Iran Red Crescent Med J. 2020 May; 22(5):e102822. doi: 10.5812/ircmj.102822.

Published online 2020 May 6. **Methods Article**

The Prediction of COVID-19 Spread in Iran From 15 March to 15 April 2020

Babak Jamshidi[1], Mansour Rezaei[2,*], Farid Najafi[1] and Azad Sheikhi[3]

[1]Kermanshah University of Medical Sciences, Kermanshah, Iran
[2]Social Development and Health Promotion Research Center, Kermanshah University of Medical Sciences, Kermanshah, Iran
[3]Tarbiat Modares University, Tehran, Iran

[*]*Corresponding author*: Department of Biostatistics, Social Development and Health Promotion Research Center, Kermanshah University of Medical Sciences, Kermanshah, Iran. Tel: +98-9183318978, Email: rezaei39@yahoo.com

Received 2020 March 17; **Revised** 2020 March 31; **Accepted** 2020 April 21.

Abstract

Background: Over 150,000 confirmed cases, around 140 countries, and about 6,000 death occurred owing to coronavirus disease 2019 (COVID-19) pandemic in China, Italy, Iran, and South Korea. Iran reported its first confirmed cases of COVID-19 in Qom City on 19 February 2020 and has the third-highest number of COVID-19 deaths after China and Italy and the highest in Western Asia.
Methods: We applied a two-part model of time series to predict the spread of COVID-19 in Iran through addressing the daily relative increments. All of the calculations, simulations, and results in our paper were carried out by using MatLab R2015b software. The average, upper bound, and lower bound were calculated through 100 simulations of the fitted models.
Results: According to the prediction, it is expected that by 15 April 2020, the relative increment (RI) falls to the interval 1.5% to 3.6% (average equal to 2.5%). During the last three days, the RI belonged to the interval of 12% to 15%. It is expected that the reported cumulative number of confirmed cases reaches 71,000 by 15 April, 2020. Moreover, 80% confidence interval was calculated as 35K and 133K.
Conclusions: The screening of suspected people, using their electronic health files, helps isolate the patients in their earlier stage, which in turn helps decrease the period of transmissibility of the patients. Considering all issues, the best way is to apply the model with no modification to model the probable increasing or decreasing acceleration of spreading.

Keywords: COVID-19, Iran, Prediction, Time Series, Model, 2020, Spreading, Cumulative Number, Daily Relative Increment

1. Background

The coronavirus disease 2019 (COVID-19) is an ongoing pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It was first identified in Wuhan, Hubei, China in late December 2019. Until 14 March, over 150,000 cases were confirmed in about 140 countries, with epicenters China, Italy, Iran, South Korea, and Europe, and about 6,000 people died owing to the disease (1, 2).

Nowadays, new infectious diseases are spreading around the world faster than before. This unprecedented rate is the result of some factors such as the increasing ease of international travel, population growth, resistance to drugs, and destruction of the environment. This high rate of emergence emphasizes the importance of research on the spreading of contagious diseases (3).

During the 2019-20 coronavirus pandemic, Iran reported its first confirmed cases of COVID-19 in Qom City on 19 February, 2020. Up to 14 March, 2020, according to Iranian health authorities, there were 611 deaths due to COVID-19 in Iran with more than 12,700 confirmed reported cases (1, 2). Until 14 March, Iran had the third-highest number of COVID-19 deaths after China and Italy and the highest in Western Asia. As an epicenter, Iran is an appropriate sample of a developing country, faced with COVID-19. The rate of growth in Iran is more than all other countries because the number of confirmed cases in this country has reached over 10000, just within 3 weeks. The daily rise in the number of casualties in Iran suggests the fight against the new coronavirus is still far from over.

The model we used for describing the time series of the daily relative increment is introduced by Jamshidi et al. (3). It is a two-part model which the first and second parts are stationary and non-stationary, respectively. In the second part, the variance is proportional to the mean and hence, can become stationary by Box-Cox transformations (4). It is noticeable that since the propagation of communicable diseases has two different behaviors in the beginning and the following days (5), we are forced to apply a two-part model (Figure 1).

