



Hybridized wavelet neuronal learning-based modelling to predict novel COVID-19 effects in India and USA

Rashmi Bhardwaj^{1,a} and Aashima Bangia²

¹ University School of Basic and Applied Sciences (USBAS), Non-Linear Dynamics Research Lab, GGS Indraprastha University, Delhi, India

² USBAS, GGS Indraprastha University, Delhi, India

Received 26 November 2021 / Accepted 5 March 2022 / Published online 18 March 2022

© The Author(s), under exclusive licence to EDP Sciences, Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract Coronavirus disease so called as COVID-19 is an infectious disease and its spread takes place due to human interaction by their pathogen materials during coughing and sneezing. COVID-19 is basically a respiratory disease as evidence proved that a large number of infected people died due to short breathing. Most widely and uncontrollably spreading unknown viral genome infecting people worldwide was announced to be 2019–2020 *nCoV* by WHO on January 30, 2020. Based on the seriousness of its spread and unavailability of vaccination or any form of treatment, it was an immediate health-emergency of concern of international-level. The paper analyses effects of this virus in countries, such as India and United States on day-to-day basis because of their greater variability. In this study, various performance measures, such as root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), mean absolute standard error (MASE) and mean absolute percentage error (MAPE) which characterize models' performances. R^2 value has been achieved to be closest to 1, i.e., 0.999 from Wavelet Neuronal Network Fuzzified Inferences' Layered Multivariate Adaptive Regression Spline (WNNFIL–MARS) for both the countries' data. It is important to capture the essence of this pandemic affecting millions of the population daily ever since its spread began from January, 2020.

1 Introduction

Corona-genomic viruses that are member of the *Coronaviridae* species have viruses consisting of large strands of RNA sequences. The RNA is sized about 27–32 kb engulfed with polyadenylated. In each cluster, viruses pigeon-hole their way into host cells genomic sequence. This so-called 'Corona' virus got first identified in animals such as mice, horses and many reptiles as well as cows, bullocks etc. that led to severe infections some of which gastric, respiratory tract ones. Predominant infections associated with this virus have been 'respiratory', 'gastric' however, 'hepatic' or 'neurological diseases' recorded.

Till date as per various studies reported the structure has been described to be as engulfed non-segmented positively oriented RNA-genomic viral classifying into the category *Corona-viridae* on the bases of characteristics. There have been a lot of research into the RNA genome of this virus. One of the key outcomes involve that its effect on the host RNA is based on individual-to-individual response which means coronavirus infections could be mild categorized as *Beta coronaviruses*.

Studied multifarious diseases regarding various characteristics such as usages, action-mechanisms and various others towards treatment of pandemic [1]. Exploration of statistical simulations to study the dynamics of HIV, estimation of river water quality through artificial intelligence, COVID-19 pandemic, complexity analysis for China region modelled to study in detail [2–6]. Studied human corona via glycoproteins in receptor-binding sites [7]. Explained time series transformations required in analysis [8]. Modelled the phases of transmission in which novel coronavirus gets transferred [9]. Discussed a dynamic model indicating transfer through hands, foots, mouths etc. [10]. Analyzed the role of 'Chloroquine' during *SARS-CoV-2* [11]. Discussed an epidemiology model for *HCV-infections* [12]. Randomness via fuzzy random set theory modelled for Plasma disruption in fusion reactors [13]. Detailed explanation for *corona-theorem* with the applicability into spectral-problems [14]. Discussed the newest variant, Omicron [15]. Explored the COVID-19 herd immunity [16]. Studied respiratory-tract in human beings for novel virus isolation [17]. Explained in detail the coronavirus related concerns [18]. Explained the AIDS mechanism so as to study viral diseases [19]. Studied as well as explained the second and third waves of coronavirus [20].

^a e-mails: rashmib@ipu.ac.in; rashmib22@gmail.com (corresponding author)

Discussed the effects of coronavirus in lower respiratory-tracts observed in infants [21]. Detailed study of Corona to understand the pneumonia infection in low immunity-patients [22]. Discussed ‘*Corona-theorem*’ for ‘*countably-many functions*’ [23]. Pathogenesis of COVID-19 epidemiology in Saudi discussed with the help of nonlinear time series [24]. WHO reported ‘*novel-coronavirus*’. Survey on ‘*MERS-CoV*’ updates [25, 26]. Studied the respiratory syndrome during human coronavirus via molecule analysis [27]. WHO stated updates on *Coronavirus* around January, 2020 along with the discussion about the ‘*pneumonia-outbreak*’ originating from bats developing into ‘*corona-virus*’. Discussed important antibiotics that would hinder the entrance of ‘*Ebola virus*’, ‘*MERS-CoV*’, ‘*SARS-CoV*’ in human-beings as well as other living beings. WHO described ‘*novel coronavirus*’ updated reports [28–31].

None of the authors designed the wavelet conjuncted neuronal network fuzzified inferences’ layered hybrid model for countrywise daily data COVID-19 genome transmission. In this study, forecasts of the datewise data of confirmed cases. In our forecasting models, we have made an attempt of artificial intelligence (AI) based learned algorithms, such as *WNNFIL*, *WNNFIL-LSSVR*, *WNNFIL-MARS*. Forecasts via validated data for longer time spans have been evaluated on the basis of error performances through RMSE, MAE, R^2 .

2 Chaotic analysis

2.1 Lyapunov characteristic exponent (LCEs)

Alexander M. Lyapunov, a Russian engineer first discussed the concept of characteristic exponents hence the name Lyapunov Characteristic Exponents (LCEs) became widely known. It is used to measure the stability of the variables considered under the system. Discussed the degree of strangeness, i.e., randomness that led to strange attractors’ formation. Such exponents have been observed to be the direct indicators as well as quantifiers for determining randomness. The number of Lyapunov exponents depend upon the dimension of the phase diagram. The LCEs < 0 indicates nearby initial conditionalities converging, i.e., delta-errors decreasing through delta-time increment. However, whenever the LCEs are positive, the infinitesimally nearby initial conditions diverging exponentially, i.e., errors increasing rapidly with respect to time increments. Such a phenomena is known as the sensitivity towards initial conditions famously known as *Chaos*. LCEs are considered as the standardizations for eigenvalues computed in various states, such as steady-state, limit-cycles or randomness. Determination of these characteristics for non-linear systems involves numerically integrating underlying differential equalities over the variation in parameters and initial conditions.

Steps to compute Lyapunov Exponents:

1. Consider $x_n = f_n(x_0)$; $y_n = f_n(y_0)$ then, the n th iteration of orbits for x_0 and y_0 under f are: $|x_0 - y_0| \ll 1, |x_n - y_n| \ll 1$,
2. Then, the divergence at that n th iteration in case of 1-dimension setup is calculated as the summation of

$$|x_n - y_n| \approx \left(\prod_{t=0}^{n-1} |f'(t)| |x_0 - y_0| \right)$$

having $|x_0 - y_0| \ll 1, |x_n - y_n| \ll 1$; $x_n = f_n(x_0), y_n = f_n(y_0)$ are n th iterations of orbits of x_0 and y_0 under f .

3. Next, exponential rate of divergence, i.e., $\log|f'(x)|$ for two localized initial conditions formulating into

$$\lambda(x_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\prod_{t=0}^{n-1} |f'(x_t)| \right)$$

$$\text{where } \prod_{t=0}^{n-1} |f'(x_t)| \approx e^{\lambda(x_0)n} \text{ for } , n \gg 1.$$

4. Hence, $\lambda(x_0)$ can be defined as LCEs for the trajectory of x_0 .
5. Thus, two trajectories in phasespace having initial-separation, δx_0 diverge $|\delta x(t)| \approx e^{\lambda t} |\delta x(0)|$ where $\lambda > 0$ denotes LCE.
6. Consider $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigen-values for the linearized-equation $\frac{du}{dt} = A(u^*)$ s.t. $m_1(t) = e^{\lambda t}$ and $\tilde{\lambda}_i = \lim_{t \rightarrow \infty} \frac{1}{t} \ln |e^{\lambda_i t}| = \text{Re}[\lambda_i]$.

Thus, LCEs can be taken as equivalent to real parts of calculated eigenvalues at all of the fixed points. Whenever these exponents become less than zero, the localized initial conditions become convergent depicting (small) delta-errors are decreasing. The subset formed is also known as attracting set. Therefore, attractor can be defined as an attracting set that consists of a dense orbit. Whereas, whenever the LCEs become positive, the initial conditions become divergent, i.e., (small) delta-errors are increasing. Behavior of attractors as per LCE values being classified into

- Equilibria: $0 > \lambda_1 \geq \lambda_2 \dots \geq \lambda_n$;
- Periodic limit-cycling: $\lambda_1 = 0, 0 > \lambda_2 \geq \lambda_3 \dots \geq \lambda_n$;
- k -periodic limit-cycling: $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0, 0 > \lambda_{k+1} \geq \lambda_{k+2} \dots \geq \lambda_{k+n}$;
- Strange-randomness, i.e., chaotic: $\lambda_1 > 0, \sum_{i=1}^n \lambda_i < 0$;
- Hyper-randomness, i.e., hyper-chaotic: $\lambda_1 > 0, \lambda_2 > 0, \sum_{i=1}^n \lambda_i < 0$.

2.2 Rescaled R/S analysis

As proposed by *H.E. Hurst*, Rescaled-range analysis (R/S analysis) involves the understanding of long-term persistence occurring in the time successions. It begins via splitting of timeseries, s_i assuming non-overlapping

sections ϕ having magnitude, p resulting in $K_p = \text{int}(K/p)$ segments altogether. Next step, this integrated data is calculated in every section as

$$\begin{aligned} X_\phi(i) &= \sum_{t=1}^i (s_{\phi p+t} - \langle s_{\phi p+t} \rangle_p) \\ &= \sum_{t=1}^i s_{\phi p+t} - \frac{i}{p} \sum_{t=1}^p s_{\phi p+t}; \\ \phi &= 0, \dots, K_p - 1 \end{aligned}$$

Now, piecewise constants' trend and various other averaging can be handled through subtraction. Next step involves recording ranges through maximum and minimum values and computing standard deviations in each section considered above denoted as

$$\begin{aligned} R_\phi(p) &= \max_{i=1}^p X_\phi(i) - \min_{i=1}^p X_\phi(i), \\ S_\phi(p) &= \sqrt{\frac{1}{p} \sum_{i=1}^p X_\phi^2(i)} \\ p(k) &= Ck^{-\alpha}; \\ C &\text{ - constant; } p(k) \text{ - ACF with lag 'k'} \\ H &= 1 - \frac{\alpha}{2} \end{aligned}$$

Lastly, as the name suggests range needs to be rescaled over all sections to capture the fluctuation in function form, $F_{R/S}$ as

$$F_{R/S}(p) = \frac{1}{K_p} \sum_{\phi=0}^{K_p-1} \frac{R_\phi(p)}{S_\phi(p)} \sim p^{H_e}, \text{ for } p \gg 1$$

having H_e —Hurst exponent as desired.

In addition, H_e calculated from slope of the regression fitline plotted for the log–log curve of the variable with respect to time.

Remark 3 From Remarks 1 and 2, H_e can be calculated in terms of multi-fractal & spectral analysis as: $2H_e \approx 1 + \eta = 2 - \varpi$.

