



COVID-19 Outbreak: Application of Multi-gene Genetic Programming to Country-based Prediction Models

Majid Niazkar¹, Hamid Reza Niazkar^{2*}

¹ Research Assistant, PhD, Department of Civil and Environmental Engineering, Shiraz University, Shiraz, IRAN

² Medical Student, Student Research Committee, Gonabad University of Medical Sciences, Gonabad, IRAN

*Corresponding Author: hmn185@hotmail.com

Citation: Niazkar M, Niazkar HR. COVID-19 Outbreak: Application of Multi-gene Genetic Programming to Country-based Prediction Models. *Electron J Gen Med.* 2020;17(5):em247. <https://doi.org/10.29333/ejgm/8232>

ARTICLE INFO

Received: 5 Apr. 2020

Accepted: 1 May. 2020

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) is a novel coronavirus that has infected more than 2,900,000 individuals worldwide. The widespread of coronavirus 2019 (COVID-19) brings about the need for a prediction model to adopt appropriate evidence-based strategies. In this study, multi-gene genetic programming (MGGP), as one of the artificial intelligence models, has been proposed for the first time for predicting the COVID-19 outbreak. Although this is a challenging task due to significant fluctuations of daily confirmed cases, the results achieved by MGGP are promising. To be more specific, the predicted confirmed cases are acceptably close to the observed values for seven countries considered in this study. Thus, MGGP is suggested for developing estimation models of COVID-19. Furthermore, similarities and differences between the proposed prediction models are presented. Finally, it is discussed why a country-based prediction model is recommended.

Keywords: COVID-19, SARS-COV-2, 2019-nCoV infection, prediction models, artificial intelligence, multi-gene genetic programming

INTRODUCTION

SARS-COV-2, the new member of coronaviruses responsible for an acute respiratory disease named COVID-19, has crossed many international borders (1). Some of the COVID-19 patients are asymptomatic, while symptomatic patients present a wide range of clinical signs and symptoms, including fever, cough, dyspnea, myalgia, confusion, headache, sore throat, rhinorrhea, chest pain, diarrhea, nausea and vomiting (2,3). According to the latest situation report of the World Health Organization (WHO), SARS-COV-2 has infected more than 2,900,000 individuals worldwide. Additionally, more than 200,000 patients have died to this date (28 April 2020) (1). To overcome this ongoing pandemic, researchers have conducted various studies focusing on different scopes, one of which is developing a prediction model (4).

Prediction models provide a historical perspective for healthcare decision-makers to adopt an evidence-based strategy to decrease morbidity, mortality, and economic losses in different levels (4,5). Through estimations, they can better evaluate the infectious capacity of pathogens and the efficacy of public health preventive measures (6). Despite numerous advantages of prediction models, some researchers believe that due to the uncertainty of official data and neglecting the infected people, who do not have access to the medical services, data provided by estimation models may be biased in some cases (6). Since COVID-19 has neither an effective vaccine nor a known treatment up to now, its rapid spread has led to

the shortcoming of hospital beds, Intensive Care Unit facilities such as ventilators, self-protection equipment such as face masks, and infecting many expert healthcare providers (5, 7). In this context, estimation models can provide an approximate number of infected individuals for future planning and management of the patients. Consequently, some mathematical, dynamical, and statistical methods have been proposed for short- and/or long-term forecasting of the COVID-19 outbreak (8,9).

According to the literature, estimation models are only proposed for a very limited number of countries. Moreover, the applicability of each prediction model to another country is questionable, while finding an appropriate mathematical model for the outbreak prediction, particularly for those countries with a short period of outbreak experience, is still challenging. In this regard, the current study aimed to investigate whether a prediction model exploited for a country applies to another one by comparing the outbreak trends in different countries. Additionally, multi-gene genetic programming (MGGP) has been applied for prediction of the COVID-19 outbreak for seven infected countries for the first-time.

