

A two-stage polynomial and nonlinear growth approach for modeling the COVID-19
outbreak in Mexico

Rafael Pérez Abreu C., PhD
ORCID: 0000-0003-3076-3714
Centro de Investigación en Matemáticas, A.C.
Calzada de la plenitud 103
Fracc. José Vasconcelos Calderon, CP 20200
Aguascalientes, Ags. Mexico

Samantha Estrada Aguilera, PhD (**Corresponding Author**)
ORCID: 0000-0002-8661-7292
University of Texas at Tyler (**Present Address**)
Psychology and Counseling
3900 University Blvd
Tyler, Texas 75799

Héctor de la Torre Gutiérrez, PhD
ORCID: 0000-0002-8343-989X
Research fellow at the Mathematics Research Centre
Avenida de la Plenitud 103, José Vasconcelos. 20200
Aguascalientes, Ags. Mexico.

Highlights

1. Nonlinear model with an autoregressive term(s) is proposed to model total cases
2. The proposed models consider the autocorrelation in the observations
3. The proposed methodology improves fit and meets the assumption of independence
4. The model parameters provide further information of the COVID-19 pandemic
5. The nonlinear growth models fit better past the peak of the pandemic

Abstract

Since December 2019, the coronavirus disease (COVID-19) has rapidly spread worldwide. The Mexican government has implemented public safety measures to minimize the spread of the virus. In this paper, the authors use statistical models in two stages to estimate the total number of coronavirus (COVID-19) cases per day at the state and national level in Mexico. Two types of models are proposed: first, a polynomial model of the growth for the first part of the outbreak until the inflection point of the pandemic curve and then a second nonlinear growth model is used to estimate the middle and the end of the outbreak. Model selection will be performed using Vuong's test. The proposed models show overall fit similar to predictive models (e.g. time series, and machine learning); however, the interpretation of parameters is less complex for decision-makers and the residuals follow the expected distribution when fitting the models without autocorrelation being an issue.

Keywords: COVID-19, epidemic modeling, time series forecasting, nonlinear growth models, Prais-Winsten estimation

Introduction

The world is currently experiencing a pandemic caused by the novel coronavirus, formally named COVID-19 by the World Health Organization (WHO). Development of a vaccine and antiviral drugs to treat COVID-19 is still ongoing resulting in hospitalization and intensive care unit management as the only option in treating COVID-19. Thus, there is a dire need for research on modeling the outbreak of COVID-19 to help officials in their decision-making processes regarding interventions and allocation of resources [22]. At the moment this manuscript was being written, the pandemic was ongoing and most of the epidemiological models developed focused on short-term predictions, identifying the daily peak of COVID-19 cases, predicting the duration of the pandemic, and estimating the possible impact of the measures implemented for minimizing exposure to the virus and decrease the fatality rate [1][3][11][12][14][15][17][25].

As of October 24, 2020, the cumulative number of COVID-19 cases in Mexico was reported 874,171 and 42, 055, 863 cases worldwide [24]. Thus, the main objective of this paper is to model the total number of COVID-19 cases per day at the national and the state level in Mexico while simultaneously providing straightforward information to decision-makers; additionally, the authors seek to determine which model provides the most stable short-term predictions. The models developed in this research support facilitate obtaining information that supports the strategic planning activities of the Mexican states, metropolitan areas, municipalities, or cities with high population density. Mexican officials can use these models to aid in the management process involving the needs and resources of the health services such as available hospital beds, intensive care units, and respirators, personal protective equipment (PPE) for health personnel. For stakeholders, such as public health officials, having access to a daily and permanent monitoring at the center of the pandemic allows them to anticipate the purchase of health needs in advance.

Further, the authors would like to share these models so that officials and statisticians outside of Mexico can make use of them for their own decision-making procedures during the length of COVID-19 pandemic. The proposed methodology in this paper can easily be applied to COVID-19 worldwide.

Literature Review

Research exists with data-driven approaches such as AR (Autoregressive), ARIMA (Autoregressive Integrated Moving Average) ranging from simple models (exponential smoothing) to more complex models such as ARIMAX, ARCH, GARCH, ARFIRMA [2][6][7][9][10]. For example, an ARIMA model with data compiled by John Hopkins University was used to predict models for the daily confirmed cases in countries where the pandemic was peaking and to predict and anticipate the resources of the healthcare systems [5]. Unfortunately, these data-driven models fail to fit the data and often lack accuracy [4][5][14]. More importantly, the parameters cannot be interpreted according to the reality of the pandemic, in other words, it is not possible to obtain information from the growth of the pandemic which causes statisticians, and officials to make their decisions based on predictive models not based on the peak of the pandemic or the growth of the pandemic. Another approach is real-time forecasting using a generalized Logistic growth model. This method has been previously used in China to generate short-term forecasting of COVID-19 cases [16][20] and with data from Canada, France, India, South Korea, and the UK to forecast daily cases [4]. These models are incredibly useful, they provide information on the current state of the pandemic; however, in this study, we aim to demonstrate that the assumption of independence of observations is not met and that its modeling performance at the earlier stages of the pandemic is not optimal [15][16].

