Forecast the Erbil International Airport Data Using Statistical Methods

Paree khan Abdulla Omer
Salahaddin University
College of Administration & Economics, Statistics and Informatics Department
Email: pareekhan.omer@su.edu.krd

Samyia Khalid Hasan
Salahaddin University
College of Administration & Economics, Statistics and Informatics Department
Email: saiya.hasan@su.edu.krd

DOI: 10.23918/ICABEP2023p15

Abstract:

This paper concentrates to use two advanced statistical methods (Box–Jenkins and Adaline Neural Network) to forecast future values using data from the past values of a variable of all people that they had taken Erbil Airport as a way to move to other countries. The real data set were collected and obtained from Erbil International Airport and The sample observation consisted of (72) month, from the Airport transfer process during a certain time period 2017-2022. For the purpose of choosing the best model fit for these two models, the values of three criteria measures Mean Square Error (MSE), Mean Absolute Deviation (MAD) and Akaike Information Criterion (AIC) have been obtained from the estimated models. Additionally, the results showed that although the parameter estimations of the two models are not directly comparable, the results of all two models are not similar. Software packages StataV.16, SPSS V.26 and Matlab (R2013a) were used to fit the models.

Keywords: Artificial Neural Network, Adaline Neural Network, Box-Jenkins methodology, ARIMA models.

1. Introduction

The purpose of this article is to compare forecasts of real-world data using a time series statistical method and an artificial neural network (ANN). This paper focuses on predicting future values of a variable based on its past behavior. Traditionally, this problem is resolved using statistical analysis; first, a time-series model is constructed, and then statistical prediction algorithms are applied to acquire future values. Time series modeling is a highly effective technique, but its construction requires the knowledge or discovery of initial conditions (Husek, 2007), (Meloun, Militký, 2004). Time-series data is widely used in the areas of statistics, economics, finance, and forecasting, as well as in domains that contain temporal information. Time-series analysis is a set of analytical methods which seek to detect and analyze the temporal component within a dataset, with the goal of uncovering significant statistics and trends.

Usually, this method includes using time-series models to effectively analyze the data and uncover its inherent properties. Time-series forecasting involves utilizing these models to make predictions

ISBN: 978-9922-9036-3-7

about future values by leveraging historical observations. (Chenhui, 2021).

There are several techniques and approaches for developing forecasting models. This study focuses solely on time series forecasting models, specifically the Auto Regressive Integrated Moving Average (ARIMA). Box and Jenkins have identified these models, it is possible to analyze and make predictions on any data set with ANNs in general. Our key objectives are to use time series neural networks and compare the results with conventional statistical approaches. In the methodology section where the author explained very well the approaches and stated that Artificial neural networks, one of the methods developed in recent years, have demonstrated their ability and efficacy, relative to conventional statistical methods, to predict and solve problems reliably and easily. The major difference between ANNs and classical models such as ARIMA is the absence of priori assumptions that are normally required for these trading applications (Box and Jenkins, 1976).

2. Time Series Analysis

Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set period of time rather than just recording the data points intermittently or randomly. Forecasting is the process of predicting future events based on past and present data. Time-series forecasting is a type of forecasting that predicts future events based on time-stamped data points. Time-series forecasting models are an essential tool for any organization or individual who wants to make informed decisions based on future events or trends. From stock market predictions to weather forecasting, time-series models help us to understand and forecast changes over time. A time series is a set of observations x_t, each one being recorded at a specific time t. A discrete time series is one in which the set T_0 of times at which observations are made is a discrete set, as is the case, for example, when observations are made at fixed time intervals. Continuous time series are obtained when observations are recorded continuously over some time interval, e.g., when $T_0 = [0, 1]$ (Brockwell and Davis, 2016).

2.1 Box-Jenkins methodology

The Box-Jenkins methodology is a strategy for identifying estimating and fore casting autoregressive integrated moving average models, The methodology consists of a four-step iterative cycle of:

- Model Identification; Identification by looking at the sample autocorrelations and the partial autocorrelations of the possible models.
- Model Estimation; Estimation by certain optimization methods of the unknown parameters. diagnostic checks on model adequacy; Testing the adequacy of the fitted model on model residuals by performing the normal probability map, ACF and PACF.
- Model forecasting stage: Prediction of future results on the basis of known data (Stencl, et.al, 2010).

