"Science Stays True Here"
Journal of Mathematics and Statistical Science, Volume 2016, 714-729 | Science Signpost Publishing

Using Box - Jenkins Models to Forecast the Cotton Crop In Iraq

Assist. Prof. Dr. Ahlam Ahmed Juma, Hala Muthanna Mohammed

Department of Statistics, College of Administration and Economy, University of Baghdad,

Central Bureau of Statistics, The Ministry of Planning, Iraq.

Abstract

The Forecasting of the time series of subjects that received attention especially effective role in contributing to the building plans for the future in all areas of life, including industrial, commercial, agricultural and other, As there were many methods of time series is the most important models developed worlds Box & Jenkins in 1970 and through the development of a methodology in the study and analysis of Autoregressive Integrated Moving Average Models ARIMA (p, d, q).

The research aims to the forecasting using linear time series and reaching a the best model can Predictable through, it was used non- seasonal Box-Jenkins models and were used the time series of the total area and the yield and production of cotton crop in Iraq for the period (1941 - 2011) was included forecasting future years (2012-2016).

1. Introduction

Time series is considered s one of the subjects that gained a special attention for their effective role in building the future plans is all fields of life industrial, commercial, agriculture and other wise as well as the methods of time series varied greatly and the most important the models put by Box and Jenkins (1970) by presenting a methodology in studying and analyzing autoregressive integrated moving average models ARIMA(p, d, q) and in them the value of time series random variable based on former values themselves and sometimes the former and present values of random error series when the random error follows a normal distribution with a mean zero and variance σ_a^2.[1][5]

Corresponding author: Assist. Prof. Dr. Ahlam Ahmed Juma, Department of Statistics, College of Administration and Economy, University of Baghdad, Central Bureau of Statistics, The Ministry of Planning, Iraq.

There is in Iraq appropriate agricultural land and climatic conditions and the availability of raining and rivers water.

The statistics refer that Iraq surface has four regions according to their to pornographic distribution and they can be arranged as top-bottom according to their proportion to the Iraq whole area (435052) kilo meters and they are as follows 39.2% desert lands, 30.24 plain land (including marshes and lakes), 21% mountains and 9.6 wavy lands and Iraq lies the north moderate region but it's climate is continental and sub optical and its rains in its system is the climate of mediteran sea and most of its rains fall in winter as well as in autumn and spring and from all that came the importance of agricultural sector that is considered one of the most important economic sector in Iraq as it provide food for the people and absorb the third of the whole employability as well as it provides the raw materials for the Iraqi industries

The cotton strategic corn is taken for its economic importance that it seeds are considered as a food for man to contains proton proportion (60%) as well as it is used in pastes industries and its using in industrial silk making and Sullivan and plastic industry.

The Objective of the Study:

The study aims at predicting the production of cotton crop according to area, production and yields on Iraq, level by using Box - Jenkins models.

2. Time series models [1][2]

Consider the classical linear model,

Most of time series models are stochastic models for non- stationary time series and they can be recognized through the two functions of auto correlation and partial auto correlation that their value aren't referenced to the zero after the second and third lags but their value remain big for a number of lags, but they have a good characteristic in their ability to transform to stationary time series and there are two conditions of stationary in the time series and they are: stationary in the mean that is achieved by taking a number of proper differences i.e.,

$$W_t = \nabla^d Z_t \qquad (2.1)$$

The symbol ∇ is called backward difference operator. And stationary in variance by taking a number of suitable transformation to reach to the stationary therein.

Autoregressive Integrated Moving Average Models [2][4]:

The formula of autoregressive integrated moving average models ARIMA(*p*,*d*,*q*) as follows:

$$\emptyset(B)(1-B)^d Z_t = \theta(B)a_t \tag{2.2}$$

Where,

$$(1-B)^d Z_t = W_t = \nabla^d Z_t \tag{2.3}$$

When p=1, d=1 and q=1 in the equation (2.2) we get ARIMA(1,1,1) models and its formula i.e.,

$$(1-\emptyset_1 B)(1-B)Z_t = (1-\theta_1 B)a_t \tag{2.4}$$

i.e.

$$Z_t = Z_{t-1} + \theta_1 Z_{t-1} - \theta_1 Z_{t-2} + a_t - \theta_1 a_{t-1} \;\; ; \; |\emptyset_1| < 1 \, ; |\theta_1| < 1 \tag{2.5}$$

When q=0 in the equation (2.2) the autoregressive integrated models ARIMA(*p*,*d*,0) ARI(*p*,*d*,0) is written according to the following formula:

$$\emptyset(B)(1-B)^d Z_t = a_t \tag{2.6}$$

When p=1, d=1, in the equation (2.6) then the model ARIMA (1,1,0) or ARI(1,1) is written according to the following formula:

$$(1-\emptyset_1 B)(1-B)Z_t = a_t \tag{2.7}$$

i.e.

$$Z_t = Z_{t-1} + \emptyset_1 Z_{t-1} - \emptyset_1 Z_{t-2} = a_t ; \; -1 < \emptyset_1 < 1 \tag{2.8}$$

