Research Article

A Study for Plausible Third Wave of COVID-19 in India through Fuzzy Time Series Modelling Based on Particle Swarm Optimization and Fuzzy c-Means

Niteesh Kumar, Harendra Kumar, and Kamal Kumar

1Department of Mathematics and Statistics, Gurukula Kangri (Deemed to be University), Haridwar 249404, Uttarakhand, India
2Department of Mathematics, Amity School of Applied Sciences, Amity University Haryana, Haryana, India

Correspondence should be addressed to Harendra Kumar; balyan.kumar@gmail.com and Kamal Kumar; kamalkumarrajput92@gmail.com

Received 16 December 2021; Revised 20 January 2022; Accepted 19 February 2022; Published 16 March 2022

Academic Editor: Juan Frausto-Solis

Copyright © 2022 Niteesh Kumar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The outbreak of COVID-19 has become a global pandemic as announced by World Health Organisation. As India has already met the two waves, named first and second wave, it is assumed that COVID-19 will again strike in India in the form of third wave. The peak during the upcoming third wave and determination of the approximated maximum number of COVID-19 infected cases and deaths at a particular day becomes crucial for India. To determine the peak of infectious curve, this article proposed a hybrid fuzzy time series forecasting model based on particle swarm optimization and fuzzy c-mean technique, named as fuzzy time series particle swarm optimization extended fuzzy c-mean technique. The proposed model works in two phases. In phase-I, particle swarm optimization extended fuzzy c-mean method is used to form initial intervals with the help of centroids, while in phase-II, these intervals are updated to form subintervals. In the present article, a fitness function is developed for particle swarm optimization to increase its convergence speed and basic fuzzy c-mean is extended by using an exponential function to tolerate the effect of outliers, named as extended fuzzy c-mean technique. The effectiveness of the proposed model has been tested based on mean square error and root mean square error on first and second wave COVID-19 data, and the obtained results are very close to the existing data of COVID-19 with less error rate. Thus, the proposed model is suitable to forecast a better approximation value of COVID-19 infected cases and deaths in India during the upcoming third wave. This study demonstrates that third wave of COVID-19 could occur in India, while also illustrating that it is unlikely for any such resurgence to be as large as the second wave. The proposed model predicts that the peak of third wave will occur approximately after 40–70 days from the mid of December. Furthermore, the impact of vaccination on infected cases and deaths during the upcoming third wave in India is also studied. With the implementation of the vaccine on the Indian people, the peak of COVID-19 infected during third wave will be shifted in forward direction. On the basis of the proposed model, government authorities will be enabling to know expected required resources such as hospital patient beds, ICU beds, and oxygen concentrators during the upcoming outspread of COVID-19 like disease in future.

1. Introduction

Currently, a novel virus has affected the whole world, named as COVID-19. The outbreak of coronavirus disease 2019 (COVID-19) has become a global pandemic as announced by WHO. At the end month of 2019, the COVID-19 infected cases show accelerated growth in most countries. Based on several reports, COVID-19 was originated from Wuhan, China, in mid-December 2019. After a few days of its appearance, it will affect the entire world’s people because of its rapid growth. As a result, there is a drastic change in the growth of COVID-19 confirmed cases and the mobility rate is at a peak in India. As time passes, the number of confirmed cases slowly decreases. But no one knows that COVID-19 shows their dangerous mode once again with more infection rate. From the last week of April 2021, the COVID-19 infected cases rapidly increased and reached their peak in India. Due to the invention of vaccine, the situation of
COVID-19 cases in India is under control now. According to the expert’s report, as the first and second waves came in India, the third wave will also predict to come in India during the upcoming months. Thus, estimating the duration and peaks of the forthcoming COVID-19 third wave outbreak in India becomes necessary. Therefore, several models have been developed for the prediction of COVID-19.

Deb and Majumdar [1] developed a time series model for analyzing the COVID-19 infected cases. Mandal et al. [2] used a mathematical model to study the impact of quarantine and travel restriction in India. Bhola et al. [3] proposed a model for prediction of COVID-19 pandemic in India. Mondal and Ghosh [4] analyzed the exponential growth of COVID-19 in India with respect to other countries and gave a future sketch of it. Mandal et al. [5] developed a mathematical model using quarantine class and studied the safety measures developed by the Indian government to control the spreading of COVID-19 and predict the trends of it in three states of India. Maleki et al. [6] developed a model based on the time series concept for forecasting the spreading and death rate of COVID-19 in the world. Mahmoudi et al. [7] compared the COVID-19 spread rate in high-risk countries using time series and fuzzy clustering algorithms. Kumar and Susan [8] proposed two approaches, named as, nested FTS-PSO and exhaustive search FTS-PSO to forecast the COVID-19. Elleuch et al. [9] proposed method for real time forecasting of COVID-19 patients health based on artificial neural networks and fuzzy interval mathematical modeling. Forecasting models and their corresponding consideration on different parameters are presented in Table 1.

Meta-heuristic algorithms have become powerful tools for modeling and optimization. Meta-heuristics have earned more popularity in solving optimization problems because of its simplicity and robustness of results produced. In past several years numerous meta-heuristic algorithms have been developed such as differential evolution, fuzzy logic, firefly algorithm, particle swarm optimization, bat algorithm, genetic programming, genetic algorithm, artificial gorilla troop’s optimization and African vulture optimization etc [18, 19]. Abdollahzadeh et al. proposed an African vulture’s optimization algorithm for the optimization of global problems based on African vultures’ lifestyle [20]. Based on gorilla troops social intelligence, another meta-heuristic algorithm, named as artificial gorilla troops optimizer is designed to perform exploration and exploitation [21]. By taking inspiration from the social behaviour of birds flocking or fish schooling, Kennedy and Eberhart [22] introduced the concept of particle swarm optimization (PSO) in 1995. In PSO, each particle is considered as individual and collection of each particle is swarm. Many researchers studied the impact of parameters on PSO technique. Tian and Shi [23] used robust update mechanism and Chaos-based initialization to modify the PSO algorithm.

