Forecasting COVID-19 cases in Algeria using logistic growth and polynomial regression models

Mohamed LOUNIS (✉ lounisvet@gmail.com)

Universite Ziane Achour de Djelfa Faculte des Sciences de la Nature et de la Vie <https://orcid.org/0000-0003-0421-2919>

Babu Malavika

Christian Medical College Department of Biostatistics

Short Report

Keywords: COVID-19, logistic growth model, polynomial regression model, Forecasting

Posted Date: February 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-223608/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The novel Coronavirus respiratory disease 2019 (COVID-19) is still expanding through the world since it started in Wuhan (China) on December 2019 reporting a number of more than 84.4 millions cases and 1.8 millions deaths on January 3rd 2021.

In this work and to forecast the COVID-19 cases in Algeria, we used two models: the logistic growth model and the polynomial regression model using data of COVID-19 cases reported by the Algerian ministry of health from February 25th to December 2nd, 2020.

Results showed that the polynomial regression model fitted better the data of COVID-19 in Algeria the Logistic model. The first model estimated the number of cases on January, 19th 2021 at 387673 cases. This model could help the Algerian authorities in the fighting against this disease.

Introduction

The COVID-19 is the third epidemic reported in the 21st century caused by a virus from the family of *Coronaviridae* after the SARS (Severe Acute Respiratory Syndrome) in 2003 and MERS (Middle-East Respiratory Syndrome) in 2012 [1]. However, the propagation of this disease is faster than the two others. Since its first report in the city of Wuhan in China on December, 2019, the disease has been widely and rapidly reported in all countries and regions of the planet gaining the statute of a pandemic since March, 2020. The number of cases continues to increase rising a total of more than 84.4 millions and near than 2 millions deaths on January 3rd, 2021 [2].

In this way, it is important to understand the epidemic evolution and evaluate the protective measures applied by the national authorities. To attempt these objectives, mathematical and statistical models represents an inescapable part not only for prediction but also for planning control and mitigation actions. These models which use various sources of data allow us opportunities to test various strategies in simulations before their application in populations or individuals [3].

Since the first reports of COVID-19, various models were used to predict the curve evolution and to estimate the final size of the epidemic. Among these approaches, logistic growth models have been previously used to describe epidemics spread [4, 5]. These simpler tools are very popular and have been applied for COVID-19 short-term forecasting [6, 7, 8, 9] and also to predict the final size of the epidemic [10, 11, 12]. Roosa et al. [13] used a generalized logistic growth model to assess the impact of containment and predict final case numbers in China. Vattay [14] used the logistic growth model to analyze the similarity in Hubei, China and Italy in term of COVID-19 death numbers, and predicted the end date in Italy. Wu et al. [15] used four logistic growth models (the classical logistic growth model, the generalized logistic model, the generalized Richards model and the generalized growth model) to analyze the growth of COVID-19 in Chinese provinces using and further applied them to predict the number sizes in other European, American and Asian countries.

Roosa et al [16] used a generalized logistic growth model, the Richards growth model, and a sub-epidemic wave model for a 5- and 10-day prediction of cumulative cases in both Guangdong and Zhejiang (China). They observed that the GLM and Richards model showed comparable predictions, while the sub-epidemic model forecasts showed significantly greater uncertainty.

Balaban, 2020 [17] compared the performance of 5 growth models like logistic growth, Von Bertalanffy growth, exponential growth, Gaussian growth and Richards growth models and used them for short term forecasting of COVID-19 in Turkey. The author showed that Von Bertalanffy model has the best performance but the exponential model has predicted the total deaths and cases better than the others. In the study of Zhou et al. [18], the authors used the logistic growth and the SEIR models to forecast the spread of COVID-19. They reported that the pandemic size estimated by the logistic model was considerably smaller than the SEIR models.

In another study, Batista [12] tried to estimate the final size of the epidemic for the whole World using logistic and SIR models. The same authors used the logistic model to forecast daily predictions and the epidemic size for China, South Korea, and the rest of the World [12].