2.2 Auto Regressive Integrated Moving Average Model

ARIMA stands for Auto-Regressive Integrated Moving Average. It is a well-known time-series model that is broadly used in statistics and econometrics. It is a generalization of the Autoregressive Moving Average (ARMA) model to overcome the disadvantage that the ARMA is only sufficient to stationary time-series data by introducing differencing into the model. ARIMA model is a form of regression analysis, and it is extensively used to understand the data and make future predictions based on historical values. (Chenhui, 2021) (Abd Saleh, 2013).

Autoregressive moving average (ARMA) models are part of the Box-Jenkins methodology for time series forecasting (Wedding, Cios, 1995).

- Autoregressive (AR): It refers to a model with a variable that regresses on its own lagged values. The number of lagged observations (i.e., the lag order) is indicated by p.
- Integrated (I): It is the time of differencing applied to the raw data that allows the time-series data to become stationary. It is also known as the degree of differencing and is indicated by d.
- Moving Average (MA): It indicates the dependency between an observation and a residual error from a moving average model applied to lagged observations. The order of the moving average is expressed as q.

ARIMA models have a wide range of applications. Javier Contreras et al. used ARIMA models to analyze the time-series data of hourly prices in the Spanish electricity markets and predict the next-day price with a reasonable average error of around 10%. (Chenhui, 2021)

The processes type are AR(p), MA(q), ARMA(p,q), ARIMA (p, d, q). The questions in mind we have are: how does one know whether it follows a purely AR process or a purely MA process or an ARMA process or an ARIMA process (Din, 2015):

$$AR_{(P)}: x_t = \sum_{i=1}^{p} \phi_i x_{t-1} + \varepsilon_t \quad , \quad MA_{(q)}: x_t = \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \quad \ldots(1)$$

$$AR_{(1)}: x_t = \phi x_{t-1} + \varepsilon_t \quad , \ which\ is\ a\ random\ walk\ for\ \phi = 1$$

$MA_{(1)}: x_t = \theta \varepsilon_{t-1} + \varepsilon_t$, where the residuals $\{\varepsilon_t\}$ are white noise distributed as:

$$\varepsilon_t \sim i.i.d. N(0, \sigma^2)$$

The stochastic process $\{\varepsilon_t\}$ is called white noise (WN), if at every moment, the random variable ε_t is normally distributed, with zero mean and constant variance, i. e. meet the conditions:

$$E(\varepsilon_t) = 0 \quad , \quad E(\varepsilon_t^2) = \sigma^2 \ ; \ Cov(\varepsilon_t, \varepsilon_{t+k}) = 0 \quad , \quad t \neq k \quad \ldots(2)$$

When using the delay operator (lag) $L(x_t) = x_{t-1}$, the processes can write (Din, 2015)::

$$AR(p): \left(1 - \sum_{i=1}^{p} \phi_i L^t\right) x_t = \varepsilon_t \ , i.e.\ \Phi(L) x_t = \varepsilon_t \quad , \quad AR(1): (1 - \phi L) x_t = \varepsilon_t \quad \ldots(3)$$

$$MA(q): x_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t \ , \qquad i.e.\ x_t = \Theta(L)\varepsilon_t \quad , \qquad MA(1): x_t = (1 + \theta L)\varepsilon_t \ldots (4)$$

Where:
ϕ , Φ = autoregressive parameters.
θ , Θ = moving average parameters

3. Adeline Neural Network

Adaline Neural Network It is a single linear unit Adaptive Linear neuron using a bipolar activation function (+1, -1). The delta rule is used for training to minimize the Mean Squared Error between the actual output and the target output. The weights and the bias are adjustable. In Fig. 1, the structure of Adaline can be seen. First, initialize weight, input data to calculate the net of the Adaline network, then apply the activation function to that output, compare with the original output if both are equal, then give the output an error back to the network and update the weight according to the error, which is calculated by the delta learning rule. Depending on equation (2)

