Mathematical Sciences Letters
An International Journal

http://dx.doi.org/10.18576/msl/110201

Using ETS State Space Model for forecasting on third wave on COVID19 in India

Khder Mohammed Alakkari[1], Pradeep Mishra[2,], Deepa Rawat[3], Mostafa Abotaleb[4], Abdullah Mohammad Ghazi Al khatib[5], Monika Devi[6] and Immad A Shah[7]*

[1]Department of Statistics and Programming, Faculty of Economics, Tishreen University, Latakkia, Syria
[2]College of Agriculture, Powarkheda,J.N.K.V.V. (M.P.), India
[3]College of Forestry, VCSG Uttarakhand University of Horticulture and forestry, Ranichauri, TehriGarhwal, Uttarakhand, India
[4]Department of System Programming, South Ural State University, 454080 Chelyabinsk, Russia
[5]Department of Banking and Insurance, Faculty of Economics, Damascus University, Damascus, Syria
[6]Department of Mathematics and Statistics, Chaudhary Charan Singh Haryana Agricultural University, Hisar-125 004, Haryana, India
[7]Division of Agricultural Statistics, SKUAST-K, India

Received: 21 Feb. 2022, Revised: 22 Mar. 2022, Accepted: 24 Mar. 2022.
Published online: 1 May 2022.

Abstract: Every day, there are new cases of COVID19 infections around the world, and a new strain of this virus known as Omicron is spreading rapidly. To contain and monitor gatherings, as well as provide assistance in areas where this infection is spreading rapidly, it is critical to identify the number of new cases of injuries and deaths. Forecasting new cases and deaths in India between the dates 9/1/2022 and 18/1/2022 was the goal of this study. Based on state-space likelihood calculations, the Error Trend Seasonal (ETS) model was used in the context of a dynamic nonlinear framework for model selection and forecast standard errors. Small sample size and non-normal distribution of data, non-stationarity, trend, and seasonality are all characteristics of these models that have never been used with data from India before. Predictions using the best-fit model from "9/1/2022" to "18/1/2022" indicate that new cases in India are expected to reach (209,275). With a growth rate of 100.43 percent, new deaths are expected to reach (369), 14.5%. The predicted values of daily new vaccination cases are expected to reach (7341793) with a growth rate of 0%. The Total cases is expected to reach (37044777) cases with a growth rate of 3.8%. The total number of deaths is expected to reach (486771) with a 0.56 percent growth rate. As the pandemic continues to spread, these estimates are essential for determining how many drugs and oxygen tanks will be needed as well as the number of ICU beds that will be needed. As a result, the government will benefit from this forecast, leading to the authorities taking the necessary precautions against this new wave of infection.

Keywords: COVID19, ETS, Forecasting, Omicron, vaccination.

1 Introduction

The coronavirus is part of a broad family of viruses that have been linked to common colds, coughs, and respiratory disorders [1]. Wuhan (Hebei Province, China) was the site of the first documented detection of the deadly new coronavirus (COVID-19). One of the family members of the respiratory pathogen responsible for the 2002–2003 outbreak (SARS-CoV-1) is the new coronavirus, also known as SARS-CoV-2 (SARS-CoV-2) [2]. However, the COVID-19 pandemic was far more widespread than the SARS-CoV-1 outbreak of 2002-2003. Despite solid preventive steps many countries took to limit the virus's transmission after its initial breakout in Wuhan, the COVID-19 virus continued to spread throughout China and fast around the world.

The first outbreaks of COVID-19 were reported from the Indian subcontinent at least nine months later. During this period, India recorded approximately 11 million cases and 0.157 million deaths, which peaked in September 2020. However, the second wave, which began in February 2021 and lasted until March 2022, was much worse than the first wave. The disease had a much more ferocious spread. SARS-CoV-2 B.1.1.7 (Alpha variant) and B.1.617.2, two highly contagious variants, emerged during the second wave, and this was a major factor in its spread (Delta variant). Several cases involving the B.1.617.2 have come to light recently [3]. Thanks to widespread vaccination, the pandemic curve in India began to flatten after October 2020.

