DOI: 10.24996/ijs.2024.65.2.22

ISSN: 0067-2904

The Impact of Overall Intervention Model on Price of Wheat

Nihad S. Khalaf Aljboori [1*], Hiba H. Abdullah[1*] , Nooruldeen A. Noori[2*]

[1]*Mathematic, College of Education for Women, Tikrit University, Tikrit, Iraq*

[2]*Mathematic, Anbar Education Directorate, Anbar, Iraq*

Received: 23/11/2022 Accepted: 7/3/2023 Published: 29/2/2024

Abstract

The analysis of observed data at different time points is prominent to single problems when the scientist studied random sampling data over time restricts the applicability of many traditional statistical methods that require random sampling. The analysis of this data is usually mentioned to the stochastic process, in a special case, if it relates to time, then it is called time series [1]. The aim of our study is to focus on the case of observations that are made at certain times of the year, specifically in the autumn season, in equal periods of time, and we estimate the effect of interference on wheat prices n order to obtain the best predictive model and the best order of this model for the overall effect on the price of wheat for the coming years.

Keywords : - Time series; Smoothing; Forecasting; Seasonal; Stationarity; Autoregressive (AR); Moving Average Model (MA); ARMA.

تأثير نموذج التدخل الشامل على سعر القمح

نهاد شريف خلف الجبوري[1]*, هبة هاني عبدالله[1]*, نورالدين اياد نوري الصعب[2]*

[1]الرياضيات, كلية التربية للبنات, جامعة تكريت, تكريت, العراق

[2]الرياضيات, المديرية العامة لتربية الانبار, الانبار, العراق

الخلاصة :

تحليل البيانات المرصودة في نقاط زمنية مختلفة بارزة لمشاكل فردية. وهو عندما العلم عند دراسة بيانات أخذ العينات العشوائية بمرور الوقت يحد من إمكانية تطبيق العديد من الأساليب الإحصائية التقليدية التي تتطلب أخذ عينات عشوائية. عادة ما يتم ذكر تحليل هذه البيانات بالعملية العشوائية وكحالة خاصة عندما يتعلق الأمر بالزمن تسمى متسلسلات الزمنية [1].

سوف تركز دراستنا على حالة المشاهدات التي تم إجراؤها في أوقات معينة من العام ، وتحديداً في فصل الخريف ، في فترات زمنية متساوية ، ونقوم بتقدير تأثير التداخل على أسعار القمح لكي نحصل على افضل أنموذج تنبؤي وافضل رتبة لهذا الأنموذج للتأثير الشامل على سعر القمح للسنوات المقبلة.

*Email: nooraldeen.nan@gmail.com

1- Introduction

The purpose of this section is to introduce the aim of this study and some basic concepts and definitions to get a full understanding of the study.

1.1 *The aim of this study*

The study aims to provide an adequate analysis of wheat price data in dollars to give future predictions for these prices.

1.2 *Time Series*

It is a phenomenon that is represented by a mathematical model of a stochastic process $\{x_t\}$ consisting of a family or a sequence of variables with the index t symbolizing time because t belongs to the indicative group T, which may be a continuous set of observations $-\infty \leq T \leq \infty$, so it is called a continuous time series. For example the river flow, temperatures, and concentration of an element in a chemical reaction. The indicative group T can also be in a discrete set $T = 0, \mp 1, \mp 2, \ldots$, so it is called discrete time series, for n examples, the population of a particular city, production of a company , and exchange rates between two different currencies that means the observations may be continuous or discrete set.[1]

There are amount of effects elements that are noticed in time series. The greatest significance of these elements can be seen in [2]:

Table 1: Elements of time series

Items	**Description**
Smoothing	The perceived (Y_1) are expected as the outcome of noise rates (ε_t) that are added to polluting a smooth indication(Π_1). $Y_t = \Pi_t + \varepsilon_t$
Modeling	We can wish to progress a regular mathematical simulation that will explain the observed model of $(X_1, X_2, \ldots)$ this pattern may depend on unidentified parameters and these will essential to be assessed.
Forecasting	On this observation $(X_1, X_2, \ldots, X_T)$, We may ask to predict any value of X_{T+P}, it will be $(P \geq 1)$, and it probably makes a proposal for the uncertainty in predicting.
Control	We can interfere with the procedure that makes values (X_t) in such an approach that the possible values change to produce a favorable result

