Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Using Time Series Models to Predict the Numbers of People Afflicted with (COVID-19) in Iraq, Saudi Arabia and United Arab Emirates

Mohammed Habeb Al-Sharoot[1] Habib Kazem Alwan[2]

Department of Statistics , College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq [1,2]
Mohammed.alsharoot@qu.edu.iq[1] *Hebabaljnaby71@gmail.com*[2]

ARTICLE INFO

Article history:
Received: 25 /11/2020
Rrevised form: 21 /12/2020
Accepted : 07 /01/2021
Available online: 11 /01/2021

Keywords:

Box- Jenkins models, ARIMA model, Coronavirus, time series analysis.

ABSTRACT

Covid-19 disease is an infectious disease caused by the newly discovered Coronavirus. There was no knowledge of this virus before an outbreak broke out in the Chinese city of Yuhan in December 2019. The Corona epidemic has caused the world to go through a major challenge as it has claimed the lives of many people and also disrupted the economy in most countries of the world. This has prompted many researchers in various disciplines to conduct studies and research to stand in the face of this epidemic. It is known that statistical methods have great importance for all sciences The other that stood against this epidemic.In this paper, we use time series ARIMA models by Box-Jenkins to predict the numbers of people afflicted with (COVID-19) in Iraq, Saudi Arabia and United Arab Emirates and compare them based on a daily time series represent the numbers of people afflicted in those countries for the period from 3/15/2020 to 4/5/2020 the emergence of that epidemic in those countries.

MSC. 41A25; 41A35; 41A36.

DOI : https://doi.org/10.29304/jqcm.2020.12.4.729

1. Introduction

Corona viruses are a wide range of viruses that may cause disease in animals and humans. It is known that a number of coronaviruses cause human respiratory diseases in severity ranging from common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and severe acute respiratory syndrome (SARS). The newly discovered Coronavirus causes Covid-19 disease.Box-Jenkins models are considered an important statistical method as they are used to represent time series data for a specific phenomenon and to predict its future values provided that the series is stationary and highly correlated. These models have been used in various economic, financial and medical sectors, and so on, as forecasting and decision-making are important matters in the planning process in all Domains.The objective of this paper is to use time series ARIMA models by Box- Jenkins to predict the numbers of people afflicted with (COVID-19) in Iraq, Saudi Arabia and United Arab Emirates and compare them based on a daily time series represent the numbers of people afflicted in those countries for the period from 3/15/2020 to 4/5/2020 the emergence of that epidemic in those countries.

∗Corresponding author : Mohammed Habeb Al-Sharoot ,Habib Kazem Alwan
Email addresses: Hebabaljnaby71@gmail.com
Communicated by : Alaa Hussien Hamadi

2. Autoregressive Integrated Moving Average Model (ARIMA)

Box, Jenkins,(1976) , described the model comprehensively and put together the method or approach of the information associated with understanding and treating stationaraty in the data and reached the model called autoregressive models and integrated moving averages in the event that the series is unstable, it can be converted into a series Stable by taking the differences of degree (d), (d = 1,2,…)and denoted by the symbol (ARIMA) and the form of the rank (p, d, q) written in the form of ARIMA (p, d, q) as the following form:

$$\emptyset(B)(1-B)^d Z_t = \theta(B)a_t, \qquad (1)$$

Where:

d: The degree of difference taken in the time series

B: backshift operator

When taking the appropriate differences to convert the non-stationary time series into stationary, then the previous model can be written as follows:

$$\emptyset(\mathrm{B})w_t = \theta(\mathrm{B})a_t, \qquad (2)$$

$$w_t = (1-B)^d Z_t, \qquad (3)$$

The time series is stationary if its data fluctuates around a constant average of the series, that is, there is no change in its mean and its variance, and therefore, the stationary series has mean and variance that does not depend on time t,

$$\mu = E(Z_t)$$

$$\sigma^2 = Var(Z_t)$$

$$= E(Z_t - E(Z_t))^2$$

The time series is either strictly stationary if the common distribution of observations $Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}$ is the same distribution for observations $Z_{t_{1+p}}, Z_{t_{2+p}}, \dots, Z_{t_{m+p}}$, this means that the distribution depends on the time period between the observations of the time series and not on the value of the real series observations, and thus the Zt series is a completely stationary series if it is

$(Z_1, Z_2, \dots, Z_n) = (Z_{1+K}, \dots, Z_{n+k})$ and for all the correct k values and at n≥1 where (=) indicates that the random vectors have the same common distribution function. The series is stationary from the second degree, meaning it has a weak stability (Weakly Stationary) if the next half is achieved the expected value to Zt is constant for all values of t_i. The covariance matrix for the variables $Z_{t_1}, \dots, Z_{t_n}$ is the same as the covariance matrix for the variables $(Z_{t_{1+k}}, \dots, Z_{t_{n+k}})$, this means that the change function depends on the time interval between the observations, i.e

$$Cov(Z_t, Z_{t+k}) = \gamma_{(k)}$$
$$= E((Z_t - \mu)(Z_{t+k} - \mu) \qquad , k = 0, \pm 1, \pm 2, \dots$$

Where k is the lag , which is the time between observations.

3. Autocorrelation Function:

The self-correlation is an indication of the strength of the relationship between the values of the variable itself at different lags (k), and its value ranges between (-1, 1) which is denoted by ρ_k, that is : -

$$-1 \leq \rho_k \leq 1$$

Where ρ_k compute by using the following formula:

$$\rho_k = \frac{E\,(Z_t - \mu)(Z_{t+k} - \mu)}{E\,(Z_t - \mu)^2}, \qquad k = 1,2,\dots,\frac{N}{4} \qquad ,\dots (4)$$

Where

Zt: time series value at time t

Z_{t+k}: observed value after lag k

N: - The size of the time series.