*Corresponding author E-mail: pradeepjnkvv@gmail.com

They have estimated 1.5 million new COVID cases since January 2022 and 327 deaths. The WHO received its first report of an Omicron variant (B.1.1.529) in South Africa on November 24, 2021. Since the B.1.1.529 strain was discovered, the number of infections has risen dramatically. The World Health Organization has identified a large number of mutations in this variation that have the potential to be harmful. The Omicron form of Covid-19 is suspected to be the cause of the recent spike in cases in India, raising concerns about the possibility of a third wave of the disease. There must be accurate data on COVID19 infections and the spread of a new strain in order to help the government formulate rules for restricting and monitoring public gatherings and providing assistance in areas where the infection is rapidly spreading. This data is critical.
In India, ARIMA and SARIMA models have been used to predict the epidemiological pattern of prevalence. New and total deaths from COVID-19's occurrence [1]. [4] formulated a simple mathematical model with a flexible transmission rate and used iterated filtering to fit the model to deaths data. Bangladesh used it to estimate the expected daily number of COVID-19 cases [5]. In Italy, a model-based method was introduced to forecast the total number of currently COVID-19 positive individuals and the occupancy of the available intensive care units [6]. It was used in Ghana to predict the number of deaths as well as the number of cases [7]. [8] used BATS, TBATS, Holt's linear trend, and ARIMA models to track the spread of COVID 19 infection cases. ARIMA models and fuzzy time series were the focus of [9]'s research. To monitor the spread of the coronavirus, three African countries, Algeria, Egypt, and South Africa are using these methods. The error Trend Seasonal (ETS) model is used to forecast (total cases, new cases, total deaths, new deaths, and new vaccinations) for India in a dynamic nonlinear model framework with state-space-based likelihood calculations, including support for model selection and the calculation of forecast standard errors. According to our previous discussion, this model outperforms other models tested on data from India in terms of predictive ability with a small sample size, non-normal data distribution, trend, seasonality, and other things [10].

2 Material and Methods:

2.1 ETS Model:

Nonlinear dynamic models can be constructed using ETS models, which combine exponential smoothing models (ETS). With the use of state-space likelihood estimates and forecast standard errors, these models are evaluated [11]

(T), seasonal (S) and error (E) are the three main components of the time series that the model studies (E). The long-term trend of the time series is represented by the trend term, while the error term shows the unpredictable time series component. The seasonal term corresponds to a pattern with known periodicity. The components we need are combined in our model in various additive and multiplicative combinations to produce y_t. We have an additive model y_t = T+E+S or multiplicative model like y_t = T. E.S where the individual components of the model are given as follows:

$$E\ [A, M]$$

$$T\ [N, A, M, AD, MD]$$

$$S\ [N, A, M]$$

AD dampened (AD), and MD (MD) are additive dampened (AD) and multiplicative dampened (MD) (damping uses an additional parameter to reduce the impacts of the trend over time). The models we want to estimate are listed in the table below:

Table 1: state-space equations for each of the models in ETS.

Trend	**Additive Error Models**	**Trend**	**Multiplicative Error Models**
N	$y_t = l_{t-1} + \varepsilon_t$ $l_t = l_{t-1} + \alpha\varepsilon_t$	**N**	$y_t = l_{t-1}(1 + \varepsilon_t)$ $l_t = l_{t-1}(1 + \alpha\varepsilon_t)$
A	$y_t = l_{t-1} + b_{t-1} + \varepsilon_t$ $l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	**M**	$y_t = (l_{t-1} + b_{t-1})(1 + \varepsilon_t)$ $l_t = (l_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(l_{t-1} + b_{t-1})\varepsilon_t$
AD	$y_t = l_{t-1} + \phi b_{t-q} + \beta\varepsilon_t$ $l_t = l_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	**MD**	$y_t = (l_{t-1} + \phi b_{t-1})(1 + \varepsilon_t)$ $l_t = (l_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(l_{t-1} + \phi b_{t-1})\varepsilon_t$
Seasonal	**Additive Error Models**	**Seasonal**	**Multiplicative Error Models**
N	$y_t = l_{t-1}s_{t-m} + \varepsilon_t$ $l_t = l_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/l_{t-1}$	**N**	$y_t = l_{t-1}S_{t-m}(1 + \varepsilon_t)$ $l_t = l_{t-1}(1 + \alpha\varepsilon_t)$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A	$y_t = (l_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t/S_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/l_{t-1}$	**A**	$y_t = (l_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $l_t = (l_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(l_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
AD	$y_t = (l_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $l_t = l_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(l_{t-1} + \phi b_{t-1})$	**MD**	$y_t = (l_{t-1} + \phi b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $l_t = (l_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(l_{t-1} + \phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$

