Current Research in Behavioral Sciences 1 (2020) 100002

Contents lists available at ScienceDirect

Current Research in Behavioral Sciences

journal homepage: www.elsevier.com/locate/crbeha

Will COVID-19 confirmed cases in the USA reach 3 million? A forecasting approach by using SutteARIMA Method

Ansari Saleh Ahmar[a,*], Eva Boj[b]

[a] Department of Statistics, Faculty of Mathematics and Natural Science, Universitas Negeri Makassar, Indonesia
[b] Department of Economic, Financial and Actuarial Mathematics, Faculty of Economics and Business, Universitat de Barcelona, Spain

ARTICLE INFO

Keyword:
Covid-19
Forecast
USA

ABSTRACT

Objectives: Forecasting the number of COVID-19 cases in the USA can provide an overview and projection of the development of COVID-19 cases in the US so that policy makers can determine the steps that must be taken. This study aimed to determine whether COVID-19 confirmed cases in the USA would reach 3 million cases with the SutteARIMA method forecasting approach.
Methods: Data from this study were obtained from the Worldometer data on 15 February 2020 to 2 July 2020. Data from 15 February 2020 to 25 June 2020 were used to performance data fitting (26 June 2020–2 July 2020). Data fitting is used to examine the extent of the accuracy of the SutteARIMA method in predicting data. To examine the level of the data accuracy, the MAPE method was used in this study.
Results: The results of forecasting data fitting on 26 June 2020 – 2 July 2020: 2,544,732; 2,590,888; 2,632,477; 2,671,055; 2,711,798; 2,755,128; 2,803,729. The accuracy of SutteARIMA for the period of 26 June 2020–2 July 2020 based on MAPE was 0.539% and the forecasting results that had been obtained were 3 million confirmed cases, namely from 05 to 06 June 2020: 1,981,299; 2,005,706; 2,030,283; 2,055,031.
Conclusions: The SutteARIMA method predicted that 2 million confirmed cases of COVID-19 will be obtained on the WHO situation report on days 168–170 or 05–07 June 2020.

1. Introduction

The first confirmed cases of COVID-19 in USA were reported on 20 January 2020, in Snohomish County, Washington DC (Holshue et al., 2020; Moghadas et al., 2020; Harcourt et al., 2020). COVID-19 is an infectious disease caused by a new coronavirus (SARS-CoV-2) discovered in China (Yang et al., 2020). Based on data presented by Worldometer on 20 April 2020, the number of confirmed cases of COVID-19 in United State was 792,759 cases or added 28,123 cases from yesterday (19 April 2020) with 42,514 total deaths and was the first highest country with confirmed cases of COVID-19 in the world (Worldometer, 2020).

To determine the case rate further in the future, it is necessary to forecast the data. COVID-19 forecasting is important to determine an overview of the increase in COVID-19 cases from time to time, so that it can be taken into consideration in the decision making process as the prevention and control of disease, this is in accordance with the statement Myers et al. (2000) who stated that an accurate disease forecasting model can improve the level of epidemic prevention and control. There were several researchers who had studied about the COVID-19 i.e. A. Ojugo and O. D. Otakore studied about Bayesian network model for detection of the Covid-19 (Ojugo and Otakore, 2020); M. Azarafza, M. Azarafza, and H. Akgün who studied about the spread pattern analysis of corona-virus (COVID-19) infection in Iran by using Cluster method (Azarafza et al., 2020); Hernández-V et al. who studied about the association of comorbidities with pneumonia and death on COVID-19 patients in Mexico (Hernández-V et al., 2020); and there were some Forecasting or predictions related to COVID-19 that had been studied by various researchers: Koczkodaj et al. who predicted about COVID-19 outside of China by using a simple heuristic (exponential curve) (Koczkodaj et al., 2020); and Roosa et al. who studied about COVID-19 real-time forecast in China with generalized logistic growth model (GLM) (Roosa et al., 2020).

2. Methods

SutteARIMA was used as the method in this study because this method is a combination of ARIMA and α-Sutte Indicator method in which ARIMA method is used by several health researchers to monitor and predict the development of a disease and SutteARIMA has a small value of MAPE when compared to ARIMA. The formula of SutteARIMA

* Corresponding author.
E-mail addresses: ansarisaleh@unm.ac.id (A.S. Ahmar), evaboj@ub.edu (E. Boj).

https://doi.org/10.1016/j.crbeha.2020.100002
Received 13 August 2020; Received in revised form 14 September 2020; Accepted 15 September 2020