Clustering is a technique to partition raw dataset into a smaller group based on similarity measures. In the literature, many researchers combine fuzzy time series (FTS) concept with fuzzy clustering techniques and PSO. Song and Chissom [24, 25] used fuzzy set theory concept to propose a time series model. Li et al. [26] presented FTS model based on fuzzy c-means clustering technique. Sang et al. [27] used IFCM clustering technique to develop forecasting FTS model. Kumar et al. [28] developed AMFCM and EMFCM clustering techniques based on two new distance function. Zhang et al. [29] used time series clustering and multiple linear regressions to present FTS forecasting model. Several authors proposed an improved version of forecasted accuracy by combining PSO with FTS [30–34]. Recently, Tinh [35] has developed hybrid FTS forecasting model based on the combination of PSO and FCM techniques together. They enhance the interval length and improve forecasted values by heuristic defuzzification method.

FCM is frequently used clustering technique but unfortunately, results of it get easily influenced by deciding the initial cluster centroid in priori. To improve the initial condition of FCM, many researchers used genetic algorithm, simulated annealing, evolution strategies and PSO-based approaches for clustering. Out of them, PSO has gained much attention of researchers due to its simplicity and high speed of convergence. PSO has an inbuilt guidance strategy which leads the solution in PSO to obtain useful information from the better solution and thereby helping them improve their own solution. This helps to recover their solutions when they get diverted to an unwanted direction. This makes the algorithm very robust, susceptible to degradation and faster convergence. While other heuristic algorithms have slow convergence rate and they are easily trapped in local optima, which lead to the undesired solution.

Basnarkov [36] studied SEAIR model using two different approaches, first one is based on differential equations and other is based on discrete-time epidemic model. In recent years, intuitionistic fuzzy sets have been preferred in the fuzzy modeling. Kumar and Gangwar [37] proposed fuzzy time series forecasting model to incorporate degree of hesitation based on intuitionistic fuzzy sets. Wang et al. [38] proposed an intuitionistic fuzzy time series forecasting model. In this model, fuzzy clustering algorithm is used to divide the whole data into unequal intervals and developed the membership function and nonmembership function of the intuitionistic fuzzy set. Kartthick and Gomathi [39] proposed galactic swarm improved whale optimization algorithm based on resource management for efficient mapping in Internet of things. Recently, Kocak et al. [40] developed an intuitionistic fuzzy time series forecasting model based on LSTM. In this method, author used intuitionistic fuzzy c-means to determine the membership and nonmembership values. Then, these values are merged by using a minimum operator.

Most researchers use mathematical modeling techniques for the prediction of COVID-19 in recent months with more error rates. The current established techniques employed for analyzing and forecasting the COVID-19 diseases are found insufficient due to the inherent uncertainties associated with the nature of the virus causing infections. The present article will focus on developing fast and efficient time series model for the ongoing and coming waves of COVID-19. This model is of urgent significance to combat the crisis effectively as it will help decision makers for estimating the number of COVID-19 infected cases and deaths. As India has already met the two waves, named first and second wave,
it is assumed that the COVID-19 will again strike in India in the form of third wave. The peak during the forthcoming third wave and determination of the approximated maximum number of COVID-19 infected cases and deaths at a particular day is crucial for India. This gave us motivation to develop a forecasting model to predict the approximated COVID-19 infected cases and deaths during the third wave and to estimate the duration and peak of the outbreak. In the present article, a fitness function is developed for PSO to increase its convergence speed and basic FCM is extended to tolerate the effect of outliers, named as extended fuzzy c-mean (EFCM). Thus, to determine the peak of the infectious curve, this article proposed a hybrid fuzzy time series forecasting model based on particle swarm optimization and extended fuzzy c-mean technique, named as fuzzy time series particle swarm optimization extended fuzzy c-mean (FTSPSOEFCM).

The effectiveness of the proposed model has been illustrated by testing it on enrolment data of Alabama University, TAIPEX data, first and second wave COVID-19 data and the obtained results are very close to the existing data with less error rate.

(vi) Proposed model predicts that the peak of third wave will occur approximately after 40–70 days from the mid of December.

(vii) The impact of vaccination on infected cases and deaths during the upcoming third wave in India is also studied.

The rest of the manuscript is categorized as follows. Section 2 defines the basic preliminaries about FCM, PSO and FTS. Mathematical notations which are used throughout the manuscript and proposed EFCM with their necessary conditions of minimization have been given in Section 3. Section 4 describes the proposed FTSPSOEFCM algorithm. Section 5 contains the performance evaluation criteria of proposed algorithm. Section 6 reveals the implementation of proposed algorithm. Finally, Section 7 concludes the present work.

2. Background Information

In this section, some basics concept about PSO, FCM, and FTS are discussed, which are used in the proposed model.

2.1. Particle Swarm Optimization. Particle swarm optimization (PSO) technique was proposed by Kennedy and Eberhart [22]. The inspiration behind this technique is the social behaviour of fishes and birds, which are flock to search the food. In this technique, each bird is treated as an individual particle, so each particle has its own position and velocity. In PSO, the global best position and personal best position play a crucial role in the upgradation of particles position and velocity. The particle velocity and position are updated by (1) and (2), respectively.

\[ v_i^{(t+1)} = \omega \times v_i^{(t)} + \gamma_1 \times \eta_1 \times (P_i^{(t)} - x_i^{(t)}) \]
\[ + \gamma_2 \times \eta_2 \times (G^{(t)} - x_i^{(t)}) \]
\[ x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)}, \]
2.2. Fuzzy c-Means Clustering Technique. FCM is a traditional algorithm among the existing clustering techniques due to its easy implementation. FCM divided the data into smaller groups, so that it can belong to more than one cluster with different membership value. Ruspini [41, 42] used the concept of fuzzy set theory to form clusters. Later on, Bezdek [43] improved the process of clustering after FCM formulation by Dunn [44]. The objective of FCM is to minimize the objective function with their necessary conditions as follows:

\[ J(x_i, C_k) = \sum_{i=1}^{n} \sum_{k=1}^{c} \mu_{ki}^{m} \| x_i - C_k \|^2; \quad 1 < m < \infty, \quad (3) \]

where \( \mu_{ki} \) is membership grade of the \( i \)th data point in \( k \)th cluster.

\[ \mu_{ki} = 1/\sum_{j=1}^{c} (\| x_i - C_j \| / \| x_i - C_k \| )^{(m-1)} \]

and \( C_k = \sum_{i=1}^{n} \mu_{ki} x_i / \sum_{i=1}^{n} \mu_{ki} \) are necessary condition that will minimize (3). Here, \( \mu_{ki} \in [0,1] \) and \( \sum_{k=1}^{c} \mu_{ki} = 1; i = 1, 2, \ldots, n. \)

2.3. Some Basic Definitions. This section contains important definitions which are used in throughout the present article as follows.