The models used in this paper are based on statistical linear models, classic time series and restricted growth called limited growth or nonlinear growth models [3][8][9][12][13][21][22], and real-time forecasting using generalized Logistic growth model [16][19]. In this paper we propose to estimate in two stages of the pandemic utilizing polynomial and nonlinear growth functions while incorporating an autoregressive component with the purpose of meeting the assumption of independence of residuals. First, we propose to use a polynomial function estimated using the Prais-Winsten methodology to estimate the first stage of the pandemic (when exponential growth of COVID-19 cases is observed). Next, in the second stage of the pandemic, (when the peak of COVID-19 cases has been reached) we propose to utilize nonlinear growth functions, Logistic and Gompertz, in to predict the total cumulative number of cases and the growth rate of the spread of COVID-19. In each of the models estimated before and after the peak of the pandemic, at the second stage of modeling we added an autoregressive component of order one (AR(1)) to compare the results to the models that do not account for the violation of independence of residuals. This approach has been successfully used to model plant and animal growth where measuring the same unit can lead to violation of independence of residuals [15][16]. Next, we will select the most useful model to modeling the pandemic by using Vuong's test [23]. We anticipate the proposed model to have good performance similar to Neural Networks (NN) or Support Vector Machines (SVM). A disadvantage of NN is that it is difficult to generate a day-to-day prediction in addition to finding the growth rate and the maximum number of cases. These are not a problem for the functions we propose. Furthermore, artificial intelligence (AI) has been used to identify, track, and forecast COVID-19 cases. However, AI models are difficult to interpret, the process by which they arrive at a decision is often referred to as a "black box" due to the complexity in understanding how AI models arrive at a certain conclusion [26]. The complex interpretation may create a barrier

for decision-makers looking for straightforward solutions in the middle of a global pandemic. Moreover, we anticipate our proposed modeling of the pandemic will meet the assumption of independence of residuals. Additionally our proposed model, in contrast to NN, will provide predictions and will make it possible to interpret parameters such as the highest number of infected, growth speed, the initial number of infected and the autoregressive parameter to measure the lag in reporting the new daily cases of infections.

Therefore, the objectives of this paper were to (1) specify the polynomial function and nonlinear growth models (Logistic and Gompertz) that include an autoregression component for dealing with the autocorrelated observations in the growth data in two stages: before and after the inflection point of the pandemic is reached, and (2) compare the different polynomial and nonlinear growth functions in their ability to describe the number of cases of the COVID-19 pandemic.

Methods

Dataset

The COVID-19 data used in this research was obtained from the publicly available data of the Mexican Secretaria de Salud Federal which contained the number of cases confirmed from February 27th to September 24th, 2020. The dataset also included the number of recovered patients and fatalities and can be downloaded in .csv format from the government's website (<https://www.gob.mx/salud/documentos/datos-abiertos-152127>). Data to test the models focused on the national level and three Mexican states: Campeche, Quintana Roo, and Tamaulipas. To demonstrate the benefit of utilizing the autoregressive term we will use data from the state of Aguascalientes. The time series beginning point ($t = 1$) established was the day where positive COVID-19 had not been not reported. For example, the state of Tamaulipas, the first positive

COVID-19 case was March 17, 2020, thus March 16, 2020 was considered ($t=1$). Analyses were conducted in R (version 3.6.2) and STATA (version 15.1)

Model Selection

This paper proposes the utilization of a two stage approach: First, Stage 1 model was fit throughout the pandemic and before reaching the peak of cases (before the inflection point), and in the second stage once the peak of cases has been reached a different type of model will be used. We will utilize Vuong test's for model selection [23].

Stage 1: Before the inflection point

In this paper, we examined a variety of models that successfully model the behavior of the pandemic (e.g. the maximum number of cases, growth rate, etc.) while simultaneously meeting the assumption of independence of residuals in the data. Using the Akaike Information Criteria (AIC) and the Root Mean Square Error (RMSE) criteria, the authors found that the models that best describe the total accumulated cases of COVID-19 in the four data examples utilized were:

- a. Polynomial model of order three with an autoregressive error component of order one (Equation 1), known as Prais-Winsten or Cochrane-Orcutt estimation, and
- b. Nonlinear growth models, including Logistic and Gompertz, had the best fit of the models examined (Equation 2 and 3).

Stage 2: After the inflection point

To examine which were the best models once the peak of positive COVID-19 cases had been reached, we focused on the same three state data which showed good fit up to the peak of the pandemic. We sampled to data points after the inflection point: 20 days after and the day before we began writing the results of this paper (September 24, 2020). The results obtained from this

stage of the pandemic will be compared with those obtained before the inflection point of the pandemic. We hypothesized that for the stage before the peak of cases one of the nonlinear growth models will have better fit.

Statistical Procedures

The complete polynomial model used in this work has the form:

$$Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \rho u_{t-1} + \varepsilon_t \quad (1)$$

where Y_t is the number of total positive COVID-19 cases reported at time t , the coefficients $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and ρ are the parameters of the model to estimate with u_{t-1} the autoregressive component and ε_t is a random error term, with $t = 0, 1, 2, \dots, i$. The value $t = 1$ is selected equal to the day of the first positive case.

The Logistic model used the following form:

$$Y_t = \frac{\beta_1}{1 + \beta_2 e^{-\beta_3 t}} + \rho u_{t-1} + \varepsilon_t \quad (2)$$

where Y_t is the number of total positive COVID-19 cases reported at time t , the coefficients $\beta_1, \beta_2, \beta_3$ and ρ are the model parameters to estimate with u_{t-1} the autoregressive component and ε_t is a random error term, with $t = 0, 1, 2, \dots, i$. β_1 models the highest number of infected, β_2 growth speed, β_3 is the initial number of infected and ρ models the autoregressive process to incorporate the delay in the process of reporting the new cases of infection as well as the inherent pandemic dynamic. Recall, the value $t = 1$ is set to the day of the first positive case observed.

