal management by using box Jenkins-ARIMA models and created applications: Case study of four waste transport and disposal service providers in Thailand. EnvironmentAsia, 13(1): 124-139.
https://doi.org/10.14456/ea.2020.12

Wu F, Niu D, Dai S, and Wu B (2020). New insights into regional differences of the predictions of municipal solid waste generation rates using artificial neural networks. Waste Management, 107: 182-190.
https://doi.org/10.1016/j.wasman.2020.04.015
PMid:32299033

Zafra C, Ángel Y, and Torres E (2017). ARIMA analysis of the effect of land surface coverage on PM10 concentrations in a high-altitude megacity. Atmospheric Pollution Research, 8(4): 660-668. https://doi.org/10.1016/j.apr.2017.01.002

Zulkipli F, Jamian NH, and Zulkifli IZ (2020). Forecasting model for organic waste generation at administration Cafe in UITM Tapah Campus. International Journal of Academic Research in Business and Social Sciences, 10(9): 1023-1032.
https://doi.org/10.6007/IJARBSS/v10-i9/7982