Definition 1. Fuzzy time series [24]: let \( f_j(t) \) be a fuzzy set defined on a universe of discourse \( X(t) = \{0, 1, 2, \ldots\} \), a subset of \( R \). Then the collection of \( f_1(t), f_2(t), \ldots \) is denoted by \( F(t) \) and it is called FTS defined on \( X(t) \).

Definition 2. First order model [25]: suppose \( F(t) \) is formed by \( F(t-1) \) i.e. \( F(t-1) \rightarrow F(t) \). Then, fuzzy relation can be expressed as \( F(t) = F(t-1) \circ R(t-1) \) where \( R(t-1) \) is fuzzy relationship between \( F(t-1) \) & \( F(t) \) and “\( \circ \)” represent max-min composition. Then, \( F(t) \) is called first order model.

Definition 3. Fuzzy logical relationship [45]: if \( F(t-1) = L(i) \) and \( F(t-1) = L(j) \) then the relationship between \( F(t-1) \) & \( F(t) \) is known as FLR and can be expressed as \( L(i) \rightarrow L(j) \) where \( L(i) \) and \( L(j) \) are previous and current state of FLR.

Definition 4. Fuzzy logical relationship group [46]: let assume that \( L(i) \rightarrow L(k1), L(i) \rightarrow L(km), \ldots, L(i) \rightarrow L(km) \) are FLR’s. Then these FLR’s can be grouped to form FLRG as \( L(i) \rightarrow L(k1), L(k2), \ldots, L(km) \).

3. Problem Formulation

In the present section, some general notations and derivation related to the proposed model is discussed.

3.1. Notations. The various notations used throughout this article are as:

\( x_i: i^{th} \) data point
\( C_k: k^{th} \) cluster centroid

\( \mu_{ki}: \) membership grade of \( i^{th} \) data point in \( k^{th} \) cluster
\( c: \) number of clusters \( (2 \leq c \leq n - 1) \)
\( m: \) fuzzy index \( (m > 1) \)
\( J(x_i, C_k): \) objective function
\( v_i^{(t)}: \) velocity of particle at time \( t \)
\( P_i^{(t)}: \) personal best position of each particle at time \( t \)
\( G_t^{(t)}: \) global best position at time \( t \)
\( R: \) basic reproductive number

EFCM: extended fuzzy c-means
FTSPSOEFCM: fuzzy times series particle swarm optimization extended fuzzy c-mean

3.2. Problem Statement. From the last year, the whole world faces a big problem in the form of COVID-19. Indian people have a drastic impact of COVID-19 due to its high population density. In India, COVID-19 already strikes two times, named as first wave and second wave. Now it is assumed that the third waves will again strike in India. Therefore, it becomes necessary to predict the duration and peak of third-wave COVID-19 in India. By considering this problem, a model has been developed by using the impact of vaccination and without vaccination on COVID-19 infected cases.

3.3. Extended Fuzzy C-Means. Due to easy implementation of FCM, it has been widely used by the researcher. But FCM is precise to the outliers. Therefore, it will get affected by them easily. To improve this disadvantage of FCM, the objective function of has been extended for better robustness and the objective function of EFCM is given below:

\[ J(x_i, C_k) = \sum_{i=1}^{n} \sum_{k=1}^{m} \mu_{ki}^{m} \| x_i - C_k \|^2 \exp \left\{ -\frac{x_i \cdot \overline{X}}{2\sigma^2} \right\}, \quad (4) \]

where, \( \overline{X} = \sum_{i=1}^{n} x_i / n \). The necessary conditions for the minimization of (4) are derived below.

3.3.1. Derivation for Cluster Centroid. Cluster centroid is the necessary condition to minimize the objective function of proposed EFCM. It is derived by using mathematical induction. For first cluster, the centroid will be as:

\[ J(x_i, C_k) = \sum_{i=1}^{n} \sum_{k=1}^{m} \mu_{ki}^{m} \| x_i - C_k \|^2 \exp \left\{ -\frac{x_i \cdot \overline{X}}{2\sigma^2} \right\}. \quad (5) \]

Differentiate the (5) with respect to \( C_1 \) by considering other variables as constant.

\[ \sum_{i=1}^{n} \mu_{i1} (x_i - C_1) \exp \left\{ -\frac{x_i \cdot \overline{X}}{2\sigma^2} \right\} = 0, \quad (6) \]

\[ C_1 = \frac{\sum_{i=1}^{n} \mu_{i1} \cdot x_i \exp \left\{ -\frac{x_i \cdot \overline{X}}{2\sigma^2} \right\}}{\sum_{i=1}^{n} \mu_{i1} \exp \left\{ -\frac{x_i \cdot \overline{X}}{2\sigma^2} \right\}}. \]

Supposed, \( j^{th} \) cluster centroid is
shown below:

\[ J(x_i, C_{j+1}) = \frac{\sum_{i=1}^{n} \mu_{m}^{i} x_i \exp \left[ -x_i \lambda_2 \sigma^2 \right]}{\sum_{i=1}^{n} \mu_{m}^{i} \exp \left[ -x_i \lambda_2 \sigma^2 \right]} \]  

(7)

Now, determination of \( k = (j+1) \)th cluster centroid is shown below:

\[ J(x_i, C_{j+1}) = \left( \sum_{i=1}^{n} \mu_{m}^{i} \| x_i - C_{j+1} \|^2 \exp \left[ -x_i \lambda_2 \sigma^2 \right] \right) \]  

(8)

Again, differentiate the above Eqs. with respect to \( C_{j+1} \) and following results are obtained.

\[ C_{j+1} = \frac{\sum_{i=1}^{n} \mu_{m}^{i} x_i \exp \left[ -x_i \lambda_2 \sigma^2 \right]}{\sum_{i=1}^{n} \mu_{m}^{i} \exp \left[ -x_i \lambda_2 \sigma^2 \right]} \]  

(9)

Hence, the general form of cluster centroid for EFCM is as follows:

\[ C_k = \frac{\sum_{i=1}^{n} \mu_{m}^{i} x_i \exp \left[ -x_i \lambda_2 \sigma^2 \right]}{\sum_{i=1}^{n} \mu_{m}^{i} \exp \left[ -x_i \lambda_2 \sigma^2 \right]} \]  