MATERIALS AND METHODS

Data of COVID-19 Outbreak

The data of confirmed cases due to COVID-19 from 20 January to 5 April 2020 were gathered from the World Health

Table 1. Controlling parameters used by MGGP

Parameter	Values
Function set	arithmetic operations
Number of generations	120
Maximum number of genes allowed in individual	5
Maximum tree depth	7
Tournament size	4
Elitism	0.05 of population
Crossover events	0.85
High-level crossover	0.2
Low-level crossover	0.8
Standard deviation of perturbation applied in mutation Gaussian perturbation of a randomly selected constant	0.1
Sub-tree mutation	0.9
Replacing input terminal with another random terminal	0.05
Gaussian perturbation of randomly selected constant	0.05
Direct reproduction	0.05
Ephemeral random constants	[-10, -10]

Organization (WHO) situation reports (1) and the National Health Commission of the People's Republic of China (NHC) official website (10). The latter is preferred for China when there was a discrepancy between these two sources. In this study, confirmed cases of China, Republic of Korea, Japan, Italy, Singapore, United States of America (USA), and Iran (Islamic Republic of) were considered.

Multi-Gene Genetic Programming

Genetic Programming (GP) is an artificial intelligence model that exploits the genetic algorithm as a search engine (11). GP is mainly suitable for when a problem with high-order complexity is under investigation (12). In a bid to overcome some of limitations of classical GP, several variants of GP such as MGGP has been introduced (13). In essence, MGGP consists of several genes, which corresponds to each GP tree (13). Basically, GP and MGGP comprise of a tree-based architecture that provides an implementation of various functions and variables in light of finding a suitable expression between input and output data (14).

GP adopts a four-step random search including initialization, selection, reproduction, and termination (11). In the beginning, it randomly generates an initial population consists of individuals (functions and terminals), which are the main candidates for the best relationship between the input and output data (12). The created population needs to be subjected by genetic operators continually in favor of achieving the best relationship between the input and output data (11, 12). In this regard, selecting appropriate functions and terminals can enable GP and MGGP to solve any complicated problem (15).

In this study, an open-source code of MGGP from the literature is used (13,16), while the controlling parameters considered in MGGP are presented in **Table 1**. The maximum number of genes allowed in individual and the maximum tree depth shown in **Table 1** are two crucial controlling parameters set by the user. The former is a multi-gene parameter, while the latter is a tree build parameter (13). There is a trade-off in selecting appropriate values for these two parameters. Particularly, developing a more precise model may be possible by increasing the values of these two parameters. However, such improvement may inevitably result in a more complicated model (13). In this study, the maximum number of genes allowed in individual is set to 5 by adopting a trial-and-error process.

RESULTS

The temporal variations of confirmed cases of seven infected countries are compared in **Figure 1**. It aims to investigate whether an estimation model developed based on the COVID-19 outbreak in one country can be applied to that of another one. As shown in **Figure 1**, the considered countries, which include China, Korea, Japan, Singapore, Italy, the USA, and Iran, have identified their first cases on different dates. According to **Figure 1**, the temporal variations of daily confirmed cases demonstrate various trends including slow rates, plateau, sharp rising and falling limbs. The combination of these variations makes it very challenging to develop a suitable prediction model. Moreover, it is postulated that the number of cases detected as positive COVID-19 varies with time differently in various countries. Consequently, the prediction model needs to be developed for each country separately.

The time-dependent records of COVID-19 confirmed cases in seven countries were used to develop country-based estimation models. To be more specific, the confirmed cases of China, Korea, Japan, Italy, Singapore, Iran and the USA were utilized as input data to MGGP, while seven different mathematical prediction models were achieved. Exponential functions and arithmetic operators, which have been already utilized to develop mathematical models for predicting the COVID-19 outbreak (17,18), were used in this study. Also, MGGP was employed and assessed for developing country-based estimation models. In the followings, the obtained prediction models for computing confirmed cases of China, Korea, Japan, Italy, Singapore, Iran, and the USA are presented:

(a) China: Since SARS-COV-2 was first detected in Wuhan, Hubei Province, China, this country has a long period of records of confirmed cases among infected countries. In this study, a 77-day record of COVID-19 confirmed cases of China was used to develop a prediction model for the outbreak in China. Application of MGGP to these data yielded to a prediction model shown in Eq. 1:

$$C(x_1) = 89.57x_1^2 \exp(x_1) - 200.1 \exp(0.4237x_1) + \frac{60.87[1.723x_1 - \exp(x_1)]}{0.9823x_1 - \frac{2.751}{x_1}} - \frac{4.736(5.462x_1 - 5.603)}{\frac{0.4209}{x_1} + x_1^2} - 171.6x_1^4 + 200.1 \quad (1)$$