2. Methods

2.1. Model

Let Y_t is a time series of the cumulative number of confirmed cases at time t. We want to study $X_t = \frac{Y_{t+1}}{Y_t - 1}$, showing the daily relative increment. The model we applied had five positive parameters (b, IR, K, θ, and a);

For t = 1, …, b-1; $X_t \sim$ normal (IR, IR^2/a),

and for = b, b+1, …; $X_t \sim$ normal (K/t^{θ}, $K^2/at^{2\theta}$).

Where; b, The length of the first days of spreading; IR, the average relative increment in the first days of spreading; θ, the acceleration of falling of the relative increments after the first days of spreading ($X_t \propto t^{-\theta}$ for t = b, b+1, b+2, …); a, the fixed ratio of the mean to the variance ($a = \frac{expectation}{(Standard\ deviation)^2}$); and finally, K, the adjusting coefficient for the curve $t^{-\theta}$ to fit the decreasing trend of the time series of the relative increment.

2.2. Estimation

This model is flexible enough to provide analysis for the propagation of a wide range of epidemics such as:

The SARS epidemic of 2003 (3), the MERS epidemic in South Korea 20 May - 7 July 2018 (3), the Ebola outbreak of 2014 - 2016 (3), the propagation of HIV/AIDS from 1990 to 2018 (3), the spreading of the Cholera of 2008 - 2009 in Zimbabwe (3), and the COVID-19 epidemic in China and four of its provinces, Beijing, Guangdong, Shanghai, and Hubei in 2020 (3).

Owing to the flexibility and transparency of the model, and the similarity between the pattern of the trend of the cumulative number of confirmed cases in Iran and the mentioned epidemics, we selected this model to conduct a study on the subjects (Figure 2).

To estimate the parameters of the model, we can:

Take b as the first point that the geometric mean of the relative increments in the previous points exceeds 3/2 times the geometric mean of the next three points or the time point that the relative increments fall irreversibly.

Calculate the geometric mean of the cumulative ratios in the previous points ($1 + X_t$) from t = 1 to t = b-1 as the estimation of the parameter IR.

Estimate the parameters θ and K according to the following linear relation:

$$\begin{aligned} X_t &\cong \frac{K}{t^{\theta}} \\ &\Rightarrow \frac{1}{X_t} \\ &\cong \frac{t^{\theta}}{K} \\ &\Rightarrow ln\left(\frac{1}{X_t}\right) \\ &\cong \theta\ ln\,(t) - \ ln\,(K) \\ &\Rightarrow \ ln\,(X_t) \\ &\cong -\theta ln\,(t) + \ ln\,(K) \\ &\Rightarrow ln\,(X_t) \\ &\cong \theta\ ln\left(\frac{1}{t}\right) + ln\ (K) \end{aligned} \tag{1}$$

Multiply all the observations after t = b-1 by t^{θ}/K to have an identical mean and variance for all of the newly obtained data (W_t).

$$X_t \sim Normal\left(\frac{K}{t^{\theta}}, \frac{K^2}{at^{2\theta}} \Rightarrow W_t = \frac{t^{\theta}}{K}X^t \sim Normal\left(1, \frac{1}{a}\right)\right) \tag{2}$$

Therefore, the variance of the newly obtained data is a good candidate for 1/a. Accordingly, $= 1/S_W^2$ (3).

The mentioned procedure leads to the following estimation (Figure 3):

$$\begin{aligned} &b\widehat{b},\ IR\widehat{IR}, K\widehat{K}, \theta\widehat{\theta}, a\widehat{a} \\ &= (13,\ 7217,\ 1.9042, 45, 2078,\ 0.0929) \end{aligned}$$

2.3. Simulation and Prediction

To predict the studied time series based on the model, using the above-mentioned parameters, we conducted the following simulations;

We ran the model with parameters (13, 0.7217, 1.9042, 45.2078, 0.0929) from Feb 19th to Apr 15th (Figure 4A).

This simulation represents the accuracy of the model to fit the available dataset using the average curve and 80% upper and lower bounds.

We ran the model with parameters (13, 0.7217, 1.9042, 45.2078, 0.0929) from 15 March to 15 April (Figure 4B).

