



Cross-Validation Comparison of COVID-19 Forecast Models

Mintodê Nicodème Atchadé¹ · Yves Morel Sokadjo² · Aliou Djibril Moussa¹ · Svetlana Vladimirovna Kurisheva³ · Marina Vladimirovna Bochenina³

Received: 5 January 2021 / Accepted: 11 May 2021 / Published online: 26 May 2021
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd 2021

Abstract

Many papers have proposed forecasting models and some are accurate and others are not. Due to the debatable quality of collected data about COVID-19, this study aims to compare univariate time series models with cross-validation and different forecast periods to propose the best one. We used the data titled “Coronavirus Pandemic (COVID-19)” from “Our World in Data” about cases for the period of 31 December 2019 to 21 November 2020. The Mean Absolute Percentage Error (MAPE) is computed per model to make the choice of the best fit. Among the univariate models, Error Trend Season (ETS), Exponential smoothing with multiplicative error-trend, and ARIMA; we got that the best one is ETS with additive error-trend and no season. The findings revealed that with the ETS model, we need at least 100 days to have good forecasts with a MAPE threshold of 5%.

Keywords COVID-19 · Cross-validation · Forecast · Statistical modeling · Time series

Introduction

Originally from China in Wuhan city, COVID-19 has been declared as a world pandemic by WHO Emergency Committee on January 30th, 2020 [1]. Since that day, we are observing rapid increases [2–4] of cases that made authorities give daily government communication to update the information of the day before. In this context, one of the main issues is to know the future number of cases [5] with the least possible bias [6], to create adaptative tools [7], to determine the reproduction number [8–12], to check Recovery Time Period

[13], or other parameters to define appropriate policies to control COVID-19 spread. Thus, it is easy to notice that numerous papers [14–18] have proposed different forecast models of COVID-19 using univariate or multivariate time series modeling.

The multivariate time series models have advantages, because they can reveal the influence of many parameters such as face mask wearing, social distancing, hand washing, airport screening, quarantine, and treatment protocols on COVID-19 spread. However, as they require numerous parameters, the goodness of fit is affected when any parameter gets to be not correct. Besides, the quality of data related to COVID-19 is debatable, because many countries were obliged to subtract some cases that they reported to the World Health Organization (WHO) [19, 20]. Actually, those issues happened, because it has been noticed that some cases were reported without any respect of the official WHO technical guidance for laboratory testing [21]. Consequently, in data sets, we can get negative numbers and it is an indicator of not high-quality data sets. Recently, due to data quality doubts, the authors were obliged to increase 20 times the cases, 40 times the recovered patients to approach real figures [15]. Multivariate time series models also face data availability issues, because during confinement, the data collected were not exhaustive. Moreover, most of the compartmental models in epidemiology depend on estimated

Mintodê Nicodème Atchadé, Yves Morel Sokadjo, Aliou Djibril Moussa, Svetlana Vladimirovna Kurisheva and Marina Vladimirovna Bochenina contributed equally.

✉ Yves Morel Sokadjo
yvesmorel17@yahoo.fr

¹ National Higher School of Mathematics Genius and Modelization, National University of Sciences, Technologies, Engineering and Mathematics, Abomey, Republic of Benin

² University of Abomey-Calavi/International Chair in Mathematical Physics and Applications (ICMPA: UNESCO-Chair), Abomey-Calavi, Republic of Benin

³ Department of Statistics and Econometrics, Saint-Petersburg State University of Economics, Saint-Petersburg, Russia

inputs such as case fatality and case recovery that are very sensitive to data quality. In the paper [15], it is easy to notice that the change in data due to uncertainty quite influenced the estimates of the Susceptible-Infectious-Recovered-Dead models. This paper [22] proposed state-of-the-art prediction models and how to design complex models using also unexploited data. A recent review [23] also compared many models and came to a conclusion about the fact that artificial intelligence (AI) and deep learning [24] technics are advised. Additionally, in these studies [25, 26], the authors used Supervised Machine Learning Models including logistic regression, naive Bayes learning algorithms, and decision tree. Finally, the latter is found to be the best when we consider the accuracy. However, considering the fact that the more complex model, the more the need of data, and the lesser the quality of COVID-19 data sets in the world; we think that a good univariate model that can give better results will help to handle the bias in forecasts.

Concerning univariate time series, there are many works that use different methods. We have the computation of forecasts with an exponential smoothing family [16] that has appreciable forecast accuracy and can fit short series. However, the choice of multiplicative trend and error in the context of time series cross-validation is much debatable. Even the multiplicative trend model takes into account the asymmetric risk, we think that a robust model should take into account an additive trend, because short-term multiplicative trend is to be additive in the long run. In that same paper, the author used 90% prediction interval and we think that it can be improved. Another recent study [17] has compared time series models in predicting COVID-19 cases, but it was not exhaustive, because it just focused on Auto Regressive models (1–3) using Maximum Likelihood, Conditional Least Squares, and Unconditional Least Squares. Additionally, there is just a work in Nepal [27] about COVID-19 forecasts that uses ETS models, Auto Regressive Integrated Moving Average (ARIMA), and Susceptible Infectious-Recovered (SIR) model for the period of 23/01/2020-30/04/2020. They found finally that ARIMA is the best among the latter models. Even this work [18] also proposed an ARIMA; however, those last two studies did not use time series cross-validation to check how robust and stable ARIMA models are in the current context of COVID-19 data.