Malhotra and Kashyap [19] calibrated the Susceptible-Infected-Recovered (SIR) model and the Logistic Growth model to forecast the endpoint of COVID-19 in India and three states.

Another tool regarded as one of the best tools to analyze and predict the pandemic growth is polynomial regression. This special type of multiple regression method was applied in several studies to analyze the behavior of COVID-19 [20, 21, 22, 23, 24]. It has shown 99.85% accuracy in the study of Yadav [25].

Prakash [26] used ANN and regression to model the COVID-19 pandemic in India and other countries like The USA, Italy and Spain. The authors observed that the results of polynomial regression follow the ground truths in India and the USA but not in Italy and Spain. They use also these models to estimate the peak of the epidemic in the cited countries.

Belfin et al, 2002 [27] have used a SEIR and polynomial regression models to estimate the pick of the COVID-19 epidemic and the basic reproduction number in India respectively.

Also, Amar et al, 2020 [28] applied seven regression models for the COVID-19 dataset including exponential polynomial, quadratic, third degree, fourth-degree, fifth-degree, sixth-degree, and logit growth. They reported that, the exponential, fourth-degree, fifth-degree, and sixth-degree polynomial regression models are excellent models (especially fourth-degree model).

In the same way Chakraborty et al, 2020 [29] compared the performance between Linear Regression model, Granular Box Regression (GBR) and the Polynomial Regression model in predicting the spread of COVID-19 in India. The authors reported that the Polynomial regression model surpassed the two other models.

Algeria reported its first case on February 25th, 2020. Since then, the actual situation shows a total number of 99897 cases and 2762 deaths [30]. Despite the different preventive measures applied since March, 2020 the number of total cases is still increasing. Thus short and long term estimation of the number of cases and prediction of the curve evolution is of great importance to understand the epidemic curve. The current work is conducted to predict the number of case using a logistic growth model.

Methods

Data sources:

In the present study, the used datasets of COVID-19 confirmed cases were taken from the daily reports of the Algerian Ministry of Health from February 25th (when the first case was reported) to December 2nd, 2020[30]. The number of confirmed cases was based on RT-PCR positive tests. The figure of the COVID-19 in Algeria on December 2nd, 2020 showed a number of 85084 cases. In parallel, the number of recovered persons was 54979 and the number of deaths was 2464 persons.

To predict the cumulative number of cases of COVID 19 in Algeria, we used both Modified Logistic Growth model and Polynomial regression.

Modified Logistic Growth Model:

Logistic Growth is characterized by increasing growth in the beginning period, but achieve stability at a later stage, as you get closer to a maximum. For example in the Coronavirus case, this maximum limit would be the total number of people in the world, because when everybody is sick, the growth will necessarily diminish. The reason to use Logistic Growth for modeling the Coronavirus outbreak is that epidemiologists have studied those types of outbreaks and it is well known that the first period of an epidemic follows Exponential Growth and that the total period can be modelled with a Logistic Growth.

The modified logistic growth model is presented as follows [31, 32]:

$$y(t) = \frac{C}{1 + a * e^{-bt}}$$

Where $y(t)$ is the number of cases at any given time t

C is the limiting value, the maximum capacity for y

$a = (C / y_0) - 1$

b rate of change.

- the number of cases at the beginning, also called *initial value* is: $C / (1 + a)$
- the maximum growth rate is at $t = \ln(a) / b$

- When y is equal to c (that is, the population is at maximum size), y / C will be 1. Therefore, the $(1 - (y/c))$ will be 0 and hence the growth will be 0.

Polynomial Regression

Polynomials are widely used in situations where the response is curvilinear, as even complex nonlinear relationships can be adequately modeled by polynomials over reasonably small ranges of the x 's. In general, the k th - order polynomial model in one variable is

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta^k x^k + \varepsilon$$

In this data we used Polynomial regression with 6 degree.

Results

Logistic Growth Model:

Figure 1 and 2 show that Logistic Growth Model is not working well for COVID-19 cases in Algeria. Model shows flattening but in real scenario it's still increasing. To get better fit using Logistic Growth Model, We have to wait for some more days to know the trend of the data in upcoming days.