This simulation is the basis of our prediction. By using the recursive formula 1 and based on the simulation, we forecasted the cumulative frequency of confirmed cases until 15 April, 2020.

It is worth mentioning that all the calculations, simulations, and results in this paper were performed using MatLab R2015b software. Throughout the paper, the average, upper bound, and lower bound were calculated through 100 simulations of the fitted models.

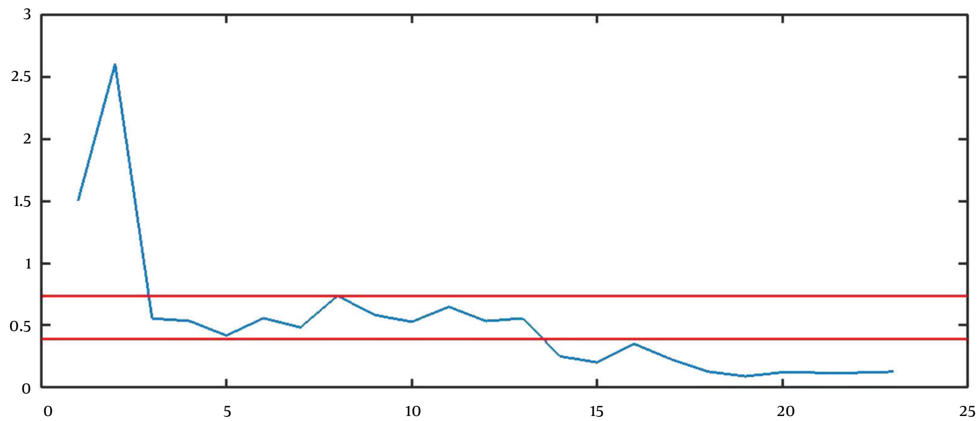

Figure 1. The time series of relative increment of COVID-19 in Iran, the lower red line determines the breakpoint, where the irreversible falling happened, and the upper red line represents the geometric mean of the relative increments before the breakpoint.

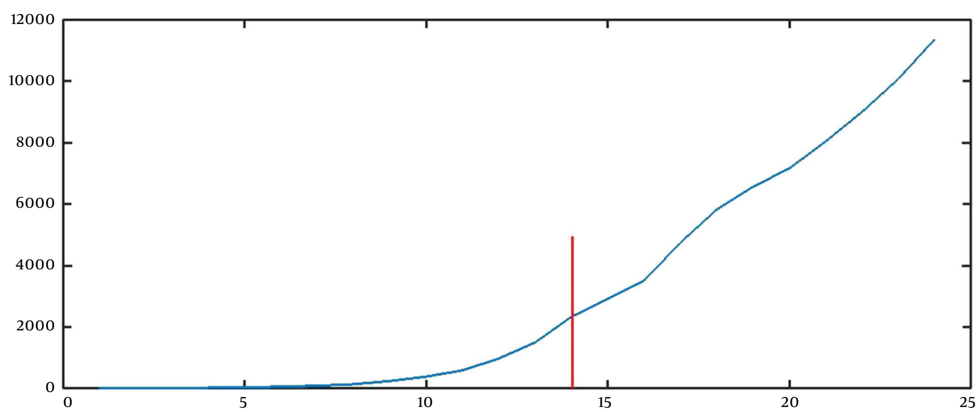

Figure 2. The time series of the cumulative number of individuals infected by COVID-19 in Iran. The red line determines the breakpoint, where the irreversible falling in the rate of growth happened.

Based on the obtained model, we attempted to simulate the model either retrospectively or prospectively, and we got the graphs of Figure 4A -C as follows: (1) realization, simulation, and prediction (Figure 4A); (2) realization and prediction (Figure 4); and (3) pure prediction (Figure 4C) of the model, respectively. According to the prediction, it is expected that with probability 80%, the daily relative increment (RI) falls to the interval of 1.5% to 3.6% (with average equal to 2.5%). It is noteworthy that during the last three days, the mentioned index related to the interval of 12% to 15%.

