ORIGINAL PAPER

Stochastic forecasting of COVID-19 daily new cases across countries with a novel hybrid time series model

Arinjita Bhattacharyya · **Tanujit Chakraborty** · **Shesh N. Rai**

Abstract An unprecedented outbreak of the novel coronavirus (COVID-19) in the form of peculiar pneumonia has spread globally since its first case in Wuhan province, China, in December 2019. Soon after, the infected cases and mortality increased rapidly. The future of the pandemic's progress was uncertain, and thus, predicting it became crucial for public health researchers. These predictions help the effective allocation of health-care resources, stockpiling, and help in strategic planning for clinicians, government authorities, and public health policymakers after understanding the extent of the effect. The main objective of this paper is to develop a hybrid forecasting model that can generate real-time out-of-sample forecasts of COVID-19 outbreaks for five profoundly affected countries, namely the USA, Brazil, India, the UK, and Canada. A novel hybrid approach based on the Theta method and autoregressive neural network (ARNN) model, named Theta-ARNN (TARNN) model, is developed. Daily new cases of COVID-19 are nonlinear, non-stationary, and volatile; thus, a single specific model cannot be ideal for future prediction of the pandemic. However, the newly introduced hybrid forecasting model with an acceptable prediction error rate can help healthcare and government for effective planning and resource allocation. The proposed method outperforms traditional univariate and hybrid forecasting models for the test datasets on an average.

Arinjita Bhattacharyya and Tanujit Chakraborty have contributed equally.

A. Bhattacharyya · S. N. Rai
Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, KY, USA

T. Chakraborty (✉)
Department of Science and Engineering, Sorbonne University Abu Dhabi, Abu Dhabi, UAE
e-mail: tanujit_r@isical.ac.in,
tanujit.chakraborty@sorbonne.ae

S. N. Rai
Biostatistics and Bioinformatics Facility, JG Brown Cancer Center, University of Louisville, Louisville, KY, USA

S. N. Rai
The Christina Lee Brown Envirome Institute, University of Louisville, Louisville, KY, USA

S. N. Rai
University of Louisville Alcohol Research Center, University of Louisville, Louisville, KY, USA

S. N. Rai
University of Louisville Hepatobiology & Toxicology Center, University of Louisville, Louisville, KY, USA

1 Introduction

In December 2019, clusters of pneumonia cases caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) were identified at Wuhan, Hubei province, in China [1–3]. Soon after the emer-

gence of the novel beta coronavirus, World Health Organization (WHO) characterized this contagious disease as a pandemic in March 2020 due to its rapid spread within and outside the highly mobile population of China which got further aggravated by the densely populated location of the sea food market and time of Chinese new year [4]. It resulted in an exponential increase in the incidence rate (IR) and case fatality rate (CFR) [5]. As of March 16, 2021, a total of 120,512,041 confirmed global cases and 2,665,742 deaths have been reported worldwide [6].

Researchers have faced significant challenges to forecast real-time COVID-19 cases with traditional mathematical, statistical, and machine learning-based forecasting tools [7–13]. Studies in March 2020 with simple and efficient forecasting methods such as exponential smoothing model predicted cases ten days ahead, with large confidence intervals, that despite the positive bias, had reasonable forecast error [14]. Previously used linear and exponential model forecasts for better preparation regarding hospital beds, ICU admission estimation, resource allocation, emergency funding, and proposing strong containment measures were conducted [15] that projected ICU admissions in Italy for March 20, 2020. ICU admission and mechanical ventilation for critically ill patients reached their peak shattering the health system of Lombardy, Italy, by end of March 2020 [16]. Health-care workers had to go through the immense mental stress left with a formidable choice of prioritizing young and healthy adults over the elderly for allocation of life support, especially unwanted ignoring of those who are extremely unlikely to survive [17,18]. Real estimates of mortality with 14-day delay demonstrated underestimating the COVID-19 outbreak and indicated a grave future with a global CFR of 5.7% [19]. Some of the countries are eventually in a second wave with cases rising in India, and the UK. The uncertain future has several unanswered questions on the daily trajectory of the pandemic, infection rate, and number of deaths. Thus, real-time forecasting with time series models could respond to these debatable quires to reach a statistically validated conjecture in this current health crisis. Some of the impacting leading-edge research concerning real-time projections of COVID-19 confirmed cases, recovered cases, and mortality using statistical, epidemiological, and machine learning models are as follows. Short-term forecasting of cumulative confirmed cases was produced using autoregressive inte-

grated moving average (ARIMA), Cubist regression, random forest, ridge, support vector regression (SVR), and stacking-ensemble learning (SEL) model in [20], whereas exponential smoothing model produced ten days ahead forecasts of actual cases within 90% CI and forecasts reflect the significant increase in the trend of global cases with growing uncertainty [14]. Forecasting and nowcasting domestic and international COVID-19 cases are also done with epidemiological compartmental models, like susceptible–exposed–infectious–recovered (SEIR) and SIR models [8,9,21]. The prediction and modeling of cumulative COVID-19 deaths in Mexico have utilized ANN and a Gompertz model with multiple optimization algorithms and emphasized on the performance of ANN with different parameter settings [22]. Parametric or data-driven modeling approaches for the virus infection and propagation and machine learning methods for predicting the virus spreading dynamics are previously used in the literature [23,24]. Modern machine learning techniques such as neural networks and black-box deep learning models were also used for COVID-19 forecasting [25,26]. Dynamics and control of COVID-19 pandemic with nonlinear incidence rates and building nonlinear models for COVID-19 forecasting have been found useful for European countries [27,28]. There is a huge literature on mathematical modeling in epidemiology starting from the early work of [29], going through the compartmental models of [30] laying down the foundations of modern epidemiology, up to the current data-driven approaches (see, e.g., [31–33] addressing COVID-19). However, forecasting COVID-19 cases is harder and this is primarily due to the following major factors [34]: *(a) Very less amount of data is available; (b) less understanding of the factors that contribute to it; and (c) model accuracy is constrained by our knowledge of the virus. With an emerging disease such as COVID-19, many biologic features of transmission are hard to measure and remain unknown; (d) another source of uncertainty, affecting all models, is that we do not know how many people are, or have been, infected; and (e) we are certainly missing a substantial number of cases due to virologic testing, so models fitted to confirmed cases are likely to be highly uncertain* [35].

Time-series forecasting models take the input of historical observations and extrapolate the patterns into the future. The past in no way will continue in the future for this pandemic, so it is challenging to model and produce predictions. There are essentially two general

approaches to forecasting a time series: (a) generating forecasts from a single model and (b) combining forecasts from many models (hybrid). In classical time series forecasting, the ARIMA model is used predominantly for forecasting linear time series [36], which has a significant strong assumption of linearity in the system and homoscedastic error distribution, typically without sudden jumps and bursts. Individual models such as ARIMA, wavelet ARIMA (WARIMA) [37], and Theta method [38, 39] are inadequate to model such situations. Usage of nonlinear time series techniques in infectious disease modeling has successfully demonstrated the success of artificial neural networks (ANNs) and autoregressive neural networks (ARNNs) [40–43]. The idea of hybridizing forecasting models is not a new concept, and their empirical applications resulted in superior forecasts to their counterparts. Most recently, some hybrid models combining linear and nonlinear time series models are proposed for COVID-19 [12,44] and performed exceptionally well for predicting these epidemics.

Motivated by this, this study considers the time series datasets of coronavirus confirmed cases that show nonlinearity, non-stationary, and non-Gaussian patterns, thus making decisions based on a discrete model critical and unreliable. Another difficulty in COVID-19 data from the modeling aspect is the unavailability of sufficient data points, which generates biased predictions and estimates, which can be well maneuvered by neural nets [45]. Most of the relevant studies focused on the outbreak's short-term and long-term forecasts of reported confirmed cases have a broad range of fluctuations, wide confidence intervals, poorly reported data and model specifics, and highly optimistic predictive performance [46]. On the other hand, this paper handles the problem of forecasting of the current pandemic cases by designing a hybrid decision-making method based on highly useful Theta forecasting method and machine learning-based ARNN model. Hybridization of two highly accurate and useful models can be regarded as the most handy solution [47] for optimizing forecasting performance, and efficiency with irregular data characteristics of COVID-19 cases [48]. The importance of hybrid methodology with a fusion of linear and nonlinear forecasting models became evident in tackling such dynamic/nonlinear time series due to its in-built time-changing variance, complex auto-correlation structure [12,49]. With the growing need for advanced parsimonious hybrid forecasting methods

accompanied by precise accuracy and accurate forecasts, this paper focuses on providing a plausible solution for COVID-19 forecasting using the newly proposed hybrid decision making system. The main objectives of this study are as follows:

1. To propose a simple and computationally efficient hybrid forecasting model which generates out-of-sample forecasts (two months) for five profoundly affected countries (the USA, Brazil, India, the UK, and Canada) with higher accuracy as compared to state-of-the-art forecasting methods.
2. To prove the model's stationarity and ergodicity properties from a statistical point of view.
3. To discuss the merits and future challenges that need attention while working with the proposal for epidemiological forecasting. We also recommend policy-making decisions and resource allocation based on these forecasts.

