

Journal of Institute of Science and Technology, 25(2), 80-92 (2020) *ISSN: 2469-9062 (print), 2467-9240 (e)* Doi: https://doi.org/10.3126/jist.v25i2.33744

AN ANALYSIS OF COVID-19 CASES IN NEPAL: A MODELING APPROACH

Surendra Raj Nepal*

*Department of Statistics, Amrit Campus, Tribhuvan University, Kathmandu, Nepal *Corresponding author: surendra.nepal@gmail.com*

(Received: August 21, 2020; Revised: November 05, 2020; Accepted: November 21, 2020)

ABSTRACT

Unlike previous coronaviruses infections, COVID-19 has badly affected not only the health of people but also the socioeconomic activities of Nepal. It would help the government of Nepal to manage this crisis if a proper mechanism to predict COVID cases has been developed. This study aims to look for patterns of confirmed, recovery and death cases. Moreover, it tries to check whether Gompertz and Logistic model would be able to read the patterns of total confirmed and death cases. It also forecasts the total number of confirmed as well as death cases. Data from January 23, 2020 to October 30, 2020 obtained from the website of Wikipedia are used for analysis. Gompertz and Logistic models were fitted to the total number of confirmed and death cases and models are compared based on various criteria. Besides, an automatic ARIMA model was used to predict cumulative confirmed and death cases and the accuracy of the model was also checked. ARIMA model forecasted 347,812 confirmed cases and 1,754 death cases till December 31, 2020. At 95 % confidence interval, the confirmed cases were expected between 273,889 and 421,734 whereas death cases were estimated from 1,387 to 2,119. Both models were fitted well in both total confirmed cases and total death cases. It was found that the Logistic model fits better in total confirmed cases whereas in total death cases, the Gompertz model was better. ARIMA model precisely forecasted the number of confirmed and death cases.

Keywords: ARIMA model, COVID-19 cases, Gompertz model, Logistic model, Nepal.

INTRODUCTION

The novel coronavirus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) responsible for coronavirus disease-19 (COVID-19) was first reported in Wuhan, Hubei Province, China on December , 2019 (Huang *et al.,* 2020; Bherwani *et al.,* 2020). Similar to severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV), SARS-CoV-2 has been recognized to be zoonotic origin and usually causes respiratory disease as an onset symptom (Guo *et al.,* 2020). COVID-19 has been spread all over the world relatively quickly than its ancestors. In January 2020, World Health Organization (WHO) declared COVID-19 as a public health emergency of international concern and as a pandemic on March 11, 2020 (WHO, 2020).

Experts have outlined three stages of transmission of COVID-19: local transmission, community transmission, and large scale transmission (epidemic). The contact among people of the different populations determines the aspect that characterizes the rate of transmission pattern in civic places and families, and this virus exhibited high transmission rate (Guo *et al*., 2020; Sarkodie & Owusu, 2020). The COVID-19 mostly attack and damage the respiratory system and alveoli therein (Gautam, 2020; Asadi *et al.,* 2020). The virus enters through the eyes, the nose, or/and the mouth infects the lung, accrues in the kidney, and can cause damage to resident renal cells (Cheng *et al.,*2020; Fan *et al.* 2020; He *et al*., 2006). The inhalation of transmittable aerosols is the substantial mode of transmission of COVID-19. The incubation period for COVID-19 is between 3-14 days (Kannan *et al.,* 2020).

Some of the characteristic features of SARS and MERS virus were similar. The mortality rate due to the failure of the respiratory organ for MERS is much higher than SARS, and older age people are more vulnerable (Du *et al.,* 2020; Rothan & Byrareddy, 2020; Hui *et al.,* 2014). The fatality rate of SARS-CoV-2 is around 2 - 3 % (Jain *et al.*, 2020). SARS and MERS showed a real CFR (Cases Fatality Rate) of 9.6 % and 34.4 %, respectively (Suwantarat & Apisarnthanarak, 2015; Majumder *et al.*, 2014). Although the fatality rate of the SARS-CoV-2 is less than its ancestors, it is causing more deaths due to high transmission rate (Guarner, 2020).

It has been around ten months since the first Covid-19 case was reported (Huang *et al.,* 2020) but the pandemic has not been controlled yet. According to Worldometer [\(www.worldometers.info\)](../../www.worldometers.info) till October 30, 2020; the total number of confirmed cases around the world has reached 45,921,794. Out of which 1,193,912 people have died from it whereas 32,252,284 recovered. In Nepal, to date total confirmed cases, total deaths and total recovered reported are 168,235; 920, and 128,958 respectively. According to the situation analysis report of the Ministry of Health and Population, Government of Nepal till October 30, 2020; positivity rate, case fatality rate, and total death/million were 11.7 %, 0.5 %, and 31.5, respectively. Confirmed cases of males were much higher than females and the most affected age group was 21-30 years.

The first case in Nepal was reported on January 23, 2020, and the first death case was observed on May 16, 2020. Most of the cases were related to the people who have returned from aboard. Various countries initiated lockdown as a measure to reduce the transmission of the virus (Gautam & Hens, 2020). Nepal government implemented nationwide lockdown from March 23, 2020 following second reported case, to control transmission of COVID-19. The lockdown lasted for around four months and was lifted from July 24, 2020. COVID-19 badly hit every sector of Nepal, especially tourism. Social and economic activities were disturbed as a result of lockdown.

