Original Research   **Open Access**

Forecasting of COVID-19 Cases in Kurdistan Region Using Some Statistical Models

Shekhmous Hassan Hussen

Department of Animal Production, College of Agricultural Engineering Sciences, University of Duhok, KR, Iraq
Email: sheikhmous.hussein@uod.ac

Abstract

Nowadays the new universal disease of the coronavirus that is called the epidemic COVID-19 is spread as geometric progression among the people around the world, so, such pathogen considered the most dangerous threat facing humanity. This study aimed to derive the best forecasting models for the close future cases of infected, recovered, and deaths in the four provinces of Kurdistan Region-Iraq to avoid more loss of human lives by applying more health care in certain province. Two forecasting methods were used including Exponential Smoothing and ARIMA models. The results indicate that both ARIMA and Exponential Smoothing models were close to each other for predicting the infected cases of COVID-19 in Kurdistan Region provinces, and the predicting models show that the pandemic might not be under control unless the people apply the government instructions for health care and keep social distances.

Keywords: Forecast; COVID-19; Kurdistan region; Statistical models.

1. Introduction

The term COVID-19 is coming from COrona VIrus Disease that has been identified in China, 2019. COVID-19 is an infectious animal and human disease caused by coronavirus group two [1], which infects the respiratory system of humans and causes infection like Pneumonia, and has become as a pandemic disease by World Health Organization (WHO) due to its fast spread (like geometric progression) around the world. Moreover, this disease has no vaccine and its genetic material (RNA) mutate periodically. The first infected case has been reported in China on November 17, 2019 [2], and after about eight months (on July 21, 2020) the total cases reached 14.7 million around the world with about 8.29 million recovered cases and 610,000 deaths [3]. In Iraq the infection, recovered and death cases on July 21, 2020, reported as 97003, 62836, and 3969 individuals, respectively. While in Kurdistan region, North of Iraq, the cases at the same date reported as 11367, 6044, and 419 individuals, respectively [4, 5]. However, the highest cases reported in some countries (such as Spain, Italy, and China) in April 2020; while the USA had a continuous peak after the last date [4]. This disease imposes a quarantine on the people around the world, which resulted in restrictions on travel and gatherings, in addition to the closure of schools, institutes, universities, companies, and factories. As a result, this pandemic disease has caused huge problems for governments and people. So, it is necessary to study the prediction of its growth or decline to know the possibility of controlling it soon due to its continuous and dramatic increase.

For forecasting such cases, often two traditional models are used; Exponential Smoothing and ARIMA models within the time series modeler procedure. In a recent study conducted by Abotaleb [6] on prediction cases of COVID-19, that reported in the USA, Italy, and China, he used both ARIMA and exponential growth models; and concluded that the exponential model is better than ARIMA for predicting the COVID-19 cases. This author added that the number of confirmed and death cases in China will continue to increase insignificantly while the number of recovered cases will increase significantly; and the best ARIMA models for China will be ARIMA (0,2,1), ARIMA (0,2,2), and ARIMA (0,2,2) for the confirmed, recovered and death cases, respectively. Regarding the comparison between ARIMA and exponential model, the author stated that the exponential model resulted in expected values higher than the ARIMA model values significantly in Italy cases. Finally, he expected that the USA will record the highest number of confirmed, recovered and death cases compared to the other two countries. In another study carried out by Benvenuto, et al. [7] on COVID-19 prediction, they performed the Auto Regressive Integrated Moving Average model (ARIMA) to predict the trend of the incidence and prevalence of COVID-19. They concluded that both the prevalence and incidence of COVID-19 are not influenced by the seasonality. They added that more data were needed to obtain a trusted trend, and they expected that the spread of COVID-19 will slightly

decrease as well as the cases will reach a plateau if the virus does not mutate. In another similar study conducted by Perc, et al. [8] on predicting COVID-19 cases for four countries (USA, Slovenia, Germany, and Iran) for two successive weeks. Therefore, they used an exponential growth method and concluded that daily case rates should be kept below 5%, and this may be achieved by restricting people's behavior. In another study by Chakraborty and Ghosh [9] on forecasting the daily infected cases of COVID-19 in Canada, France, India, South Korea, and the UK in which they used ARIMA, Wavelet-based Forecasting (WBF) and hybrid ARIMA-WBF models. They proposed that the last model can explain the non- linear and non-stationary behavior present in the studied datasets of COVID-19 cases. They added that the fitted ARIMA models for studied datasets were; ARIMA(1,2,1), ARIMA(1,1,2), ARIMA(0,1,1), ARIMA(2,1,0) and ARIMA(2,2,2) for India, Canada, France, South Korea and UK, respectively. Also, several previous investigations predicted the future COVID-19 cases for China using traditional time series forecasting models [10-12].

The present study aims to derive the best forecasting models for infected cases in the four provinces of Kurdistan region (Erbil, Sulaymaniyah, Duhok, and Halabja), to avoid more loss of human lives by applying more health care in certain governorates.

2. Methodology

The present study was carried out on the daily and accumulative data of infected, recovered and death cases of COVID-19 in the four provinces (Erbil, Sulaymaniyah, Duhok, and Halabja) of Kurdistan region (KR-Iraq) during the period from March 1, 2020, up to July 31, 2020 (153 days). The raw data were obtained from the official website of Kurdistan region government [5]. The objective was to predict the future of COVID-19 cases for upcoming days (half of the previous period, which is about 76 days). The study was performed using the two forecasting methods (ARIMA and Exponential Smoothing models) of time series analysis within SPSS software [13]. Such procedures resulted often in precise forecasting especially for short-run predictions on a relatively significant amount of raw data [14].

