Scholars Journal of Physics, Mathematics and Statistics OPEN ACCESS
Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com

Artificial Neural Network (ANN) and Arima Models for Better Forecast of the Air Pollution Data in Malaysia

Bashir Ahmed Albashir Abdulali[1, 2*], Nurulkamal Masseran[1]

[1]School of Mathematical Sciences, UKM, Bangi, 43600, Malaysia
[2]Department of Statististic, Misurata University, 2478, Libya

DOI: 10.36347/sjpms.2021.v08i10.001 | **Received:** 06.11.2021 | **Accepted:** 09.12.2021 | **Published:** 12.12.2021

***Corresponding author:** Bashir Ahmed Albashir Abdulali

Abstract Original Research Article

The latest trend of air pollution and variables influencing the air quality in Malaysia are studied in this research since there have been changes recently. Living conditions and health have been negatively impacted by air pollutants. An important method utilised nowadays is time series modelling, which is able to forecast events in the future. In this research, forecasting used one-year hourly Air Pollution Index (API) information originating from a station in Klang, Malaysia. The API values were predicted via the Artificial Neural Network model (ANN) and Autoregressive Integrated Moving Average model (ARIMA). Each of the approach's performance was assessed via the root means square error (RMSE), mean square error (MSE), and mean absolute error (MAE). The outcomes highlight the fact that compared to ARIMA, the ANN provided the lowest forecasting error to predict API. As such, the ANN may be regarded as a reliable predictive method to generate data for the general public regarding the status of air quality at a particular time.

Keywords: Air Pollution, API, ARIMA, ANN, Malaysia.

1. INTRODUCTION

(Breuer & Bower, 1999) Air pollution simply refers to air contamination via harmful gases (for instance sulfur, nitrogen and oxides of carbon) and smoking (STERN *et al.*, 1973). According to (Gurjar *et al.*, 2010) suggested that air pollution includes biological or physicochemical materials in the atmosphere that may be harmful or cause uneasiness to live organisms or deteriorate the environment. Monitoring of air pollution is a suitable manner of comprehending air pollution issues, evaluating and reviewing environmental procedures in order to appropriately address air pollution issues at all levels (Breuer & Bower, 1999).

1.1 Air pollution in Malaysia

It is undeniable that air pollution is among the main problem affecting human health nowadays; therefore, understanding their characteristics can help to improve the air quality monitoring system and the health system. Beginning from 1980, Malaysia experienced annual episodes of haze, in which the worst took place in 1997. As such, the Malaysian government set up the Haze Action Plan, Malaysia Air Quality Index (MAQI), and the Air Pollution Index (API) to improve air quality. In Malaysia, the three primary sources of air pollution are open burning, stationary and mobile sources. Recent years show that emissions originating from mobile sources such as vehicles produced a severe amount of air pollution - approximately 70 to 75% of the total air pollution. On the contrary, emissions from stationary sources make up 20 to 25% of the air pollution, whereas forest fires and open burning make up 3 to 5% of the air pollution (Department of the Environment, 1996, 2001).

1.2 The Air Pollution Trend in Klang, Malaysia

(Abdullah, Armi, Samah, & Jun, 2012; Afroz, Hassan, & Akma, 2003; Rahman *et al.*, 2015). The Klang Valley comprises of Putrajaya, Kuala Lumpur and adjoining towns of Selangor for instance Gombak, Shah Alam, Sepang, Petaling Jaya and Klang. Its population is approximately 3.98 million people, as of the year 2000. The severe air pollution within Klang Valley has contributed to rising cases of respiratory diseases. The pollutants affecting Klang Valley are attributed to various sources that include industrial and commercial developments. In most developing nations, the main source of air pollution in urban regions are motor vehicles. Other sources include power plants, industrial waste incinerators, quarries, urban construction projects, and open burning. Coincidentally,

back in 1997, Klang Valley was among the severely impacted areas due to Malaysia's widespread forest fires originating from Indonesia. These are boundary haze; the Klang valley's fast transformation into a large urban region has contributed to several environmental problems and significant pollution throughout the last decade (Abdullah, Armi, Samah, & Jun, 2012; Afroz, Hassan, & Akma, 2003; Rahman *et al*., 2015).

1.3 Air Pollutant Index (API)

The air pollution index (API) is a general approach in describing the quality of air in the environment. It is based on the greatest average value of individual indices for all variables namely ozone (O3), nitrogen dioxide (NO2), sulfur dioxide (SO2), suspended particulate matter (PM10) and carbon monoxide (CO) at a specific time. These characteristics can be explained using the information of statistical models. The seasonal variation is also influencing the concentration of pollution[9]. As illustrated in the following table, the API reference value was recorded following the Malaysia Ambient Air Quality Index (MAQI) of 1989:

Table 1: API values for air quality status in Malaysia

API	Diagnosis
0 – 50	Good
51 – 100	Moderate
101–200	Unhealthy
201- 300	Very Unhealthy
301 -500	Hazardous

Source: (Department of Environment, 2000)

The hourly average of API air pollution index for Klang Valley in the year 2014 is explored in this research. We have about 8760 hours of API air pollution index for Klang station. Besides, this study interpolates the time series modeling ARIMA and Artificial Neural Network Model (ANN) distribution pattern across Klang valley during the whole period of 2015 to determine the best prediction of these two models satisfying the following objectives; (1) investigate the suitability of the ARIMA Model on the air pollution data, (2) investigate the suitability of the Artificial Neural Network (ANN) on the air pollution data, (3) compare and find a better prediction model. In section 2, we review the statistical methods, which are both the ARIMA model and Artificial Neural Network are outlined, including the comparison between two models and a prediction for new hours. Section 3 will also include a review of the theory of the models and selection criteria between them, where section 4 discusses the main results from this study and assess its contribution to the literature.

2. LITERATURE REVIEW

2.1 Box-Jenkins (ARIMA) Model

(Pagan & Schwert, 1990) (Brockwell & Davis, 1987) (Hibon & Makridakis, 1997) The autoregressive (AR) models were first suggested by Yule in 1926 (Hibon & Makridakis, 1997). Then it was supplemented by (Slutzky, 1937), who presented the Moving Average (MA) schemes in 1937 and integrated both MA and AR schemes. All stationary time series can be modeled by the ARMA processes as long as the amount of AR terms, q (the amount of MA terms), and the appropriate order of p were specified. Box and Jenkins suggested such approach and termed it as the Box-Jenkins method to ARIMA models, in which the letter "I", set between MA and AR, represented the word "Integrated". The popularity of the Box-Jenkins methodology and ARIMA models among academicians rose during the 1970s (Brockwell & Davis, 1987). Evidence of the ARIMA model's superiority compared to non-parametric and Markov switching models were found by (Pagan & Schwert, 1990). The Box-Jenkins time-series analyses (ARIMA models) were utilised by (Chavez *et al*., 1999) to model and forecast future consumption and production of energy in Asturias. A stochastic autoregressive integrated moving average ARIMA model was also utilised by (Slini *et al.*, 2002) in Greece for maximum ozone concentration forecasting.