Remark 4 Clearly, whenever $0 < \varpi < 1$, then $0.5 < H_e < 1$. In general, this relation may not hold for all multifractal analysis.

As and when the data is stationary, H_e using rescaled R/S. Thus, $H_e < 1/2$ specify long-term anti-correlation among data, and $H_e > 1/2$ shows positive correlations. In addition, power-law correlations decay sooner than $1/p$, taking $H = 1/2$ for larger p values.

3 Machine learning-based forecast models

Variable considered with respect to time successions' corpora that needs to be put together for real-time

response variables' approximation towards *nCOVID-19*. Furthermore, it is taken as mentioned by the designated authorities that these are the daily-confirmed cases nearly from last week of January for the study of variations in time-periods in accordance with various data sets for simulations. Everyday active cases data for United States and India. For this forecast model case study, datewise data of Confirmed cases recorded in India & USA is taken as available on the designated authorities' reports and bulletins.

The unprecedented and unforeseen *nCov-19* spread had its own course world over. Both the countries studied in this article had their own different timelines for the spread of outbreak. The data sets of USA seemed to have greater variability as compared to that of India. Therefore, both the countries' confirmed cases became important towards the study of this pandemic. Uncertainty associated with the available data invites a model that can forecast the spread of an infectious disease. To understand the dynamics of the disease including its mode of transmission, we need to have intelligent forecasting models with respect to realistic human data (confirmed, deaths and recovered cases). Accordingly, we have developed innovative forecasting models for the country-based datewise data of confirmed cases.

3.1 Discrete waveform signal analysis (DWS)

Theorem Sequence consisting of embedded-approximation-subsets to be taken as

$$0 \leftarrow \dots V_{j-i} \subset V_j \subset V_{j+1} \dots \rightarrow R$$

having $f(x) \in V_j - f(2x) \in V_{j+1}$

$$\begin{aligned} f(t) \in V_0 &\Leftrightarrow f(t - k) \in V_0, k \in Z \\ \varphi(t - k)_{k \in Z} &\text{ forming an orthonormal basis} \end{aligned}$$

plus, the sequence of orthogonal complements, details' subspaces:

$$\begin{aligned} W_n \text{ s.t. } W_{n+1} &= W_n \oplus W_n. \\ \varphi &\text{ - scaling function, i.e., a low-pass sieve.} \end{aligned}$$

Remark Basis, $V_{\{j\}}$ to be understood as

$$\varphi_{j,k}(t) = 2^{j/2} \varphi(2^j t - k), k \in Z$$

For non-stationary time-dependent series, detrended fractal investigation of time successions can be easily implemented by considering wavelet coefficients of the type $X_\phi(i)$. In this case, the convolution corresponds to the basic arithmetic such as addition/subtraction for mean values of $X_\phi(i)$ within sections having magnitude, p .

The wavelet coefficients $W_\varphi(\kappa, p)$ are based on both time position κ & scaling, p . The waveform transforms required for the data signals, d_s is calculated through

the following summation for discrete waveform signals of the time series, $s_i, i = 1, \dots, K$:

$$W_\varphi(\kappa, p) = \frac{1}{p} \int_{-\infty}^{\infty} d_s(t) \varphi[(t - \kappa)/p] dt, \text{ (continuous)}$$

$$= \frac{1}{p} \sum_{i=1}^K d_s(i) \varphi[(t - \kappa)/p], \text{ (discrete)}$$

In addition, the local frequency decomposition of the signal is described by the time resolution apt for the considered frequency $f = 1/p$.

3.2 Least squares support vector regression (LSSVR)

LS-SVRs handle higher complicacies as compared to SVR as described by Suykens. Objective-function does not change much as compared to that of the classical one. Difference arises when ε -based loss function replaces the classical squared-loss function. It has been observed that regression through this approach annihilates noise alongwith monitoring of computational-labour.

Mathematically, it can be understood as

$$\phi_2(\omega, h, \varepsilon) = \gamma A_W + \zeta A_D$$

with

$$A_W = \frac{1}{2} \omega^T \omega \text{ and}$$

$$A_D = \frac{1}{2} \sum_{i=1}^K \varepsilon_i^2$$

$$= \frac{1}{2} \sum_{i=1}^K (y_i - (\omega^T \sigma(x_i) + h))^2$$

where γ, ζ - hyper-parameters tuning with respect to SSE

Hence, the LSSVR regressor resultant that can be achieved from the Lagrangian-function as followed:

$$\left\{ \begin{array}{l} L_m(\omega, h, \varepsilon, \alpha) = \phi_2(\omega, \varepsilon) \\ \quad - \sum_{i=1}^K \alpha_i \{ [\omega^T \sigma(x_i) + h] + \varepsilon_i - y_i \} \\ = \frac{1}{2} \omega^T \omega + \frac{\mu}{2} \sum_{i=1}^K \varepsilon_i^2 \\ \quad - \sum_{i=1}^K \psi_i \{ [\omega^T \sigma(x_i) + h] + \varepsilon_i - y_i \} \end{array} \right\}$$

where $\psi_i \in \mathbb{R}$ as Lagrange multipliers.

3.3 Wavelet neuronal network fuzzified inferences' layered least-squares support vector regression (WNNFIL-LSSVR)

Beginning with the NNFIL model as follows, Fig. 1:

The following theorem and lemma for the conjuncted model are proposed as:

Theorem *The simulation of the responses obtained after training through DWTS into approximations and details trained through NNFIL.*

Lemma *The results are then transferred towards least squares support vector regressors' model having 'Gaussian function' as kernel forming an important part for the formulation of optimization-based solution.*

Thus, the flowchart in Fig. 2 clearly puts forward every step involved in the proposed model.

3.4 Multivariate adaptive regression spline (MARS)

Conceptually, MARS is the process of creating intricate pattern of correlations among the multi-variables known as the predictor-response variables not considering or taking into account the already known knowledge of relations that need to be hypothesized.

Thus, MARS build this model of single and multi-response data then validated and tested using separate sets. Finally, then the model can be utilized for estimation. The generalized-form MARS given as in the equation below:

$$y = f(x) = \beta_0 + \sum_{m=1}^M \beta_m H_{km}(x_{v(k,m)})$$

y - output parameter;
 β_0 - constant;
 M - no. of basis-function;
 $H_{km}(x_{v(k,m)}) - m^{th}$ basis function;
 β_m - Corresponding coefficient of $H_{km}(x_{v(k,m)})$.
 $x_{v(k,m)}$ - predictor in the k^{th} of m^{th} product.

The basis-function, H is defined as

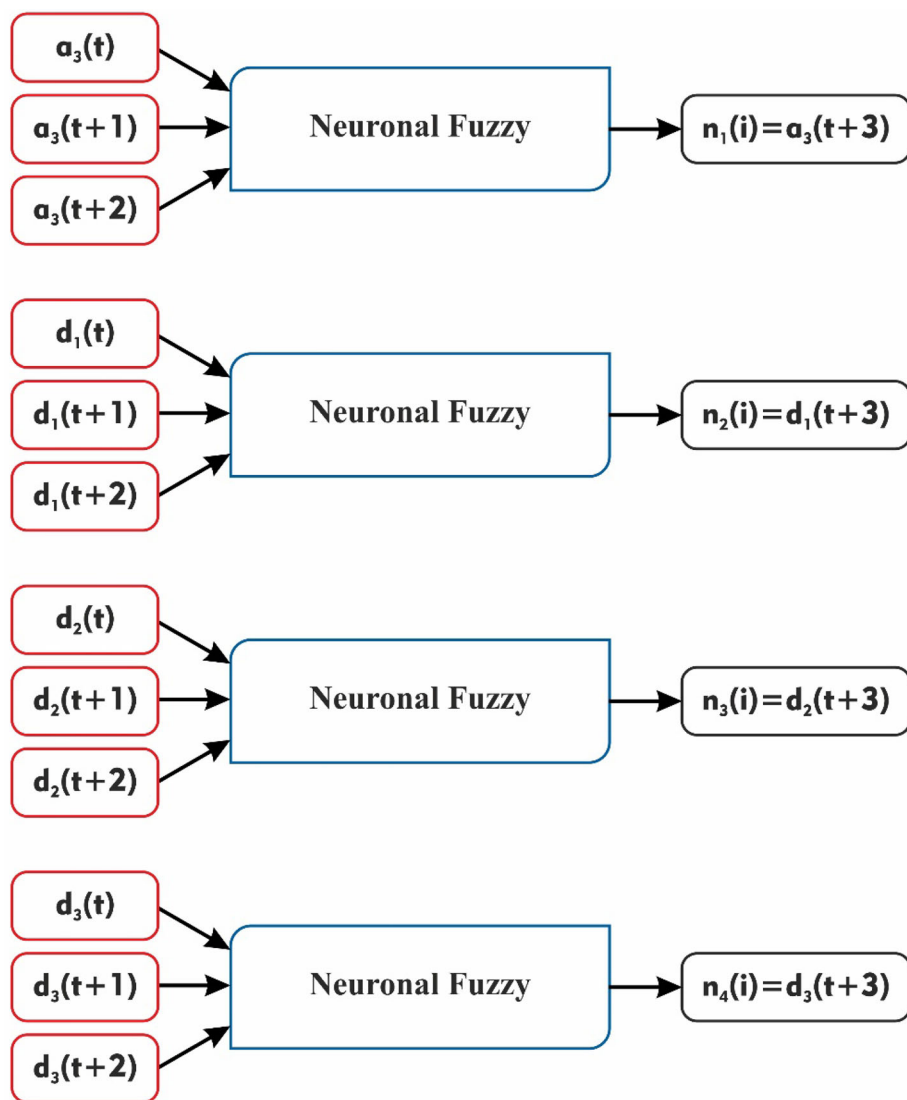
$$H_{km} = \prod_{k=1}^k h_{km}$$

For the order $K = 1$, model becomes additive, while in case of $K = 2$, model becomes pairwise interactive

3.5 Wavelet neuronal network fuzzified inferences' layered multivariate adaptive regression spline (WNNFIL-MARS)

The following theorem and lemma for the conjuncted model are proposed as:

Fig. 1 Working of the NNFIL conjuncted model



Theorem Simulation of the resultant values achieved from DWT decomposition into cAs and cDs trained through NNFIL.

Lemma Thus, results are then transferred towards MARS technique having each feature, f mined leading to the formation of Adaptive Regression Splines.

Thus, the flowchart of Fig. 3 demonstrates every step involved in this proposed model.