(10)

3.3.2. Derivation for Membership Function. Membership function is another necessary condition, which will minimize (4). The membership function of proposed EFCM will be obtained by forming the Lagrangian function [43] for EFCM. So, the Lagrangian function for (4) is

\[ J_k(\lambda, \mu_k) = \sum_{k=1}^{c} \| x_i - C_k \|^2 \exp \left[ -x_i \lambda_2 \sigma^2 \right] - \lambda \left( \sum_{k=1}^{c} \mu_k - 1 \right). \]  

(11)

For obtaining necessary condition, differentiates the (11) with respect to \( \lambda \) by considering other variables as constant.

\[ \frac{dJ_k}{d\lambda} = \left( \sum_{k=1}^{c} \mu_k - 1 \right) = 0. \]  

(12)

Again, differentiate (11), but this time, with respect to \( r \mu \) and obtain the following form.

\[ \mu_{ts} = \left( \frac{\lambda}{m} \right)^{1/(m-1)} \]  

(13)

From (12) and (13), we get

\[ \left( \frac{\lambda}{m} \right)^{1/(m-1)} = \frac{1}{\left( \sum_{j=1}^{c} \| x_i - C_j \|^2 \exp \left[ -x_i \lambda_2 \sigma^2 \right] \right)^{1/(m-1)}}. \]  

(14)

By putting the value of \( \lambda/m \) in (13), the general form of membership value is obtained.

\[ \frac{1}{\sum_{j=1}^{c} \| x_i - C_j \|^2 \exp \left[ -x_i \lambda_2 \sigma^2 \right]^{1/(m-1)}}. \]  

(15)

4. Description of the Proposed Model

As the first wave and second wave already came in India, the third wave also predicted to comes in India during the forthcoming days, according to experts’ reports. Thus, it becomes necessary to estimate the duration and peaks of the outbreak of upcoming COVID-19 third wave in India. So, the Indian government prepare in advance for public health and economic decision. The present article addresses a model for the prediction of COVID-19 third wave in India using FTS technique based on the combination of PSO and EFCM algorithm. The present FTSPSOEFCM model predicts third wave in two phases.

Phase-I: PSO-EFCM has been implemented to obtain upgraded centroids, which are further used to form intervals

Phase-II: FTS model is used to forecast COVID-19 third wave in India

4.1. Phase-I. In phase-I, intervals are formed based on cluster centroid which are formed through PSO-EFCM technique. To start the process of PSO-EFCM, form the initial centroid by using PSO technique, and then upgrade these centroids through the proposed EFCM. In the present technique, centroids are treated as a swarm particle. Before running PSO, randomly select the centroids and initialize each particle velocity. Now, calculate the fitness value of each particle by using

\[ f_k = \frac{1}{J_k + \eta} \]  

(16)

where \( \eta \) is constant and \( J_k = \sum_{i=1}^{n} \| x_i - C_{k} \|^2 \exp \left[ x_i X_2 \sigma^2 \right] \). On the basis of fitness value, select gbest (G) and pbest (P_k) to update the velocity and position by (17) and (18), respectively.

\[ V_{k}^{t+1} = \omega \times V_{k}^{t} + k_1 \times p_{k}^{t} \times (P_{k}^{t} - C_{k}^{t}) \]  

(17)

\[ C_{k}^{t+1} = C_{k}^{t} + V_{k}^{t+1}, \]  

(18)

where \( \omega \) varies from 0.9 \( \rightarrow \) 0.4, \( k_1 = k_2 = 1.49, \phi_1 = 0.5949, \phi_2 = 0.0855 \). Repeat the process, until it reaches the termination condition (here termination condition is either it reaches maximum number of iteration 100 or gbest does not change for successive 20 iteration).
Now, PSO gives the centroids, which are further upgraded by using EFCM clustering technique for more stability of these centroids. Calculate the membership value and centroid by using (19) and (20) respectively.

\[
\mu_{ki}^{(t+1)} = \frac{1}{\sum_{j=1}^{m} \|x_i - C_k^j\|^2 \exp[-\|x_i - C_k^j\|^2 / \|x_i - C_k\|^2 / 2 \sigma^2]}^{1/(m-1)}
\]

(19)

\[
C_k^{(t+1)} = \frac{\sum_{i=1}^{n} (\mu_{ki}^{(t+1)})^n x_i \exp[-\|x_i - C_k\|^2 / 2 \sigma^2]}{\sum_{i=1}^{n} (\mu_{ki}^{(t+1)})^n \exp[-\|x_i - C_k\|^2 / 2 \sigma^2]}
\]

(20)

Repeat the process of EFCM until the difference between successive membership values is negligible for all datasets. Finally, the updated and stable centroids are obtained. Arrange these centroids in ascending order and form the initial intervals with the help of them.

Phase-I is elaborated in the form of Algorithm 1 with involved steps.

4.2. Phase-II. In phase-II, FTS technique is used to upgrade the intervals, which are formed in phase-I. Initially, assign each data point to the formed intervals and partition each interval into subintervals of equal length. Among these subintervals, select only those subintervals in which historical data belong and referred them as \(d_i; i = 1, 2, \ldots, n\). These selected subintervals called upgraded intervals. Now, define linguistic variables for upgraded intervals and allocate them to given data to fuzzified it as. \(L_1, L_2, \ldots, L_n; n \in N\).

These fuzzified values can be defuzzified by using (21), for known linguistic variable, is given as follows:

\[
F(t) = \frac{\alpha_1}{\alpha_1 + \alpha_2} E_{GI} + \frac{\alpha_2}{\alpha_1 + \alpha_2} E_{Li},
\]

(21)

where \(E_{GI} = \sum_{i=1}^{s} m_i / s; E_{Gi} = b_{it} + (b_{it} - b_{ut})(m_i - m_{i-1}) / 2 \) \((b_{it} - b_{ut})(m_i + m_{i-1}) / 2\), \(b_{it}\) and \(b_{ut}\) represent the lower and upper bound of the selected intervals & \(m_i\) represent the mid-values of intervals at any time \(t\).