2.2 Artificial Neural Network Model

(Gardner & Dorling, 1998) The neuron's first model was proposed in 1943 by the neurophysiologist Warren McCulloch and a mathematician, Walter Pitts (Piccinini, 2004). They created a model that had two inputs and a single output (Mehrotra et al., 1997). Artificial Neural Networks are utilised in a wide array of applications nowadays such as pattern recognition and image processing. It has an outstanding performance and advanced mathematical computation power especially in its flexibility adapting to parallelism technique. The Artificial Neural Networks Model has been utilised to estimate and classify matters related to Medicine, Health, Manufacturing, General Applications, Accounting and Finance, Engineering and Marketing (Gardner & Dorling, 1998).

2.3 A comparison of Neural Network with ARIMA Models in Predicting Events

A comparison for prediction ability using Artificial Neural Network (ANN) and ARIMA models for time series predictions revealed that the ANN performed better than ARIMA model for prediction of the direction of stock movement as the former can pinpoint hidden patterns within the information utilized (Hansen *et al*., 1999). An investigation was also carried out in Stockholm (Kolehmainen *et al*., 2001) to ascertain the distinct neural network approaches that have potential utility and study them with regression with periodic components based on an hourly time series of NO2. The neural network methods had a much better forecast. A hybrid approach was utilised by (Zhang, 2003) that integrates ANN models and ARIMA models for non-linear and linear modeling to produce better outcomes with the hybrid method. The Malaysian

economy's gross domestic product (GDP) was projected by (Junoh, 2004) to utilise data based economic indicators via ANN. The econometric approach highlighted the fact that ANN fared better in the GDP forecasts. On the other hand, a comparison of the variable autoregressive models and neural network model was made by (Düzgün, 2008). The outcomes of forecasting performances in Turkish GDP show that the outcomes from ARIMA models were more favourable than those of the ANN model.

The ANN and ARIMA models were utilised by (Hilovska, 2010) to obtain the estimation for aggregate water consumption. The ANN has superior performance in estimating nonlinear time series while the ARIMA model is better in terms of estimating linear time series. The performance of ARIMA and ANN models for stock price estimation was discussed by (Comrie, 2016). Compared to the traditional ARIMA model, the ANN model's performance was more predictive in nature. Three approaches were utilised to estimate values of API: the fuzzy time series, autoregressive integrated moving average (ARIMA) and Artificial neural network (ANN). The outcomes show that the ANN generated the least forecasting error in forecasting API, in comparison to the ARIMA and fuzzy time series (Haizum *et al*., 2013).

The number of newly infected Covid-19 cases daily for the following 30 days was predicted via application of the ARIMA and ANN models. Both algorithms predicted a surge in the number of patients who were newly infected. Models attested to the fact that the ARIMA's projection was more accurate than the ANN (Moftakhar *et al*., 2020).

Based on the consumer price index data by ANN and Box-Jenkins (ARIMA) models, the information in the following period was estimated by (Işığıçok *et al*., 2020). The predictive performance of both techniques were compared and studied. Outcomes from both approaches almost resembled one another.

3. METHODOLOGY

3.1 ARIMA Model

Building the ARIMA forecasting model requires three important steps that must be given consideration namely parameter estimation, tentative identification and diagnostic checking. In an ARIMA model, future values of a variable are presumed to be a linear function of numerous previous observations and random errors which are independently and identically distributed. The generated time series will have the following structure:

$$y_t = \theta_0 + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad \ldots\ldots\ldots (1)$$

(Hibon & Makridakis, 1997) Where y_t and ϵ_t are the actual data and random errors at the time of period t; Φ_i where (i = 1,2,...,p) and θ_j where (j = 0,1,2,...,q) are model parameters, whereas p and q are integers that are usually noted as the order of the model. Random errors, ϵ_t are presumed to be identically and independently distributed (iid) with a 0 mean and constant variance of σ^2 then we can compute the autocorrelation (ACF) and partial autocorrelation coefficients (PACF) for the residuals of the trend model. The PACF and ACF are the only useful indicators of the order p and q of the ARIMA (p,q,d) model if the series that they are computed is stationary (Hibon & Makridakis, 1997).

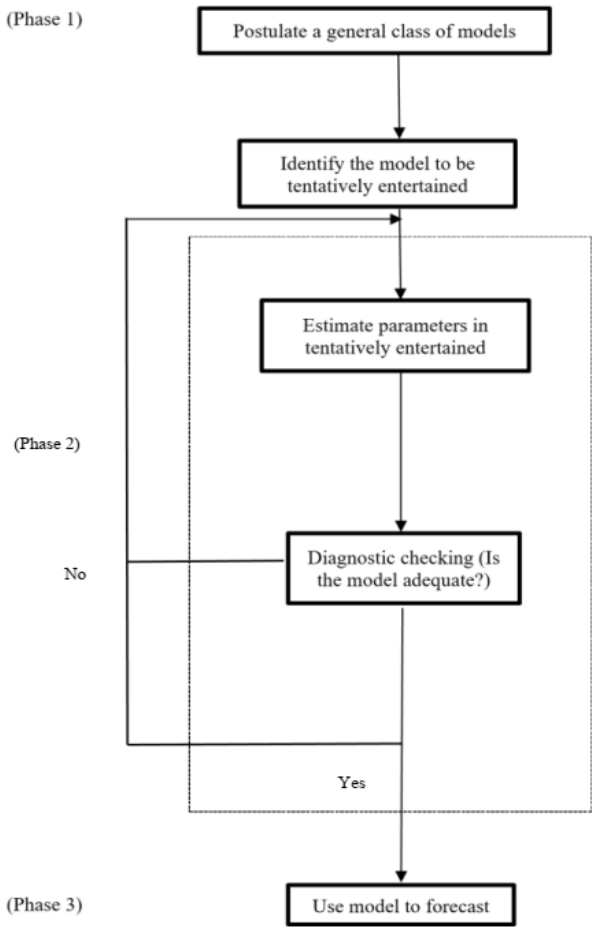

Fig 1: The Diagram of ARIMA Model

These are the process involved in the three-step model-building. It is a typical process, which could be repeated a few times till a suitable model is found. Then, the selected model may be utilised for the purpose of predicting.

