Forecasting exports and imports through artificial neural network and autoregressive integrated moving average

Teg Alam*

*Corresponding author.
E-mail address: t.alam@psau.edu.sa (T. Alam)

© 2019 by the authors; licensee Growing Science, Canada.
doi: 10.5267/j.dsl.2019.2.001

1. Introduction

The Kingdom of Saudi Arabia preserves the largest amount of export of petroleum and it has the second-largest proven petroleum and the fifth-largest proven natural gas reserves in the world. The economy of the country depends primarily on oil and gas products. Saudi Arabia exported SAR 611.48B and imported SAR491.43B in 2016, yielding a positive trade balance of SAR 119.29B. The growth domestic product (GDP) of Saudi Arabia was SAR 2423.40B and its GDP per capita was SAR 204.08K.

2. Methods and Materials

2.1 Artificial Neural Network

Artificial Neural Network (ANN) is a well-organized data mining technique which is achieved from a biological neural networks. ANN collects a large amount of data interconnected in some specific patterns to help communication among various units normally called nodes or neurons and each of these is joint with other neurons through some connection links. Each association is joint with a particular weight, which gives some feedback about the input data. This is an essential part of neurons...
to come up with a particular problem. Each neuron maintains a combined state or an activation signal. Output signals, produced after joining the input signals and activation rule, are dispatched to other units. Some important developments of ANN are given in Table 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>1943</td>
<td>Warren McCulloch and Walter Pitts</td>
<td>Physiologist, and mathematician’s ideas are used for ANN purposes</td>
</tr>
<tr>
<td>1956</td>
<td>Taylor</td>
<td>An associative memory network</td>
</tr>
<tr>
<td>1964</td>
<td>Taylor</td>
<td>Winner-take-all circuit with association with output units</td>
</tr>
<tr>
<td>1969</td>
<td>Minsky and Papert</td>
<td>Multilayer perceptron concept</td>
</tr>
<tr>
<td>1971</td>
<td>Kohonen</td>
<td>Associative memories</td>
</tr>
<tr>
<td>1986</td>
<td>Rumelhart, Hinton, and Williams</td>
<td>Generalised Delta Rule</td>
</tr>
<tr>
<td>1988</td>
<td>Kosko</td>
<td>A hybrid of Binary Associative Memory and Fuzzy Logic ANN</td>
</tr>
</tbody>
</table>

2.1.1. Basic Model of Artificial Neural Network

An ANN mode can be expressed in Fig. 1 as follows,

![Neural Network Diagram](image-url)

Fig. 1. Neural Network

Any ANN configuration can be computed as \( y_w = \sum W_i x_i \). The output is measured using the function \( Y \) based on the net input \( (F(y_{in})) \). ANN has been widely used for predicting different incidents (Gaida et al., 2017; Kotur & Žarković, 2016; Sözen et al., 2011; Deng, 2010; Tektaş, 2010). Kavaklıoğlu et al. (2009) applied ANN method to estimate the electricity consumption using the historical data over the period 1976-2006. Ardakani and Ardehali (2014a,b) used ANN for prediction of electrical energy consumption for some countries. Li et al. (2007) applied ANN for estimating crop yield and compared their results with multivariate regression (Chamberlain, 1982). Zeng et al. (2017) implemented enhanced back-propagation for energy consumption predicting using neural network. Sokolov-Mladenović et al. (2016) estimated economic growth using ANN with a learning system based on trade, import and export parameters. Kankal and Uzlu (2017) applied ANN with a metaheuristics method to forecast demand for electricity. Tsai and Huang (2017) applied ANN for forecasting container flows among some Asian ports. According to Aydin et al. (2016), some of the world’s largest energy consumer (HC) consume approximately 62% of the world energy consumption. Thus, it is essential to find a predicted the future of world energy consumption. They proposed an ANN based model for HCs’ energy consumption. Olgun et al. (2012) applied ANN for predicting the demand of natural gas in Turkey and compared their results with support vector machines (Joachims, 1998; Uddin, 2009; Hsu & Lin, 2002; Chang & Lin, 2011). Liu et al. (2017) applied ANN to forecast the Chinese energy consumption. Mollaiy-Berneti (2015) used a hybrid of ANN based on the back-propagation (BP) type neural network and some metaheuristics. The method offered some advantage over the local search capabilities of BP technique and some search capability of some metaheuristics algorithm. Panda et al. (2010) used ANN method for the prediction of the agricultural crop yield...
prediction. This paper uses ANN technique to estimate the import and export of the Kingdom of Saudi Arabia.