2.1. ARIMA Model

It is a classical time series analysis method, used for predicting linear tendencies in stationary data for time series (ARIMA is a non-stationary series enabling the differencing of the data series arising from the models by changing first-difference of yt). ARIMA is followed by (p, d, q); the p and q parameters represent the order of the Autoregressive (AR) and Moving-Average (MA) model, respectively, and d represents the level of differencing [9]. Mathematically, the ARIMA model can be expressed as the main model (Model 1) as follow:

$$\Delta y_t = c + a_1 \Delta y_{t-1} + a_2 \Delta y_{t-2} + \cdots + a_p \Delta y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

(Model 1)

Where $\Delta y_t = y_t - y_{t-1}$ = the change in actual observations at time t; εt is the random error at the same time t; c and a are ARIMA model's coefficients. ARIMA model assumes that the error tends to be zero and the variance is constant $((\varepsilon t) = 0$ and $(\varepsilon t) = \sigma 2)$; and satisfies the sequence of independently and identically distributed (iid) condition. In the case of p = q = 0, the ARIMA called the random walk model and described as ARIMA (0, 1, 0). Deriving an ARIMA model for given time series data could be summarized in three frequent steps: Identification of the Model (achieving stationary); estimation of the parameter (Autocorrelation function –ACF- and the partial autocorrelation function –PACF-), also plots are utilized to choose the values of both parameters (p and q); and diagnostics the model by checking the candidate model (s) for adequacy (finding the 'optimum' fitted predicting model by using Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC)) [15].

However, the ARIMA models that is a combination of three procedures; autoregressive (AR); the differencing; and the moving-average (MA), have been described by Abonazel and Abd-Elftah [16] as:

The Autoregressive form of order p (AR (p)) is write as:

$$y_t = c + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \varepsilon t \qquad \text{(Model 2)}$$

t =1, 2, 3…T

Where: c = constant; a1, a2, ap = coefficients; εt = the error $((\varepsilon t) = 0$ and $var\ (\varepsilon t) = \sigma 2)$; such model, called a long-term memory model.

While the Moving-Average (MA) model is time series (yt) and called the moving-average process of order q, MA (q), as:

$$y_t = \varepsilon t - \theta_1 \varepsilon_{t-1} - \theta_2\ \varepsilon_{t-2} - \cdots - \theta_q\ \varepsilon_{t-q} \qquad \text{(Model 3)}$$

This model is a short memory model because q errors only affect xt.

Moreover, Autoregressive Moving-Average (ARMA) Model is a procedure that has both p and q order (ARIMA (p, q)), it is a combination of both previous models, as follow:

$$y_t = c + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \varepsilon t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \qquad \text{(Model 4)}$$

Anyway, there is also the Box-Jenkins method [17] for analyzing time series. This method of analysis permit ARIMA models to derive the optimum fit model depending on the iteration of the previous values of a time series, until obtaining the satisfied model. So, this method has four stages instead of three, as follow:

Model identification: identifying series seasonality, insure the stationary of the variables as well as using the Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) plots for determination the best autoregressive (AR) or moving-average (MA) component in the series model.

Model estimation: using Maximum Likelihood Estimation (MLE) or non-linear least-squares estimation for best fit ARIMA model coefficients.

Model checking: checking whether the estimated model complies with the stationary univariate characteristics, where the residuals should be characterized as independent and constant in average and variance; and the residual ACF and PACF plots are indicating misspecification. So, if the resulted estimation is unsuitable, it should be returned to the first stage to drive a sufficient model. The last model will compare with other ARIMA models to select the best fit. However, two criteria used in model choosing: Akaike's Information Criterion (AIC) and Bayesian Information Criteria (BIC).

Forecasting: when the chosen ARIMA model is satisfied for a stationary univariate procedure in which it can be used for prediction.

2.2. Exponential Smoothing Model

This is the average of smoothing variability as a random process, and also its called a moving average of exponentially weighted, using Holt's Linear Trend Method [18]. This model has the following properties: (i) the weight is declined by older data; (ii) it is the simplest way for calculation and (iii) small size data is required. Also, it deals with the average current value by calculating the weighted average from the present value of a certain variable and the average value from the past. However, this model takes a weighted average of all past data and re-use it as the forecast of the final mean of the distribution; and create an exponential smoothing average to enable forecasting data. This way involves a prediction model and also two smoothing equations (level and trend equations):

The prediction model is as follow:

$$\ddot{Y}_{t+h|t} = l_t + hb_t \qquad\qquad \text{(Model 5)}$$

While the level equation is as:

$$l_t = \alpha\, y_t + (1-\alpha)\,(l_{t-1} + b_{t-1})$$

And the trend equation is as:

$$b_t = \gamma\,(l_t - l_{t-1}) + (1-\gamma)b_{t-1}$$

Where l_t denotes the series level estimation at the time, t; b_t is the estimate of the series slope at the same time, t; α is the factor of smoothing in level equation ($0 \le \alpha \le 1$) and γ is the smoothing parameter in trend equation ($0 \le \gamma \le 1$). For simple exponential smoothing, the level equation appears that the l_t is a weighted average of y_t observation and the simple predicted time will be as ($l_{t-1} + b_{t-1}$); while in the case of trend equation, the b_t is the weighted average based on ($l_t - l_{t-1}$) and b_{t-1} at time t. So, the prediction function changes into linear shape and equivalent to the previous value plus h (times) of the trend.

2.3. Running the Analysis

Forecasting for time series within SPSS program Ver. 26 [13] was used for running the analysis. Two forecasting models (ARIMA and Exponential smoothing) were applied on the data of accumulative daily infected cases of the four provinces (Sulaymaniyah, Erbil, Halabja, and Duhok) of Kurdistan region. The analysis was conducted from March 1, 2020, to July 31, 2020 (153 days); and the forecasting was carried out for the next interval that equal half of the previous period (76 days), which started from August 1, 2020, up to October 15, 2020. Descriptive statistics were computed for the daily cases, and Chi-square test was applied for recovered, deaths and gender proportions, using the same software.

