Application Of The Statistical Analysis For Prediction Of The Jordanian GDP By Using ARIMA Time Series And Holt's Linear Trend Models For The Period (2003-2013)

Dr. Hasan Yasien Touama
Faculty of Economics and Administrative Sciences, Zarqa University, Jordan
E-mail: dralfaisel @ yahoo.com, Mobile: 00962/772542991
* This research is funded by the Deanship of Research in Zarqa University /Jordan

Abstract

This study aimed to apply of the statistical analysis for prediction of the Jordanian Gross Domestic Product (GDP) at (market prices), by using ARIMA Time Series and Holt's Linear Trend models for the period (2003-2013), also to compare between the forecasting values of (Box-Jenkins) methodical ARIMA model, and the Holt's linear trend Exponential Smoothing.

To achieve the study objectives, the study is mainly based on the secondary data related to GDP selected from the annual reports of the Jordanian Economics for the period (2003-2013). The study findings a number of results, including:

1. The Jordanian Gross Domestic Product (GDP) will see a rise in the nearest future.
2. There were a statistically significant differences at the significance level ($\alpha = 0.05$), between the forecasting values of the Jordanian Gross Domestic Product (GDP), by using (Box-Jenkins) methodical ARIMA model for time series and the Holt's linear trend Exponential Smoothing model, in favor of the ARIMA model for time series.

Upon the foregoing results, the study reached to a number of conclusions, recommendations, and suggestions.

Keywords: (Box-Jenkins) methodical, ARIMA model, Exponential Smoothing, Holt's linear trend**,** Jordanian GDP.

1. Introduction

A time series is a set of values of a particular variable that occur over a period of time in a certain pattern. The most common patterns are increasing or decreasing trend, cycle, seasonality, and irregular fluctuations (Bowerman, et al, 2005). To model a time series event as a function of its past values, analysts identify the pattern with the assumption that the pattern will persist in the future.

To identify the appropriate ARIMA model for a Time Series, you begin by identifying the order (s) of differencing needing to stationeries the series and remove the gross features of seasonality, perhaps in conjunction with a variance-stabilizing transformation such as deflating. If you stop at this point and predict that the differenced series is constant, you have merely fitted a random walk or random trend model (Recall that the random walk model predicts the first difference of the series to be constant, and a random trend model predicts that the first difference of the series is a random walk, rather than a constant). However, the best random walk or random trend model may still have autocorrelated errors, suggesting that additional factors of some kind are needed in the forecasting equation.

2. Methodology

2.1. The Study Problem

The study problem is try to find an method that combine the forecasted values from the selected methods, as ARIMA time series model in such a way that the privileges of (Box-Jenkins) methodical and the Holt's linear trend Exponential Smoothing, which can be used to improve the forecasting process.

2.2. The Study Importance

The forecasting process considered of the topics that are gaining a great importance as it can through the forecasting for decision-makers in order to take the correct decisions, also it helps the management levels in the process of decision-making in the fields of industry, agriculture, economics, and administration, etc..., and the forecasting of the Jordanian Gross Domestic Product (GDP) is very important in the posterior periods as a result of the economic situation in Jordan, as well as the existence of the economic crisis that threaten the continuation of the crash of the situation of the Jordanian economy.

2.3. The Study Objectives

The main objective of this study is to apply of the statistical analysis for prediction of the Jordanian Gross Domestic Product (GDP) by using ARIMA model for Time Series and Holt's Linear Trend model for the period (2003-2013). This main objective will be achieved through achievement of the following sub-objectives:

1. To identify at the ARIMA model for Time Series.

3. To identify at the Exponential Smoothing, and the Holt's Linear Trend model.
4. Compare between the forecasting values of the Jordanian GDP using ARIMA model for Time Series and Holt's Linear Trend model for the period (2014-2020).
5- Compare between the outcomes of applying the two models of Time Series, in order to select the best model.

2.4. The Study Hypotheses

To achieve the study objectives, it has been putting one hypothesis as a null form (H_0), as follows:
H_0: There are no a statistically significant differences at the significant level ($\alpha = 0.05$), between the forecasting values of the Jordanian Gross Domestic Product (GDP), by using (Box-Jenkins) methodical ARIMA model for time series and the Holt's linear trend Exponential Smoothing model during the period (2014-2020).

3. The Theoretical Framework

3.1. ARIMA model for Time Series

The ARIMA models are the most general class of models to forecasting a time series which can be made to be “stationary” by differencing, perhaps in conjunction with nonlinear transformations such as logging or deflating. A random variable that is a time series is stationary if its statistical properties are all constant over time. A stationary series has no trend, its variations around its mean have a constant amplitude, and it wiggles in a consistent fashion. An ARIMA model can be viewed as a “filter” that tries to separate the signal from the noise, and the signal is then extrapolated into the future to obtain forecasts.