When p=0 in the equation (2.2) then the integrated moving average modelARIMA(0,d,q) or IMA(0,*d*,*q*) will be in the following formula:

$$(1-B)^d Z_t = \theta(B)a_t \tag{2.9}$$

i.e.

$$\nabla^d Z_t = \left(1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)a_t \tag{2.10}$$

When d=1 and q=1 in the equation (2.9) the model ARIMA(0,1,1) or IMA(1,1) is written in the following formula:

$$(1-B)Z_t = (1-\theta B)a_t \tag{2.11}$$

i.e.

$$Z_t = Z_{t-1} + a_t - \theta_1 a_{t-1}; \ -1 < \theta < 1 \tag{2.12}$$

The following algorithm shows the levels of building time series model:

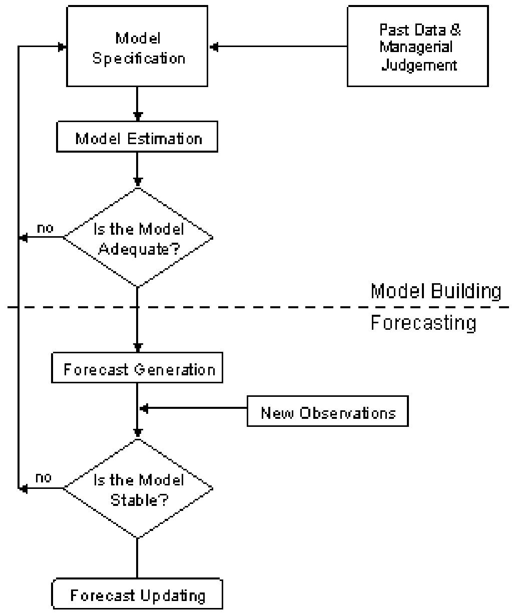

Figure 1. Algorithm Methodology Box - Jenkins

Comparison Criteria:

The criteria depended in practice are:

1- Bayesian information criterion[1][8]:

The Bayesian information criterion (BIC) takes the following formula:

$$BIC(M) = n \ \ln \hat{\sigma}_a^2 - (n - M) \ln(1 - \frac{M}{n}) + M \ln n + M \ln \left(\frac{\sigma_z^2}{\sigma_a^2} - 1\right) / M \tag{2.13}$$

And when some terms are neglected, the result will be the following abbreviated formula:

$$BIC(M) = n \ \ln \hat{\sigma}_a^2 + M \ln n \tag{2.14}$$

To get the standard formula, the equation (2.14) is divided by sample size (n)

$$NBIC(M) = n \ \ln \hat{\sigma}_a^2 + \frac{M \ln(n)}{n} \tag{2.15}$$

And the order is chosen for the model that meets the least value of the criterion.

2- Mean Square Error [7]:

The mathematic formula mean square (MSE) is:

$$\sum_{t=1}^{n} e_t^2 \ MSE = \frac{1}{n} \tag{2.16}$$

When the MSE is near zero this is an indicator that the series estimating value are very near from the real observations of the time series.

3- Mean Absolute Error [7]:

The mean absolute error (MAE) is the average of deviations from the real values and the formula:

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t| \tag{2.17}$$

4- Mean Absolute Percentage Error [7]:

The mean absolute percentage error (MAPE) is one of the accurate measures of the common usage in quantitative methods to predict and it formula:

$$PET = \left(\frac{Z_t - e_t}{Zt}\right) * 100 \tag{2.18}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} |PET| \tag{2.19}$$

Test for Diagnostic Checking:

1 - Testing the Confidence Intervals [2][8]:

If the auto correlation coefficients for the residuals of the model diagnosed lie within the two confidence intervals and on the level of confidence 95%, i.e.

$$pr\left\{\hat{r_k}(a)\right| < 1.96\frac{1}{\sqrt{n}}\right\} = 1 - \alpha \tag{2.20}$$

That refers to the errors is randomize and thus, the model diagnosed with be appropriate for the time series data.

2 - Box and Pierce [3]:

$$Q_{B\&P} = (n-d)\sum_{k=1}^{m} \hat{r_k}^2(a) \sim \chi^2_{(m-j),\alpha} \tag{2.21}$$

Where:

d: represents the differences taken. k: represents the number of lags and $k = 1, 2, \ldots, m$.

And then the calculated value $Q_{B\&P}$ is compared with the tabulate value x^2 with degree of freedom $m-j$ and significance level α.

3 - **Ljung and Box Test [6]:**

$$Q_{L\&B} = n(n+2)\sum_{k=1}^{m}\frac{\hat{r}_k^2(a)}{n-k} \sim \chi^2_{(m-j),\alpha} \tag{2.22}$$

The value $Q_{L\&B}$ is compared with the tabulate value x^2 with degree of freedom $m-j$ and significance level α.

3. The Practical Side

3-1: Data Collection

The data that the researcher depends there up and concerning the total of the planted area and production and the total average of the crop are issued from ministry of planning-Central Bureau of Statistics-the directorate of agricultural statistics for the cotton crop the period (1941- 2011) and an Iraq level, and the researcher depend in analyzing data on SPSSVar.20 and STATISTICAVar.8 programs.