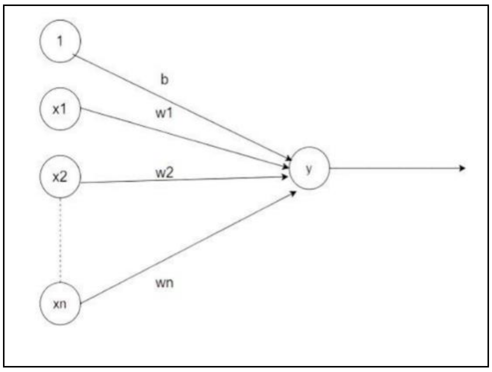


Fig.1: Structure of Adeline Network

The difference between the expected value and the forecast value is squared to determine the error in the Adaline in a single iteration, or $(y - \hat{y})^2$. For all of the scenarios with j = 1, 2,..., n in a given dataset, this procedure of comparing the expected and Forecast values is repeated.

3.1 Algorithm of Adaline Neural Network

Step 1: Initialize weight. Not zeros, but small random values are used. Set the learning rate α.

Step 2: While the stopping condition is false, do steps 3 to 7. Step 3: For each training set, perform steps 4 to 6.

Step 4: Set the activation of the input unit $x_i = s_i$ for $i = 1,2,..n$.

Step 5: Compute the net input to the output unit.

$$y_{im} = \sum_{i=1}^{n} w_i x_i + b \qquad \ldots (5)$$

Here, b is the bias, and n is the total number of neurons.

Step 6: Update the weights and bias for i = 1 to n.

$$w_i(new) = w_i(old) + (t\_y_{im})x_i \quad \dots (6)$$

$$b(new) = b(old) + (t\_y_{im}) \quad \dots (7)$$

Where wi, yim and t are the weight, forecast output, and true value respectively, and calculate

$$Error = (t - t\_y_{im})^2 \quad \dots (8)$$

When the forecast output and the true value are the same then the weight will not change.

Step 7**:** Test the stopping condition. The stopping condition may be when the weight changes at a low rate or no change.

3.2 Measurements of Selecting Model

a. Mean Squared Error (MSE): it is measures to forecast the error variance.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \quad \dots (9)$$

Where Fi : the forecasted values, n: is the sample size, Xi: the actual observation at time t.

b. Mean Absolute Deviation: it is the average distance of a dataset between each data point and the mean. It gives us an idea of the variability in a dataset.

$$MAD = \frac{\sum_{i=1}^{n}\lfloor x_i - \bar{x}\rfloor}{n} \quad \dots (10)$$

c. Akaike's Information Criterion (AIC): It is a general criterion in the construction and modeling of time series created by Japanese scientist Akaike, and it is employed in diagnosing the rankings of time series models. It used to compare the quality of a set of statistical models to each other. The following formula is used to compute the AIC value.

$$AIC_{(k)} = 2k - nln(\sigma_\varepsilon^2) \quad \dots (7)$$

Where n: is the sample size, σ_ε^2: Residuals variance. k : The number of estimated parameters.

4. Application Part

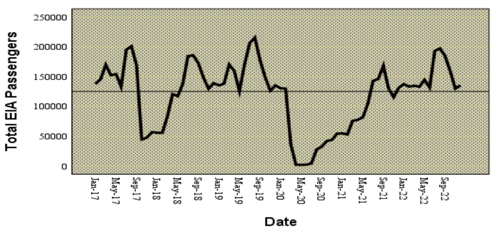
In this Part, we analyze actual data by using statistical methods to forecast the total passengers by two powerful techniques. Eight variables (EIA Pax aircraft Movement, Male arrivals, Male departures, Female arrivals, Female departures, Domestic passenger, Infant Passengers less than 2 years old and adult passengers older than 2 years old) can be defined as X_1, X_2 ,.., X_8. And sued three criterions for selecting the best model; Mean Square Error (MSE), Mean Absolute Deviation (MAD) and Akaike Information Criterion (AIC).