The time series is supposed to be affected by four main components that affect the observed data, which are as follows:[1]

- Trend: It means the general tendency of the series to increase or decrease on the $x - axis$
- Seasonal: Seasonal changes are fluctuations within a year or during the season.
- Cyclical Variation**:** It describes the periodic variance in the time series over the medium term in the successful series according to the conditions that are repeated in cycles.
- Irregular Components**:** The lack of regularity in a series of times is the result of unprecedented effects and it is neither condemned nor emanating in a certain pattern and these differences are caused by other signed incidents.

1.3 Stationarity and Non-Stationarity

It is a characteristic of coincidences, which is based on the assumption that the coincidence process is in a state of statistical balance, and it is said that the coincidence process $\{x_t\}$ "strictly stationary" if the random variables $(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$ and $(x_{t_{m+1}}, x_{t_{m+2}}, \ldots, x_{t_{m+k}})$ have the same distribution of joint probability to any $k \in \mathbb{N}$ and $m \in \mathbb{Z}$, which is a rare or a weak condition in our work in practical. Therefore, studies are directed and paid attention to a state of stability with a satisfactory, and the weakness of complete stationarity. The weak stationarity or the process $\{x_t\}$ is stationary from the rank m can be defined as follows: If the joint moment function is from the rank m for the vector $< x_1, x_2, \dots, x_k >$ equal and exists to the same vector after h of steps $< x_{1+h}, x_{2+h}, \dots, x_{k+h} >$ for any $k \in \mathbb{N}$ and $m \in \mathbb{Z}$.

The stability of the time series requires the second rank, namely the following two basic conditions:

- Stationary in the mean, it is one of the weak stationary conditions $E(x_t) = \mu = \text{constant}\ for\ all\ t.$
- The variance is fixed to all the values of t.

In contrast to these two conditions, the series is considered unstable. It is indicated that most of the series is unstable, and its instability occurs in different ways. It is possible that the time series is unstable in mean or variance, or both [3].

It is clear that not all time series that we encounter are stationary, although many time series are related in simple methods of stationary series to two important cases: Firstly, the trend models detect a set of deterministic trends and constant noise. The usual form of this model is $x_t = \beta_0 + \beta_1 t + \varepsilon_t$. Secondly, the integrated models which are the observation time series $x_{t+1} + x_t = \varepsilon_{t+1}$ where (ε_t) is stationary series. The main pattern which is important in this case is a random walk [4].

1.3.1 Autoregressive (AR)

An autoregressive (AR) model forecasts future behavior established on past behavior. It is operated for forecasting while there is some correlation among the values in a time series and the values which lead to and accomplish them. We only use historical or previous data to expect the behavior, therefore the identify autoregressive (AR). The process is essentially a linear regression of the data in the present series versus one or more previous assessments in the same series [5].

In an autoregressive (AR) model, the rate of the result alterable (x) at a certain theme (t) in time is a similar "regular" linear regression that is straight associated with the forecaster variable(y). Where the regular linear regression and the autoregressive (AR) models vary that means (x) is dependent on (y) and preceding rates of (x) [5].

The autoregressive (AR) procedure is a sample of a stochastic method, which has grades of indecision or uncertainty built in. Autoregressive (AR) models are also named as following conditional methods, Markov methods, or transition methods [4].

In 1926, Yule was considered one of the first researchers to study the stationarity time series. He studied the model of regression and the complete world of Walkers in 1931 to reach the world model of Auto-regression. The general formula for order (p) of this model is AR(p) [3].

