



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# Forecasting daily Covid-19 cases in the world with a hybrid ARIMA and neural network model

Lucas Rabelo de Araújo Morais<sup>\*</sup>, Gecynalda Soares da Silva Gomes

Department of Statistics, Federal University of Bahia, Salvador, Bahia, Brazil

## ARTICLE INFO

### Article history:

Received 5 February 2022  
 Received in revised form 11 June 2022  
 Accepted 7 July 2022  
 Available online 15 July 2022

### Keywords:

MLP  
 ARIMA  
 Hybrid model  
 Covid-19  
 Forecasting

## ABSTRACT

The use of models to predict disease cases is common in epidemiology and related areas, in the context of Covid-19, both ARIMA and Neural Network models can be applied for purposes of optimized resource management, so the aim of this study is to capture the linear and non-linear structures of daily Covid-19 cases in the world by using a hybrid forecasting model. In summary, the proposed hybrid system methodology consists of two steps. In the first step, an ARIMA model is used to analyze the linear part of the problem. In the second step, a neural network model is developed to model the residuals of the ARIMA model, which would be the non-linear part of it. The neural network model was superior to the ARIMA when considering the capture of weekly seasonality and in two weeks, the combination of models with the capture of seasonality in two weeks provided a mixed model with good error metrics, that allows actions to be premeditated with greater certainty, such as increasing the number of nurses in a location, or the acceleration of vaccination campaigns to diminish a possible increase in the number of cases.

© 2022 Elsevier B.V. All rights reserved.

## 1. Introduction

Covid-19 (Sars-CoV-2), was already previously present in nature, however, until now due to a mutation in its Spike protein, the virus has exponentially increased its transmissibility [1], being spread all over the world. In general, the incubation period for the virus is 1 to 14 days. In addition to personal protective measures, such as social distancing and the use of masks and hand sanitizers, it is also recommended that the infected, or people that could possibly be infected, avoid contact with pets, due to the possibility of human–animal transmission, given as a suggestion that some animals such as cats may act like intermediaries for the virus [2].

Despite the symptomatic cases of Covid-19, there are also asymptomatic cases, which can also contaminate their surroundings, even though there are no symptoms, they also pose risks for those who come into contact with these individuals [3]. In studies to understand the historical behavior of the pandemic such as in [4] the historical process in which the World Health Organization (WHO) initially identified a cluster of pneumonia cases in China from December 31 in the city of Wuhan is minutiously detailed and it gets to the point from November 2020 to January 2021 when there was growth concerning the new variants of Sars-CoV-2.

The use of models to predict disease cases is very common in epidemiology and related areas, in the context of Covid-19, some works chose to investigate risk prediction deciding to work with Neural Network models and X-ray images in the field of computer vision like the works of [5,6], the main contribution of this paper is to discuss and present the use of hybrid modeling to predict the number of cases in order to help planning and resource optimization. [7] classified the scenarios of main application of AI in the Covid-19 pandemic in three main fields, where works that deal with computer vision and image classification are classified in the field of early detection and diagnosis of infection, the field of potential treatment, containing studies and applications that deal with the use of AI for predicting effective and safe drugs, and the field of outbreak and prediction of virus spread, with studies and applications of forecasting models, which is also the subject of this work.

In the context of Covid-19 both ARIMA and Neural Network models can be applied for purposes of optimized resource management, such as purchasing masks, ICU beds or to guide the adoption of public policies, however there may be limitations in applying only one of these models separately [8], while reviews on Covid-19 forecasting literature such as [9] suggests that studies addressing the development of hybrid approaches to forecast the pandemic should be proposed, the literature review of [10] revealed that studies based on the conventional machine learning models such as Support Vector Machine, Random Forest and Multilayer Perceptron Neural Networks showed poor performance while implemented standalone, even though the authors identify

<sup>\*</sup> Corresponding author.

E-mail address: [lucasrabeloufba@gmail.com](mailto:lucasrabeloufba@gmail.com) (L.R. de Araújo Morais).

these models as the mostly used in the pandemic, pointing that the performance appeared to be better for hybrid models coupled with optimization techniques.