4 Performance measures

The estimation provided from the conjuncted models is evaluated on the basis of following performances, such as ‘root mean squared error (RMSE)’, ‘mean absolute error (MAE)’, ‘coefficient of determination (R^2)’:

$$(i) \quad \text{MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$$

having y_i is the recorded data; \hat{y}_i is the predicted assessment; n is the number of days in prediction. The square root for MSE is referred as ‘root mean squared error’:

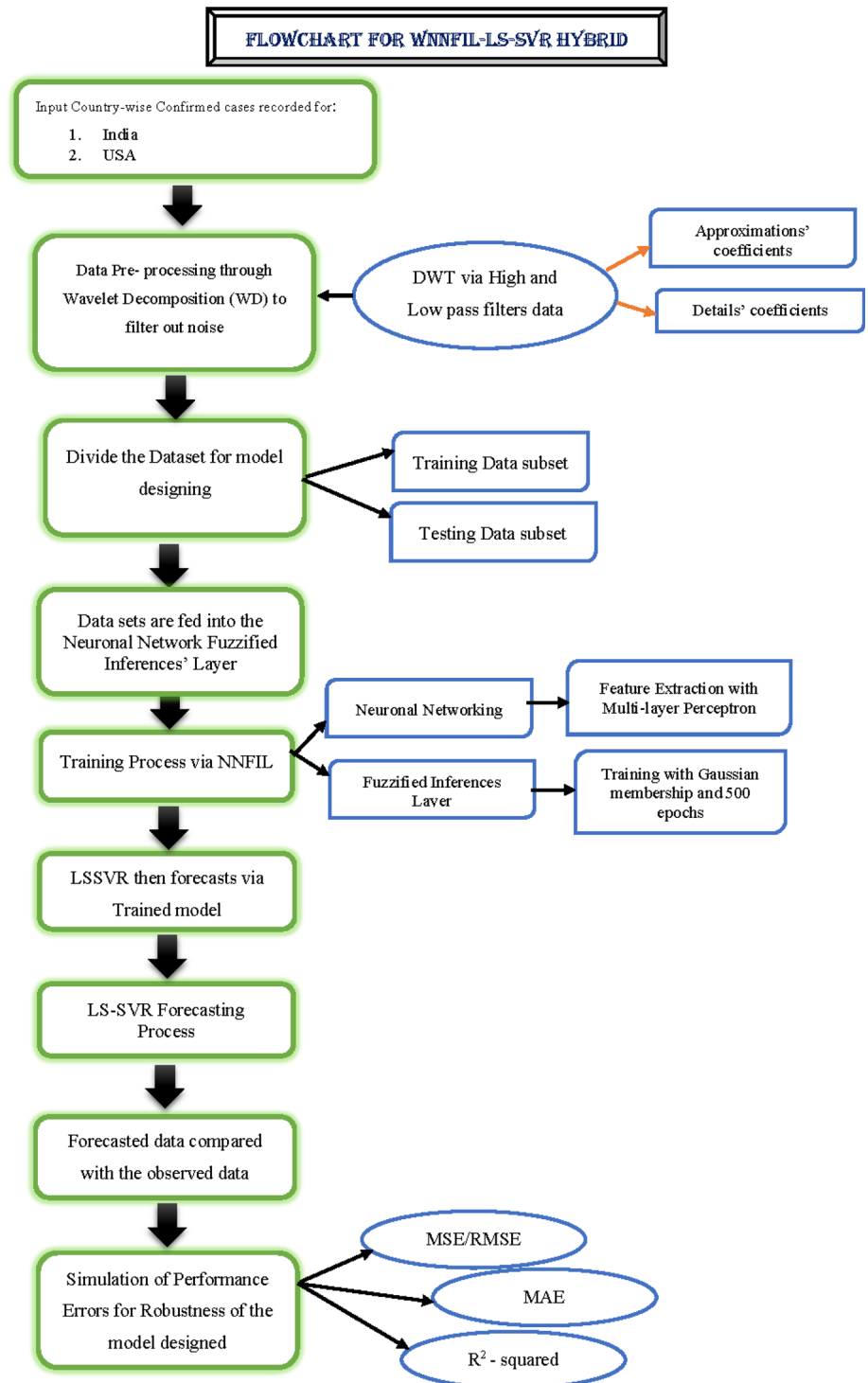
$$(ii) \quad \text{RMSE} = \sqrt{\text{MSE}}$$

$$(iii) \quad R^2 = 1 - \frac{\text{SSE}}{\text{SST}}$$

having $\text{SSE} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$; $\text{SST} = \sum_{i=1}^n (y_i - \bar{y}_i)^2$,

$$(iv) \quad \text{MAE} = \frac{1}{N} \left(\sum_{i=1}^N |y_i - \hat{y}_i| \right),$$

Fig. 2 Flowchart of WNNFIL–LSSVR hybrid model for COVID-19



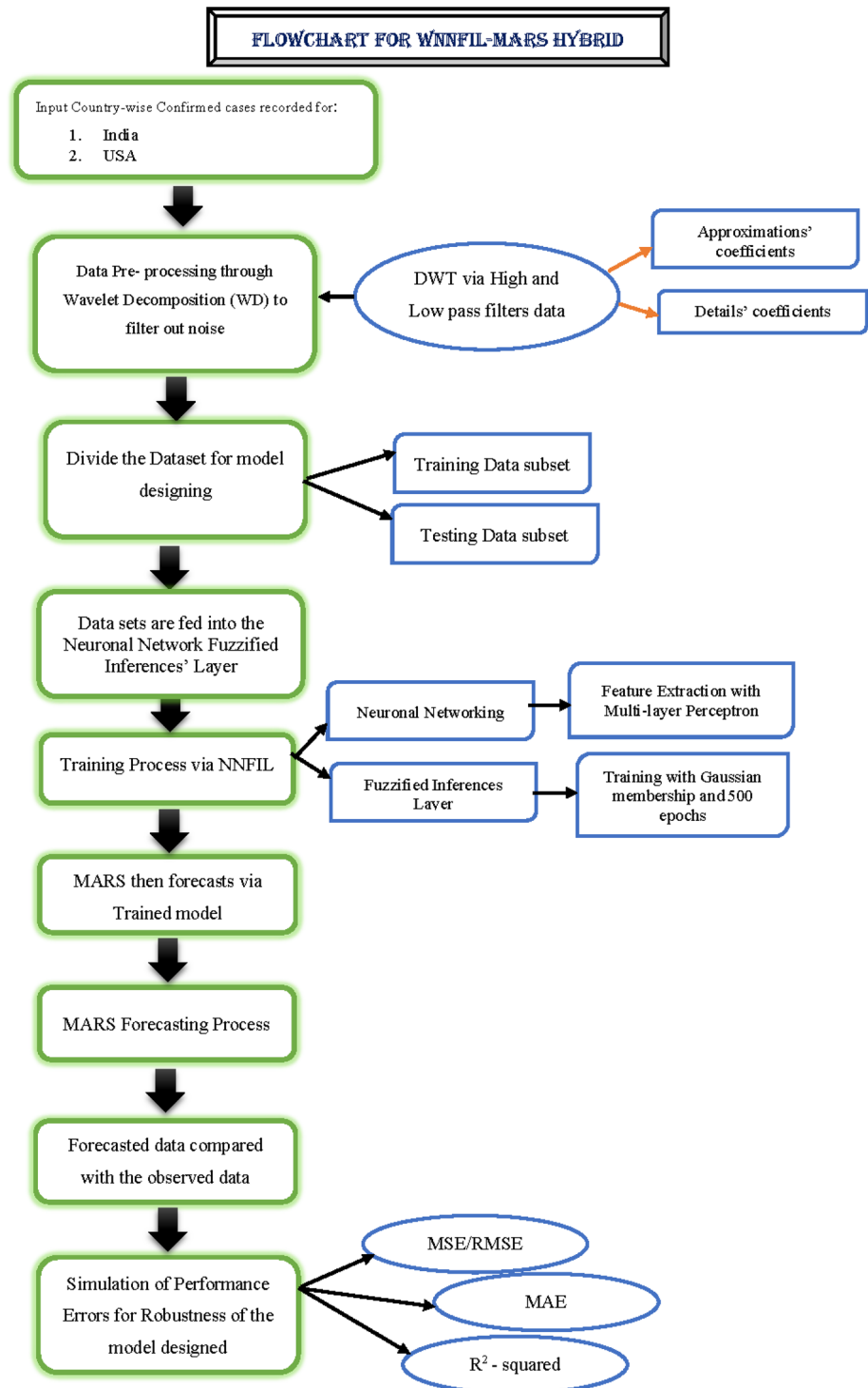
having y_i is the actual datavalues, \hat{y}_i is the predicted assessment, n is the number of days in prediction.

5 Discussion of results

The results for the confirmed cases arising in USA and India have been simulated provided by design-

ated sources. Figure 4 shows linear fitting of WNNFIL trained model for USA data set. Model trained values increase at 60th day and goes upto 80th day after which it almost remains stabilized for the rest of the days. Whereas linear regression plots a straight increasing line leaving out all the data values. Figure 5 depicts linear regression of WNNFIL–LSSVR trained model for USA data set. Model trained values remain stabilized till 60 days the increases and goes upto 80th

Fig. 3 Flowchart of WNNFIL–MARS hybrid model for COVID-19



day after which it almost remains stabilized for the rest of the days. Whereas linear regression plots a straight increasing line leaving out all the data values, Fig. 6 illustrates linear regression of WNNFIL–MARS trained model for USA data set. Figure 7 shows that Lyapunov exponents for the confirmed cases in USA. Whenever Lyapunov exponents can be seen < 0 , it showed that

the nearby initial conditions (ICs) converging to one another leading towards small epsilon-errors decreasing with respect to time. For any of the Lyapunov characteristics > 0 (+ve), it showed infinitesimally nearby initial conditions(ICs) diverging from one another exponentially fast implying that the errors in initial conditions would rise with respect to time. As and when tra-