1.1 Data Collection

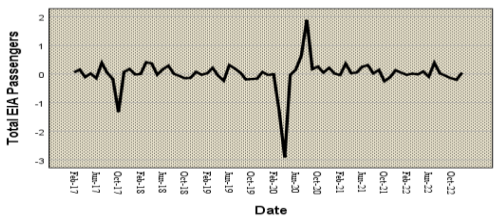
The data were used for time series analysis and ANN in this paper, considered monthly data and these are delivered as monthly totals. Often, the seasonal effect becomes much cleaner and easier to understand if we switch to the daily average per month rather than considering the monthly total. From simplified patterns, we humans as well as prediction models usually are more successful in extracting the relevant information. The numbers of international passenger bookings (in thousands) monthly on an airline in the Erbil were obtained from the Erbil International Airport for the period time 2017-2022. Where the sample dataset of this study was limited of (6) year's ago; beginning from 1st January 2017 through to 31st December 2022 of all people that they had taken Erbil International Airport as a way to move to other countries. The sample observation consisted of (72) month, total international Passengers is a main variable under study (dependent variable) during period of six years was measured.

1.2 *Apply of Time Series analysis*

The most important means of visualization is the time series plot, where the data are plotted versus time/index.



a. Monthly original Observed value



b. the natural logarithm, difference (1)

Fig. 2: Graphic representation of the original and transformed time series.

Figure (2, a.) above shows number of total passengers distributed for seven years ago in Erbil airport; where the Y-axis represents the total EIA against the X- axis which is the time. The line has different behavior look at the (2017and 2020), this express that the air movement internationally demand for travelling did not show highly significant improvement as a new epicenter of COVID-19 emerged in several countries, leading to a re-imposition of travel restrictions in (2020). The struggle against ISIS, which began in 2017, has had a significant impact (affected) on Kurds' daily lives, particularly through its effects on the KRG economy, which leaded to stopped travelling on Erbil airport during period of time. as we saw in line of 2017.

Figure (2, a.) shows that the process is not stationary, and it is necessary to differentiate it first. we checked the stationary condition for the original data series before transformed data series, and the Autocorrelation Function (ACF) and the Partial Auto Correlation Function (PACF) also confirmed the series is not stationary look at the figure (3).

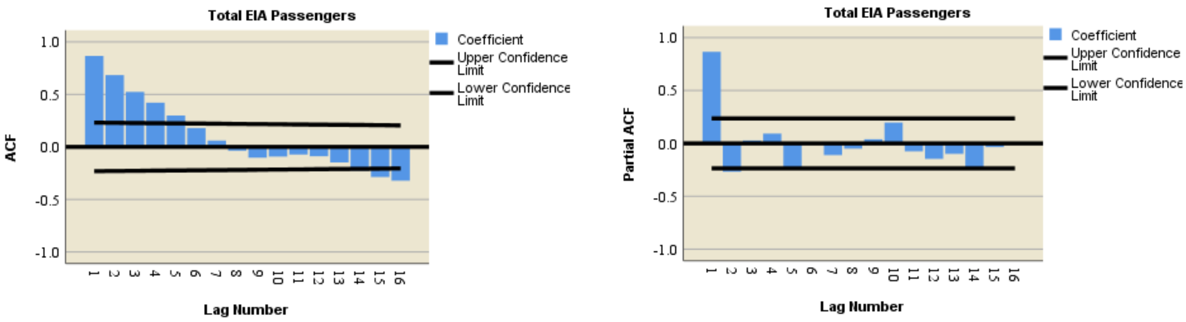


Fig. 3: Graphic representation of ACF and Partial ACF of the original time series data

Figure (3) gives a plot of the ACF and PACF of the original series, here we notice that the ACF does not tail off. the series is non-stationary and differencing will be needed. we suggest to apply some transformations, such as logarithms and one differentiating data series to ensure that the assumption of stationary of the ARIMA model is approximately satisfied as figure (2, b.), look at the ACF and PACF of the transformed data look at the figures (4) and Appendixes.