Works about hybrid AI modeling in the context of the Sars-CoV-2 pandemic were also traced by reviews such as [11], that showed that this method was also used in studies of the field of computer vision with X-ray images, but also traces one study that used a hybrid machine learning model for time series forecasting with data from China, apart from that the authors also identify CNN, RF, ResNet and SVM approaches as the most used against Covid-19, and found that from 264 studies 84% used either computational tomography or X-ray images. In this context the aim of this study is to capture the linear and non-linear structures of daily Covid-19 cases in the world, using a classical approach proposed by [8], to evaluate the construction of a hybrid model for forecasting daily Covid-19 cases around the world and discussing its use to help the management of resources and action-taking in the healthcare sector.

## 2. Methodology

The database is publicly available and it is owned by the organization Our World In Data (OWID) [12], containing information on the number of daily cases in several countries around the world, the database is real-time updated and the information is collected from various government and press agencies and it was used to build the time series of daily Covid-19 cases in the world until April 25th 2021, the R software (Version 4.0.3) was used to build forecast models and database processing.

### 2.1. Forecasting models

A hybrid model based on [8], using a multilayer perceptron neural network (MLP) and an ARIMA model was constructed in order to make predictions for the daily number of Covid-19 infections in the world. ARIMA models and MLP Neural Network structures were initially adjusted, then the models with the best error metrics were used to build the hybrid model, the error metrics used to evaluate the models are the MASE (Mean Absolute Scaled Error), the SMAPE (Symmetric Mean Absolute Percentage Error) and the R2 (Coefficient Of Determination).

The MASE is considered by some authors the best error measure available for predictions [13], since it is less sensitive to outliers, when  $MASE < 1$ , on average, the errors are smaller than a simple naïve model. The SMAPE was used as an alternative to the MAPE (Mean Absolute Percentage Error), since problems that could arise from small values are less severe for the SMAPE. The coefficient of determination R2 close to 1 indicates an almost perfect relationship between the fitted model and the data, while an R2 close to 0 implies that just fitting the mean is similar to the fitted model [14].

### 2.2. ARIMA model

There is great utility in using the Arima model to guide disease management so that the effects of the intervention on the disease studied can be predicted. It is stated by [15] that “Clinical studies have benefited from the use of these techniques, given the great applicability of the ARIMA model”. In this work, we chose the Arima model to predict the number of daily cases of Covid-19 in the world in the year 2021, for such, since we dealt with real-time updated data, a model was fitted in 4 time windows of 7 days, the model that presented the best average MASE was considered as the best model for later prediction of cases. The study by [15] briefly describes the ARIMA model and also suggests that “since the periodic structures of infectious disease epidemics change

over time. In this case, it is suggested to use short periods of time segment to analyze the effects of each segment”. In [16] this technique is used, the paper also mentions a few studies that used time intervals between epidemics in Japan.

Although [8] summarizes the ARIMA model, the model is also detailed in other studies such as [17], where the forecasting process is mathematically described, however the Box–Jenkins methodology used to adjust the Arima(p,d,q) model and the autoregressive models ARMA, ARIMA, SARIMA and ARMAX, may be found with a more developed definition in studies like the one developed by [18], which also defines the autoregressive model of order p AR(P), the process of moving averages MA(q) of order q, the order d in which the process is stationary and SARIMA models (p,d,q)x(P,D,Q)[m] where m is the seasonal period of the model, the model parameters and the Box–Jenkins methodology can also be found briefly in [19].

$y'_t$  is the differenced series in the full ARIMA(p,d,q) model, which may be written as:

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (1)$$

For the analysis we used the `auto.arima()` function of R, its entire methodology, in addition to being available on R's CRAN, is well explained in chapter 8 in [20], for the seasonal part, for example, the work in chapter 2.1 demonstrates that for daily cases “If the frequency of observations is greater than once a week, there is usually more than one way to deal with the frequency. For example, data with daily observations can have a weekly seasonality (frequency = 7) or an annual seasonality (frequency = 365.25)”. In addition the `include.drift` argument, which was used in our practical part is also explained in chapter 8.7, the drift method is defined in chapter 3.1 and is a variation of the naïve method allowing forecasts to grow or decrease over time, the amount of change over time (drift) is defined as the average of the changes seen in historical data.