Fig. 4 Linear regression fit for WNNFIL trained model for USA

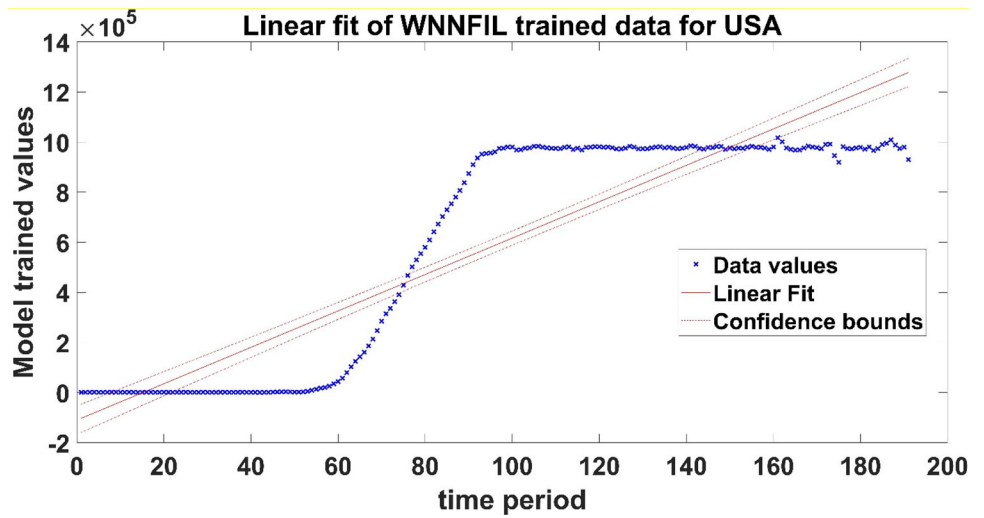


Fig. 5 Linear regression fit for WNNFIL-LSSVR trained model for USA

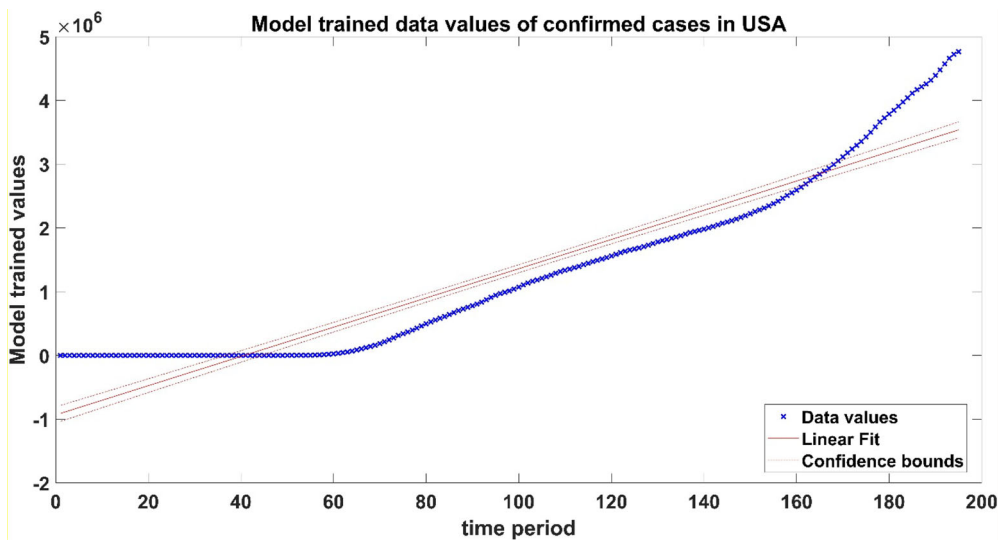
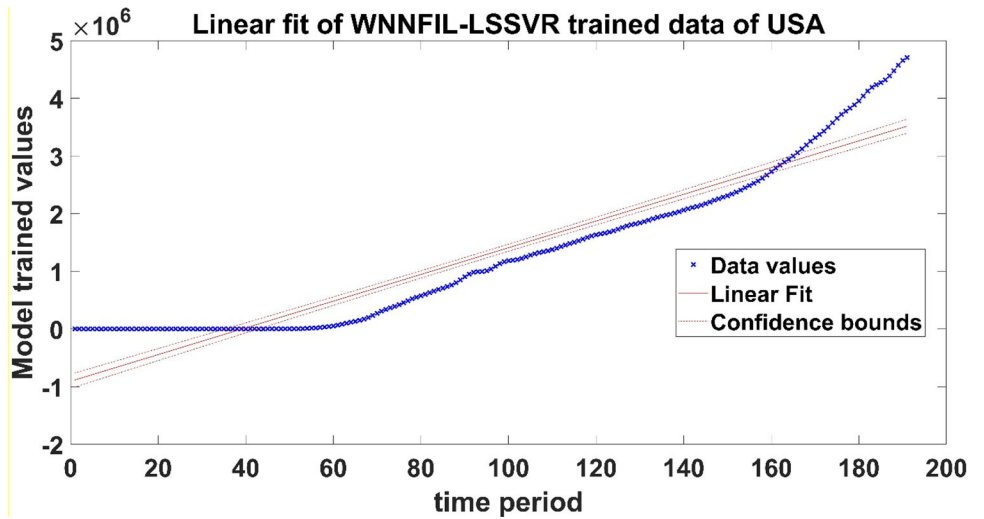


Fig. 6 Linear regression fit for WNNFIL-MARS trained model for USA

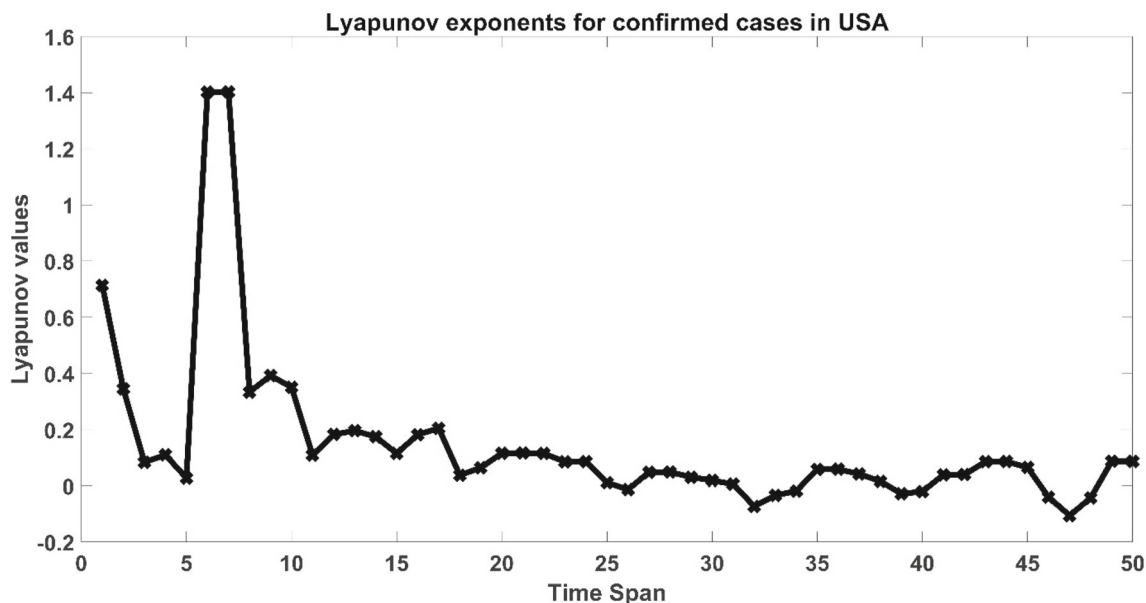


Fig. 7 Lyapunov exponent values for Confirmed cases in USA

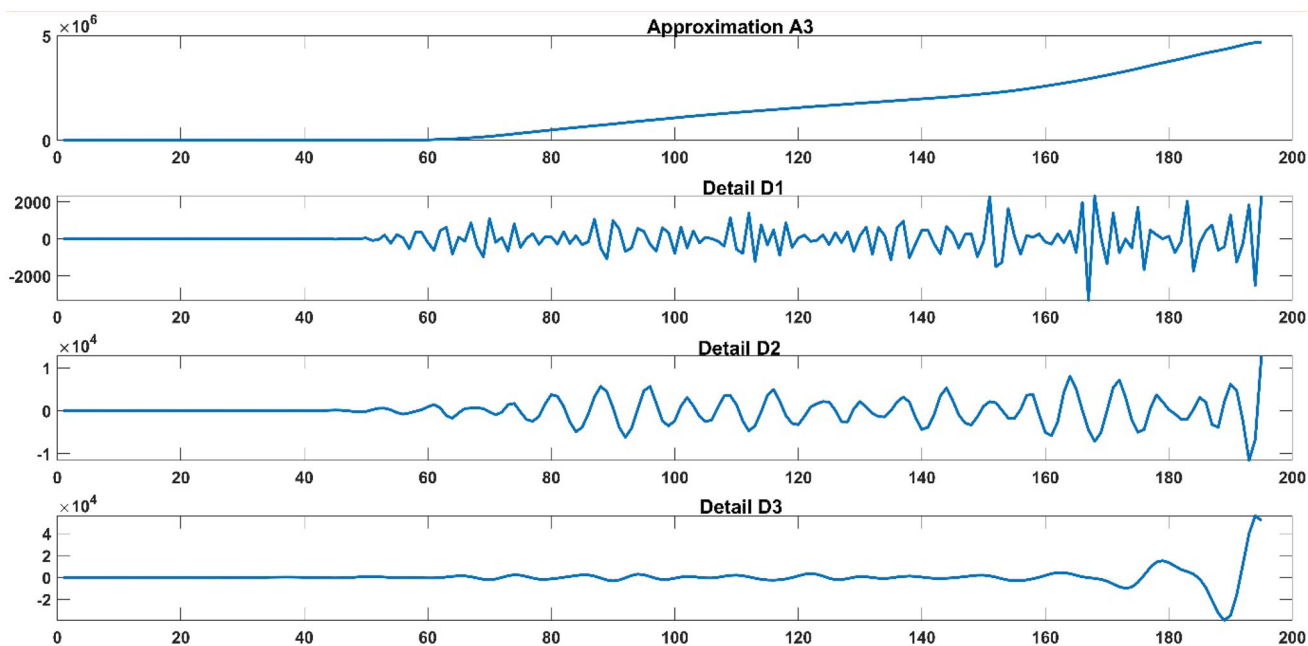


Fig. 8 Wavelet denoising of confirmed cases for USA

jectories indicate diverging with respect to time being known as the sensitive dependence on ICs coined as *Chaos*. Figure 8 clearly tracks the wavelet decomposition into A_3, D_1, D_2, D_3 to feed as responses into trained values. It filtered the noise that extracted characteristics trained and tested to forecast the predictor variables. Figure 9 shows the actual values in comparison to trained forecasts obtained from wavelet decomposition values fed into WNNFIL trained model for the confirmed cases in USA. Similarly, Fig. 10 demonstrates the actual values in comparison to tested fore-

casts obtained from WNNFIL tested model for the confirmed cases. The predicted values based upon the confirmed cases data observed in comparison to the actual ones through WNNFIL–LSSVR using RBF kernel in Fig. 11. Similarly, estimated values upon the confirmed cases observed in comparison to the actual ones through WNNFIL–MARS model as can be seen in Fig. 12.

For India, confirmed cases have been analyzed. Figure 13 shows linear fitting of WNNFIL trained model for India confirmed data set. Model trained values remain almost in the neighbourhood of zero upto 70