Taking into account the aforementioned issues about data quality related to COVID-19 cases, we hypothesized the use of univariate time series forecast, because it clearly spares us from great biased estimates. In addition, this work focuses on the world data and it means the more countries, the more bias in data collect. Using time series cross-validation with ETS models gives a variety of 30 models times the number of cross-validation, and it is an adaptive method of parameter estimation which is quite appropriate to the COVID-19 situation as a fast-changing one. ETS models do not require

stationarity and it varies the components to check the best model. Consequently, we hypothesized that ETS models might be the right choice to work with the world data. To test that assumption with updated data, this work compared several univariate time series models (with a cross-validation process) and proposed appropriate statistical models to predict the number of cases in the world. This paper will help international institutions to have a good model to forecast COVID-19 cases and to take sustainable decisions about the world economy and response to the pandemic. It can also be used in AI algorithms as this work [28] has proposed it. The following sections are organized in the following structure: the second section regroups the data description and methodology, and the third section contains the modeling and forecasting results, discussion, and conclusion.

Materials and Methods

Materials

As introduced, we focused on COVID-19 total cases from 2019-12-31 to 2020-11-21. They are time series, and we collected daily cumulated total cases in the dataset titled “Coronavirus Pandemic (COVID-19)” from “Our World in Data” [29]. The variable of interest is the count of laboratory-confirmed infections and they are indexed by their respective date. The data were accessed on 2020-11-21 and we will make the dataset available if requested. The accuracy and reliability of those numbers are linked to daily verification and change. Actually, in the context of this work, we divided the period and considered the first 30 days and after, iteratively, added 14 days to the last period to compare the real and predicted values.

Methods

The classical univariate models in this work are ARIMA, exponential smoothing model with multiplicative error and multiplicative trend components (ESM), and ETS. Each of them will be computed using cross-validation methods, and finally, we will select the most appropriate in terms of forecast errors using the mean absolute percentage error (MAPE).

ARIMA Modeling

ARIMA or autoregressive integrated moving average is a form of statistical modeling that uses time series data to either predict a future trend or to output latent information to understand how a variable of interest changes within a period. It has three parameters and the first one p is about the order of the Autoregressive (AR) model, the second one d concerns the level of

differentiating, and the third one q shows the Moving average order. The functional is Eq. (1):

$$(1 - \rho_1 B - \dots - \rho_p B^p) \nabla^d y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t. \tag{1}$$

With $\varepsilon_t \sim \text{white noise}(0, \sigma_\varepsilon^2)$, $By_t = y_{t-1}$ and $B^p = y_{t-p}$, $\nabla^d = (1 - B)^d$ is the differentiation parameter of order d ($d \geq 0$), (ρ_1, \dots, ρ_p) , $(\theta_1, \dots, \theta_q)$ are the coefficients and σ_ε^2 the residual variance to be estimated.

ETS Modeling

ETS model is set to capture different components (Error, Trend, Season) of a time series. Indeed, it makes short-term forecasts, which is appropriate in the case of strong dynamics. This model focuses on trend, seasonal components of different traits [30]. The possible combinations of the trend and season give the 15 following models in Table 1. Consequently, 30 different models are possible (15 with additive errors and 15 with multiplicative ones). In other words, in combination with the error that can be Additive or Multiplicative, the models in the Table 1 can be extended to 30 models in total. We can recall that the paper [16] used ETS (M, M, N) that is already included in Table 1. For instance, ETS (A, A, N) is defined by:

$$\begin{aligned} y_t &= l_{t-1} + b_{t-1} + \varepsilon_t \\ l_t &= l_{t-1} + b_{t-1} + \alpha \varepsilon_t \\ b_t &= b_{t-1} + \beta \varepsilon_t. \end{aligned} \tag{2}$$

Besides, with h , a step ahead forecast parameter, the particular case related to damped trend has as a recurrence form [31] the following equations:

$$\begin{aligned} \hat{y}_{t+h|t} &= l_t + (\phi + \phi^2 + \dots + \phi^h) b_t \\ l_t &= \alpha y_t + (1 - \alpha)(l_{t-1} + \phi b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)\phi b_{t-1}, \end{aligned} \tag{3}$$

where l_t and b_t are, respectively, the level and trend components at time t , ε_t the error term with the smoothing parameter $0 < \alpha < 1$ for level, and $0 < \beta < 1$ for trend, and

Table 1 Different models in ETS modeling

	Season		
Trend	N	A	M
N	N, N	N, A	N, M
A	A, N	A, A	A, M
AD	AD, N	AD, A	AD, M
M	M, N	M, A	M, M
MD	MD, N	MD, A	MD, M

N none, A additive, M multiplicative, D damped, AD additive damped, MD multiplicative damped

ϕ the damping parameter. The initial coefficients l_0 and b_0 and the latter smoothing parameters are obtained with the maximum-likelihood estimate (MLE) method [32].

Mean Absolute Percentage Error

The mean absolute percentage error (MAPE) is also called mean absolute percentage deviation (MAPD), and it is a statistic that quantifies whether a forecast is accurate or not. To make that estimate easier of interpretation, MAPE is set as a percentage of the errors and it is the formula (4):

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100; \tag{4}$$

such that n is the number of predicted values, y_t is the actual value at time t , and \hat{y}_t is the forecast. Actually, in the context of this work, the retained threshold of MAPE is 5%. MAPE will be used to check the best model in terms of forecast in our study.