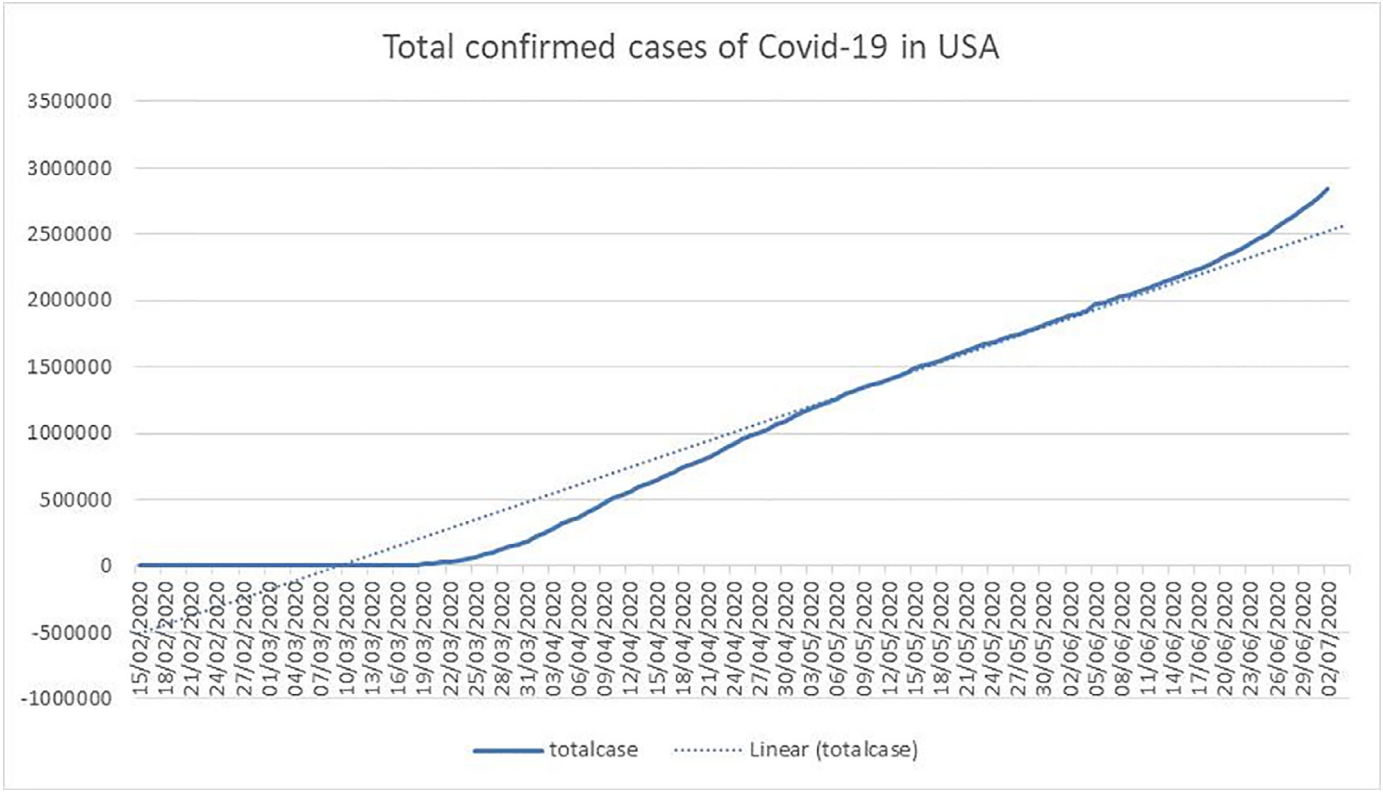

Fig. 1. Total confirmed cases and linear trend forecasts (total cases) of COVID-19 in USA (15 February 2020–02 July 2020).