Therefore, this study proposes a novel hybrid Theta autoregressive neural network model (TARNN) model combining Theta and ARNN models that can capture complex COVID-19 data structures. The proposed hybrid decision-making model captures complex data structures and linear plus nonlinear behavior of pandemic datasets. In the first phase of our proposed model, Theta method catches the linear patterns of the dataset. Then, the ARNN model is employed to capture the nonlinear patterns in the data using residual values obtained from the base Theta model. The proposed TARNN model has easy interpretability, produces robust predictions, and is adaptable to seasonality. Through experimental evaluation, we have shown the excellent performance of the proposed hybrid model for the pandemics forecasting for several datasets. An RShiny application is also built for TARNN which can help practitioners reproducing and updating the forecasts as further data become available.

The rest of the paper is organized as follows: Sect. 2 describes the formulation of the proposed hybrid TARNN model. The ergodicity and stationarity of the proposed hybrid model are discussed in Sect. 3. In Sect. 4, we discuss the country-wise COVID-19 confirmed case datasets and experimental results. The discussions about the results and practical implications are given in Sect. 5. Finally, we conclude the paper in Sect. 6.

2 Methodology

We start by discussing the single forecasting models to be used in the hybridization, followed by the detailed formulation of the proposed hybrid Theta-ARNN model.

2.1 Theta method

The 'Theta method' or 'Theta model' is a univariate time series forecasting technique that performs particularly well in M3 forecasting competition and of interest to forecasters [39]. The method decomposes the original data into two or more lines, called theta lines, and extrapolates them using forecasting models. Finally, the predictions are combined to obtain the final forecasts. The theta lines can be estimated by simply modifying the 'curvatures' of the original time series. This change is obtained from a coefficient, called θ coefficient, which is directly applied to the second differences of the time series:

$$Y''_{new}(\theta) = \theta Y''_{data}, \tag{1}$$

where $Y''_{data} = Y_t - 2Y_{t-1} + Y_{t-2}$ at time t for $t = 3, 4, \ldots, n$ and $\{Y_1, Y_2, \ldots, Y_n\}$ denote the observed univariate time series. In practice, the coefficient θ can be considered as a transformation parameter which creates a series of the same mean and slope with that of the original data but having different variances. Now, Eq. (1) is a second-order difference equation and has solutions of the following form [38]:

$$Y_{new}(\theta) = a_\theta + b_\theta(t-1) + \theta Y_t, \tag{2}$$

where a_θ and b_θ are constants and $t = 1, 2, \ldots, n$. Thus, $Y_{new}(\theta)$ is equivalent to a linear function of Y_t with a linear trend added. The values of a_θ and b_θ are computed by minimizing the sum of squared differences:

$$\sum_{i=1}^{t} [Y_t - Y_{new}(\theta)]^2$$
$$= \sum_{i=1}^{t} [(1-\theta)Y_t - a_\theta - b_\theta(t-1)]^2. \tag{3}$$

Forecasts from the Theta model are obtained by a weighted average of forecasts of $Y_{new}(\theta)$ for different values of θ. Furthermore, the prediction intervals and

likelihood-based estimation of the parameters can be obtained based on a state space model which is demonstrated in [38]. The generalized version of the Theta method is suitable for automatic forecasting of time series [50].

2.2 ARNN model

Artificial neural network-based forecasting methods received increasing interest in various applied domains in the late 1990s. A wide variety of neural nets are popularly used for supervised classification, prediction, and nonlinear time series forecasting [51]. The architecture of a simple feedforward neural network can be described as a network of neurons arranged in an input layer, hidden layer, and output layer in a prescribed order. Each layer passes the information to the next layer using weights that are obtained using a learning algorithm [52]. ARNN model is a modification to the simple ANN model especially designed for prediction problems of time series datasets [52]. ARNN model uses a pre-specified number of lagged values of the time series as inputs, and the number of hidden neurons in its architecture is also fixed [53]. ARNN(p, k) model uses p lagged inputs of the time series data in a one hidden layered feedforward neural network with k hidden units in the hidden layer. Let $\underline{x}$ denote a p-lagged inputs and f be a neural network of the following architecture:

$$f(\underline{x}) = c_0 + \sum_{j=1}^{k} w_j \phi \left(a_j + b'_j \underline{x} \right); \tag{4}$$

where c_0, a_j, w_j are connecting weights, b_j is p-dimensional weight vector, and ϕ is a bounded nonlinear sigmoidal function (e.g., logistic squasher function or tangent hyperbolic activation function). These weights are trained using a gradient descent backpropagation [54]. Standard ANN faces the dilemma to choose the number of hidden neurons in the hidden layer, and the optimal choice is unknown. However, for ARNN model, we adopt the formula $k = [(p+1)/2]$ for nonseasonal time series data where p is the number of lagged inputs in an autoregressive model [53].

2.3 Proposed TARNN model

In this section, we describe the proposed hybrid model based on Theta method and ARNN model and we

name it TARNN model. The proposed TARNN model is based on an error remodeling approach, and there are broadly two types of error calculations popular in the literature which are given below [55].

Definition 1 In the additive error model, the forecaster treats the expert's estimate as a variable, $\hat{Y}_t$, and thinks of it as the sum of two terms:

$$\hat{Y}_t = Y_t + e_t,$$

where Y_t is the true value and e_t is the additive error term.

Definition 2 In the multiplicative error model, the forecaster treats the expert's estimate $\hat{Y}_t$ as the product of two terms:

$$\hat{Y}_t = Y_t \times e_t,$$

where Y_t is the true value and e_t is the multiplicative error term.

Now, even if the relationship is of product type, in the log–log scale, it becomes additive. Hence, without loss of generality, we may assume the relationship to be additive and expect the errors (additive) of a forecasting model to be random shocks. However, this is violated when there are complex correlation structures in the time series data and less amount of knowledge is available about the data generating process. A simple example is the daily confirmed cases of the COVID-19 cases for various countries where very little is known about the structural properties of the current pandemic. Thus, we need two-stage modeling approach to formulate this complex time series problem. The proposed TARNN model is a hybrid model based on the additive error remodeling approach. The hybrid TARNN approach consists of three basic steps:

– In the first step of the TARNN model, the Theta method is applied to the time series data to model the linear components of a given time series dataset.
– Theta model generates in-sample forecasts, and the error series is calculated.
– In the next phase, the residuals (additive errors) generated by the Theta method are remodeled using a nonlinear ARNN model. Finally, both the forecasts obtained from the Theta and ARNN models are combined together to get the final forecasts for the given time series.

The mathematical formulation of the proposed hybrid TARNN model (Z_t) is as follows:

$$Z_t = L_t + N_t,$$

where L_t is the linear part and N_t is the nonlinear part of the hybrid model. We can estimate both L_t and N_t from the available time series data. Let $\hat{L}_t$ be the forecast value of the Theta model at time t and ϵ_t represent the error residuals at time t, obtained from the Theta model. Then, we write

$$\epsilon_t = Z_t - \hat{L}_t.$$

These left-out residuals are further modeled by ARNN model and can be represented as follows:

$$\epsilon_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-p}) + \varepsilon_t,$$

where f is a nonlinear function and the modeling is done by the ARNN model as defined in Eq. (4) and ε_t is supposed to be random shocks. Therefore, the combined forecast can be obtained as follows:

$$\hat{Z}_t = \hat{L}_t + \hat{N}_t,$$

where $\hat{N}_t$ is the forecasted value of the ARNN model. An overall flow diagram of the proposed TARNN model is given in Fig. 1.

