Application of Neural Network Time Series (NNAR) and ARIMA to Forecast Infection Fatality Rate (IFR) of COVID-19 in Brazil

by Ansari Saleh Ahmar

Submission date: 07-May-2023 04:02PM (UTC-0500)
Submission ID: 2086739515
File name: 2 - 372-1238-1-PB NNAR JOIV.pdf (3.27M)
Word count: 2025
Character count: 9304

JOIV : Int. J. Inform. Visualization, 5(1) - March 2021 8-10

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Application of Neural Network Time Series (NNAR) and ARIMA to Forecast Infection Fatality Rate (IFR) of COVID-19 in Brazil

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

[a]Department of Statistics, Universitas Negeri Makassar, Daeng Tata Street, Makassar, 90223, Indonesia
[b]Department of Economic, Financial and Actuarial Mathematics, Faculty of Economics and Business, Universitat de Barcelona, Spain
Corresponding author: *ansarisaleh@unm.ac.id

***Abstract*— Forecasting is a method that is often used to view future events using past time data. Past time data have useful information to use in obtaining the future. The aim of this study was to forecast infection fatality rate (IFR) of COVID-19 in Brazil using NNAR and ARIMA. ARIMA and NNAR are used because (1) ARIMA is a simple stochastic time series method that can be used to train and predict future time points and ARIMA also capable of capturing dynamic interactions when it uses error terms and observations of lagged terms; (2) the Artificial Neural Network (ANN) is a technique capable of analyzing certain non-linear interactions between input regressor and responses, and Neural Network Time Series (NNAR) is one method of ANN in which lagged time series values were used as inputs to a neural network. Data included in this study were derived from the total data of confirmed cases and the total data of death of COVID-19. The data of COVID-19 in Brazil from February 15, 2020 to April 30, 2020 were collected from the Worldometer (https://www.worldometers.info/coronavirus/) and Microsoft Excel 2013 was used to build a time-series table. Forecasting was accomplished by means of a time series package (forecast package) in R Software. Neural Network Time Series and ARIMA models were applied to a dataset consisting of 76 days. The accuracy of forecasting was examined by means of an MSE. The forecast of IFR of COVID-19 in Brazil from May 01, 2020 to May 10, 2020 with NNAR (1,1) model was around in 6,85% and ARIMA (0,2,1) was around in 7.11%.**

***Keywords*— Infection disease; COVID-19; forecasting; NNAR; ARIMA.**

Manuscript received 7 Sep. 2020; revised 11 Jan. 2021; accepted 28 Feb. 2021. Date of publication 31 Mar. 2021.

I. Introduction

Wuhan, Hubei Province, China is the location that COVID-19 first reported in December 2019. Based on WHO [1] data, as of May 1, 2020, there were 3,175,207 confirmed cases and 224,172 confirmed deaths. Based on the data presented by Worldometer on May 1, 2020, the number of confirmed cases of COVID-19 in Brazil was amounted to 78,162 people with 5,466 deaths and was the second highest country in the America based on confirmed cases of COVID-19 [2]. There have been many researchers who have studied about COVID-19, specifically on Forecasting e.g. Forecasting data of COVID-19 in Spain [3][4], USA [5], and World [6]; and other studies about COVID-19 [7]–[9] This forecasting was important to find out the spread of COVID-19 cases and the rate of fatality.

II. Material and Method

The neural network can be thought of as a network of "neurons" that are organized in layers. The inputs form the bottom layer, and the projections (or outputs) form the top layer. Intermediate layers containing "hidden neurons" can also be present. The neural network autoregression or NNAR model is the lagged time series values can be used as inputs to a neural network, just like we used lagged values in a linear autoregression model. The NNAR(p, k) model indicates that there are p lagged inputs and k nodes in the hidden layer. For example, the NNAR(6,2) model is a neural network with the last six observations used as inputs to predict the output Z_t and two neurons in the hidden layer [10].

The Z_t process is an autoregressive-moving average or ARMA (p,q) model model if it fulfilled [3][4]:

$$\phi_p(B)Z_t = \theta_q(B)a_t, a_t \sim WN(0,\sigma^2), \phi_p, \theta_q \in \mathbb{R}, t \in \mathbb{Z}.$$

with $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$ (for AR(p)) and

$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q)$ (for MA(q)).