3-2: Statistic Analysis

The three time series data have been drawn for the cotton crop according to the area, production and yields as in the illustrations cambered squinty (3-1), (3-2) and (3-3) to know the behavior data it is obvious that the data non stationary in variance as well as in the mean.

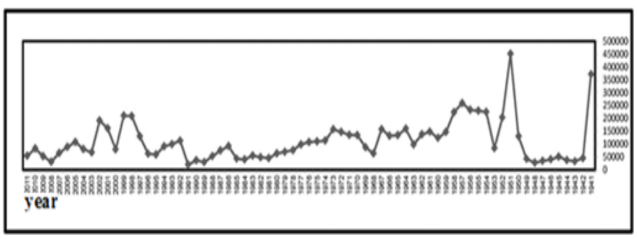

Figure (3-1)
The observations of cotton area time series in Iraq

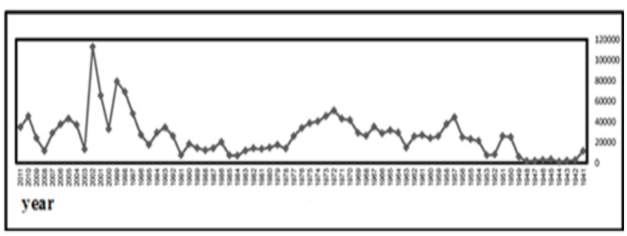

Figure (3-2)
Observations of cotton crop production time in Iraq

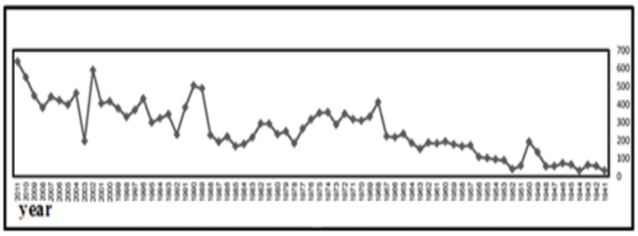

Figure (3-3)
Observations of the cotton yields time series in Iraq

The coefficients of the two functions ACF and PACF have been drawn as in the figures (3-4) and (3-5) concerning the area series and figures (3-6) and (3-7) concerning production series as well as figures (3-8) and (3-9) for the yields series.

By drawing the two functions, it has been shown that the series are non-stationary and the coefficients values lie outside the confidence interval {± 0.233}with confidence level 95%.