### 2.3. MLP neural networks

Neural networks are flexible computing structures to model a wide range of nonlinear problems, a significant advantage of neural network models over other nonlinear models is that neural networks are universal approximators that can approximate a large class of functions with a high degree of accuracy. It is further stated by [8] that the single hidden layer feedforward network is the most widely used form of modeling for forecasting time series, this affirmative is one of the reasons we used our neural network models with only one layer. In addition, the author of the `mlp()` function that we use in the R software to adjust the model, mentions in his work [21] that the most commonly used form of neural networks for prediction is the MLP neural network and that the one step ahead prediction is calculated using inputs that are lagged observations of the time series or other explanatory variables.

Mathematically the form of our  $\hat{y}_{t+1}$  forecasts are represented in Eq. (2), where  $I$  is the number of inputs  $p_i$  of the NN,  $H$  is the number of hidden nodes in the network,  $g(\cdot)$  is a non-linear transfer function,  $\beta_0$  and  $\gamma_{0i}$  are the biases of each neuron,  $w = (\beta, \gamma)$  are the network weights with  $\beta$  for the output and  $\gamma$  for the hidden layers:

$$\hat{y}_{t+1} = \beta_0 + \sum_{h=1}^H \beta_h g \left( \gamma_{0i} + \sum_{i=1}^I \gamma_{hi} p_i \right) \quad (2)$$

One of the concerns with the definition of neural network models was the structure of the model, as they would later be compared with Arima models and used to generate a mixed model, the inputs were built based on the statements given

by [22] that the number of neurons in the input layer indicates the maximum cycle size or seasonality that the model can detect, a model of six neurons in the input layer could therefore capture a period seasonality equal to or less than six, in this case the models with similar seasonality will be compared in ARIMA and MLP, in addition the order of differences of the MLP models that capture seasonality were built based on the D parameter of the SARIMA models, the best MLP will be used to build the mixed model.

#### 2.4. Hybrid MLP-ARIMA model

Proposed by [8] as a solution to the possibility that the approximation of ARIMA models to complex nonlinear problems does not get to be adequate enough and the use of neural networks when used to model linear problems can produce mixed results. In general, it is difficult to fully know the characteristics of the data in a real problem, the hybrid methodology proposed by Zhang has linear and non-linear modeling capabilities, it can be a good strategy for practical use, when combining different models, different aspects can be captured. In summary, the proposed hybrid system methodology consists of two steps. In the first step, an ARIMA model is used to analyze the linear part of the problem. In the second step, a neural network model is developed to model the residuals of the ARIMA model, which would be the non-linear part of it.

Since the ARIMA model cannot capture the nonlinear structure of the data, the residuals from the linear model will contain information about nonlinearity. The Neural network's results can be used as residual predictions for the ARIMA model. The hybrid model uses the unique capability and strength of the ARIMA Model and the neural network model in determining different patterns. Therefore, according to Zhang, it may be advantageous to model linear and non-linear patterns separately using different models and then combine the predictions to improve overall modeling and prediction performance.

Basically the predictions can be summarized with the following equation, where  $\hat{L}_t$  is the linear part and  $\hat{N}_t$  is the non-linear part of the forecast:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \quad (3)$$

The ARIMA and MLP models chosen were those with the lowest MASE, with predictions for the short, medium and long time lengths (respectively 7, 14 and 28 days) made considering the daily cases until April 25, 2021. The prediction of the linear part was made using the ARIMA model, the MLP neural network was trained on the residuals of the ARIMA model to predict the non-linear part of our time-series, finally our predicted variable would then be the linear part of the model plus the non-linear part, just like in Eq. (3).