1.1 Artificial Neural Network (ANN) Model

Before entering the training process, the structure of the neural network must be defined since many different neural network structures are available. The best structure of a neural network relies on the type of data and system to be modeled. Moreover, learning mechanisms and activation rules are useful in determining the structure of the neural network. One of the well-known structures is the backpropagation

mechanism based on a multiplayer neural network model. A typical ANN architecture has three distinct interconnected layers namely (i) hidden layers, (ii) an input layer and (iii) an output layer. The layers are connected by neurons, and a numeric weight value represents the strength of each connection. This numeric weight value which corresponds to the decision boundary is predicted utilizing different optimization algorithms on the dataset for training involved in the prediction tasks. When the predicted values are stabilized after validation, the trained ANN is tested against a dataset to measure its predicting power. Most commercial backpropagation tools offer the most influence on training time and neural network performance. The following Equation gives the output value for a unit:

$y = f(h_j)$ …….. (2)

$=f(\sum_i^n w_{ij}\, x_i\, \theta_j)$ …….. (3)

$$= \begin{cases} 0, w_{ij}\, x_i < \theta \\ 1\, w_{ij}\, x_i \geq \theta \end{cases} (i = 1,2\ldots,n) \text{ …….. (4)}$$

Whereby the output value y is computed from a set of input patterns, x_i of i^{th} unit in a previous layer, w_{ij} is the weight on the connection from the neuron i^{th} to unit j, θ_j is the threshold value of the threshold function f, and the number of units in the previous layer is n. In general, the logistic or sigmoid function is utilised as an activation function. If h_j, is the net input to unit j, subsequently output (θ_j) of unit j *is* measured as:

$\theta_j = \frac{1}{1+e^{-h_j}}$ …….. (5)

Given a large sample of training data, the Multilayer Feedforward Neural Network (MLFFNN) model can perform nonlinear regression and provide an estimation for the weight values. After validation, when the estimated values are stabilized, the data set can be predicted by applying trained MLFFNN. The Backpropagation algorithm is utilised as a training algorithm of MLFFNN for estimating the parameters to capture the nonlinear pattern of a set of inputs and outputs for forecast modeling (Blake, 2004; Chaudhuri *et al*., 2016).

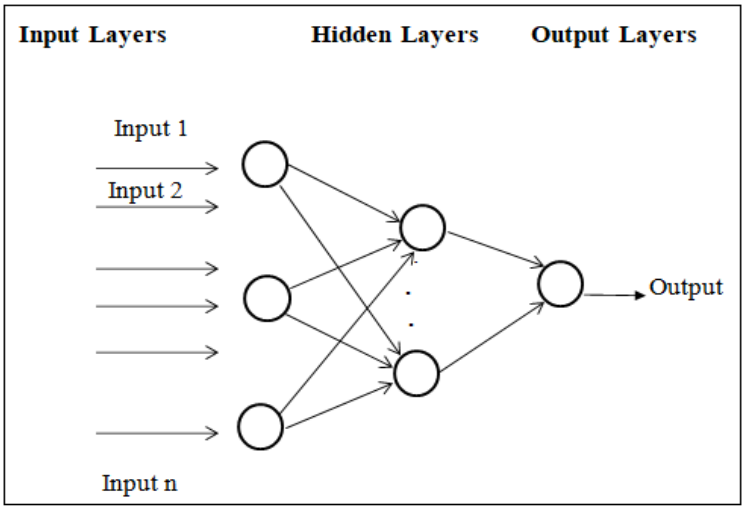

Fig 2: Diagram of ANN Model

The input layer contains concurrently fed input units. After that, weighted inputs are fed into the hidden layer. The hidden layers weighted outputs are the input to the output layer and represent the network's prediction. This network topology is known as Multilayer Feed-Forward network (MLFFN) since it has at least three distinct layers and all the weights do not cycle back to an input unit or a previous layer's output unit. The network output can then be reverse transformed back into the original target data units when the networks are used in the field (Blake, 2004).

3.3 Comparison of ANN and ARIMA Findings

In this study, the criteria that are utilised to evaluate the accuracy of the forecast are the mean squared error (MSE), the root means squared error (RMSE) and mean absolute error (MAE) as listed below.

$$\boldsymbol{MSE} = \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N} \text{ …….. (6)}$$

$$MAE = \frac{\sum_{i=1}^{N}|x_i - \hat{x}_i|}{N} \text{ …….. (7)}$$

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}} \text{ …….. (8)}$$

4. RESULTS AND DISCUSSION

This preliminary analysis consists of the computation of the descriptive statistics to the data. The table below displays the descriptive statistics of the dataset.

Table 2: The descriptive statistics of the API dataset

-	N	Min	Max	1*st* QU	Median	Mean	3*st* QU	Variance	Std
API	8760	0	255	51	57	62.1	66	670.03	25.88

The Air Pollution Index (API)'s overall values for the Klang Valley are concluded by the Air Pollution Index (API), which is based mainly on five main pollutants namely CO, O3, PM10, NO2 and SO2 in the ambiance. Hourly measurements for PM10 and SO2 are averaged over 24 hours of the running period, with 8 hours for CO. Meanwhile, NO and O3 measurements are noted every hour before an hourly index is computed using sub-index functions for each pollutant. Next, the highest index value recorded is noted as the API for the hour. The following graph shows the time series of API hourly rate in 2015 collected from Klang station.