The acronym ARIMA stands for Auto-Regressive Integrated Moving Average. Lags of the stationeries series in the forecasting equation are called "autoregressive" terms, lags of the forecast errors are called "moving average" terms, and a time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series. A non-seasonal ARIMA model is classified as an ARIMA (p,d,q) model, where:

p : Number of autoregressive parameter in the model.
d : Number of non-seasonal was differenced.
q : Number of moving average parameter in the model.

That is mean, the Autoregressive Integrated Moving Average ARIMA model generates non-stationary series that are integrated of order *d*, denoted ***I(d)***. A non-stationary ***I(d)*** process is one that can be made stationary after taking ***d*** differences. Such processes are often called **difference-stationary** or **unit root** process.

A series that you can model as a stationary ARMA(p,q) process after being differenced ***d*** times is denoted by ARIMA(p,d,q). Therefore, the form of the ARIMA(p,d,q) model is given by the following formula (Heino, 2005: p. 29):

$$\Delta^d y_t = c + \Phi_1 \Delta^d y_{t-1} + \dots + \Phi_p \Delta^d y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \qquad ,\dots(1)$$

where:

Δ^d : *d* th differenced time series.

y_t : Forecasting value.

c : Constant term.
ε_t : White Noise (forecasting errors), with mean zero.

That is mean, in lag operator notation, $L^i y_{t-i}$, can write the ARIMA (p,d,q) model as follows:

$$\Phi(L)\, y_t = \Phi(L)\,(1-L)^d\, y_t = c + \theta(L)\, \varepsilon_t \qquad ,\dots(2)$$

And $\Phi(L), \theta(L)$ are given as follows:

$$\Phi(L) = (1 - \Phi_1 L - \dots - \Phi_p L^P) \qquad ,\dots(3)$$
$$\theta(L) = (1 - \theta_1 L - \dots - \theta_q L^q) \qquad ,\dots(4)$$

Where:
Φ(L) **:** A stable degree *p* AR lag operator polynomial.
θ(L) **:** Invertible degree *q* MA lag operator polynomial.

3.2. Holt's Exponential Smoothing

The exponential smoothing is a procedure for continually revising a forecast in the light of more recent experience. Exponential smoothing assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more weight in forecasting than the older observations. And the exponential smoothing be on three methods are as follows (Prajakta, 2004: pp. 3-4):

a. Single Exponential Smoothing.
b. Double Exponential Smoothing (Holt's).
c. Triple Exponential Smoothing (Winter's).

In Single Exponential Smoothing, the forecast function is simply the latest estimate of the level. If a slope component is now added which itself is updated by exponential smoothing, the trend can be taken into account.

For a series $y_1, y_2 \dots y_t$, the forecast function at time, ***t*** for the value ***t+k*** is given by the following formula:

$$\hat{Y}_{t+k} = L_t + b_t k \quad , k = 1, 2, 3, \dots \quad , \dots(5)$$

Where:

L_t : The current level.
b_t : The current slope.

Since there are now two terms to the exponential smoothing, two separate smoothing constants are required, θ for the level and β for the slope. As in single exponential smoothing, the updated estimate of the level L_t is a linear combination of $\hat{y}_{t/t-1}$ and y_t :

$$L_t = \theta y_t + (1-\theta)(L_{t-1} + b_{t-1}) \quad , 0 < \theta < 1 \quad , t = 1,2,\dots,n \quad , \dots(6)$$

This provides the level at time, ***t*** . Since the level at time ***t-1*** is already known, it is possible to update the estimate of the slope:

$$b_t = \beta (L_t - L_{t-1}) + (1-\beta) b_{t-1} \quad , 0 < \beta < 1 \quad , t = 1,2,\dots,n \quad , \dots(7)$$

The parameters θ, β should lie in the interval **(0, 1)**, and can be selected by minimizing the criterions (*RMSE*, *MAE* or *MAPE*).