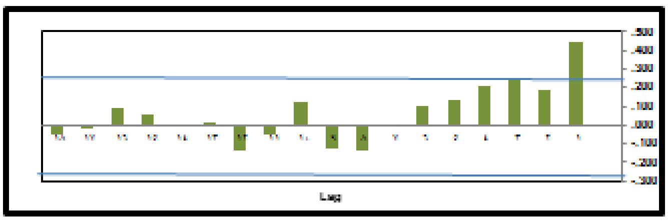

Figure (3-4)
ACF for cotton area

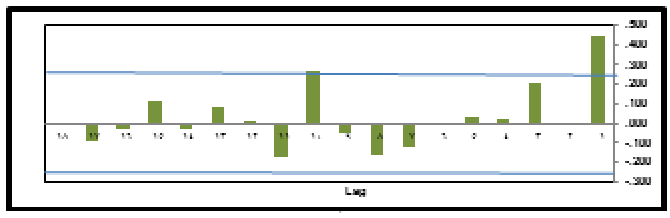

Figure (3-5)
PACF for cotton area

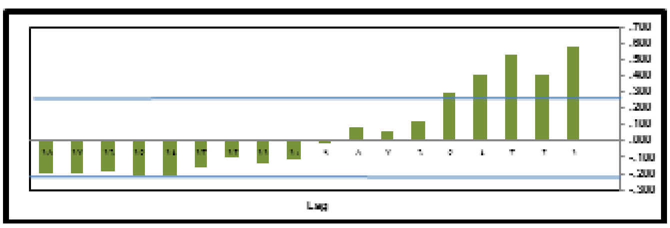

Figure (3-6)
ACF for cotton production

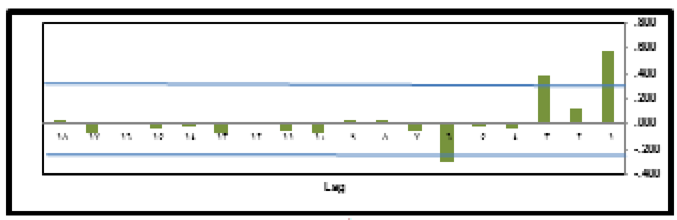

Figure (3-7)
PACF for cotton production

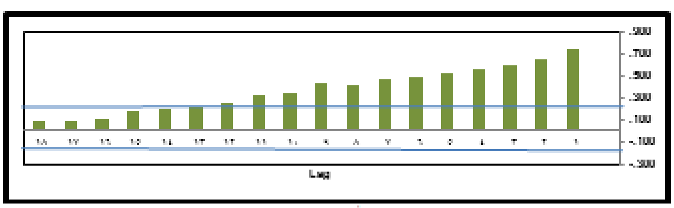

Figure (3-8)
ACF for cotton yields

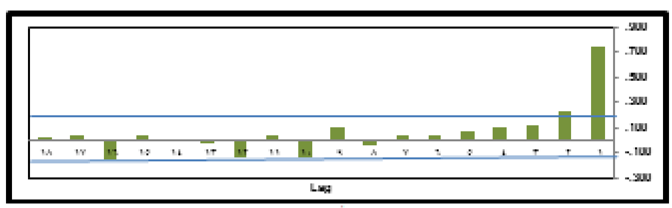

Figure (3-9)
PACF for cotton yields

The series of cotton area, the logarithm transformation has been taken to get the stationary in variance with the taking of the first difference of data to get the stationary in the mean and thus the series become stationary in variance and mean and through the two figures (3-10) and (3-11) concerning the coefficients of the two functions ACF and PACF, it has been shown that those coefficients lie within the confidence intervals{± 0.233} with confidence level 95% and this indicates that the data become stationary.

Regarding the cotton production, the logarithm transformation has been taken to reach to the stationary in variance as well as taking the first difference to get the stationary in the mean, thus, the series become stationary and this can be known from the values of the two functions ACF and PACF, and the two figures (3-12) and (3-13) showing that their values lie within the confidence intervals and the data become stationary.

And we take the first difference for the cotton yields has been taken to reach to the stationary in the mean and thus the figure (3-14) for the values of ACF function and figure (3-15) for the PACF which show that the data of this series are stationary.

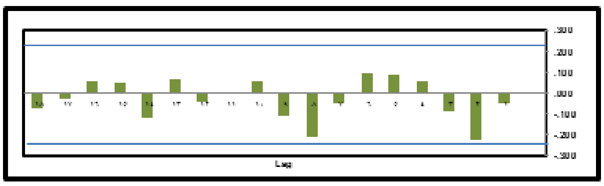

Figure (3-10)
ACF for cotton area after taken log. transf. & 1st diff.

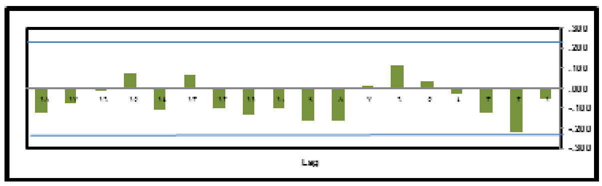

Figure (3-11)
PACF for cotton area after taken log. transf. & 1st diff.

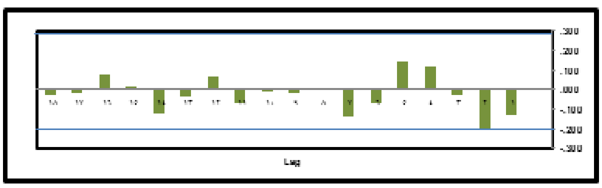

Figure (3-12)
ACF for cotton production after taken log transf. & 1st diff.

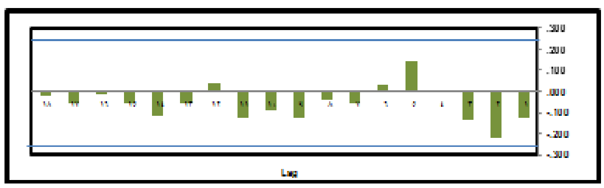

Figure (3-13)
PACF for cotton production after taken log. transf. & 1st diff.

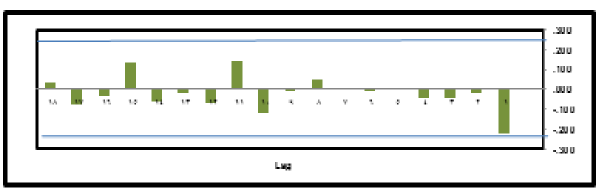

Figure (3-14)
ACF for cotton yields after taken 1st diff.

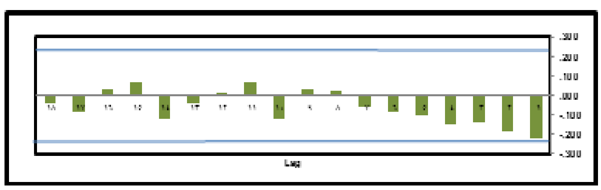

Figure (3-15)
PACF for cotton yields after taken 1st diff.

Cotton Area Series:

The model appropriate for the cotton crop area data has been defined by suggesting a number of models and illustrated in the table (3-1) as it is very difficult to define the model by the values of the two functions ACF and PACF.

This table shows that the model ARIMA (2,1,2) is the better one which has the least values in the criteria (NBIC, MAE, MAPE and RMSE) as well as the model coefficients have been estimated by using the approximate maximum likelihood method and the results are as follows:

$$\emptyset_1 = 0.715, \quad \theta_1 = 0.892, \quad \emptyset_2 = -O.747, \quad \theta_2 = -0.592$$

$$S.E = 0.205, \quad S.E = 0.244, \quad S.E = 0.153, \quad S.E = 0.220$$

Thus, the model is written in the following formula,

$$(1 - 0.715\, B + 0.747B^2)(1 - B)Z_t = (1 - 0.892\, B + 0.592\, B^2)a_t$$

The values of the two functions ACF and PACF related with the series of the diagnosed model residuals have been calculated and checking their coefficients calculated and shown in the figure (3-16) and (3-17) respectively that shows the randomness of those coefficients and they are lie within the two confidence intervals with a level of 95%.