The Gompertz model used the following form:

$$Y_t = \beta_1 e^{(-e^{(\beta_2(t-\beta_3))})} + \rho u_{t-1} + \varepsilon_t \quad (3)$$

where Y_t is the number of total positive COVID-19 cases reported at time t , the coefficients $\beta_1, \beta_2, \beta_3$ and ρ are the model parameters to estimate with u_{t-1} the autoregressive component and ε_t is a random error term, with $t = 0, 1, 2, \dots, i$. The β_1 parameter estimates the

top number of infected, β_2 growth speed, β_3 is the initial number of infected, and ρ models the autoregressive component. Similarly, the value $t = 1$ is set to the day of the first positive case observed.

Recall, part of the proposed work in this paper is to fit nonlinear growth models and the polynomial models in two stages. The models were fitted in six time points, the end date of the study (September 24, 2020), 20 days before the peak of the pandemic for each of the Mexican states examined, peak day, and 10, 20 and 30 days after the peak of cases was reached. The models were compared in pairs utilizing Vuong's test, a classical likelihood ratio approach to model selection for nested and non-nested models, which uses the Kulback-Leiber information criterion [23]. The hypotheses for Vuong's model selection test are:

H_0 : Model fits are equal for the focal population

H_{1A} : Model 1 fits better than Model 2

H_{1B} : Model 2 fits better than Model 1

Thus, the results will provide a p-value from Vuong's test which can be compared to a significance level α and aid in selecting the model that best fits the data. Graphically, we can observe in Figure 1 the approximate modeling of the pandemic, identify the peak (when there is a change in the growth rate), and the steps where the models of interest in the project will be used.

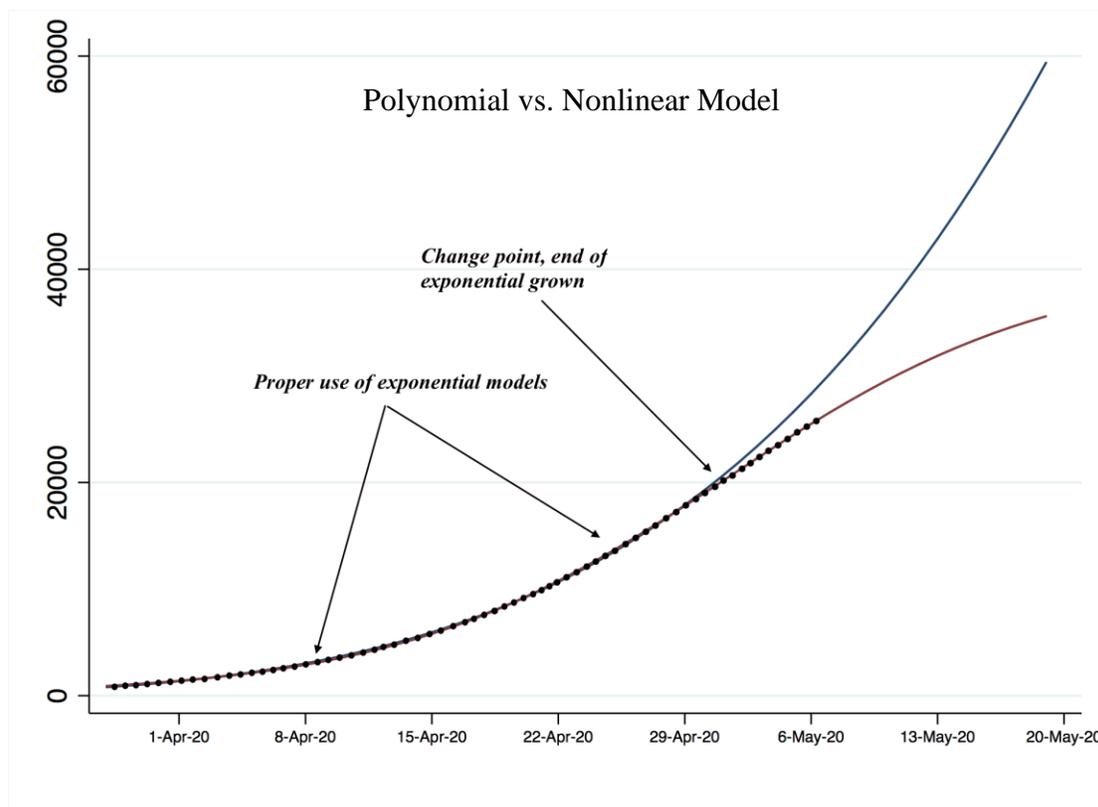


Figure 1. Polynomial versus a nonlinear growth model

Results

The dates when the peak of COVID-19 cases was reached at the national level and in the three Mexican states selected for the study can be seen in Table 1.

Table 1

COVID-19 Peak dates.

	Peak date
México (National Level)	August 1st
Campeche	July 25th
Quintana Roo	July 23rd
Tamaulipas	August 1st

As previously mentioned, the proposed models were adjusted to six time points. We will use Young's test, for each time point, to see which model fits best. The p-values obtained via

Young's test for the alternative hypothesis A1 (H_{1A}) can be seen in Table 2. Note that, the p-values for the alternative hypothesis B1 (H_{1B}) are the complement of the p-values for the alternative hypothesis A1 (H_{1A}). For example, comparing the polynomial and Logistic models, in Table 2 we can see the p-value obtained for the state of Tamaulipas on September 24, was 0.9703 for the H_{1A} and its complement 0.0397 is the p-value for the H_{1B} . At the significance level $\alpha = 0.05$, we reject the null hypothesis in favor of H_{1B} implying the Logistic model fits the data better than the polynomial model.

