International Journal of Difference Equations (IJDE).
ISSN 0973-6069 Volume 16, Number 2 (2021), pp. 179-187
© Research India Publications
https://www.ripublication.com/ijde.htm

Analysis of Confirmed Cases of Covid-19 in Saudi Arabia Using Poison Autoregressive and Holt Winter Models

Salem Mubarak Alzahrani*, Fath E. I. Elsmih,**
Khalid Salem Al-Zahrani*, Sayed Saber*****

**Faculty of Arts and Science in Almandaq, Al-Baha University, Al-Baha, Saudi Arabia.*

***Department of Mathematics, Faculty of Science and Arts of Baljurashi, Al-Baha University, Al-Baha, Saudi Arabia.,*

***Department of Statistical Study, Peace University, Sudan.*

****Department of Biology, Faculty of Sciences and Arts in Almandaq, Al-Baha University, Al-Baha, Saudi Arabia,*

*****Department of Mathematics and Computer Science, Faculty of Science, Beni-Suef University, Egypt.*

*****Department of Mathematics, Faculty of Science and Arts of Baljurashi, Al-Baha University, Al-Baha, Saudi Arabia.*

Abstract

We applied two statistical models, namely Poison Autoregressive and Holt Winter models. These models were used to better explain the Covid-19 contagion dynamics in Saudi Arabia, which can have an impact on health, economics, finance, and other disciplines. Both models were tested based on daily count time series data of confirmed cases. Moreover, we evaluate the precision of a Poison Autoregressive method with Holt-Winters method. The result showed that a Poison Autoregressive method gives better results forecasting.

Keywords: Covid-19, Count Time Series, log linear Poison Autoregressive, Holt Winter, Saudi Arabia.

Mathematics Subject Classification: 92C50, 34K45, 34D20, 92D25

1. INTRODUCTION

Saudi Arabia has a population of 34 million people, with the majority living in Riyadh, Mecca, Jeddah, Dammam, Medina, and other major cities (The General Authority for Statistics - Saudi Arabia, 2020). Saudi Arabia is ranked second in the Eastern Mediterranean region in terms of the infected and active cases, according to the WHO, 2020a. There were 80,185 total cases, 54,553 recoveries, and 441 deaths in Saudi Arabia as of May 28, 2020 []. The first Covid-19 case in Saudi Arabia was detected on March 2nd in Qatif city in the Eastern area ; the patient had previously visited an endemic region of Iran.

The goal of this work is to apply a Poison Autoregressive and Holt-Winters approach to predict and forecast Covid-19 daily confirmed cases. The Poison Autoregressive has been considered by many studies. These studies dealing with modeling and forecasting Covid-19 around the world such as: [1], [2], [3]. The Holt-Winter technique [4] is a univariate time series forecasting approach. It's a time-series data smoothing process in which the weights and values of previous data are exponentially diminishing. Exponential smoothing comes in three types. The first type for univariate data is single exponential smoothing time series forecasting. When the data from a time series has no systematic pattern. No seasonality or trends were seen in the data [5]. This type of exponential smoothing just employs a single smoothing factor ranging from 0 to 1. Charles Holt and Peter Winters developed the Holt-Winters forecasting technique. Smoothing time series data before using it to forecast other characteristics of the data [6] and [7]. This method is straightforward and does not necessitate a large amount of data storage. It uses the maximum likelihood function for parameter estimation and is based on short forecasting [4,6,8, 9, 10]. There are two Holt-Winter models based on the seasonal component: additive and multiplicative models [6]. The additive models are used to create a model with a linear and exponential trend. It is ideal for data with no trend or seasonality that does not grow with time [4].