Where parameters: α: smoothing factor for the level, β: smoothing factor for the trend,γ:smoothing factor for the seasonal, ϕ: damping coefficient. And initial states: l: initial level components, b: initial growth components,s: initial seasonal components, which is estimated as part of the optimization problem.ETS model, it is possible to compare a likelihood-based information criterion (AIC):

$$-2log\ L(\hat{\theta}) + 2k \qquad (1)$$

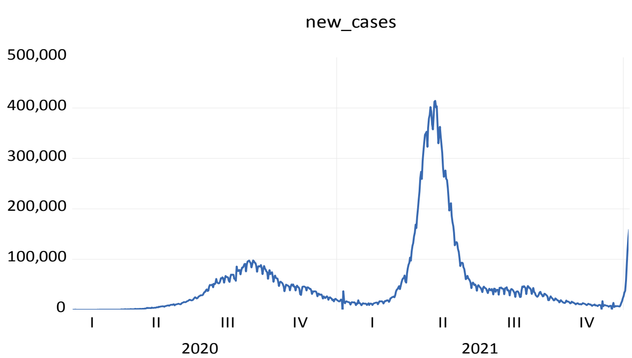
Where $\hat{\theta} the$ maximum value of the likelihood function. The out-of-sample Root mean squared error (RMSE) is calculated as before:

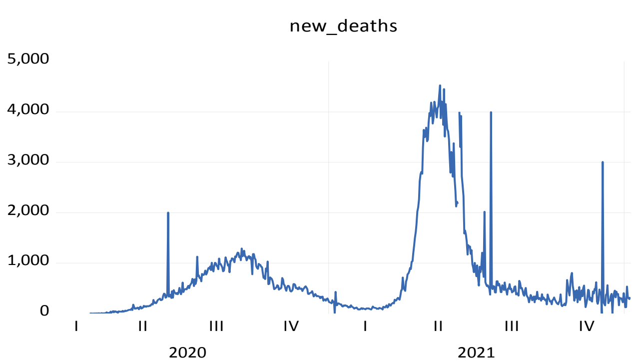
$$\sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \qquad (2)$$

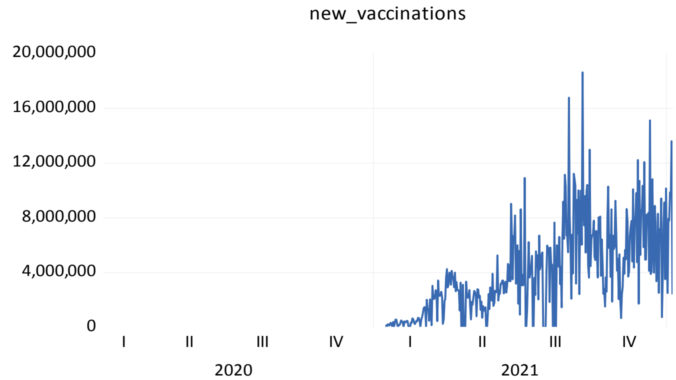
Where$\hat{y}_t$: the forecast value, y_t: the actual value, n: number of fitted observed, n: the number of correctly fitted data points that have been observed. This indication can improve prediction accuracy because it is more closely tied to the variable's value.