In order to predict the cumulative number of the confirmed cases during the study (Y_t), we applied the following recursive equation:

$$X_t = Y_{t+1}/Y_t - 1 \Gamma\Rightarrow Y_{t+1} = Y_t (X_t + 1) \; and \, Y_1 = 2 \qquad (3)$$

Therefore, Figure 5 was obtained as the predicted path of the propagation of COVID-19 in Iran within a month. Correspondingly, it is expected that the number of cumulative confirmed cases reaches 71000 by 15 April, 2020. The 80% confidence interval was calculated as 35K and 133K.

3. Discussion

We applied a two-part model of time series to predict the spread of COVID-19 in Iran through addressing the daily RI. According to our model and related calculations and simulations, we expect that by 15 April, 2020, the daily relative increment will decrease to around 2.5% (80% confidence interval: 1.5% to 3.6%). Accordingly, it is predicted that the cumulative number of confirmed cases increases

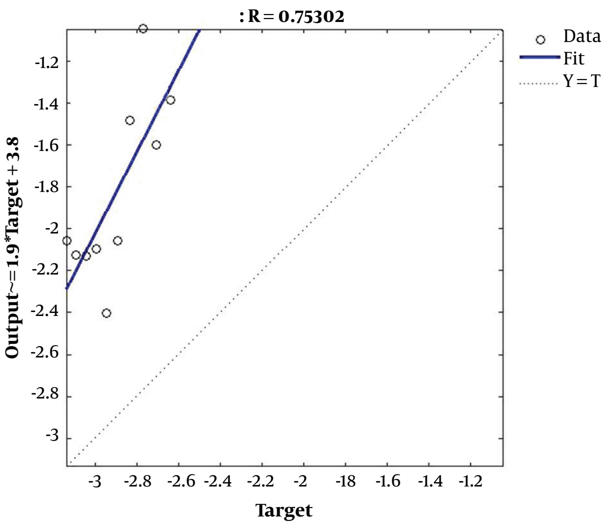

Figure 3. The graph of the regression to estimate the parameters of the model 1.9, and 3.8 are the estimates of θ and ln (K), respectively.

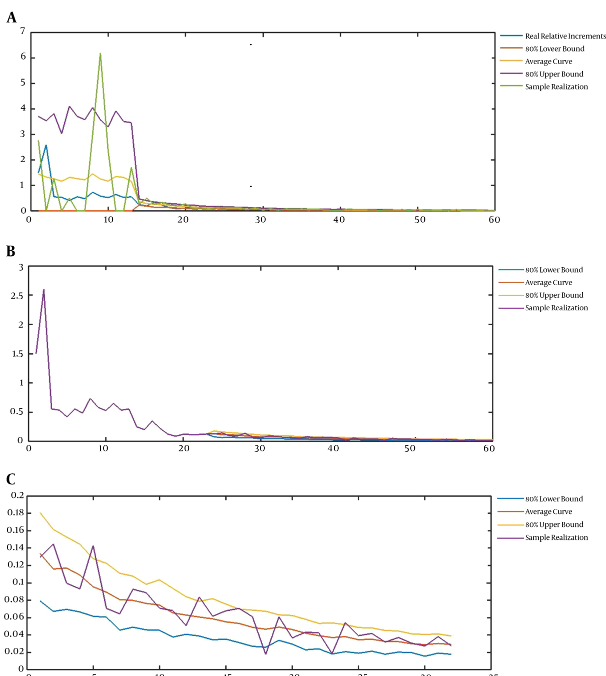

Figure 4. A, The real data from 19 February to 14 March, 2020, and average curve, 80% upper and lower bounds, and a sample realization of the fitted model from 19 February to 15 April, 2020; B, the real data from 19 February to 14 March, 2020, and average curve, 80% upper and lower bounds, and a sample realization of the fitted model from 15 March to 15 April, 2020; C, average curve, 80% upper and lower bounds, and a sample realization of the fitted model for RI from 15 March to 15 April, 2020.

to 71,000 (80% confidence interval: 35,000 - 133,000) by 15 April 2020.