### 3. Results

For the prediction models adjusted to the daily number of Covid-19 cases in the world (Table 1), the lesser computationally intensive model in all considered time periods was the ARIMA model both without and with seasonality being considered. Although the MLP(1,5,1) model without capturing the seasonal effect also required a higher level of computational power in comparison to the adjusted ARIMA, the MLP(7,5,1), MLP(14,5,1) and hybrid MLP(14,5,1)-Arima(2,0,4)(0,1,2)[14] were the most computationally intensive models.

When considering a short term forecast horizon (7 days) the best models were the ones which considered a seasonal effect of 14 days, the Arima(2,0,4)(0,1,2)[14] and the MLP(14,5,1) neural network had the best MASE and SMAPE error measures. In prediction horizon of medium (14 days) and long (28 days) term the

behavior was similar, the best models were the ARIMA and the MLP neural network with a seasonality of 14 days, even though in long time forecast horizons the error measure MASE was higher than 1 in both models by using the hybrid MLP-ARIMA model in these 28 days prediction horizon the error measures improved, and the MASE was way lower in short time term forecasts with the mixed model.

For medium term forecasts, despite the good performance of the Arima(2,0,4)(0,1,2)[14] and MLP(14,5,1) models in the error metrics, the combination of the models generated mixed results, there was not a good capture of the linear and non-linear part of the models resulting in a high value for the MASE statistic in the hybrid MLP(14,5,1)-Arima(2,0,4)(0,1,2)[14] model, the R2 measure had a negative value, while the SMAPE was one of the highest recorded in comparison to the other ARIMA and MLP fitted models, only inferior to the MLP(1,5,1) model without capturing seasonality.

### 4. Discussion

It is important to point out that database corrections can generate changes in error measures, there are corrections made to information prior to April 25th for example, this changed some error measures, some for less and some for more, which may reveal a limitation of the work, even though there is a slight degree of instability in all models because we are dealing with real-time updated data and already started with a great amount of data in the midst of the pandemic, which differs our work from some that also used from hybrid methodologies such as [23] that proposed a hybrid methodology using a polynomial neural network (PNN) with corrective feedback to forecast in early stage epidemic, but did not deal with constant updates in the data. In the database corrections throughout our work the ARIMA and MLP models calculated with the functions in R kept their structures, revealing that the changes in the database were not enough to change the structure of the models.

The SMAPE measure proved to be more resilient to forecast errors, hence the importance of using it with the support of MASE and the R2 coefficient, for the mixed model MLP(14,5,1)-Arima(2,0,4)(0,1,2)[14] in short-term forecasts the use of the three error metrics demonstrate that the MASE was adequate for being close to 0 and much below 1, being the model therefore better than a simple naïve, the R2 coefficient was close to 1, so the model fit was quite good, then we can go on to analyze the SMAPE, because despite being more resistant to errors, its value was low, even after analyzing the support error metrics.

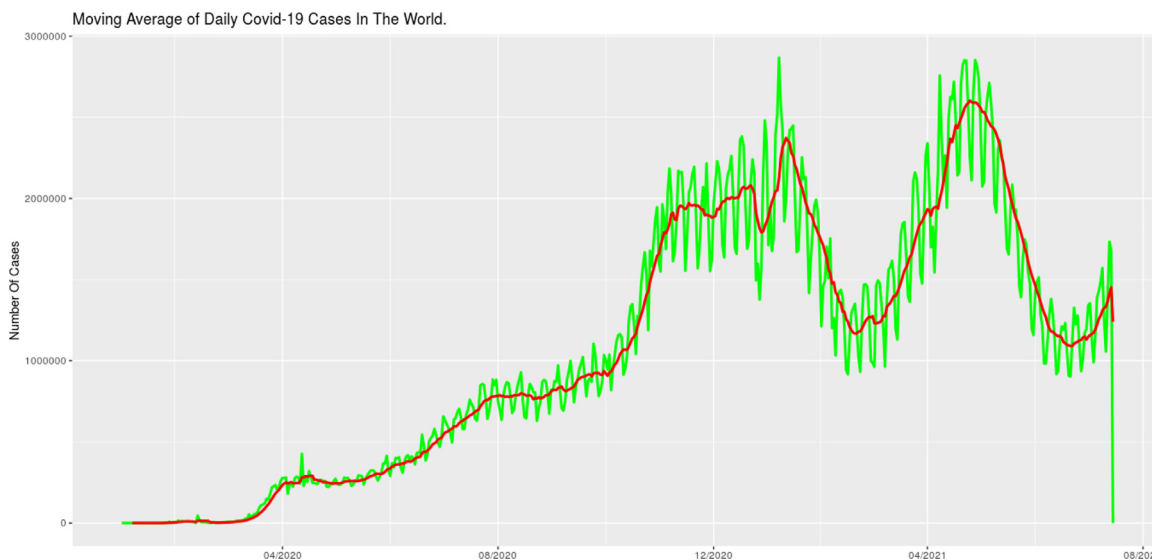
The role of prediction models in the management of Covid-19 would be to create a preparation to deal with the behavior of daily cases of the disease in a specific location, and to forecast the consequence of adopted measures, considering the world in (Fig. 1), with the use of models considering the use of each with its specificity or need, ARIMA could be used in cases where a low amount of computational power is needed, but even if its error metrics are inferior to neural network or mixed models, ARIMA allows for planning for small health units in the short and/or medium term.