The AR(p) model is described by the equation:

$$x_t = b_1 x_{t-1} + \cdots + b_p x_{t-p} + z_t \qquad (1)$$

Where $\mathrm{b_1, b_2, \dots, b_p}$ are the parameters of the model and $\mathrm{z_t}$ is a white noise process. There are countless developed cases of this model, for example, Fourier's autoregressive model.[14]

1.3.2 Moving Average Model (MA)

The stochastic process $\{x_t\}$ is said to be a moving average model of order (q) and is denoted by MA(q) if the equation below is satisfied:

$$x_t = \tau + z_t + \vartheta_1 z_{t-1} + \cdots + \vartheta_q z_{t-q}. \qquad (2)$$

Or it can be written the equation (1.3.2) as follows:

$$x_t = \tau + \theta(B) z_t.$$

Where B is a shift operator, τ is arbitrary constant and $\theta(B) = 1 + \vartheta_1 B + \cdots + \vartheta_q B^q.$ [6].

It is stated that the stability of the model according to the Wold is always satisfied because θ(B) is a polynomial[8]. However, Hamilton says that the inevitability of the model must be checked by checking whether the roots of the polynomial θ(B) lie outside the unit circle that means there is no condition for stability but it is limited by the condition of reversibility. [7]

1.3.3 Auto Regressive - Moving Average Models (ARMA)

The time series $\{x_t\}$ of the autoregressive model and the moving average of order (p, q) can be defined as a linear combination of the autoregressive model of the order p AR(p) and the moving average of order (q) MA(q) .Its equation can be written as:

$$.x_t = b_1 x_{t-1} + \cdots + b_p x_{t-p} + \vartheta_1 z_{t-1} + \cdots + \vartheta_q z_{t-q} + z_t \qquad (3)$$

Where $z_t \sim \text{iid } N(0, \sigma^2)$ and $(a_1, a_2, \ldots, a_p, \theta_1, \theta_2, \ldots, \theta_q)$ are constants[6].

The best example of scaling the autoregressive moving averages model (ARMA) is the autoregressive model integrated moving averages model which is known as (ARIMA), and the general formula of this model of order (p, d, q) is ARIMA(p, d, q) which is described by form. [6],[13]

$$\left(1 - \sum_{i=1}^{p} a_i B^i\right)(1 - B)^d x_t = \left(1 + \sum_{j=1}^{q} \beta_j B^j\right) z_t, \qquad (4)$$

so that q, d and p are integers greater than or equal to zero and refer to the autoregressive, integral and moving average parts of the model, respectively. [6]

1.4 Prediction in time series

After validating the suitability of the model, it finds a prediction of the studied data, in particular, phenomenon in the future values.

1.4.1 Exponential smoothing

The subject of the exponential smoothing of important statistical and semantic procedures that address confusion or random errors may be defined as a process of refinement of data that is confusing, and it is a type of estimation process that has proved successful in many cases. Therefore, one of the most important methods in predicting time series is considered. The double exponential smoothing (DES) model is one of the methods that study the prediction under the exponential smoothing list. It gives the previous observations weights of unequal values, as these weights are decreasing exponentially from the most recent data points [9].

The double exponential smoothing model is described by the following equation:

$$\check{Y} = \left(2 + \frac{\alpha}{1-\alpha} l\right) S_t' - \left(1 + \frac{\alpha}{1-\alpha} l\right) S_t''.$$

Where S_t^r is the local smoothing level in time (t) and S_t'' is the smoothing level in the general direction of the time series.

1.4.2 Intervention

Suppose that at the time $t = T$, where T will be recognized, there was interference towards the time series. By intervention, we mean a change in the procedure, law, politics, etc., which aims to change the values of the x_K chain. We need to assess the amount of intervention that changed the series (if any). For the sample, assume that the region has created a new maximum speed on its roads and requirements to learn the amount of the new limit that guarantees accident rates [10].

The intervention analysis in the time series is raised to the analysis of how to mean the level series after the intervention, as soon as the ARIMA structure is expected the x_t series that keeps them before and after the intervention.

Overall Intervention Model

Assuming that ARIMA model for x_i (The monitored series) with no intervention is

$$x_t - \mu = \frac{\theta(B)}{\Phi(B)} \dot{\omega}_t.$$

Where $\dot{\omega}_t$ is the usual assumptions on error storage, $\theta(B)$ and $\Phi(B)$ are polynomials of MA and AR, respectively.

Let $Z_t$is the amount of change at the time due to the intervention. By definition, we have $E(Z_t) = 0$ before the time T (Intervention time).