In the proposed TARNN model, ARNN is applied to remodel the left-over autocorrelations in the residuals which Theta model individually could not model. Thus, the proposed TARNN model can be considered as an error remodeling approach. This is important because due to model mis-specification and disturbances in the pandemic time series, the linear Theta model may fail to generate white noise behavior for the forecast residuals. The TARNN approach eventually can improve the predictions for the epidemiological forecasting problem as shown in Sect. 4 .

Remark 1 The idea of additive error modeling is useful for modeling complex time series for which achieving random shocks based on individual forecasting models is difficult. More precisely, the TARNN approach is developed for forecasting the COVID-19 confirmed cases for which the data generating process and the various characteristics of the epidemic are still unknown. The proposed TARNN model only assumes that the linear and nonlinear components of the epidemic time series can be separated individually.
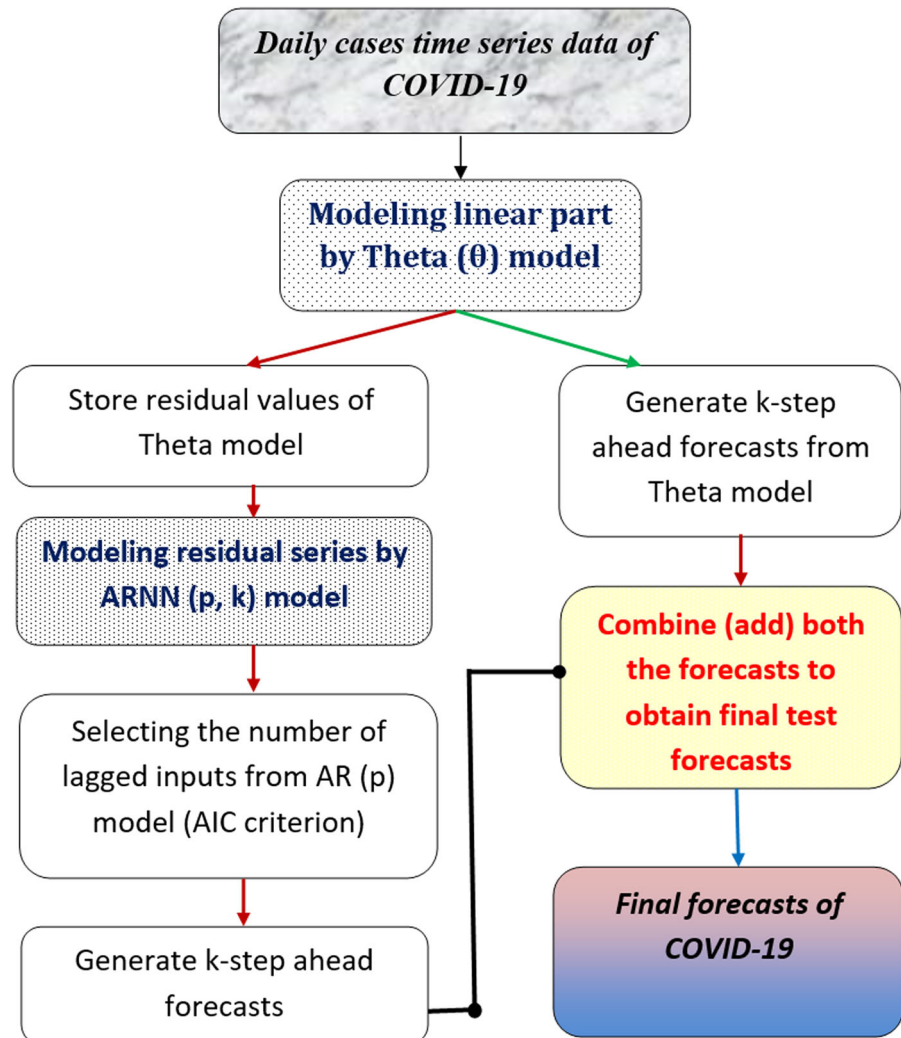
3 Ergodicity and stationarity of the proposed TARNN model

In this section, we derive the results for the ergodicity and asymptotic stationarity of the proposed TARNN model. The ergodicity and stationarity are of particular importance from a statistician's point of view in time series analysis since for such processes a single realization displays the whole probability of the data generating process. We use several previous results on nonlinear time series and Markov chains to find sufficient conditions for which the overall process is ergodic and stationary [56–59].

To start, we write the underlying stochastic model of the Theta method using the state space approach. We initialize the model by setting $Y_1 = l_1$ and then for

$t = 2, 3, \ldots$; and drift term b, let

$$Y_t = l_{t-1} + b + \epsilon_t \quad \text{and} \quad l_t = l_{t-1} + b + \alpha \epsilon_t$$

where $\{\epsilon_t\}$ is assumed to follow Gaussian white noise with mean zero and variance σ^2 and α is the smoothing parameter for the simple exponential smoothing (SES) model. Now, Y_t follows a state space model which gives forecasts equivalent to SES with drift [38]. Also, Y_t can be written in the following form:

$$Y_t = Y_{t-1} + b + (\alpha - 1)\epsilon_{t-1} + \epsilon_t. \quad (5)$$

The above is an ARIMA$(0, 1, 1)$ process with drift term [38]. The left-out residuals of Theta model are further modeled by ARNN process. We consider the ARNN process generated by the additive noise of the ARIMA$(0, 1, 1)$ process with drift. Let ϵ_t be the process defined by the stochastic difference equation of

the following form:

$$\epsilon_t = f(\epsilon_{t-1}, \epsilon_{t-2}, \cdots, \epsilon_{t-p}, \Theta) + \varepsilon_t, \tag{6}$$

where ε_t is an i.i.d. noise process and $f(\cdot, \Theta)$ is a feed-forward neural network with weight (parameter) vector Θ and inputs $\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-p}$. The definition of f is given in Eq. (4).

3.1 Time series as Markov chains

We start by defining the following notations:

$$z_{t-1} = \left(\epsilon_{t-1}, \ldots, \epsilon_{t-p}\right)'$$
$$F(z_{t-1}) = \left(f(z_{t-1}), \epsilon_{t-1}, \ldots, \epsilon_{t-p+1}\right)'$$
$$e_t = \left(\varepsilon_t, 0, \ldots, 0\right)'$$

Then, Eq. (6) can be written as follows [56]:

$$z_t = F(z_{t-1}) + e_t \tag{7}$$

with $z_t, e_t \in \mathbb{R}^p$. In this section, we show the (strict) stationarity of the state space form, as defined in Eq. (7). The problem of showing $\{z_t\}$ to be stationary is closely related to the ergodicity of the process [60]. A Markov chain $\{z_t\}$ is called geometrically ergodic if there exists a probability measure π on the state space $(\mathbb{R}^p, \mathbb{B})$ and a constant $\rho > 1$ such that

$$\lim_{n \to \infty} \rho^n \|P^n(z, A) - \pi(A)\| = 0 \tag{8}$$

for each $z \in \mathbb{R}^p$, $\mathbb{B}$ is the Borel σ-algebra on $\mathbb{R}^p$ and $\|\cdot\|$ denotes the total variation norm. If Eq. (8) holds for $\rho = 1$, then $\{z_t\}$ is called ergodic.

The definition for $P^n(z, A)$ can be given as the probability that $\{z_n\}$ moves from z to the set $A \in \mathbb{B}$ in n steps: $P^n(z, A) = P(z_{t+n} \in A | z_t = z)$. Also, expression for $\pi(A)$ is as follows:

$$\pi(A) = \int P^1(z, A) \pi(dz) \quad \text{for all} \quad A \in \mathbb{B}.$$

Thus, we call π as the stationary measure and the distribution of z_t converges to π if $\{z_t\}$ is ergodic. Then, we say $\{z_t\}$ is asymptotically stationary [61]. To establish the ergodicity of TARNN processes, we need the concept of irreducibility and aperiodicity. A Markov chain $\{z_t\}$ is called irreducible if

$$\sum_{n=1}^{\infty} P^n(z, A) > 0 \quad \text{for all} \quad z \in \mathbb{R}^p,$$

whenever $\lambda(A) > 0$ and λ denotes the Lebesgue measure on $(\mathbb{R}^p, \mathbb{B})$. Thus, for an irreducible Markov chain, all parts of the state space can be reached by the Markov chain irrespective of the starting point. Now, an irreducible Markov chain is aperiodic if there exists an $A \in \mathbb{B}$ with $\lambda(A) > 0$ and for all $C \in \mathbb{B}$, $C \subseteq A$ with $\lambda(C) > 0$, there exists a positive integer n such that

$$P^n(z, C) > 0 \quad \text{and} \quad P^{n+1}(z, C) > 0 \quad \text{for all} \quad z \in \mathbb{R}^p.$$

Hence, it is possible that the Markov chain returns to given sets only at specific time points for an aperiodic Markov chain. For most general time series models, irreducibility and aperiodicity cannot be assumed automatically. However, for a TARNN process, these conditions can be checked. In general, it is sufficient to assume the distribution of the noise process to be an absolutely continuous component with respect to Lebesgue measure and the support of the probability density function (PDF) is sufficiently large.