This study examined patterns of total confirmed cases, total active cases, total recovery cases, total death cases, total PCR tests as well as newly infected cases, new active cases, new recovery cases, new death cases, and new PCR tests. It also tested whether Gompertz and Logistic models will be able to read the patterns of total confirmed cases and total death cases. Moreover, it compared two models using different criteria. Furthermore, it estimated the time period of the maximum daily confirmed and death cases. Besides, it also estimated the saturation point of both total confirmed and death cases. Gompertz and Logistic models can be used for fitting rather than forecasting. So, automatic ARIMA especially known for forecasting time series data was used to forecast the total confirmed and death cases.

The most common global COVID-19 analyses are Graphical, Descriptive, Projection, Bayesian, and Modeling. From literature review, it was found that nonlinear models was better than the linear model for investigating the total number of confirmed cases of COVID-19. Growth curves were generally fitted by nonlinear regression or linear regression if the model can be linearized by transformation. However, a linear form of the most widely used growth models does not exist (Blasco *et al.,* 2003). Nonlinear functions are particularly suitable for modeling growth data, since predictions outside the range of the data set can be obtained more reliably than by linear models, and few parameters having a biological interpretation can be used to describe the entire growth process (Vuori *et al*., 2006).

Verity *et al*. (2020) have estimated the severity of COVID-2019 and it showed a case fatality ratio in China to be 1.38 (95 % confidence limit of 1.23-1.53). Silva *et al.* (2020) investigated COVID-19 through Bayesian analysis of the total number of cases in Goias, Brazil where they found outbreak peak 60 days after the onset with 95 % limits from 51 to 68 days. They also estimated the total confirmed cases as 3180 and prevalence rate 4.53 per 10,000. Following the outbreak in Wuhan, several modeling groups around the world have estimated and the modeling results have shown a wide range of variations

(Cyranoski, 2020). Estimated basic reproduction number varied from 2 to 6, peak time estimated from mid-February to late March, and the total number of infected people ranged from 50,000 to millions. Roda *et al.* (2020) found a linkage between the transmission rate and the case-infection ratio, which resulted in a variety of best-fit parameter values, and can create significantly different model predictions of the epidemic.

Ahmadi *et al.* (2020) looked into the trend of the COVID-19 epidemic in Iran until May 13, 2020, using Gompertz and other growth models. They predicted the number of patients on April 3, 2020 by Gompertz model with 95 % confidence interval (CI) as 47 500 (38 907-52 640). They also made predictions on the flat epidemic curve and the number of patients based on the Gompertz model as 67 000 (61 500-87 000) cases. According to their report based on Gompertz model 4620 (3930-5550) deaths might occur from May 13 to June 1, 2020, respectively, and then the curve would be flattened. Jia *et al.* (2020) analyzed COVID-2019 using three different models: Logistic, Bertalanffy, and Gompertz. According to them, COVID-19 and SARS virus both being coronaviruses, the infection pattern might be similar. So, they first tested all three models for SARS where they found Logistic and Gompertz were better than Bertalanffy. They again applied these models using data till February 29, 2020to predict the epidemic situation of COVID-19 in the later stage of the epidemic. According to their results, the Logistic model was better than Bertalanffy, and Gompertz models in fitting all the data of Wuhan, while the Gompertz model was better in fitting the data outside Wuhan. They estimated the final cumulative number of confirmed cases of COVID-19 in Wuhan was between 49852 and 57447, and turning point February 9, 2020 with the total death toll of 2502.

Torrealba-Rodriguez *et al.* (2020) made a prediction of COVID-19 in Mexico by taking data from February 27 to May 8, 2020. According to them, the Gompertz model was slightly better than the Logistic model. The Gompertz model predicted a total of 47,576 cases, while as a total of 42,131cases from the Logistic model on May 16. They also forecasted the total number of COVID-19 infection until the end of the epidemic, from the Gompertz and Logistic model, predicting 469,917 and 59,470 cases, respectively, and maximum daily new cases on June 25 and May 8 estimated by corresponding models. Martelloni and Martelloni (2020) studied the temporal evolution of the SARS-Cov-2 in Italy where among four different models; the generalized logistic model best described the situation in Italy.

Martinez *et al.* (2020) have researched 'short-term forecasting of daily COVID-19 cases in Brazil by using the Holt's model'. They have calculated MAPE (mean absolute percentage error) for each model. According to

their results, the MAPE of Gompertz is less than that of Logistic which means the Gompertz model is more accurate than Logistic. Asadi *et al.* (2020) found the generalized Gompertz model as a good fit for measuring the number of individuals infected in Italy and Iran. Kriston (2020) investigated COVID-19 cases by taking data till March 29 from John Hopkins University and made projections for six countries: Hubei in China, South Korea, Germany, United States, Brazil, and South Africa using Hierarchical Logistic model. It was observed that the model approximated the cases very well. Castorina and Iorio (2020) analyzed coronavirus data by microscopic growth laws. They studied cases of China, South Korea, Singapore, and Italy and found that Gompertz laws a less effective containment effort, predict a much larger maximum number of infected than Logistic laws. Harvey and Kattuman (2020) forecasted COVID-19 in the UK and Germany using Gompertz and Logistic models where they summarized dynamics Gompertz model worked extremely well and superior to Logistic. They projected saturation level in the UK as 186,000. Razzak (2020) studied New Zealand COVID-19 infection rate by fitting the Gompertz model to data from February 28 to March 27, 2020, where it was observed the model fit well and lockdown significantly reduced the infection rate.