μ: - represents the mean.

Whereas, plotting the autocorrelation coefficients with lag k, then is called Autocorrelation Function.

4. Partial Autocorrelation Function:

It is an indicator that measures the relationship between (Z_t) and (Z_{t+k}) for the same series, assuming that the series values are fixed and can be calculated according to the following formula:

$$\rho_j = \emptyset_{k1}\,\rho_{j-1} + \emptyset_{k2}\,\rho_{j-2} + \cdots + \emptyset_{kk}\,\rho_{j-k}\,, For\ j = 1,2,3,\dots,k, \qquad (5)$$

when plotting the partial autocorrelation coefficients (ϕkk) with lags (k), then the is called partial autocorrelation function.

5. Box- Jenkins Model building:

a- Identification: the first step in time series analysis is to draw it to determine whether the time series is stationary or non-stationary as they fluctuate around several averages, seasonal effects, or outliers values.

For this purpose we compute the autocorrelation and partial autocorrelation coefficients. To represent the time series data by the models ARIMA (p, d, q) we must diagnosis the order of p,d,q which can be made using several methods as follows:

i. Akaike Information(AIC):

It is a standard used to diagnose the degree of the ARMA model (p, q) and according to the following formula

$$AIC = n\ln(\hat{\sigma}_a^2) + 2m, \qquad (6)$$

ii. Schwarz Criterion(SIC):

$$BIC(m) = n\ln(\hat{\sigma}_a^2) + m\ln(n)\,,, \dots (7)$$

-

b- Estimation: after the proposed model that represents the time series data under study has been identified and the appropriate rank has been determined for it, the parameters of the chosen model are estimated and often the main reason for estimating the model is to use it to calculate future predictions of the time series, there are several methods of estimation

- Maximum likelihood

To estimate the parameter parameters (ARMA) (p, q), we use the maximum and the aggregate function of the parameters by validating the observations, which are

$$L(\emptyset,\theta,\sigma_a \backslash Z) = (2\pi\sigma_a^2)^{-n/2} \left|M_n^{(p,q)}\right|^{1/2} exp\left\{-\frac{S(\emptyset,\theta)}{2\sigma_a^2}\right\}, \qquad (7)$$

$$\left(M_n^{(p,q)}\right)^{-1}\sigma_a^2 = \Gamma_n$$

$$S(\emptyset,\theta) = \sum_{t=1-Q}^{n} [at\backslash Z,\emptyset,\theta]^2 , \qquad (8)$$

- The Method Of Moments
- Ordinary Least Squares (OLS)

c- Diagnostic Checking: For the purpose of diagnostic checking the proposed model that representing the time series data, we must calculate the residual is according to following : -

$$a_t = Z_t - \hat{Z}_t, \qquad (9)$$

Which should be random, unbound variables, by testing then the null hypothesis: $H0{:}\,\rho = 0$ against the alternative hypothesis $H1{:}\,\rho \neq 0$. There are several tests as follow:

1. Box-Pierce (Q)

In 1970 Box and Pierce reached a statistic by which the ARIMA (p, d, q) model diagnostic validity could be tested and assuming that we had m of the estimated autocorrelations of the residuals $r_k(\hat{a})$ that distributed a normal distribution with a mean of zero and variance 1 / N it is misfit

$$Q = n\sum_{k=1}^{m} r_k^2(\hat{a}) \sim x_{(m-p-q)}^2, \qquad (10)$$

where, n: represents the number of observations for the identified model, n = N - d

N: original number of time series observations

d: represents the numer of differences taken to achieve the stationarity

m:represent √ (n)

Then Q calculated is compared with (X^2) tabular with a degree of freedom (m-p-q). If the calculated Q is smaller than tabular, it does not reject the null hypothesis, that is, random errors are not correlated and therefore the model is appropriate and good, but if it is larger than the model is inappropriate and in this case the stage must be repeated The first is to diagnose another model to represent the time series, estimate its parameters, and check it.

2. Box-Ljung (Q)

Ljung & Box modified the original Q-test formula proposed by (Box & Pierce) as follows:

$$Q = n(n+2)\sum_{k-1}^{m}(n-k)^{-1} r_k^2(\hat{a}) \sim x_{(m)}^2 , \qquad (11)$$

and they proved that they have the advantage in use, because are close to the expected values.

3. If the residuals autocorrelation coefficients are within the confidence limits of the 95% confidence level, $-1.96\frac{1}{\sqrt{n}} \leq \hat{r}_k(\hat{a}) \leq 1.96\frac{1}{\sqrt{n}}$, (12)

d- Forecasting: One of the primary objectives of time series analysis is prediction. When diagnosing the model and estimating its parameters and then the stage of examining the relevance of the model to time series data, it becomes ready to use for prediction as it is appropriate and matches the original data if it has the minimum mean squares of the prediction error

For example, if we want to predict the value of the time series in the period (t + L), which is $\hat{Z}$t(L), this value is calculated by taking the conditional prediction of (Z) at time (t + L).