The value of Z_t may or may not be zero after time T[10]. Then the general model, including the effect of the intervention, maybe a pattern as follows:

$$x_t - \mu = z_t + \frac{\theta(B)}{\Phi(B)} \dot{\omega}_t.$$

2. Methodology

The effect of the intervention by using the mean absolute deviation (MAD), the mean square error (MSE) and mean absolute percentage error (MAPE) of the wheat product in our paper is stuied. Annual data of wheat prices were obtained from site Federal Reserve Bank (www.fred.stlouisfed.org) in period of (Jan 1980 - July 2017) [11].

2.1 The Mean square error

The sum of the mean square error is divided by the number of time series observations, it is given by:

$$MSE = \frac{\sum a_t^2}{n},$$

where a_t^2 are errors that represent the real values of the time series subtracted from the estimated values[12].

2.2 The Mean Absolute Deviation

The sum of the mean absolute deviation of the error divided by the number of observations of the time series, and its equation is given by:

$$MAD = \frac{\sum |a_t|}{n}, \qquad [12]$$

2.3The Mean Absolute Percentage Error

It is defined as follows:

$$MAPE = \frac{\sum PE_t}{n-m},$$

where $PE = \left(\frac{Y_t - Y_t'}{Y_t}\right) * 100$, Y_1 is true values, and Y_t' is the estimated values.[12]

2.4 Wheat annual prices data

The annual data on wheat prices are obtained from the site Federal Reserve Bank (www.fred.stlouisfed.org) in period of Jan 1980 - July 2017 [11].

Figure1: Global price of Wheat (source: www.fred.stlouisfed.org)

Units: US dollars for a per-metric ton
Frequency: Annual

The value represents the standard prices that represent the global market. It is determined by the largest source of a specific commodity. The prices are the average of the period in nominal US dollars.

3. Data Analysis

The time series is converted to (0) and (1) where the effect of the interceptor event is a sudden start and a temporary effect of the intervention. The intervention was of the pulse type, where the intervention was due to the high wheat prices as shown in Figure 2.

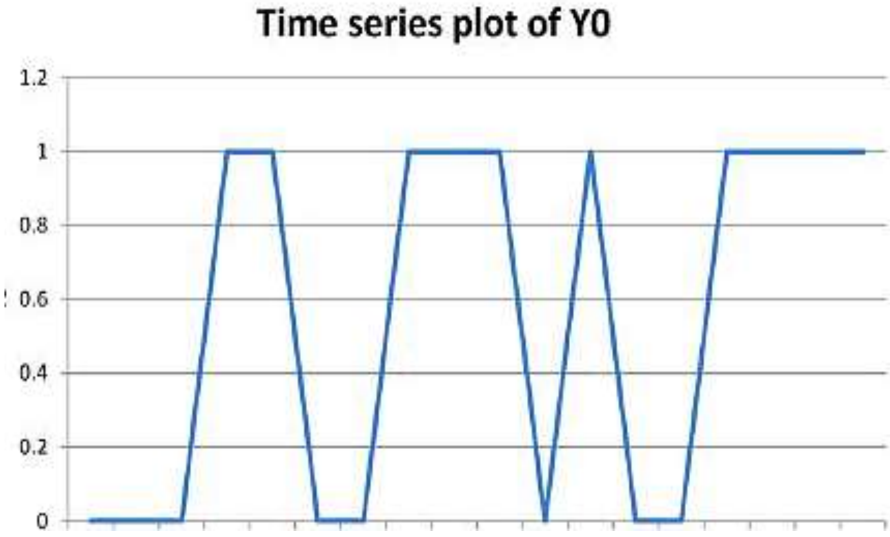
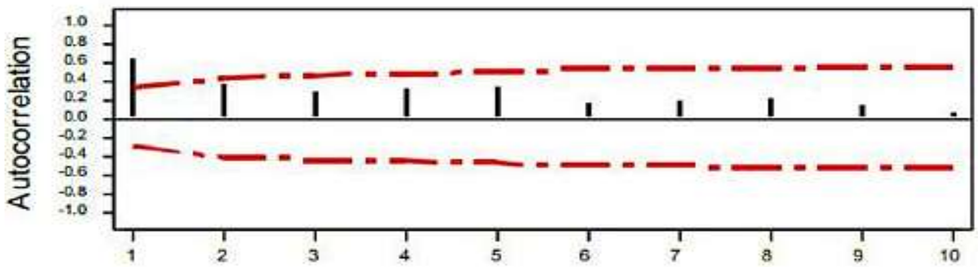