3. Results and Discussion
3.1. Descriptive Statistics

The descriptive statistics of the daily data for the infected cases of the four studied provinces within the studied period are presented in Table (1).

It could be noticed from Table (1) that Sulaymaniyah province had the highest means of infected cases compared to other provinces. This is because the COVID-19 pandemic firstly begins in Sulaymaniyah. The values of both skewness and kurtosis and also standard deviation estimates indicating that the distribution of all provinces cases are non-normally distributed, therefore, the median values were taken to express the general view of such cases which stated that Sulaymaniyah province had the highest value of cases (6), followed by Erbil province (4); while both Duhok and halabja provinces were in the safe mode compared to the other two previous provinces. Consequently, such finding represents dangerous bell to attend more health care for both Sulaymaniyah and Erbil provinces.

Table-1. Descriptive statistics of the daily infected data

Province	Mean	Median	Std. Dev.	Min.	Max.	Skewness	Kurtosis
Erbil	28.1	4	47.354	0	211	2.256	4.847
Sulaymaniyah	55.26	6	76.062	0	286	1.263	0.507
Duhok	3.48	0	6.956	0	48	3.315	13.898
Halabja	3.1	0	5.836	0	36	2.257	6.443

Regarding to the percentages of recovered and death cases, the analysis of Chi-square showed that proportions of both were highly significant (p<0.01), where the chi-square values of recovered and death percentages were (39.34 and 58.24, respectively). Halabja province was recorded the highest recovered percentage (85.3 %) of infected cases; while Duhok province was recorded the lowest death percentage (0 %) of infected cases (Figure 1).

However, the overall expected percentages of recovered and deaths cases were (65.55 and 3.9 %, respectiverl). Moreover, the rest percent represent the undertreatment cases. Here, the people of both Sulaymaniyah and Erbil, should apply the instructions that followed by people of both Duhok and Halabja provinces to minimize the deaths cases and maximize the recovered ones.

In respect to the infected gender percentage, the Chi-square test showed that there is no significant difference ($p>0.05$) between the ratios of both sexes for the four provinces (Figure 2), where the Chi-square value was (6.7). But in general the male percentage was higher than that of female one, this may due to the more active movement for males than females in Kurdistan during the quarantine period.

The accumulated observed cases for the infected of the four provinces are illustrated in Figures (3), in which could be noticed that Sulaymaniyah province had the highest cases (8522) followed by Erbil province (4328); while both Halabja and Duhok provinces had the lowest numbers of cases (475 and 532, respectively). The trend of curves had fluctuated in similarity among the four provinces.

3.2. ARIMA model Results

The results of ARIMA models for the accumulated infected cases of the four provinces are presented in Table (2). Such models involve also sub-models like autoregressive model (AR) and moving average model (MA), and their significance levels, ARIMA models were used non-transformed data for all provinces except for Sulayamniah province where the square root way was used for transformation the data to be as normal distribution data. As shown from the Table (2), the best forecasted model of Erbil province is ARIMA (1,2,2) for the infected cases; while the parallel model of Sulaymaniyah is ARIMA (0,2,1), and this last model is also suitable for Duhok infected cases; but in respect to Halabja's infected cases, the model ARIMA (0,2,4) is reported as the best forecasted model.

Figure-1. Recovered and deaths percentages in the studied provinces
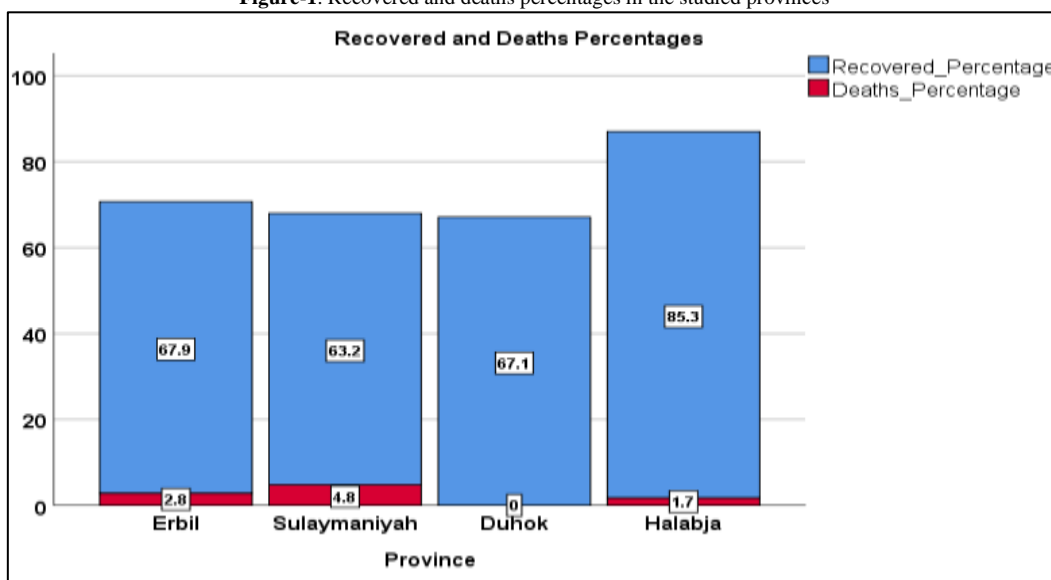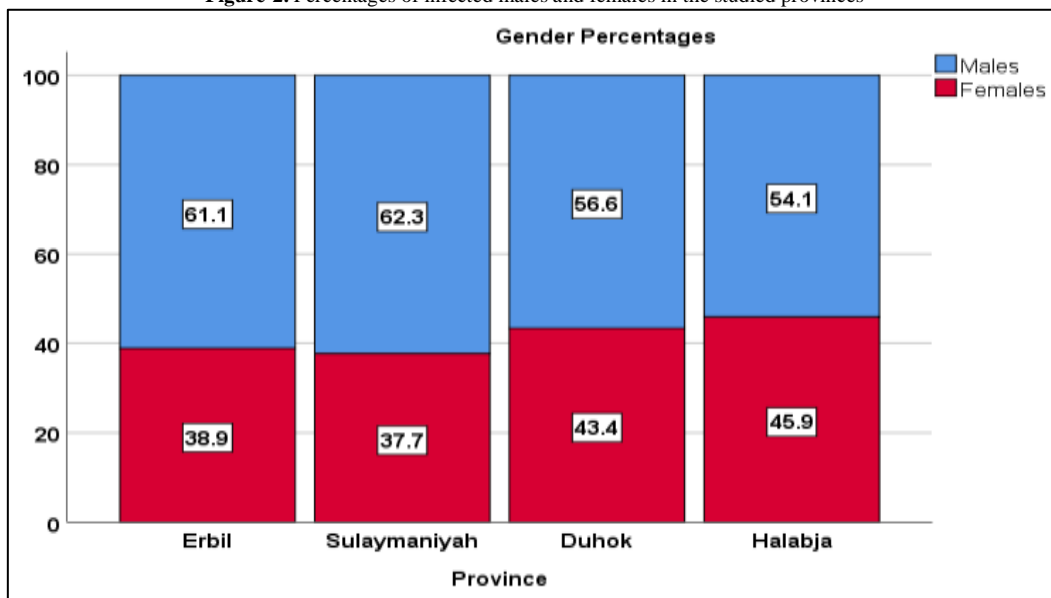