Analysis Process

We used R software [33] (version: 4.0.0) for the whole work and the library “forecast” with the function ets() for ETS model, auto.arima for ARIMA model, and ets() with the parameters (model = “MMN”, damped = False) for the exponential smoothing model with multiplicative trend and error. The models use the algorithm of Hyndman and Khandakar [32] that combines unit root tests, minimisation of the Corrected Akaike Information Criterion (AICc), and Maximum Likelihood Estimate (MLE) to propose the best model. Besides, we have two main points in the analysis process before stepping in the final forecast. The first one is that we will select the appropriate model among the ETS, ARIMA, and Exponential smoothing (Multiplicative trend and error). Actually, for each training data set, we computed three training data sets (TDS) (1 week, 2 weeks, and 3 weeks) of forecasts to check after what level of short-term prediction is good enough to forecast COVID-19 cumulative cases. Every testing data set was increased of 2 weeks until the end of the data. It helps to check how many days we can trust the forecast values. To select the most appropriate model, we checked each model assumption and used the MAPE. The second point is, once the final model is selected, we used also the MAPE to propose the smallest data training and the period of forecasts to consider for good prediction. Generally, the analysis process in the work is in Fig. 1.

Results, discussion, and conclusion

Results

In the world, the number of new cases of COVID-19 is counted and put at the disposal of everyone by the WHO. Actually, the cumulative number of daily confirmed cases of COVID-19 on a given day is the sum of the new cases on that day and the total number of cases on the eve. It is illustrated in Fig. 2.

When we look at Fig. 2, we can understand that there was a flat part of the graph until around the mid of March and after, we noticed a rapid increase in the number of cases. To forecast this time series, we need a model with the least bias in terms of MAPE. Thus, we used different training and test data sets to check the stochastic assumptions and accuracy of the three models to forecast COVID-19 cumulative cases. The results are set in the Tables 2, 3, 4, and 5. Actually, in Tables 2, 3, and 4, values in bold are for $MAPE \leq 0.05$ and values in italic mean a problem of residuals autocorrelation (p -value < 0.05). Additionally, in Table 5, values in bold mean smallest MAPEs level for 1, 2 and, 3 weeks forecasts respectively. We also have in the Tables 2, 3, and 4 different MAPEs and the verification of each model assumptions. To choose the best model, we proceeded as follows:

- Point 1: The best fitting model can be selected with the MAPE in Tables 2, 3, and 4. We have three different periods of forecast, and for each one, we output descriptive statistics in the Table 5 to select the models with the smallest MAPE. In terms of range, whatever the period, we can notice that ETS models have the smallest maxima and mean of MAPEs. To sum up, the best model, considering the fitting, is ETS.
- Point 2 : About the assumptions for time series, residual autocorrelation p value > 0.05 is essential. In Tables 2, 3, and 4, we can count 6 autocorrelated issues for ARIMA model, 2 for ETS, and 17 for MMN. The residuals do not follow a normal distribution (p value $\simeq 0.00$), and this was predictable, because the dependent variable is the cumulated variable. Besides, we got stationary residuals with higher TDS and this result is also understandable, because with great sample size, the variability becomes stationary for the models. MMN and ARIMA models have most of their residuals that are homoscedastic (p value > 0.05), but ETS does not. Actually, the presence of heteroscedasticity in ETS is a point that shows the significant change among the daily numbers of cases about the pandemic. In the context of the current work, it is not important as the autocorrelation of residuals.

Table 2 MAPE of ARIMA with different training data sets

TDS	1 week MAPE	2 week MAPE	3 week MAPE	ARIMA	AC	Norm	Hetero	Stat
1-30	17.76	28.82	36.70	020	0.35	0.00	0.09	0.99
1-44	22.24	17.87	13.77	020	0.65	0.00	0.21	0.71
1-58	2.67	8.40	18.36	021	0.93	0.00	0.33	0.02
1-72	13.96	30.13	43.43	021	0.99	0.00	0.67	0.06
1-86	9.74	16.58	21.32	020	0.07	0.00	0.25	0.26
1-100	1.05	0.97	0.96	020	0.29	0.00	0.04	0.01
1-114	1.40	2.00	2.27	020	0.92	0.00	0.03	0.01
1-128	0.78	1.37	2.30	021	0.92	0.00	0.02	0.01
1-142	0.68	1.40	2.46	121	0.97	0.00	0.05	0.01
1-156	0.45	0.61	0.56	422	0.42	0.00	0.05	0.01
1-170	0.84	1.76	3.01	222	0.22	0.00	0.14	0.01
1-184	0.19	0.58	1.23	222	0.34	0.00	0.04	0.01
1-198	0.39	1.06	1.76	520	0.06	0.00	0.03	0.01
1-212	0.61	0.80	1.01	020	0.53	0.00	0.00	0.01
1-226	0.13	0.35	0.50	222	0.00	0.00	0.16	0.01
1-240	0.09	0.21	0.36	222	0.00	0.00	0.11	0.01
1-254	0.37	0.78	1.19	222	0.00	0.00	0.10	0.01
1-268	0.06	0.04	0.20	222	0.00	0.00	0.04	0.01
1-282	0.18	0.50	1.16	222	0.00	0.00	0.01	0.01
1-296	0.39	0.98	1.76	222	0.00	0.00	0.01	0.01

TDS training dataset, AC autocorrelation, Norm normality, Hetero heteroscedasticity, Stat stationarity, values in bold mean $MAPE \leq 0.05$ and values in italic mean a problem of residuals autocorrelation (p -value < 0.05)