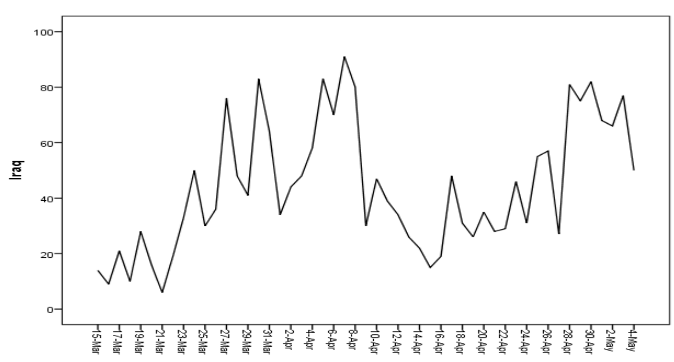
6. The application side:

Data was collected, of three time series, each series consists of (51) observations , for the period from 3/15/2020 to 4/5/2020, and that these data represent the numbers of people afflicted with coronavirus disease in Iraq, Saudi Arabia and the United Arab Emirates , taken from data sheets World Health Organization.

i- Coronavirus afflicted series for Iraq:

At this stage, data is prepared by drawing the time series, evaluating the autocorrelation and partial correlation coefficients, as well as the confidence limits of the autocorrelation function of the original data to know the behavior of that data, using the statistical program ( ), through Figure (1), which represents the time series data for the number of injuries in Iraq we note an increasing trend with time and that the variance tends to be stable, which indicates that the series is non-stationary and to make the series to be stationary ,we take the first difference as shown in Figure (3).

Thus, it is noted that there is no trend, no seasonal effects, and all autocorrelation coefficients for the sample within the confidence limits $(-0.27 \leq r_k \leq 0.27)$as in Figure (2) and test the significance of the coefficients for the autocorrelation function using (Ljung & box) after taking the first difference so its value (24.161) was less than the tabular at the significance level(0.05) of (24.996) so we accept the null hypothesis.



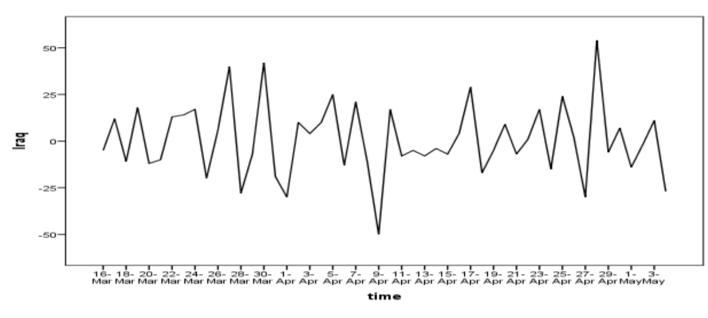


Figure (1): original time series of coronavirus (covid-19) to Iraq. figure(2): first different time series of coronavirus (covid-19) to Iraq .

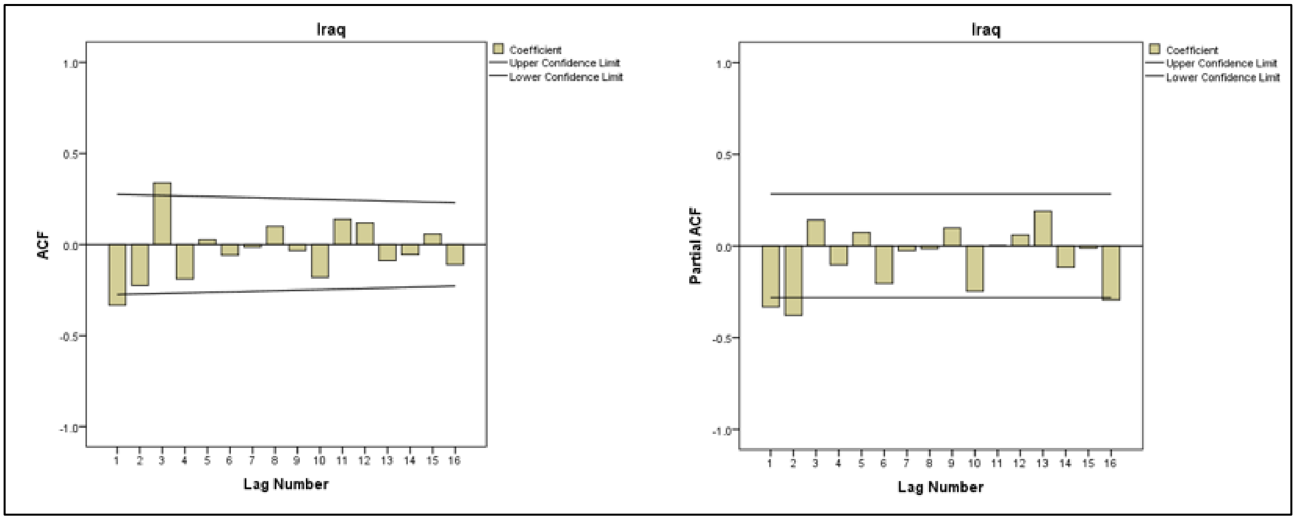


Figure (3): autocorrelation and partial autocorrelation coefficients after taking the first difference.