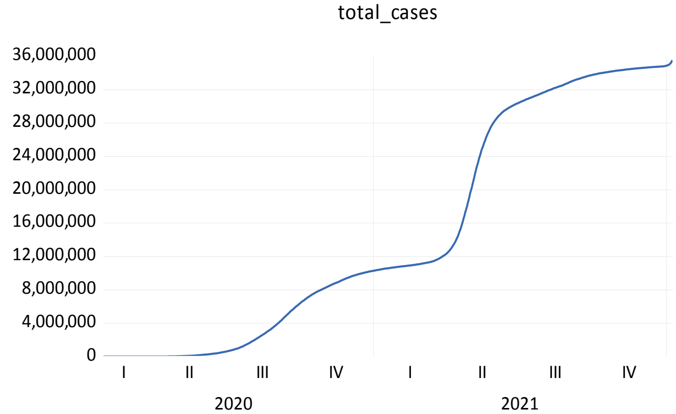
3 Results and Discussion:

We visualize the data and get the following results:









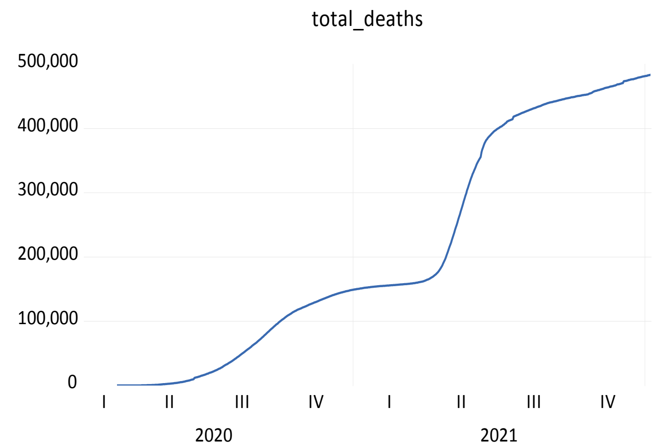


Fig.1 Visual perception of the evolution of (new cases, new deaths, new vaccinations, total cases, total deaths) during the corona pandemic.

In the visualization, we see a large increase in new cases during the first quarter of 2020 and an increase in new fatalities during the same period, both of which contributed to the increase in cumulative numbers of cases and total deaths during the first quarter of 2019. The increased number of new vaccines has also led to a decrease in the number of new cases and fatalities, which has continued until the beginning of 2022 when new cases have risen once more with a new wave of infection and mortality. The following table provides a high-level summary of the most significant descriptive statistics:

Table 2: Normal distribution and descriptive statistics for Index COVID19 in India.

	Normality	Maximum	Minimum	Skewness	Kurtosis
New cases	Non	414188	0	3.01	12.25
New deaths	Non	4592	0	2.49	8.71
New vaccinations	Non	18627269	0	0.99	4.02
Total cases	Non	35528004	1	0.38	1.51
Total deaths	Non	483790	1	0.35	1.56

We can't use the mean and standard deviation for characterizing because there isn't a normal distribution for all variables. We can observe that the distribution is skewed to the right for each variable, with leptokurtic distributions (new cases, new deaths, new vaccinations). The most significant number of new deaths, 4592, was recorded on 18/5/2021 when the number of new cases reached 414188. On the other hand, this figure dropped after 18/9/2021, when the maximum number of new vaccines (18627269) was reached. In order to forecast (new cases, new deaths, new vaccines, total cases, and total deaths), we estimate the ETS model for all variables. The following are the forecasting findings.

Table 3: the results of estimating ETS models.

	Model	**Parameters ETS**				**Initial stats**				**AIC**	**RMSE**
		α	β	ϕ	γ	l	b	$s1$	$s2$		
New cases	(M,A,N)	0.135069	0.029023	-	-	46.99119	-3.83958	-	-	16905.33	0.4102
New deaths	(M,AD,N)	0.109758	0.038004	0.967714	-	0.198014	0.23731	-	-	11445.05	0.5758
New vaccinations	(M,N,N)	0.142627	-	-	-	228772	-	-	-	12369.34	0.619024
Total cases	(A,MD,N)	1	1	0.983756	-	0.8	1.34090	-	-	17276.75	7179.872
Total deaths	(A,A,M)	0.987891	0.37929	-	0.0018	2.553	8.4148	0.9	1.1	12129.49	331.4