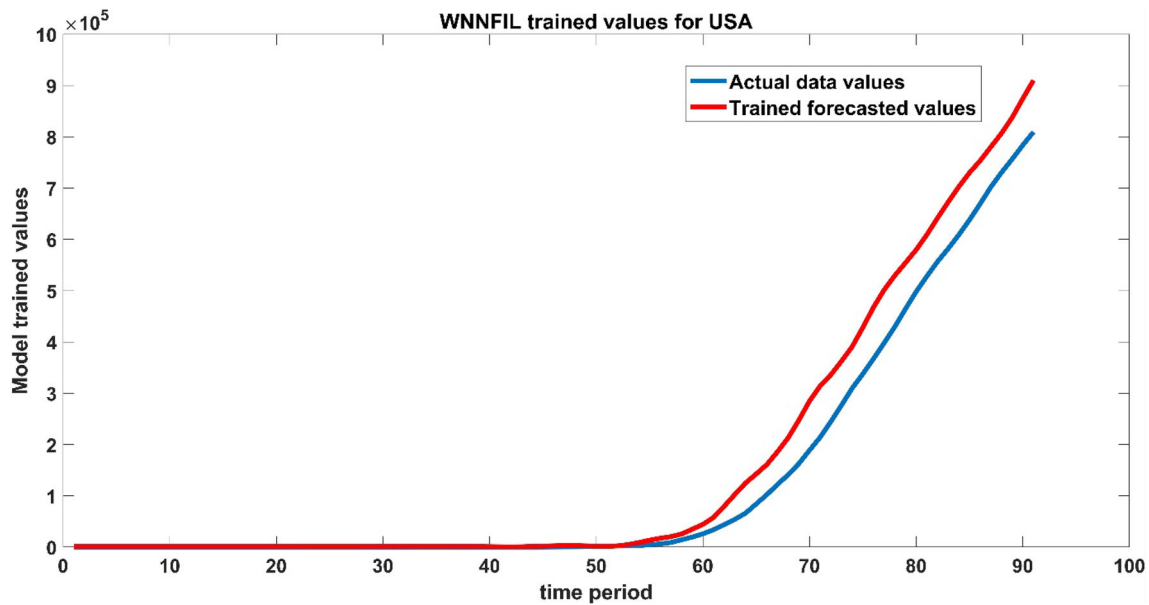


Fig. 9 WNNFIL trained model for confirmed USA

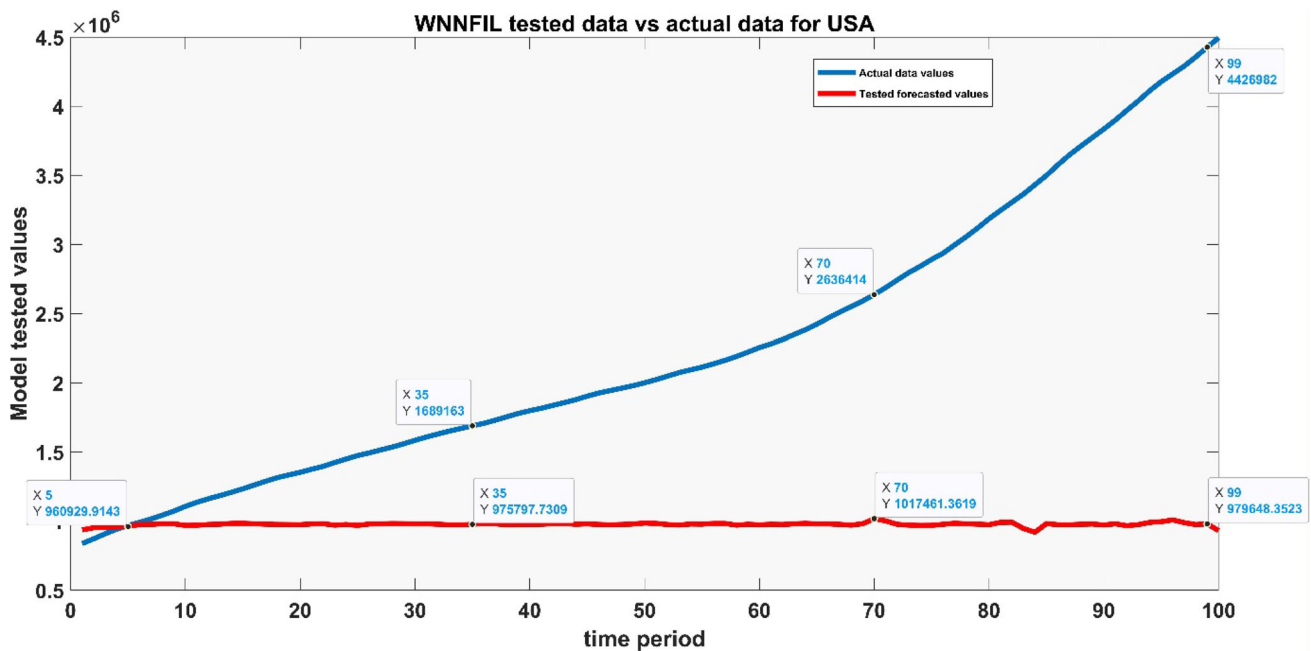


Fig. 10 WNNFIL tested model for confirmed USA

days then increase for rest of the days. Whereas linear regression plots a straight increasing line leaving out almost all the data values. Figure 14 depicts linear regression of WNNFIL–LSSVR trained model for India. Model trained values remain almost stabilized upto 120 days then increases for the rest of the period. Whereas linear regression plots a straight increasing line leaving out almost all the data values. Figure 15 illustrates linear regression of WNNFIL–MARS trained model for India. Model trained values remain almost stabilized for almost 120 days then increases nearly

exponentially. Figure 16 shows that Lyapunov exponents for the confirmed cases in India. Whenever Lyapunov exponents can be seen < 0 (-ve), it showed that the nearby initial conditions(ICs) converging to one another leading towards small epsilon-errors decreasing with respect to time. For any of the Lyapunov characteristics > 0 (+ ve), it showed infinitesimally nearby initial conditions(ICs) diverging from one another exponentially fast implying that the errors in initial conditions would rise with respect to time. As and when trajectories indicate diverging with respect to time being

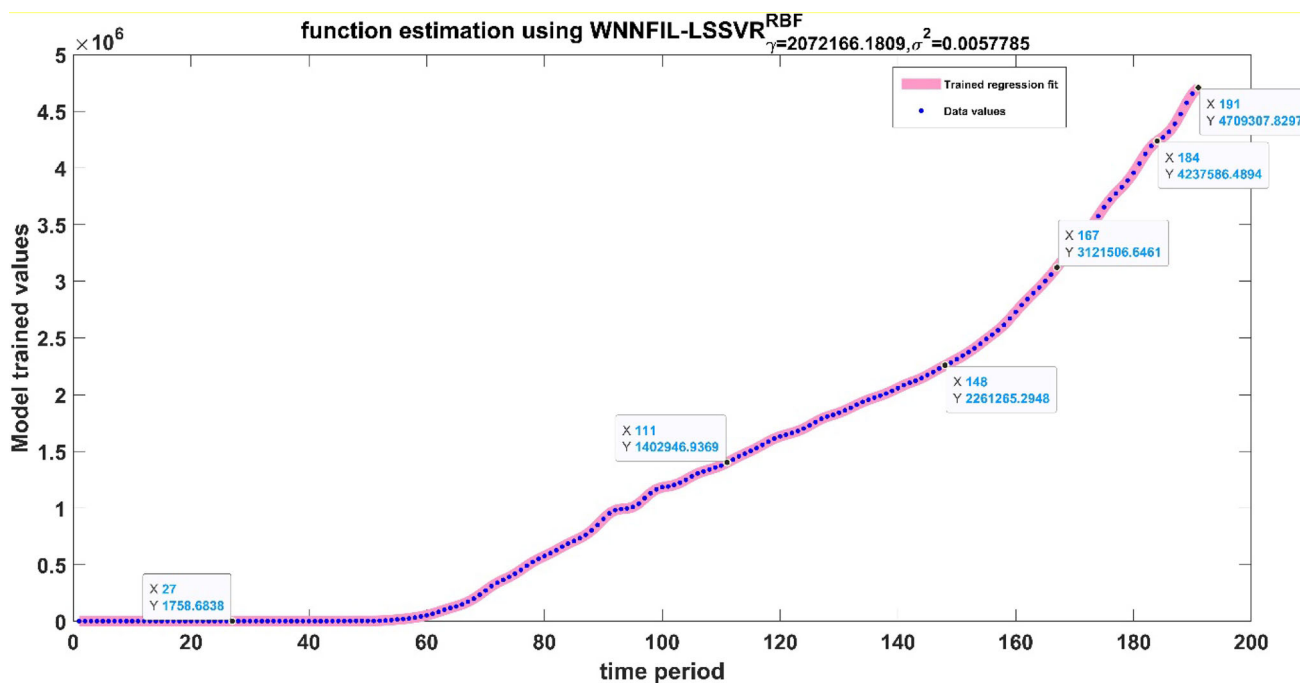


Fig. 11 WNNFIL-LSSVR model for confirmed USA

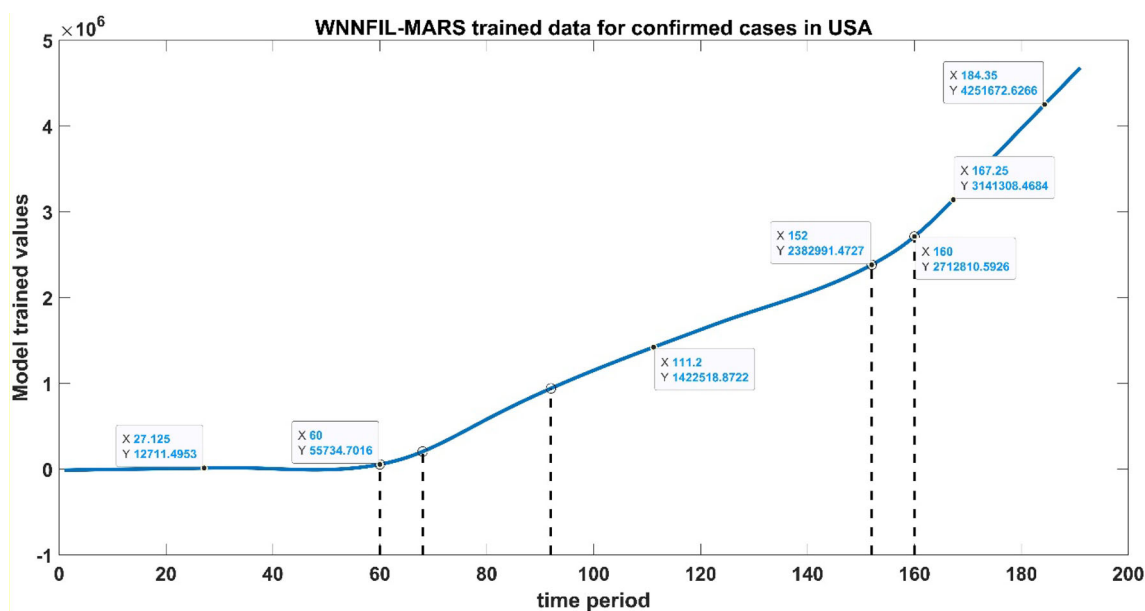


Fig. 12 WNNFIL-MARS model for confirmed USA

known as the sensitive dependence on ICs coined as *Chaos*. Thus, these values indicated towards the randomness in the system. Figure 17 clearly tracks the wavelet decomposition into A_3, D_1, D_2, D_3 to feed as responses into trained values. It shows the decomposition of signal into approximations and details at various levels. It filtered the noise that extracted characteristics trained & tested to forecast the predictor variables. Figure 18 shows the actual values in comparison to trained forecasts obtained from wavelet decom-

position values fed into WNNFIL trained model for the confirmed cases in USA. Similarly, Fig. 19 demonstrates the actual values in comparison to tested forecasts obtained from WNNFIL tested model for the confirmed cases. The two trajectories are entirely different as the testing under this model is unable to capture the peaks of future forecasts. The predicted values of the confirmed cases data observed in comparison to the actual ones through WNNFIL-LSSVR using RBF kernel in Fig. 20. Similarly, estimated values up to of the