Table (1):

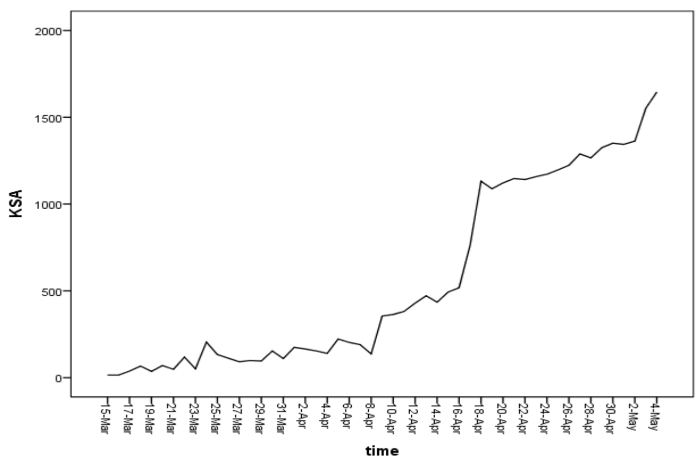
Lag	Partial Autocorrelations		Autocorrelations				
	Partial Autocorrelation	Std. Error	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
					Value	df	Sig.[b]
1	-.333	.141	-.333	.137	5.879	1	.015
2	-.378	.141	-.225	.136	8.629	2	.013
3	.141	.141	.337	.134	14.905	3	.002
4	-.103	.141	-.190	.133	16.945	4	.002
5	.074	.141	.024	.132	16.979	5	.005
6	-.205	.141	-.059	.130	17.183	6	.009
7	-.026	.141	-.015	.129	17.196	7	.016
8	-.015	.141	.099	.127	17.806	8	.023
9	.097	.141	-.035	.126	17.884	9	.037
10	-.246	.141	-.180	.124	19.991	10	.029
11	.001	.141	.138	.122	21.256	11	.031
12	.061	.141	.118	.121	22.213	12	.035
13	.190	.141	-.088	.119	22.763	13	.045
14	-.116	.141	-.055	.118	22.984	14	.061
15	-.013	.141	.058	.116	23.230	15	.079
16	-.294	.141	-.110	.114	24.161	16	.086

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

ii- Coronavirus afflicted series for Saudi Arabia

The second time series related to the number of injuries in Saudi Arabia, we note an increasing trend with time as

shown in Figure (4) so we can say that the series is non-stationary in mean , the first difference was taken the series to make it stationary as in Figure (6) and all the autocorrelation and partial autocorrelation coefficients of the sample within Confidence limits as in Figure (5) and the significance of the autocorrelation and partial correlation coefficients using the (Ljung & BOX) test where the test demonstrated its relevance.



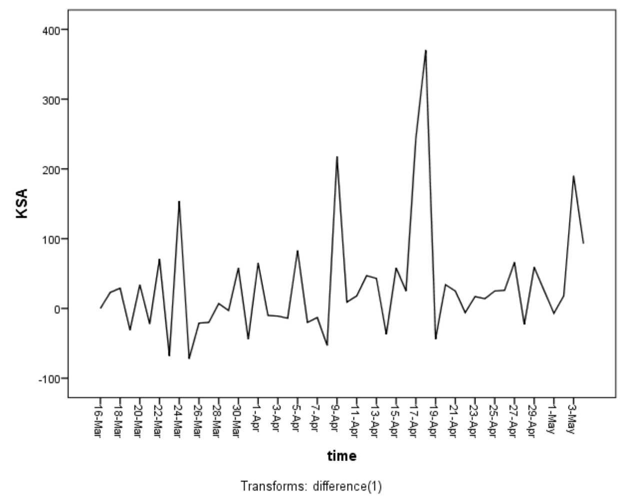
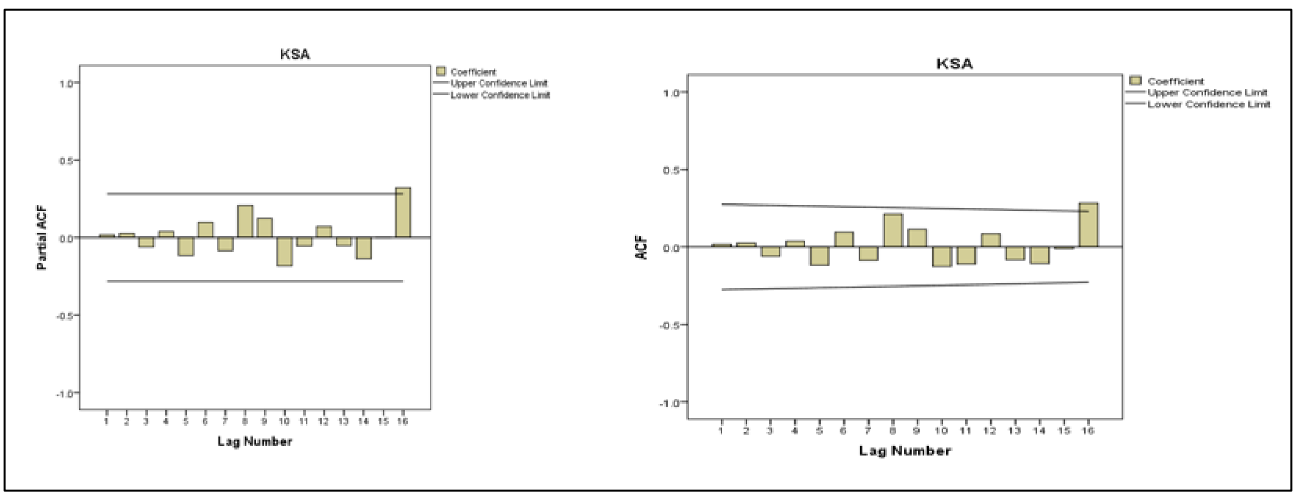


Figure (4): original time series of coronavirus(covid-19) to KSA. Figure(5): first different time series of coronavirus(covid-19) to KSA Table (2):

a. The underlying process assumed is independence (white noise).