If the linguistic variable does not exist, then the forecasted values can be calculated by using (22)

\[
F(t) = F(t-1) + \sum_{i=1}^{n} \frac{F(t-i)}{n} R \frac{e^{\lambda t}}{(a/k)e^{\lambda t} - 1},
\]

(22)

where \(F(t-i); i = 1, 2, \ldots\) represent the previous forecasted values and \(F(t)\) stands for the current forecasted value at any time \(t\), \(R = aS / (b + c); a, b,\) and \(c\) represent the transmission rate, recovery rate and death rate, respectively, \(\lambda\) is any random constant and \(S\) represent the total number of susceptible. To study the vaccination impact on the proposed model, the vaccination factor \(dv\) is introduced in variable \(R\) in reciprocal as. \(R = aS / (b + c)(dv + 1)\).

Phase II is elaborated in the form of Algorithm 2 with summarized steps.

(1) partition the intervals, obtained in phase-I, into subintervals based on COVID-19 data

(2) calculate the midvalues of selected subintervals, according to the belonging of COVID-19 data

(3) for \(i\) \((1 \leq i \leq a)\) \% \(a = \) no. of subintervals

(4) define the linguistic variables only for selected subintervals

as

\[
\mu_i = \sum_{j=1}^{n} d_j / \beta_{ij}; i \in N
\]

where \(\beta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0.5 & \text{if } i = j - 1 \text{ or } j + 1 \\ 0 & \text{otherwise} \end{cases}
\]

(5) end for

(6) assign linguistic variable to all historical data according to the belonging of data to their respective subinterval

(7) define FLR and FLRG

(8) defuzzify the historical data by using (21) for known linguistic variable

(9) determine the forecasted value by using (22) for unknown linguistic variable

(10) return forecasted values

Both phases I and II are depicted through flow chart in Figure 1.

5. Performance Measure

Forecasted accuracy of the proposed model is evaluated on the basis of two different parameters which are mean square error (MSE) and root mean square error (RMSE).

5.1. Mean Square Error [47]. MSE is measures as an average of squared difference between forecasted and actual value. The obtained value of MSE should be lower for better results and calculated by using (23).

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (AV_i - F_i)^2,
\]

(23)

where \(n\) is the total number of data points.

5.2. Root Mean Square Error [48]. RMSE is used to calculate that how much the forecasted value differs with actual value. The value of RMSE is obtained by (24).
where $n$ is the total number of data points.

6. Performance Evaluation

In this section, first we show that the EFCM tolerate the outliers then, the proposed FTSPSOEFCM model is implemented on COVID-19 data. The epidemic data of COVID-19 infected cases and deaths in India have been obtained from authorized portal [49]. For simplicity, the whole data of COVID-19 is divided into two groups i.e., first wave and second wave. The duration of third wave of COVID-19 in India will be predicted. But before the prediction, two different data sets are implemented to evaluate the performance of the proposed model and the obtained results are compared to those produced by other models. The enrolment and TAIFEX data sets are used as the test set. Then, the official data of first wave and second wave of COVID-19 in India have been implemented on the proposed model as a training set for prediction of the forthcoming COVID-19 third wave in India. By the analysis of official data of COVID-19 in India, it may be concluded that the time period of first wave and second wave was 1st April 2020–31st January 2021 and 1st February 2021–15th December 2021, respectively.

6.1. Effect of Outliers. FCM is an effective algorithm but it gets easily affected in the presence of outliers. FCM is extended by introducing a factor in the objective function of FCM to conquer this problem. Let us consider a set of $m$ data points $X = \{x_1, x_2, \ldots, x_m\}$ defined on $\mathbb{R}^n$. The estimated value of $C$ is obtained by minimizing $\sum_{j=1}^{m} \|x_j - C\|^2$ with respect to $C$ and calculated by

$$C = \frac{\sum_{j=1}^{m} x_j}{m}.$$  \hspace{1cm} (25)

Let $\{3.95, 4, 4.5, 4.7, 4.9, 5, 5.1, 5.3, 5.9, 6.2, 7\}$ be randomly generated artificial data, which has to be tested by the procedure of least-square method. Before and after adding the noisy point 30 in artificial data set, the estimated value of $C$ by (25) is 5.1409 and 7.6292, respectively. The result shows that minimizer is highly affected by noisy point. The same data set was tested by proposed model for calculating minimizer and the value of minimizer by (20) are 5.0896 and 6.1249 before and after adding outlier respectively. Hence, proposed model is able to tolerate the outliers.

6.2. Implementation of Proposed Model on Test Set. Student enrolment data of Alabama University and TAIFEX data sets are implemented to evaluate the performance of the proposed model & the obtained results are compared to those produced by other models.
Example 1. A real-world historical data of student enrolment in Alabama University from 1971 to 1992 years has been considered as shown in Table 2.

Let $Y$ be the universe of discourse which contain the historical data. Now, randomly choose the number of clusters and randomly assign the centroids, here the
numbers of clusters are preferred 7, to start the process of PSO algorithm. Apply the steps from 3 to 19 of Algorithm 1, completing the iterative process involved in Algorithm 1, until the termination condition is not reached. After final updated centroids, in ascending order, will be

\[ c_1 = 13446.0962, c_2 = 14691.7843, c_3 = 15374.2656, c_4 = 15900.0612, \]
\[ c_5 = 16408.5900, c_6 = 16886.1949, c_7 = 19068.1174. \]  

(26)

Now apply the steps from 21 to 27 of Algorithm 1, the following basic intervals are obtained:

\[
\begin{align*}
    a_1 &= [12823.2522, 14068.9402], a_2 = (14608.9402, 15033.0249], a_3 = (15033.0249, 15637.1634], \\
    a_4 &= (15637.1634, 16154.3256], a_5 = (16154.3256, 16647.3924], a_6 = (16647.3924, 17977.1562], \\
    a_7 &= (17977.1562, 20159.0786].
\end{align*}
\]

(27)