Polynomial Regression:

As shown in Figure 3, the data fitted well in polynomial regression model. Predictions of table 1 showed that the number of cases will be estimated at 387673 (372809- 402538) cases on January 19th 2021.

Table 1: Projected cumulative number of COVID-19 cases for Algeria

Days	Date	Observed cases	Predicted cases	Lower	Upper
282	02-Dec-20	85084	89340	88543	90136
283	03-Dec-20	NA	91433	90566	92300
284	04-Dec-20	NA	93629	92685	94572
285	05-Dec-20	NA	95930	94903	96956
286	06-Dec-20	NA	98341	97226	99455
287	07-Dec-20	NA	100865	99655	102074
288	08-Dec-20	NA	103506	102196	104817
289	09-Dec-20	NA	106270	104851	107688
290	10-Dec-20	NA	109158	107625	110691
291	11-Dec-20	NA	112177	110523	113831
292	12-Dec-20	NA	115330	113547	117113
293	13-Dec-20	NA	118621	116702	120540
294	14-Dec-20	NA	122056	119994	124118
295	15-Dec-20	NA	125638	123425	127851
296	16-Dec-20	NA	129373	127000	131745
297	17-Dec-20	NA	133264	130724	135804
298	18-Dec-20	NA	137318	134602	140034
299	19-Dec-20	NA	141540	138639	144441
300	20-Dec-20	NA	145933	142838	149028
301	21-Dec-20	NA	150504	147205	153802
302	22-Dec-20	NA	155257	151746	158769
303	23-Dec-20	NA	160199	156464	163934
304	24-Dec-20	NA	165335	161366	169303
305	25-Dec-20	NA	170670	166457	174882
306	26-Dec-20	NA	176209	171742	180677
307	27-Dec-20	NA	181960	177227	186694
308	28-Dec-20	NA	187928	182916	192940
309	29-Dec-20	NA	194119	188817	199421
310	30-Dec-20	NA	200539	194935	206143
311	31-Dec-20	NA	207194	201275	213113
312	01-Jan-21	NA	214092	207844	220339
313	02-Jan-21	NA	221237	214649	227826
314	03-Jan-21	NA	228638	221694	235582
315	04-Jan-21	NA	236301	228987	243615
316	05-Jan-21	NA	244233	236534	251931
317	06-Jan-21	NA	252440	244342	260538
318	07-Jan-21	NA	260930	252418	269443
319	08-Jan-21	NA	269711	260768	278654
320	09-Jan-21	NA	278790	269399	288180
321	10-Jan-21	NA	288173	278319	298028
322	11-Jan-21	NA	297870	287534	308206
323	12-Jan-21	NA	307887	297053	318722
324	13-Jan-21	NA	318233	306882	329585
325	14-Jan-21	NA	328916	317029	340804
326	15-Jan-21	NA	339944	327501	352386
327	16-Jan-21	NA	351325	338308	364342
328	17-Jan-21	NA	363068	349455	376680

329	18-Jan-21	NA	375181	360953	389409
330	19-Jan-21	NA	387673	372809	402538

Discussion

In this work, we used two types of models to predict COVID-19 cases in Algeria: Logistic Growth model and polynomial growth model.

Results showed that logistic growth model despite its wide use in forecasting COVID-19 curve has shown discordance between the real and the fitted data from March 25th to December 2nd, 2020. In the same way, Abusam et al, 2020 showed that both Velhust and Richards models are not well adapted and needs more parameterization for COVID-19 predictions in Kuwait[33].

Regarding polynomial regression model, Results showed that this model fitted well COVID-19 data in Algeria. These results are in accordance of those of Amar et al, 2020 who showed that the exponential, fourth-degree, fifth-degree, and sixth-degree polynomial regression models are excellent models and especially fourth-degree model.

However, the limitation of our model is that we cannot predict for more than 60 days. Usually polynomial regression is used for the short term prediction. This type of model was widely used in the case of COVID-19[20, 21, 22, 23, 24] and has shown an excellent accuracy in certain cases [25].