2. METHODS

Recent studies found that time series forecasting frameworks that focus on the past behavior of a random phenomenon are the best at capturing the underlying trends and patterns. So, it is then used to forecast the underlying random variable's future behavior. Over the last few years, a lot of effort has gone into developing various time series forecasting algorithms for predicting pandemics. Poison Autoregressive model for count time series and the Holt winter method are utilized to modelling and predicting the daily confirmed cases of Coronavirus in Saudi Arabia

2.1. The Poisson Autoregressive method

The Poisson autoregressive framework for count time series was used to model the confirmed cases of Coronavirus in Saudi Arabia, which is a count variable [11], [12], [13], [14], [19]. The statistical distribution of confirmed Coronavirus cases in Saudi Arabia at time t (day), conditional on data up to $t-1$, is characterized by a Poisson distribution with autoregressive intensity and follows a Poisson distribution with autoregressive intensity. and is defined by:

$$X|\mathcal{F}_{t-1} \sim Poissn(\gamma_t),$$
$$\log(\gamma_t) = \delta + \partial \log(1 + X_{t-1}) + \beta \log(\gamma_{t-1}), \quad (1)$$

where $\delta, \partial, \beta \in R$ and $X \in N$. Here, we used $\log(1 + X_{t-1})$, instead of $\log(\gamma_{t-1})$, permits to allow with zero values. In fact, δ is the short-term dependence on the previous period and it is the intercept in Eq. (1), and β is the long-term component dependence on all previous values of the observed process. Notice that ∂ and β depended on the expected number of confirmed cases γ_t based on previous confirmed cases. In the greater ranking of the persistence of shocks in the counting process, the total $(\partial + \beta)$ is close to 1. The sign of the autocorrelation of the confirmed cases are verified by the value of $(\partial + \beta)$. The confirmed cases show a positive autocorrelation at lag 1 when the sum $(\partial + \beta)$ is positive. The decrease in autocorrelation after the first lag is slower if the sum of $(\partial + \beta)$ is closer to 1. If the sum $(\partial + \beta)$ close to 1, the autocorrelation slower decrease after the first lag. The sum $(\partial + \beta)$ effects on the unconditional mean of the process.

2.2. Holt Winter Method:

Holt Winter is a triple exponential smoothing method which is a univariate time series extension of exponential smoothing. It is used when the series exhibit seasonal pattern, i.e., the method allow seasonality [15]. This method is straightforward and does not require a lot of data storage [18]. The Holt's method is appropriate for short-term prediction and estimates parameters using the maximum likelihood estimator, see [15,16, 17]. Based on the seasonal component, there exist two types of Holt-Winter, namely additive and multiplicative models [17]. For Holt's Winter, there are separate equations combined to generate the ultimate future prediction. Holt's winter method was concentrated on the following equations:

$$S_t = \alpha x_t + (1-\alpha)(S_{t-1} + b_{t-1}) \quad , \quad 0 < \alpha < 1,$$
$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \quad , \quad 0 < \gamma < 1,$$

$$F_{t+m} = b_t + b_t m,$$

where S_t represents the smoothed value, b_t represents the trend smoothed value, and m represents the forecast period ahead.

3. RESULTS

This section presents the empirical results by applying Poison Autoregressive model and Holt's Holt Winter method to analyze Covid-19 confirmed cases in Saudi Arabia from March 3, 2020, to June 10, 2021.

3.1. Data Availability:

For this investigation, Covid-19 daily confirmed cases in Saudi Arabia from March 3, 2020, to June 10, 2021, were used. The data was obtained from the Saudi Arabian Ministry of Health. As seen in Figure 1, the daily confirmed cases are increasing each day. On the other hand, declines between October 2020 and April 2021. Because of the Saudi Administrator's large-scale social restriction agenda, this is possible.