Yang *et al.* (2020) used the ARIMA model to predict the number of cases and death in Hubei, China. They claimed the model was accurate having low Mean Absolute Error and high R-square value. Malki *et al.* (2020) forecasted that there might be a second round of pandemic in a year using ARIMA and SARIMA. Hariharan and Prakash (2020) predicted the number of infected cases for the next few days using the ARIMA model. From their findings, it was observed that the model was accurate and forecasted values were closer to actual values. Roy *et al.* (2020) forecasted the COVID-2019 epidemic pattern and compared the actual and predicted values. From their study, it was seen that the west and south of Indian districts are most vulnerable for COVID-2019. Sahai *et al.* (2020) used the ARIMA model and forecasted total infected cases in the top five affected countries for the next 77 days. It was found from their study about forecasting accuracy within acceptable agreement.

MATERIALS AND METHODS

Data of infected cases provided by the Ministry of Health and Population, Government of Nepal on daily basis were archived in Wikipedia (MoHP, 2020). Data from January 23, 2020 to October 30, 2020 were used for analysis. At first, the trends of confirmed, death and recovery cases were observed through graphs. Besides, Summary statistics of daily new cases, death cases, recovery cases, and the number of PCR tests were calculated. Nonlinear Gompertz and Logistic models were fitted to the total number of confirmed cases as well as total death cases and estimated timeline at which maximum daily new confirmed and death cases would occur (Bates & Watts, 1988) by using nlsLM function of R package minpack.lm (Timur *et al.,* 2016) of R statistical software (Team, 2019). Models were compared by using Akaike information criterion (AIC), Deviance information criterion (DIC), Bayesian information criterion (BIC), and Loglikelihood. Each criterion has some limitations so all those criteria were computed in this study to select the model correctly. But one can compute BIC only for a large data set. The goodness of fit of the models was assessed through the value of R-square and test statistic values along with the above mention criteria. The Gompertz model (1825) was first given by Gompertz to study hazards in a life table (Seber & Wild, 2003). The model can be stated, as given in equation (1).

$$
y_t(D_t) = A \exp[-\exp\{-K(t-T)\}] + e_t
$$
 (1)

Where, y_t and D_t are the total number of confirmed and death cases at time t, A is the upper asymptote, K is the growth coefficient, e_t is the error term and T is the time at inflection which represents time at maximum daily (confirmed/death) cases.

The Logistic model was first proposed by Verhulst in 1838 to describe the growth in the size of the population or organ (Seber & Wild, 2003). The model can be expressed as;

$$
y_t(D_t) = \frac{B}{[1 + \exp(-K(t - T))] + E_t}
$$
 (2)

Where, y_t and D_t are the total number of confirmed and death cases at time t, B is the upper asymptote, K is the growth coefficient, E_t is the error term and T is the time at inflection.

Even though A of equation (1) and B of equation (2) both represent upper asymptote, the point of inflection of equation (2) is B/2 which is not true in the case of equation (1) so A is greater than B.

ARIMA (Autoregressive integrated moving average)

ARIMA is specially designed for forecasting so it can predict more precisely than Gompertz and Logistic models. ARIMA models are normally denoted by ARIMA (p, d, q) where p is the number of time lags, d is the degree of differencing and q is the order of moving average. Automatic ARIMA was used to forecast the total number of confirmed and death cases by using R software. The model can be stated as

$$
y_{t} = \sum_{1}^{p} \alpha_i y_{t-i} + \sum_{1}^{q} \delta_j \varepsilon_{t-j} \tag{3}
$$

$$
D_{t=}\sum_{1}^{p} \alpha_i D_{t-i} + \sum_{1}^{q} \delta_j \varepsilon_{t-j} \tag{4}
$$

Where, α , δ and ϵ are the parameters of the model.

The Gompertz and Logistic models can estimate growth coefficient and time when outbreak will slow down but it cannot forecast the number of cases precisely. On the other hand, automatic ARIMA can forecast accurately but it cannot estimate the time when the outbreak will reach maximum.

RESULTS

Data obtained from secondary source was first organized and cleaned then analyzed using R software.

Trend analysis

From Fig. 1, it can be observed that the confirmed cases' pattern is almost the same till the third week of April and started to change from the third week of May. The trend from June to the first week of July is alike then after September, the slope has drastically changed. As far as daily new cases are concerned, the increment was mostly in October. After the first week of July, the trend started to decrease and from the fourth week, again it marched upward. From the last week of May to July, the number of active cases rocketed and then it started to fall steeply. After September it started to increase again.