If there is a differencing then the ARIMA model becomes as follows [3][4]:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t, a_t \sim WN(0,\sigma^2), \phi_p, \theta_q \in \mathbb{R}, t \in \mathbb{Z}.$$

with $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p)$ (for AR(p)), $(1 - B)^d$ (for differencing non-seasonal) and $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q)$ (for MA(q)).

The fatality rate of infection, the possibility of death for an infected person, is one of the most significant aspects of the coronavirus disease pandemic of 2019 (COVID-19). The infection fatality risk is directly linked to the estimated overall mortality burden of COVID-19. Furthermore, the infection fatality rate is used to justify a variety of non-pharmacological public health measures [11].

The infection fatality rate (IFR) data of COVID-19 in Brazil from February 15, 2020 to April 30, 2020 were collected from the Worldometer website (https://www.worldometers.info/coronavirus), and Microsoft Excel 2013 was used to build a time-series database. Forecasting was accomplished by means of a time series package (*forecast package*) in R Software. Neural Network Time Series (NNAR) and ARIMA models were applied to a dataset consisting of 76 days. ARIMA and NNAR are used because (1) ARIMA is a simple stochastic time series method that can be used to train and predict future time points and ARIMA also capable of capturing dynamic interactions when it uses error terms and observations of lagged terms; (2) the Artificial Neural Network (ANN) is a technique capable of analyzing certain non-linear interactions between input regressor and responses, and Neural Network Time Series (NNAR) is one method of ANN in which lagged time series values were used as inputs to a neural network. The accuracy of forecasting was examined by means of an MSE [12].

III. Result and Discussion

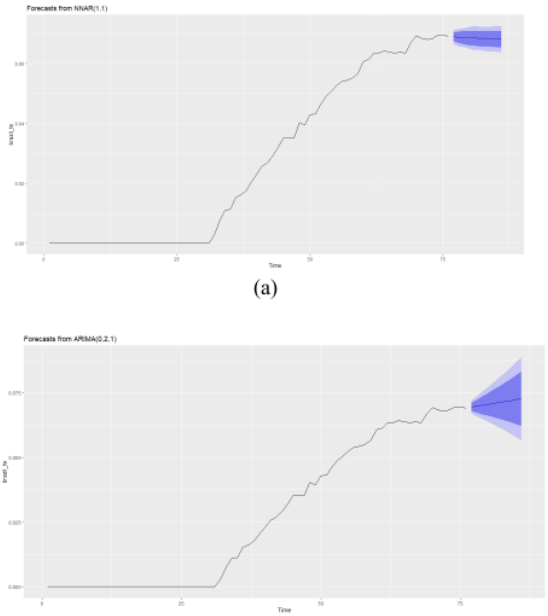
Table 1 shown that the average total confirmed, total death, and infection fatality rate in Brazil was amounted to 14242 cases, 887 cases, and 2,69% (from February 15, 2020 to April 30, 2020) and the maximum total of confirmed cases and death were amounted to 85380 cases and 5901 cases.

TABLE I
Descriptive Statistics of COVID-19 in Brazil

	Total cases	Total death	IFR (%)
Mean	14242	887	2,69
Median	2086	40	1,91
Standard Deviation	21902	1496	2,74
Minimum	0	0	0,00
Maximum	85380	5901	6,95

Based on the analysis data using forecast package (R Software), NNAR(1,1) and ARIMA(0,2,1) models for COVID-19 were obtained in Brazil. The forecast value, lower and upper limit 80% confidence interval, and lower and upper limit 95% confidence interval of infection fatality rate (IFR) for 5-day period from May 12, 2020 to May 16, 2020 using NNAR(5,3) and ARIMA(0,2,1) models were reported in Fig. 1, table 1, and table 2.

Fig. 1 shown that the forecast of NNAR(1,1) was constant linear around 6,8% but forecast of ARIMA(0,2,1) increased until 7,2%.