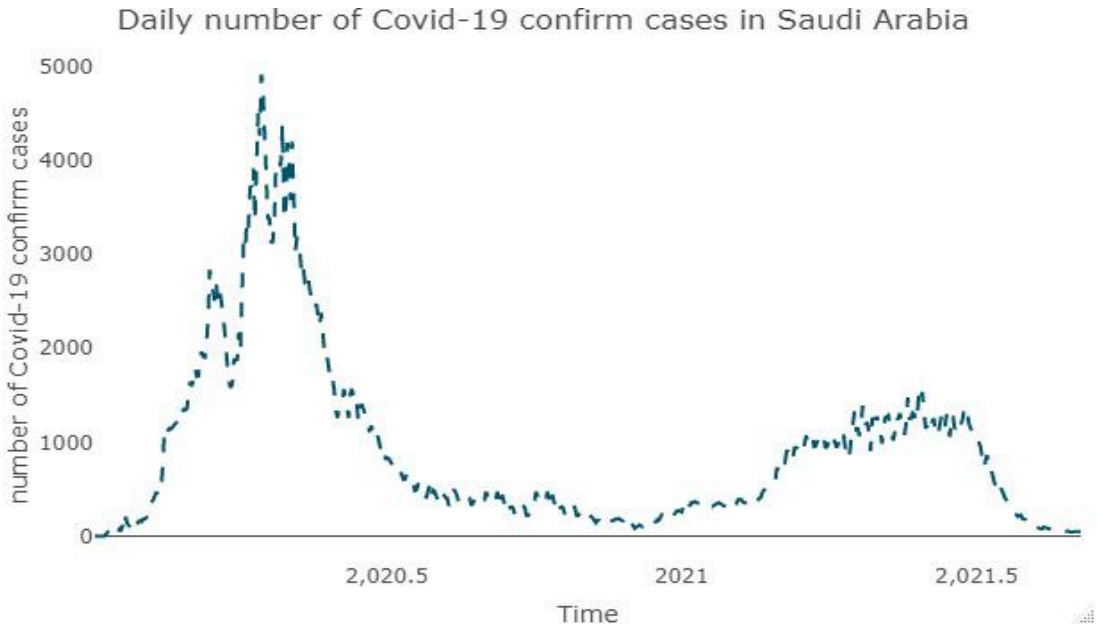

Figure 1. Sequence chart of new series of Covid-19 confirmed cases during the period 3/3/2020 – 6/10/2021

The data's suitability is next assessed by examining the stationarity of the daily confirmed cases. Based on the Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

Stationarity test findings from Table 1, we concluded that daily confirmed cases are stationary at the level in lag 6 with KPSS Level = 1.658 (p-value = 0.01). Because the stationary assumption is satisfied, we may construct a prediction model utilizing daily confirmed cases based on these findings.

Table 1. The parameter estimates of Poisson autoregressive model of Covid-19 confirmed cases data.

Variable	*KPSS* Test	
	Level	*prob*
Covid-19 Confirmed Cases	0.39748	0.01

Following the stationarity test, we divided the data into two sections: training data (459 days) and testing data (153 days). By comparing the training data to multiple parameter combinations based on the validation data, the optimum parameter can be determined. The best parameter combination is then discovered, and the entire training data is retrained with it. Following that, the effectiveness of each strategy is assessed and compared to one another. Various evaluation criteria were utilized in the study, namely Mean Absolute Error (MAE), Root Means Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) are used to evaluate the correctness of the two models. The flowing formula gives the MAE, RMSE, and MAPE formulas for a given number of observations [19]:

$$\mathrm{MAE} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \left|\mathrm{Y_i} - \widehat{\mathrm{Y}}_{\mathrm{i}}\right|}{\mathrm{n}},$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{T} (\mathrm{Y_i} - \widehat{\mathrm{Y}}_{\mathrm{i}})^2}{T}},$$

$$\mathrm{MAPE} = \frac{1}{T} \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \left|\mathrm{Y_i} - \widehat{\mathrm{Y}}_{\mathrm{i}}\right|}{\mathrm{Y_i}} \times 100.$$

3.2. Applying Poisson Autoregressive model

Table 1 illustrates the 95 % credible intervals for the computed Poisson autoregressive coefficients for Eq. (1). All coefficients confirm the presence of both a short-term reliance and a long-term trend, and their credible intervals exclude zero.

Table 2. Parameter estimates of Poisson autoregressive Model for confirmed cases.