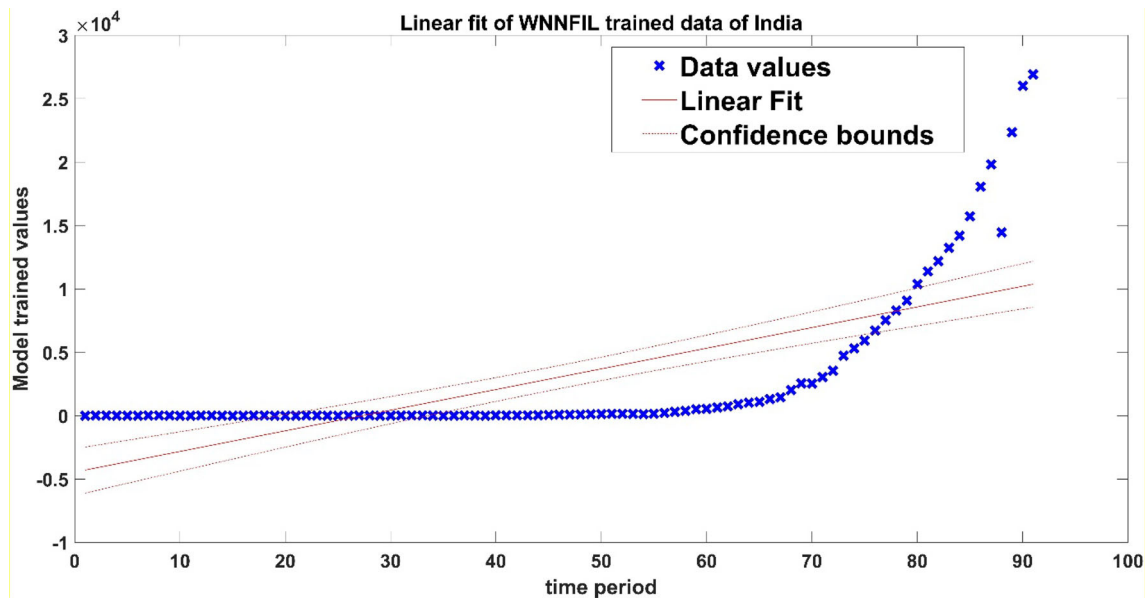


Fig. 13 Linear regression fit for WNNFIL trained model for India

Fig. 14 Linear regression fit for WNNFIL–LSSVR trained model for India

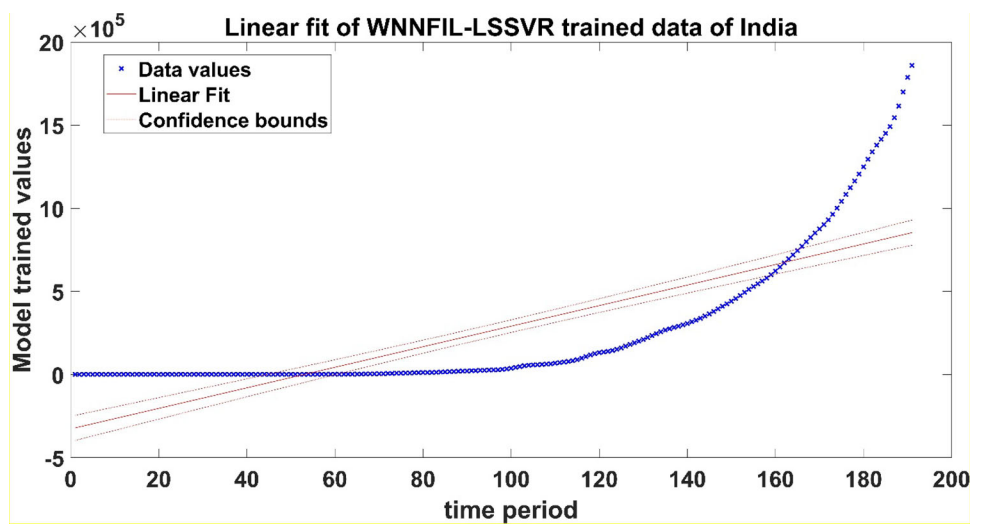
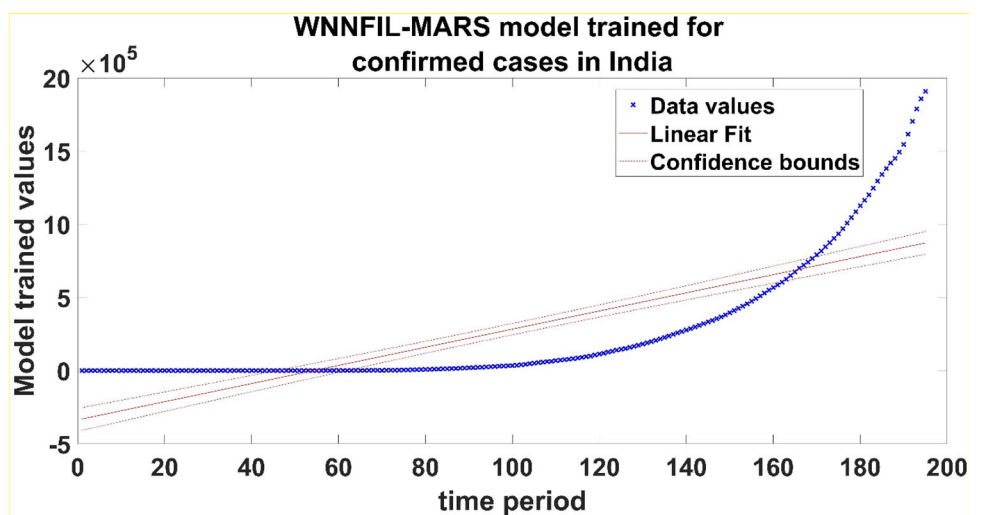


Fig. 15 Linear regression fit for WNNFIL–MARS trained model for India



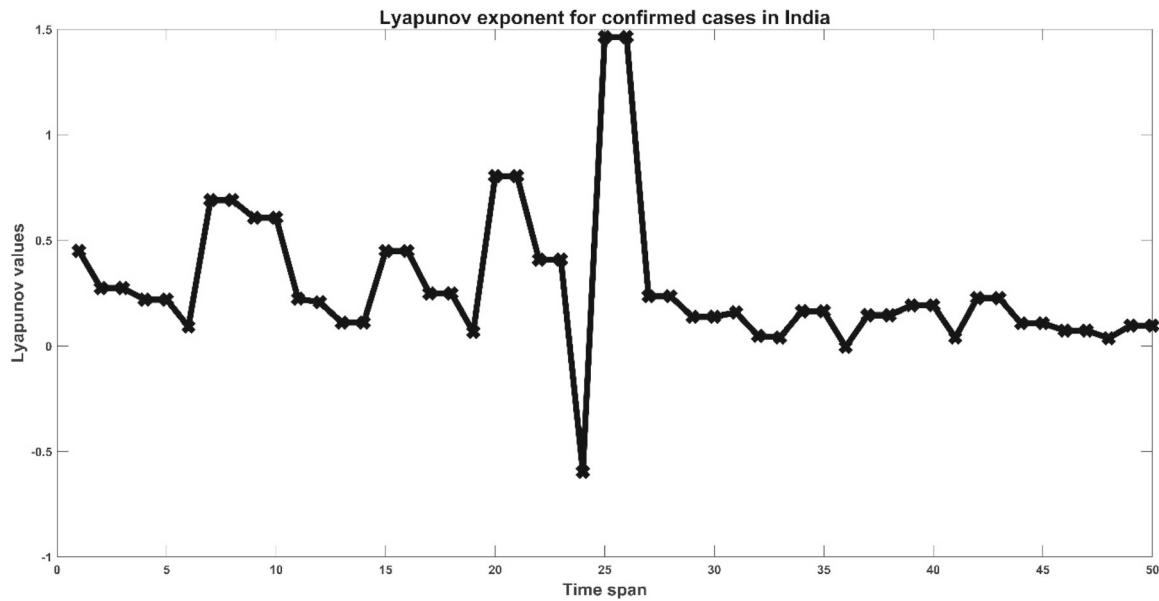


Fig. 16 Lyapunov exponent values for confirmed cases in India

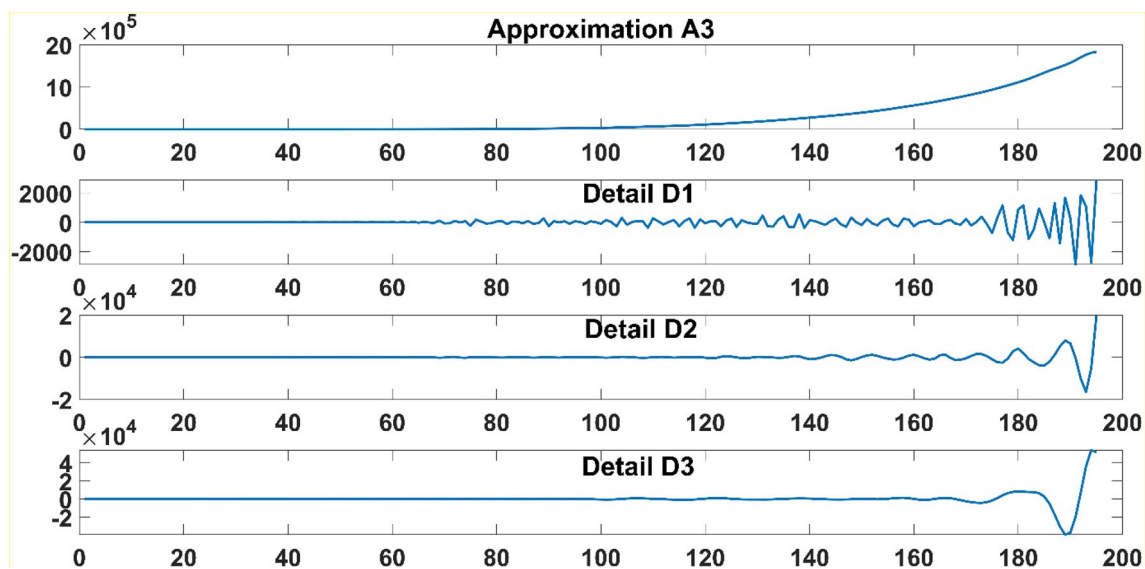


Fig. 17 Wavelet decomposition of daily data of confirmed cases for India

confirmed cases observed in comparison to the actual ones through WNNFIL–MARS intelligent hybrid model in Fig. 21. Table 1 tabulates the statistical characteristics, such as LCEs, Hurst Exponents (HE), R/S value, entropy, fractal dimension (FD), predictability index (PI) and the behaviour of two data sets according to PI. The behaviour observed as anti-persistence based on all the statistical parameters. Determination of these characteristics for non-linear systems involves numerically integrating underlying differential equalities over the variation in parameters and initial conditions. Table 2 shows performance measures that are RMSE, MSE, MAE, R^2 for training data models which are WNNFIL, WNNFIL–LSSVR, WNNFIL–MARS intelligent hybrid

for confirmed cases of India and USA. WNNFIL–MARS model performs best across all errors and fits best in case of India, i.e., 0.987 and comparable to WNNFIL–LSSVR in case of USA, i.e., 0.995. Table 3 encapsulates various studies carried out towards COVID-19 spread.