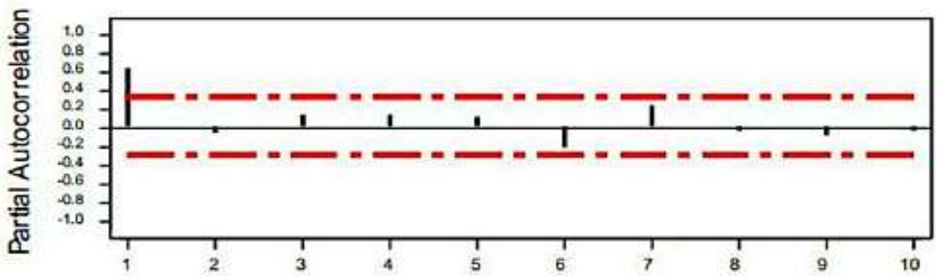
Figure 2: Time series plot of $\boldsymbol{Y0}$

Based on the computer applications of the wheat price series to obtain the values of the impact of the intervention, which was chosen according to the criteria of Mean Square Error ($MSE = 0.194$), Mean Absolute Deviation ($MAD = 0{,}393$), Mean Absolute Percentage ($MAPE = -6.122$) and value of ($\Phi = 0.668$).

We use Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) to determine the suitable models as it is shown in Figure 3.



(a)



(b)

Figure 3: (a) Autocorrelation Function for Y_0 (b) Partial Autocorrelation Function for Y_0

While Tables 2 and 3 represent the values of Autocorrelation and partial Autocorrelation as well as $T - test$ values for ten lags, respectively.

Table 2- of Autocorrelation and $\boldsymbol{T-test}$ values, and LBQ for $\boldsymbol{Y_0}$

Lags	Corr	$T-test$	LBQ
1	0.63	4.11	18.17
2	0.37	1.77	24.39
3	0.29	1.32	28.48
4	0.32	1.37	33.38
5	0.34	1.41	39.20
6	0.17	0.66	40.69
7	0.19	0.76	42.69
8	0.22	0.84	45.28
9	0.14	0.55	46.45
10	0.07	0.27	46.75

Table 3-of Partial Autocorrelation and $\boldsymbol{T-test}$ values for $\boldsymbol{Y_0}$

Lags	Corr	$T-test$
1	0.63	4.11
2	−0.06	−0.39
3	0.14	0.92
4	0.14	0.94
5	0.11	0.71
6	0.21	−1.37
7	0.24	1.58
8	−0.03	−0.23
9	−0.10	−0.64
10	−0.03	−0.17

Figure 3 shows that the time series was the (AR1) model and the coefficients of the selected model were estimated by the program. The value of ($\Phi = 0{,}668$), which has the lowest value of ($MSE = 0.14609$), where the intervention model is as follows:

$$Y_t = wP_1^T$$

Where Y_t outcome of the intervention, W unknown information value, and P_1^T Pulse function.

For the purpose of checking the suitability of the selected model, we draw the Partial Autocorrelation Function (PACF) for errors as in Figure 6.