Figure-2. Percentages of infected males and females in the studied provinces

From the models in Table (2), it could be noticed that the estimate of all difference parameters (d) are equal two ($d=2$); that is mean, the second difference is active.. However, autoregressive model (AR) which represent (p) in the

ARIMA model, is active just for Erbil model (Lag 1); while moving average (MA) models which represent (q) in the ARIMA model are active and significant for all foretasted models and have been represented in (Lag 2, Lag 1, Lag1 and Lag 4, for Erbil, Sulaymaniyah, Duhok and Halabja province, respectively). Similar results and trends were found by Abotaleb [6] for forecasting ARIMA models for USA and China, but his finding for Italy was didn't agree with the present results.

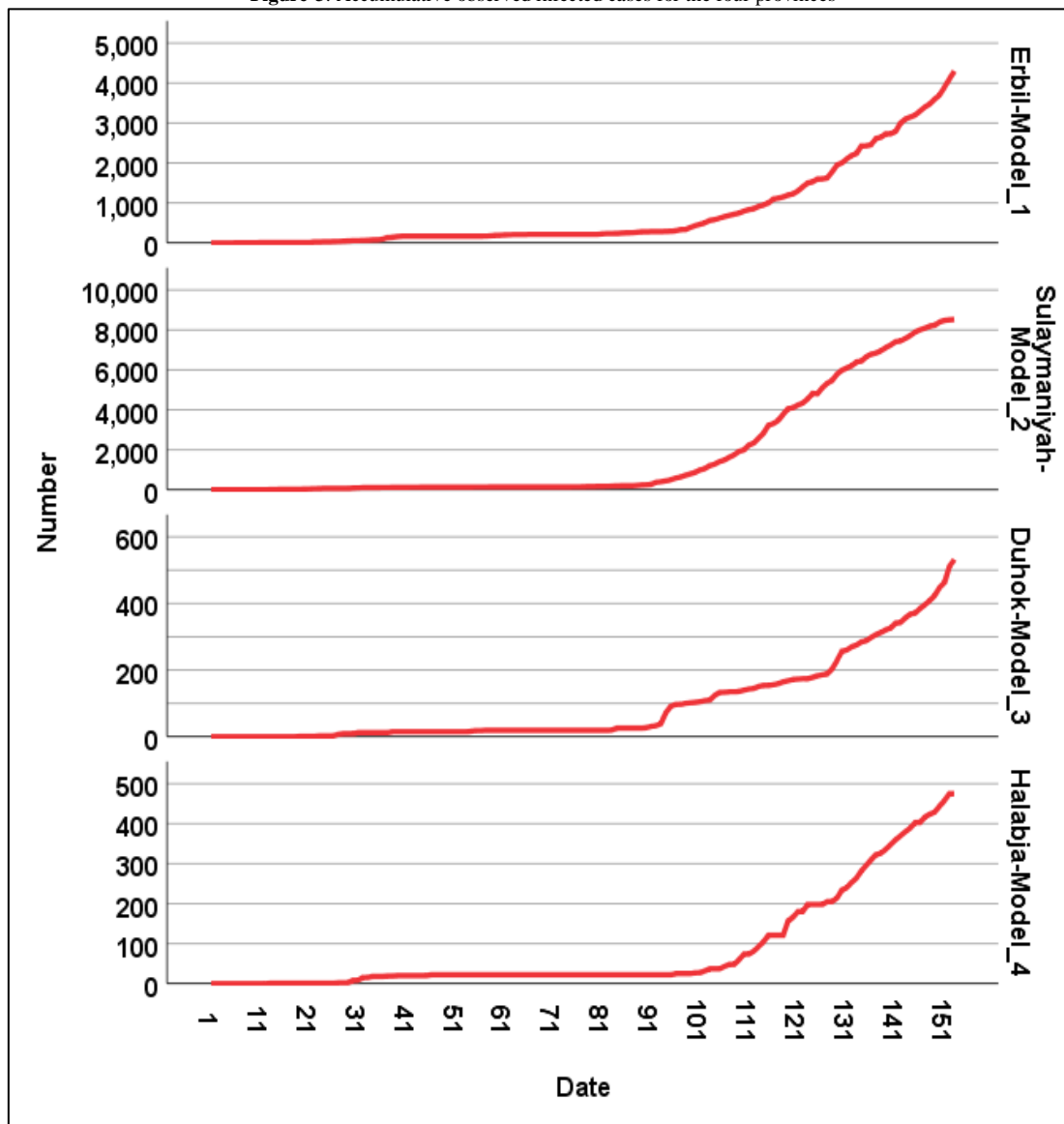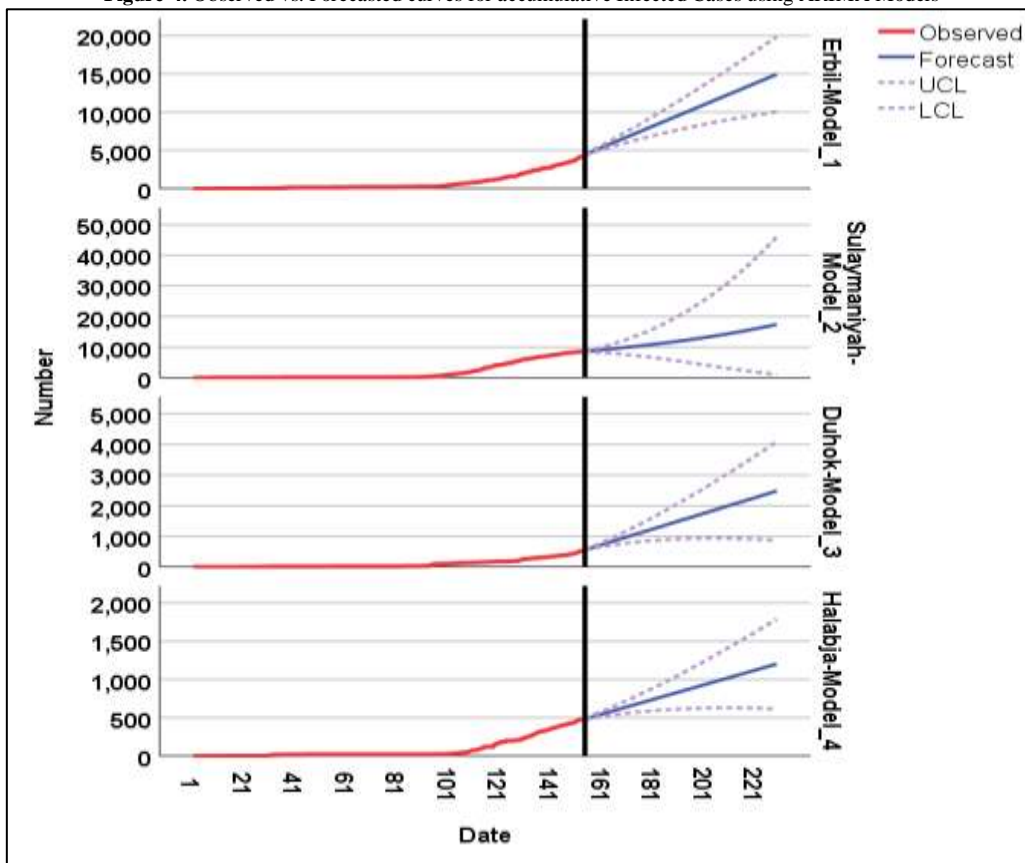**Figure-3**. Accumulative observed infected cases for the four provinces