Parameter	Estimate	t- Statistic	Std. Error	p-value	CI (lower)	CI (upper)
δ	1.691	12.524	0.5372	0.031	0.6384	2.7441
∂	0.070	24.379	0.0120	0.019	0.0465	0.0937
β	0.915	25.088	0.0118	0.014	0.8921	0.9386

Figure 2 shows the residuals chart, residuals autocorrelations and residuals of Poisson autoregressive model of Covid-19 confirm cases. these figures confirm that residuals of the model are random.

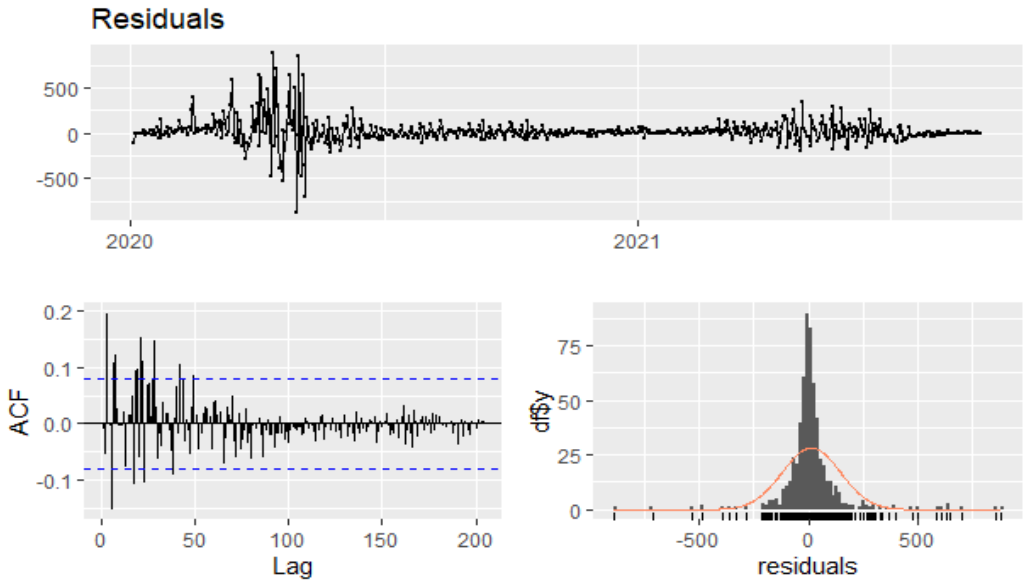

Figure 2. Residuals of Poisson autoregressive model

3.3. Applying Holt Winter method

The Holt Winter technique was applied to daily confirm cases from March 3, 2020, to June 10, 2021, and the estimated parameters are displayed in Table 3.

Table 3. Paramereters estimated of Holt Winter model.

Model	Initial states		Smoothing parameters	
	l	b	alpha	beta
Holt-Winter	-1.556	0.0736	0.9005	1e-04

Figure 3 shows the residuals diagram, residuals autocorrelations and residuals Holt Winter, Figures confirm that residuals of the model are random.