Conclusions

The study aims to model the features of COVID-19 that can be identified to reduce its spread and develop design strategies such as strictly following social distancing,

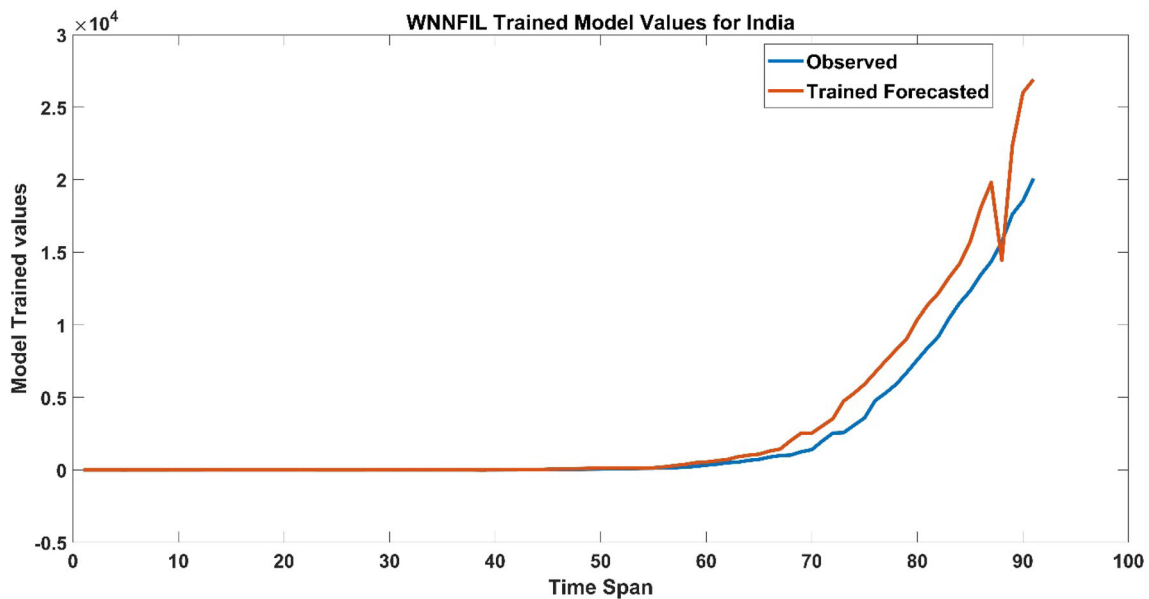


Fig. 18 WNNFIL trained model for confirmed India

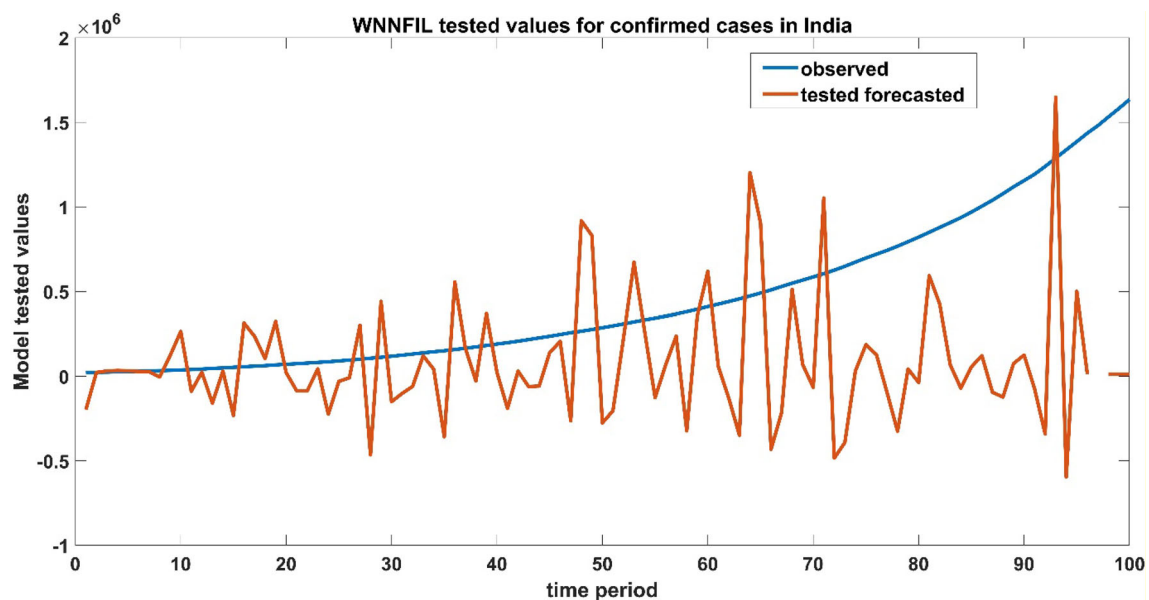


Fig. 19 WNNFIL tested model for confirmed India

sanitizing etc. Thus, it is necessary to study ongoing scenario of the pandemic. Modelling outcomes provide an insight on the extent of interventions required to circumvent or to include so as to predict and estimate future trends. In addition, this study concludes that in both the number of confirmed people with corona have seen a gradual dip which indicates contrary to the perception of pandemic causing havoc in one of the most developed countries. Performance error measures of training models have been simulated through Lyapunov, WNNFIL, WNNFIL–LSSVR, WNNFIL–MARS, Hurst and Entropy for both India and USA. For India, WNNFIL–MARS perform better

compared to WNNFIL–LSSVR. This can be validated through RMSE, MSE, MAE and R^2 values. Similarly, for USA, WNNFIL–MARS have lower error values as compared to WNNFIL–LSSVR. The goodness of fit, R^2 value has been achieved to be closest to 1, i.e., 0.999 from WNNFIL–MARS for both the countries' data. This indicates that wavelet neuronal multivariate adaptive regression splines training-learned model fits best among all other prediction models in this study. In the present-day scenario, a rise leading to a second wave of corona-affected confirmed cases in the month of November have been observed in both the countries. This indicates we need to follow extra-precautious

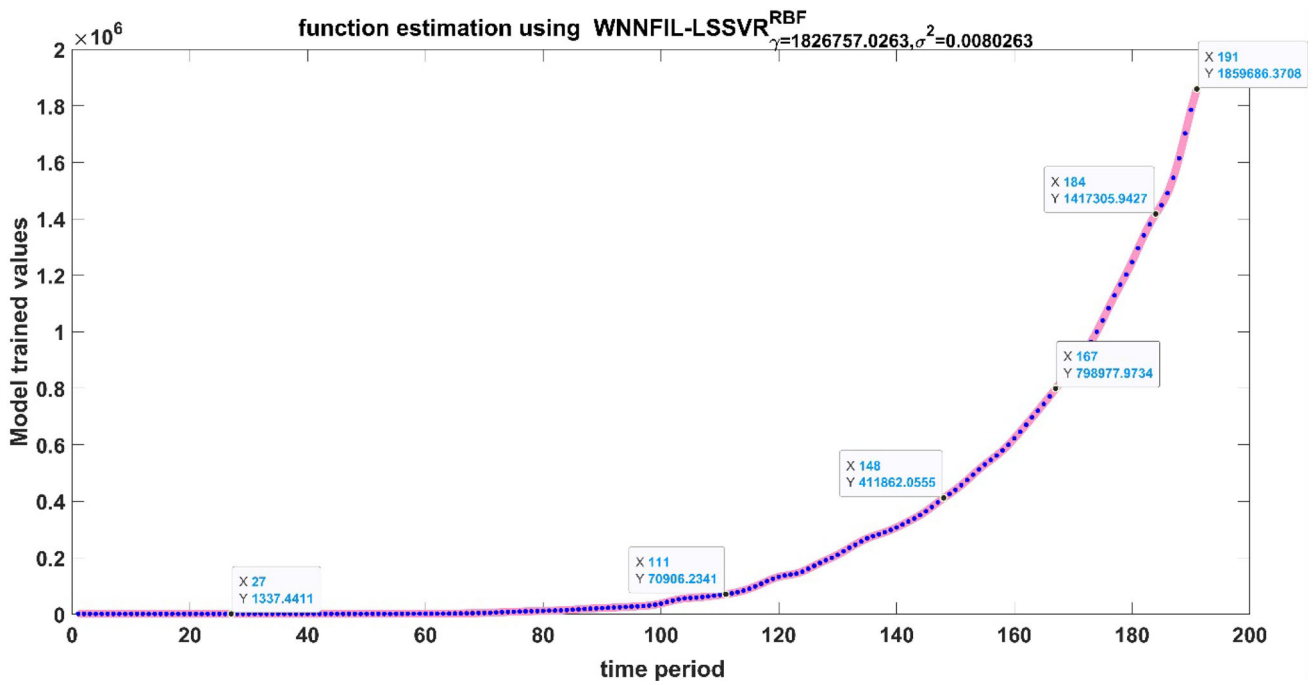


Fig. 20 WNNFIL-LSSVR model for confirmed India

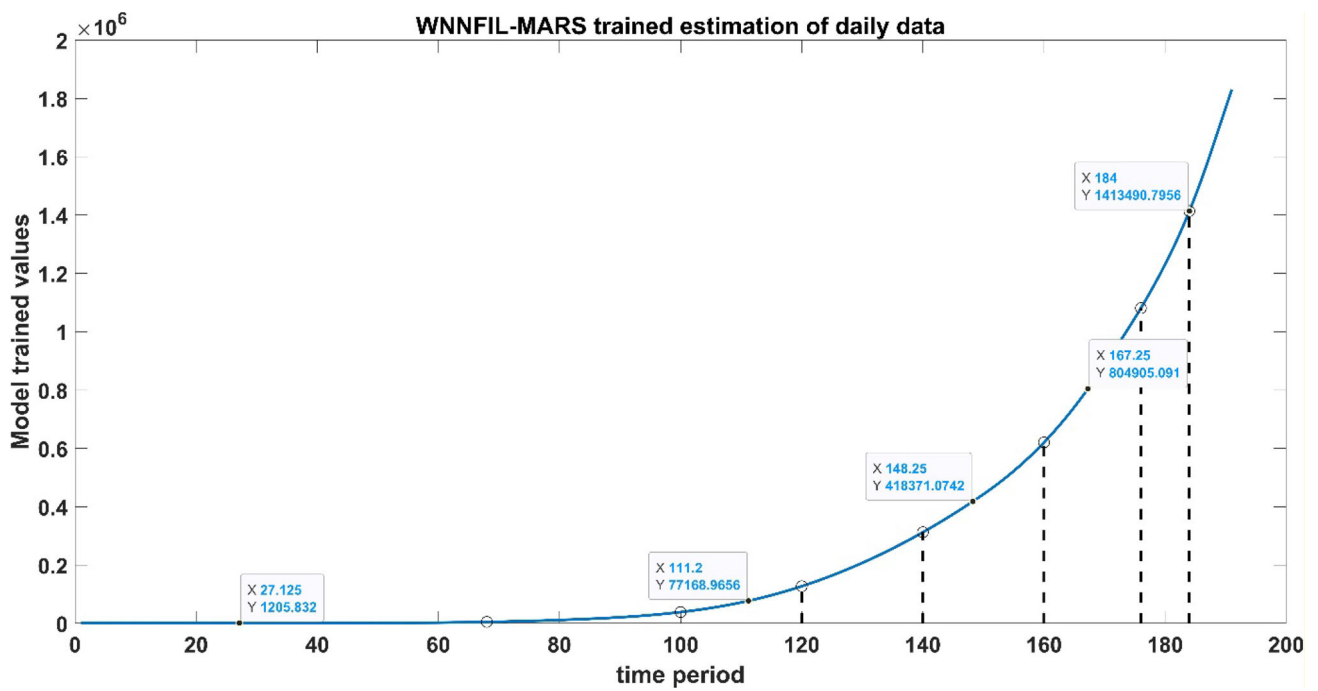


Fig. 21 WNNFIL-MARS model for confirmed India

Table 1 Various statistical characteristics for the two data types

Data types	Lyapunov exponent	Hurst exponent (H_e)	R/S value	Entropy	Fractal dimension ($F_d = 2 - H_e$)	Predictability Index ($P_1 = 2 F_d - 1.5 $)	Behaviour based on P_1
India confirmed	0.0085	0.7806	24.8813	6.7593	1.2194	0.5612	Anti-persistence
USA confirmed	0.0027	0.7810	24.9198	7.1708	1.2190	0.5620	Anti-persistence