Among the most computationally intensive models, the MLP(14,5,1) should be considered for planning actions that should be taken for the medium term planning, and the mixed for short and long term planning, in which the level of preparation to deal with an atypical situation in the health sector is higher, by capturing the two natures (linear and nonlinear) the mixed model allows actions to be premeditated with greater certainty, such as increasing the number of nurses in a location, or the acceleration of vaccination campaigns to diminish a possible increase in cases.

**Table 1**  
Error measures of the adjusted forecasting models.

Forecast horizon	Short term(7 days)			Medium term(14 days)			Long term(28 days)		
	MASE	SMAPE	R2	MASE	SMAPE	R2	MASE	SMAPE	R2
Arima(2,1,2) <sup>a</sup>	0.8147	0.0666	0.3655	0.7893	0.0663	0.5368	1.3533	0.1553	-0.1938
Arima(2,0,3)(0,1,2)[7] <sup>a</sup>	0.6033	0.0508	0.7426	0.6122	0.0506	0.7061	1.4004	0.1640	-0.2233
Arima(2,0,4)(0,1,2)[14] <sup>a</sup>	0.5441	0.0453	0.7113	0.5179	0.0439	0.7593	1.2644	0.1467	0.0401
MLP(1,5,1)	1.2615	0.1073	-0.1680	1.4714	0.1114	-1.0805	1.6767	0.1609	-1.0199
MLP(7,5,1)	0.4250	0.0355	0.8442	0.8180	0.0655	0.5554	1.5633	0.1887	-0.3784
MLP(14,5,1)	0.3584	0.0302	0.8851	0.3181	0.0267	0.9098	1.0262	0.1182	0.2815
MLP(14,5,1)Arima(2,0,4)(0,1,2)[14] <sup>a</sup>	0.2904	0.0236	0.9192	1.4191	0.1101	-0.3109	0.5865	0.0710	0.7596

<sup>a</sup>With the drift technique.



**Fig. 1.** Moving average of daily cases of Covid-19 in the world.

**5. Conclusion**

The ARIMA models proved to be less computationally intensive compared to neural networks and the hybrid model, when there is no capture of the seasonal effect, the ARIMA model was better than the MLP, both in terms of computational performance and error metrics, however, as expected, the neural network model was superior to the ARIMA when considering the capture of weekly seasonality and in two weeks, the combination of models with the capture of seasonality in two weeks provided a mixed model with good error metrics.

The use of models should therefore be considered to assist in decision-making and in short, medium and long-term planning, representing a greater use for the preparation than an objective forecast, mainly due to the dynamic update character of the data and its corrections like [24] that also cited the lack of accurate information in the beginning of the pandemic as a problem, the important thing is to know whether the trend in the coming days will be a growth or a decrease in the daily number of cases, to then guide decision making, in fact the models that obtained a MASE < 1 can be considered for this, being the MLP(14,5,1)-Arima(2,0,4)(0,1,2)[14] model a good option for short and long-term planning and the MLP(14,5,1) for medium-term planning.