Fig. 1. Total confirmed cases, daily new cases, and total active cases

Fig. 2 shows that the total Recovery rate improved from June and it kept on going at a greater pace. There was not much variation in the daily new recovery in June as in contrast to September and October.

Fig. 2. Total recovery and daily new recovery

After observing the first death in the middle of May, the total number of deaths started to climb up with almost the same pattern, as shown in Fig. 3. Likewise, a similar pattern in new daily death was observed most of the time from the first week of June to the second week of July then it started to go up and continued till October.

Fig. 3. Total deaths and daily new death

Fig. 4 depicts the total number of RT-PCR (Reverse transcription-polymerase chain reaction) tests has increased after the middle of May and it increased with the same pattern. Similarly, new PCR tests' trend was on the rise from middle of May to end of June, but started to fall after July and increased after last week of July which continued till third week of October.

Fig. 4. Total number of RT-PCR and daily new RT-PCR tests

Summary statistics are presented in Appendix1. After analyzing data from January 23 to October 30, it was found that on average, 597 cases and 4 deaths per day were reported, with 458 people per day recovered, and 5668 PCR tests were done. Although there was variation in all the variables listed in Appendix1, among four the most consistent is the PCR test and the most deviated is the recovery. All the variables were positively skewed. There were more peaks in the middle in case of recovery,

Table 1. Parameters of the Gompertz model for total confirmed cases

Parameters	Estimate	Std. Error	t-value	R-square
A	1464000000	2491000000	0.588	99.80%
K	0.002443	0.0004164	5.867*	
т	1185	230.4	$5.141*$	
.				

**Significant at 1 %*

Table 2. Parameters of the Logistic model for total confirmed cases

Parameters	Estimate	Std. Error	t-value	R-square
B	11050000	27230000	0.406	99.91%
K	0.0273	0.0004605	53.709*	
т	449.4	104	$4.32*$	

confirmed cases, and death whereas there was more flatness in PCR tests.

Table 1 depicts that all parameters except the upper asymptote of the Gompertz model are significant at 1 % level of significance and the value of R-square is 99.80 %. Higher values of R-square and a highly significant majority of parameters indicate the validity of the model. Such a higher value of R-square was observed in Gompertz and Logistic models in previous research as well (Jia, *et al*., 2020). According to the model, the growth coefficient was expected to 0.2443 % and maximum daily cases were estimated to be in the next year. According to this model, the events will get stable only after January 2022.

Like in the Gompertz model, all parameters except the upper asymptote of the Logistic model were significant at 1 % level of significance which can be observed in Table 2. In this case, the majority of parameters are highly significant with a higher R-square value confirmed the validity of the model. The model showed the growth coefficient as 2.73 % and maximum daily cases was expected in March 2021. According to this model, the events will get stable after April 2021.

Model selection

Model having lowered AIC, BIC and DIC and greater loglikelihood was considered as a better model. According to all criteria mentioned in Table 3, the Logistic model was better than the Gompertz model when the total number of confirmed cases was used as a dependent variable.

In the case of total death in Gompertz model, Table 4 shows that all parameters are significant at 1 % level of significance and the R-square value of 99.71 % confirms the validity of the model even in total death cases. The model depicted the growth rate coefficient of 1.17 % and death cases will be flat after 3729. The model estimated maximum daily death cases around the last week of December.

**Significant at 1 %*

Logistic 1687615584 5208.80 5223.37 -2600.4

Table 3. Model selection criteria for total confirmed cases

Table 4. Results of parameters of the Gompertz model for total death cases

Parameters		Estimate Std. Error t-value		R-square
А	3729	341.8	$10.91*$	99.71%
K	0.0117	0.0005056	23.14*	
$\mathbf T$	311	6.593	$47.16*$	
1.771 \sim \sim \sim				

**Significant at 1%*

In the case of total death, all parameters were significant at 1 % level of significance and the R-square value was 99.71 %. All highly significant parameters with higher Rsquare values could not be obtained for not valid model. The model depicted the growth rate coefficient of 1.17 % and death cases will be flat after 3729. The model estimated maximum daily death cases around the last week of December and the cases will get plateau after January 2021.

According to Table 5, the parameters B, K, and T were significant at 1 % and the R-square value was 99.60 % indicating the validity of the model. These evidences validated the model for the total death cases. The model showed a growth coefficient of 3.94 % and maximum daily death cases were expected around October 15. The model estimated the death cases will be flat after 1346 and the cases will get plateau after November 2020. Based on all criteria mentioned in Table 6, the Gompertz model was better than the Logistic model while analyzing total death cases which can be observed from the table.