Table 2

Vuong's model selection test p-values corresponding to the H_{1A}

Date	Models (Model 1 – Model 2)	Mexico	Campeche	Quintana Roo	Tamaulipas
24th September	Polynomial - Logistic	0.8002	1.0000	0.9999	0.9703
	Polynomial - Gompertz	0.9688	1.0000	0.9176	0.9672
	Logistic - Gompertz	0.9650	0.0000	0.0000	0.0838
20 days after the peak day	Polynomial - Logistic	0.1301	0.9676	0.9948	0.9534
	Polynomial - Gompertz	0.6976	0.9718	0.9904	0.9672
	Logistic - Gompertz	0.9414	0.0525	0.0044	0.1852
Peak day	Polynomial - Logistic	0.4070	0.5492	0.0519	0.7105
	Polynomial - Gompertz	0.8021	0.6610	0.0043	0.7793
	Logistic - Gompertz	0.7601	0.4918	0.3863	0.4909
10 days before peak day	Polynomial - Logistic	0.1448	0.7669	0.3247	0.4770
	Polynomial - Gompertz	0.9668	0.6496	0.1247	0.8257
	Logistic - Gompertz	0.9944	0.2151	0.2550	0.6823
20 days before peak day	Polynomial - Logistic	0.2096	0.9336	0.8939	0.9325
	Polynomial - Gompertz	0.9005	0.9712	0.7853	0.9458
	Polynomial - Gompertz	0.9641	0.4513	0.0592	0.0791
30 days before peak day	Polynomial - Logistic	0.4802	0.8805	0.2162	0.7986
	Polynomial - Gompertz	0.9474	0.9262	0.3218	0.7929
	Logistic - Gompertz	0.9800	0.2506	0.8606	0.7532

Table 3 summarizes which models fit better for each region at each of the six time points examined. In general, it is easy to see that the nonlinear growth models do not fit better than the

polynomial models on dates after the peak of the pandemic has been reached. On the other hand, when examining the dates prior the peak of the pandemic there was a negligible difference between the models. More importantly, examining the model comparisons in Table 3 for dates before the peak and during the peak of the pandemic it is clear that no model worked better than the others. In other words, when modeling growth rate of COVID-19 cases there is no difference between the models. However, in the late phase of the pandemic, after the inflection point, the nonlinear growth models perform better than the polynomial model fit.

Table 3

Model comparison summary table according to Vuong's test

Date	Mexico	Campeche	QRR	Tamaulipas
24th September	Gompertz	Logistic	Logistic	Gompertz, Logistic
20 days after the peak day	Any	Gompertz, Logistic	Logistic	Gompertz, Logistic
Peak day	Any	Any	Any	Any
10 days before peak day	Gompertz	Any	Any	Any
20 days before peak day	Any	Any	Any	Any
30 days before peak day	Any	Any	Any	Any

Polynomial and nonlinear growth models were useful for modeling the beginning of the epidemic, see Figure 1, until reaching the maximum peak of daily cases for the pandemic [4]. On the other hand, nonlinear growth models are more accurate and effective when more information is available, and the maximum peak of daily cases has been reached. Furthermore, time series models allow practical real-time monitoring of when (a) exponential growth is beginning, (b) exponential growth is in effect, and (c) exponential growth is about to end which indicates that the epidemic is reaching its end. Finally, the nonlinear growth models are useful to describe the behavior at the end of the pandemic and allow monitoring and detection of a possible double peak of the epidemic. In general, the Logistic and Gompertz models had the better fit. However, as the pandemic advanced the Gompertz estimated the late phases of the pandemic better.

Once the models are estimated, it is possible to calculate the cumulative total case forecasts for COVID-19. Forecasting the total cases and the most recent information on rates (or percentage) of positive active COVID-19, outpatients, stable hospitalized patients, seriously hospitalized, and intubated hospitalized patients. Likewise, with the information from the SENTINEL Prevention Model, the total number of asymptomatic COVID-19 positives are estimated. The fit of these models can be examined graphically below in Figures 3 to 6 for three Mexican states of Campeche, Quintana Roo, and Tamaulipas and the estimate at the national level. These results are based on the official data source of the daily releases issued by the Ministry of Health on its website <https://coronavirus.gob.mx/> as of September 24th, 2020.