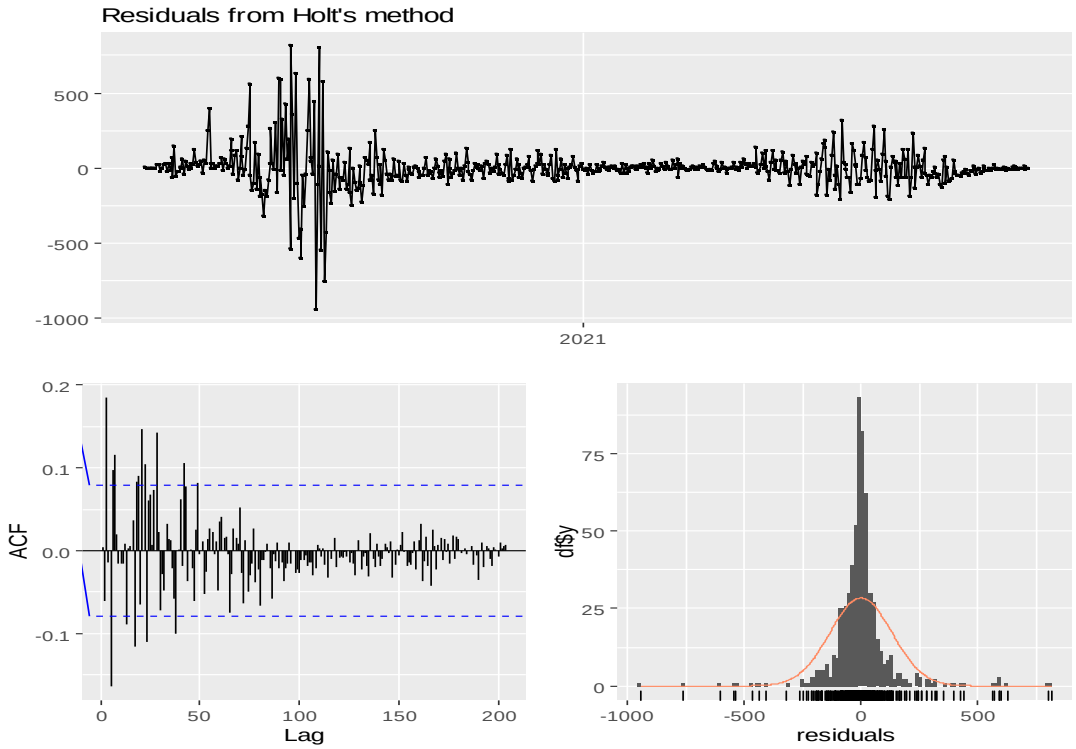

Figure 3. Residuals of Holt Exponential Smoothing Model

4. DISCUSSION

The first model, a Poison Autoregressive as a function of both a short-term and long-term dependence, provides the best fit to the data for this study. The use of a Poisson autoregressive model that includes short- and long-term memory effects can greatly improve the estimation of confirmed cases and can indicate whether the disease is trending upward or downward, which can help public decision-makers better plan health policy interventions and take the appropriate actions to contain the virus to the extent possible. To evaluate the predictions, we looked to visual difference between Poisson Autoregression with a Holt's method. The assessment the above measurements are expected to be as minimum as possible. Minimum values suggests that the prediction is accurate.

Table 4. The evaluation of the analytical accuracy of our two models.

Method	MAE	RMSE	MAPE
Poisson Autoregressive method	17666.4	132.15	69.1
Holt Winter method	17516.37	132.3494	69.795

5. CONCLUSIONS

This study applied **two** methods, namely Poisson Autoregressive model and Holt's winter model. The results of this work revealed that a Poison Autoregressive model produces more accurate predictions. Besides, it forestalls future Coronavirus pandemic occurrences. Poison Autoregressive prediction finding can be employed as a baseline for additional interference. As a result, it is necessary to extend the application of Poisson Autoregressive to other scenarios in Saudi Arabia and develop methods to solve estimate problems such as underestimating and merging in long-term forecasting. Table 4, showed that the evaluation of the analytical accuracy of our two models. The results showed that a Poisson Autoregressive model outperforms a Holt's winter model.

FUNDING:

This research is a part of a project entitled "Using statistics and mathematical modelling to understand infectious disease outbreaks: A case study of the Covid19 epidemic and its impact on the Al-Baha region". The Deanship of Scientific Research, Al- Baha University, and KSA (Grant No. 1442/21) funded this project.

ACKNOWLEDGEMENTS:

All authors wish to express their gratitude to the assistance of the Deanship of Scientific Research, Al-Baha University, and KSA (Grant No. 1442/21) funded this project.

REFERENCES

[1] Konstantinos Fokianos, Anders Rahbek & Dag Tjøstheim (2009) Poisson Autoregression, Journal of the American Statistical Association, 104:488, 1430-1439, DOI: 10.1198/jasa. 2009.tm08270

[2] Ahmed Msmali, Zico Mutum, Idir Mechai, Abdullah Ahmadini, Modeling and Simulation: A study on predicting the outbreak of COVID-19 in Saudi Arabia medRxiv 2021.01.17.21249837; doi: https://doi.org/10.1101/2021.01.17.21249837

[3] Agosto A, Campmas A, Giudici P, Renda A. Monitoring COVID-19 contagion growth. Stat Med. 2021 Aug 15;40(18):4150-4160. doi: 10.1002/sim.9020. Epub 2021 May 11. PMID: 33973656; PMCID: PMC8242489.