To upgrade these initial intervals, the steps of Algorithm 2 are applied. Now partition these intervals into subintervals, according to the belonging of enrolment data. Repeat the above procedure of partitioning the intervals, until all the intervals were partitioned. Apply step 4 of Algorithm 2, define the linguistic variable for each interval \( d_i \), \( i = 1, 2, \ldots, 16 \) are expressed as

\[
L(1) = \frac{d_1}{1} + \frac{d_2}{0.5} + \frac{d_3}{0} + \cdots + \frac{d_{14}}{0} + \frac{d_{15}}{0} + \frac{d_{16}}{0},
\]

\[
L(2) = \frac{d_1}{0.5} + \frac{d_2}{1} + \frac{d_3}{0.5} + \cdots + \frac{d_{14}}{0} + \frac{d_{15}}{0} + \frac{d_{16}}{0},
\]

\[
\cdots
\]

\[
L(15) = \frac{d_1}{0} + \frac{d_2}{0} + \frac{d_3}{0} + \cdots + \frac{d_{14}}{0.5} + \frac{d_{15}}{0.5} + \frac{d_{16}}{0.5},
\]

\[
L(16) = \frac{d_1}{0} + \frac{d_2}{0} + \frac{d_3}{0} + \cdots + \frac{d_{14}}{0.5} + \frac{d_{15}}{0.5} + \frac{d_{16}}{1},
\]

(28)

where denominator of \( d_i \) denotes the membership value of each subinterval to linguistic variable. After applying the remaining steps involved in Algorithm 2, the forecasted enrolment data with their linguistic variable are shown in Table 3. The performance of the proposed model is tested against various existing model in literature based on two parameters MSE and RMSE.

### Table 2: Student enrolment data of Alabama University.

| Year | Enrolments | Year | Enrolments | Year | Enrolments | Year | Enrolments |
|------|------------|------|------------|------|------------|------|------------|
| 1971 | 13055      | 1977 | 15603      | 1983 | 15497      | 1989 | 18970      |
| 1972 | 13563      | 1978 | 15861      | 1984 | 15145      | 1990 | 19328      |
| 1973 | 13867      | 1979 | 16807      | 1985 | 15163      | 1991 | 19337      |
| 1974 | 14696      | 1980 | 16919      | 1986 | 15984      | 1992 | 18876      |
| 1975 | 15460      | 1981 | 16388      | 1987 | 16859      | —    | —          |
| 1976 | 15311      | 1982 | 15433      | 1988 | 18150      | —    | —          |

Table 2: Student enrolment data of Alabama University.

From Table 3, the value of MSE and RMSE are 10816.20 and 104.0009, which are very less in comparison of another existing model. It shows that the forecasted values obtained by proposed model are near about to the actual value. Therefore, it can be concluded that the proposed model produced conventional result than other existing models in literature. The student enrolments obtained in Table 3 by the proposed model and other existing model also depicted graphically in Figure 2.

**Example 2.** A real-world historical data of TAIFEX has been considered as shown in Table 4.

The proposed forecasting model has been tested on TAIFEX data and obtained the forecasted value. The obtained results are shown in Table 5. This data set was implemented on the 4 existing FTS techniques and compared them with the proposed forecasting model. By the analysis of Table 5, it can be concluded that the proposed algorithm produces the better results than the existing FTS techniques.

According to Tables 3 and 5, it is clear that the proposed model produces the best results in terms of two criteria (MSE & RMSE). As a result of all comparison, it can be concluded that the proposed model gives better predicted results for all-time series with less error rate.

### 6.3. Predicted COVID-19 Infected Cases during the Upcoming Third Wave in India. In India, the first case of COVID-19 was reported on 30 January 2020 [55]. After that, the number...
Table 3: Comparison of proposed model with other existing models.

| Year | Actual enrolments | Linguistic variable by proposed model | Song and Chisson [24] | Bas et al. [50] | Gupta and Kumar [51] | Sang et al. [27] | Iqbal and Zhang [52] | Li and Cheng [53] | Proposed model |
|------|-------------------|--------------------------------------|-----------------------|----------------|---------------------|-----------------|---------------------|-----------------|----------------|
| 1971 | 13055             | L(1)                                 | —                     | —              | —                   | —               | —                   | —               | 13166.40       |
| 1974 | 14696             | L(4)                                 | 14000                 | 15016          | 14951.36            | 14335           | 14557.95            | 14500           | 14516.78       |
| 1977 | 15603             | L(9)                                 | 16000                 | 15458          | 15533.19            | 15589           | 15544.27            | 15500           | 15614.98       |
| 1980 | 16919             | L(13)                                | 16813                 | 16881          | 17113.79            | 16882           | 16894.05            | 16500           | 17058.15       |
| 1983 | 15497             | L(8)                                 | 16000                 | 15462          | 15532.34            | 15510           | 15500.11            | 15500           | 15507.46       |
| 1986 | 15984             | L(11)                                | 16000                 | 15942          | 15532.34            | 15630           | 15724.11            | 15500           | 16051.18       |
| 1989 | 18970             | L(15)                                | 19000                 | 19055          | 19000               | 18956.86        | 18500               | 18946.89       |

MSE 423026.65 26108.45 186316.66 106543.67 27467.64 85040 10816.20
RMSE 650.40 161.58 431.64 326.41 165.73 291.61 104.00

Figure 2: Graphical representation of student enrolments.

Table 4: Historical data of the TAIFEX under 03/08/1998–30/09/1998.

| Date       | Actual data | Date       | Actual data | Date       | Actual data | Date       | Actual data |
|------------|-------------|------------|-------------|------------|-------------|------------|-------------|
| 03/08/1998 | 7552        | 18/08/1998 | 7220        | 02/09/1998 | 6430        | 17/09/1998 | 6906        |
| 04/08/1998 | 7560        | 19/08/1998 | 7285        | 03/09/1998 | 6200        | 18/09/1998 | 6842        |
| 05/08/1998 | 7487        | 20/08/1998 | 7274        | 04/09/1998 | 6403.2      | 19/09/1998 | 7039        |
| 06/08/1998 | 7462        | 21/08/1998 | 7225        | 05/09/1998 | 6697.5      | 21/09/1998 | 6861        |
| 07/08/1998 | 7515        | 24/08/1998 | 6955        | 07/09/1998 | 6722.3      | 22/09/1998 | 6926        |
| 10/08/1998 | 7365        | 25/08/1998 | 6949        | 08/09/1998 | 6859.4      | 23/09/1998 | 6852        |
| 11/08/1998 | 7360        | 26/08/1998 | 6790        | 09/09/1998 | 6769.6      | 24/09/1998 | 6890        |
| 12/08/1998 | 7330        | 27/08/1998 | 6835        | 10/09/1998 | 6709.75     | 25/09/1998 | 6871        |
| 13/08/1998 | 7291        | 28/08/1998 | 6695        | 11/09/1998 | 6726.5      | 28/09/1998 | 6840        |
| 14/08/1998 | 7320        | 29/08/1998 | 6728        | 14/09/1998 | 6774.55     | 29/09/1998 | 6806        |
| 15/08/1998 | 7300        | 31/08/1998 | 6566        | 15/09/1998 | 6672        | 30/09/1998 | 6787        |
| 17/08/1998 | 7219        | 01/09/1998 | 6409        | 16/09/1998 | 6952.75     | —           | —           |