3.2 Main results

It is clear from the above discussion that if the Markov chain is geometrically ergodic, then its distribution will converge to π and the corresponding time series will be called asymptotically stationary, see also [58]. Lemma 1 states that the state space of the Markov chain cannot be reduced depending on the starting point.

Lemma 1 *Let $\{z_t\}$ be defined by* (7), *and let $E|\varepsilon_t| < \infty$ and the PDF of ε_t be positive everywhere in $\mathbb{R}$. Then, if f is defined by* (4), *the Markov chain $\{z_t\}$ is φ-irreducible and aperiodic.*

Proof Since the support of the PDF of ε_t is the whole real line, that is, the PDF is positive everywhere in $\mathbb{R}$, then we can say that $\{z_t\}$ is φ-irreducible by using [56]. In our case, every non-null p-dimensional hypercube can be reached in p steps with positive probability (and hence every non-null Borel set A). A necessary and sufficient condition for $\{z_t\}$ to be aperiodic is to have a set A and positive integer n such that $P^n(z, A) > 0$ and $P^{n+1}(z, A) > 0$ for all $z \in A$ [58]. In this case, this is true for all n due to consideration of the unbounded additive noise. $\square$

The theorem below states the necessary conditions for geometric ergodicity of a Markov chain. This can be obtained using the decomposition technique and ergodicity of stochastic difference equations [56,59].

Theorem 1 *Suppose $\{z_t\}$ is defined as in (6) and (7), F is a compact set that can be decomposed as $F = F_h + F_d$, and the following conditions hold:*

(i) *$F_h(.)$ is continuous and homogeneous, and $F_d(.)$ is bounded;*

(ii) *$E|\varepsilon_t| < \infty$, and probability distribution function of ε_t is positive everywhere in $\mathbb{R}$; then, $\{z_t\}$ is geometrically ergodic.*

Proof $\{z_t\}$ satisfies the following equation:

$$z_t = F(z_{t-1}) + e_t \quad t \geq 1, \quad \text{and} \quad F : \mathbb{R}^p \to \mathbb{R}^p. \quad (9)$$

Let F_h be continuous and homogeneous, viz. $F_h(cz) = cF_h(z)$ for all $c > 0$, $z_t \in \mathbb{R}^p$, and F_d be bounded. Let the origin, O, be a fixed point of F_h. It is important to note that ε_t satisfies the condition (ii) in Theorem 1. We are going to show the existence of a continuous Lyapunov function, V, in a neighborhood of the origin which will ensure the geometric ergodicity of (7).

To start with we let $W \subseteq \mathbb{R}^p$, the closure of W by $\overline{W}$ and its boundary by $\widehat{W}$. We also let V be defined over the closure of the unit ball. We let $p_0 = \inf_{\|z\|=1} V(z)$, where $\|\cdot\|$ denotes the Euclidean norm. Also, let G be the maximal connected component of $\{z : V(z) < \frac{p_0}{2}\}$ that contains the origin. Then, we have $F(\overline{G}) \subset G$.

Let $g(z) = \inf\{r \geq 0, z \in rG\}$, $z \in \mathbb{R}^p$, where $rG = \{rz, z \in G\}$. Then, $g(z)$ is well defined and it can easily be checked that g has the following properties:

1. $g(cz) = cg(z)$, for all $c > 0$.
2. There exists $0 < c < \mathbb{C} < \infty$ such that $c\|z\| \leq g(z) \leq \mathbb{C}\|z\|$.
3. $\left(z/g(z)\right) \in \widehat{G}$.
4. There exists $\epsilon > 0$, $0 < \theta < 1$ such that for all $z \in \widehat{G}$, $y \in \mathbb{R}^p$, we have $\|y - F(z)\| < \epsilon \Rightarrow y \in G$ and $g(y) < \theta$.

Now, for Eq. (7), ε_t satisfies $E|\varepsilon_t| < \infty$. Let $A \in \mathbb{B}$ and $z \in \mathbb{R}^p$, we define $P(z, A)$ be the transition probability function as:

$$P(z, A) = \int_{A - F(z)} f(t)\mathbb{P}(dt).$$

Thus, it holds that

$$\int g(y) P(z, dy)$$
$$= \int g\big(F_h(z) + t\big) f(t)\mathbb{P}(dt)$$

$$= \int_{\|t/g(z)\| < \epsilon} g(z) g\left(F_h\left(\frac{z}{g(z)}\right) + \frac{t}{g(z)} \right) f(t)\mathbb{P}(dt)$$
$$+ \int_{\|t/g(z)\| \geq \epsilon} g(z) g\left(F_h\left(\frac{z}{g(z)}\right) + \frac{t}{g(z)} \right) f(t)\mathbb{P}(dt)$$
$$< \big[\theta g(z)\big] + \big[M + g(z)\beta(z)\big] \quad \text{for some } M > 0.$$

Here, $|\beta(z)| < B < \infty$ for all z and $\beta(z) \to 0$ as $\|z\| \to \infty$. We let $h(z) = g(z) + 1$ and r be such that

$$|\beta(z)| < \frac{(1 - \theta)}{4} \quad \text{for} \quad \|z\| > r.$$

Then, for $r_0 = \max\left(r, \frac{4(C+1)}{(1-\theta)p}\right)$, there exists B' such that

1. $\int f(y) P(z, dy) < B' < \infty$ when $\|z\| \leq r_0$.
2. $\int f(y) P(z, dy) < \frac{1}{2}(1 + \theta)h(z)$ when $\|z\| > r_0$.

Using Theorem 4 of [60], we can conclude that $\{z_t\}$ is geometric ergodic. $\qquad\square$

The next theorem gives the main result for asymptotic stationary of the TARNN model.

Theorem 2 *Let $E|\varepsilon_t| < \infty$ and the PDF of ε_t be positive everywhere in $\mathbb{R}$, and $\{\varepsilon_t\}$ and $\{z_t\}$ are defined as in (6) and (7), respectively. Then, if f is a nonlinear neural network as defined in (4), then $\{z_t\}$ is geometrically ergodic and $\{\varepsilon_t\}$ is asymptotically stationary.*

Proof The noise process ε_t satisfies $E|\varepsilon_t| < \infty$ by assumption (e.g., Gaussian noise). It is also important to note that neural network activation functions, more precisely logistics or tan-hyperbolic activation functions, are continuous compact functions and have a bounded range. Thus, $\{z_t\}$ satisfies all the criteria to be geometrically ergodic and using Theorem (1), one can write that for the ARNN process with $F_h \equiv 0$ and $F_d \equiv F$. Thus, the series $\{\varepsilon_t\}$ is asymptotically stationary. $\qquad\square$

Remark 2 Some interpretations and practical implications of the theoretical results are given below:

– The geometric rate of convergence in Theorem 2 implies that the memory of TARNN process vanishes exponentially fast. This implies that the simplest version of the proposed model converges to a Wiener process.

Table 1 Description of COVID-19 datasets

Country name	Start–end date	No. of days	Minimum value	Maximum value	Train size	Test size
USA	22/01/2020–26/02/2021	402	0	251,161	342	60
Brazil	26/02/2020–26/02/2021	367	0	70,574	307	60
India	30/01/2020–26/02/2021	394	0	97,894	334	60
UK	31/01/2020–26/02/2021	393	0	41,460	333	60
Canada	26/01/2020–26/02/2021	398	0	10,100	338	60

– This is important for predictions over larger intervals of time; for example, one might train the network on an available sample and then use the trained network to generate new data with similar properties like the training sample. The results for the asymptotic stationarity can guarantee that the proposed hybrid model cannot have growing variance over time.