Table 3 MAPE of the best ETS with different training data sets

TDS	1 week MAPE	2 week MAPE	3 week MAPE	ETS	AC	Norm	Hetero	Stat
1-30	13.99	24.68	32.62	M,A,N	0.17	0.00	0.62	0.10
1-44	21.90	17.36	13.13	M,A,N	0.16	0.00	0.77	0.10
1-58	3.23	9.30	19.41	M,A,N	0.15	0.00	0.50	0.03
1-72	12.30	28.07	41.41	M,A,N	0.14	0.00	0.36	0.04
1-86	9.12	15.75	20.37	M,A,N	0.09	0.00	0.28	0.01
1-100	1.24	1.24	1.29	M,A,N	0.10	0.00	0.23	0.01
1-114	1.40	2.00	2.27	A,A,N	0.92	0.00	0.03	0.01
1-128	0.78	1.37	2.30	A,A,N	0.90	0.00	0.02	0.01
1-142	0.63	1.29	2.30	A,A,N	0.82	0.00	0.03	0.01
1-156	0.75	1.44	2.51	A,A,N	0.95	0.00	0.02	0.01
1-170	0.71	1.48	2.60	A,A,N	0.93	0.00	0.01	0.01
1-184	0.43	0.96	1.72	A,A,N	0.51	0.00	0.00	0.01
1-198	0.39	0.94	1.56	A,A,N	0.97	0.00	0.00	0.01
1-212	0.61	0.80	1.01	A,A,N	0.53	0.00	0.00	0.01
1-226	0.21	0.17	0.17	A,A,N	0.07	0.00	0.00	0.01
1-240	0.23	0.36	0.52	A,A,N	0.06	0.00	0.00	0.01
1-254	0.61	1.10	1.58	A,A,N	<i>0.03</i>	0.00	0.00	0.01
1-268	0.24	0.38	0.68	A,A,N	0.05	0.00	0.00	0.01
1-282	0.18	0.45	1.05	A,A,N	0.05	0.00	0.00	0.01
1-296	0.62	1.29	2.16	A,A,N	<i>0.01</i>	0.00	0.00	0.01

TDS training dataset, AC autocorrelation, Norm normality, Hetero heteroscedasticity, Stat stationarity, values in bold mean MAPE ≤ 0.05 and values in italic mean a problem of residuals autocorrelation (p-value < 0.05)

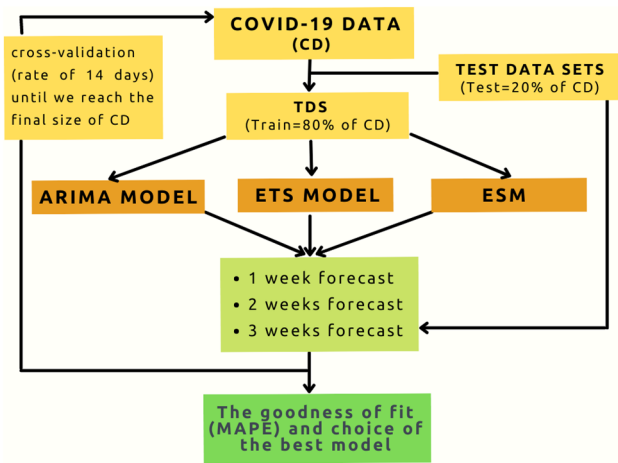


Fig. 1 Study analysis process

The rule of thumb is to keep the best model and considering the Point 1, Point 2, and the results in Table 5, we think that using ETS (A,A,N) will help to forecast with the least bias (realistic forecasts). Now, as we have chosen the best model, we can now visualize how it works and how to use it for a good forecast. Let us note that in Fig. 3; we should normally have 20 images (the number of TDS), but in terms of commodities (numerous images), we decided to show the beginning A (from 2020-01-21 to 2020-02-19), B (from 2020-01-21 to 2020-05-13), C (from 2020-01-21

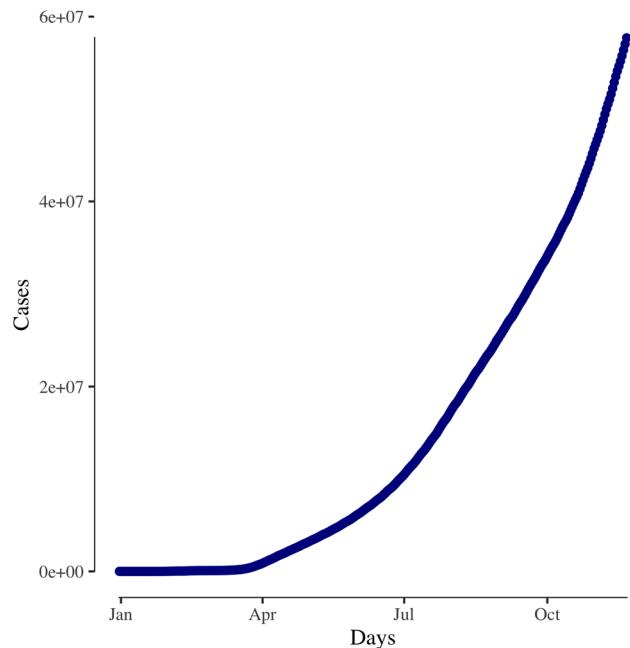


Fig. 2 Cumulative number of daily confirmed cases of COVID-19

to 2020-08-19), and the last (from 2020-01-21 to 2020-11-11). It helps to visualize the good fitting of our best model. When we look at the Fig. 3, first, we cannot trust