The table shows that the only model chosen with a seasonal component is for total deaths (A, A, M). We can also see from table.3 that the significant changes in trend were for each of the three models chosen with a seasonal component (total cases and deaths) where α and β are large, Where a preference was expressed for selecting the best model with the lowest AIC value, the following models were considered:

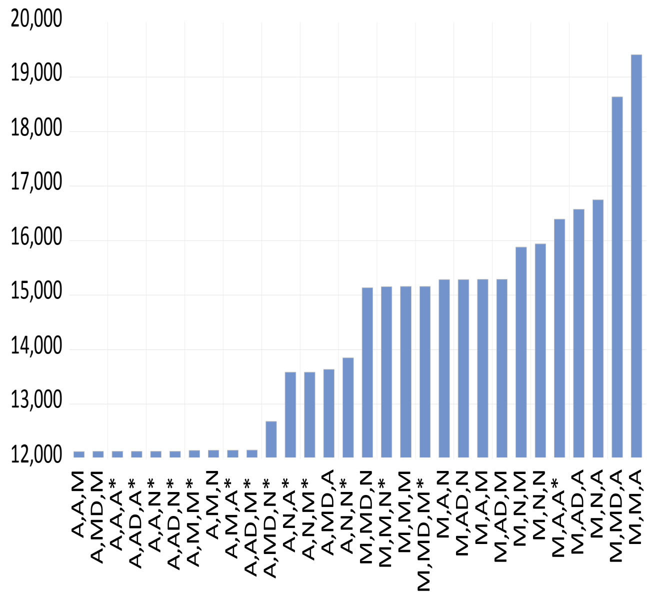
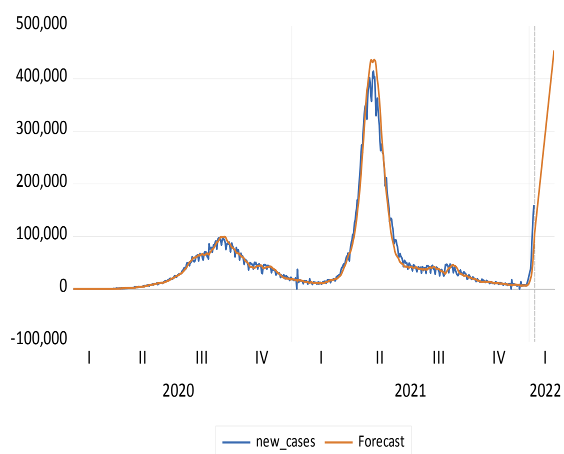
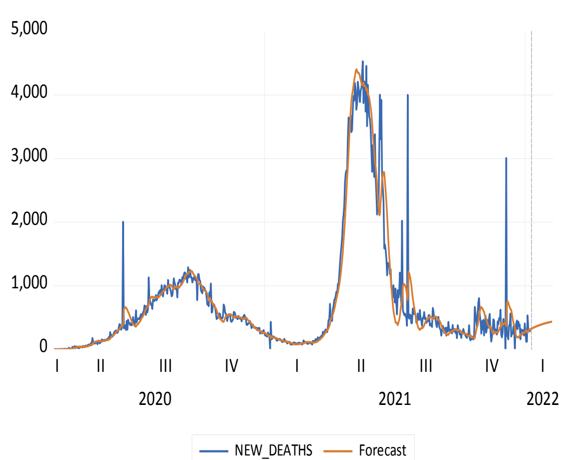


Fig 2. Selected best model for total deaths.

At this point we are creating a forecast for COVID19 data for an entire month and we get the following visualizations:





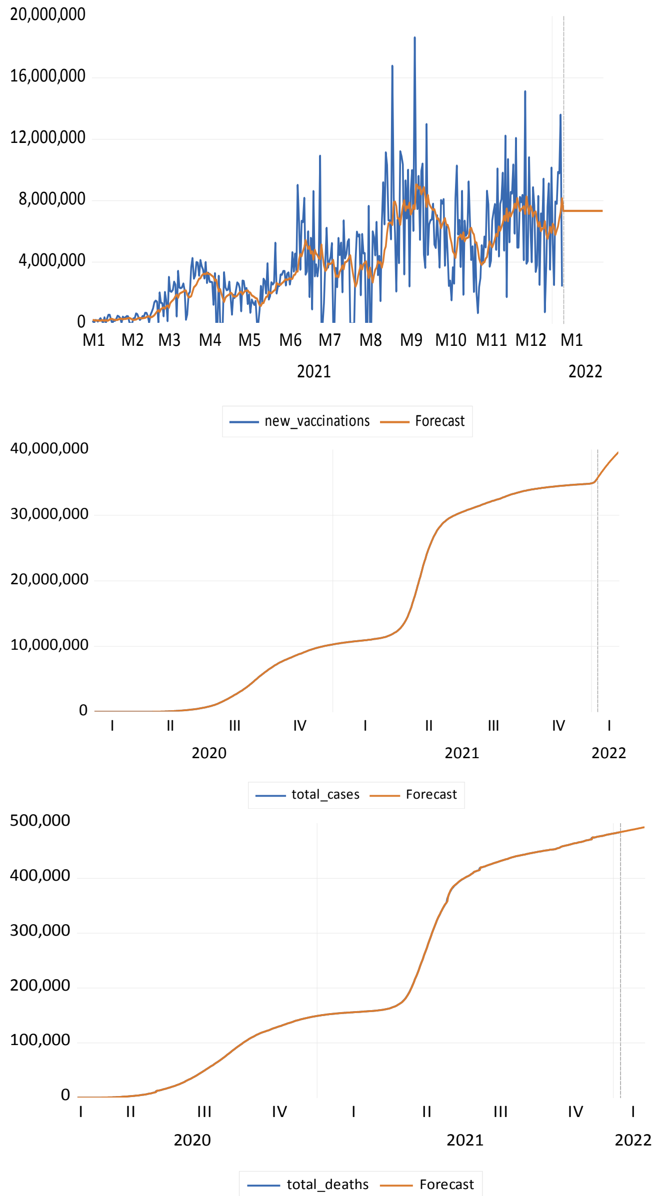


Fig 3. Forecasting for COVID19 data.

The visualization demonstrates that there is a good convergence between the actual and estimated values using the model, which is consistent with selecting the lowest values for information criteria according to table3. We also note that out-of-sample prediction values are the best for new cases because they have the lowest RMSE. In the table below, we show our expectations for (new cases, new deaths, new vaccinations, total cases, and total deaths) over the next ten days:

Table 4. Forecasting points in COVID19 for 10 next day.

	New cases	New deaths	New vaccinations	Total cases	Total deaths
9/1/2022	104410.83	322.45	7341793.05	35685745.86	484025.64
10/1/2022	116062.46	328.36	7341793.05	35844188.09	484386.73
11/1/2022	127714.10	334.07	7341793.05	36000742.92	484621.74
12/1/2022	139365.73	339.60	7341793.05	36155421.79	484982.91
13/1/2022	151017.36	344.95	7341793.05	36308236.44	485217.85
14/1/2022	162668.99	350.13	7341793.05	36459198.92	485579.10
15/1/2022	174320.63	355.14	7341793.05	36608321.53	485813.95
16/1/2022	185972.26	359.99	7341793.05	36755616.87	486175.28
17/1/2022	197623.89	364.69	7341793.05	36901097.74	486410.05
18/1/2022	209275.53	369.23	7341793.05	37044777.21	486771.46

The table shows that new cases are expected to increase significantly over the next ten days (209275.53) with a relatively stable number of new deaths recorded (369), as daily new vaccination cases contribute significantly to reducing deaths with a significant increase in total cases, prompting authorities to take the necessary measures with this new wave.