– From practitioners point of view, when the data are generated by the irreducible TARNN process, the estimated weights are not too far from the true weights. Then, one can draw an indirect conclusion on the statistical nature of the observed data based on the estimated weights.

4 Experimental analysis

Five time series COVID-19 datasets are considered for assessing various forecasting models (individual and hybrid). The datasets are mostly nonlinear, non-stationary, and non-Gaussian in nature. We have used root-mean-square error (RMSE), mean absolute error (MAE), and mean absolute scaled error (MASE) [62] to evaluate the predictive performance of the models used in this study. Since the number of data points in all the datasets is limited, advanced deep learning techniques will overfit the datasets [63].

4.1 Data description

Data are collected from the 'Our World in Data' public repository (Link: https://ourworldindata.org/coronavirus) on five countries: the USA, Brazil, India, the UK, and Canada that have the leading number of confirmed cases of COVID-19. These are univariate time series from inception—February 26, 2021, of confirmed cases that are analyzed for generating future out-

break predictions. We have taken last 60 data points as test points on which the forecasting models are evaluated. A summary of the COVID-19 datasets of confirmed cases and their descriptions is presented in Table 1. A pictorial view of the training datasets along with auto-correlation function (ACF) and partial ACF (PACF) plots is given in Table 2.

4.2 Performance evaluation metrics

The performance of different forecasting models are evaluated based on RMSE, MAE, and MASE metrics for these five time series [53]:

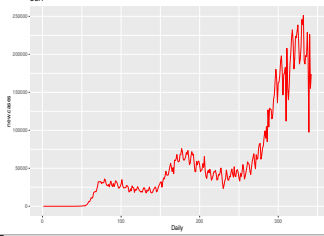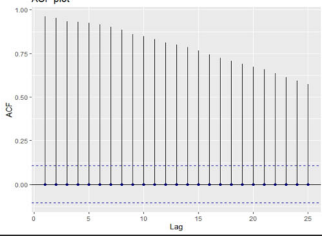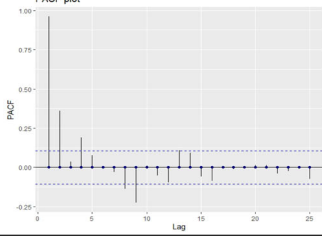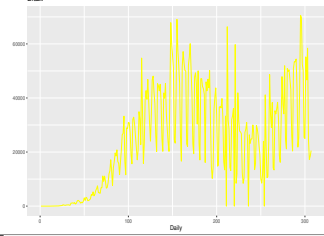$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2};$$
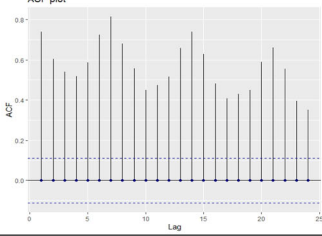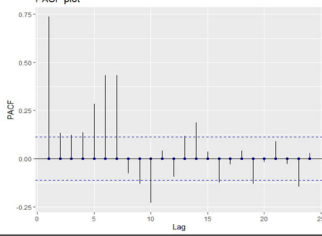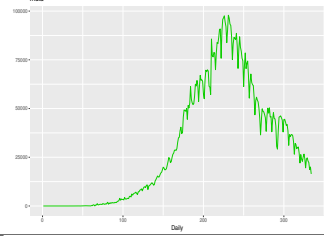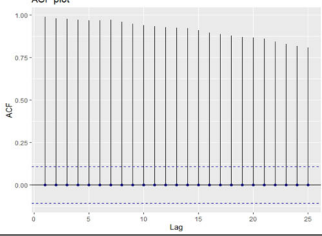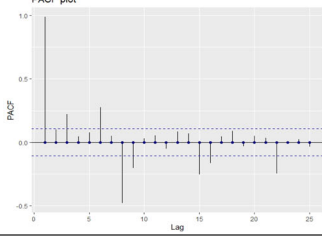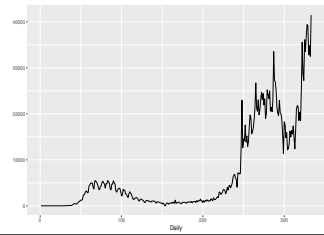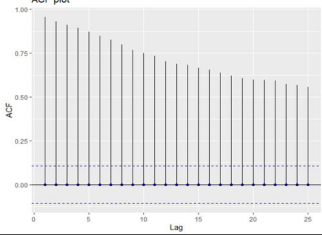
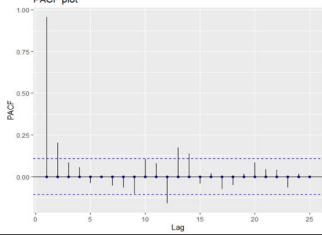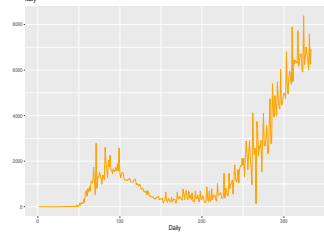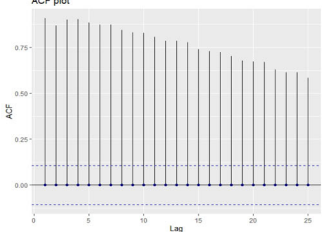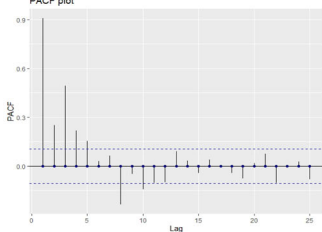$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n denotes the number of data points. MASE is computed using 'mase' function available under the 'Metrics' package in R statistical software. By definition, the lower the value of these performance metrics, the better is the performance of the concerned forecasting model.

4.3 Analysis of results

A schematic diagram is presented in Fig. 2 to give an outline of the models to be used in this section. We start the experimental evaluation for all datasets with the classical ARIMA(p, d, q) using '*forecast*' [64] statistical package in the R statistical software [65]. The nonlinear and non-seasonal time series of these countries were modeled with traditional and

Table 2 Training datasets (new daily cases) and corresponding ACF and PACF plots for all the countries considered in this study

Country	Training data	ACF plot	PACF plot
USA			
Brazil			
India			
UK			
Canada			

advanced methods such as ARIMA, Theta, WARIMA [66], ETS [67,68], TBATS, ARNN, and two popularly used hybrid models, namely hybrid ARIMA-ARNN [69], hybrid ARIMA-WARIMA [12], and the newly introduced hybrid Theta-ARNN (TARNN) models. For the accuracy of prediction, the data were partitioned into training and test sets, where the test set consisted of last 60 data points, and the training set included (inception - 60+1) data points. Table 3 gives the essential details about the functions and packages required for the implementation of standard individual models.
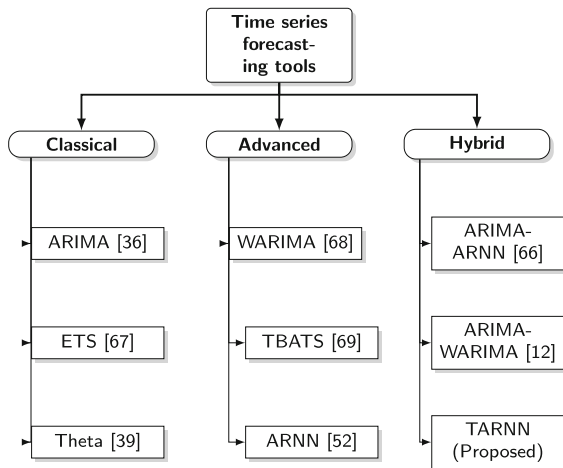
Fig. 2 Time series forecasting tools (available and proposed) used in this study

In the proposed TARNN model, linear modeling is conducted with Theta model using 'thetaf' function under the 'forecast' package in R statistical software. Nonlinear modeling with ARNN approach is carried out with 'caret' package using 'nnetar' function in R statistical software. Furthermore, Theta residuals are modeled with ARNN(p, k) model having a pre-defined Box-Cox transformation set $\lambda = 0$ to ensure the forecast values to stay positive. The values of p and k are obtained by training the network, which is a data-dependent approach as discussed in Sect. 2.2. Further, both the linear and nonlinear forecasts are summed up to obtain the final forecasts. Theta model was fitted to five datasets, namely the USA, Brazil, India, the UK, and Canada. Theta model residuals for these five countries were trained using ARNN(15, 8), ARNN(21, 11), ARNN(25, 13), ARNN(24, 12), ARNN(12, 6), and ARNN(9, 5) models with an average of 500 networks for all five datasets, each of which is a 15–8–1, 21–11–1, 25–13–1, 24–12–1, 12–6–1, and 9–5–1 networks with 137, 254, 352, 313, 85, and 56 weights and with 5,807,711, 309,816, 20,735, 883.6, 209,396, and 59,477 estimated σ^2, respectively. Finally, the predicted results of both Theta and ARNN models are added together to obtain the estimated forecasts of the proposed TARNN model. Rest of the hybrid models were compared in the similar fashion.