Table 4 MAPE of ESM with different training data sets

TDS	1 week MAPE	2 week MAPE	3 week MAPE	MMN	AC	Norm	Hetero	Stat
1-30	61.09	292.57	1241.40	M,M,N	0.05	0.00	0.64	0.07
1-44	20.02	13.31	16.42	M,M,N	0.05	0.00	0.74	0.01
1-58	3.18	9.05	18.95	M,M,N	0.05	0.00	0.48	0.01
1-72	10.66	24.56	36.58	M,M,N	<i>0.01</i>	0.00	0.33	0.01
1-86	2.14	10.29	29.51	M,M,N	<i>0.01</i>	0.00	0.26	0.01
1-100	2.12	8.12	17.22	M,M,N	<i>0.00</i>	0.00	0.21	0.01
1-114	2.35	5.47	9.54	M,M,N	<i>0.00</i>	0.00	0.18	0.01
1-128	0.38	0.58	1.15	M,M,N	<i>0.00</i>	0.00	0.16	0.01
1-142	0.31	0.36	0.62	M,M,N	<i>0.00</i>	0.00	0.15	0.01
1-156	0.59	0.69	0.82	M,M,N	<i>0.00</i>	0.00	0.14	0.01
1-170	0.61	0.81	1.04	M,M,N	<i>0.00</i>	0.00	0.13	0.01
1-184	0.24	0.39	0.76	M,M,N	<i>0.00</i>	0.00	0.12	0.01
1-198	0.24	0.27	0.39	M,M,N	<i>0.00</i>	0.00	0.11	0.01
1-212	0.55	0.44	0.61	M,M,N	<i>0.00</i>	0.00	0.11	0.01
1-226	0.27	0.47	0.96	M,M,N	<i>0.00</i>	0.00	0.10	0.01
1-240	0.33	0.34	0.26	M,M,N	<i>0.00</i>	0.00	0.10	0.01
1-254	0.68	1.10	1.40	M,M,N	<i>0.00</i>	0.00	0.10	0.01
1-268	0.21	0.19	0.23	M,M,N	<i>0.00</i>	0.00	0.09	0.01
1-282	0.27	0.45	0.87	M,M,N	<i>0.00</i>	0.00	0.10	0.01
1-296	0.60	1.11	1.69	M,M,N	<i>0.00</i>	0.00	0.09	0.01

TDS training dataset, AC autocorrelation, Norm normality, Hetero heteroscedasticity, Stat stationarity, values in bold mean MAPE ≤ 0.05 and values in italic mean a problem of residuals autocorrelation (p-value < 0.05)

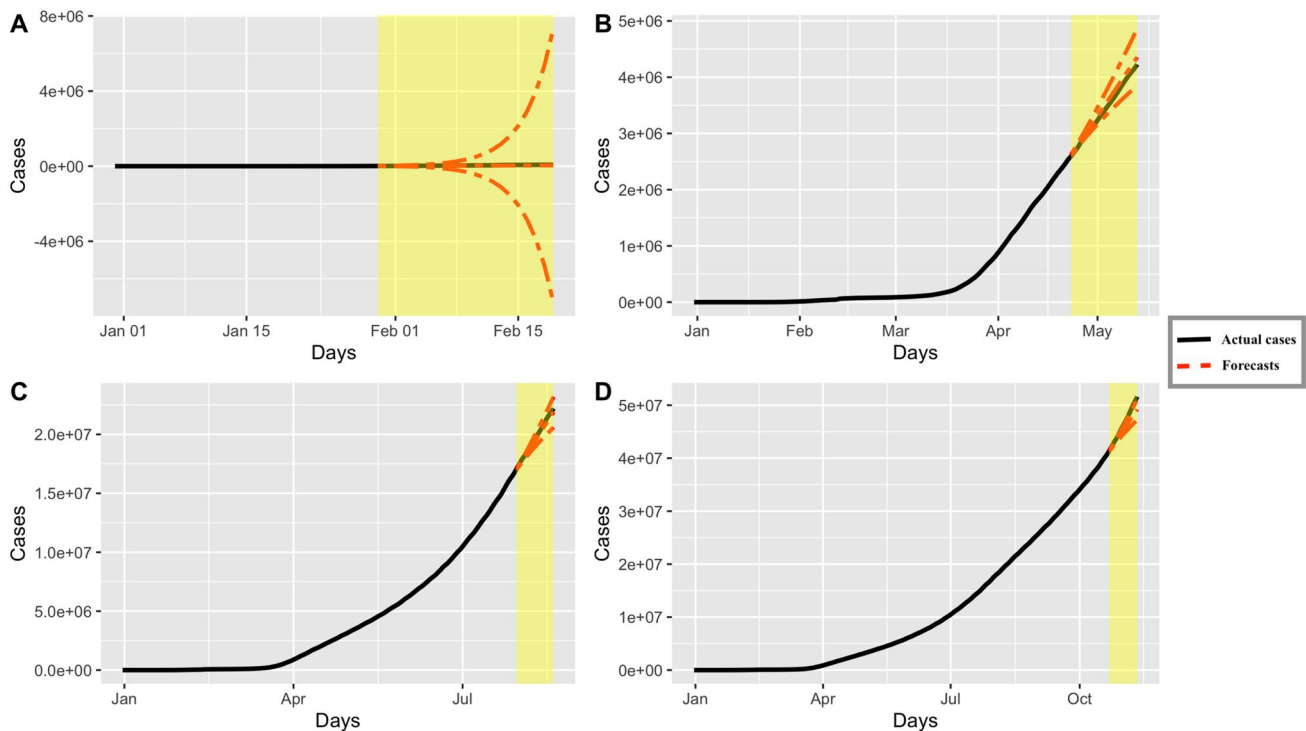


Fig. 3 Different forecasts of the kept ETS (A,A,N) model using each training data set

the forecasts in Fig A, because the real data (black line) are in a very large confidence interval. The uncertainty bounds are very big due to the small sample size.