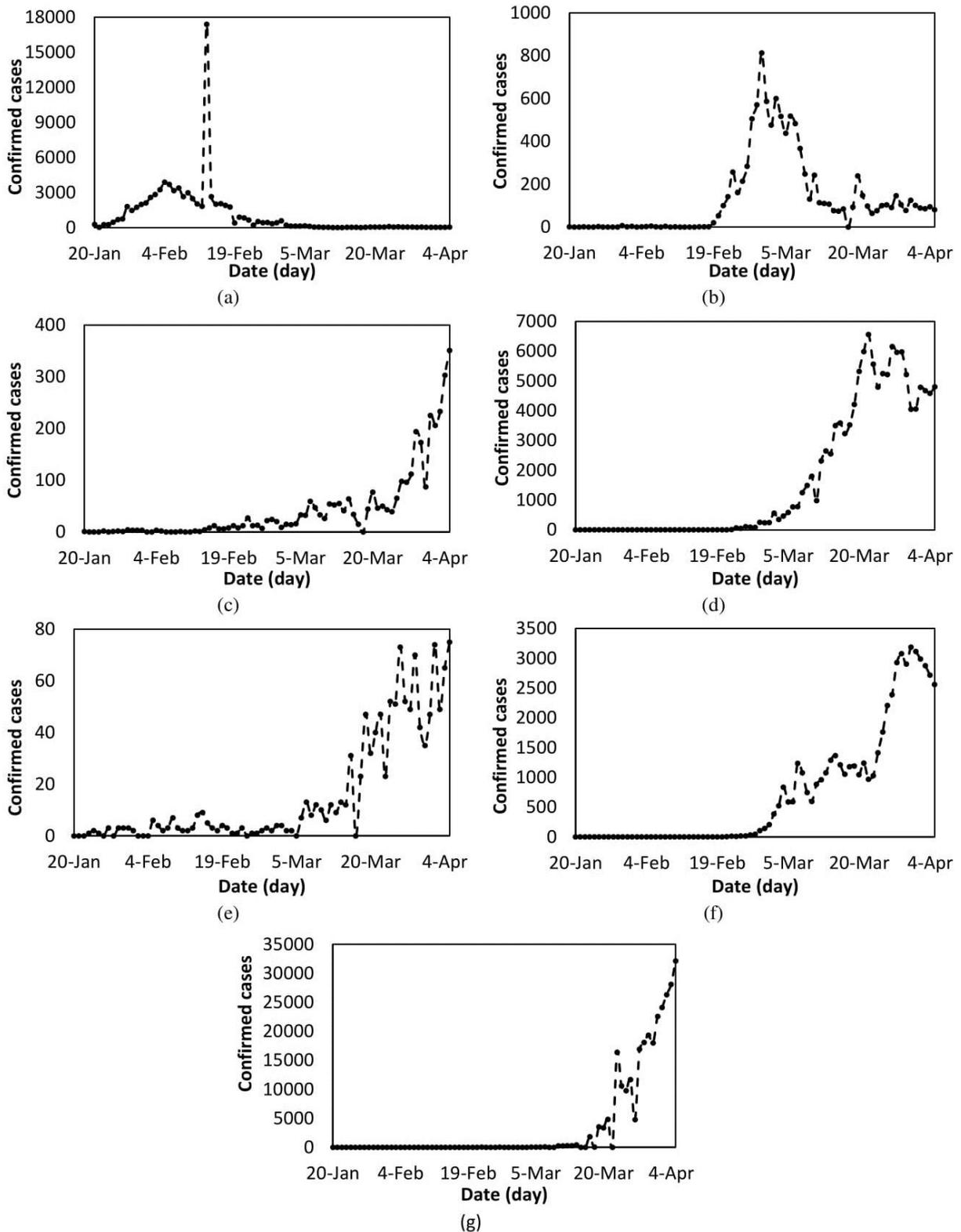


Figure 1. Comparison of trends of confirmed cases in (a) China, (b) Korea, (c) Japan, (d) Italy, (e) Singapore, (f) Iran and (g) the USA

where C is accumulative confirmed cases, $x_1 = \frac{t-1}{76}$, t is time starting from one, and \exp is the exponential function.