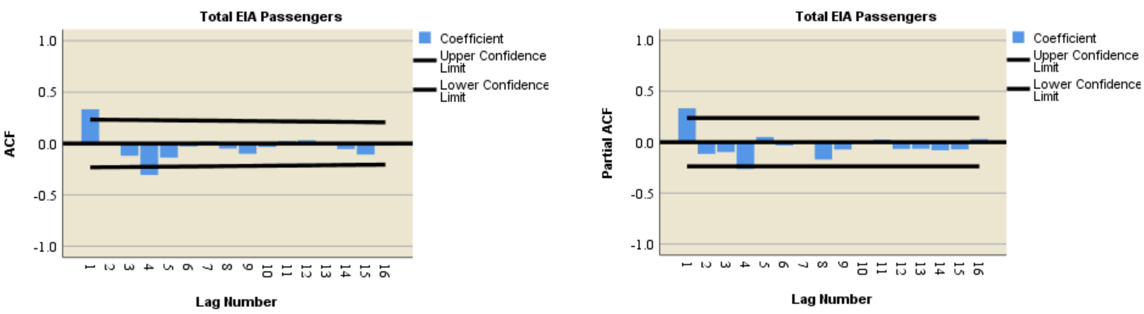


Fig. 4: Graphic representation of ACF and Partial ACF of the transformed time series data

The procedure of fitting is summarized by the following steps (Box and Jenkins, 1976):

1. Transform data using natural log transformation which was found the most appropriate.
2. Removing trend and fluctuations component by using the first order differencing and natural logarithm.
3. Model identification by plotting ACF and PACF of monthly observations.

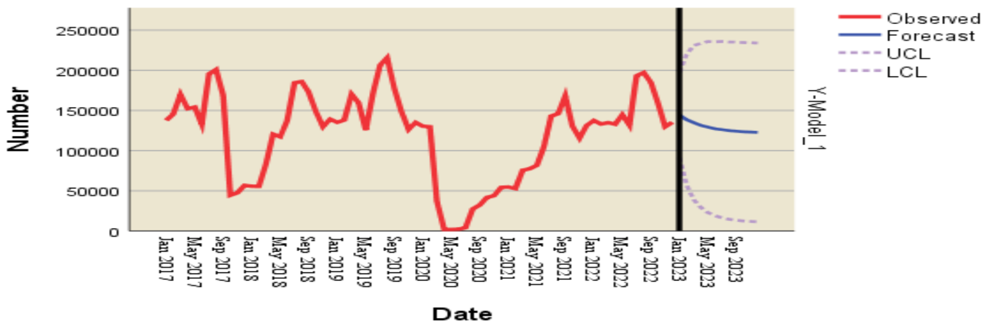


Fig. 5: ARIMA Forecast model of (2,0,2)

Graph above shows the red line represents values of the original time-series data and the blue line shows the time-series forecasting for the twelve months of (2023) with taking the difference and natural log to remove the component. The prediction was constructed for 12 months of year 2023 and confidence level was set to 95% for forecasting future values. The values computed using statistical analysis have to be compared to results of neural network experiments.

The MSE, MAE and AIC metrics for all models in this paper are listed in Table (1). It is obvious from the table that one model overall has the lowest MSE, MAE and AIC for the variable under the study (total Passengers), proving that it has the best performance when predicting the case trend (and also the case number) of total passengers. It predicts almost the exact values to the actual values, meaning that it can play a very important role in predicting the case trend in the future. Results of the analyses of forecasting the data are comparable in terms of prediction accuracy. Some of the expected differences were proofed. The first difference is that both statistical methods computed worse results, when compared to the neural networks results. The cause could be the type of the prediction, focusing on precise point prediction.