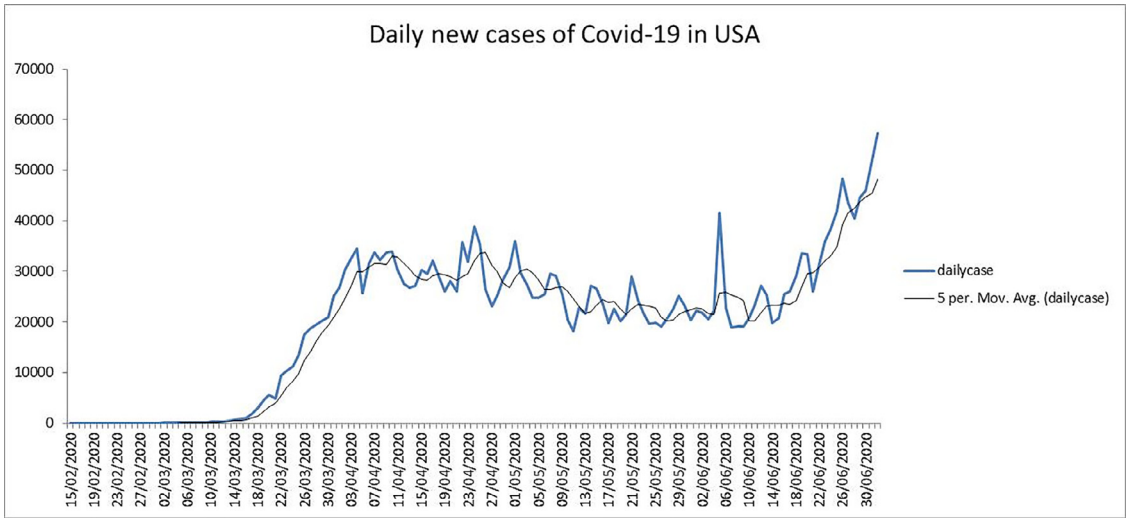

Fig. 2. Daily new cases of COVID-19 in USA (15 February 2020–02 June 2020).

(Ahmar and Boj, 2020a; Ahmar and Boj, 2020b; A.S. Ahmar and Boj, 2020):

$$Z_t = \alpha\left(\frac{\phi_1}{2} + \frac{\Delta z}{3\alpha + 3\beta}\right) + \beta\left(\frac{\phi_3}{2} + \frac{2\Delta y}{3\beta + 3\gamma}\right) + \gamma\left(\frac{\phi_3}{2} + \frac{2\Delta x}{3\gamma + 3\delta}\right) + \frac{\phi_4\delta}{2} + \cdots + \frac{\phi_p Z_{t-p}}{2} + \frac{a_t}{2} - \frac{\theta_1 a_{t-1}}{2} - \frac{\theta_2 a_{t-2}}{2} - \cdots - \frac{\theta_q a_{t-q}}{2}$$

where:

$\delta = Z_{t-4}$

$\gamma = Z_{t-3}$

$\beta = Z_{t-2}$

$\alpha = Z_{t-1}$

$\Delta x = \gamma - \delta = Z_{t-3} - Z_{t-4}$

$\Delta y = \beta - \gamma = Z_{t-2} - Z_{t-3}$

$\Delta z = \alpha - \beta = Z_{t-1} - Z_{t-2}$

Z_t = data at t time
$Z_{t\text{-}p}$ = data at $(t - p)$ time

Table 1
Results of fitting confirmed cases of COVID-19 in USA.

Date	Actual	Forecast	APE
26/06/2020	2,552,956	2,544,732	0.00322
27/06/2020	2,596,537	2,590,888	0.00218
28/06/2020	2,637,077	2,632,477	0.00174
29/06/2020	2,681,811	2,671,055	0.00401
30/06/2020	2,727,853	2,711,798	0.00589
01/07/2020	2,779,953	2,755,128	0.00893
02/07/2020	2,837,189	2,803,729	0.01179
		MAPE	0.00539

Table 2
Forecast of 2 million confirmed cases of COVID-19 in USA.

Date	Forecast	Lower boundary (95% confidence)	Upper boundary (95% confidence)
03/07/2020	2,898,803	2,753,863	3,043,743
04/07/2020	2,965,866	2,817,573	3,114,159
05/07/2020	3,038,127	2,886,221	3,190,033
06/07/2020	3,115,591	2,959,811	3,271,371
07/07/2020	3,198,540	3,038,613	3,358,467

Data from this study were obtained from the Worldometer data on 15 February 2020 to 02 July 2020. Data from 15 February 2020 to 25 June 2020 were used to perform data fitting (26 June 2020–2 July 2020). Data fitting is used to examine the extent of the accuracy of the SutteARIMA method in predicting data. To examine the level of the data accuracy, the MAPE method was used in this study.

3. Ethics statement

The forecasting methods used freely available data from the public platform source with no primary data collection to human and animal subjects. Therefore, no institutional review board approval is needed.

4. Results

The total confirmed cases and daily new cases in USA can be seen in Figs. 1 and 2.

To further examine the likelihood of achieving the 3 million confirmed cases in the USA, it is necessary to forecast the data. Forecasting method applied in this study used the SutteARIMA method. The first stage carried out in this study was to determine the level of forecasting accuracy. Testing the accuracy of forecasting from the SutteARIMA method was done by fitting the data on the period of 26 June 2020–2 July 2020 by using the data from the period of 15 February 2020 to 25 June 2020. The level of forecasting accuracy can be seen from the mean absolute percentage error (MAPE). The results of this forecast are presented in Table 1 and the results of the forecast to 3 million cases are presented in Table 2.

5. Discussion

Based on Fig. 1 it can be seen that the COVID-19 confirmed cases have an upward trend from time to time. This trend gives an indication that 2 million cases will be achieved in a short time. This is also reinforced from Fig. 2 which shows that the average number of daily confirmed cases is around 40,000–60,000 cases (last 3 days) and if it refers to the moving average chart (5 periods) then the addition of cases per day will be around 40,000–50,000 cases (last 3 days).