Considering the MAPE smaller than 5% in Table 3, we can focus on TDS having at least 100 days with 1 or 2 or 3 weeks as periods of forecast (PF). In the case of this paper, we have 327 days in total and the aforementioned checking makes us compute the final model estimates and the other models with 2 week forecasts. We finally kept this, because in many works, the latency period is around 14 days [34–36] and the cross-validation rate was also 2 weeks. Although we recommend to keep 2 weeks for a forecast and repeat it if needed, people can still consider 3 weeks, because it also gave good results after 100 days of TDS. The model outputs are in Fig. 4. First, ESM has very big uncertainty bounds and it explains how bad were its MAPEs in the computation of the cross-validation. Second, we have ARIMA and ETS that got similar graphs and it was also noticeable in Table 5. However, ARIMA was not better than ETS, because it had higher MAPE. The actual and fitted values are quite similar, and this is why, it is even difficult to get the difference between them in Fig. 4.

Table 5 Mimima, means, and maxima of the MAPEs

Model	Min	Mean	Max	Error
ARIMA	0.06	3.70	22.24	1 week
ETS	0.18	3.48	21.90	1 week
MMN	0.21	5.34	61.09	1 week
ARIMA	0.04	5.76	30.13	2 weeks
ETS	0.17	5.52	28.07	2 weeks
MMN	0.19	18.53	292.57	2 weeks
ARIMA	0.20	7.72	43.43	3 weeks
ETS	0.17	7.53	41.41	3 weeks
MMN	0.23	69.02	1241.40	3 weeks

Values in bold mean smallest MAPEs level for 1, 2 and, 3 weeks forecasts respectively

Discussion

Daily decisions on COVID-19 have been influencing the spread of the pandemic and an adapted forecasting tool is required for better policy. Many studies have been proposing multi- or univariate models to forecast COVID-19 cases, but most of them failed to predict well the upcoming situations [37]. This study checked among the propositions, the one that is the most appropriate concerning daily realities about COVID-19 in the world. The quality of collected data about

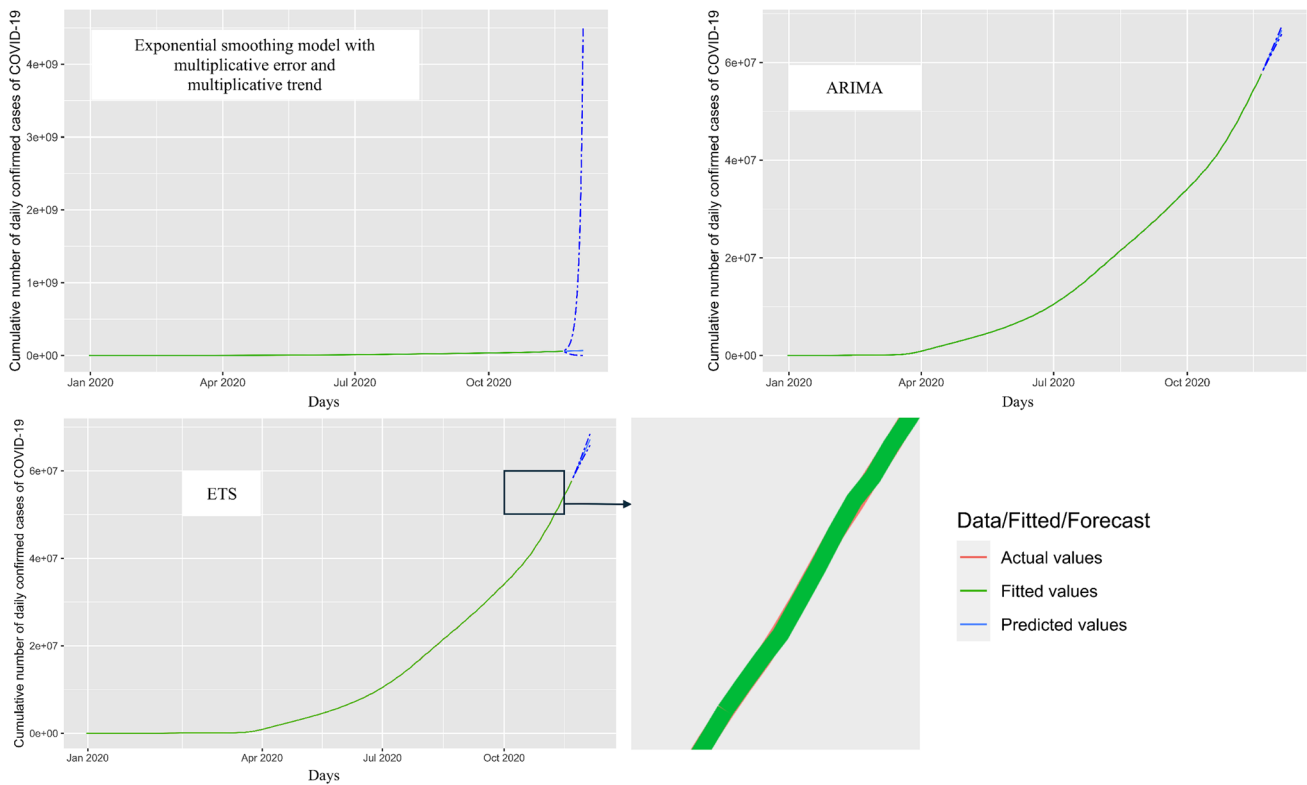


Fig. 4 Best model ETS (A,A,N) actual, fitted, and predicted values

the pandemic is still debatable for some countries [15, 19, 20]; and they are not also exhaustive due to confinement during outbreak periods. To reduce at most as possible the bias in our model, we proposed to avoid complex models due to the issues with current data and focused on univariate time series modeling. Considering the principle of “garbage-in, garbage-out”, using one time series analysis and having good forecasts is advisable.