(b) Republic of Korea: The total number of confirmed cases in Korea was lower than a hundred cases from 20 January to 19 February 2020, while it increases from lower than 100 to more than 1000 after the next seven days. Eq. 2 shows the prediction model developed by MGGP using a 42-day record of confirmed cases in Korea:

$$\begin{aligned}
 C(x_2) = & 80.83x_2 + 4.826\exp(\exp(x_2)) \\
 & + \frac{785x_2}{\exp(\frac{3.579}{x_2})} - \frac{373.2\exp(x_2^2)}{\exp(1.01x_2)} \\
 & + 76x_2^2 + \frac{1281\exp(x_2^2)}{\exp(\exp(x_2))} - 111.2
 \end{aligned} \tag{2}$$

where $x_2 = \frac{t-1}{50}$ and t is starting from ten.

(c) Japan: Although Japan has infected by COVID-19 soon after China, the variations of confirmed cases reported in this country, as depicted in **Figure 1**, illustrate a slow rate and even plateau in the first month of the outbreak in this country. Using a 77-day record of total confirmed cases of Japan by MGGP resulted in Eq. 3, while t is starting from one in Eq. 3.

$$C(x_1) = 0.0003198 \exp[\exp(2x_1)] - 1.65x_1 + 1.014 \exp(x_1) + \frac{2.199(x_1 + 0.9187)}{x_1 + 3.023} - 1.674 \quad (3)$$

(d) Italy: The COVID-19 outbreak in Italy has a slow rate as the number of confirmed cases was lower than 10 for 24 days from 30 January to 22 February 2020. However, it starts to increase rapidly from 23 February 2020. The 41-day record of the COVID-19 outbreak in Italy was used to develop the prediction model shown in Eq. 4:

$$C(x_3) = -19.15x_3 - \frac{0.8795}{\exp(1.898x_3^3)} - 2.226 \exp(x_3)(x_3 - 9.732) - 2.032 \exp(x_3^2) - 8.297x_3^2 - 18.76 \quad (4)$$

where $x_3 = \frac{t-1}{42}$ and t is starting from three.

(e) Singapore: The 74-day record of confirmed cases of Singapore was used as input data to develop a prediction model for this country. Eq. 5 is achieved by applying MGGP to the outbreak in Singapore:

$$C(x_4) = 3.474x_4 - 29.1 \exp(2x_4) + 20.17 \exp[\exp(x_4)] - 15.97 \exp(2x_4^2) + 34.59x_4^2 - 9.759 \quad (5)$$

where $x_4 = \frac{t}{74}$ and t is starting from one.

(f) Iran: The 37-day record of confirmed cases of Iran was used to develop a prediction model using MGGP. The obtained equation is presented in Eq. 6:

$$C(x_5) = 41.71x_5 - 56.95 \exp(x_5) + 5.321x_5 \exp[\exp(x_5)] + 26.91x_5^2 - 3.326x_5^2 \exp[\exp(x_5)] + 56.98 \quad (6)$$

where $x_5 = \frac{t}{46}$ and t is starting from ten.

(g) USA: From 20 January to 3 March 2020, the total confirmed cases identified in the USA were lower than 100, while it becomes more than 1000 in the next nine days. The 22-day record of confirmed cases from 15 March 2020 was used for developing the following prediction model:

$$C(x_6) = 48.29x_6 \exp(0.8439x_6) - \frac{0.0006372}{\exp(0.912x_6)(x_6^2 - 0.43)} - \frac{14.86x_6}{\exp(x_6^2)} - 33.25x_6 - 41.95x_6^2 - 30.62x_6^3 - 0.004418 \quad (7)$$

where $x_6 = \frac{t-1}{27}$ and t is starting from seven.