And we can forecast of the future values $Y_n(1)$ by the following formula:

$$Y_n(1) = L_n + b_t 1 \quad , \quad 1 > 0 \quad , \dots(8)$$

Also, we calculate the reasonable starting values for L_1 and b_1 , as follows:

$$L_1 = Y_1 \quad , \dots(9)$$
$$b_1 = Y_2 - Y_1 \quad , \dots(10)$$

3.3. Forecasting Accuracy Tests

There are three important criterions to test the **Accuracy** of the forecasting, and to compare between the forecasting values of the **ARIMA** and **Holt's** models**,** in order to determine the best model, as follows:

a. Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\sum e_t^2 / n} \quad , \dots(11)$$

b. Mean Absolute Error (MAE):

$$MAE = \sum |e_t| / n \quad , \dots(12)$$

c. Mean Absolute Percentage Error (MAPE):

$$MAPE = \sum [|e_t| / Y_t] / n \quad , \dots(13)$$

4. Statistical Analysis and Results Discussion

4.1. Estimation the Parameters of ARIMA (0,1,2) model

The results in Table (1), refers to estimate the parameters of ARIMA (0,1,2) model, as follows:

Table 1. Results of Estimation the Parameters of ARIMA (0,1,2) model

ARIMA Model Parameters Estimation					
Model	**Difference (1) / MA**	**Estimate**	**SE**	**t - value**	**Sig.**
GDP- model	Lag 1	- 0.721	0.125	- 5.768	0.000
	Lag 2	- 0.803	0.134	- 5.993	0.000

The estimation of the parameters related to ARIMA (0,1,2) model, which stated in Table (1), explained that the estimation of the parameters were statistically significant, because the statistical significant (Sig.) are equals to (0.000) were less than the significant level ($\alpha = 0.05$).

According to the estimation of the parameters of ARIMA (0,1,2) model, the **forecasting model** was explained by the following formula:

$$Y_t = \eta - Y_{t-1} + 0.721\ \varepsilon_t + 0.803\ \varepsilon_{t-1}$$

4.2. Test the adequate of ARIMA (0,1,2) model

The results in Table (2), refers to the criterions of testing the forecasting accuracy of ARIMA (0,1,2) model, as follows:

Table 2. Results of the Criterions for testing the Forecasting Accuracy

Model	**Accuracy Criterions**			**Box- Ljung Statistic (Q*)**		
	RMSE	**MAE**	**MAPE**	**Q***	**df.**	**Sig.**
GDP- model	5.956	3.178	6.387	0.405	9	0.817

Critical value of (χ^2) with (df. = 9), at the significant level ($\alpha = 0.05$) = 16.919

The results of the criterions for testing the forecasting accuracy of ARIMA (0,1,2) model incoming in Table (2), indicates to:

a. The ARIMA (0,1,2) model is **adequate** to forecast the Jordanian GDP, because the value of the Box -Ljung (Q*) statistic which is equals to (0.405) is less than the critical value of Chi-square (χ^2) amount (16.919) at the significant level ($\alpha = 0.05$), and the significant (p-value) (0.817) is greater than ($\alpha = 0.05$).

b. The **forecasting accuracy** by the ARIMA (0,1,2) model is very high**,** because the values of the accuracy criterions (RMSE, MAE, and MAPE) were very low.

4.3. Forecasting of the Jordanian GDP by ARIMA (0,1,2) model

The results in Table (3), indicates to the **forecasting** results of the Jordanian GDP by ARIMA (0,1,2) model during the period (2014 -2020), as follows:

Table 3. Forecasting Results of the Jordanian GDP by ARIMA (0,1,2) model

Years	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Forecasting	25558.7	27324.3	29090.0	30855.6	32621.3	34386.9	36152.5

4.4. Estimation the Parameters of the Holt's Linear Trend Model

The results in Table (4), refers to estimate the parameters of the Holt's Linear Trend model, as follows:

Table 4. Results of Estimation the Parameters of the Holt's Linear Trend model

Holt's Linear Trend Model Parameters Estimation					
Model	**Smoothing constants**	**Estimate**	**SE**	**t - value**	**Sig.**
GDP- model	Theta (Level) / $\theta = 0.2$	0.88	0.193	4.559	0.001
	Beta (Trend) / $\beta = 0.1$	0.91	0.308	2.955	0.003

The estimation of the parameters related to Holt's Linear Trend model, which stated in Table (4), explained that the estimation of the parameters were statistically significant, because the significant (p-value) are equals to (0.001), (0.003) were less than the significant level ($\alpha = 0.05$).