Table 1: Comparison of different forecasting statistical models

Models	MES	MAE	AIC
ARIMA (2,0,0)	**26977.7**	**17829.7**	**20.4889**
ARIMA (1,0,1)	**27056.4**	**17788.7**	**20.4947**
ARIMA (2,0,2)	**26545.1**	**17722.2**	**20.3121**
ARIMA (1,0,2)	**26993.4**	**18068.3**	**20.5178**
ARIMA (0.1.1)	**28563.7**	**18806.3**	**20.5198**

In table (1), We can conclude that the better result is the smallest values of these criterion of all *ARIMA* models suggestions. Obtained forecast result of the *ARIMA (2,0,2)* copying the real data more precisely that in the others *ARIMA* models. Future work of total passengers will focus on this model is much better than the others forecasting futures value. here, the method identifies an *ARIMA (2,0,2)* with drift term as the best fitting model and produces a forecast that is approximately linearly decreasing. Nevertheless, in practice, where we do not have intimate knowledge about the data generating process, careful modelling with *ARIMA* (and potentially adding drift terms) is important for producing successful forecasts and figure (6) below is the ACF and PACF for the residuals of the model ARIMA (2,0,2).

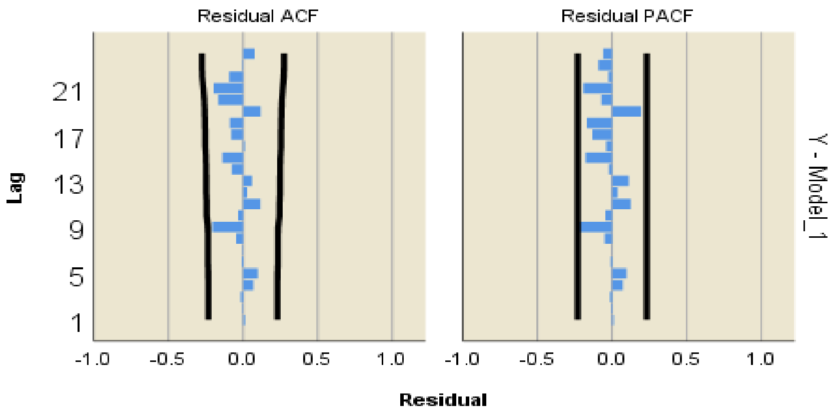


Fig. 6: ACF and PACF of the Residual Forecast of ARIMA (2,0,2)

Figure (6) shows the plot of ACF and PACF diagrams residual for Forecast values of ARIMA (2,0,2) model. The ACF and PACF should be considered together for the whole estimation period. The mean of the residuals is close to zero and there is no significant correlation in the residual's series.

5.3 Apply Adaline Neural Network

The structure of the Adaline network is initializing weight between 0.5, -0.5, learning rate (α) is equal to 0.1, 0.01and 0.005 input is 72, the number of hidden layer is 4 notes and the output there are 72 months of data sets; 50 data sets are selected as the training set, 11 for validation set and 11 data sets are used for the testing set. After normalizing by using Activation Function and depending on the single hidden layer and output.

Table 2: The Measurement of Adaline Network

Measurement Learning Rate	MSE	MAD	AIC
α=0.1	0.353	0.054	-6.916
α=0.01	0.0521	0.321	-12.93
α=0.005	0.0034	0.041	-16.631

in Table (2), we can see that the minimum values of MSE, MAD, and AIC with the difference in learning rate of training the network show that learning rate 0.005 is best than others depending on these minimum values of measurement.

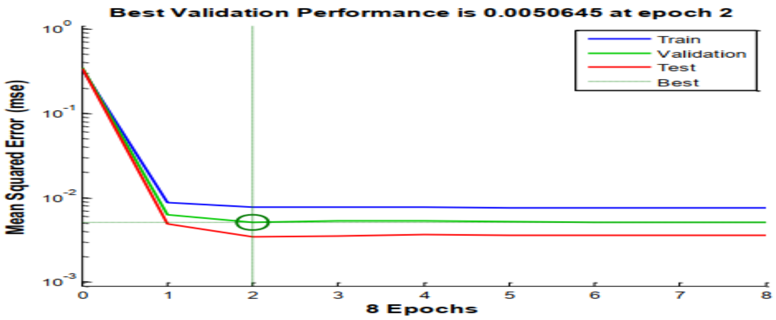


Fig. 7: The Performance of Adaline Network

Figure (7) shows the best performance of the Adeline network. The mean square error is divided by three sets: the training set, the validation set, and the testing line. The data used in this research consists of the best possible result. Consisted of 72 months: 70% for the training set (50 observations), 15% for the validation set (11) and 15% for validation. The validation set is equal to the best performance, the training set and test set are near, with the best stopping at iteration 2 of epochs 8, and the minimum value of the mean square error is equal to 0.0050645 at the time of training data, which is 0.00:04.