The performances of the prediction models developed by MGGP are assessed in **Figure 2**. As shown, the observed total confirmed cases were plotted versus the estimated ones for seven countries: (a) China (from 21 January to 5 April 2020), (b) Republic of Korea (from 24 February to 5 April 2020), (c) Japan (from 12 February to 5 April 2020), (d) Italy (from 25 February to

5 April 2020), (e) Singapore (from 23 January to 5 April 2020), (f) Iran (from 29 February to 5 April 2020), and (g) the USA (from 15 March to 5 April 2020). To be more specific, the x and y points in **Figure 2** represent the observed and predicted confirmed cases in one specific date. Based on **Figure 2**, the points are so much close to the $y=x$ line, which indicates the high precision of predicted models developed by MGGP. Furthermore, by comparing the mathematical models shown in Eq. 1 to Eq. 7, it can be concluded that the variation of confirmed cases and subsequently, the COVID-19 outbreak are country-based. Consequently, MGGP for developing country-based prediction models of the COVID-19 outbreak is suggested.

DISCUSSION

As illustrated in **Figure 1**, each country has an exclusive trend of confirmed cases. Consequently, the prediction model needs to be developed separately for each country. The discrepancy among the trends of confirmed cases in different countries may be due to the following reasons:

First of all, healthcare decision-makers have adopted different strategies: (a) Quarantine: Some countries, such as China and Italy, quarantined the major cities infected. This strategy may prevent the generation of further focal infected centers. However, it may bring about concerns regarding psychological aspects (7). (b) Travel bans and border screening measures, particularly for international communications from the epicenter of the COVID-19 outbreak (Wuhan, Hubei Province): Most countries have applied strict travel bans for China flight transportations. However, some countries were in content with border screening measures such as temperature control (1). (c) Public education on disease prevention: The public health education is one of the effective measures to control SARS-COV-2, like any other transmissible disease. This education includes not only broadcasting sufficient information using media and online services about the way COVID-19 could transmit but also transparency of data flow and up-to-date reports of confirmed and death cases. (d) Active case finding: Developed countries track the confirmed cases and active and/or retroactively search for infected cases (1).

Secondly, demographic characteristics, personal and environmental hygiene, and social determinants of health factors are different for each country. For instance, the hygiene level is quite different across various countries. Moreover, countries have different demographic composition, which this discrepancy becomes more distinguished when a widespread disease like COVID-19 threatens the worldwide health community. In this regard, elderly, immunocompromised individuals or those with preexisting medical conditions are known to be at a higher risk of experiencing severe COVID-19 (19-21).

Third of all, health care facilities, availability of diagnostic kits, and the diagnosis approaches may vary in different countries. For instance, Japan has developed in house PCR-Assay since 16 January 2020 for the diagnosis of COVID-19 (1,22).

Finally, according to the literature, when a virus spreads to an uninfected region, any mutations in the initial viral infections will rapidly become very common, even if they were initially rare in the epicenter of the outbreak. This may lead to minor changes in the outbreak (23).