Table-2. ARIMA model with its autoregressive and moving average estimates

Province	Model	AR(1)	MA(1)	MA(2)	MA(3)	MA(4)
Erbil	ARIMA(1,2,2)	-0.618 **	-	0.69 **	-	-
Sulaymaniyah	ARIMA(0,2,1)	-	0.78 **	-	-	-
Duhok	ARIMA(0,2,1)	-	0.613 **	-	-	-
Halabja	ARIMA(0,2,4)	-	0.834 **	-	0.223 *	- 0.229 **

*= significant ($p<0.05$); **= highly significant ($p<0.01$)

As curves, the forecasted ARIMA models for the accumulative number of infected cases for the four provinces are illustrated in Figure (4). As shown in Figure (4), the number of infected cases for Erbil province will be increased linearly for a considerable time, and the observed numbers may be multiplied (3) times. While for Sulaymaniyah province, the expected number of infected cases will be increased slightly and less than Erbil cases. In respect to Duhok cases, the predicted number of infected cases will be increased linearly and more than Erbil cases. Regarding Halabja expected cases, the expected number will increase gradually and linearly for infected cases and will be similar to Erbil cases. Contrary findings were found by Benvenuto, *et al.* [7], who stated that the the spread of COVID-19 will slightly decrease, and the cases will reach a plateau if the virus does not mutate.

Figure-4. Observed vs. Forecasted curves for accumulative Infected Cases using ARIMA Models

3.3. Exponential Smoothing Model Results

The forecasting results of infected cases for the four provinces of Kurdistan region by exponential smoothing models using Holt's method are presented in Table (3). It could be noticed from the mentioned table that estimates of all model's parameters (using both Alpha-level and Gamma-trend equations) of accumulative infected cases for the four provinces are highly significant ($p<0.01$), and the Exponential Smoothing Model was used square root way as correction factor for data transformation to be normally distributed data. The significant parameters of Exponential Smoothing Model indicate that the prediction of future cases for the studied areas and cases are active, and the government should take a specific health care instruction to control the spread of this pandemic disease (COVID-19). However, a similar study by Perc, *et al.* [8] on forecasting COVID-19 cases of USA, Slovenia, Germany, and Iran countries for 14 successive days using exponential smoothing method reported that the rate of cases should not exceed 5 %.