4.5. Test the adequate of the Holt's Linear Trend Model

The results in Table (5), refers to the criterions of testing the forecasting accuracy of Holt's Linear Trend model, as follows:

Table 5. Results of the Criterions for testing the Forecasting Accuracy

Model	**Accuracy Criterions**			**Box- Ljung Statistic (Q*)**		
	RMSE	**MAE**	**MAPE**	**Q***	**df.**	**Sig.**
GDP- model	7.345	4.921	8.649	1.682	9	0.673

[Critical value of (χ^2) with (df. = 9), at the significant level ($\alpha = 0.05$)] = 16.919

The results of the criterions for testing the forecasting accuracy of the Holt's Linear Trend model incoming in Table (5), indicates to:

a. The Holt's Linear Trend model is **adequate** to forecast the Jordanian GDP, because the value of the Box - Ljung (Q*) statistic which is equals to (1.682) is less than the critical value of Chi-square (χ^2) amount (16.919) at the significant level ($\alpha = 0.05$), and the significant (p-value) (0.673) is greater than ($\alpha = 0.05$).

b. The forecasting accuracy by the Holt's Linear Trend model is very high**,** because the values of the accuracy criterions (RMSE, MAE, and MAPE) were low.

4.6. Forecasting of the Jordanian GDP by Holt's Linear Trend model

The results in Table (6), indicates to the forecasting results of the Jordanian GDP by the Holt's Linear Trend model during the period (2014 -2020), as follows:

Table 6. Forecasting Results of the Jordanian GDP by Holt's Linear Trend model

Years	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
Forecasting	23804.7	24155.5	24789.3	25649.4	26690.6	27876.7	29178.7

4.7. Comparative between the results of ARIMA (0,1,2) and Holt's models

The results in Table (7), indicates to the values of the criterions related for ARIMA (0,1,2) model and Holt's Linear Trend model, as follows:

Table 7. Comparative between the Criterions of ARIMA (0,1,2) and Holt's models

Criterion	ARIMA (0, 1, 2) model	Holt model
RMSE	5.956	7.345
MAE	3.178	4.921
MAPE	6.387	8.649

The results contained in Table (7), explained that the ARIMA (0,1,2) model is the **best** of the Holt's Linear Trend model, because of all the criteria of ARIMA (0,1,2) model values are less than those of the Holt's Linear Trend model.

4.8. Test the study hypothesis

H_0: There are no a statistically significant differences at the significant level (α = 0.05), between the **forecasting values** of the (GDP), by using ARIMA model for time series and the Holt's linear trend model during the period (2014-2020).

To test the study hypothesis validity, **Mann-Whitney Test was** used as showed in table (8):

Table 8. Results of the Mann-Whitney Test

Model	Mann-Whitney U	Mean Rank	Ranks Sum	Exact Sig.
ARIMA	7	10	70	0.026
Holt		5	35	

[Critical value of Mann-Whitney (U_0) at (n=7, m=7), and (α = 0.05)] = 12

The results in Table (8), indicates to there exist a statistically significant differences at the significant level (α = 0.05), between the **forecasting values** of the (GDP), by using ARIMA (0, 1, 2) time series and the Holt's linear trend models during the period (2014-2020), and these differences were in favor of ARIMA (0,1,2) time series model. Which is supported by the calculated value of **Mann-Whitney U (7)** is less than the critical value of **Mann-Whitney U_0** (12) at the significant level (α = 0.05), and the Exact significant (p-value) (0.026) is less than (α = 0.05). This means that will be **rejected** the null hypothesis (H_0).

5. Conclusions and Recommendations

5.1. Conclusions

a. The Jordanian Gross Domestic Product (GDP) will see a rise in the nearest future.

b. For the time series under study, the use of ARIMA (0,1,2) model is the **best** of the Holt's Linear Trend model.

c. There were a statistically significant differences at the level (α = 0.05), between the forecasting values of the (GDP), by using ARIMA (0,1,2) model for time series and the Holt's linear trend model, in favor of the ARIMA model for time series during the period (2014-2020).

d. The method of Holt's linear trend predictive was achieved accuracy less than of ARIMA (0.1,2) model, shows that by the predicting results of which were in the case of using ARIMA (0,1,2) more than reasonable to predict in the case of using the Holt's Linear Trend model.

5.2. Recommendations

Based on the previous results, the study recommends the following:

a. Necessary to use ARIMA (0,1,2) time series model for the purposes of prediction, because it is more efficient than the Holt's Linear Trend model.

b. The study suggests conducting future relevant studies including other models, as Holt –Winters exponential smoothing model for time series forecasting.