5. Conclusion

1. The line has different behavior look at the (2017and 2020) in the original time series sequence, these due to a new epicenter of COVID-19 emerged in several countries, leading to a re-imposition of travel restrictions in (2020). The struggle against ISIS, which leaded to stopped travelling on Erbil airport during period of time. as we saw in line of 2017.
2. The empirical of this study revealed the best ARIMA validated model to forecast future values of the total passengers of time series for the nest year (2023).
3. Five *ARIMA* models validated for the same time series of total passengers, where the MES, MAE and AIC of *ARIMA* (2,0,2) is the best one. Because it has the lowest MES, MAE and AIC values.
4. The diagnostic checking models show that ANN model is more adequate than Time series models.
5. linear regression and Convolutional Neural Network (CNN) [24]. In the future work, such models can be included to broaden the comparison.
6. It can be concluded that the minimum value of the Adaline Network's learning rate is 0.005 of the measurement to select the best model.
7. The best performance of the Adaline network depending on a single hidden layer and six notes with the training, validation, and testing set at the time is 0.00:04.

6. Recommendations

1. In this paper, only five time-series analysis methods are used to predict the total passengers as a case study and compare with each other. There are actually many more models that can be used in this time-series.
2. Only univariate data is considered in this research. Many researchers have proven that more sources of data can help increase the accuracy of prediction.
3. In recent years, usually combine both the statistical models (ARIMA) and machine learning approaches (ANN) to produce better results.
4. Using the Adaline network and comparing it with Box and Jenkins.
5. Depending on the model 1-6-1, it has less time to forecast the future.
6. The Adaline network is a highly effective forecast for this research.

References

Abd Saleh, Z. (2013). *Forecasting by Box-Jenkins (ARIMA) Models to Inflow of Haditha Dam.* Journal of Babylon University/Engineering Sciences/ No. (5)/ Vol. (21), *Babylon Univ., Babil, Iraq*.

Bergson A., Whitehead, B. & Choate, T. (1996). *Cooperative- Competitive Genetic Evolution of Radial Basis Function Centers and Widths for Time Series Prediction.* IEEE Trans. on Neural Networks, Vol. 7, no. 4, pp. 869-880.

Bianchini, M. & Scarselli, F. (2014*). On the Complexity of Neural Network Classifiers: A Comparison between Shallow and Dee Architectures.* IEEE Transactions on Neural Networks and Learning Systems. University of Siena. Via Roma 56, I-53100, Siena, Italy.

Box, G. E. & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control San Francisco. Calif: Holden-Day*.

Box, G. Jenkins, G. & Reinsel, G. (1994). *Time Series Analysis, Forecasting and Control*. 3rd edition, Prentice Hall, Englewood Clifs, NJ.

Bogdan, M. (2009). *Neural Network Architectures and Learning algorithms*. IEEE Industrial Electronics Magazine, Vol 3, no 4, pp.56-63.

Bogdan, M., Cotton, N. & Kaynak, O. (2007). *Neural Network Trainer with Second Order Learning Algorithms*. 11th INES -International Conference on Intelligent Engineering Systems, Budapest, Hungary, June 29, pp. 127-132.

Brockwell, P. J. & Davis, R. A. (2016). *Introduction to Time Series and Forecasting*. Third Edition, Department of Statistics Columbia University, New York, NY, USA, ISBN 978-3-319-29852-8

Chenhui, X. (2021). *A Comparative Study: Time Series Analysis Methods for Predicting COVID-19 Case Trend.* Degree Project in Computer Science and Engineering, Second Cycle, Stockholm, Sweden.

Din, M. A. (2015). *ARIMA by Box Jenkins Methodology for Estimation and Forecasting Models in Higher Education*. Athens Institute for Education and Research ATINER, ATINER's Conference Paper Series EMS015-1846

Husek, R. (2007). *Ekonometrická analýza*. 1st edition. Praha: Oeconomica. 2007, pp. 368. ISBN 978-80-245-1300-3. In Czech.