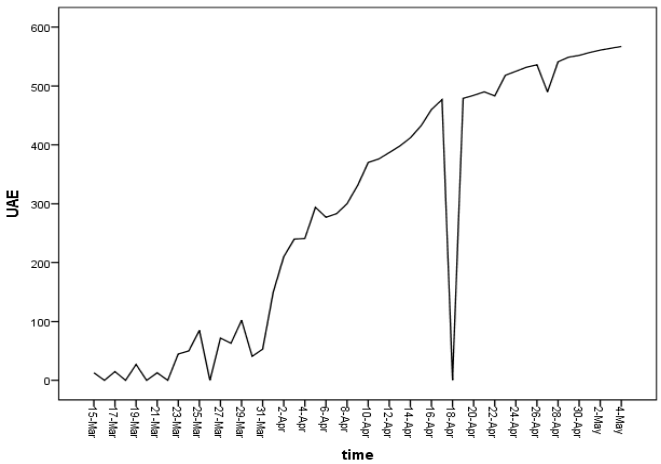
Lag	Partial Autocorrelations		Autocorrelations				
	Partial Autocorrelation	Std. Error	Autocorrelation	Std. Error[a]	Box-Ljung Statistic		
					Value	df	Sig.[b]
1	.017	.141	.017	.137	.015	1	.901
2	1	.141	.026	.136	.053	2	.974
3	-.061	.141	-.060	.134	.252	3	.969
4	.040	.141	.038	.133	.336	4	.987
5	-.116	.141	-.117	.132	1.124	5	.952
6	.097	.141	.096	.130	1.664	6	.948
7	-.086	.141	-.088	.129	2.128	7	.952
8	.207	.141	.212	.127	4.903	8	.768
9	.124	.141	.115	.126	5.738	9	.766
10	-.182	.141	-.125	.124	6.755	10	.748
11	-.054	.141	-.109	.122	7.548	11	.753
12	.071	.141	.084	.121	8.034	12	.782
13	-.053	.141	-.084	.119	8.530	13	.807
14	-.138	.141	-.106	.118	9.347	14	.808
15	-.002	.141	-.012	.116	9.357	15	.858
16	.320	.141	.282	.114	15.439	16	.493

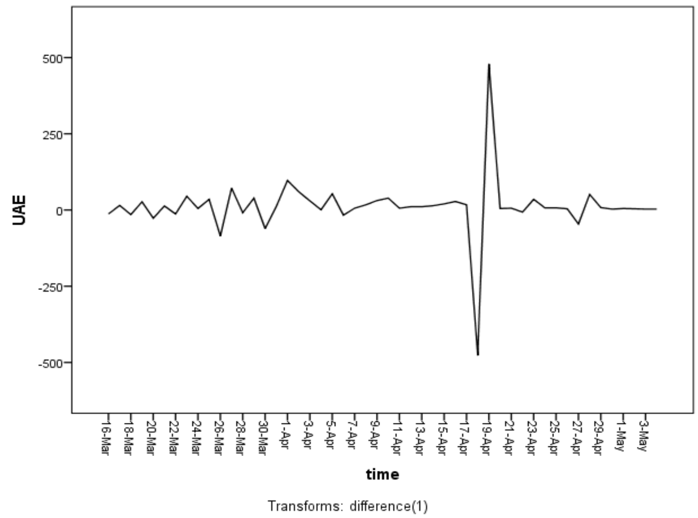
b. Based on the asymptotic chi-square approximation.



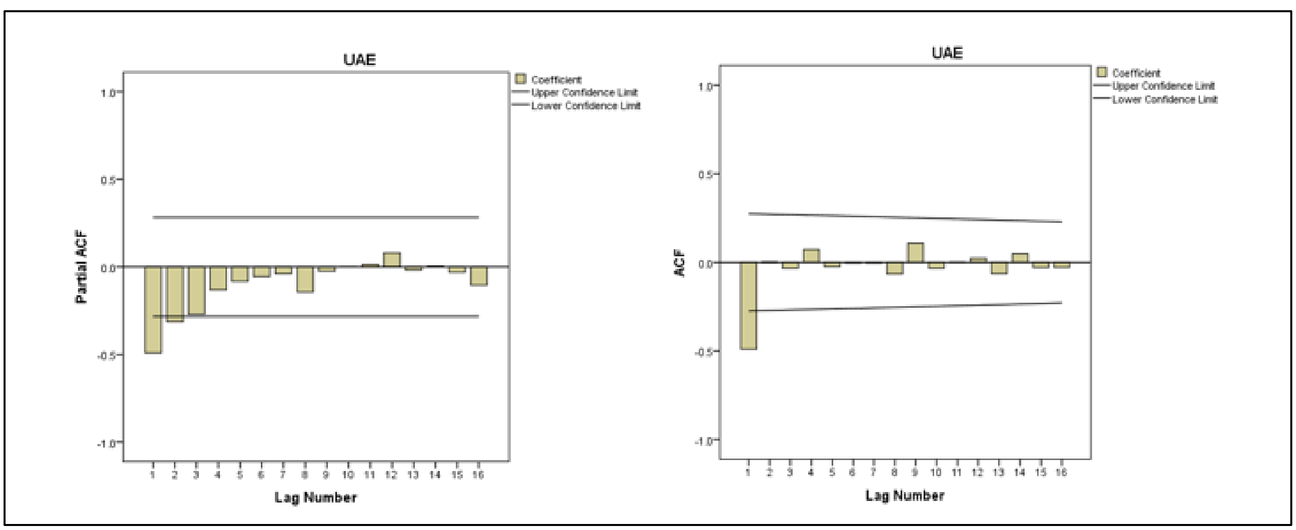
Figure(6): autocorrelation and partial autocorrelation coefficients after taking the first difference iii- Coronavirus afflicted series for United Arab Emirates

the third time series related to the number of injuries in the United Arab Emirates, where an increasing trend was observed as shown in Figure (7) and thus the series is non-stationary , the first difference was taken to the series to make it stationary in mean as in Figure (9),we note that all the autocorrelation and partial correlation coefficients of the sample within the confidence limits as in Figure (8) and the significance of the autocorrelation and partial correlation coefficients using the (Ljung & BOX) test, where the test demonstrated its relevance.