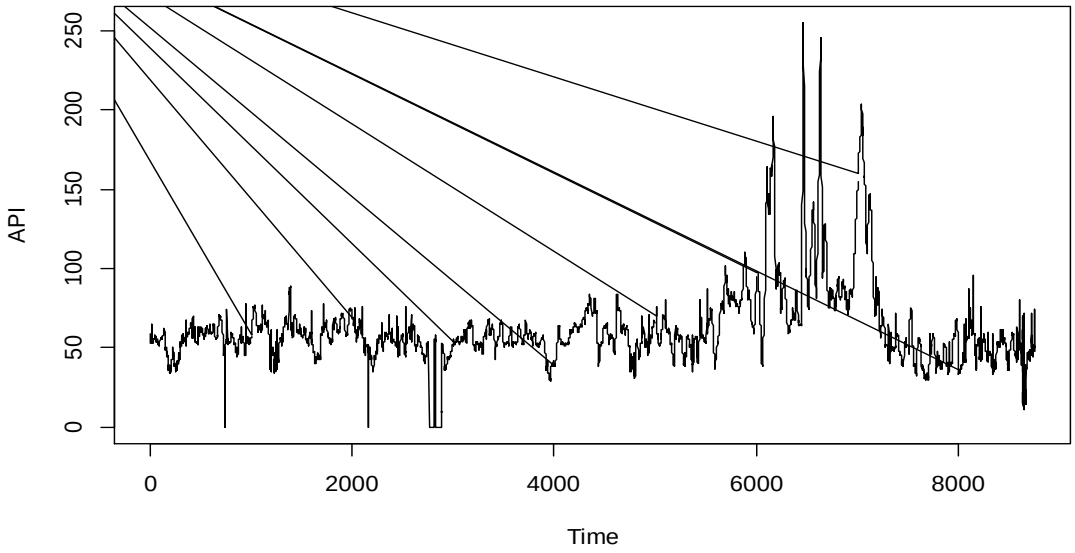

Fig 3: API hourly rate in 2015 collected from Klang station

4.1 ARIMA Model Output

4.1.1 Model Identification

The time series plot of the API index displayed in Figure 3 shows several violations in the dataset which may not be stationary. Thus, the Dickey-Fuller Test will be applied to the dataset.

Null Hypothesis: API has a unit root
Exogenous: Constant
Lag Length: 7 (Automatic - based on SIC, maxlag=40)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.021686	0.0000
Test critical values:	1% level	-3.430921	
	5% level	-2.861677	
	10% level	-2.566884	

*MacKinnon (1996) one-sided p-values.

Fig 4: Augmented Dickey-Fuller Test

From the output in Figure 4, we reject H_0 since the p-value is small. This implies that the time series of API data is stationary. However, we will look for the partial autocorrelation and autocorrelation functions of the air pollutant index (API) to confirm it's stationary.

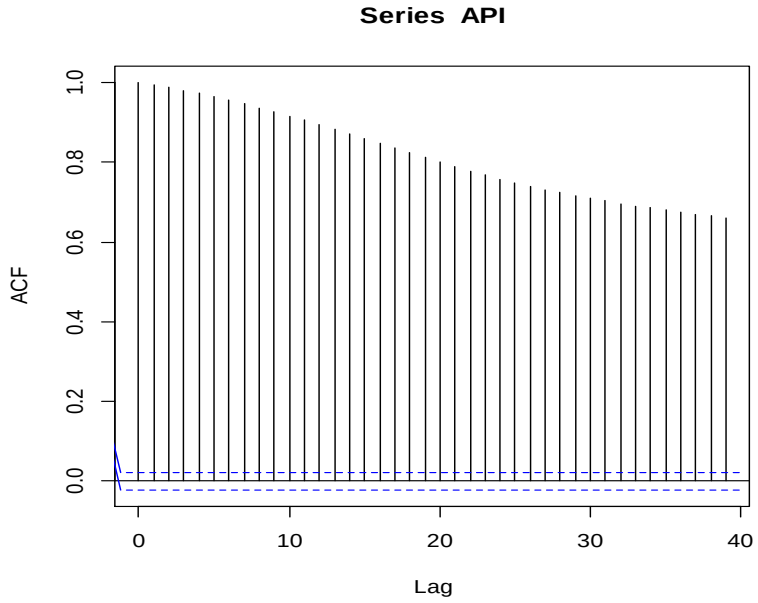

Fig 5: ACF plot for API dataset

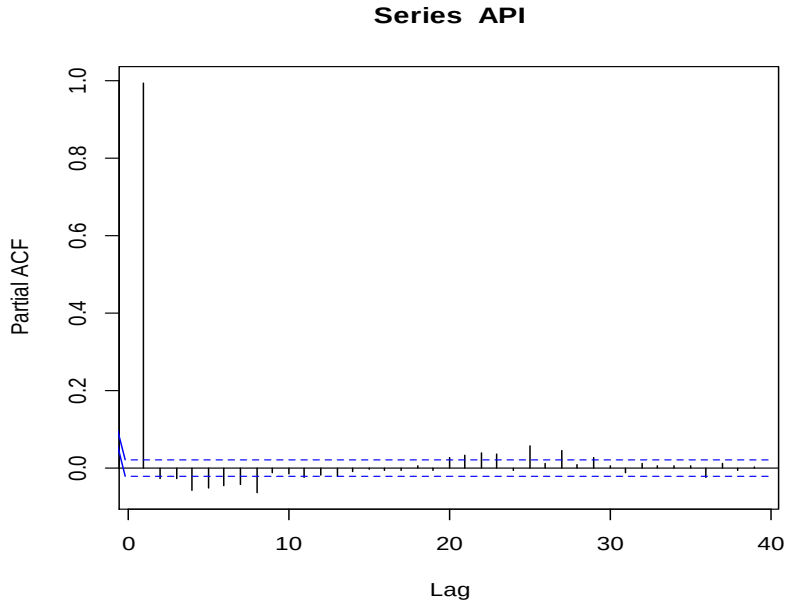

Fig 6: PACF plot for API dataset

The figures above show the partial autocorrelation and autocorrelation functions of the air pollutant index (API) for the period under consideration. The autocorrelation function dies down at a prolonged rate, confirming a mean trend, while the partial autocorrelation function truncates after the first lag.

4.1.2 Model Selection

The autocorrelation function tails o to zero from the correlograms above, while the partial autocorrelation function truncates after lag 1. This indicates an AR (1) model. Therefore, the actual model identified is ARIMA (1,0,0). The identified model might not be the best model for the data; consequently, it is compared with other probable models, and the best model will be chosen for the forecast. After several trials, it is clear that ARIMA (2,0,2) model was chosen when we made the comparison to other models based on the SSR (Sum Square of Residual), S.E of regression (Standard Error of regression), and AIC (Akaike Information Criterion) which has the least error value as well as the highest the R-Squared. Hence, the most statistically preferred model is ARIMA (2, 0, 2)

Sample (adjusted): 3 8760
Included observations: 8758 after adjustments
Convergence achieved after 16 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	62.11821	3.316162	18.73196	0.0000
AR(1)	1.891767	0.021674	87.28336	0.0000
AR(2)	-0.893187	0.021490	-41.56364	0.0000
MA(1)	-0.887424	0.023990	-36.99104	0.0000
MA(2)	0.042087	0.011548	3.644501	0.0003

R-squared	0.987904	Mean dependent var	62.10493
Adjusted R-squared	0.987899	S.D. dependent var	25.88794
S.E. of regression	2.847828	Akaike info criterion	4.931561
Sum squared resid	70987.91	Schwarz criterion	4.935602
Log likelihood	-21590.30	Hannan-Quinn criter.	4.932938
F-statistic	178721.9	Durbin-Watson stat	1.996407
Prob(F-statistic)	0.000000		

Fig 7: The Output for ARIMA (2,0,2)

4.1.3 Model Adequacy

The selected model is now diagnostically tested. Unfortunately, none of the models satisfied the diagnosed test due to the violation in the real dataset. We also tried the 1^{th} or 2^{nd} difference for the dataset, but none of the models are diagnostically test satisfied in the 1st or 2nd difference. Nevertheless, based on goodness of t criteria as shown in Table 4, we believe that ARIMA (2,0,2) is a good approximated model for our dataset.