Table-3. Exponential Smoothing Model Estimates using Holt's method

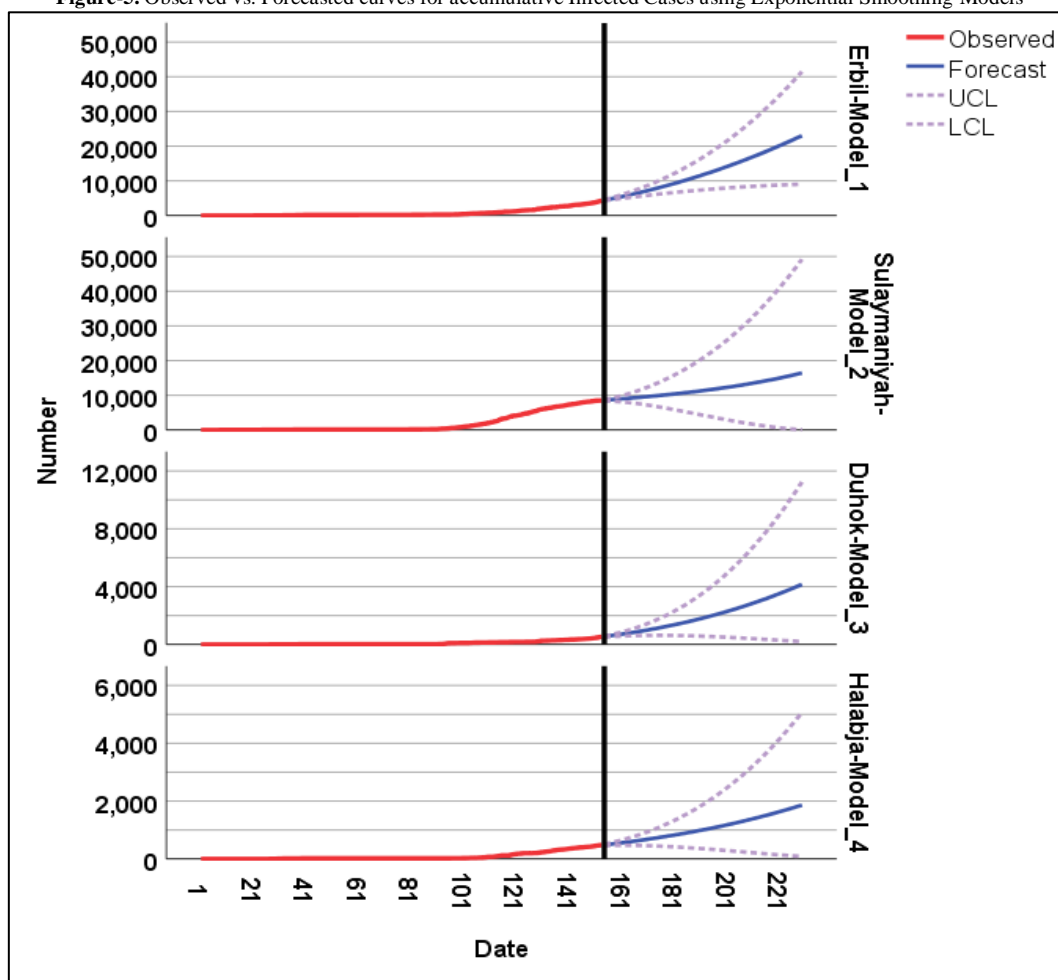| Model		Estimate	SE	t	Sig.
Erbil	α	1.000	.083	12.117	.000
	γ	.147	.052	2.824	.005
Sulaymaniyah	α	.727	.079	9.254	.000
	γ	.375	.093	4.040	.000
Duhok	α	1.000	.083	12.082	.000
	γ	.200	.062	3.247	.001
Halabja	α	.902	.081	11.116	.000
	γ	.163	.052	3.131	.002

As curves, the forecasted exponential smoothing models (using Holt's method) for the accumulative infected cases with confidence intervals (95 %) for the studied provinces are illustrated in Figure (5). As it's obvious from the mentioned Figure, the expected number of infected cases for Erbil province will be non-linearity increased with narrow fluctuation (depending on the government instructions and restrictions), and the observed numbers may be multiplied (5.5) times; while for Sulaymaniyah province infected cases, the expected number of cases will slightly increase with wide logical ranges, this trend is similar to ARIMA model which may be related with data transformation way; for Duhok province, the expected number of accumulative infected cases will increase tightly and non-linearity (here the government should apply more health care and restrictions), where the observed numbers may be multiplied (8) times; regarding Halabja province, the expected trend will be similar to Erbil one, but less intensity, and the observed cases may be multiplied (4) times. However, the hard changes over time for the slope of curve lines might be from the small estimates of Gamma models. These results ensure that the government should

apply the restrictions with health care instructions for more time to control the COVID-19 pandemic. Moreover, several studies have been done by Chakraborty and Ghosh [9]; Li, *et al.* [10]; Roosa, *et al.* [11]; Wu, *et al.* [12] in different countries on forecasting COVID-19 cases, and they reported findings differed in the shape of curves due to the structure of data distribution.

3.4. ARIMA vs. Exponential Smoothing Models

To compare both forecasting models (ARIMA vs. Exponential smoothing) for selecting the best prediction models, it should be depended on the fit statistic accuracy measures which include: Auto-Correlation Errors at first lag (ACF1), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Maximum Absolute Percentage Error (MaxAPE), Maximum Absolute Error (MaxAE), Bayesian Information Criterion (BIC) and Ljung-Box Significance. Table (4) shows the comparison between both main studied forecasting models for the fit statistics of COVID-19 infected cases of studied provinces of Kurdistan region-Iraq. It could be observed from the table that both Erbil and Sulaymaniyah forecasting within ARIMA models are significant at the level of ($p < 0.01$) and ($p < 0.05$), respectively according to Ljung-Box fit parameter estimate, while Duhok and Halabja models were insignificant. Regarding Exponential Smoothing Models, just Duhok province model was highly significant ($p < 0.01$) according to Ljung-Box fit parameter estimate. In general, both studied methods are close to each other for forecasting time series of the studied COVID-19 cases in the studied provinces, since almost half of fit statistics (accuracy measures) are lower in ARIMA and the rest half are lower in Exponential Smoothing model (Table 4). The present results disagree with the findings of Abotaleb [6] who found that the Exponential growth models were better than that of ARIMA ones for forecasting COVID-19 confirmed cases of the USA, Italy, and China, which may be attributed to the data structure.