(a)

(b)

Fig. 1. (a) Forecast of NNAR(1,1) and (b) ARIMA(0,2,1) for COVID-19 in Brazil

TABLE II
Forecast value of NNAR(1,1) for the 10 days of Infection Fatality Rate (IFR) of COVID-19 in Brazil

Date	Infection Fatality Rate (%)				
	Forecast	Low 80	High 80	Low 95	High 95
01/05/2020	6,89	6,74	7,05	6,64	7,12
02/05/2020	6,88	6,69	7,07	6,57	7,16
03/05/2020	6,87	6,64	7,09	6,54	7,19
04/05/2020	6,86	6,61	7,08	6,50	7,22
05/05/2020	6,85	6,59	7,09	6,44	7,25
06/05/2020	6,84	6,58	7,08	6,44	7,23
07/05/2020	6,83	6,56	7,07	6,42	7,24
08/05/2020	6,83	6,55	7,08	6,39	7,22

Date	Infection Fatality Rate (%)				
	Forecast	Low 80	High 80	Low 95	High 95
09/05/2020	6,83	6,55	7,09	6,39	7,26
10/05/2020	6,82	6,54	7,09	6,36	7,23
Mean	6,85	6,60	7,08	6,47	7,21

TABLE III
FORECAST VALUE OF ARIMA(0,2,1) FOR THE 10 DAYS OF INFECTION FATALITY RATE (IFR) OF COVID-19 IN BRAZIL

Date	Infection Fatality Rate (%)				
	Forecast	Low 80	High 80	Low 95	High 95
01/05/2020	6,95	6,78	7,11	6,69	7,20
02/05/2020	6,98	6,72	7,24	6,59	7,38
03/05/2020	7,02	6,67	7,37	6,48	7,55
04/05/2020	7,05	6,61	7,49	6,38	7,73
05/05/2020	7,09	6,56	7,62	6,27	7,91
06/05/2020	7,13	6,50	7,76	6,16	8,09
07/05/2020	7,16	6,43	7,89	6,05	8,28
08/05/2020	7,20	6,36	8,03	5,92	8,47
09/05/2020	7,23	6,29	8,17	5,79	8,67
10/05/2020	7,27	6,22	8,32	5,66	8,88
Mean	7,11	6,51	7,70	6,20	8,01

Based on table 1, forecast of IFR of COVID-19 in Brazil from May 01, 2020 to May 10, 2020 with NNAR(1,1) model was around in 6,85% and ARIMA(0,2,1) was around in 7,11%. Modeling COVID-19, in particular Brazil, is very complicated as it related to the number of assessments, procedures, randomness, stay at home, social restrictions/social distancing/physical distancing, and several other considerations.

IV. CONCLUSIONS

Based on the result and discussion, it showed that the model forecasting of IFR of COVID-19 in Brazil were NNAR(1,1) and ARIMA(0,2,1). The results of this study were obtained with an error rate of 6,85% for NNAR and 7,11% for ARIMA. The findings of this modeling will only be used for learning and research purposes, as well as an overview of decision-makers for the next 10 days relevant to the growth of the COVID-19 situation.

REFERENCES

[1] WHO, "Situation report - 77 Coronavirus disease 2019 (COVID-19," 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/. [Accessed: 07-Apr-2020].

[2] Worldometer, "Spain Coronavirus," 2020. [Online]. Available: https://www.worldometers.info/coronavirus/country/spain/. [Accessed: 08-Apr-2020].

[3] A. S. Ahmar and E. Boj, "SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock market in Spain," *Sci. Total Environ.*, p. 138883, Apr. 2020.

[4] A. S. Ahmar and E. Boj, "The date predicted 200.000 cases of COVID-19 in Spain," *J. Appl. Sci. Eng. Technol. Educ.*, vol. 2, no. 2, Jun. 2020.

[5] A. S. Ahmar and E. Boj, "Will COVID-19 confirmed cases in the USA reach 3 million? A forecasting approach by using SutteARIMA Method," *Curr. Res. Behav. Sci.*, vol. 1, p. 100002, 2020.

[6] A. S. Ahmar and R. Rusli, "Will COVID-19 cases in the World reach 4 million? a forecasting approach using SutteARIMA," *JOIV Int. J. Informatics Vis.*, vol. 4, no. 3, p. 159, 2020.

[7] E. Yeboah and R. G. Kyeremeh, "COVID-19 in Ghana: An Overview of Government Contributions in Fighting this Pandemic," *Quant. Econ. Manag. Stud.*, vol. 1, no. 6, pp. 375–382, 2020.