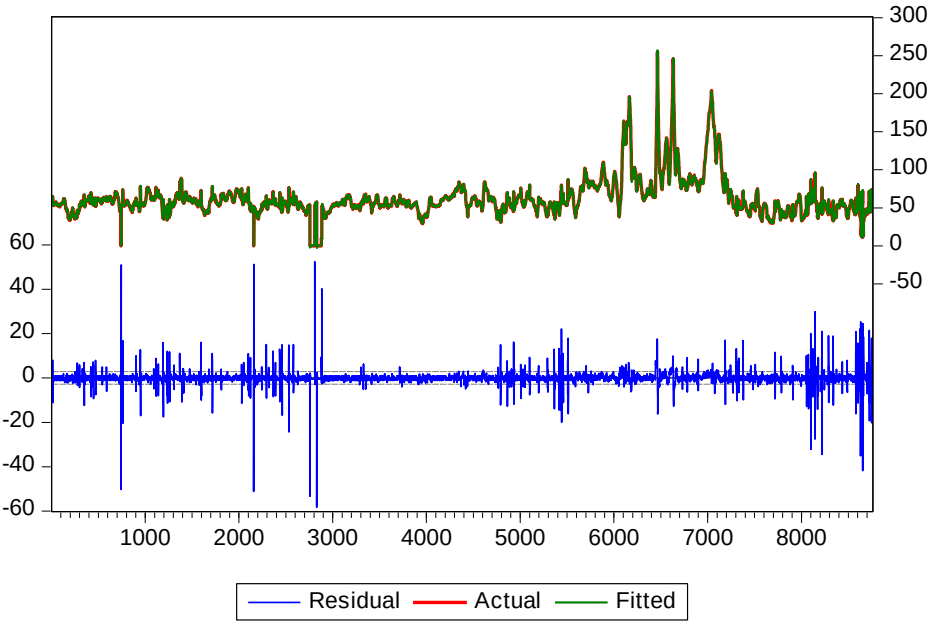

Fig 8: The plot for the actual, fitted, and residual model for ARIMA (2,0,2)

The plot shows the actual, fitted, and residual model of ARIMA (2,0,2). It has the lowest S.E of regression (Standard Error of regression) and AIC (Akaike Information Criterion). In the residual error plot, it could be said that the error is constant at most of the dataset, where the fitted values are very close to the actual observations. Thus, ARIMA (2,0,2) is the best t model and it can be used to forecast API to compare with the ANN model.

4.1.4 Parameter Estimate

In general: On the left-hand side

$(1 - AR(1)L - AR(2)L^2 - AR(3)L^3 - \cdots)[z_t - C]$ …… (9)

On the right-hand side

$(1 + MA(1)L + MA(2)L^2 + MA(3)L^3 + \cdots)$

…… (10)

Where $AR(1), AR(2), AR(3), \cdots$ and $MA(1), MA(2), MA(3), \cdots$ are coefficients from the E-

Views software. Equating the left and the right-hand sides, we can rearrange the estimated model to be:

$z_t = \delta + \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_1 a_{t-q}$ ……. (11)

The point is that in general $AR(1), AR(2), AR(3), \cdots$ are not $\phi_1, \phi_2, \phi_3, \cdots$ and $MA(1), MA(2), MA(3), \cdots$ are not $\theta_1, \theta_2, \theta_3, \cdots$.

For ARIMA (2,0,2)

Therefore, from Table 4 which contains the output for ARIMA (2,0,2) model, the estimation result corresponds to the following specification:

$62.11821 + (1 - 1.891767\mathrm{L} + 0.893187\mathrm{L}^2)\mathrm{u_t} = (1 - 0.887424\mathrm{L} + 0.042087\mathrm{L}^2)\mathrm{a_t}$ …….. (12)

Where L is the lag operator

- L(constant) = constant ……. (13)
- $\mathrm{L}(\mathrm{z_t}) = \mathrm{z_{t-1}}$ ……(14)
- $\mathrm{L}^2(\mathrm{z_t}) = \mathrm{z_{t-2}}$ ……. (15)
- $\mathrm{L}^{\mathrm{d}}(\mathrm{z_t}) = \mathrm{z_{t-d}}$ ……. (16)

Thus, the combination for both sides would be as the following:

$z_t = 1.891767 z_{t-1} - 0.893187 z_{t-2} + 0.0882 - a_t + 0.887424 a_{t-1} - 0.042087 a_{t-2}$ …….. (17)

4.1.5 FORECASTING FOR ARIMA MODEL

The obtained model is utilised in forecasting some future API values in Klang. The prediction was made by writing R code to forecast some future values.

Table 3: The forecasted values of API using ARIMA Model

Future Hours	Forecast
8761	51.07591
8762	51.15859
8763	51.24793
8764	51.34309
8765	51.44333
8766	51.54796
8767	51.65635
8768	51.76795
8769	51.88227
8770	51.99884