It is also important to point out that there are a few limitations in this work, such as the heterogeneity which is characteristic of the data, since we dealt with data from all over the world, the paper also does not consider the impact caused by the great number of variants of Covid-19, therefore not capturing its effect, so the usability of the model built in this work may not be ideal,

but when used in certain controlled conditions this technique may bring up good results for longer forecast horizons.

While some state of art works conducted in the Sars-CoV-2 pandemic have not focused mainly in predicting the random component of Covid-19 data, but also dealt with real-time updated data like [25] that proposed a model that showed accuracy even after 3 weeks working with data, there are still plenty of possibilities to be explored in this field and future research should be conducted with the use of hybrid methodologies and support of previous studies that proposed deep learning or classical statistical models, to evaluate the impact of forecasting in controlled conditions.

**CRedit authorship contribution statement**

**Lucas Rabelo de Araújo Morais:** Methodology, Writing – original draft, Visualization, Investigation, Writing – review & editing.  
**Gecynalda Soares da Silva Gomes:** Supervision, Writing – review & editing.

**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.

**Acknowledgments**

This work was supported by The Brazilian National Council for Scientific and Technological Development(CNPQ) under Grant 144790/2020-3.

## References

- [1] S. Platto, Y. Wang, J. Zhou, E. Carafoli, History of the COVID-19 pandemic: Origin, explosion, worldwide spreading, *Biochem. Biophys. Res. Commun.* 538 (2021) 14–23, <http://dx.doi.org/10.1016/j.bbrc.2020.10.087>.
- [2] T. Kulkarni, P. Sharma, P. Pande, R. Agrawal, S. Rane, A. Mahajan, COVID-19: A review of protective measures, *Cancer Res. Statist. Treatment* 3 (2) (2020) 244–253, [http://dx.doi.org/10.4103/CRST.CRST\\_172\\_20](http://dx.doi.org/10.4103/CRST.CRST_172_20).
- [3] L. Wei, J. Lin, X. Duan, W. Huang, X. Lu, J. Zhou, Z. Zong, Asymptomatic COVID-19 patients can contaminate their surroundings: an environment sampling study, *MSphere* 5 (3) (2020) <http://dx.doi.org/10.1128/mSphere.00442-20>.
- [4] T. Carvalho, F. Krammer, A. Iwasaki, The first 12 months of COVID-19: a timeline of immunological insights, *Nat. Rev. Immunol.* 21 (2021) 245–256, <http://dx.doi.org/10.1038/s41577-021-00522-1>.
- [5] J. Prada, Y. Gala, A.L. Sierra, COVID-19 mortality risk prediction using X-Ray images, *Int. J. Interact. Multimedia Artif. Intell.* 6(Regular Issue) (2021) 7–14, <http://dx.doi.org/10.9781/ijimai.2021.04.001>.
- [6] M.I. Khattak, M. Al-Hasan, A. Jan, N. Saleem, E. Verdú, N. Khurshid, Automated detection of COVID-19 using chest X-Ray images and CT scans through multilayer-spatial convolutional neural networks, *Int. J. Interact. Multimedia Artif. Intell.* 6(Regular Issue) (2021) 15–24, <http://dx.doi.org/10.9781/ijimai.2021.04.002>.
- [7] M. Alsharif, Y. Alsharif, S. Chaudhry, M. Albreem, A. Jahid, E. Hwang, Artificial intelligence technology for diagnosing COVID-19 cases: a review of substantial issues, *Eur. Rev. Med. Pharmacol. Sci.* 24 (17) (2020) [http://dx.doi.org/10.26355/eurrev\\_202009\\_22875](http://dx.doi.org/10.26355/eurrev_202009_22875).