We compared our proposed TARNN model with traditional single models (ARIMA, TBATS, Theta, ARNN, WARIMA, ETS) along with hybrid ARIMA-

WARIMA model and hybrid ARIMA-ARNN model, and the experimental results are reported in Table 4. The performance of the proposed hybrid Theta-ARNN (TARNN) model is superior compared to all traditional individual and hybrid models on average. The theoretically proven asymptotic stationarity of the proposed hybrid model also suggests that the model cannot have a growing variance over time. The consistency and adequacy in the experimental results empirically approve the same. Thus, the efficacy of the proposed methodology of the proposed hybrid model is experimentally validated. Nevertheless, no model can have dominant advantage and this is related to no free lunch theorem in statistical learning [63]. Forecast of COVID-19 cases for March 20–29, 2021, is generated using proposed TARNN model and shown in Fig. 3.

5 Discussions

We now provide a brief discussion on some time series characteristics that will guide a potential modeler to opt for their hybrid scheme. A time series is usually classified into discrete or continuous, deterministic or stochastic, stationary or non-stationary, and linear or nonlinear. In real-life epidemiological forecasting problems, it is often challenging to determine whether a time series under study is generated from a linear or nonlinear underlying process. Since all the available pandemic datasets are complex and often contain both linear and nonlinear patterns, a single model may be insufficient to meet the whole data characteristic adequately. Thus, a hybrid model combining both the linear and nonlinear components is indeed useful. The basic assumption in the hybrid methodology is the additive relationship between the linear and nonlinear components of the time series. Thus, the proposed hybrid decision-making system is best suited for both stationary or non-stationary time series and situations when the data exhibit both linear and nonlinear patterns. In the development of the hybrid methodology, we need the component model to be sub-optimal, and it will be useful to combine individual forecasts based on different information sets to produce superior forecasts. Based on our experience on COVID-19 forecasting in this work, we comment that our proposal is best suited in situations where the datasets will have enough nonlinearity and non-stationarity. The experimental finding also suggests that the presence of the concept shift in

**Table 3** R functions and packages for standard forecasting model implementations

Model	R function	R package	References
ARIMA	Auto.arima	Forecast	[64]
ETS	Ets	Forecast	[64]
TBATS	Tbats	Forecast	[64]
Theta	Thetaf	Forecast	[64]
ARNN	Nnetar	Forecast	[64]
WARIMA	WaveletFittingarma	WaveletArima	[70]

Table 4 Quantitative measure of performances of forecasting methods on the COVID-19 test datasets for five countries

Methods	USA	UK	India	Canada	Brazil
MASE values					
ARIMA	3.55	5.53	0.67	2.08	1.55
Theta	3.24	5.88	1.89	2.08	1.31
WARIMA	4.97	6.46	5.26	1.8	1.53
ETS	2.47	5.62	1.17	1.65	1.86
TBATS	**2.46**	5.58	1.44	1.81	4.77
ARNN	3.29	5.67	3.16	2.23	2.83
ARIMA-ARNN	3.59	5.50	**0.71**	2.06	1.53
ARIMA-WARIMA	3.64	5.69	0.72	2.05	1.58
Proposed TARNN	3.27	**5.07**	1.60	**1.53**	**1.05**
RMSE values					
ARIMA	85,378.21	19,598.5	5325.44	2879.54	28,661.69
Theta	77,538.27	21,033.15	9159.85	2976.44	22,950.91
WARIMA	116,791.17	43,298.22	25,609.64	2764.52	28,139.06
ETS	60,473.05	20,017.92	6527.45	2425.78	33,242.02
TBATS	**604,72.43**	19,855.08	7262.95	2647.24	86,202.46
ARNN	79,694.81	16,796.84	14,859.34	3063.08	48,693.24
ARIMA-ARNN	86,626.36	19,557.8	**5323.38**	2876.86	27,825.87
ARIMA-WARIMA	86,938.13	20,249.21	5428.55	2845.54	28,908.41
Proposed TARNN	78,320.37	**16,559.04**	7854.66	**2323.95**	**20,423.69**
MAE values					
ARIMA	72,823.9	17,672.58	2833.76	2390.47	24,714.92
Theta	66,491.15	18,797.76	8024.95	2497.19	20,853.47
WARIMA	101,923.84	36,890.4	22,279.76	2069.01	24,294.93
ETS	50,537.67	17,953.84	4972.05	1893.77	29,506.42
TBATS	**50,533.39**	17,839.37	6091.65	2074.26	75,834.12
ARNN	67,409.69	14,941.82	13,389.13	2556.93	44,939.99
ARIMA-ARNN	73,690.5	17,585.56	**3023.91**	2386.66	24,259.25
ARIMA-WARIMA	74,637.97	18,187.05	3056.05	2353.5	25,162.57
Proposed TARNN	66,939.65	**14,770.08**	5487.11	**758.82**	**16,701.18**

Best results are made 'bold'. Proposed TARNN beats all the models for three out of five datasets and remains competitive for other two datasets

Fig. 3 Out-of-sample forecast of COVID-19 cases for March 20–29, 2021, using proposed TARNN model
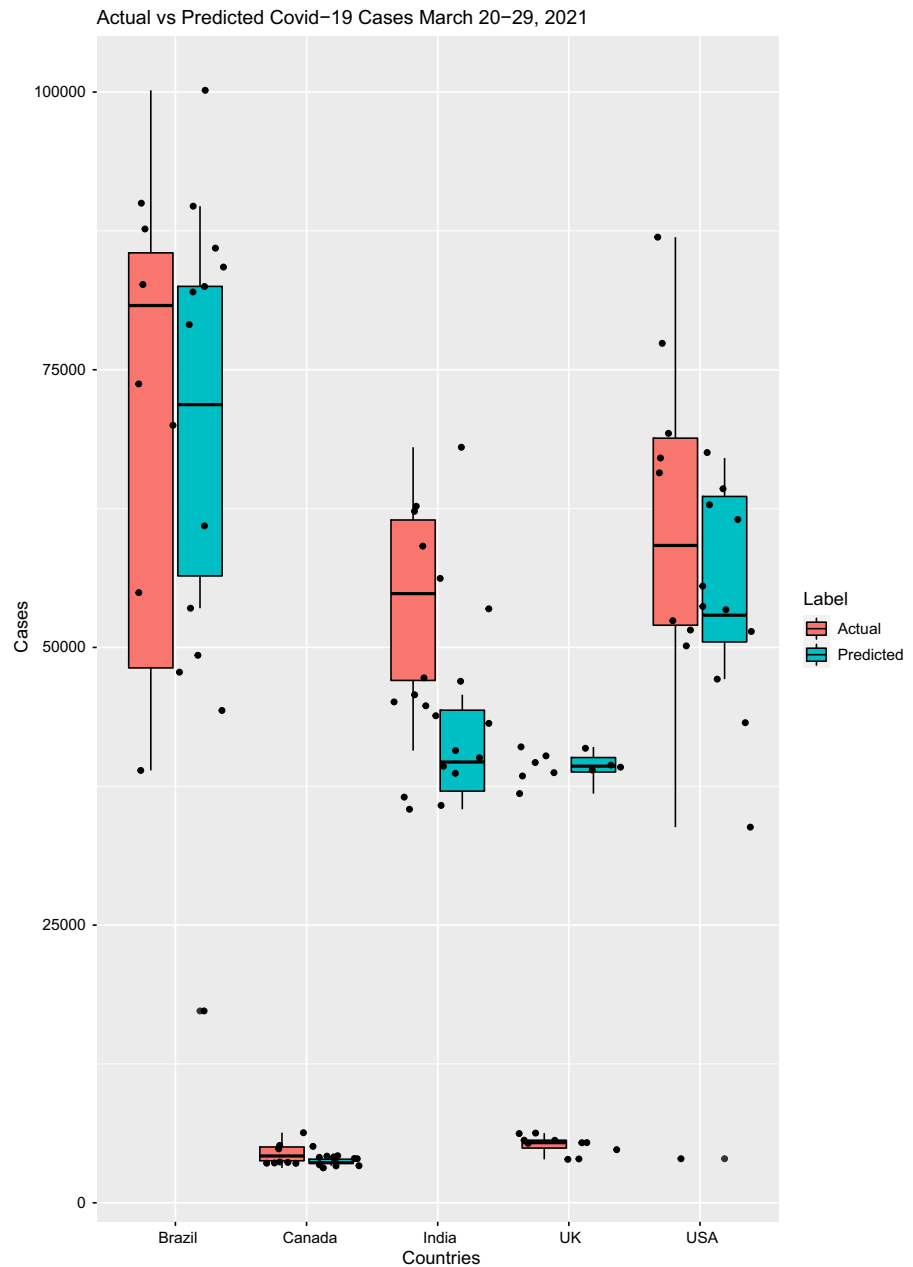
time series leads to inconsistency in the performance among different time series models (see, for example, all the hybrid models including our proposed one fail to capture the concept shift in the US data as reported in Table 4). Even in Fig. 3, the out-of-sample forecasts and actual values are worse for the UK data, whereas for other countries, we could capture the data irregularities quite well. Experimental finding also suggests the presence of the concept shift in time series data which