4 Conclusions

Noting that the sudden increase in cases is due to the omicron variant, India is now facing a third wave of COVID-19. The government must remain vigilant to ensure that the health system does not become overburdened. Infections are still on the rise in many Indian cities. We could see a significant increase in numbers over the next few weeks. In this paper, we used the Error Trend Seasonal (ETS) model within a dynamic nonlinear model framework with state-space based likelihood calculations to predict (total cases, new cases, total deaths, new deaths, and new vaccinations) in India, with support for model selection and forecast standard errors. These models have never been used with data from India before. They outperform the others in small sample size, non-normal data distribution, non-stationarity, trend, and seasonality. Prediction of (new cases, new deaths, new vaccinations, total cases, and total deaths) in India for the next ten days based on the best model developed using the lowest value of the information criterion. Predictions using the best-fit model from "9/1/2022" to "18/1/2022" indicate that new cases in India are expected to reach (209,275). With a growth rate of 100.43 percent, new deaths are expected to reach (369), with a growth rate of 14.5%. With a growth rate of 0%, the predicted values of daily new vaccination cases are expected to reach (7341793). The total number of cases is expected to reach (37044777) with a 3.8 percent growth rate. With a 0.56 percent growth rate, the total number of deaths is expected to reach (486771). These estimates are critical for understanding what healthcare systems will require to respond to the pandemic, such as the number of drugs to purchase, the amount of oxygen required, and the number of ICU beds needed, among other things. Proper preparation and effective planning to defeat this virus are desperately needed, so this forecast will assist the government. It may lead to the authorities taking the necessary measures against this new wave.

Availability of data and materials

All data are included within this paper.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

The authors received no financial support for the research

Acknowledgement

The work was supported by Act 211 of the Russian government, contract No. 02.A03.21.0011. The work was supported by the Russian Ministry of Science and Higher Education (government order FENU-2020-0022).

References

[1] Mishra P, Al Khatib AMG, Sardar I, Mohammed J, Ray M, et al. (2020) Modelling and Forecasting of COVID-19 in India. J. DOI: 10.23937/2474-3658/1510162.

[2] Pal M, Berhanu G, Desalegn C, Kandi V. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. Cureus. 2020;12(3):e7423. Published 2020 March 26. doi:10.7759/cureus.7423.

[3] Mandal, S., Arinaminpathy, N., Bhargava, B, and Panda, S. (2021). Plausibility of a third wave of COVID-19 in India: A mathematical modelling based analysis. Indian J Med Res 153, May & June 2021, pp 522-532. DOI: 10.4103/ijmr.ijmr_1627_21.

[4] Song, H., Fan, G., Zhao, S., Li, H., Huang, Q., He, D. (2021). Forecast of the COVID-19 trend in India: A simple modelling approach. Mathematical Biosciences and Engineering. 6(18). 9775-9786. doi: 10.3934/mbe.2021479.

[5] Kibria, H., Jyoti, O., Matin, A. (2022). Forecasting the spread of the third wave of COVID-19 pandemic using time series analysis in Bangladesh, Informatics in Medicine Unlocked, 28, 1-8. doi.org/10.1016/j.imu.2021.100815.

[6] Fenga, L. (2021). Forecasting the COVID-19 Diffusion in Italy and the Related Occupancy of Intensive Care Units. Journal of Probability and Statistics. 2021. 1-9. doi.org/10.1155/2021/5982784.

[7] Mohammed et al: Modeling and forecasting of COVID-19 from the context of Ghana. African Review of Economics and Finance | ISSN 2042-1478 |.

[8] Abotaleb, M. Makarovskikh, T, et al. (2020). Modeling Covid-19 Infection Cases and Vaccine in 5 Countries Highly Vaccinations. 13(2). 403-417. DOI: 10.47000/tjmcs.905508.

[9] Fatih, C. ,A.dHamimes, P. Mishra, Vol. 7 (2), 2020, pages 967 - 982. Covid-19 Statistics, Strange trend and Forecasting of Total Cases in the most Infected African Countries: An ARIMA and Fuzzy Time Series Approaches.

[10] Hyndman, Rob J., Koehler, Anne B., Ord, J. Keith, and Simone Grose (2008). Forecasting with Exponential Smoothing: The State Space Approach. Berlin: Springer-Verlag.

[11] Hyndman, Rob J., Koehler, Anne B., Snyder, Ralph D., and Simone Grose (2002). A State Space Framework For Automatic Forecasting Using Exponential Smoothing Methods, International Journal of Forecasting, 18, 439–454.