Automatic ARIMA suggests ARIMA (3, 2, 3) for estimating the total confirmed cases. Table 7 depicts that the values of standard errors are less than the coefficients of moving average (ignoring sign): Ar1, Ar2, Ar3, Ma1, Ma2, and Ma3 indicating the model was not bad for forecasting. Moreover, the accuracy of the model can be checked through ME (mean error), RMSE (root mean square error), MAE (mean absolute error), MPE (Mean Percentage Error), MAPE (mean absolute percentage

error), and MASE (mean absolute square error). These values presented in Table 7 are not so high and MAPE suggests that the model maintains 97.05 % accuracy in prediction. Furthermore, the Box-Pierce test showed that residuals were distributed independently over time. Forecasted values with 80 % and 95 % confidence intervals for October 31, 2020 to December 31, 2020, are presented in Appendix-2. Till October 31, 171,712 confirmed cases were expected according to this ARIMA model and at 95 % confidence interval the values lied between 171,015 and 172,409. The model predicted 347,812 cases till December 31 and the values ranged from 273,889 to 421,734 at 95 % confidence interval.

In case of total death, ARIMA (1, 2, 2) model was suggested by Automatic ARIMA. In this case also, the standard error was less than the coefficient and the values of ME, RMSE, MAE, MPE, MAPE, and MASE, shown in Table 8 are even less than that of confirmed cases recommend that the model can be used for predicting total death cases. According to MAPE, the model seemed to be 96.39 % accurate. Box-Pierce test indicated that there was no evidence of autocorrelation among residuals. Forecasted death cases from October 31, 2020 to December 31, 2020 are shown in Appendix 3. The model predicted 934 death cases till October 31 and at 95 % confidence interval the values lie between 929 and 938. 1,754 death cases are expected till December 31 and at 95 % confidence interval the value range from 1,387 to 2,119.

Table 5. Results of parameters of Logistic model for total death cases.

B 1346	39.	$34.41*$	99.60 %
0.0394 K	0.0007185	54.97*	
264.30 т	1.54	$171.61*$	

**Significant at 1%*

An analysis of COVID-19 cases in Nepal: a modeling approach

Table 6. Model selection criteria for total death cases

Model	DIC.	AIC	BIC-	Log-likelihood
Gompertz	33799.11	1375.87	1388.36	-683.93
Logistic	38072.33	1395.87	1408.37	-693.93

Table 7. Results of ARIMA model for total confirmed cases

	Ar1	Ar2	Ar3	Ma1	Ma ₂	Ma3
Coefficient	-0.4164	-0.2271	0.2455	-0.2729	0.3634	-0.6049
Std. error	0.1916	0.1652	0.1072	0.1893	0.0887	0.0708
ME	RMSE	MAE	MPE	MAPE	MASE	
29.44	350.60	143.90	0.641	2.95	0.2403	
	Box-Pierce test					
Chi-square	Df		p-value			
0.0079			0.9291			

Table 8. Results of ARIMA model for total death cases

DISCUSSION

Ahmadi *et al.* (2020) have researched 'Modeling and forecasting trend of COVID-19 epidemic in Iran until May 13, 2020'using three different models Gompertz, Von Bertalanffy, and least squared error. Their Gompertz model predicted growth coefficients as 0.1 and 0.066 respectively for the number of infected and death cases, which were higher than this study's result in both infected and death cases. The time to reach the total confirmed cases and total death cases' curves flat were much lower than their results. However, the total confirmed cases at that time were almost the same in both studies. Jia *et al.* (2020) investigated COVID-2019 in three different areas of China by using Gompertz, Bertalanffy, and Logistic models. Their results matched with this study showing that the Logistic model is better than the Gompertz model in confirmed cases and the Gompertz model is better than the Logistic model in death cases. Interestingly, the values of the R-square of the Logistic model as well as the Gompertz model for the total confirmed cases were identical to this study. The growth coefficient of the Logistic model of total infected cases shown by their

study was much higher, but the corresponding figure of the Gompertz model for total number of deaths was almost same as Wuhan's result. On the other hand, the inflection points for their study are much earlier than the points shown by this research.

Torrealba-Rodriguez *et al.* (2020) used the Gompertz and Logistic model for analyzing confirmed cases of COVID-19 in Mexico by collecting data till May 8, 2020. They found Gompertz model superior to the Logistic model which does not support the finding of this research. Compared to this study, R-square was more whereas the point of inflection was less. The prediction for the total cases till the end of the epidemic made by the model was around a similar figure. Asadi *et al.* (2020) analyzed COVID-19 cases in Spain using the Gompertz model and found a growth coefficient higher than the results of this study. Faranda *et al.* (2020) studied COVID-19 data of different countries. They applied the Logistic model to the Chinese number of infections and it was observed that both growth coefficient and inflection point were more but R-square was greater than 0.99 in both researches. Kriston (2020) investigated COVID-19 cases by using the

Hierarchical Logistic model from data till March 31, 2020. According to his findings regarding Nepal, the upper asymptote was estimated at 468 which was very less than this study and the growth rate coefficient was negative which was opposite of this research. The inflection point was 60 days earlier than the finding of this study. Azad and Hussain (2020) looked into COVID-19 infected cases of Bangladesh by using Simple exponential, Gompertz, Logistic, and Richards models. They found the Gompertz model better than Logistic with R-square 0.99 and their results do not match with the results of this investigation.