leads to inconsistency in the performance among different time series models for a few datasets.

Many parameters associated with COVID-19 transmission are still poorly understood. The resulting model uncertainty is not always calculated or reported in a standardized way. Once we can incorporate these variables, we can improve our estimates and update the TARNN model accordingly. Since purely statistical approaches do not account for how transmission

occurs, they are generally not well suited for long-term predictions about epidemiological dynamics (such as when the peak will occur and whether a resurgence will happen) or for inferences about intervention efficacy. Most forecasting models therefore limit their projections to one week or a few weeks ahead. Moreover, the problem of using confirmed cases to fit models is further complicated by the fact that the fraction of cases that are confirmed is spatially heterogeneous and time-varying. Amid enormous uncertainty about the future of the COVID-19 pandemic, the proposed TARNN model yields quantitative projections that policymakers may need in the short term to allocate resources or plan interventions. Guided by the short-term forecasts reported in this paper, the lockdown period can be adjusted accordingly and vaccination project can be enhanced. Since we presented a real-time forecast system unlike an ex post analysis, thus one can regularly update the actual confirmed cases and update the out-of-sample forecasts, just like it happens in temperature forecasting. To conclude, this model can further be easily extended for similar nonlinear and non-Gaussian forecasting problems arising 99in other applied domains.

6 Conclusions

In this study, we proposed a novel hybrid Theta-ARNN (TARNN) model using residual modeling approach that performs considerably well for confirmed cases of COVID-19 forecasting for the countries that include the ones with the highest number of cases: the USA, followed by Brazil, India, the UK, and Canada. The proposed TARNN model filters linearity using the Theta model and can better explain the linear, nonlinear, and non-stationary tendencies present in the selected COVID-19 datasets compared to the traditional single and hybrid models. It also yields better forecast accuracy than various traditional single and hybrid models for three out of five countries. The proposal will be useful in decision and policy making for government officials and policymakers to allocate adequate healthcare resources for the coming days in responding to the crisis. Time series of epidemics can oscillate heavily due to various epidemiological factors, and these fluctuations are challenging to be captured adequately for precise forecasting. This newly developed model can still predict with better accuracy provided the condi-

tions of asymptotic stationarity of the hybrid model are satisfied. This method can be used to update real-time forecasts as more data become available. The study covering multiple countries can be utilized without geographical borders and reflects the impact of social distancing, wearing masks, lock down, shutdown, quarantine, and sanitizing proper measures implemented by authorities. Prevalent techniques in the literature were unable to completely capture the nonlinear behavior of stochastic time series containing inherent random shock components. This new method has significant theoretical (established ergodicity and stationarity of the proposed TARNN process) as well as practical implications. Authorities and health care can modify their planning in stockpiles and hospital beds depending on these forecasts of the COVID-19 pandemic.

Declaration

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al.: Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet **395**(10223), 497–506 (2020)
2. Guan, W.J., Ni, Z.Y., Hu, Y., Liang, W.H., Ou, C.Q., He, J.X., Liu, L., Shan, H., Cl, Lei, Hui, D.S., et al.: Clinical characteristics of coronavirus disease 2019 in china. N. Engl. J. Med. **382**(18), 1708–1720 (2020)
3. Trilla, A., Trilla, G., Daer, C.: The 1918 Spanish flu in Spain. Clin. Infect. Dis. **47**(5), 668–673 (2008)
4. Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J., Yan, P., Chowell, G.: Real-time forecasts of the COVID-19 epidemic in china from February 5th to February 24th, 2020. Infect. Dis. Model. **5**, 256–263 (2020)
5. Organization, W.H., et al.: Who director-general's opening remarks at the media briefing on Covid-19-11 march 2020 (2020)
6. Boulos, M.N.K., Geraghty, E.M.: Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic

and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics (2020)

7. Li, Q., Feng, W., Quan, Y.H.: Trend and forecasting of the COVID-19 outbreak in China. J. Infect. **80**(4), 469–496 (2020)

8. Wu, J.T., Leung, K., Leung, G.M.: Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet **395**(10225), 689–697 (2020)

9. Fanelli, D., Piazza, F.: Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos, Solitons Fractals **134**, 109761 (2020)

10. Kucharski, A.J., Russell, T.W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., Sun, F., Jit, M., Munday, J.D., et al.: Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect. Dis. (2020)

11. Zhuang, Z., Cao, P., Zhao, S., Lou, Y., Wang, W., Yang, S., Yang, L., He, D.: Estimation of local novel coronavirus (COVID-19) cases in Wuhan, China from off-site reported cases and population flow data from different sources. medRxiv (2020)

12. Chakraborty, T., Ghosh, I.: Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis. Chaos Solitons Fractals 109850 (2020)

13. Perc, M., Gorišek Miksić, N., Slavinec, M., Stožer, A.: Forecasting COVID-19. Front. Phys. **8**, 127 (2020)

14. Petropoulos, F., Makridakis, S.: Forecasting the novel coronavirus COVID-19. PLoS ONE **15**(3), e0231236 (2020)

15. Grasselli, G., Pesenti, A., Cecconi, M.: Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. JAMA **323**(16), 1545–1546 (2020)

16. Grasselli, G., Zangrillo, A., Zanella, A., Antonelli, M., Cabrini, L., Castelli, A., Cereda, D., Coluccello, A., Foti, G., Fumagalli, R., et al.: Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy region, Italy. JAMA **323**(16), 1574–1581 (2020)

17. Emanuel, E.J., Persad, G., Upshur, R., Thome, B., Parker, M., Glickman, A., Zhang, C., Boyle, C., Smith, M., Phillips, J.P.: Fair allocation of scarce medical resources in the time of COVID-19 (2020)

18. Rosenbaum, L.: Facing COVID-19 in Italy-ethics, logistics, and therapeutics on the epidemic's front line. N. Engl. J. Med. **382**(20), 1873–1875 (2020)

19. Baud, D., Qi, X., Nielsen-Saines, K., Musso, D., Pomar, L., Favre, G.: Real estimates of mortality following covid-19 infection. Lancet Infect. Dis. (2020)

20. Ribeiro, M.H.D.M., da Silva, R.G., Mariani, V.C., dos Santos Coelho, L.: Short-term forecasting COVID-19 cumulative confirmed cases: perspectives for Brazil. Chaos Solitons Fractals 109853 (2020)

21. Ray, D., Salvatore, M., Bhattacharyya, R., Wang, L., Du, J., Mohammed, S., Purkayastha, S., Halder, A., Rix, A., Barker, D., et al.: Predictions, role of interventions and effects of a historic national lockdown in India's response to the COVID-19 pandemic: data science call to arms. Harv. Data Sci. Rev. **2020**(Suppl 1) (2020)