Figure-5. Observed vs. Forecasted curves for accumulative Infected Cases using Exponential Smoothing Models
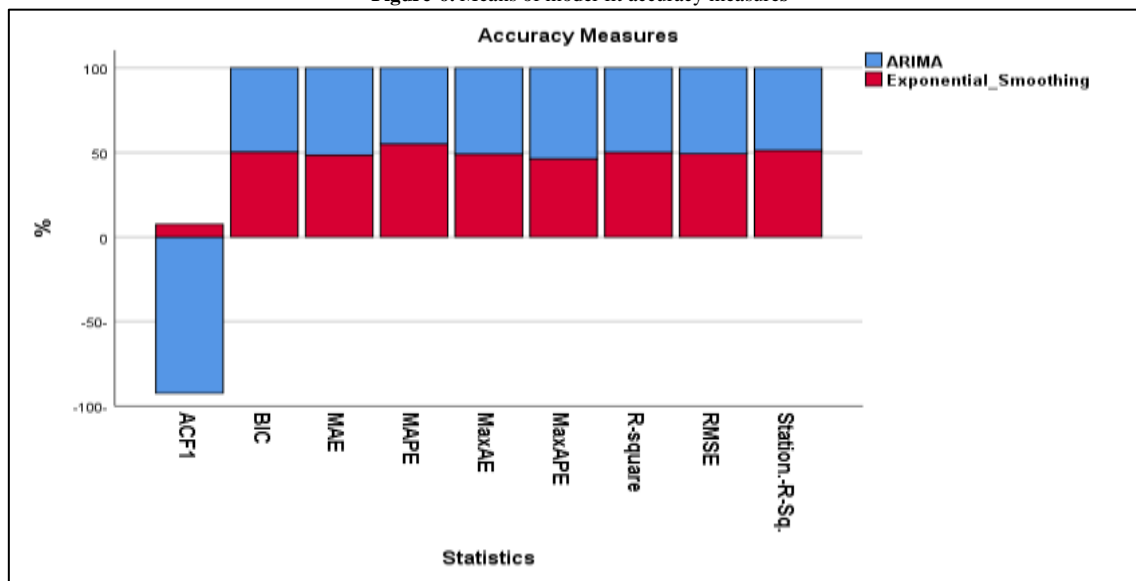

Also, Figure (6) illustrates the overall mean of the fit statistics criteria as the ratio between both used models (ARIMA and Exponential Smoothing). The chart shows that the R-square values of both models are equal (0.999); but half of the rest eight accuracy measures of the ARIMA model are lower than their Exponential Smoothing model counterparts including (ACF1, BIC, MAPE and stationary R-squared); and on the contrary, the other half of the rest accuracy measures of the ARIMA are higher than their Exponential Smoothing model counterparts involving (MAE, MaxAE, MaxAPE and RMSE). However, this stocked bar confirms that both ARIMA and Exponential Smoothing models are close to each other and both suitable for forecasting time series of COVID-19 cases in Kurdistan Region provinces.
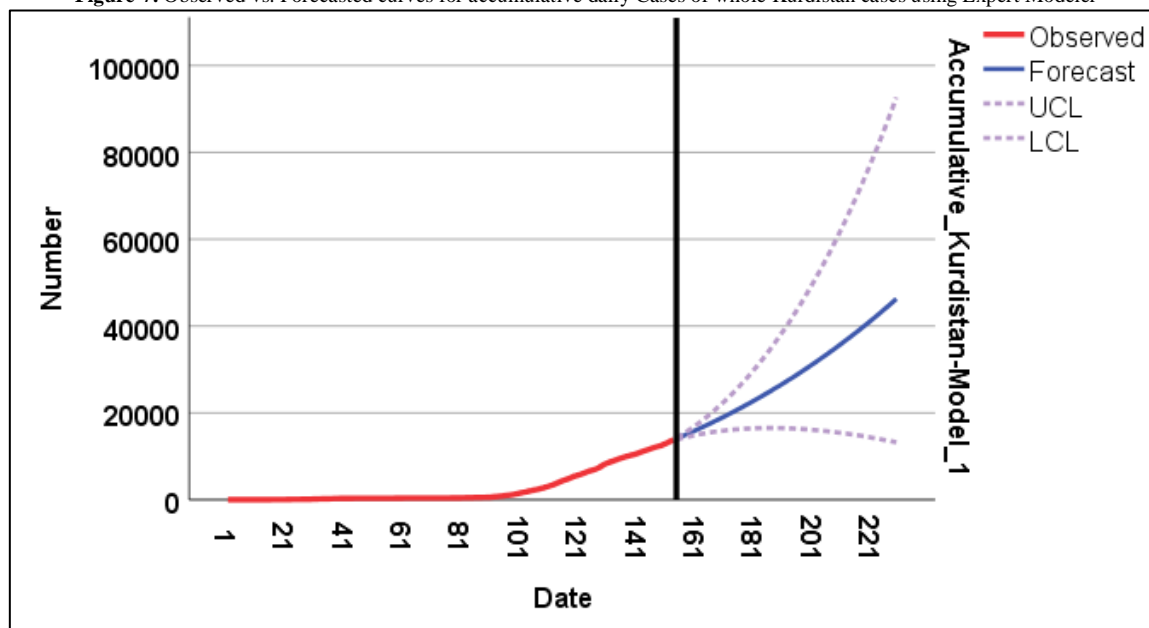
Table-4. Forecast Model statistics