Table 5: Comparison of proposed model with other existing models.

| Date       | Actual data | Chen [45] | Haurng [54] | Haurng [46] | Kuo et al. [31] | Proposed model |
|------------|-------------|-----------|-------------|-------------|-----------------|----------------|
| 07/08/1998 | 7515        | 7500      | 7500        | 7500        | 7518            | 7520           |
| 14/08/1998 | 7320        | 7183      | 7100        | 7188        | 7332            | 7301           |
| 20/08/1998 | 7274        | 7183      | 7100        | 7188        | 7280            | 7278           |
| 27/08/1998 | 6835        | 6775      | 6650        | 6775        | 6843            | 6838           |
| 02/09/1998 | 6430        | 6450      | 6550        | 6550        | 6417            | 6435           |
| 08/09/1998 | 6859.4      | 6775      | 6850        | 6850        | 6852            | 6836           |
| 15/09/1998 | 6762        | 6775      | 6650        | 6775        | 6771            | 6760           |
of infected cases accelerated rapidly from the mid of April 2020 and reaches the highest in the month of September 2020 which started declined later and comes under control at the end of February 2021 [56]. During the increment of number of COVID-19 infected cases in first wave, the Indian government implemented a period of lockdown and other safety guidelines. That's why the peak in first wave was observed after a long time from the existence of first case in India. Due to the awareness of Indian people, the impact of the first wave was less in India in comparison to other countries. The COVID-19 data of the first wave is implemented on the proposed model as follows.

Let $Y$ be the universe of discourse which contain the COVID-19 infected case during the first wave in India. Now, randomly choose the number of clusters and randomly assign the centroids, here the numbers of clusters are preferred 9, to start the process of PSO algorithm. Apply the steps 3 to 19 of Algorithm 1, until the termination condition is not reached. After completing the iterative process involved in Algorithm 1, final updated centroids, in ascending order, will be

$$C_1 = 2772.6854, C_2 = 10089.9974, C_3 = 15835.5487, C_4 = 20919.5460, C_5 = 28644.6761,$$
$$C_6 = 38874.2052, C_7 = 48343.5275, C_8 = 62602.1088, C_9 = 79334.8218. \tag{29}$$

Now, apply the steps from 21 to 27 of Algorithm 1, the following basic intervals are obtained:

$$\alpha_1 = [-885.9706, 6431.3414], \alpha_2 = (6431.3414, 12962.7731], \alpha_3 = (12962.7731, 18377.5474],$$
$$\alpha_4 = (18377.5474, 24782.1111], \alpha_5 = (24782.1111, 33759.4407], \alpha_6 = (33759.4407, 43608.8664],$$
$$\alpha_7 = (43608.8664, 55472.8182], \alpha_8 = (55472.8182, 70968.4653], \alpha_9 = (70968.4653, 99529.2559]. \tag{30}$$

To upgrade these initial intervals, the steps of Algorithm 2 are applied. Now, form the partition of these intervals into subintervals, according to the belonging of first wave COVID-19 data. Repeat the above procedure of partitioning the intervals, until all the intervals were partitioned. Now, select those subintervals which contain the COVID-19 data. After applying the steps 3, 4 and 5 of Algorithm 2, define the linguistic variables $L(i)$ and assign them to the identified COVID-19 existing data, according to their belongings into the subintervals. Along with this, define first order FLR/FLRG in the following form

$$\text{FLR}(x) = \{x_k, L(i) \rightarrow L(j)\}; \ \forall x = \text{days},$$
$$\text{FLRG}(i) = \{G(i), L(i) \rightarrow L(j)\}; \ \forall G = \text{group}. \tag{31}$$

After applying the remaining steps involved in Algorithm 2, the forecasted infected cases during first wave of COVID-19 with their linguistic variable are shown in Table 6. After the implementation of the proposed model on first wave of COVID-19, it is found that the forecasted results are much closer to the actual trend of the COVID-19 infected cases in India.

As the first wave comes to end, immediately after that the COVID-19 infected cases swiftly increase once again in the middle of March 2021, this is the chirm of the second wave. It becomes out-of-control at the end of April 2021 and it reaches more than 4 Lakh cases per day. As the daily cases rapidly increases, once again the Indian government implemented partial lockdown in different phases. But this time the implementation of lockdown was late as compared to the implementation of it during the first wave. As a result, the number of infected cases decreases suddenly. One of the major reasons for decreasing the number of infected people is that India has invented the vaccine and started a test of it on the infected people and also on susceptible persons. As a result of it, at the end of October 2021 this situation comes under control in India.

Again, repeat the whole process as defined in both Algorithms 1 and 2 for the data of COVID-19 infected cases during second wave, the forecasted results for second wave in India with their linguistic variables is also shown in Table 6. Therefore, the forecasted values of infected cases during first wave and second wave are very close to the existing data of COVID-19. The results of the proposed

| Date       | Actual data | Chen [45] | Haurng [54] | Haurng [46] | Kuo et al. [31] | Proposed model |
|------------|-------------|-----------|-------------|-------------|-----------------|----------------|
| 21/09/1998 | 6861        | 6850      | 6850        | 6850        | 6852            | 6858           |
| 28/09/1998 | 6840        | 6850      | 670         | 6750        | 6843            | 6841.51        |
| MSE        | 9668.94     | 7856.5    | 5437.58     | 66.84       | 63.96           |                |
| RMSE       | 98.33       | 88.64     | 73.74       | 8.18        | 7.98            |                |

Table 5: Continued.
FTSPSOEFCM model of first wave and second wave are also depicted by graphical representation in Figures 3 and 4, which contain the forecasted values and actual values of infected cases during first wave and second wave, respectively.