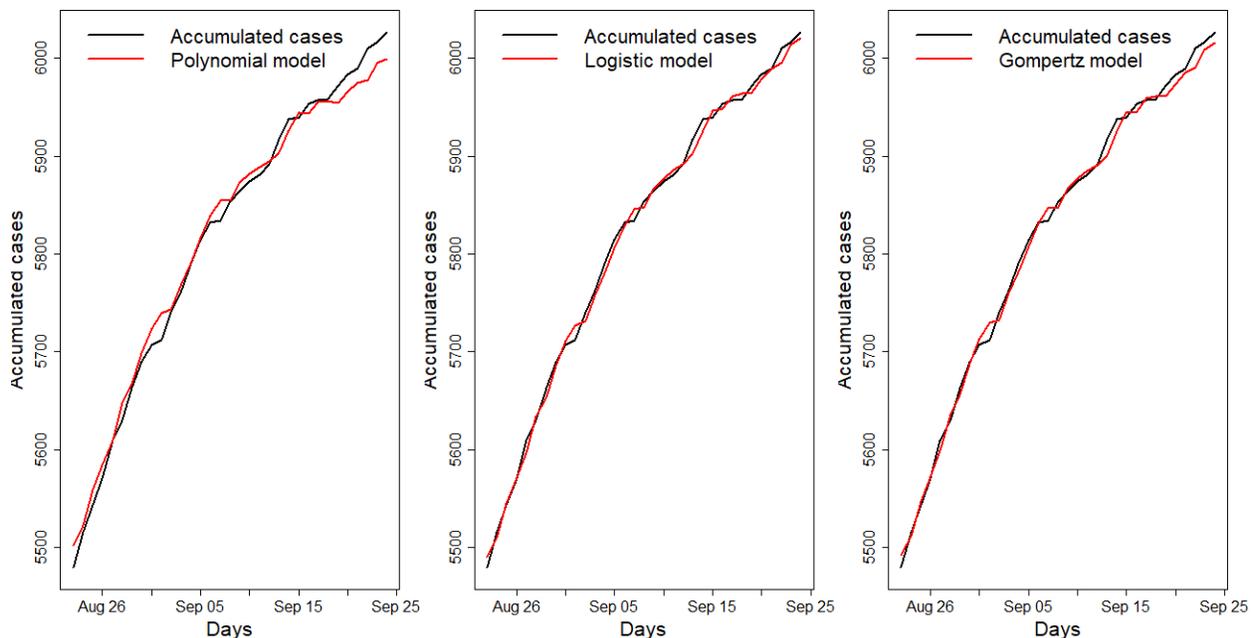


Figure 3. Fit of the three models eight weeks before the end of data collection on September 24 for the state of Campeche

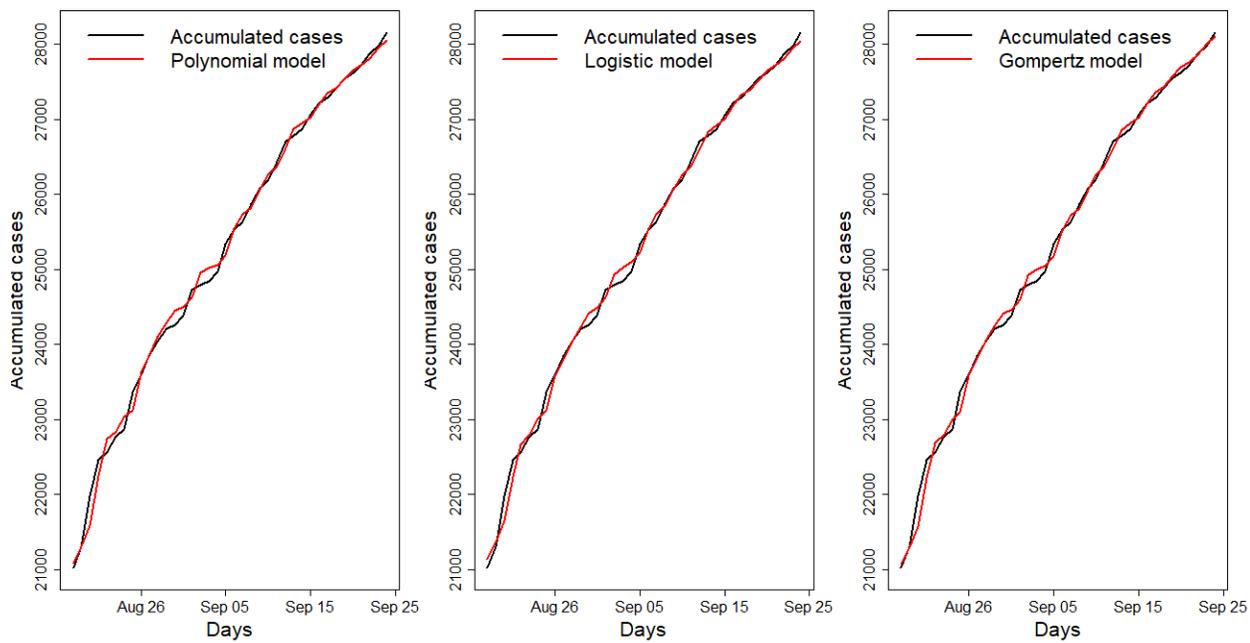


Figure 4. Fit of the three models eight weeks before the end of data collection on September 24 for the state of Tamaulipas