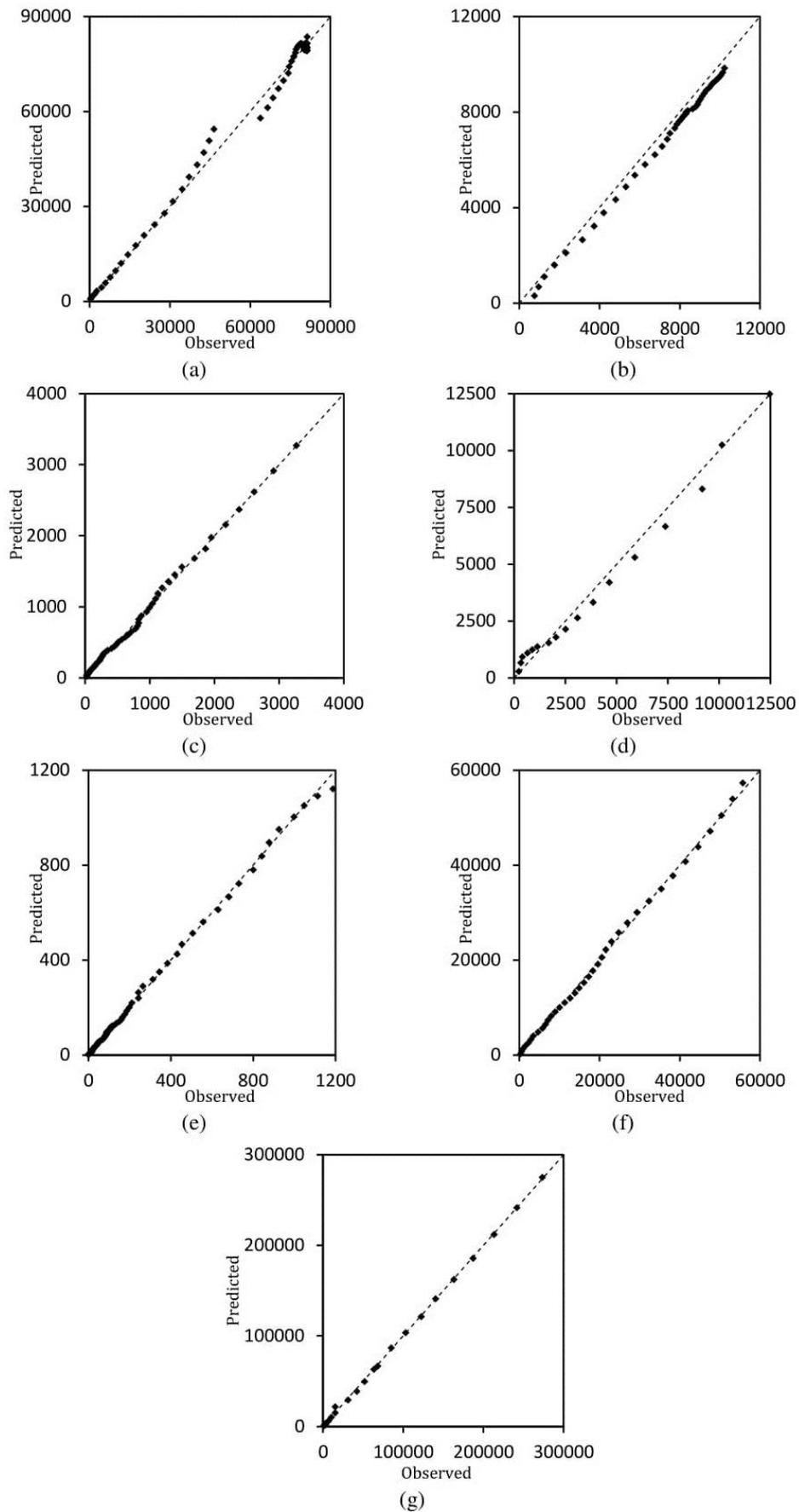


Figure 2. Comparison of observed and predicted confirmed cases in (a) China, (b) Korea, (c) Japan, (d) Italy, (e) Singapore, (f) Iran, and (g) the USA

The estimation models, like those proposed in this study, can provide a possible approximation of COVID-19 threat in the future, while such information inevitably brings about a data-based awareness for healthcare decision-makers. As a result, the implication of precise prediction models may reduce the negative consequences of the COVID-19 outbreak. Furthermore, it emphasizes that an exclusive prediction model is required to be developed for each country. The proposed models in this study, like many other ones available in the literature for the COVID-19 outbreak, have inevitable limitations. First, they are based on the reported number of confirmed cases (2,4,18), which may underestimate the positive cases due to the limited sources for identifying all positive cases. Moreover, prediction models may not take into account asymptomatic patients (patients with positive for COVID-19 with no clinical symptoms) because they are mostly neglected in the reported data. Obviously, the longer a country has experienced COVID-19, the more number of data records is available for the country. This provides enough data to train artificial intelligence models including MGGP, which may result in more accurate prediction models.

The advantage of MGGP in comparison with nonlinear regression models is that the structure of a prediction model does not require to be assumed in advance (24,25). To be more precise, both structure and parameters of a prediction model can be achieved by MGGP. As a result, this advantage enables developing a prediction model without shape limitation, while the user can decide the trade-off between the accuracy and complexity of the prediction model by controlling over the maximum genes allowed in each individual and depth of trees in MGGP (25).