Conclusion

A forecast of COVID-19 cases in Algeria was conducted in this study using logistic growth and polynomial regression models based on data from February, 25th, and December 2nd, 2020. Results showed that polynomial regression model is more adapted in COVID-19 forecasting in Algeria. We used then this model to predict the short term future cases until January 19th, 2020. The total number of cases in this date is estimated at 387,673 cases. This model could help the Algerian Government in adapting the best strategies against the COVID-19 epidemic.

Declarations

Acknowledgements

Not applicable.

Conflict of interest statement:

The authors declare that they have no conflict of interest

References

1. Nikhat S and Fazil M. Overview of COVID-19: Its prevention and management in the Light of Unani Medicine. *Science of the Total Environment*. 2020. 728, 138859. doi: <https://doi.org/10.1016/j.scitotenv.2020138859>
2. Johns Hopkins University of Medicine (JHUM), Coronavirus resource center: <https://coronavirus.jhu.edu/map.html> (Accessed on January 3, 2021)
3. Kotwal A, Yadav AK, Yadav J, Kotwal J, Khune S. Predictive models of COVID-19 in India: A rapid review. *Medical journal armed forces India*. <https://doi.org/10.1016/j.mjafi.2020.06.001>
4. Pell B, Kuang Y, Viboud C, Chowell G. Using phenomenological models for forecasting the 2015 Ebola challenge. *Epidemics* 2018; 22: 62-70. doi:10.1016/j.epidem.2016.11.002.
5. Chowell G, Hincapie-Palacio D, Ospina J, Pell B, Tariq A, Dahal S, Moghadas S, Smirnova A, Simonsen L, Viboud C. Using Phenomenological Models to Characterize Transmissibility and Forecast Patterns and Final Burden of Zika Epidemics. *PLOS Currents Outbreaks*. 2016 May 31 .Edition 1. doi: 10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583.
6. Jia L, Kewen L, Jiang Y, Guo X, Zhao T. Prediction and analysis of Coronavirus Disease 2019. *arXiv: Populations and Evolution*. 2020. <https://arxiv.org/abs/2003.05447>.
7. Chen DG, Chen X, Chen JK. Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. *Glob Health Res Policy*.2020; 5:25.doi: 1186/s41256-020-00152-5.
8. Malavika B., Marimuthu S., Joy M, Nadaraj A, Asirvathamb E S, Jeyaseelanc L. Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models. *Clinical Epidemiology and Global Health*.<https://doi.org/10.1016/j.cegh.2020.06.006>
9. Shen CY. Logistic growth modelling of COVID-19 proliferation in China and its international implications.*Int J Infect Dis*. 2020 Jul; 96: 582– 589.doi: 1016/j.ijid.2020.04.085.
10. Alboaneen D, Pranggono B, Alshammari D, Alqahtani N, Alyaffer R. Predicting the Epidemiological Outbreak of the Coronavirus Disease 2019 (COVID-19) in Saudi Arabia. *Int J Environ Res Public Health*. 2020 Jun 25;17(12):4568. doi: 3390/ijerph17124568.
11. Batista M, Estimation of the final size of the COVID-19 epidemic, medRxiv, (2020) 2020.2002.2016.20023606.
12. Batista M. Estimation of the final size of the second phase of coronavirus epidemic by the logistic model. medRxiv (2020) doi: <https://doi.org/10.1101/2020.03.11.20024901>
13. Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, Yan P, and Chowell G. Real-time forecasts of the COVID-19 epidemic in china from february 5th to february 24th, 2020. *Infectious Disease Modelling*, 5:256–263, 2020.
14. Vattay G. Predicting the ultimate outcome of the COVID-19 outbreak in Italy. *arXiv preprint arXiv:2003.07912*, 2020.