22. Conde-Gutiérrez, R., Colorado, D., Hernández-Bautista, S.: Comparison of an artificial neural network and Gompertz model for predicting the dynamics of deaths from COVID-19 in México. Nonlinear Dyn. 1–15 (2021)

23. Machado, J.T., Ma, J.: Nonlinear dynamics of COVID-19 pandemic: modeling, control, and future perspectives. Nonlinear Dyn. **101**(3), 1525–1526 (2020)

24. Han, C., Li, M., Haihambo, N., Babuna, P., Liu, Q., Zhao, X., Jaeger, C., Li, Y., Yang, S.: Mechanisms of recurrent outbreak of COVID-19: a model-based study. Nonlinear Dyn. 1–17 (2021)

25. Sujath, R., Chatterjee, J.M., Hassanien, A.E.: A machine learning forecasting model for COVID-19 pandemic in India. Stoch. Environ. Res. Risk Assess. 1 (2020)

26. Rustam, F., Reshi, A.A., Mehmood, A., Ullah, S., On, B.W., Aslam, W., Choi, G.S.: COVID-19 future forecasting using supervised machine learning models. IEEE Access **8**, 101489–101499 (2020)

27. Rohith, G., Devika, K.: Dynamics and control of COVID-19 pandemic with nonlinear incidence rates. Nonlinear Dyn. **101**(3), 2013–2026 (2020)

28. Péni, T., Csutak, B., Szederkényi, G., Röst, G.: Nonlinear model predictive control with logic constraints for COVID-19 management. Nonlinear Dyn. **102**(4), 1965–1986 (2020)

29. Bernoulli, D.: Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. Histoire de l'Acad., Roy. Sci.(Paris) avec Mem pp. 1–45 (1760)

30. Kermack, W.O., McKendrick, A.G.: Contributions to the mathematical theory of epidemics. II.-The problem of endemicity. Proc. R. Soc. Lond. Ser. A **138**(834), 55–83 (1932). Containing papers of a mathematical and physical character

31. Das, P., Nadim, S.S., Das, S., Das, P.: Dynamics of COVID-19 transmission with comorbidity: a data driven modelling based approach. Nonlinear Dyn. 1–15 (2021)

32. Huang, N.E., Qiao, F., Tung, K.K.: A data-driven model for predicting the course of COVID-19 epidemic with applications for China, Korea, Italy, Germany, Spain, UK and USA. medRxiv (2020)

33. Quaranta, G., Formica, G., Machado, J.T., Lacarbonara, W., Masri, S.F.: Understanding COVID-19 nonlinear multi-scale dynamic spreading in Italy. Nonlinear Dyn. **101**(3), 1583–1619 (2020)

34. Ioannidis, J.P., Cripps, S., Tanner, M.A.: Forecasting for COVID-19 has failed. Int. J. Forecast. (2020)

35. Holmdahl, I., Buckee, C.: Wrong but useful-what COVID-19 epidemiologic models can and cannot tell us. N. Engl. J. Med. (2020)

36. Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: Time Series Analysis: Forecasting and Control. Wiley, Hoboken (2015)

37. Conejo, A.J., Plazas, M.A., Espinola, R., Molina, A.B.: Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans. Power Syst. **20**(2), 1035–1042 (2005)

38. Hyndman, R.J., Billah, B.: Unmasking the theta method. Int. J. Forecast. **19**(2), 287–290 (2003)

39. Assimakopoulos, V., Nikolopoulos, K.: The theta model: a decomposition approach to forecasting. Int. J. Forecast. **16**(4), 521–530 (2000)

40. Zhang, G.P.: Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing **50**, 159–175 (2003)

41. Cadenas, E., Rivera, W.: Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. Renew. Energy **35**(12), 2732–2738 (2010)

42. Milačić, L., Jović, S., Vujović, T., Miljković, J.: Application of artificial neural network with extreme learning machine for economic growth estimation. Phys. A **465**, 285–288 (2017)

43. Maleki, A., Nasseri, S., Aminabad, M.S., Hadi, M.: Comparison of ARIMA and NNAR models for forecasting water treatment plant's influent characteristics. KSCE J. Civ. Eng. **22**(9), 3233–3245 (2018)

44. Ghosh, I., Chakraborty, T.: An integrated deterministic—stochastic approach for forecasting the long-term trajectories of COVID-19. Int. J. Model. Simul. Sci. Comput. 2141001 (2021)

45. Pasini, A.: Artificial neural networks for small dataset analysis. J. Thorac. Dis. **7**(5), 953 (2015)

46. Wynants, L., Van Calster, B., Bonten, M.M., Collins, G.S., Debray, T.P., De Vos, M., Haller, M.C., Heinze, G., Moons, K.G., Riley, R.D., et al.: Prediction models for diagnosis and prognosis of COVID-19 infection: systematic review and critical appraisal. BMJ **369** (2020)

47. Oliveira, M.R., Torgo, L.: Ensembles for time series forecasting. J. Mach. Learn. Res. **39**, 360–370 (2014)

48. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley, Hoboken (2004)

49. Percival, D.B., Walden, A.T.: Spectral Analysis for Univariate Time Series, vol. 51. Cambridge University Press, Cambridge (2020)

50. Spiliotis, E., Assimakopoulos, V., Makridakis, S.: Generalizing the theta method for automatic forecasting. Eur. J. Oper. Res. **284**(2), 550–558 (2020)

51. Zhang, G., Patuwo, B.E., Hu, M.Y.: Forecasting with artificial neural networks: the state of the art. Int. J. Forecast. **14**(1), 35–62 (1998)

52. Faraway, J., Chatfield, C.: Time series forecasting with neural networks: a comparative study using the air line data. J. Roy. Stat. Soc.: Ser. C (Appl. Stat.) **47**(2), 231–250 (1998)

53. Hyndman, R.J., Athanasopoulos, G.: Forecasting: principles and practice. OTexts (2018)

54. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation. California Univ San Diego La Jolla Inst for Cognitive Science, Tech. rep (1985)

55. Mosleh, A., Apostolakis, G.: The assessment of probability distributions from expert opinions with an application to seismic fragility curves. Risk Anal. **6**(4), 447–461 (1986)

56. Chan, K.S., Tong, H.: On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. Adv. Appl. Probab. **17**(3), 666–678 (1985)

57. Brockwell, P.J., Lindner, A.: Strictly stationary solutions of autoregressive moving average equations. Biometrika **97**(3), 765–772 (2010)

58. Trapletti, A., Leisch, F., Hornik, K.: Stationary and integrated autoregressive neural network processes. Neural Comput. **12**(10), 2427–2450 (2000)

59. Chakraborty, T., Chakraborty, A.K., Biswas, M., Banerjee, S., Bhattacharya, S.: Unemployment rate forecasting: a hybrid approach. Comput. Econ. **57**(1), 183–201 (2021)

60. Tweedie, R.: The existence of moments for stationary Markov chains. J. Appl. Probab. **20**(1), 191–196 (1983)

61. Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer, Berlin (1993)

62. Hyndman, R.J., Koehler, A.B.: Another look at measures of forecast accuracy. Int. J. Forecast. **22**(4), 679–688 (2006)

63. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, Berlin (2009)

64. Hyndman, R.J., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E.: Package 'forecast'. https://cran.r-project.org/web/packages/forecast/forecast.pdf (2020)

65. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2017). https://www.R-project.org/

66. Percival, D.B., Walden, A.T.: Wavelet Methods for Time Series Analysis, vol. 4. Cambridge University Press, Cambridge (2000)

67. Hyndman, R., Koehler, A.B., Ord, J.K., Snyder, R.D.: Forecasting with Exponential Smoothing: The State Space Approach. Springer, Berlin (2008)

68. De Livera, A.M., Hyndman, R.J., Snyder, R.D.: Forecasting time series with complex seasonal patterns using exponential smoothing. J. Am. Stat. Assoc. **106**(496), 1513–1527 (2011)

69. Chakraborty, T., Chattopadhyay, S., Ghosh, I.: Forecasting dengue epidemics using a hybrid methodology. Phys. A Stat. Mech. Appl. 121266 (2019)

70. Paul, R.K., Samanta, S., Paul, M.R.K., LazyData, T.: Package 'waveletarima'. Seed **500**, 1–5 (2017)