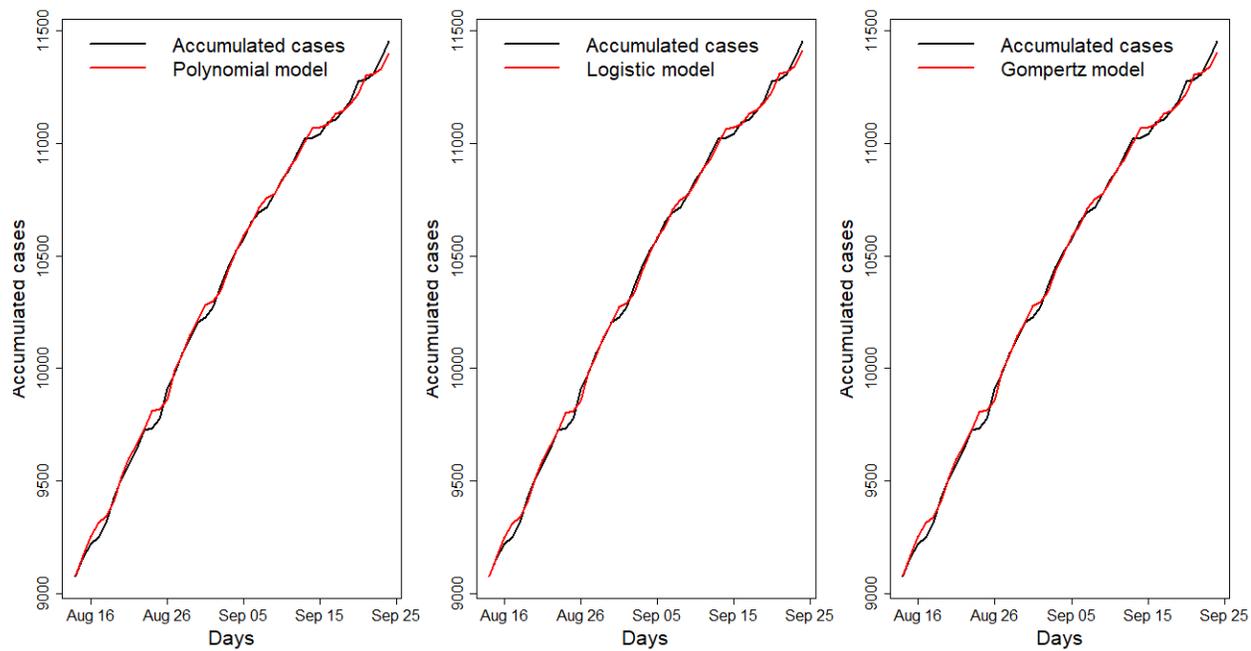


Figure 5. Fit of the three models eight weeks before the end of data collection on September 24 for the state of Quintana Roo

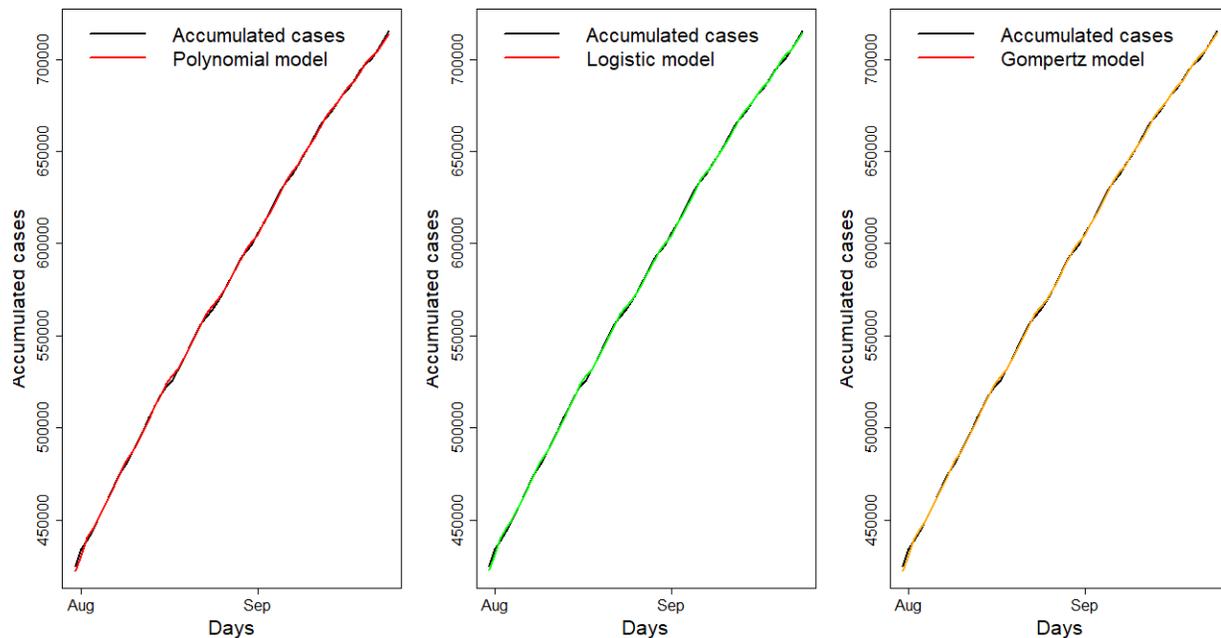


Figure 6. Fit of the three models eight weeks before the end of data collection on September 24 at the national level in Mexico

Autocorrelation

An autoregressive model was fit to the Logistic and Gompertz functions with the purpose of eliminating the residual autocorrelation. The analysis revealed that the inclusion of an autoregressive component of the second order improved the fit of the data. For example, for the state of Aguascalientes, Figures 7a and 7b show the autocorrelation residual and partial autocorrelation plots for models without the autoregressive terms. In these figures, it is clear that the residuals show a pattern of dependency. Furthermore, Figures 8a and 8b demonstrate a good fit and that the assumption of independence has been met. The patterns in the data reveal that nonlinear models with autoregressive terms met assumptions of independence. The models showed good fit, fixed the issue of autocorrelation of residuals while providing better interpretability of the model coefficients in terms of growth rate, maximum number of cases and

initial number of cases. Thus, the model proposed in this paper produces a substantial improvement of the predictions.