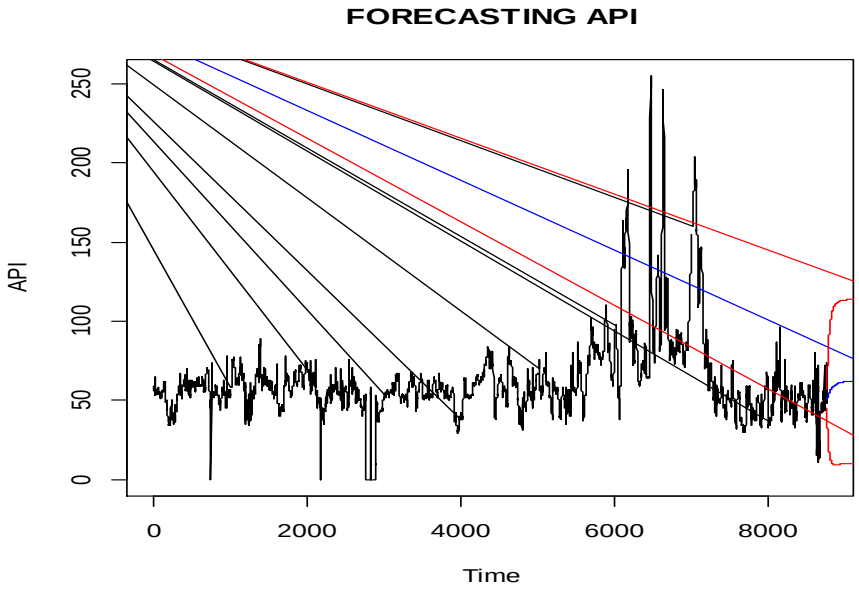

Fig 9: The predicted value of API with its interval using ARIMA (2,0,2)

4.2 Artificial Neural Network (ANN) Model

By using the nonlinear autoregressive neural networks in MATLAB software which can be trained to predict events from past values. The nonlinear autoregressive neural networks model is utilised in predicting future events of the air pollutant index. There are two series involved; the first is an input series x (t), and the second one is an output series y (t). To forecast the events of y(t) from previous values of x(t) but without knowledge of prior events of y(t) the input-output model is noted as the following:

$y(t) = f(x(t-1), \ldots, x(t-d))$ …… (18)

The Nonlinear autoregressive neural networks model offers better predictions than this input-output model since it utilised the additional information containing prior values of *y*(*t*). There are three kinds of targets time steps:

- Training
- Validation
- Testing

The following diagram of the ANN was utilised for 1000 hidden layers by MATLAB software. Training multiple times will

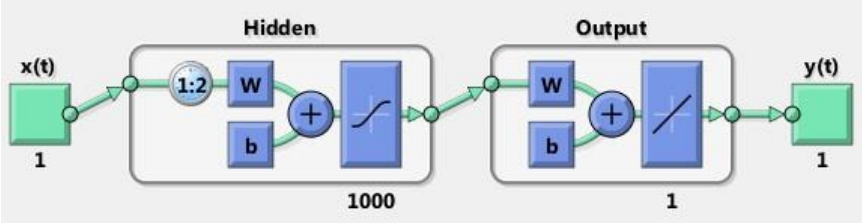

Fig 10: The Diagram of the Artificial neural network

The following diagram of the ANN was utilised for 1000 hidden layers by MATLAB software. Training multiple times will generate different results due to different sampling and initial conditions. After many trials, the best result that could be found is as the following:

Results

	Target Values	MSE	R
Training:	6132	9.26076e-4	9.55353e-1
Validation:	1314	8.38667e-4	9.53854e-1
Testing:	1314	7.77056e-4	9.60300e-1

Fig 11: The Output of the Artificial neural network (ANN) model

From Figure 11, we can note that the Mean Squared Error (MSE) is 8.38667e−4, which is a good indication for the ANN model. R values are proposed to measure the correlation between outputs and targets values, it is evident that R's value is very close to 1, which reveals a strong correlation between outputs and target values.

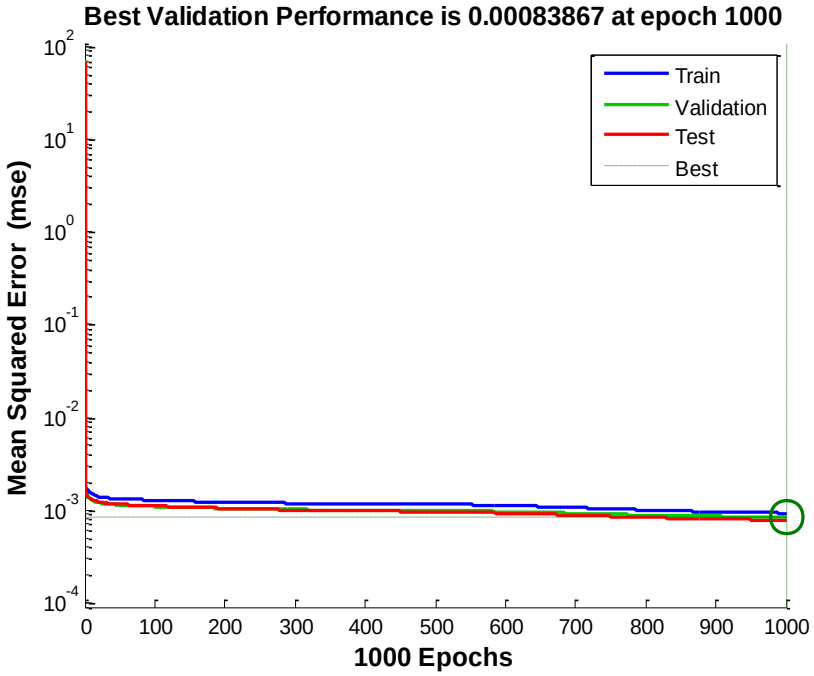

Fig 12: The Mean Square Error for Best Validation of ANN model

This figure revealed that validation, testing errors and training are reduced till iteration 1000. There is no overfitting because both testing and validation errors did not increase before iteration 1000. Training is carried out in an open loop which is also known as series-parallel architecture that includes the testing and validation steps.