We take two tests $Q_{L\alpha B}$ = 5.764 and $Q_{B\alpha p}$ = 5.595 have been calculated and when their value is compared with tabular value to x^2 with d.f (14) and significance level 0.05 and equal to (6.571) it, has been shown no significant differences and the model diagnosed model is appropriate for the cotton area series data.

Table (3-1)

The models suggested for the cotton area data series

Fit Statistic	MODELS					
	ARIMA(1,1,0)	**ARIMA(2,1,0)**	**ARIMA(0,1,1)**	**ARIMA(0,1,2)**	**ARIMA(1,1,1)**	**ARIMA(1,1,2)**
RMSE	87794.751	79960.754	84273.418	77768.872	90116.025	77363.085
MAPE	60.833	57.591	60.886	57.507	61.539	60.191
MAE	51405.126	47175.157	49152.088	44664.430	52477.470	45155.311
NBIC	22.826	22.700	22.744	22.644	22.939	22.695
Fit Statistic	**MODELS**					
	ARIMA(2,1,1)	**ARIMA(2,1,2)**	**ARIMA(3,1,0)**	**ARIMA(0,1,3)**	**ARIMA(3,1,1)**	**ARIMA(3,1,3)**
RMSE	79535.033	77177.005	80997.037	77517.926	81082.946	79499.359
MAPE	57.630	55.286	58.000	59.783	58.009	59.934
MAE	46376.725	43534.002	47796.986	45541.517	47171.702	45990.046
NBIC	22.750	22.5883	22.786	22.699	22.849	22.931

Fit Statistic	MODELS		
	ARIMA(1,1,3)	**ARIMA(2,1,3)**	**ARIMA(3,1,2)**
RMSE	77496.112	84711.578	84415.408
MAPE	59.119	55.973	56.225
MAE	43910.547	47597.524	47774.018
NBIC	22.759	22.997	22.990

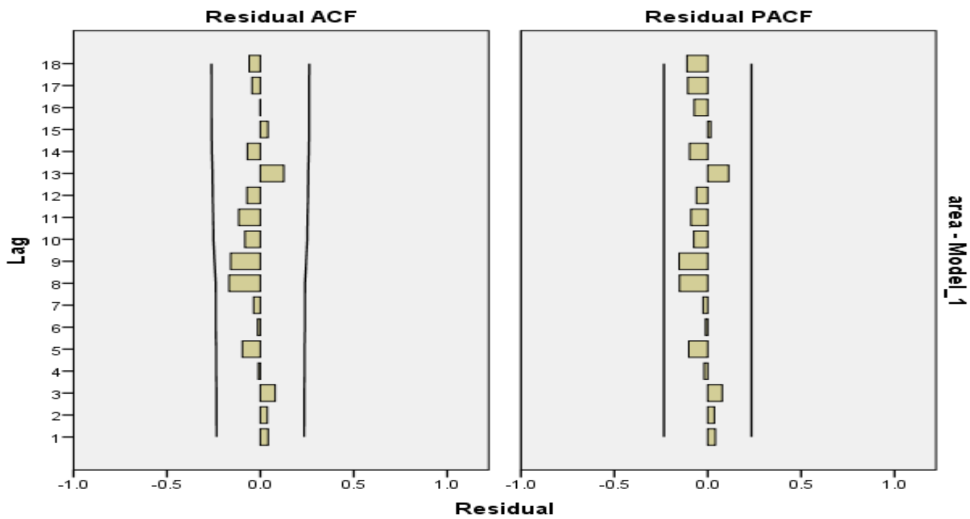

Figure (3-16) ACF for residualsARIMA(2,1,2)

Figure (3-17) PACF for residualsARIMA(2,1,2)

The forecasting values of the series as a result of the application of the selected model are shown in the table (3-2).

Table (3-2)

The forecasting values with the probability terms and significance interval 95% for the selected model and cotton area series in Iraq

Year	Lower Limit	Forecasting	Upper Limit
2012	18071.85	53152.03	156328.1
2013	14154.44	55635.62	218682.1
2014	12906.69	59390.03	273282.8
2015	11548.33	61489.04	327398.3
2016	9380.28	60592.38	391399.7

Cotton Production Series:

Was calculated number of models suggested and described in table (3-3) and the model ARIMA (2,1,2) was the best one as it has been given the following criteria (RMSE, MAPE, MAE, NBIC).