Figure(7): original time series of coronavirus(covid-19)to UAE . Figure(8): first different times of coronavirus(covid-19)to UAE.



Figure(8): autocorrelation and partial autocorrelation coefficients after taking the first difference.

Table(3):

Lag	Partial Autocorrelations		Autocorrelations				
	Partial Autocorrelation	Std. Error	Autocorrelation	Std. Error[a]	Box-Ljung Statistic Value	df	Sig.[b]
1	-.490	.141	-.490	.137	12.762	1	.000
2	-.312	.141	.003	.136	12.763	2	.002
3	-.270	.141	-.034	.134	12.827	3	.005
4	-.131	.141	.073	.133	13.125	4	.011
5	-.081	.141	-.024	.132	13.159	5	.022
6	-.055	.141	-.004	.130	13.160	6	.041
7	-.040	.141	-.004	.129	13.161	7	.068
8	-.143	.141	-.066	.127	13.428	8	.098
9	-.025	.141	.108	.126	14.171	9	.116
10	-.002	.141	-.034	.124	14.246	10	.162
11	.015	.141	.001	.122	14.246	11	.220
12	.079	.141	.023	.121	14.282	12	.283
13	-.018	.141	-.064	.119	14.572	13	.335
14	.004	.141	.048	.118	14.737	14	.396
15	-.030	.141	-.029	.116	14.801	15	.466
16	-.103	.141	-.028	.114	14.860	16	.535

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

Diagnosis:

The first step in the construction stages of the time series model is to diagnose the model. Diagnostic criteria have been applied that depend on the curve shape of the sample partial autocorrelation function (ACF) and the shape of the partial autocorrelation function curve (PACF) and when matching the values of the autocorrelation and partial autocorrelation coefficients of the time series After taking the first difference with theoretical behavior, a function curve (ACF) is observed that gradually decreases with increasing displacement periods K.

1. The appropriate form for the first time series whose data represent the number of casualties in Iraq is ARIMA (2,1,2)

2. The appropriate model for the second time series whose data represents the number of cases of coronary disease in Saudi Arabia is ARIMA (1,1,1).

3. The appropriate model for the third time series whose data represent the number of cases of corona disease in the United Arab Emirates is ARIMA (0,1,1).

Table(4):

Model Description			Model Type
Model ID	Iraq	Model_1	ARIMA(2,1,2)
Model ID	KSA	Model_1	ARIMA(1,1,1)
Model ID	UAE	Model_1	ARIMA(0,1,1)

Table (5):

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.	
Iraq-Model_1	0	.315	17.317	37.438	13.002	6.095	16.610	14	.278	0
KSA-Model_1	0	9.185E-5	83.023	31.786	53.386	9.073	15.763	16	.470	0
UAE-Model_1	0	.398	79.366	26.976	44.814	8.905	4.578	17	.999	0

Estimation

After verifying the suitability of the model, testing the significance of the parameter, and testing the homogeneity of variance, the next step comes from the stages of building a model of time series is to estimate the models for those series and by applying the most estimate methods where the following results were obtained:

1. Estimation of Iraqi Time Series

Table (6)

ARIMA Model Parameters								
					Estimate	SE	t	Sig.
Iraq-Model_1	Iraq	Square Root	Constant		.082	.105	.783	.438
			AR	Lag 1	-.943	.246	-3.834	.000
				Lag 2	-.370	.245	-1.507	.139
			Difference		1			
			MA	Lag 1	-.607	.259	-2.340	.024
				Lag 2	.282	.265	1.064	.293

Model		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Iraq-Model_1	ACF	-0.02	0.01	-0.03	0.03	-0.04	-0.06	0.07	-0.02	-0.12	-0.12	0.13	0.19	-0.01	-0.08	-0.18	-0.18	-0.24	0.10	-0.10	-0.03	-0.02	-0.01	-0.05	0.16
	SE	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.17

Model		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Iraq-Model_1	PACF	-0.02	0.01	-0.03	0.03	-0.04	-0.06	0.07	-0.02	-0.12	-0.12	0.13	0.20	0.00	-0.10	-0.23	-0.22	-0.25	0.08	-0.12	-0.05	0.05	0.01	-0.16	0.06
	SE	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14

Model Fit											
Fit Statistic	Mean	SE	Minimum	Maximum	Percentile						
					5	10	25	50	75	90	95
Stationary R-squared	.315	.	.315	.315	.315	.315	.315	.315	.315	.315	.315
R-squared	.477	.	.477	.477	.477	.477	.477	.477	.477	.477	.477
RMSE	17.317	.	17.317	17.317	17.317	17.317	17.317	17.317	17.317	17.317	17.317
MAPE	37.438	.	37.438	37.438	37.438	37.438	37.438	37.438	37.438	37.438	37.438
MaxAPE	149.304	.	149.304	149.304	149.304	149.304	149.304	149.304	149.304	149.304	149.304
MAE	13.002	.	13.002	13.002	13.002	13.002	13.002	13.002	13.002	13.002	13.002
MaxAE	44.791	.	44.791	44.791	44.791	44.791	44.791	44.791	44.791	44.791	44.791
Normalized BIC	6.095	.	6.095	6.095	6.095	6.095	6.095	6.095	6.095	6.095	6.095

2-Estimation of Saudi Arabia Model

Table(7):