2.1.2. Auto-Regressive Integrated Moving Average (ARIMA)

The early time series are concentrated on stochastic processes by Walker (1931). Udny Yule (1927) and Wold (1938) are among the first who introduced Autoregressive Moving Average (ARMA) models for time series, but was not able to determine the likelihood function for maximum likelihood (ML) estimation of the parameters (Ljung & Box, 1978). Then Box and Jenkins (Box et al., 2015) developed their methods for time series for forecasting purposes. Today, most important tools based on Box-Jenkins models (Kendall, 1995; Olajide et al., 2012) are commonly used for forecasting. ARIMA models are the most comprehensive times series for forecasting purposes. The linear type ARIMA model is considered as a primary forecasting one for a stationary data where the predictors includes of lags of the dependent variable and/or lags of the forecast errors. A typical ARIMA model is stated as an ‘ARIMA \((p, d, q)\)’ model.

where,
- \(p\) represents the number of the autoregressive terms,
- \(d\) denotes the number of non-seasonal differences required for stationarity
- \(q\) is associated with the number of lagged prediction errors in the forecasted equation.

Let \(y\) denote the \(d\)th difference of \(Y\), then, the ARIMA can be stated as follows,

- If \(d=0\): \(y_t = Y_t\)
- If \(d=1\): \(y_t = Y_t - Y_{t-1}\)
- If \(d=2\): \(y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}\)

and a general forecasting is formulated as follows,

\[
\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}
\]

Here the moving average parameters (\(\theta\)’s) are defined so that their signs are negative in the equation, following the convention introduced by Box and Jenkins. Various ARIMA models that are commonly encountered are given in Table 2.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Various ARIMA models</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(1,0,0)</td>
<td>First-order autoregressive model</td>
</tr>
<tr>
<td>ARIMA(0,1,0)</td>
<td>Random walk</td>
</tr>
<tr>
<td>ARIMA(1,1,0)</td>
<td>Differenced first-order autoregressive model</td>
</tr>
<tr>
<td>ARIMA(0,1,1) Without constant</td>
<td>Simple exponential smoothing</td>
</tr>
<tr>
<td>ARIMA(0,1,1) With constant</td>
<td>Simple exponential smoothing with growth</td>
</tr>
<tr>
<td>ARIMA(0,2,1) Without constant</td>
<td>Linear exponential smoothing</td>
</tr>
<tr>
<td>ARIMA(1,1,2) Without constant -</td>
<td>Damped-trend linear exponential smoothing</td>
</tr>
</tbody>
</table>

ARIMA method has been extensively used for forecasting method (Montanari et al., 1997; Reikard, 2009). Valipour et al. (2013) in a novel work compared ARIMA, and artificial neural network for predicting the inflow of Dez dam reservoir using the monthly discharges from 1960 to 2007. They compared root mean square error and mean bias error and reported that their proposed ANN was the
best model for inflow prediction of the Dez dam reservoir. Khashei and Bijari (2010) used an ANN model for forecasting purposes and compared their results with ARIMA method. Khashei and Bijari (2011) in other assignment presented a hybridization of ANN and ARIMA models for time series forecasting. Pedro and Coimbra (2012) made an assessment using prediction technique such as ARIMA, k-Nearest-Neighbors (kNNs), Artificial Neural Networks (ANNs), and ANNs optimized by Genetic Algorithms (GAs/ANN) for solar power production with no exogenous inputs. Wang et al. (2015) implemented ARIMA for improving forecasting accuracy of annual runoff time series. Kazem et al. (2013) applied support vector regression with chaos-based firefly algorithm (Feng et al., 2013) for stock market price forecasting. Liu et al. (2014) used genetic algorithm for short-term wind speed forecasting.