Table 1 shown the accuracy of SutteARIMA for the period 26 June 2020–2 July 2020 is 0.539%. By looking at the value of the small level of accuracy of the SutteARIMA method, the SutteARIMA method was used to predict when the USA confirmed cases reach 3 million cases. According to Table 2, SutteARIMA method predicted that 3 million cases of COVID-19 in the USA will be obtained on days 168–170 from the WHO situation report or on 05–07 July 2020.

The SutteARIMA method is a combination of the ARIMA method and the α-Sutte Indicator. In this study, the SutteARIMA method had an accuracy rate on data fittings of 0.539%. Based on the results of this study, the SutteARIMA method predicted that 3 million cases will be obtained on days 168–170 from the WHO situation report or on 05–07 July 2020. It is expected that from the results of this study, it can provide input to the USA government to be able to take policies, so that the spread of COVID-19 did not become severe, because based on the results of the continued analysis of the SutteARIMA method, if the addition of these cases is stagnant, it is predicted that there will be around 4 million cases on end of July 2020.

Funding

None.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Ansari Saleh Ahmar: Conceptualization, Data curation, Formal analysis, Methodology, Writing - original draft, Writing - review & editing. **Eva Boj:** Conceptualization, Methodology, Writing - review & editing.

References

Ahmar, A.S., Boj, E., 2020a. SutteARIMA: short-term forecasting method, a case: Covid-19 and stock market in Spain. Sci. Total Environ., 138883 doi:10.1016/j.scitotenv.2020.138883.

Ahmar, A.S., Boj, E., 2020b. The date predicted 200.000 cases of Covid-19 in Spain. J. Appl. Sci. Eng. Technol. Educ. 2. doi:10.35877/454RI.asci22102.

Azarafza, M., Azarafza, M., Akgün, H., 2020. Clustering method for spread pattern analysis of corona-virus (COVID-19) infection in Iran. J. Appl. Sci. Eng. Technol. Educ. 3, 1–6. doi:10.35877/454ri.asci31109.

Harcourt, J., Tamin, A., Lu, X., Kamili, S., Sakthivel, S.K., Murray, J., et al., 2020. Isolation and characterization of SARS-CoV-2 from the first US COVID-19 patient. BioRxiv doi:10.1101/2020.03.02.972935.

Hernández-V, A., Azañedo, D., Vargas-Fernánde, R., Bendezu-Quispe, G., 2020. Association of comorbidities with pneumonia and death among COVID-19 patients in Mexico: a nationwide cross-sectional study. Korean J. Prev. Med. 0. doi:10.3961/jpmph.20.186.

Holshue, M.L., DeBolt, C., Lindquist, S., Lofy, K.H., Wiesman, J., Bruce, H., et al., 2020. First case of 2019 novel coronavirus in the United States. N. Engl. J. Med. 382, 929–936. doi:10.1056/NEJMoa2001191.

Koczkodaj, W.W., Mansournia, M.A., Pedrycz, W., Wolny-Dominiak, A., Zabrodskii, P.F., Strzaška, D., et al., 2020. 1,000,000 cases of COVID-19 outside of China: the date predicted by a simple heuristic. Global Epidemiol., 100023 doi:10.1016/j.gloepi.2020.100023.

Moghadas, S.M., Shoukat, A., Fitzpatrick, M.C., Wells, C.R., Sah, P., Pandey, A., et al., 2020. Projecting hospital utilization during the COVID-19 outbreaks in the United States. Proc. Natl. Acad. Sci. 117, 9122–9126. doi:10.1073/pnas.2004064117.

Myers M.F., Rogers D.J., Cox J., Flahault A., Hay S.I. Forecasting Disease Risk for Increased Epidemic Preparedness in Public Health, 2000, p. 309–30. 10.1016/S0065-308X(00)47013-2.

Ojugo, A., Otakore, O.D., 2020. Forging an optimized bayesian network model with selected parameters for detection of the coronavirus in Delta State of Nigeria. J. Appl. Sci. Eng. Technol. Educ. 3, 37–45. doi:10.35877/454RI.asci2163.

Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J.M., et al., 2020. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. Infect. Dis. Model. 5, 256–263. doi:10.1016/j.idm.2020.02.002.

Worldometer. Spain Coronavirus 2020. https://www.worldometers.info/coronavirus/country/spain/ (accessed April 8, 2020).

Yang, Y., Peng, F., Wang, R., Guan, K., Jiang, T., Xu, G., et al., 2020. The deadly coronaviruses: the 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China. J. Autoimmun., 102434 doi:10.1016/j.jaut.2020.102434.