KSA-Model_1	KSA	No Transformation	Constant		32.663	12.621	2.588	.013
			AR	Lag 1	.842	4.232	.199	.000
			Difference		1			
			MA	Lag 1	.836	4.296	.195	.000

	Model	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
KSA-Model_1	ACF	.011	.022	-.064	.035	-.120	.094	-.091	.211	.114	-.127	-.111	.084	-.085	-.108	-.012	.283	-.034	.032	-.050	.005	-.106	-.043	-.153	.093
	SE	.141	.141	.142	.142	.142	.144	.146	.147	.153	.154	.156	.158	.159	.160	.161	.161	.171	.171	.171	.171	.171	.173	.173	.176

	Model	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
KSA-Model_1	PACF	.011	.022	-.065	.037	-.119	.094	-.089	.206	.125	-.182	-.055	.070	-.053	-.139	-.004	.320	-.146	.012	.099	-.057	-.139	-.026	-.005	-.134
	SE	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141	.141

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile		
					5	10	
Stationary R-squared	.398	.	.398	.398	.398	.398	.398
R-squared	.864	.	.864	.864	.864	.864	.864
RMSE	79.366	.	79.366	79.366	79.366	79.366	79.366
MAPE	26.976	.	26.976	26.976	26.976	26.976	26.976
MaxAPE	281.172	.	281.172	281.172	281.172	281.172	281.172
MAE	44.814	.	44.814	44.814	44.814	44.814	44.814
MaxAE	460.754	.	460.754	460.754	460.754	460.754	460.754
Normalized BIC	8.905	.	8.905	8.905	8.905	8.905	8.905

3.Estimation of United Emirates Arabia Model :Table(8):

ARIMA Model Parameters								
					Estimate	SE	t	Sig.
United Emirates Arabia Model	Estimation	Square Root	Constant		12.585	2.202	5.716	.000
			Difference		1			
			MA	Lag 1	.825	.090	9.161	.000

Model		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UAE Model_1	ACF	0.01	0.01	0.01	0.08	0	-0.02	-0.04	-0.05	0.08	-0.01	-0.03	-0.05	-0.12	-0.06	-0.1	-0.09	-0.08	0.08	0.07	-0.07	-0.02	-0.06	0.07	-0.12
	SE	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15

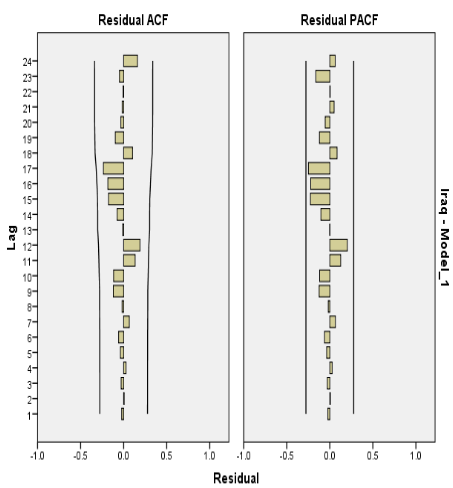
Model		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UAE Model_1	PACF	0.01	0.01	0.01	0.08	0	-0.02	-0.04	-0.05	0.08	0	-0.02	-0.04	-0.13	-0.06	-0.1	-0.08	-0.05	0.08	0.08	-0.07	-0.03	-0.08	0.04	-0.12
	SE	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14

Model Fit							
Fit Statistic	Mean	SE	Minimum	Maximum	Percentile		
					5	10	
Stationary R-squared	9.185E-5	.	9.185E-5	9.185E-5	9.185E-5	9.185E-5	9.185E-5
R-squared	.977	.	.977	.977	.977	.977	.977
RMSE	83.023	.	83.023	83.023	83.023	83.023	83.023
MAPE	31.786	.	31.786	31.786	31.786	31.786	31.786
MaxAPE	217.755	.	217.755	217.755	217.755	217.755	217.755
MAE	53.386	.	53.386	53.386	53.386	53.386	53.386
MaxAE	336.265	.	336.265	336.265	336.265	336.265	336.265
Normalized BIC	9.073	.	9.073	9.073	9.073	9.073	9.073

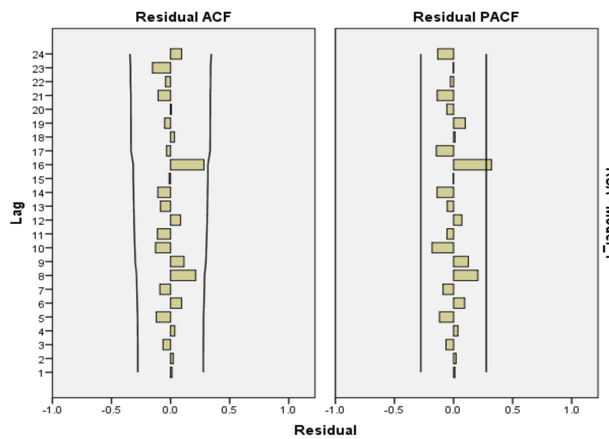
The test

For the random residual series test, the autocorrelation and partial autocorrelation coefficients of the estimated residues were calculated as shown in Figure (10), all the coefficients ($r_k(\hat{a})$) fall within the confidence limits ($-0.27 \leq r_k(\hat{a}) \leq 0.27$). For the purpose of making sure of the suitability of the model, the test statistics (Ljung & Box) were applied and showed that the tabular value is greater than the calculated value and for all models, and this indicates the randomness of the residuals of these models.