References:

1. Andres, M. A., and Carolina, G. M., (2012), ***Time Series Analysis: Forecasting with ARIMA Models,*** **Universidad Carlos III de Madrid.**
2. Asteriou, D., Hall, S. G., (2011), ARIMA Models and the Box-Jenkins Methodology, ***Applied Econometrics***, 2nd Ed., Palgrave MacMillan, pp. 265-286.
3. Armstrong, J. S., (2003), ***Principles of Forecasting***, Kluwer Academic Publishers.
4. Al-Aani, A. H., (2005), ***Usage of the ARIMA model in Economic Forecasting***, Faculty of Management and Economics Journal, Al-Anbar University, Iraq.
5. Bachini, F., Ciammola, A., Iannaccone, R., and Marini, M., (2010), Combining Forecasts for a flash Estimate of Euro Area GDP, 5th Ed., ***Colloquium on Modern Tools for Business Cycle Analysis***, ISTAT.
6. BuHamra, S., Smaoui, N., and Gabr, M., (2003), The Box-Jenkins Analysis and Neural Networks Prediction and Time Series Modeling, ***Applied Mathematical Modeling***, 27, pp. 805-815.
7. Biery, A. M., (2002), ***Statistical Forecasting Methods***, Suaad King University, Riyadh, Saudi Arabia.
8. Bowerman, B. L., O'Connell, R. T., and Anne, B. K., (2005), ***Forecasting Time Series and Regression***, 4th Ed., Belmont, CA: Thomson Brooks / Cole.
9. Box, G. E., Jenkins, G. M. and Reinsel, G. C., (2008), ***Time Series Analysis: Forecasting and Control***, 4th Ed., Hoboken , New Jersey, John Wiley & Sons.
10. Brocklebank, J. C. Dickey, D. A., (2003), ***SAS for Forecasting and Time Series***, SAS Institute Inc., Cary, Nc, USA.
11. Brockwell, P. J. and Davis, R. A., (2002), ***Introduction to Time Series and Forecasting***, 2nd Ed., New York, Springer.
12. Chatfield, C., (2002), ***Time- Series Forecasting***, Boca Raton, Chapman & Hall ⁄ CRC.
13. Fatima, S., (2007), Hybrid System of Simple Exponential Smoothing and Neural Network for KSE100 Index, ***Market forces***, 2.
14 Hamad, R.A., (2010), Enhancing the Efficiency of Forecasting using Combined Time series Models and Neural Networks: An Applied Study, ***PH. Degree in Statistics***, Faculty of Commerce, Por-Sa'aid University, Egypt.
15. Heino, B. N., (2005), ***Univariate Time Series Analysis; ARIMA Models***, Econometrics, 2, p. 29.
16. Marisa, A., (2008), ***Comparative Study of Artificial Neural Network and Box –Jenkins ARIMA model for Stock Price Indexes***.
17 Pena, D., Tiao, G. C. and Tsay, R. S., (2001), ***A course in Time Series Analysis***, New York, John Wiley & Sons.
18. Prajakta, S. K., (2004), ***Time Series Forecastin*g *using Holt – Winters Exponential Smoothing***, Kanwal Rekhi School of iinformation Technology.
19. Sallehuddin, R., Shamsunddin, S.M., Hashim, S.Z., and Abraham, A., (2007), Forecasting Time Series Data Using Hybrid Relational Artificial Neural Network and Auto Regressive Integrated Moving Average (ARIMA) model, ***Neural Network World***, 6, pp. 573-605.
20. Shumway, R. H. and Stoffer, D. S., (2011), ***Time Series Analysis and its Applications: with examples***, 3rd Ed., New York, Springer.
21. Theresa, H. D., Warner, B., Entertainment, G., & Burbank, C.A., (2013), ***The Box-Jenkins Methodology for Time Series Models***, Statistics & Data Analysis, SAS Global, Forum.
22. Touama, H. Y., Emman, H. H., (2009), ***Methods of Applied Statistics***, Dar Al- Safa for Printing, Publishing and Distribution, Jordan, Amman.
23- Yaffee, R. A., and McGee, M., (1999), ***An Introduction to Time Series Analysis and Forecasting***, Academic Press, Inc.
24. Zou, H., and Yang, Y., (2004), Combining Time Series Models for Forecasting, ***International Journal of Forecasting*** , 20, pp. 69-84.

Biographic

Hasan Yasien Touama, Associate Professor of Applied Statistics and Quantitative Methods at Faculty of Economics and Administrative Sciences of Zarqa University in Jordan**,** he teaches and achieves the researches in Statistics, Quantitative Methods (Operations Researches & Production Management), and Econometrics. He took his BA in Statistical Sciences at (1977-1978), Master Degree in Designing & Analyzing Experiments at (1980-1981), and his Ph. D. in Applied Statistics and Quantitative Methods (Major) at (1994-1995) from Baghdad University. He is the author of (12) twelve books and over (31) thirty one article and research of the following subjects and fields, Applied Statistics, Bio-statistics, Experimental Design, Econometrics, Computerize Systems, Total Quality Management and Quantitative Methods.

Appendix 1. The Charts of Autocorrelation Function (ACF)