15. Wu K, Darcet D, Wang Q, Sornette D. Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in China and in the rest of the world. *Nonlinear Dyn* 2020; 1-21. doi :10.1007/s11071-020-05862-6.
16. Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, Yan P and Chowell G. Short-term Forecasts of the COVID-19 Epidemic in Guangdong and Zhejiang, China: February 13–23, 2020. *J. Clin. Med.* **2020**, 9, 596; doi:10.3390/jcm9020596
17. Balaban M, Growth Models for Covid-19 Death Figures of Turkey. *Journal of Advances in Medicine and Medical Research* 32(20): 1-11, 2020.
18. Zhou X, Ma X, Hong N, Su L, Ma Y, He J, Jiang H, Liu C, Shan G, Zhu W, Zhang S, Long. Forecasting the worldwide spread of COVID-19 based on logistic model and SEIR model. *medRxiv* (2020). <https://doi.org/10.1101/2020.03.26.20044289>
19. Malhotra B, Kashyap V. Progression of covid-19 in indian states - forecasting endpoints using sir and logistic growth models. *medRxiv* (2020). doi: <https://doi.org/10.1101/2020.05.15.20103028>
20. Da Rocha AF. Regression Polynomial Analysis of the COVID-19 epidemics: some initial findings. Preprint · April 2020.doi: 10.13140/RG.2.2.28338.56001
21. Da Rocha AF. Regression Polynomial Analysis of the COVID-19 epidemics: Regional dynamics in Spain and Germany. Preprint June 2020.doi: 10.13140/RG.2.2.33177.65121
22. Da Rocha AF. Regression Polynomial Analysis of the COVID-19 epidemics: An alternative infection modeling. Available at SSRN 3682535 (2020).
23. Sharma VK and Nigam U. Modelling of COVID-19 cases in India using Regression and Time Series models.*medRxiv* (2020). doi: <https://doi.org/10.1101/2020.05.20.20107540>
24. Ekum, M., and A. Ogunsanya. Application of hierarchical polynomial regression models to predict transmission of COVID-19 at global level. *Int J ClinBiostatBiom*,6.1 (2020): 027. doi: 10.23937/2469-5831/1510027
25. Yadav, Milind, MurukessanPerumal, and M. Srinivas. "Analysis on novel coronavirus (COVID-19) using machine learning methods. *Chaos, Solitons& Fractals*139 (2020): 110050.
26. Prakash A, Sharma P, Sinha I K and Singh UP. Spread & Peak Prediction of Covid-19 using ANN and Regression (Workshop Paper).2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM). 356-365. DOI 10.1109/BigMM50055.2020.00062
27. Belfin RV, Piotr B, Radhakrishnan BL, Rejula V. COVID-19 peak estimation and effect of nationwide lockdown in India. *medRxiv* (2020). doi: <https://doi.org/10.1101/2020.05.09.20095919>
28. Amar L A, Taha AA, Mohamed MY. Prediction of the final size for COVID-19 epidemic using machine learning: A case study of Egypt. *InfectiousDisease Modelling*, 2020, 5 : 622e634. <https://doi.org/10.1016/j.idm.2020.08.008>
29. Chakraborty M, Mukhopadhyay A, Maulik U. A Comparative Analysis of Different Regression Models on Predicting the Spread of Covid-19 in India. 2020 IEEE, 5th International Conference on Computing Communication and Automation (ICCCA) Galgotias University, Greater Noida, UP, India. Oct 30-31, 2020

30. Algerian health and hospital reform minister: Carte épidémiologique. Available at: <https://www.covid19.gov.dz/carte/> (Accessed on January 3rd, 2021).
31. Ji LQ. Analysis of a modified logistic model for describing the growth of durable customer goods in China. Math Comput Appl. 2013, 4; 18:30–37.
32. Korstanje J. Modeling logistic growth. Available from: <https://towardsdatascience.com/modeling-logistic-growth-1367dc971de2>; 2020.
33. AbusamA, Abusam R, Al-AnziB. Adequacy of Logistic models for describing the dynamics of COVID-19 pandemic. Infectious Disease Modelling 5 (2020) 536e542.doi : <https://doi.org/10.1016/j.idm.2020.08.006>

Figures

Figure 1

Modeling of COVID-19 cases in Algeria using Logistic growth model

Figure 2

Prediction of COVID-19 cases in Algeria using Logistic Growth Model

Covid-19 Prediction for Algeria

Figure 3