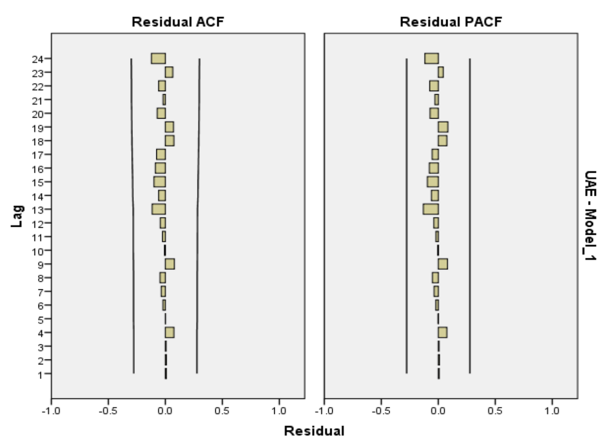
1



2



3

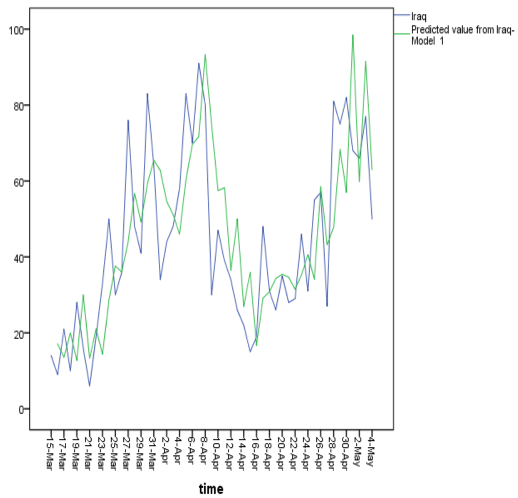


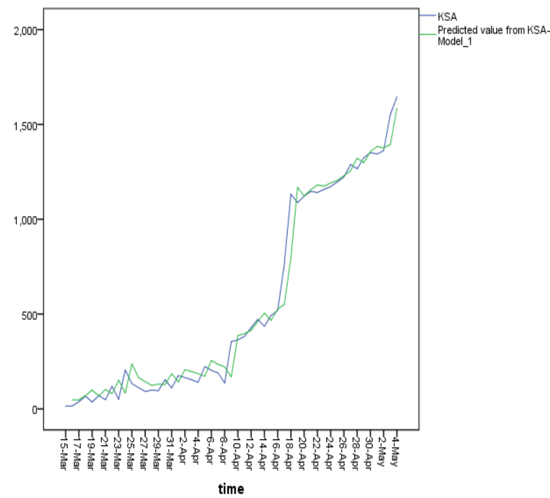
Figure(10): residual Autocorrelation and partial Autocorrelation cofficients for the three models (1- Iraq Residuals series 2-Saudi Arabia Residuals series 3. United Emirates Arabia)

Forecasting

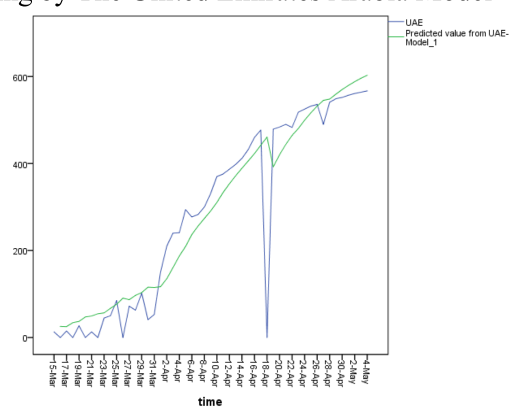
To predict numbers of people with coronavirus and using validated models, each of these models has been drawn in figure(11):

1. Forecasting by The Iraq model : 2. Forecasting by Saudi Arabia Model





3. Forecasting by The United Emirates Arabia Model



Figure(11): forecasting number of coronavirus for the three series in Iraq, Saudi Arabia and United Emirates Arabia

7. Conclusions:

1. We note, through the study of the three time series, including the number of cases of corona disease in these countries, that they are non-stationary time series on mean and that there were all have a clear trend.

2. stationary was achieved in all of these chains after taking the first differences and after matching the parameters of autocorrelation and partial autocorrelation with the theoretical behavior of the autocorrelation and partial autocorrelation functions.

3. By comparison criteria it was found that

• The first time series model for the number of cases of corona disease in Iraq is ARIMA (2,1,2) which is considered the best and can be used for prediction.

• The second time series model for the number of cases of corona disease in Saudi Arabia is ARIMA (1,1,1) and can be used to predict

•The third time series model for the number of cases of corona disease in the UAE is ARIMA (0,1,1) and it can be used to predict.

References:

[1] Bisgaard, S. and Kulahci, M. (2011),"Time Series Analysis and Forecasting By Example ", Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

[2] Box G, E.P & Jenkins, G. M.,(1976), "Time series analysis forcasting and control sanfrancisco Helden-day" .

[3] Brockwell, P.J. & Davis, R.A. (1987). Time series. Theory and methods, Springer.Cao, R. (1999). An overview of bootstrap methods for estimating and predicting in time series, Test, 8, 1, 95-116.

[4] Chatfield, (1975), "The analysis of time series, An introduction" second edition chapman and hall , London-New york

[5] Cryer, J.D., and Chan, K.S. (2008), "Time Series Analysis With Applications in R", 2nd ed., Springer, New York

[6] Fan,J.,and Y., Qiwei.(2003) .Nonlinear Time Series Nonparametric and Parametric Methods, "Springer, New York,USA".

[7] McLeod, A.I. & Li, W.K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations, Journal of Time Series Analysis, 17, 571—599.