[8] S. R. Saha and D. M. M. H. Khan, "Prevalence and determinants of mental distress during COVID-19 outbreak in Bangladesh: evidence from an online survey," *J. Appl. Sci. Eng. Technol. Educ.*, vol. 3, no. 1, pp. 90–103, 2021.

[9] R. E. Grewelle and G. A. De Leo, "Estimating the Global Infection Fatality Rate of COVID-19," *medRxiv*, p. 2020.05.11.20098780, Jan. 2020.

[10] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.

[11] J. P. A. Ioannidis, "Infection fatality rate of COVID-19 inferred from seroprevalence data," *Bull. World Health Organ.*, vol. 99, no. 1, pp. 19-33F, Jan. 2021.

[12] A. S. Ahmar, "Forecast Error Calculation with Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE)," *JINAV J. Inf. Vis.*, vol. 1, no. 2, 2020.

Application of Neural Network Time Series (NNAR) and ARIMA to Forecast Infection Fatality Rate (IFR) of COVID-19 in Brazil

ORIGINALITY REPORT

16%	9%	11%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	İbrahim Demir, Murat Kirişci. "Forecasting COVID-19 disease cases using the SARIMA-NNAR hybrid model", Cold Spring Harbor Laboratory, 2021 Publication	2%
2	Emanuel Montero, Carlos Rocha, Helder Oliveira, Eduardo Cerqueira, Paulo Mendes, Aldri Santos, Denis Rosário. "Proactive radio- and QoS-aware UAV as BS deployment to improve cellular operations", Computer Networks, 2021 Publication	2%
3	ifjir.com Internet Source	2%
4	biomedres.us Internet Source	1%
5	Tanujit Chakraborty, Swarup Chattopadhyay, Indrajit Ghosh. "Forecasting dengue epidemics using a hybrid methodology", Cold Spring Harbor Laboratory, 2018	1%

	Publication	
6	pubmed.ncbi.nlm.nih.gov Internet Source	1 %
7	repository.ubaya.ac.id Internet Source	1 %
8	Rashmi Priya, Dharavath Ramesh, Venkanna Udutalapally. "NSGA-2 Optimized Fuzzy Inference System for Crop Plantation Correctness Index Identification", IEEE Transactions on Sustainable Computing, 2022 Publication	1 %
9	ejurnal.politeknikpratama.ac.id Internet Source	1 %
10	www.ncbi.nlm.nih.gov Internet Source	1 %
11	eudl.eu Internet Source	1 %
12	static.bsu.az Internet Source	1 %
13	Martin Neil, Norman Fenton, Magda Osman, Scott McLachlan. "Bayesian network analysis of Covid-19 data reveals higher infection prevalence rates and lower fatality rates than widely reported", Journal of Risk Research, 2020 Publication	1 %

14	Janmenjoy Nayak, Pandit Byomakesha Dash, Bighnaraj Naik, Subhashree Mohapatra, A. R. Routray. "Deep Learning-Based Trend Analysis on Indian Stock Market in COVID-19 Pandemic Scenario and Forecasting Future Financial Drift", Journal of The Institution of Engineers (India): Series B, 2022 Publication	<1 %
15	Mahdi Kalantari. "Forecasting COVID-19 pandemic using optimal singular spectrum analysis", Chaos, Solitons & Fractals, 2021 Publication	<1 %
16	Ovidiu-Dumitru Ilie, Roxana-Oana Cojocariu, Alin Ciobica, Sergiu-Ioan Timofte, Ioannis Mavroudis, Bogdan Doroftei. "Forecasting the Spreading of COVID-19 across Nine Countries from Europe, Asia, and the American Continents Using the ARIMA Models", Microorganisms, 2020 Publication	<1 %
17	thetattyjournal.org Internet Source	<1 %
18	John P.A. Ioannidis, Cathrine Axfors, Despina G. Contopoulos-Ioannidis. "Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in	<1 %

pandemic epicenters", Environmental Research, 2020
Publication

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		

Application of Neural Network Time Series (NNAR) and ARIMA to Forecast Infection Fatality Rate (IFR) of COVID-19 in Brazil

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3