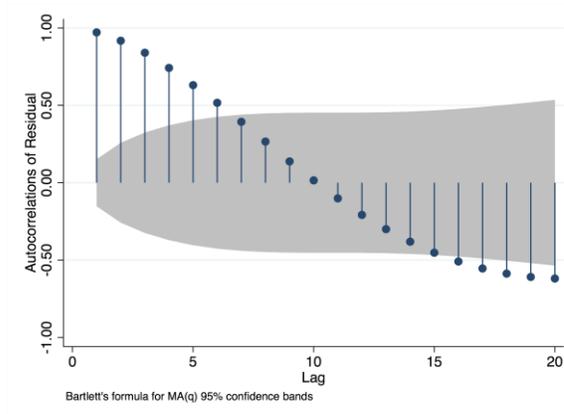


Figure 7(a)

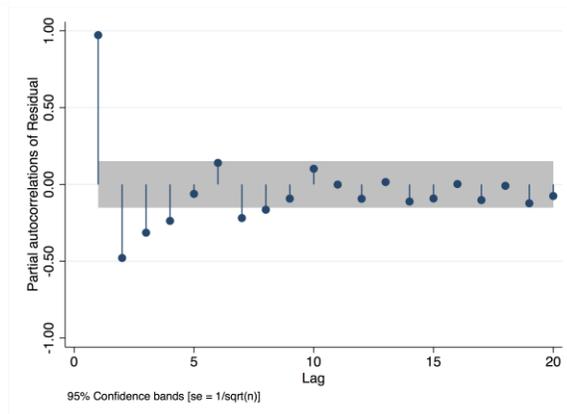


Figure 7(b)

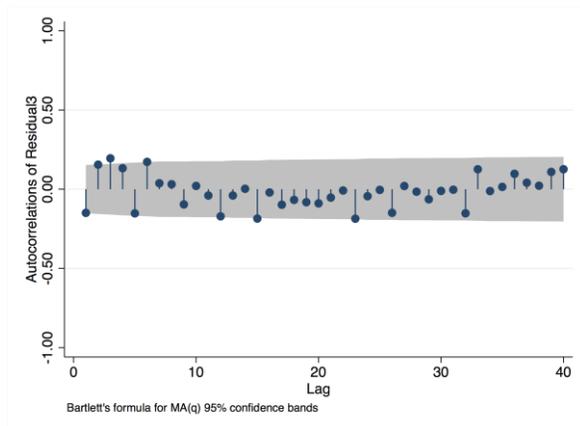


Figure 8(a)

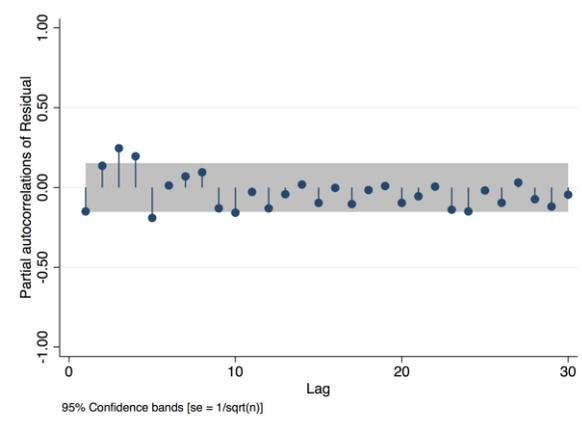


Figure 8(b)

Figures 7-8. Autocorrelation residuals

Limitations

The research we presented in this paper is not without limitations. The projections resulting from the models were estimated without considering any type of intervention. In other words, the intervention effects such as “stay at home” and “social distancing” are not considered in the model.

Due to the characteristics of the models used and the atypical characteristics of the COVID-19 pandemic, the results and information derived from the monitoring strategy, well known in epidemiology science as “SENTINEL contention model,” are considered to adjust:

1. The actual data shows a lot of variability from one day to the next; and
2. Projections are very successful in a time frame of 1 to 10 days after which it is necessary to re-update the models to maintain the desired monitoring, precision, and reliability.
3. In some cases, such as the state of Aguascalientes, it was necessary to add a second order autoregressive component.

The benefit of utilizing a second order autoregressive component is shown Table 4, for the state of Aguascalientes (before August 20th, 2020). For the Gompertz nonlinear growth model, we corrected for the dependency between observations by including an autocorrelation term of the second order in the model which improved considerably the overall fit criteria. Thus, we recommend the inclusion of an autoregressive term of the second order when modeling COVID-19 case growth.

Table 4

Comparison models by goodness of fit criteria for the state of Aguascalientes.

Model Type	Durbin-Watson	AIC	BIC
Gompertz	(3, 139) = 0.0569	1509.95	1519.08
Gompertz + AR(1)	(4, 139) = 1.1040	968.83	980.56
Gompertz + AR(1) + AR(2)	(5, 139) = 2.2986**	937.6	952.27

Note. Compared to Durbin-Watson statistical table values.

* p-value > 0.01

** p-value > 0.05

Discussion

The modeling of the COVID-19 outbreak is a unique challenge for statisticians considering the data is limited, and there are often delays in updating the data. Our approach to modeling the pandemic can provide assistance to decision-making officials for containing and anticipating a

“double peak” of the COVID-19 pandemic in Mexico. We divided the pandemic data into two stages, in stage one we implemented a polynomial model of order three with an autoregressive error component of order one applying the Prais-Winsten or Cochrane-Orcutt estimation methodology to control the autocorrelation in the data. In stage two, when the peak of COVID-19 cases has been reached we fit nonlinear growth functions, Logistic and Gompertz, to predict the total cumulative number of cases and the growth rate of the spread of COVID-19. Further, the models used in this paper are different than those used in predictive models using time series or machine learning; however, our model meets the assumption of independence of residuals and the interpretability of our model is superior to those of machine learning particularly for government officials without a statistics or machine learning background.