The estimation of model coefficients equals:

$$\Phi_1 = 0.482\,, \quad \Phi_2 = -0.694\,, \quad \theta_1 = 0.743, \quad \theta_2 = -0.528$$

$$S.E = 0.234, \quad S.E = 0.166 \quad S.E = 0.270, \quad S.E = 0.234$$

Accordingly, the model writes the following formula:

$$Z_t = Z_{t-1} + 0.482(Z_{t-1} - Z_{t-2}) - 0.694(Z_{t-2} - Z_{t-3}) + a_t - 0.743a_{t-1} + 0.528\, a_{t-2}$$

The autocorrelation and partial autocorrelation coefficients for residuals values for the model are calculated and shown in the two figures (3-18) and (3-19) respectively showing the randomness of the model that the coefficients lie within the two confidence interval {± 0.233} with confidence level 95%.

Table (3-3)

The models suggested for the cotton production data series

Fit Statistic	MODELS					
	ARIMA(1,1,0)	ARIMA(2,1,0)	ARIMA(0,1,1)	ARIMA(0,1,2)	ARIMA(1,1,1)	ARIMA(1,1,2)
RMSE	19721.917	16218.078	17322.900	16245.212	17055.799	16175.204
MAPE	68.156	61.694	64.896	61.602	63.236	61.750
MAE	11702.085	10494.033	10970.811	10515.694	10800.509	10473.365
NBIC	19.840	19.509	19.580	19.512	19.610	19.565
Fit Statistic	**MODELS**					
	ARIMA(2,1,1)	**ARIMA(2,1,2)**	**ARIMA(3,1,0)**	**ARIMA(0,1,3)**	**ARIMA(3,1,1)**	**ARIMA(3,1,3)**
RMSE	16251.273	15184.786	16674.184	15911.884	16814.068	17498.916
MAPE	60.747	60.632	60.699	61.580	60.911	60.835
MAE	10404.214	10309.209	10541.146	10369.780	10558.698	10328.455
NBIC	19.574	19.507	19.625	19.532	19.703	19.904

Fit Statistic	MODELS		
	ARIMA(1,1,3)	ARIMA(2,1,3)	ARIMA(3,1,2)
RMSE	16473.646	17362.316	15638.857
MAPE	61.741	60.865	57.883
MAE	10513.651	10354.817	9857.530
NBIC	19.662	19.828	19.618

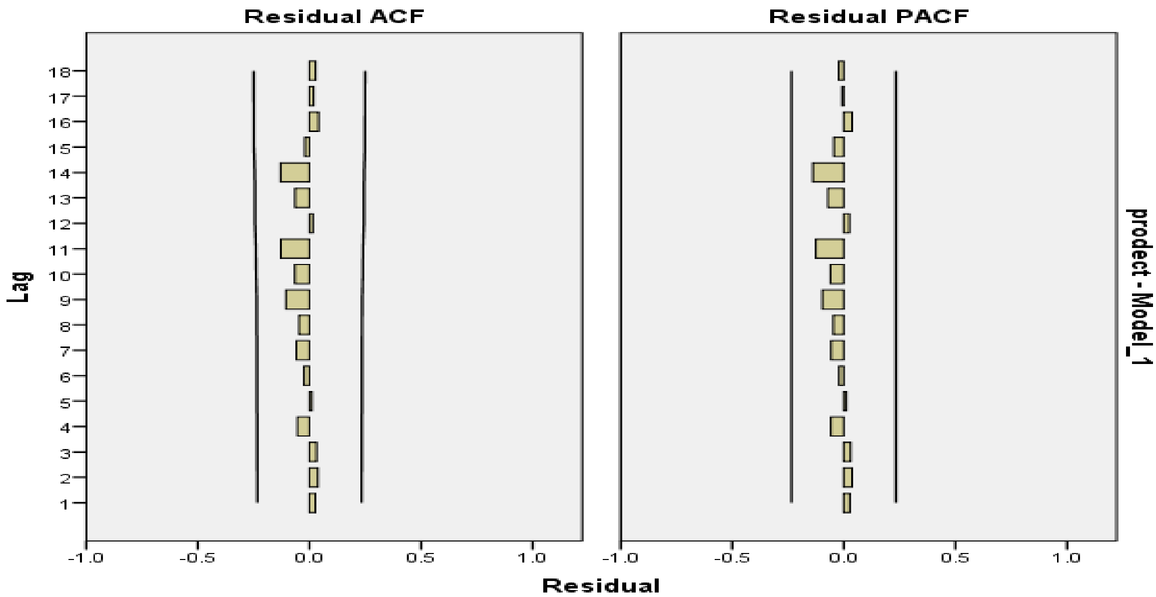

Figure (3-18)
ACF for residualsARIMA(2,1,2)

Figure (3-19)
PACF for residuals ARIMA(2,1,2)

The following two tests $Q_{L\alpha B}$ = 5.971 and $Q_{B\alpha P}$ = 5.869 have been calculated and when compared with thetabular value to x^2 in d.f (14) and significance level 0.05 is equals (6.571). Emerges that there are no significant differences and diagnosed model is appropriate data series.

The forecasting values of the series as a result of the application of the selected model are shown in the table (3-4).