De Natale *et al.* (2020) examined COVID-19 cases in Italy using a Logistic model where they predicted peak of infection around mid-March and saturation after the first week of April. Castorina *et al.* (2020) analyzed coronavirus spreading by using macroscopic growth laws. They compared the results of four countries: China, South Korea, Italy, and Singapore obtained from Gompertz and Logistic models. Both models showed the highest growth rate in South Korea and the lowest in Italy. The growth rate coefficient of the Gompertz model for total confirmed cases of this study was near the coefficient of Italy and Singapore. Razzak (2020) investigated the New Zealand COVID-19 infection rate by using the Gompertz model and estimated infected cases to peak on March 28 with a growth coefficient 0.5. This study has a lower coefficient and a longer period to reach maximum.

Hariharan and Prakash (2020) judged the accuracy of the ARIMA model by MAPE, as in this study. Roy *et al.* (2020) used ARIMA (2,2,2) for forecasting total confirmed cases and their MAE and RMSE values were slightly less than this study's findings. Hernandez-Matamoros *et al.* (2020) forecasted COVID-19 cases per region using ARIMA of different order for each region. Their RMSE values of all regions were much higher than the RMSE of this study. Sahai *et al.* (2020) used the ARIMA model of different orders for different countries to predict COVID-19 in the top five affected countries. They also predicted confirmed cases at 95 % confidence interval for 77 days. Their MAPE value in India was slightly higher than the MAPE of this research.

CONCLUSION

After analyzing COVID-19 data from January 23, 2020 to October 30, 2020, it was found that the patterns of total recovery and the total death were almost similar but in the case of total confirmed cases, it was slightly different. Both the Gompertz and Logistic models were fitted well in analyzing the total number of confirmed cases. Also, based on AIC; the Logistic model was found to be better than the Gompertz model. Likewise, both models would read the pattern of the number of death cases well. The appropriate of these models was justified by higher values of R-square. Gompertz model seemed to be better for

fitting the total number of death cases based on AIC and other criteria as well. After fitting models, confirmed and death cases were forecasted by using the ARIMA model and the accuracy of the model was judged by various criteria and was found to be a good forecaster.

REFERENCES

- Ahmadi, A., Fadai, Y., Shirani, M., & Rahmani, F. (2020). Modeling and forecasting trend of COVID-19 epidemic in Iran until May 13, 2020. *Medical Journal of the Islamic Republic of Iran*, *34*(1), 183- 195.
- Asadi, M., Di Crescenzo, A., Sajadi, F. A., & Spina, S. (2020). A generalized Gompertz growth model with applications and related birth-death processes. *Research Square,1*(1), 1-31.
- Azad, A. K., & Hussain, A. M. (2020). Modeling and analysis of the early growth dynamics of COVID-19 transmission. *Preprints* 2020, 202005037[2.](../../.%20https:/doi.org/10.20944/preprints202005.0372.v1) [https://doi.org/10.20944/preprints202005.0372.v1](../../.%20https:/doi.org/10.20944/preprints202005.0372.v1)
- Bates, D. M., & Watts, D. G. (1988). *Nonlinear regression analysis and its applications* (vol. 2). New York, USA: Wiley.
- Bherwani, H., Anjum, S., Kumar, S., Gautam, S., Gupta, A., Kumbhare, H., ... & Kumar, R. (2020). Understanding COVID-19 transmission through Bayesian probabilistic modeling and GIS-based Voronoi approach: a policy perspective. *Environment, Development and Sustainability* (2020). https://doi.org/10.1007/s10668-020-00849-0.
- Blasco, A., Piles, M., & Varona, L. (2003). A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. *Genetics Selection Evolution*, *35*(1), 21-41.
- Castorina, P., Iorio, A., & Lanteri, D. (2020). Data analysis on coronavirus spreading by macroscopic growth laws. *International Journal of Modern Physics C*, 2050103. <https://doi.org/10.1142/S012918312050103X>
- Cheng, Y., Luo, R., Wang, K., Zhang, M., Wang, Z., Dong, L., ... & Xu, G. (2020). Kidney impairment is associated with in-hospital death of COVID-19 patients. *MedRxiv*. <https://doi.org/10.1101/2020.02.18.20023242>
- Cyranoski, D. (2020). When will the coronavirus outbreak peak? *Nature*, 18 February 2020. *https://doi.org/10.1038/d41586-020-00361-5*
- De Natale, G., Ricciardi, V., De Luca, G., De Natale, D., Di Meglio, G., Ferragamo, A., ... & Spina, E. (2020). The COVID-19 infection in Italy: a statistical study

of an abnormally severe disease. *Journal of Clinical Medicine*, *9*(5), 1564.