Model	ARIMA Models fit statistics							
	ACF1	RMSE	MAPE	MAE	Max APE	Max AE	Normalized BIC	Ljung-Box Sig.
Erbil	-0.122	30.558	5.373	16.099	100	124.22	6.906	0.007
Sulaymaniyah	-0.192	53.765	3.782	28.32	34.44	234.89	8.002	0.041
Duhok	0.052	5.298	5.59	2.439	100	31.703	3.368	0.353
Halabja	-0.048	4.43	5.306	2.255	100	28.429	3.077	0.058
Exponential Smoothing Models (Holt's Method) fit statistics								
Erbil	-0.008	30.828	5.185	16.017	89.979	137.996	6.923	0.527
Sulaymaniyah	0.005	51.645	3.777	26.672	34.275	210.442	7.955	0.274
Duhok	0.218	5.43	6.597	2.562	92.254	32.862	3.45	0.008
Halabja	-0.01	4.561	6.67	2.298	93.716	28.813	3.101	0.348

Figure-6. Means of model fit accuracy measures

Finally, the accumulative daily cases of all provinces could be summarized in Figure (7) as overall infected cases maight be expected in Kurdistan region. However, the Expert Modeler within SPSS software was used for Kurdistan cases, after the previous results indicate that both ARIMA and Exponential Smoothing are almost resulted in close findings. The expert modeler selected ARIMA(0,2,1) model (R^2=1); with MA lag1 = 0.684 (p<0.01), and ACF value = -0.100, with square root transformation. As it is shown from the Figure (7), the forecasted number of infected cases will be increased as exponential growth; so the people should take health care and keep social distances.

Figure-7. Observed vs. Forecasted curves for accumulative daily Cases of whole Kurdistan cases using Expert Modeler

4. Conclusions

It could be concluded from the present study that the ARIMA and Exponential Smoothing models which used for studied data are close to each other for the forecasted infected cases of COVID-19. The trends of predicted cases for the four provinces indicate that the COVID-19 epidemic could not be controlled in Kurdistan region provinces unless the people apply all government instructions, because the present results indicate the dangerous situation.

5. Suggestion

This study suggests that the negative effects of COVID-19 (the bad situation for health, commercial relationship, economic, education, technology, traveling…etc.) must be challenged by applying the instructions of the government, take more health care and applying social distances to avoid more live loss and to reserve the human activities normally.

Acknowledgment

I thank the government of Kurdistan Region for their suitable instructions that kept the COVID-19 under control. Also, I appreciate the restriction of Kurdistan people by the laws and instructions. Many thanks to Mr. Renas H. Isa for his effort in revising the draft.

Declaration

I declare that I have no competed financial, interests or personal relationships that might have appeared to affect the work of this paper.

References

[1]   Hui, D. S., Azhar, E. I., Madani, T. A., Ntoumi, F., Kock, R., and Dar, O., 2020. "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-The latest 2019 novel coronavirus outbreak in Wuhan, China." *International Journal of Infectious Diseases,* vol. 91, pp. 264-266.

[2]   Josephine, M., 2020. *Coronavirus: China's first confirmed Covid-19 case traced back to November 17.* South China Morning Post.

[3]   COVID-19 Dashboard, 2020. *The Center for Systems Science and Engineering (CSSE).* ArcGIS, Johns Hopkins University.

[4]   COVID-19 Map Website, 2020. Available: https://google.com/covid19-map/?hl=en-IQ&mid=/m/02j71

[5]   The Official Website of Kurdistan Region Government. Available: https://gov.krd/coronavirus-en/dashboard/

[6]   Abotaleb, M. S. A., 2020. "Predicting COVID-19 cases using some statistical models: An application to the cases reported in China Italy and USA." *Academic Journal of Applied Mathematical Sciences,* vol. 6, pp. 32-40.

[7]   Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., and Ciccozzi, M., 2020. "Application of the ARIMA model on the COVID- 2019 epidemic dataset." *Data in Brief,* vol. 29, pp. 1-4. Available: www.elsevier.com/locate/dib

[8]   Perc, M., Gorišek, M. N., Slavinec, M., and Stožer, A., 2020. "Forecasting COVID-19." *Front. Phys.,* vol. 8, p. 127.

[9]   Chakraborty, T. and Ghosh, I., 2020. "Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis." *Chaos, Solitons and Fractals,* vol. 135, p. 109850. Available: https://doi.org/10.1016/j.chaos.2020.1098500960-0779/©2020ElsevierLtd

[10]   Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., and Tong, Y., 2020. "Early transmission dynam- ics in Wuhan, China, of novel coronavirus-infected pneumonia." *N. top N. Engl. J. Med.,* Available: https://www.nejm.org/doi/full/10.1056/nejmoa2001316

[11]   Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., and Hyman, J., 2020. "Real-time forecasts of the covid-19 epidemic in china from February 5th to February 24th, 2020. ." *Infect. Dis. Modell.,* vol. 5, pp. 256-63.

[12]   Wu, J. T., Leung, K., and Leung, G. M., 2020. "Nowcasting and forecasting the potential domes- tic and international spread of the 2019-ncov outbreak originating in Wuhan, China: a modelling study." *Lancet,* vol. 395, pp. 689–97.

[13]   SPSS, 2019. *Statistical package for social Sciences.* IBM statistics software, Ver.

[14]   Granger, C. W. J. and Newbold, P., 2014. *Forecasting economic time series.* Academic Press.

[15]   Hyndman, R. J. and Athanasopoulos, G., 2018. *Forecasting: Principles and practice.* OTexts.

[16]   Abonazel, M. R. and Abd-Elftah, A. I., 2019. "Forecasting egyptian gdp using ARIMA models." *Reports on Economics and Finance,* vol. 5, pp. 35-47. Available: www.m-hikari.com https://doi.org/10.12988/ref.2019.81023

[17]   Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M., 2015. *Time series analysis: forecasting and control.* John Wiley and Sons.

[18]   Holt, C. E., 1957. *Forecasting seasonals and trends by exponentially weighted averages (O.N.R. Memorandum No. 52).* Pittsburgh USA: Carnegie Institute of Technology.