In this survey, we have accomplished a survey on forecasting methods and the frequencies of the words used in Web of Science database. There are approximately 700 research works published and indexed in this database and Fig. 2 demonstrates the frequencies of the words used in research areas that use forecasting.

According to Fig. 2, ANN and ARIMA have been extensively used in forecasting techniques. When co-word analysis is implemented for scientometrics purposes, we apply clusters of keywords and their interconnections and the clusters are addressed as themes. Each theme obtained here is specified in terms of two perspectives; namely “density” and “centrality” and some basic statistics for density and centrality are implemented for classification of the themes into various groups. In a theme, the keywords and their intercorrelations draw a network graph, called a “thematic network” where “centrality” is considered as the horizontal axis and “density” is taken into account as the vertical axis. In a network, if there is a big correlation from one node with other nodes, we consider a higher centrality for it and it is considered as an important part in the network. Centrality is thus implemented to measure the correction degree among various topics. Thematic map is a kind of plot which makes it possible to analyze themes based on the quadrant in which they are placed. Themes in the upper-right quadrant are both well developed and important for the structuring of a research field such as “big data” and “big data analytics”. Themes in the upper-left quadrant have well developed internal ties but unimportant external ties and so are of only marginal importance for the field such as “social network”. Themes in
the lower-left quadrant are both “weakly developed and marginal”, mainly representing either emerging or disappearing. Themes in the lower-right quadrant are “important for a research field but are not developed”, so this quadrant groups transversal and general, basic themes such as “ARIMA” and “ANN” (See Fig. 3) (Esfahani et al., 2019; Salimi et al., 2019; Alavi et al., 2019; Gilani et al., 2019; Pourkhani et al., 2019; Tayebi et al., 2019; Javid et al., 2019).

Fig. 3. Thematic Map

Next, we present the results of the implementation of ARIMA and ANN methods.

3. Result and Findings

3.1. Data

The data regarding the total annual exports and imports of the Kingdom was collected from Saudi Arabian Monetary Authority (SAMA). The information were on yearly basis and in Saudi Arabian Riyal (SAR) from years 1968 to 2017. The summary statistics for exports and imports data of the Kingdom are given below in the Table 3 by using Software XLSAT.

Table 3
Summary Statistics of Export and Imports of the Kingdom

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exports</td>
<td>50</td>
<td>9118.000</td>
<td>1456502.000</td>
<td>384270.267</td>
<td>407367.661</td>
</tr>
<tr>
<td>Imports</td>
<td>50</td>
<td>2578.000</td>
<td>655033.364</td>
<td>181662.394</td>
<td>189268.157</td>
</tr>
</tbody>
</table>
3.2. Models for exports

It is evident from the Fig. 4 and Fig. 5 that the exports of the Kingdom were gradually increasing and decreasing over time up till the year 2009 and after that they were continually increasing up to the year 2012. Also, it is evident from the figures that exports had a decreasing trend after 2012 up to 2016. After that, they were gradually increasing.

Fig. 4. The results of Export using Neural Network

Fig. 5. The results of Export using ARIMA (1, 1, 2)

Fig. 6. Plot of ACF of Exports

Fig. 6 shows that the autocorrelations were positive, strong and deteriorated slowly, which indicates that there were possible shifts in both the mean and the variability over time for this series. It means the arithmetic mean may be edging upwards, and the variability may be increasing and after that for time being edging downwards.

3.2.1 Prediction of the export

ANN and ARIMA models are used to create the predictions for future values of the time series by using the XLSTAT package. From the Table 4, it is observed that the best prediction is provided by ANN and then ARIMA (1, 1, 2). Thus, we conclude that ANN model predicted the exports more efficiently than ARIMA models did.
### Table 4
Prediction of Export of the Kingdom using different model