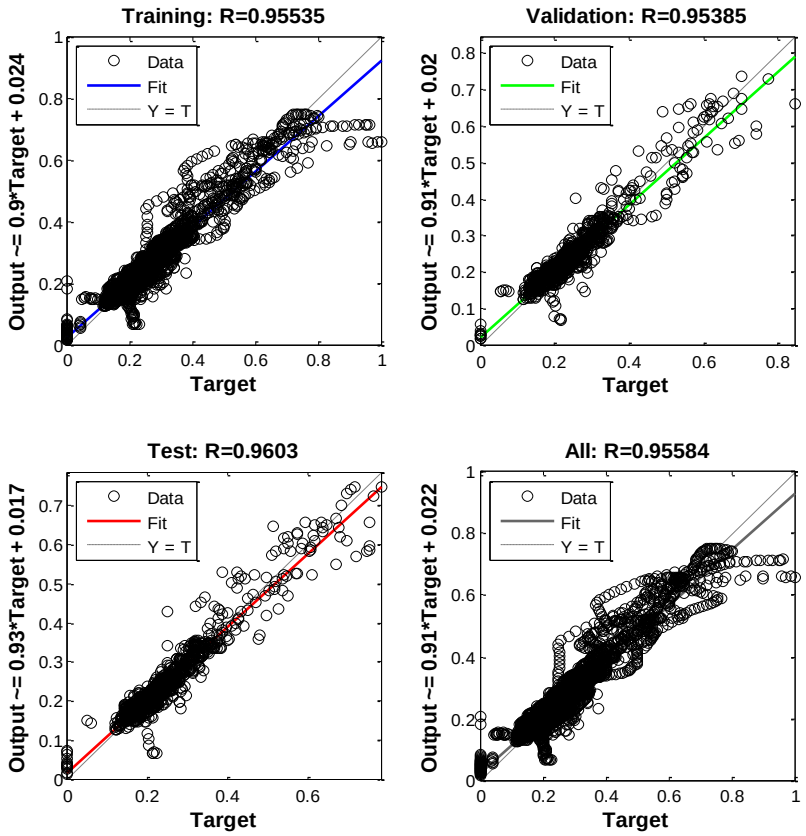

Fig 13: The R-Squared for ANN model

Figure 13 shows the correlations plots between the targets and outputs. It reveals that there are also strong correlations between the target values and outputs.

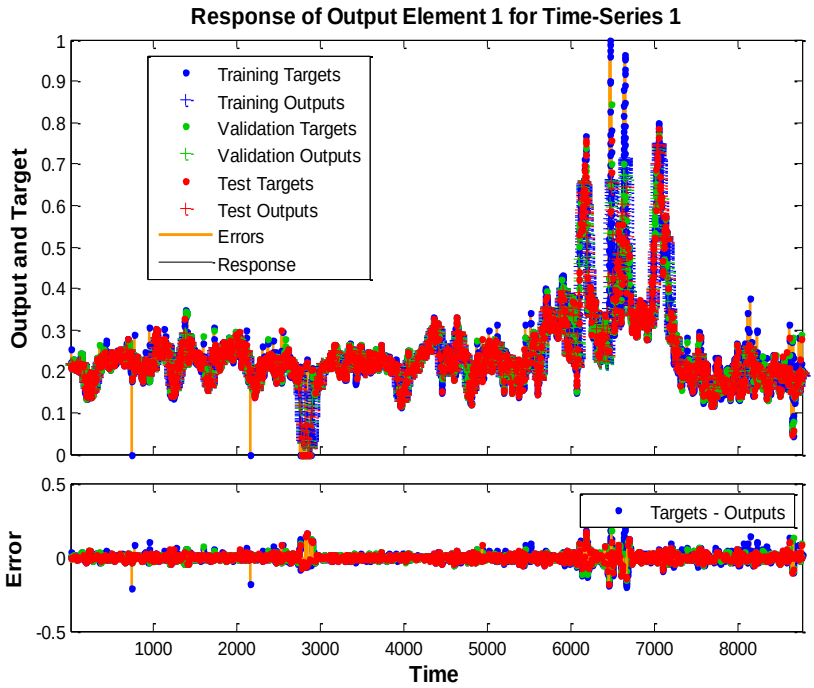

Fig 14: The response of output and the Error of the Targets for (ANN) model

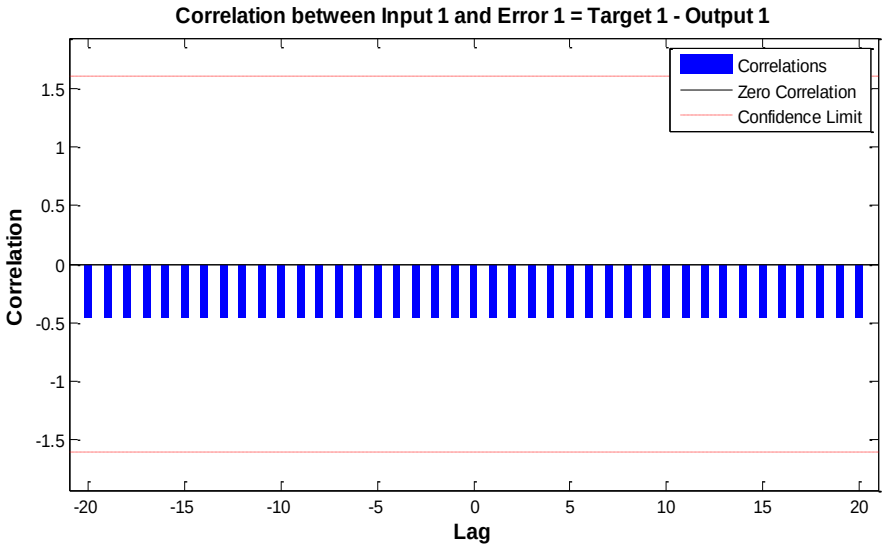

Fig 15: The correlation between Input and Error

The input-error cross-correlation function illustrates how the errors are correlated with the input sequence $y(t)$. All the correlation values must be zero to get the best prediction model. Figure 12 displays that all the correlation values range within the confidence bounds around 0.

4.2.1 FORECASTING FOR ANN MODEL

The obtained model is utilised to predict some future values of API in Klang Valley. The result was obtained by Matlab software to forecast some future values.

Table 4: The forecasted values of API for ANN Model

Future Hours	Forecast
8761	51.28815
8762	51.28815
8763	51.28560
8764	51.28560
8765	51.28560
8766	51.28305
8767	51.28050
8768	51.27795
8769	51.29580
8770	51.29580

4.3 Comparison of ANN and ARIMA Findings

Table 5: The Comparison of ANN and ARIMA Models

	ARIMA	ANN
MSE	8.105490	0.00083867
RMSE	2.847014	0.02896000
MAE	1.092132	0.04212000

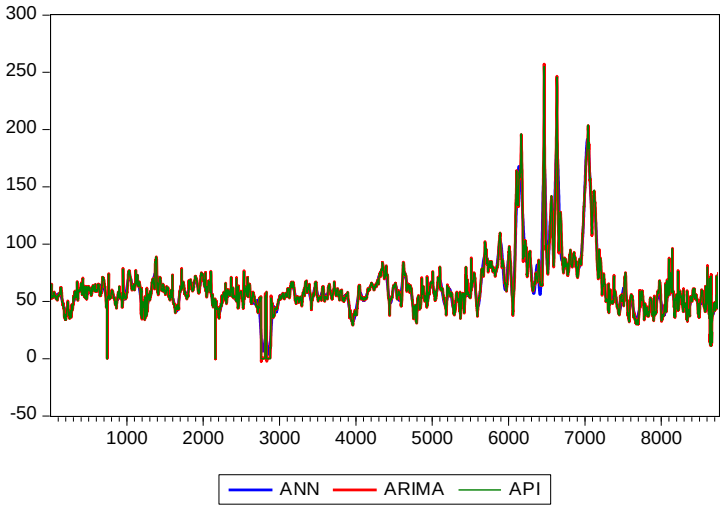

Fig 13: The Comparison of Actual API data, ANN and ARIMA Mode

Typically, the results based on residual error evaluation performance give similar results to the best model utilised for forecasting API. Nevertheless, the RMSE is the most common evaluation performance based on the difference between the forecasted and observed values with high sensitivity in extreme values due to the power term. Table 5 displays the ANN model performance testing data set utilised in forecasting each station based on the smallest RMSE values. The predicted and actual values of the time series plot for Klang are displayed in figure 16. The best result for forecasting API is provided by the ANN model.