Lin, F., Yu, X., Gregor, S. & Irons, R. (1995). *Time Series Forecasting with Neural Networks.* Complexity International, Vol. 02, ISSN 1320-0682, Australia.

Meloun, M., Militky, J. (2004). *Statistická analýza experimentálních data*. 2nd edition. Praha: Academia. 2004, pp. 953. ISBN 80-200-1254-0. In Czech.

Štencl, M. Popelka, O. & Štastny, J. (2010). *Comparison of Time Series Forecasting with Artificial Neural Network and Statistical Approach.* Acta univ. agric. et silvic. Mendel. Brun., 2011, LIX, No. 2, pp. 347–35.

7. Appendixes

ACF for time Series: Total EIA Passengers (Autocorrelations)

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	.866	.115	56.274	1	.000
2	.683	.115	91.785	2	.000
3	.525	.114	113.029	3	.000
4	.422	.113	126.957	4	.000
5	.299	.112	134.044	5	.000
6	.180	.111	136.666	6	.000
7	.060	.110	136.963	7	.000
8	-.037	.110	137.075	8	.000
9	-.103	.109	137.972	9	.000
10	-.093	.108	138.709	10	.000
11	-.072	.107	139.160	11	.000
12	-.088	.106	139.855	12	.000
13	-.147	.105	141.811	13	.000
14	-.225	.104	146.454	14	.000
15	-.286	.103	154.110	15	.000
16	-.321	.103	163.897	16	.000

ACF for time Series: after transformation Total EIA Passengers (Autocorrelations)

Lag	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
			Value	df	Sig.[b]
1	.333	.116	8.220	1	.004
2	.009	.115	8.226	2	.016
3	-.120	.115	9.328	3	.025
4	-.306	.114	16.572	4	.052
5	-.137	.113	18.041	5	.073
6	-.030	.112	18.110	6	.066
7	.020	.111	18.142	7	.011
8	-.050	.110	18.350	8	.019
9	-.100	.109	19.190	9	.084
10	-.035	.108	19.293	10	.037
11	.023	.108	19.339	11	.055
12	.032	.107	19.427	12	.079
13	.011	.106	19.437	13	.110
14	-.055	.105	19.712	14	.139
15	-.107	.104	20.777	15	.144
16	-.016	.103	20.800	16	.186

1. ARIMA (2,0,0)

Model Statistics

Model	Number of Predictors	Model Fit statistics					Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	RMSE	MAPE	MAE	Normalized BIC	Statistics	DF	Sig.	
Total EIA Passengers-Model_1	0	.769	27047.690	65.017	18033.689	20.589	14.658	16	.550	0

2. ARIMA (1,0,1)

Model Statistics										
Model	Number of Predictors	Model Fit statistics					Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	RMSE	MAPE	MAE	Normalized BIC	Statistics	DF	Sig.	
Total EIA Passengers-Model_1	0	.767	27122.810	68.752	18015.874	20.594	15.410	16	.495	0

3. ARIMA (2,0,2)

Model Statistics										
Model	Number of Predictors	Model Fit statistics					Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	RMSE	MAPE	MAE	Normalized BIC	Statistics	DF	Sig.	
Total EIA Passengers-Model_1	0	.781	26687.820	64.210	17913.235	20.681	10.328	14	.738	0

4. ARIMA (1,0,2)

Model Statistics										
Model	Number of Predictors	Model Fit statistics					Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	RMSE	MAPE	MAE	Normalized BIC	Statistics	DF	Sig.	
Total EIA Passengers-Model_1	0	.772	27059.964	64.983	18105.851	20.649	13.295	15	.580	0

5. ARIMA (0,1,1)

Model Statistics										
Model	Number of Predictors	Model Fit statistics					Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	RMSE	MAPE	MAE	Normalized BIC	Statistics	DF	Sig.	
Total EIA Passengers-Model_1	0	5.551E-16	28766.988	45.378	18811.468	20.594	20.756	18	.292	0