The comparison of proposed models (Eq. 1 to Eq. 7) implies that MGGP resulted in a unique estimation model for each country. It mainly suggested a country-based rather than a universal estimation model. The significant similarity between these equations is that they all exploit exponential function, which can capture a rapidly-increasing trend like that of the COVID-19 outbreak. Since the outbreak trends and the estimation models achieved by MGGP (Eq. 1 to Eq. 7) are significantly different, this study recommends a country-based prediction model rather than a universal one for predicting the COVID-19 outbreak. Furthermore, the exponential function was used by MGGP to develop the COVID-19 prediction models. As a result of the characteristics of the exponential function, the proposed prediction models will be more accurate when they are used to capture a rapidly-increasing trend. However, the beginning period of records of confirmed cases in many countries shows a low rate of increase or even a plateau. Thus, the experience of working with the exponential function using MGGP indicates that more precise prediction models can be achieved when the mentioned periods are excluded. As a result, MGGP captured the trends of confirmed cases of COVID-19 better in countries with significant fluctuations. According to the obtained results, MGGP can predict COVID-19 infected cases very accurately, and it is suggested for estimation of future infected cases, while the availability of data with longer time intervals unarguably helps to provide a more accurate estimation model.

CONCLUSIONS

The trend analysis of the COVID-19 outbreak can be beneficial to healthcare systems as well as planning for national and international measures. In this study, prediction models for the COVID-19 outbreak are developed by multi-gene genetic programming (MGGP) for China, Korea, Japan, Italy, Singapore, Iran, and the USA. The confirmed cases estimated by the proposed models were acceptably close to the observed values. This indicates that the proposed models developed by MGGP yielded to promising results. Comparing the trends of daily confirmed cases of COVID-19 in seven countries demonstrates that each one of these infected countries has a different trend. This exclusiveness requires developing a country-based prediction model. Therefore, a prediction model for one specific country, e.g., China, may not be applicable for other infected ones, while the outbreak of each country needs to be investigated separately.

ACKNOWLEDGEMENTS

The authors would like to dedicate this article to all healthcare providers who devotedly lost their lives during the COVID-19 outbreak. Also, the authors would like to thank the WHO and NHC for providing the online data.

REFERENCES

1. WHO Coronavirus disease (COVID-2019) situation reports. 2020. Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>
2. Bai Y, Yao L, Wei T, Tian F, Jin D-Y, Chen L, et al. Presumed asymptomatic carrier transmission of COVID-19. *Jama*. 2020. <https://doi.org/10.1001/jama.2020.2565> PMID:32083643 PMCID:PMC7042844
3. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*. 2020. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)
4. Niazkar HR, Niazkar M. COVID-19 international outbreak and the need for a suitable estimation model: A second-order polynomial equation with constant coefficients based on imported infected cases seems inadequate. *Asian Pacific Journal of Tropical Medicine*. 2020;13(4):185-6. <https://doi.org/10.4103/1995-7645.280234>
5. Cakir Z, Savas H. A Mathematical Modelling Approach in the Spread of the Novel 2019 Coronavirus SARS-CoV-2 (COVID-19) Pandemic. *Electron J Gen Med*. 2020;17(4):em205. <https://doi.org/10.29333/ejgm/7861>
6. Bai Z, Gong Y, Tian X, Cao Y, Liu W, Li J. The Rapid Assessment and Early Warning Models for COVID-19. *Virologica Sinica*. 2020;1. <https://doi.org/10.1007/s12250-020-00219-0> PMID:32239446 PMCID:PMC7110270
7. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*. 2020;395(10228):931-4. [https://doi.org/10.1016/S0140-6736\(20\)30567-5](https://doi.org/10.1016/S0140-6736(20)30567-5)

8. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*. 2020. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9)
9. Gao Y, Zhang Z, Yao W, Ying Q, Long C, Fu X. Forecasting the cumulative number of COVID-19 deaths in China: a Boltzmann function-based modeling study. *Infection Control & Hospital Epidemiology*. 2020;1-16. <https://doi.org/10.1017/ice.2020.101>
10. Williams AR, Bisaga A. From AIDS to opioids—how to combat an epidemic. *New England Journal of Medicine*. 2016;375(9):813-5. <https://doi.org/10.1056/NEJMp1604223> PMID:27579632 PMCID:PMC5517310
11. Niazkar M, Talebbeydokhti N, Afzali SH. Novel grain and form roughness estimator scheme incorporating artificial intelligence models. *Water resources management*. 2019;33(2):757-73. <https://doi.org/10.1007/s11269-018-2141-z>
12. Niazkar M. Revisiting the estimation of Colebrook friction factor: A comparison between artificial intelligence models and CW based explicit equations. *KSCE Journal of Civil Engineering*. 2019;23(10):4311-26. <https://doi.org/10.1007/s12205-019-2217-1>
13. Searson D. GPTIPS: Genetic programming and symbolic regression for MATLAB. 2009.
14. Niazkar M, Talebbeydokhti N, Afzali S-H. Bridge backwater estimation: A Comparison between artificial intelligence models and explicit equations. *Scientia Iranica*. 2020. <https://doi.org/10.24200/SCI.2020.51432.2175>
15. Niazkar M, Talebbeydokhti N, Hosein Afzali S. Application of Artificial Intelligence Models to Estimate Discharge over Semicircular Weirs. *Advance Researches in Civil Engineering*. 2019;1(4):22-34. <https://doi.org/10.30469/ARCE.2019.96032>
16. Searson DP, Leahy DE, Willis MJ, editors. GPTIPS: an open source genetic programming toolbox for multigene symbolic regression. *Proceedings of the International multicongference of engineers and computer scientists*; 2010.
17. Fu X, Ying Q, Zeng T, Long T, Wang Y. Simulating and forecasting the cumulative confirmed cases of SARS-CoV-2 in china by Boltzmann function-based regression analyses. *Journal of Infection*. 2020. <https://doi.org/10.1101/2020.02.16.20023564>
18. Li Q, Feng W, Quan Y-H. Trend and forecasting of the COVID-19 outbreak in China. *Journal of Infection*. 2020;80(4):469-96. <https://doi.org/10.1016/j.jinf.2020.02.014>
19. Liu K, Chen Y, Lin R, Han K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *Journal of Infection*. 2020. <https://doi.org/10.1016/j.jinf.2020.03.005> PMCID:PMC7102640
20. Abdulmir AS, Hafidh RR. The Possible Immunological Pathways for the Variable Immunopathogenesis of COVID--19 Infections among Healthy Adults, Elderly and Children. *Electronic Journal of General Medicine*. 2020;17(4):em202. <https://doi.org/10.29333/ejgm/7850>
21. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of autoimmunity*. 2020:102433. <https://doi.org/10.1016/j.jaut.2020.102433> PMID:32113704 PMCID:PMC7127067
22. Corman V, Bleicker T, Brünink S, Zambon M. Diagnostic detection of Wuhan coronavirus 2019 by real-time RT-PCR. Geneva: World Health Organization, January. 2020;13.
23. Tang X, Wu C, Li X, Song Y, Yao X, Wu X, et al. On the origin and continuing evolution of SARS-CoV-2. *National Science Review*. 2020;nwaa036. <https://doi.org/10.1093/nsr/nwaa036> PMCID:PMC7107875
24. Mohammadi Bayazidi A, Wang G-G, Bolandi H, Alavi AH, Gandomi AH. Multigene genetic programming for estimation of elastic modulus of concrete. *Mathematical Problems in Engineering*. 2014;2014. <https://doi.org/10.1155/2014/474289>
25. Lee J-S, Suh K-D. Development of Stability Formulas for Rock Armor and Tetrapods Using Multigene Genetic Programming. *Journal of Waterway, Port, Coastal, and Ocean Engineering*. 2020;146(1):04019027. [https://doi.org/10.1061/\(ASCE\)WW.1943-5460.0000540](https://doi.org/10.1061/(ASCE)WW.1943-5460.0000540)