It seems that the rate of growth in Iran is more than all other countries. The number of confirmed cases in Iran has reached over 10,000 within the first three weeks. There are some probable explanations for such a high rate of transmission in Iran. COVID-19 is a viral disease with no vaccine, and therefore, herd immunity is the only factor for stopping the transmission of the disease. For countries such as China, where the government applies very intensive policies for quarantine of affected people and cities, such huge increment stops in a shorter period. Although some European countries such as Italy follow the same strategy, Iran had a very flexible policy in both international and between cities traffic. Iran did not ban the air traffic between Iran and China during the outbreak of COVID-19 in China. Nevertheless, the famous health network for providing primary health care in all urban and rural areas provide opportunities for health education to all people. In addition, it seems that social distancing in Iran has not been successful yet. While the disease became an epidemic in China, Iran was involved in the parliamentary election with a large number of election meetings.

In fact, because of the coming Iranian New Year holidays, an increase could be expected due to traveling between big industrialized cities and small towns. Unfortunately, there is no obligation for banning such trips in Iran, and it does not seem that the government has any intension for such a policy. Such social gathering has an increasing effect on both of the number and probability of transmission of the disease. However, the screening of suspected people using their electronic health files, started from the last week, helps isolate the patients in their earlier stage, which, in turn, helps decrease the period of transmissibility of the patients. It is worth mentioning that our prediction for the RI in Iran is based on the number of confirmed cases by real-time PCR. However, we acknowledge that there are frequent patients diagnosed as COVID-19 by lung CT scan. We did not take into account such cases as they are not recorded as positive in the formal reports. On the other hand, the government has pledged to make more strict measures. Considering all such issues, the best way is to apply the model with no modification to model the probable increasing or decreasing acceleration of spreading.

Footnotes

Authors' Contribution: Idea, simulation, first draft: BJ. Final draft: MR. Discussion: FN. Literature review and revi-

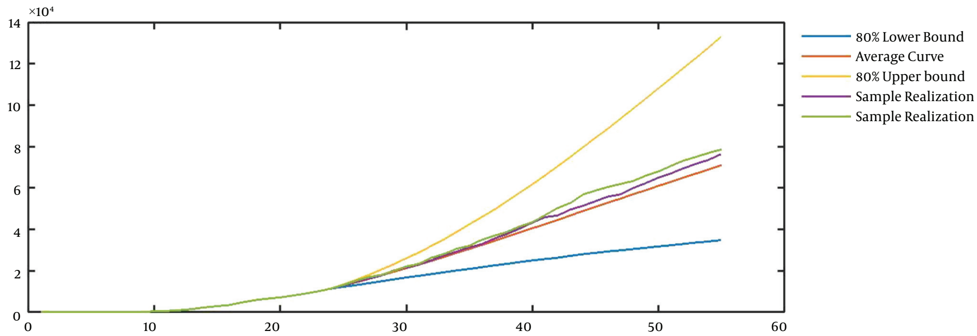

Figure 5. The predicted average curve for the number of total confirmed cases, 80% upper bound, 80% lower bound, and two sample realizations of the fitted model for 15.03.2020 - 15.04.2020.

sion: AS.

Conflict of Interests: All authors declare that they have no conflict of interest.

Funding/Support: None.

References

1. WHO. *Coronavirus disease 2019 (Covi-19) situation reports: 1 to 51*. 2020. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports.
2. Worldometer website. *COVID-19 coronavirus pandemic*. 2020. Available from: https://www.worldometers.info/coronavirus/#countries.
3. Jamshidi B, Rezaei M, Bekrizadeh H, Zargaran SJ. *A new family of time series to model the decreasing relative increment of spreading of an outbreak. Focused on Covid-19 in China, Preprint*. 2020. Available from: https://www.researchgate.net/publication/340236985_Modeling_decrease_rate_of_increment_by_time_series.
4. Box GEP, Cox DR. An analysis of transformations. *J Royal Statistic Soc*. 1964;**26**(2):211–43. doi: 10.1111/j.2517-6161.1964.tb00553.x.
5. Britton T, House T, Lloyd AL, Mollison D, Riley S, Trapman P. Five challenges for stochastic epidemic models involving global transmission. *Epidemics*. 2015;**10**:54–7. doi: 10.1016/j.epidem.2014.05.002. [PubMed: 25843384]. [PubMed Central: PMC4996665].