- Du, R. H., Liang, L. R., Yang, C. Q., Wang, W., Cao, T. Z., Li, M., ... & Hu, M. (2020). Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. *European Respiratory Journal*, *55*(5). <https://doi.org/10.1183/13993003.00524-2020>
- Fan, C., Lei, D., Fang, C., Li, C., Wang, M., Liu, Y., ... & Yu, Y. (2020). Perinatal transmission of COVID-19 associated SARS-CoV-2: should we worry? *Clinical infectious diseases*, ciaa226. <https://doi.org/10.1093/cid/ciaa226>
- Faranda, D., Castillo, I. P., Hulme, O., Jezequel, A., Lamb, J. S., Sato, Y., & Thompson, E. L. (2020). Asymptotic estimates of SARS-CoV-2 infection counts and their sensitivity to stochastic perturbation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *30*(5), 051107.
- Gautam, S. (2020). COVID-19: air pollution remains low as people stay at home. *Air Quality, Atmosphere & Health*. [https://doi.org/10.1007/s1186 9-020-00842 -](https://doi.org/10.1007/s1186%209-020-00842%20-6%20) [6](https://doi.org/10.1007/s1186%209-020-00842%20-6%20)
- Gautam, S., & Hens, L. (2020). SARS-CoV-2 pandemic in India: what might we expect? *Environment, Development and Sustainability,. https://doi.org/10.1007/s1066 8-020-00739 -5.*
- Guarner, J. (2020). Three emerging coronaviruses in two decades: The story of SARS, MERS, and now COVID-19. *American Journal of Clinical* Pathology. [https://doi.org/10.1093/ajcp/aqaa0 29](https://doi.org/10.1093/ajcp/aqaa0%2029)
- Guo, Y. R., Cao, Q. D., Hong, Z. S., Tan, Y. Y., Chen, S. D., Jin, H. J., ... & Yan, Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak–an update on the status. *Military Medical Research*, *7*(1), 1-10.
- Hariharan, B. & Prakash, P.N. (2020). Prediction of COVID-19 infected cases using ARIMA: A case study.<https://ssrn.com/abstract=3641767>
- Harvey, A., & Kattuman, P. (2020). Time series models based on growth curves with applications to forecasting coronavirus. *Covid Economics, Vetted and Real-Time Papers*, (24). 126-157.
- He, L., Ding, Y., Zhang, Q., Che, X., He, Y., Shen, H., ... & Deng, Y. (2006). Expression of elevated levels of pro-inflammatory cytokines in SARS‐CoV‐infected ACE2+ cells in SARS patients: relation to the acute lung injury and pathogenesis of SARS. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, *210*(3), 288-297.

Hernandez-Matamoros, A., Fujita, H., Hayashi, T., & Perez-Meana, H. (2020). Forecasting of COVID19 per regions polynomial functions. *Applied Soft Computing*, 96, 106610. <https://doi.org/10.1016/j.asoc.2020.106610>