[4] V. Papastefanopoulos, P. Linardatos, S. Kotsiantis, COVID-19: a comparison of time series methods to forecast percentage of active cases per population, Appl. Sci. Basel. 10 (11) (2020) 3880.

[5] Djauhari M A, Asrah N M, Li L S and Djakaria I 2020 Forecasting model of electricity consumption in Malaysia: A geometric Brownian motion approach Solid State Technology 63 3 40-46

[6] Panda M 2020 Application of ARIMA and Holt-Winters forecasting model to predict the spreading of COVID-19 for India and its states medRxiv

[7] Makatjane K and Moroke N 2016 Comparative study of holt-winters triple

exponential smoothing and seasonal Arima: forecasting short term seasonal car sales in South Africa Risk governance & control: financial markets & institutions 6 1 Winter 2016

[8] Bezerra A K L and Santos É M C 2020 Prediction the daily number of confirmed cases of COVID19 in Sudan with ARIMA and Holt Winter exponential smoothing International Journal of Development Research 10 08 39408-39413

[9] A.M. Awajan, M.T. Ismail, S. Al Wadi, Improving forecasting accuracy for stock market data using EMD-HW bagging, PLoS One 13 (7) (2018), e0199582.

[10] Nasrin Talkhi, Narges Akhavan Fatemi, Zahra Ataei, Mehdi Jabbari Nooghabi, Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods, Biomedical Signal Processing and Control 66 (2021) 102494

[11] Agosto A, Campmas A, Giudici P, Renda A. Monitoring COVID-19 contagion growth. Stat Med. 2021 Aug 15;40(18):4150-4160. doi: 10.1002/sim.9020. Epub 2021 May 11. PMID: 33973656; PMCID: PMC8242489.

[12] P. J. Brockwell, R. A. Davis, Time Series: Data Analysis and Theory, 2nded., Springer, New York,1991.

[13] Kharroubi SA (2020) Modeling the Spread of COVID-19 in Lebanon: A Bayesian Perspective. Front. Appl. Math. Stat. 6:40. doi: 10.3389/fams.2020.00040

[14] Barría-Sandoval C, Ferreira G, Benz-Parra K, López-Flores P (2021) Prediction of confirmed cases of and deaths caused by COVID-19 in Chile through time series techniques: A comparative study. PLoS ONE 16(4): e0245414. https://doi.org/10.1371/journal.pone.0245414

[15] Roosa K.; Lee Y.; Luo R.; Kirpich A.; Rothenberg R.; Hyman J.; et al. Real-time forecasts of the COVID19 epidemic in China from 5 February to 24 February 2020. Infect. Dis. Model. 2020, 5, 256–263. https://doi.org/10.1016/j.idm.2020.02.002 PMID: 32110742

[16] Zhang S., Diao M., Yu W., Pei L., Lin Z., & Chen D. (2020). Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. International Journal of Infectious Diseases, 93, 201–204. https://doi.org/10. 1016/j.ijid.2020.02.033 PMID: 32097725

[17] Ali M, Khan DM, Aamir M, Khalil U, Khan Z (2020) Forecasting COVID-19 in Pakistan. PLoS ONE 15(11): e0242762. https://doi.org/10.1371/ journal. pone.0242762

[18] Stapper, M. Count Data Time Series Modelling in Julia—The CountTimeSeries.jl Package and Applications. *Entropy* 2021, *23*, 666. https://doi.org/10.3390/e23060666

[19] Salem Mubarak Alzahrani, Fath E. I. Elsmih, Khalid Salem Al-Zahrani, Sayed Saber, A Log Linear Poisson Autoregressive Model to Understand COVID-19 Dynamics in Saudi Arabia, submitted.