Table (3-4)
The forecasting values with the probability terms and significance interval 95% for the selected model and cotton production series in Iraq

Year	Lower Limit	Forecasting	Upper Limit
2012	7714.936	24728.65	79262.6
2013	5472.181	23441.39	100416.8
2014	5966.115	28816.43	139183.8
2015	6290.155	34386.53	187981.6
2016	4973.994	33303.35	222982.4

Series Cotton Yields:

Been determined appropriate model for the data after the cotton yields suggest a number of models. Through the results in table (3-5) show that the model ARIMA (1,1,1) is appropriate because give less value in the criteria (MAE, MAPE, RMSE, NBIC).

Parameters estimation for the model is equal:

$$\Phi_1 = 0.167, \quad \theta_1 = 0.678$$

$$S.E = 0.220, \quad S.E = 0.174$$

And the model estimated is written in the following formula:

$$Z_t = Z_{t-1} + 0.167(Z_{t-1} - Z_{t-2}) + a_t - 0.678a_{t-1}$$

We calculate the values of ACF & PACF for the residuals of the model ARIMA (1,1,1) and shown in the two figures (3-20) for ACF and figures (3-21) for PACF and wenoted that those coefficients lie within the confidence interval and this refers to the random model

Table (3-5)

The models suggested for the cotton yields data series

Fit Statistic	MODELS					
	ARIMA(1,1,0)	ARIMA(2,1,0)	ARIMA(0,1,1)	ARIMA(0,1,2)	ARIMA(1,1,1)	ARIMA(1,1,2)
RMSE	83.811	83.027	80.937	81.165	80.127	81.728
MAPE	27.551	25.917	25.000	24.618	24.609	24.727
MAE	55.952	53.713	52.974	53.339	51.430	53.416
NBIC	8.918	8.960	8.948	8.914	8.913	8.989
Fit Statistic	**MODELS**					
	ARIMA(2,1,1)	**ARIMA(2,1,2)**	**ARIMA(3,1,0)**	**ARIMA(0,1,3)**	**ARIMA(3,1,1)**	**ARIMA(3,1,3)**
RMSE	81.728	83.537	82.741	81.683	82.211	81.103
MAPE	24.765	25.942	24.970	24.863	24.683	25.688
MAE	53.372	53.005	53.032	53.406	53.297	54.635
NBIC	8.989	9.093	9.014	8.988	9.061	9.156

Fit Statistic	MODELS		
	ARIMA(1,1,3)	ARIMA(2,1,3)	ARIMA(3,1,2)
RMSE	81.330	80.746	82.754
MAPE	24.299	24.666	24.725
MAE	51.812	51.697	52.978
NBIC	9.040	9.086	9.135

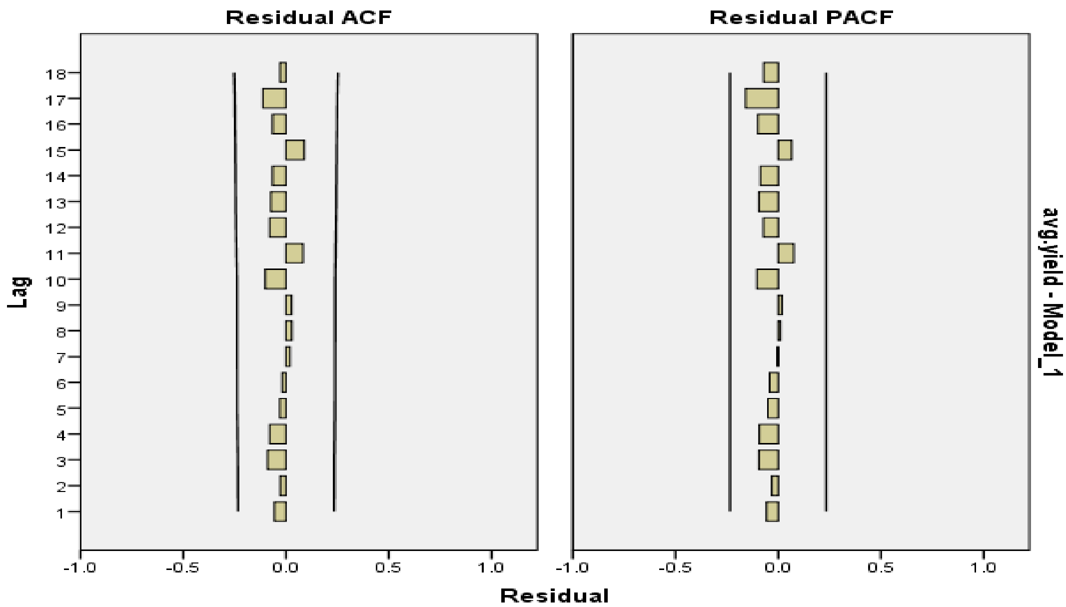

Figure (3-20)
ACF for residualsARIMA(1,1,1)

Figure (3-21)
PACF for residuals ARIMA(1,1,1)

The following two tests $Q_{L\alpha B}$ = 6.958 and $Q_{B\alpha P}$ = 6.876 have been calculated and when compared with the tabular value to x^2 in d.f (16) and significance level 0.05 is equals (7.961). In other words, the calculated value is smaller than the tabular, this refers to that no significant differences and diagnosed model is appropriate data series.