5. CONCLUSIONS

In the past decade, time series analysis and forecasting have become actively researched areas. The accuracy of time series prediction is relying on the fundamental of decision processes. Thus, the research is improving in terms of the effectiveness of the forecasting models. Box and Jenkins' ARIMA model is one of the most widespread methods in the study of forecasting models. Lately, the ANN with its nonlinear modeling capabilities has revealed its potential in time series forecasting applications. Even though ANNs and ARIMA have flexibility in modeling various issues, both are not the best universal models which can be applied in every forecasting application. In general, the result shows that the ANN is a flexible intelligence forecasting method that provides an effective and useful tool for modeling poorly understandable and complex processes. Hence, the neural network is utilised as the benchmark for modern forecasting methods in creating novel techniques to enhance the accuracy in forecasting.

REFERENCES

- Abdullah, A. M., Armi, M., Samah, A., & Jun, T. Y. (2012). *An Overview of the Air Pollution Trend in Klang Valley, Malaysia*. 13–19.
- Afroz, R., Hassan, Ã. M. N., & Akma, N. (2003). *Review of air pollution and health impacts in Malaysia*. *92*, 71–77. https://doi.org/10.1016/S0013-9351(02)00059-2
- Breuer, D., & Bower, J. (1999). *Monitoring ambient air quality for health impact assessment* (Vol. 85). WHO Regional Office Europe.
- Brockwell, P. J., & Davis, R. A. (1987). *Time series: theory and methods*. Springer Science-Veriag Inc.
- Comrie, A. C. (2016). *Comparing Neural Networks and Regression Models for Ozone Forecasting Comparing Neural Networks and Regression Models for Ozone Forecasting*. *2247*(November). https://doi.org/10.1080/10473289.1997.10463925
- Department of the Environment, M. (1996). *Malaysia Environmental Quality Report* (M. of S. Department of Environment Technology and Environment, Malaysia, Ed.).
- Department of the Environment, M. (2001). *Clean Air Regional Workshop—Fighting Urban Air Pollution: From Plan to Action.* (M. of S. Department of Environment Technology and Environment Malaysia, Ed.).
- Düzgün, Dr. R. (2008). A Comparison Of Artificial Neural Networks ' And Arima Models ' Success In GDP Forecast. *Marmara Üniversitesi İ.İ.B.F. Dergisi YIL*, 165–176.
- Gardner, M. W., & Dorling, S. R. (1998). *Artificial Neural Networks (The Multilayer Perceptron)— A Review Of Applications In The Atmospheric Sciences*, *32*(14), 2627–2636.
- Gurjar, B. R., Molina, L. T., & Ojha, C. S. P. (2010). *Air pollution: health and environmental impacts*. CRC press.
- Haizum, N., Rahman, A., Hisyam, M., & Talib, M. (2013). *Jurnal Teknologi Forecasting of Air Pollution Index with Artificial Neural Network*. *2*, 59–64.
- Hansen, J. V, McDonald, J. B., & Nelson, R. D. (1999). Time Series Prediction With Genetic-Algorithm Designed Neural Networks: An Empirical Comparison With Modern Statistical Models. *Computational Intelligence*, *15*(3), 171–184.
- Hibon, M., & Makridakis, S. (1997). *ARMA models and the Box–Jenkins methodology*.
- Sterba, J., & Hilovska, K. (2010). The implementation of hybrid ARIMA neural network prediction model for aggregate water consumption prediction. *Aplimat—Journal of Applied Mathematics*, *3*(3), 123-131.
- Işığıçok, E., Öz, R., & Tarkun, S. (2020). Forecasting and Technical Comparison of Inflation in Turkey With Box-Jenkins (ARIMA) Models and the Artificial Neural Network. *International Journal of Energy Optimization and Engineering (IJEOE)*, *9*(4), 84–103.
- Junoh, M. Z. H. M. (2004). Predicting GDP growth in Malaysia using knowledge-based economy indicators: a comparison between neural network and econometric approaches. *Sunway Academic Journal*, *1*, 39–50.
- Kolehmainen, M., Martikainen, H., & Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment*, *35*(5), 815–825.
- Mehrotra, K., Mohan, C. K., & Ranka, S. (1997). *Elements of artificial neural networks*. MIT press.
- Moftakhar, L., Seif, M., & Safe, M. S. (2020). Exponentially Increasing Trend of Infected Patients with COVID-19 in Iran: A Comparison of Neural Network and ARIMA Forecasting Models. In *Iran J Public Health* (Vol. 49). http://ijph.tums.ac.ir
- Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, *45*(1), 267–290.
- Piccinini, G. (2004). The first computational theory of mind and brain: A close look at McCulloch and Pitts's "logical calculus of ideas immanent in

nervous activity." *Synthese*, *141*(2), 175–215. https://doi.org/10.1023/B:SYNT.0000043018.52445.3e

- Rahman, S. R. A., Ismail, S. N. S., Raml, M. F., Latif, M. T., Abidin, E. Z., & Praveena, S. M. (2015). The Assessment of Ambient Air Pollution Trend in Klang Valley, Malaysia. *World Environment*, *5*(1), 1–11.
- Slini, T., Karatzas, K., & Moussiopoulos, N. (2002). Statistical analysis of environmental data as the basis of forecasting: an air quality application. *Science of the Total Environment*, *288*(3), 227–237.
- Slutzky, E. (1937). The Summation of Random Causes as the Source of Cyclic Processes. *Econometrica*, *5*(2), 105. https://doi.org/10.2307/1907241
- Stern, A. C., Wohlers, H. C., Boubel, R. W., & Lowry, W. P. (1973). *Fundamentals Of Air Pollution* (Fourth, Issue (1973)).
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159–175.