Although we hypothesized that ETS model might be well adapted because of its capacity to vary in 30 different models and as much adaptive as possible in terms of COVID-19 evolution, we compared the classical univariate time series models. Among them, the best one is ETS model, because it respects the residuals autocorrelation assumption and has the smallest MAPE in Table 5. The study in Nepal [27] used 99 days and found ARIMA (MAPE 4.18%) and ETS (M,A,N) (MAPE 4.55%) for 2 week forecast. In our case, the MAPE (with 100 days) of ARIMA is 0.97 and ETS (A,A,N) is 1.24, and we can notice that the trend estimate for Nepal is also additive and the difference is about the error that is multiplicative and it can be related to the fact that we are working with the world data. This current study and the one in Nepal are similar in terms of trend type, while the study [16] with at most 50 days of training data set proposed a multiplicative trend. In Table 3, it is easy to notice that we only got an additive trend and this was our hypothesis. In addition, we used a cross-validation technique, and in their work [27], they did not. This point might explain their finding, because our process is more robust than theirs.

Actually, short-term forecasts (at most 2 weeks) are globally advised to maintain short-term forecasts, because their MAPEs are smaller than 5%. Even in the papers about COVID-19 forecast [16, 27, 38, 39], the authors proposed short-term (10, 14) days. Although we advise to keep 2 weeks for a forecast and repeat it if needed, people can still consider 3 weeks, because it also gave good results after 100 days of TDS. Many works [40–42] are interested in long forecasts such as the end period of the peak of COVID-19. The best forecasts in our model are from the training data having at least 100 days and it is understandable when you look at the Fig. 3. Actually, it has two parts, one that is flat and another one that shows a high increase of cases. Especially for every TDS (1–72), there are high MAPEs due to the fact that the train data set is at the transition part of the change between the flat part and the high increase part. The forecast model should take into account both parts, because from 23/03/2020, new cases started duplicating (becoming additive with time) compared to the past number of cases (26,069 new cases on 21/03/2020 [43] and 40788 on 23/03/2020 [43]) and even until 13/05/2020 to have good forecasts (small MAPEs).

When we assume that future decisions will follow the past structure, we think that the best model to forecast COVID

cumulative cases in the world is an ETS with additive error and trend without any season. The main limitation of this work is quite related to the predictions of the world figures about COVID-19, because they are aggregated and heterogeneous data sets. These remarks have also been mentioned in [37], because this kind of work does not take into account particular changes in small or big countries. However, the final decision with short-term forecasts can help in the decrease of bias in this study and allow international institutions to adjust decisions about the pandemic.

Conclusion

Many researchers have been computing classical univariate time series models to forecast COVID-19 cases. However, Error Trend Season (ETS) is let, while we can have 30 models in it to handle better short-term forecast of the pandemic cases. The use of cross-validation techniques and Mean Absolute Percentage Error (MAPE) to compare those models (ARIMA, ETS, and Exponential smoothing with multiplicative error-trend) helps to come out with ETS as the best model with the smallest MAPE and number of residuals correlations for all TDS. Actually, we advise other studies to also think of ETS models that are flexible for short-term forecasts and it provides realistic results. To get robust outputs, we propose to have at least a data set of 100 observations to expect good estimates.

Acknowledgements The authors are grateful to the lecturers at the International Chair in Mathematical Physics and Applications (ICMPA - UNESCO Chair), who were quite helpful in the improvement of this paper. Furthermore, we thank Mr. Innocent Koovi, an independent researcher, who helped to proof-read this paper.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. World Health Organization, et al. Statement on the second meeting of the International Health Regulations. Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV) (internet). 2020.
2. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science*. 2020;368(6492):742–6.
3. Sokadjo YM, Atchadé MN. The influence of passenger air traffic on the spread of COVID-19 in the world. *Transp Res Interdiscip Perspect*. 2020;8:100213.
4. Lau H, Khosrawipour V, Kocbach P, Mikolajczyk A, Ichii H, Schubert J, et al. Internationally lost COVID-19 cases. *J Microbiol Immunol Infect*. 2020.
5. Massonnaud C, Roux J, Crépey P. COVID-19: forecasting short term hospital needs in France. medrxiv. 2020.