The forecasting values of the series as a result of the application of the selected model are shown in the table (3-6).

Table (3-6)
The forecasting values with the probability terms and significance interval 95% for the selected model and cotton yield series in Iraq

Year	Lower Limit	Forecasting	Upper Limit
2012	383.9094	545.7955	707.6817
2013	350.9293	530.7336	710.5379
2014	337.2456	528.2462	719.2468
2015	326.9770	527.8354	728.6939
2016	317.6240	527.7676	737.9112

4. Conclusion & Recommendations

4.1. Conclusion

1. The three time series data of cotton are drawn according to (area, production and yield) to know the data nature and behavior of the data and it is clear the fluctuation of the data and non stationary in variance as well as in mean. And A.C and P.A.C coefficients have been drawn and values calculation it is clear that the series are non-stationary and the values of their coefficients lie within the two confidence intervals {± 0.233} with confidence level 95%.
2. The cotton area time series become stationary after taking the natural logarithm and the first difference of data and the ACF and PACF lie between the two confidence intervals. And ARIMA (2,1,2) model is the best one and thus has the least values in criteria (RMSE, NBIC, MAE, and MAPE). The formula is equal to:

$$(1- 0.715\, B + 0.747B^2)\,(1-B)Z_t = (1-0.892B + 0.592B^2)\, a_t$$

   Using the model accuracy tests show that the model is appropriate and randomly.
3. The predicting values of cotton area series (2012- 2016) were referring to the observed fluctuations and the proportion in comparison with 2011 were (2.7, - 1.9, 11.9, 13.5 and 9.7).
4. The time series for cotton production became stationary after taking the natural logarithm and the first difference of data and the two functions ACF and PACF lie between the two confidence terms. And the model ARIMA(2,1,2) is the best one suggested to give it the least values in the following criteria (RMSE, NBIC, MAE, and MAPE) and the diagonized model is

$$\mathrm{Z_t} = \mathrm{Z_{t-1}} + 0.482(\mathrm{Z_{t-1}} - \mathrm{Z_{t-2}}) - 0.694(\mathrm{Z_{t-2}} - \mathrm{Z_{t-3}}) + \mathrm{a_t} - 0.743\mathrm{a_{t-1}} + 0.528\, \mathrm{a_{t-2}}$$

   Using the model accuracy tests for randomness, we show that the model is appropriate and randomly.
5. The predicting values for the cotton production series (2012- 2016) were referring to the observed fluctuations and the proportion in comparison with 2011 were(-28.3,-32.1,-16.5, -0.3, -3.5).
6. The data of cotton yield becomes stationary after taking the first difference of data and ACF & PACF lie between the two confidence intervals with confidence level 95%.And ARIMA (1,1,1) model is the best one and thus has the least values in criteria (RMSE, NBIC, MAE, MAPE). The formula is equal to:

$$Z_t = Z_{t-1} + 0.167(Z_{t-1} - Z_{t-2}) + a_t - 0.678a_{t-1}$$

   Using the model accuracy tests show that the model is appropriate and randomly

7. The cotton series predicting values (2012-2016) refer to the oscillation observed and the proportions compared with 2011 (-14.3,-17.2,-17.2, -17.1 & - 16.7).

8. The oscillation observed in the predicting values for cotton crop was a result of land salinity desertification, erosion and the problem of drowning the local market with imported agricultural products and non sufficiency of the requirements of production (Chemical Fertilizers, agricultural machines, high production seeds) to cover the -vegetarian production.

4.2. Recommendations

1. A comparison between Carrying out a comparison between Box-Jenkins and other predicting methods like markov chains and multiple regression especially that there are many factor affecting the production increasing.
2. Conducting research on the most important factors affecting the production of cotton crop.
3. Studies predictive seasonal time-series for agricultural sector data.

References

[1]. Anderson, T.W. (1971) The statistical Analysis of Time Series – John Wily& Sons.

[2]. Box, George E.P and Jenkins, Gwilym M.and Reinsel,Gregory C. (2008) Time Series Analysis: Forecasting and Control - 4th Edition - John Wiley& Sons.

[3]. Box, G.E.P and Pierce, David A., (1970) Distribution of Residual Autocorrelation in Autoregressive – Integrated Moving Average Time series Models– JASA,Vol. 65,No.332,PP.(1509-1526).

[4]. Fuller,Wayne A. (1976) Introduction to Statistical Time Series – John Willy &Sons, Inc.

[5]. Hamilton,James. D. (1994) Time Series Analysis – Princeton University Press, New Jersey.

[6]. Ljung,G.M.and Box,G.E.P (1978) On aMeasure of Lack of Fit in Time Series Models – Biometrical, Vol. 65, N0.2, PP. (297-303).

[7]. Makridakis Spyros and other, (1998) Forecasting: method and application–3rd. John Wiley and Sons&Inc.

[8]. Wei,William W.S (1990) Time Series Analysis Univarite and Multivariate Methods – Addison,Wesley publishingcomp.

Published: Volume 2016, Issue 11 / November 25, 2016