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual Value of Export</th>
<th>Predicted Value for Export (in million SAR)</th>
<th>ARIMA(1,0,0)</th>
<th>ARIMA(0,1,1)</th>
<th>ARIMA(1,1,2)</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>763313.06</td>
<td>1227127.829</td>
<td>1280858.878</td>
<td>1196259.604</td>
<td>767425.7889</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>688423</td>
<td>739801.554</td>
<td>700605.619</td>
<td>679106.385</td>
<td>673640.1108</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>831881.29</td>
<td>669726.153</td>
<td>703692.124</td>
<td>752574.815</td>
<td>842459.3463</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>803961.547</td>
<td>851660.629</td>
<td>868907.714</td>
<td>1062343.541</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>777836.759</td>
<td>851660.629</td>
<td>875109.992</td>
<td>1113639.628</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td>753391.526</td>
<td>851660.629</td>
<td>879892.482</td>
<td>952126.2365</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE</td>
<td>143006.113</td>
<td>143237.4797</td>
<td>134996.6626</td>
<td>46122.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From Fig. 7, we concluded that the plot gives us an observed and predicted exports values for the years 1968 to 2017, as well as, for the next 3 years forecasted values of exports using ANN, ARIMA (1, 0, 0), ARIMA (0, 1, 0) and ARIMA (1, 1, 2) models.

![Fig. 7. Predicted Value for Export (in million SAR)](image)

### 3.3 Models for Imports

From the Fig. 8 and Fig. 9, it is concluded that the imports of the kingdom is also increasing and decreasing slowly over time up till year 2010 and after that it was continuously increasing up to year 2015. Also, it is evident from the above figures that imports are decreasing after 2015 up to 2017. After that, they are gradually increasing. Fig. 10 shows that the autocorrelations are strong, positive and deteriorating slowly which also indicates that there are possible shifts in both the mean and the variability over time for this series and the trend can be removed by differencing once or twice.
4. Prediction of Import

From the Table 5, it is observed that the best prediction is provided by ANN followed by the arima model; ARIMA (0, 1, 1).

**Table 5**

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual Value of Import</th>
<th>Predicted Value for Import (in million SAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ARIMA(1,0,0)</td>
</tr>
<tr>
<td>2015</td>
<td>655033.36</td>
<td>646286.769</td>
</tr>
<tr>
<td>2016</td>
<td>525636.01</td>
<td>649401.771</td>
</tr>
<tr>
<td>2017</td>
<td>504446.62</td>
<td>521740.477</td>
</tr>
<tr>
<td>2018</td>
<td>500835.375</td>
<td>511040.068</td>
</tr>
<tr>
<td>2019</td>
<td>497272.583</td>
<td>511040.068</td>
</tr>
<tr>
<td>2020</td>
<td>493757.591</td>
<td>511040.068</td>
</tr>
<tr>
<td>RMSE</td>
<td>37588.68409</td>
<td>34054.36786</td>
</tr>
</tbody>
</table>

It is evident from the Fig. 11 that the plot could give us an observed and predicted imports values for the years 1968 to 2017, as well as for the next 3 years forecasted values of imports using ANN, ARIMA (1, 0, 0), ARIMA (0, 1, 0) and ARIMA (1, 1, 2) models.
5. Residual Analytic for ANN and ARIMA (1, 1, 2)

It is evident from the Figs. 12-15, the residuals of the exports and imports of the models show analytic plots that are useful in making decision. From the aforementioned figures, we can say that the time series plot of the residuals model allow us to look for trends in the residuals. It is clear from the time series plot shown in the Figs. 12-15 that the series of residuals are a stationary series.
6. Conclusion

In this paper, we have presented a comprehensive review on the applications of two models; namely artificial neural network (ANN) and ARIMA models for the purpose of forecasting. We have also mentioned that the historical review has shown a significant progress made in the field of ANN. Now, we can say that today is an epoch of evolution for neural network tools and techniques. An ARIMA model is one of the best techniques for predicting the level of any time series data with any pattern of change and is suitable for at least 50 observations. The primary objective of the study was to predict the total annual exports and imports of the Kingdom of Saudi Arabia. It has been evident from the analysis that ANN, ARIMA (0,1,1) and ARIMA (1,1,2) were the most appropriate models for forecasting the total annual exports and imports of the Kingdom of Saudi Arabia. The forecasted values of total annual exports and imports of the Kingdom will be 952126.2365 (in million SAR) and 510767.219 (in million SAR) respectively for the year 2020, which would be fluctuating as compared with 2014-15.

References