https://www.covid19.mohp.gov.np

https://www.wikepedia.org

https://www.worldometers.info

- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., ... & Cheng, Z. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, *395*(10223), 497-506.
- Hui, D. S., Memish, Z. A., & Zumla, A. (2014). Severe acute respiratory syndrome vs. the Middle East respiratory syndrome. *Current opinion in pulmonary medicine*, *20*(3), 233-241.
- Jain, N., Choudhury, A., Sharma, J., Kumar, V., De, D., & Tiwari, R. (2020). A review of novel coronavirus infection (Coronavirus Disease-19). *Global Journal of Transfusion Medicine*, *5*(1), 22.
- Jia, L., Li, K., Jiang, Y., & Guo, X. (2020). Prediction and analysis of Coronavirus Disease 2019. *arXiv preprint arXiv:2003.05447*.
- Kannan, S., Ali, P. S. S., Sheeza, A., & Hemalatha, K. (2020). COVID-19 (Novel Coronavirus 2019)-recent trends. *Eur. Rev. Med. Pharmacol. Sci*, *24*(4), 2006- 2011.
- Kriston, L. (2020). Projection of cumulative coronavirus disease 2019 (COVID-19) case growth with a hierarchical logistic model. *Bull World Health Organ COVID-19 Open Preprints. http://dx. doi. org/10.2471/BLT*, *20*.
- Majumder, M. S., Rivers, C., Lofgren, E., & Fisman, D. (2014). Estimation of MERS-coronavirus reproductive number and case fatality rate for the spring 2014 Saudi Arabia outbreak: insights from publicly available data. *PLoS currents*, *6*. [https://doi.org/10.1371/currents.outbreaks.98d2f8f33](https://doi.org/10.1371/currents.outbreaks.98d2f8f3382d84f390736cd5f5fe133c) [82d84f390736cd5f5fe133c](https://doi.org/10.1371/currents.outbreaks.98d2f8f3382d84f390736cd5f5fe133c)
- Malki, Z., Atlam, E. S., Ewis, A., Dagnew, G., Alzighaibi, A. R., ELmarhomy, G., ... & Gad, I. (2020). ARIMA models the risk of second rebound. *Neural Computing and Applications*, 1-20. [https://doi.org/10.1007/s00521-020-05434-0.](https://doi.org/10.1007/s00521-020-05434-0)
- Martelloni, G., & Martelloni, G. (2020). Analysis of the evolution of the Sars-Cov-2 in Italy, the role of the asymptomatics and the success of Logistic model. *arXiv Preprint arXiv:2004.02224*. <https://doi.org/10.1016/j.chaos.2020.110150>
- Martinez, E. Z., Aragon, D. C., & Nunes, A. A. (2020). Short-term forecasting of daily COVID-19 cases in Brazil by using the Holt's model. *Revista da Sociedade Brasileira de Medicina Tropical*, *53*.
- MoHP (2020).*COVID-19 situation report.* Kathmandu, Nepal: Ministry of Health and Population.
- Razzak, W. A. (2020). *Modeling New Zealand COVID-19 infection rate, and the efficacy of social distancing policy* (No. 2004). New Zealand: School of Economics and Finance, Massey University.
- Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling, 5*(1), 271-281.
- Rothan, H. A., & Byrareddy, S. N. (2020). The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of Autoimmunity*, 102433. <https://doi.org/10.1016/j.jaut.2020.102433>
- Roy, S., Bhunia, G. S., & Shit, P. K. (2020). Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. *Modeling Earth Systems and Environment*, 1-7. [https://doi.org/10.1007/s40808-](https://doi.org/10.1007/s40808-020-00890-y) [020-00890-y](https://doi.org/10.1007/s40808-020-00890-y)
- Sahai, A. K., Rath, N., Sood, V., & Singh, M. P. (2020). ARIMA modeling & forecasting of COVID-19 in top five affected countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(5), 1419-1427.
- Sarkodie, S. A., & Owusu, P. A. (2020). Global assessment of environment, health and economic impact of the novel coronavirus (COVID-19). *Environment, Development and Sustainability*, 1-11. [https://doi.org/10.1007/s10668-020-00801-2](../../%20https:/doi.org/10.1007/s10668-020-00801-2)
- Seber, G. A., & Wild, C. J. (2003). *Nonlinear regression*. New Jersey, USA: John Wiley & Sons, pp. 62-63.
- Silva, R. R., Velasco, W. D., da Silva Marques, W., & Tibirica, C. A. G. (2020). A Bayesian analysis of the total number of cases of the COVID 19 when only a few data is available. A case study in the state of Goias, Brazil. *medRxiv*.
- Suwantarat, N., & Apisarnthanarak, A. (2015). Risks to healthcare workers with emerging diseases: lessons from MERS-CoV, Ebola, SARS, and avian flu. *Current opinion in infectious diseases*, *28*(4), 349-361.
- Team, R. C. (2019). R: A language and environment for statistical computing (version 3.1. 2). Vienna, Austria: R Foundation for Statistical Computing.
- Timur, E. V., Mullen, K. M., Spiess, A. N., & Bolker, B. (2015). minpack. lm: R Interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK, plus support for bounds. <https://doi.org/10.1038/s41598-018-27594-9>
- Torrealba-Rodriguez, O., Conde-Gutiérrez, R. A., & Hernández-Javier, A. L. (2020). Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos, Solitons & Fractals*, 109946. https://doi.org/10.1016/j.chaos.2020.109946.
- Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., ... & Dighe, A. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases*, *20*(1), 669-677.
- Vuori, K., Strandén, I., Sevón-Aimonen, M. L., & Mäntysaari, E. A. (2006). Estimation of non-linear growth models by linearization: a simulation study using a Gompertz function. *Genetics Selection Evolution*, *38*(4), 1-16.
- WHO. (2020). *Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel regulations-(2005)-emergency-committeeregarding-the-outbreak-of-novel-coronavirus-(2019 ncov) accessed on February 21, 2020.* WHO.
- Yang, Q., Wang, J., Ma, H., & Wang, X. (2020). Research on COVID-19 based on ARIMA modelΔ- taking Hubei, China as an example to see the epidemic in Italy. *Journal of Infection and Public Health*, *13*(10), 1415-1418.

APPENDIX

Appendix 1. Summary statistics

Appendix 2. Forecasted confirmed cases with confidence intervals

14-Dec	327	298782.9	268019.0	329546.8	251733.6	345832.3	
15 -Dec	328	301667.0	269943.8	333390.2	253150.6	350183.4	
16 -Dec	329	304551.0	271858.9	337243.2	254552.7	354549.4	
17 -Dec	330	307435.1	273764.3	341105.9	255940.1	358930.1	
18 -Dec	331	310319.2	275660.2	344978.1	257312.9	363325.4	
19 -Dec	332	313203.2	277546.7	348859.8	258671.2	367735.2	
20 -Dec	333	316087.3	279423.8	352750.8	260015.3	372159.3	
21 -Dec	334	318971.4	281291.5	356651.2	261345.1	376597.6	
22 -Dec	335	321855.4	283150.1	360560.7	262660.8	381050.0	
23 -Dec	336	324739.5	284999.6	364479.3	263962.6	385516.3	
24 -Dec	337	327623.5	286840.1	368407.0	265250.6	389996.5	
25 -Dec	338	330507.6	288671.6	372343.6	266524.9	394490.3	
26 -Dec	339	333391.7	290494.2	376289.2	267785.6	398997.7	
$27 - Dec$	340	336275.7	292307.9	380243.5	269032.8	403518.7	
28-Dec	341	339159.8	294113.0	384206.6	270266.6	408052.9	
29 -Dec	342	342043.8	295909.4	388178.3	271487.2	412600.5	
30 -Dec	343	344927.9	297697.1	392158.7	272694.7	417161.1	
31 -Dec	344	347812.0	299476.4	396147.6	273889.0	421734.9	

Appendix 3. Forecasted death cases with confidence interval