6. Shinde GR, Kalamkar AB, Mahalle PN, Dey N, Chaki J, Hassanien AE. Forecasting models for coronavirus disease (COVID-19): a survey of the state-of-the-art. *SN Comput Sci.* 2020;1(4):1–15.
7. Rajasekar SJS. An enhanced IoT based tracing and tracking model for COVID-19 cases. *SN Comput Sci.* 2021;2(1):1–4.
8. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ. Presymptomatic transmission of SARS-CoV-2-Singapore, January 23-March 16, 2020. *Morbidity Mortality Wkly Rep.* 2020;69(14):411.
9. Luo L, Liu D, Liao XI, Wu Xb, Jing QI, Zheng Jz, et al. Modes of contact and risk of transmission in COVID-19 among close contacts. *medRxiv.* 2020.
10. Patrikar S, Kotwal A, Bhatti V, Banerjee A, Chatterjee K, Kunte R, et al. Incubation period and reproduction number for novel coronavirus (COVID-19) infections in India. *MedRxiv.* 2020.
11. Khailaie S, Mitra T, Bandyopadhyay A, Schips M, Mascheroni P, Vanella P, et al. Development of the reproduction number from coronavirus SARS-CoV-2 case data in Germany and implications for political measures. *medRxiv.* 2020.
12. Riou J, Althaus CL. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance.* 2020;25(4):2000058.
13. Bhapkar H, Mahalle PN, Dey N, Santosh K. Revisited COVID-19 mortality and recovery rates: are we missing recovery time period? *J Med Syst.* 2020;44(12):1–5.
14. Wynants L, Van Calster B, Bonten MM, Collins GS, Debray TP, De Vos M, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ.* 2020;369.
15. Anastassopoulou C, Russo L, Tsakris A, Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PloS one.* 2020;15(3):e0230405.
16. Petropoulos F, Makridakis S. Forecasting the novel coronavirus COVID-19. *PloS one.* 2020;15(3):e0231236.
17. Koçak M. A comparison of time-series models in predicting COVID-19 cases. *Turk Klinikleri J Biostat.* 2020;12(1).
18. Benvenuto D, Giovanetti M, Vassallo L, Angeletti S, Ciccozzi M. Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in brief.* 2020; 105340.
19. World Health Organization. Coronavirus disease (COVID-2019) situation reports. 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
20. Coronavirus (COVID-19) Data Hub. COVID-19 activity.csv; 2020. <https://data.world/covid-19-data-resource-hub/covid-19-case-counts/workspace/file?filename=COVID-19+Activity.csv>.
21. World Health Organization. Global surveillance for human infection with coronavirus disease (COVID-19). 2020. [https://www.who.int/publications-detail/global-surveillance-for-human-infection-with-novel-coronavirus-\(2019-ncov\)](https://www.who.int/publications-detail/global-surveillance-for-human-infection-with-novel-coronavirus-(2019-ncov)).
22. Santosh K. COVID-19 prediction models and unexploited data. *J Med Syst.* 2020;44(9):1–4.
23. Rahimi I, Chen F, Gandomi AH. A review on COVID-19 forecasting models. *Neural Comput Appl.* 2021; 1–11.
24. Asraf A, Islam MZ, Haque MR, Islam MM. Deep learning applications to combat novel coronavirus (COVID-19) pandemic. *SN Comput Sci.* 2020;1(6):1–7.
25. Muhammad L, Islam MM, Usman SS, Ayon SI. Predictive data mining models for novel coronavirus (COVID-19) infected patients' recovery. *SN Comput Sci.* 2020;1(4):1–7.
26. Muhammad L, Algehyne EA, Usman SS, Ahmad A, Chakraborty C, Mohammed I. Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. *SN Comput Sci.* 2021;2(1):1–13.
27. Bhandary S, Shrestha SL, Khatriwada RP, Shah DN, Munankarmi NN, Banjara MR, et al. Trend analysis, modelling and impact assessment of COVID-19 in Nepal. *medRxiv.* 2020.
28. Santosh K. AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on multi-tudinal/multimodal data. *J Med Syst.* 2020;44(5):1–5.
29. Max Roser EOO, Ritchie H, Hasell J. Coronavirus pandemic (COVID-19). *Our world in data.* 2020.
30. Jofipasi CA, et al. Selection for the best ETS (error, trend, seasonal) model to forecast weather in the Aceh Besar District. In *IOP conference series: materials science and engineering.* vol 352. IOP Publishing; 2018. p. 012055.
31. Gardner ES Jr, McKenzie E. Forecasting trends in time series. *Manag Sci.* 1985;31(10):1237–46.
32. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. *J Stat Softw.* 2008;26(3):1–22.
33. RStudio Team. RStudio: integrated development environment for R. 2020. <http://www.rstudio.com/>.
34. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann Intern Med.* 2020;172(9):577–82.
35. Jin YH, Cai L, Cheng ZS, Cheng H, Deng T, Fan YP, et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Military Med Res.* 2020;7(1):4.
36. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung SM, et al. Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. *J Clin Med.* 2020;9(2):538.
37. Jewell NP, Lewnard JA, Jewell BL. Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *JAMA.* 2020;323(19):1893–4.
38. Roosa K, Lee Y, Luo R, Kirpich A, Rothenberg R, Hyman JM, et al. Short-term forecasts of the COVID-19 epidemic in Guangdong and Zhejiang, China: February 13–23, 2020. *J Clin Med.* 2020;9(2):596.
39. Al-qaness MA, Ewees AA, Fan H, Abd El Aziz M. Optimization method for forecasting confirmed cases of COVID-19 in China. *J Clin Med.* 2020;9(3):674.
40. COVID I, Murray CJ, et al. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. *MedRxiv.* 2020.
41. Kuniya T. Prediction of the epidemic peak of coronavirus disease in Japan, 2020. *J Clin Med.* 2020;9(3):789.
42. Batista M. Estimation of the final size of the COVID-19 epidemic. *MedRxiv.* 2020;10:16–20023606.
43. (WHO) WHO, et al. Coronavirus disease (COVID-2019) situation report-62, 22 March 2020 (2020).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.