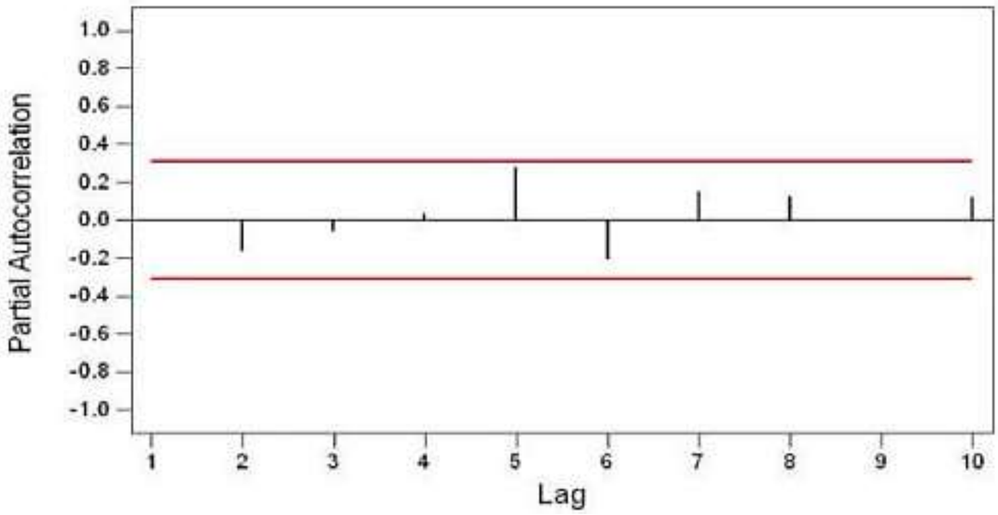


Figure 6: Partial Autocorrelation Function (PACF) for errors

4. Conclusion

❖ The best model that is obtained for prediction is AR1.
❖ The value of ($\Phi = 0{,}668$), which has the lowest value of ($MSE = 0.14609$).
❖ the intervention model is as follows:

$$Y_t = wP_1^T$$

❖ From the plot of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of the study and the effect of the intervention, it is shown that the series of model AR1 was the model chosen:

$$Y_t = (0.6680)S^T$$

❖ Using some predictive parameters (MSE, MAD), it was found that the effect of the intervention is more efficient for the AR1 model which is the best to predict the time series because it has the least value of (MSE, MAD).

References

[1] C. Chatfield, "*The analysis of time series: An introduction, sixth edition, 6th ed. UK*: *Chapman and Hall/CRC,* 2003.
[2] S.-M. Chow, L. Ou, J. F. Cohn, and D. S. Messinger, "Representing self-organization and nonstationarities in dyadic interaction processes using dynamic systems modeling techniques," in Innovative Assessment of Collaboration, Cham: Springer International Publishing, 2017, pp. 269–286.
[3] S.-M. Chow, J. Zu, K. Shifren, and G. Zhang, "Dynamic factor analysis models with time-varying parameters," *Multivariate Behav. Res*., vol. 46, no. 2, pp. 303–339, 2011.
[4] R. Dahlhaus, "Fitting time series models to nonstationary processes," *Ann. Stat*., vol. 25, no. 1, pp. 1–37, 1997.
[5] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *J. Am. Stat. Assoc.*, vol. 74, no. 366a, pp. 427–431, 1979.
[6] E. Ferrer, J. S. Steele, and F. Hsieh, "Analyzing the dynamics of affective dyadic interactions using patterns of intra- and interindividual variability," *Multivariate Behav. Res.*, vol. 47, no. 1, pp. 136–171, 2012.

[7] J. D. Hamilton, *Time Series Analysis. Princeton, NJ: Princeton University Press*, 2020.
[8] Wei, William. "*Time Series Analysis: Univariate and Bivariate*." (2006).
[9] F. Sylvia, and F. Sylvia, *Finite mixture and Markov switching models*. *Springer New York*, 2006.
[10] E. L. Hamaker, R. M. Kuiper, and R. P. P. P. Grasman, "A critique of the cross-lagged panel model," *Psychol. Methods*, vol. 20, no. 1, pp. 102–116, 2015.
[11] Bank, Federal Reserve. Federal Reserve Bank. [Online] 2018. (www.fred.stlouisfed.org).
[12] R. S. Tsay, *Analysis of financial time series*, *2nd ed. Newy York: Wiley-Interscience*, 2005.
[13] A. Ali Mudhir, "Mixing ARMA models with EGARCH models and using it in modeling and analyzing the time series of temperature," *Iraqi J. Sci*., pp. 2307–2326, 2021.
[14] A. I. Taiwo, T. O. Olatayo, A. F. Adedotun, and K. K. Adesanya, "Modeling and forecasting periodic time series data with Fourier Autoregressive model," *Iraqi J. Sci*., pp. 1367–1373, 2019.