- [8] G.P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing* 50 (2003) [http://dx.doi.org/10.1016/S0925-2312\(01\)00702-0](http://dx.doi.org/10.1016/S0925-2312(01)00702-0).
- [9] I. Rahimi, F. Chen, A. Gandomi, A review on COVID-19 forecasting models, *Neural Comput. Appl.* (2021) <http://dx.doi.org/10.1007/s00521-020-05626-8>.
- [10] M. Khan, M.T. Mehran, Z.U. Haq, Z. Ullah, S.R. Naqvi, M. Ihsan, H. Abbas, Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review, *Expert Syst. Appl.* 185 (2021) <http://dx.doi.org/10.1016/j.eswa.2021.115695>.
- [11] O. Dogan, S. Tiwari, M.A. Jabbar, S. Guggari, A systematic review on AI/ML approaches against COVID-19 outbreak, *Complex Intell. Syst.* 7 (2021) 2655–2678, <http://dx.doi.org/10.1007/s40747-021-00424-8>.
- [12] H. Ritchie, E. Mathieu, L. Rodés-Guirao, C. Appel, C. Giattino, E. Ortiz-Ospina, J. Hasell, B. Macdonald, D. Beltekian, M. Roser, Coronavirus pandemic (COVID-19), in: *Our World in Data*, 2020, Retrieved from: <https://ourworldindata.org/coronavirus>.
- [13] R.J. Hyndman, A.B. Koehler, Another look at measures of forecast accuracy, *Int. J. Forecast.* 22 (4) (2006) 679–688, <http://dx.doi.org/10.1016/j.ijforecast.2006.03.001>.
- [14] L.J. Saunders, R.A. Russell, D.P. Crabb, The coefficient of determination: What determines a useful R2 statistic? *Invest. Ophthalmol. Vis. Sci.* 53 (11) (2012) 6830–6832, <http://dx.doi.org/10.1167/iovs.12-10598>.
- [15] R.C. SATO, Gerenciamento de doenças utilizando séries temporais com o modelo ARIMA, *Einstein (São Paulo)* 11 (1) (2013) 128–131, <http://dx.doi.org/10.1590/S1679-45082013000100024>.
- [16] A. Sumi, K. Kamo, N. Ohtomo, K. Mise, N. Kobayashi, Time series analysis of incidence data of influenza in Japan, *J. Epidemiol.* 21 (1) (2011) 21–29, <http://dx.doi.org/10.2188/jea.je20090162>.
- [17] D. Pissarenko, *Neural networks for financial time series prediction: Overview over recent research*, 2002.
- [18] A. dos Santos Gomes, F. Cribari Neto, *Modelagem e previsão da arrecadação do imposto de renda no Brasil*, 2003.
- [19] R. Adhikari, R. Agrawal, An introductory study on time series modeling and forecasting, 2013, <http://dx.doi.org/10.48550/arXiv.1302.6613>.
- [20] R. Hyndman, G. Athanasopoulos, *Forecasting: principles and practice*, (2) 2018, OTexts: Melbourne, Australia.
- [21] N. Kourentzes, D.K. Barrow, S.F. Crone, *Neural network ensemble operators for time series forecasting*, 2014, <http://dx.doi.org/10.1016/j.eswa.2013.12.011>.
- [22] J. Flores, *Comparação de modelos MLP/RNA e modelos Box-Jenkins em séries temporais não lineares*, 2009.
- [23] R. González-Crespo, E. Herrera-Viedma, N. Dey, S.J. Fong, G. Li, Finding an accurate early forecasting model from small dataset: A case of 2019-nCoV novel coronavirus outbreak, *Int. J. Interact. Multimedia Artif. Intell.* 6(Special Issue on Soft Computing) (2020) 132–140, <http://dx.doi.org/10.9781/ijimai.2020.02.002>.
- [24] M. Dur-e Ahmad, M. Imran, Transmission dynamics model of coronavirus COVID-19 for the outbreak in most affected countries of the world, *Int. J. Interact. Multimedia Artif. Intell.* 6(Regular Issue) (2020) 4, <http://dx.doi.org/10.9781/ijimai.2020.04.001>.
- [25] M. Gupta, R. Jain, S. Taneja, G. Chaudhary, M. Khari, E. Verdú, Real-time measurement of the uncertain epidemiological appearances of COVID-19 infections, *Appl. Soft Comput.* 101 (2021) <http://dx.doi.org/10.1016/j.asoc.2020.107039>.