Table 2 Performance measures of training models of India & USA daily Confirmed cases data

Data types	Training model	Performance measures			
		RMSE	MSE	MAE	R^2
Daily confirmed cases in India	WNNFIL	6.8109×10^3	4.6389×10^7	3.0366×10^3	0.502
	WNNFIL–LSSVR	5.0707×10^5	2.5712×10^{11}	2.6635×10^5	0.624
	WNNFIL–MARS	5.4639	36.5944	2.8954	0.987
Daily confirmed cases in USA	WNNFIL	7.3783×10^5	5.4440×10^{11}	5.8735×10^5	0.805
	WNNFIL–LSSVR	1.8827×10^6	3.5446×10^{12}	1.3166×10^6	0.999
	WNNFIL–MARS	3.5371	14.3904	2.0331	0.995

Table 3 Different studies carried out towards COVID-19 spread

Published-year	Authors	Brief-Explanation	Characteristics recorded	Simulated results
March, 2020	Fang Y. et al.	Data-driven analysis of transmission of COVTD-19 outbreak	SEIR model, Data fitting	MAE; MSE; R^2
May, 2020	Salgotra R. et al.	Time-Series Analysis and Forecast for COVTD-19 in India	Genetic Algorithm	RMSE; R^2
June, 2020	Rustam F. et al.	Supervised Machine Learning forecasts for COVTD-19	Machine Learning methods	RMSE; MAE; R^2
2020	Bhardwaj; Bangia	Wavelet–NNFEL technique for the daily data of China, India, USA for COVID-19 spread	WNNFIL	MSE, MASE, sMAPE, MAE, RMSE, R^2
2021	Present study Bhardwaj; Bangia	Hybridized Wavelet Neuronal learning-for the data of India and USA to predict novel COVTD-19 spread	WNNFIL; WNNFIL–LSSVR; WNNFIL–MARS	MSE, MAE, RMSE, R^2 , alpha

behaviour. Thus, the social distancing in the infected parts of different cities and countries might show substantial improvement in real and as well as analytical front. In addition, it may aid to employ similar control strategies and various other measures taken by authorities and citizens against this pandemic. Furthermore, this study helped to understand the rise in transmission and spread through surface contacts or humans coming in contact with each other.

Acknowledgements Corresponding author is thankful to G.G.S.I.P. University for providing research facilities.

Declarations

Conflict of interest No conflict-of-interest for publication of this article.

References

- M.A. Al-Bari, Chloroquine analogues in drug discovery: new directions of uses, mechanisms of actions and toxic manifestations from malaria to multifarious diseases. *J. Antimicrob. Chemother.* **70**, 1608–1621 (2015)
- R. Bhardwaj, A. Bangia, Statistical time series analysis of dynamics of HIV. *JNANABHA. Special Issue* **48**, 22–27, (2018). http://docs.vijnanaparishadofindia.org/jnanabha/jnanabha_volume_special_issue_2018/4.pdf
- A. Bangia, R. Bhardwaj, K.V. Jayakumar, Water quality analysis using Artificial Intelligence conjunction with Wavelet Decomposition, in *Advances in Intelligent Systems and Computing (AISC) Volume 979 Numerical Optimization in Engineering and Sciences*. ed. by J. Kacprzyk, D. Dutta, B. Mahanty (Springer, Singapore, 2020), pp. 159–166. https://doi.org/10.1007/978-981-15-3215-3_11
- R. Bhardwaj, A. Bangia, Data driven estimation of novel COVID-19 transmission risks through hybrid soft-computing techniques. *Chaos Soliton Fractals.* **140**, 110152 (2020). <https://doi.org/10.1016/j.chaos.2020>
- R. Bhardwaj, A. Bangia, Machine learned regression assessment of the HIV epidemiology development in Asian region. In: M. Jyoti, A. Ritu, A. Abdon (eds) Chapter 4, *Mathematical Modeling and Soft Computing in Epidemiology*, pp 52–79. Taylor & Francis Publisher (2020). <https://www.taylorfrancis.com/chapters/edit/10.1201/9781003038399-4/machine-learned-regression-assessment-hiv-epidemiological-development-asian-region-rashmi-bhardwaj-aashima-bangia>
- R. Bhardwaj, A. Bangia, J. Mishra, Complexity Analysis of Pathogenesis of Coronavirus Epidemiology Spread in the China region. Chapter 13, *Mathematical Modeling and Soft Computing in Epidemiology*, pp. 247–271, (2020). <https://www.taylorfrancis.com/edit/10.1201/9781003038399-13/complexity-analysis-pathogenesis-coronavirus-epidemiological-spread-china-region-rashmi-bhardwaj-aashima-bangia-jyoti-mishra>
- A. Bonavia, B.D. Zelus, D.E. Wentworth, P.J. Talbot, K.V. Holmes, Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E. *J. Virol.* **77**, 2530–2538 (2003)
- G.E.P. Box, D.R. Cox, An analysis of transformations. *J. R. Stat. Soc. Ser. B* **26**, 211–252 (1964). <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
- T.-M. Chen, J. Rui, Q.-P. Wang, Z.-Y. Zhao, J.-A. Cui, L. Yin, A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infect. Dis. Poverty* **9**(24), 1–8 (2020). <https://doi.org/10.1186/s40249-020-00640-3>
- S. Chen, D. Yang, R. Liu, J. Zhao, K. Yang, T. Chen, Estimating the transmissibility of hand, foot, and mouth disease by a dynamic model. *Public Health* **174**, 42–48 (2019)
- P. Colson, J.M. Rolain, D. Raoult, Chloroquine for the 2019 novel coronavirus SARS-CoV-2. *Int. J. Antimicrob. Agents* **55**(3), 105923 (2020). <https://doi.org/10.1016/j.ijantimicag.2020.105923>
- J.-A. Cui, S. Zhao, S. Guo, Y. Bai, X. Wang, T. Chen, Global dynamics of an epidemiological model with acute and chronic HCV infections. *Appl. Math. Lett.* **103**, 106–203 (2020). <https://doi.org/10.1016/j.aml.2019.106203>
- D. Datta, R. Bhardwaj. Fuzziness-Randomness modeling of Plasma Disruption in First Wall of Fusion Reactor Using Type I Fuzzy Random Set. Chapter 5. *An Introduction to Fuzzy Sets*. pp. 91–113 (2020)
- P.A. Fuhrmann, On the Corona theorem and its application to spectral problems in Hilbert space. *Trans. Am. Math. Soc.* **132**(1), 55–66 (1968). <https://www.math.bgu.ac.il/~paf/corona.pdf>
- A. Gowrisankar, T.M.C. Priyanka, S. Banerjee, Omicron: a mysterious variant of concern. *Eur. Phys. J. Plus* **137**(1), 1–8 (2022)
- A. Gowrisankar, L. Rondoni, S. Banerjee, Can India develop herd immunity against COVID-19? *Eur. Phys. J. Plus* **135**(6), 1–9 (2020). (**Art. No. 526**)
- D. Hamre, J.J. Procknow, A new virus isolated from the human respiratory tract. *Proc. Soc. Exp. Biol. Med.* **121**, 190–193 (1966)
- K.V. Holmes, Coronaviruses. In D.M. Knipe, P.M. Howley (ed.), *Fields virology*, volume 1, 4th edn. Lippincott-Raven Publishers, New York. pp. 1187–1203 (2001). <https://www.worldcat.org/title/fields-virology/oclc/45500371>
- B. Jubelt, J.R. Berger, Does viral disease underlie ALS? Lessons from the AIDS pandemic. *Neurology* **57**, 945–946 (2001). <https://n.neurology.org/content/57/6/945.short>
- C. Kavitha, A. Gowrisankar, S. Banerjee, The second and third waves in India: when will the pandemic be culminated? *Eur. Phys. J. Plus* **136**(5), 1–12 (2021)
- K. McIntosh, R.K. Chao, H.E. Krause, R. Wasil, H.E. Mocega, M.A. Mufson, Coronavirus infection in acute lower respiratory tract disease of infants. *J. Infect. Dis.* **130**, 502–507 (1974). <https://academic.oup.com/jid/article/130/5/502/2189345>
- F. Pene, A. Merlat, A. Vabret, F. Rozenberg, A. Buzyn, F. Dreyfus, A. Cariou, F. Freymuth, P. Lebon, Coronavirus 229E-related pneumonia in immunocompromised patients. *Clin. Infect. Dis.* **37**, 929–932 (2003). <https://doi.org/10.1086/377612>
- M.A. Rosenblum, Corona theorem for countably many functions. *Integral Equ. Operat. Theory* **3**, 125–137 (1980). <https://doi.org/10.1007/BF01682874>
- S.K. Sharma, S. Bhardwaj, R. Bhardwaj, M. Alowaidi, Nonlinear time series analysis of pathogenesis of COVID-19 pandemic spread in Saudi Arabia. *CMC-Comput. Mater. Continua* **66**(1), 805–825 (2021). <https://www.techscience.com/cmc/v66n1/40482>
- World Health Organization. Novel Coronavirus—Japan (ex-China). World Health Organization. <https://www.who.int/csr/don/17-january-2020-novel-coronavirus-japan-ex-china/en/>. Accessed 20 Jan 2020
- WHO. Middle East respiratory syndrome coronavirus (MERS-CoV)—update:2 DECEMBER 2013. http://www.who.int/csr/don/2013_12_02/en/
- C.-Y. Wu, J.-T. Jan, S.-H. Ma, C.-J. Kuo, H.-F. Juan, Y.-S.E. Cheng et al., Small molecules targeting severe acute respiratory syndrome human coronavirus. *Proc. Natl. Acad. Sci.* **101**, 10012–10017 (2004)

28. World Health Organization. Coronavirus. World Health Organization. <https://www.who.int/health-topics/coronavirus>. Accessed 19 Jan 2020
29. P. Zhou, XL. Yang, XG. Wang, B. Hu, L. Zhang, W. Zhang *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. **579**(7798), 270–273 (2020). <https://www.nature.com/articles/s41586-020-2012-7>
30. N. Zhou, T. Pan, J. Zhang, Q. Li, X. Zhang, C. Bai, H. Zhang, Glycopeptide antibiotics potently inhibit cathepsin 1 in the late endosome/lysosome and block the entry of Ebola virus, middle east respiratory syndrome coronavirus (MERS-CoV), and severe acute respiratory syndrome coronavirus (SARS-CoV). *J. Biol. Chem.* **291**(17), 9218–9232 (2016). <https://doi.org/10.1074/jbc.M116.716100>
31. WHO World Health Organization. Novel coronavirus—China. <http://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/>. Accessed 12 Jan 2020