In summary, this paper showed the efficiency of utilizing two different types of models estimated in two stages (stages depending on the state of the pandemic). The majority of the efforts to model the COVID-19 pandemic are nonlinear models, such as Logistic and Gompertz [4][5]. However, these models do not take into account autoregression thus making short, medium, and long-term predictions that could possibly be skewed. We recommend, due to parsimony in fitting the model, that in the early stages of the pandemic where there is an exponential growth (during or before the peak) to utilize a polynomial model of the third order estimated in with an autoregressive component. For the later, more advanced stages when the peak of the pandemic has been reached, we recommend utilizing a nonlinear growth model (Logistic or Gompertz) estimated with an autoregressive component.

References

- [1] M Batista. Estimation of the final size of the COVID-19 epidemic, MedRxiv.doi. 10 (2020) 16-20023606.
- [2] D Benvenuto, M Giovanetti, L Vassallo, S Angeletti, M Ciccozzi. Application of the ARIMA model on the COVID-2019 epidemic dataset, Data in Brief. 29 (2020) 105340.
- [3] GE Box, GM Jenkins, GC Reinsel, GM Ljung, Time series analysis: forecasting and control, John Wiley & Sons 2015.
- [4] T Chakraborty, I Ghosh. Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis, Chaos, Solitons & Fractals. 135 (2020) 109850.
- [5] VKR Chimmula, L Zhang. Time series forecasting of COVID-19 transmission in Canada using LSTM networks, Chaos, Solitons & Fractals. 135 (2020) 109864.
- [6] S Choi, E Jung, BY Choi, Y Hur, M Ki. High reproduction number of Middle East respiratory syndrome coronavirus in nosocomial outbreaks: mathematical modeling in Saudi Arabia and South Korea, J.Hosp.Infect. 99 (2018) 162-168.
- [7] G Coutin Marie. Utilización de modelos ARIMA para la vigilancia de enfermedades transmisibles, Revista Cubana de Salud Pública. 33 (2007) 0-0.
- [8] A Gordillo, Estimación de Curvas de Crecimiento en Ovinos Ccriollos Cchapingo a través de Modelos Lineales y No-Lineales, (2019).
- [9] A Grzegorzcyk. Application of the Richards function to the description of leaf area growth in maize (*Zea mays* L.), Acta Soc.Bot.Pol. 63 (1994) 5-7.
- [10] K Kim, T Tandi, JW Choi, J Moon, M Kim. Middle East respiratory syndrome coronavirus (MERS-CoV) outbreak in South Korea, 2015: epidemiology, characteristics and public health implications, J.Hosp.Infect. 95 (2017) 207-213.
- [11] Q Lin, S Zhao, D Gao, Y Lou, S Yang, SS Musa, et al. A conceptual model for the outbreak of Coronavirus disease 2019 (COVID-19) in Wuhan, China with individual reaction and governmental action, International journal of infectious diseases. (2020).
- [12] J Mazurek, Z Nenickova. Predicting the number of total COVID-19 cases and deaths in the USA by the Gompertz curve. (2020).
- [13] VK Menon. Prediction of number of cases expected and estimation of the final size of coronavirus epidemic in India using the Logistic model and genetic algorithm, arXiv preprint arXiv:2003.12017. (2020).

- [14] DC Montgomery, EA Peck, GG Vining, Introduction to linear regression analysis, John Wiley & Sons 2012.
- [15] SZd Pinho, LRd Carvalho, MM Mischán, Passos, José Raimundo de Souza. Critical points on growth curves in autoregressive and mixed models, *Scientia Agricola*. 71 (2014) 30-37.
- [16] T Porter, E Kebreab, HD Kuhl, S Lopez, AB Strathe, J France. Flexible alternatives to the Gompertz equation for describing growth with age in turkey hens, *Poult.Sci.* 89 (2010) 371-378.
- [17] MHDM Ribeiro, RG da Silva, VC Mariani, LdS Coelho. Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil, *Chaos, Solitons & Fractals*. 135 (2020) 109853.
- [18] K Roosa, Y Lee, R Luo, A Kirpich, R Rothenberg, J Hyman, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, *Infectious Disease Modeling*. 5 (2020) 256-263.
- [19] K Roosa, Y Lee, R Luo, A Kirpich, R Rothenberg, JM Hyman, et al. Short-term forecasts of the COVID-19 epidemic in Guangdong and Zhejiang, China: February 13–23, 2020, *Journal of clinical medicine*. 9 (2020) 596.
- [20] CY Shen. A Logistic growth model for COVID-19 proliferation: experiences from China and international implications in infectious diseases, *International Journal of Infectious Diseases*. (2020).
- [21] MM Tariq, F Iqbal, E Eydurán, MA Bajwa, ZE Huma, A Waheed. Comparison of non-linear functions to describe the growth in Mengali sheep breed of Balochistan, *Pak.J.Zool.* 45 (2013) 661-665.
- [22] DAM Villela, Discrete time forecasting of epidemics, *Infectious Disease Modeling*. 5 (2020) 189-196.
- [23] QH Vuong. Likelihood ratio tests for model selection and non-nested hypotheses, *Econometrica: Journal of the Econometric Society*. (1989) 307-333.
- [24] World Health Organization, Coronavirus disease (COVID-19) pandemic. .,
- [25] X Zhang, R Ma, L Wang. Predicting turning point, duration and attack rate of COVID-19 outbreaks in major Western countries, *Chaos, Solitons & Fractals*. 135 (2020) 109829.
- [26] J Zhou, G Tse, S Lee, T Liu, WK Wu, D Zeng, et al. Identifying main and interaction effects of risk factors to predict intensive care admission in patients hospitalized with COVID-19: a retrospective cohort study in Hong Kong, *medRxiv*. (2020).