Hence, the proposed FTSPSOEFCM model is well trained and suitable to forecast the possible third wave [57] of COVID-19 in India. In order to forecast the third wave of COVID-19, the forecasted data of second wave from the proposed model are considered as original data. When the number of infected cases and duration of the third wave are the approximate, which may vary with the actual cases in upcoming days. Furthermore, the impact of vaccination on infected cases and deaths during the upcoming third wave in India is also studied. The predicted newly COVID-19 infected cases during third wave for the upcoming days without vaccination impact with different values of $R (\approx 1.5, 2.0, 2.5)$, is shown in Table 7. The results of Table 7 for infected cases are also depicted graphically in Figure 5.

By the implementation of proposed model, it is observed that there is a sudden increment in the COVID-19 infected cases. As the value of $R$ increases from 1.5 to 2.5, the number of COVID-19 infected cases increases. From Table 7, when the value of $R$ is 1.5, the total number of infected cases as of after 15 days is around approx. 13081 and as per the proposed model, it is expected to increase upto 267187 for infected persons. Similarly, when the value of $R$ are 2.0 and 2.5, the maximum number of infected person are 353253 and 487941, respectively.

From both Figures 5 and 6, it can be concluded that, with increasing the value of $R$, the number of COVID-19 infected

### Table 6: Forecasted COVID-19 infected cases during first wave and second wave in India.

| Date       | Linguistic variable | Forecasted COVID-19 infected cases | Date       | Linguistic variable | Forecasted COVID-19 infected cases |
|------------|---------------------|-----------------------------------|------------|---------------------|-----------------------------------|
| 01/04/2020 | L(1)                | 419                               | 10/02/2021 | L(10)               | 15696                             |
| 11/04/2020 | L(3)                | 865                               | 20/02/2021 | L(13)               | 18465                             |
| 21/04/2020 | L(8)                | 1642                              | 02/03/2021 | L(15)               | 19544                             |
| 01/05/2020 | L(12)               | 2249                              | 12/03/2021 | L(29)               | 29391                             |
| 11/05/2020 | L(17)               | 3845                              | 22/03/2021 | L(66)               | 48182                             |
| 21/05/2020 | L(26)               | 6215                              | 01/04/2021 | L(109)              | 88944                             |
| 31/05/2020 | L(35)               | 8559                              | 11/04/2021 | L(126)              | 179809                            |
| 10/06/2020 | L(43)               | 11198                             | 21/04/2021 | L(141)              | 325646                            |
| 20/06/2020 | L(62)               | 16102                             | 01/05/2021 | L(156)              | 405470                            |
| 30/06/2020 | L(70)               | 20396                             | 11/05/2021 | L(147)              | 361449                            |
| 10/07/2020 | L(90)               | 29902                             | 21/05/2021 | L(135)              | 270193                            |
| 20/07/2020 | L(104)              | 38946                             | 31/05/2021 | L(120)              | 139777                            |
| 30/07/2020 | L(139)              | 57108                             | 10/06/2021 | L(112)              | 101743                            |
| 09/08/2020 | L(148)              | 66237                             | 20/06/2021 | L(97)               | 62872                             |
| 19/08/2020 | L(158)              | 73316                             | 30/06/2021 | L(93)               | 58500                             |
| 29/08/2020 | L(169)              | 84268                             | 10/07/2021 | L(69)               | 51388                             |
| 08/09/2020 | L(183)              | 95644                             | 20/07/2021 | L(72)               | 52022                             |
| 18/09/2020 | L(188)              | 98762                             | 30/07/2021 | L(69)               | 51393                             |
| 28/09/2020 | L(159)              | 75458                             | 09/08/2021 | L(35)               | 37322                             |
| 08/10/2020 | L(161)              | 76587                             | 19/08/2021 | L(54)               | 41696                             |
| 18/10/2020 | L(141)              | 62308                             | 29/08/2021 | L(77)               | 48470                             |
| 28/10/2020 | L(131)              | 49763                             | 08/09/2021 | L(77)               | 46860                             |
| 07/11/2020 | L(120)              | 45386                             | 18/09/2021 | L(43)               | 34589                             |
| 17/11/2020 | L(107)              | 38123                             | 28/09/2021 | L(24)               | 25357                             |
| 27/10/2020 | L(111)              | 40728                             | 08/10/2021 | L(23)               | 20125                             |
| 07/12/2020 | L(87)               | 25602                             | 18/10/2021 | L(9)                | 11214                             |
| 17/12/2020 | L(88)               | 26129                             | 28/10/2021 | L(12)               | 14982                             |
| 27/10/2020 | L(76)               | 19935                             | 07/11/2021 | L(8)                | 11915                             |
| 06/01/2021 | L(77)               | 20074                             | 17/11/2021 | L(8)                | 10957                             |
| 16/01/2021 | L(59)               | 14652                             | 27/11/2021 | L(2)                | 8875                              |
| 26/01/2021 | L(48)               | 11747                             | 07/12/2021 | L(2)                | 8589                              |

| MSE        | 9.05E+06            | MSE                               | 5.65E+07   |
| RMSE       | 3.01E+03            | RMSE                              | 7.52E+03   |
Figure 3: Graphical representations of forecasted and actual infected cases of COVID-19 during first wave.

Figure 4: Graphical representations of forecasted and actual infected cases of COVID-19 during second wave.
cases also increases. With the impact of vaccine on Indian people, the number infected cases are much lesser than the without the vaccination effect. Thus, the peak of COVID-19 infected during third wave is observed between 52 and 65 days, 45–60 days, and 40–55 days for 1.5, 2.0, and 2.5 respectively values of $R$, without the implementation of vaccination. Similarly, the peak of COVID-19 infected during third wave is observed between 70 and 85 days, 62–79 days and 52–70 days for 1.5, 2.0 and 2.5, respectively, values of $R$, with the implementation of vaccine.
6.4. Predicted COVID-19 Deaths during the Upcoming Third Wave in India. The number of predicted approximate COVID-19 deaths during the third wave in India is obtained. In the same way as processed above, before the forecasting of COVID-19 deaths during the upcoming third wave in India, the proposed FTSPSOEFCM model has been tested against official deaths data of COVID-19 during the first and second wave. The proposed model is applied to first and second wave data independently. According to the step 1 of Algorithm 1, the number of randomly generated cluster will be
same for first wave as well as second wave. The forecasted results of deaths during the first and second wave COVID-19 with their linguistic variables are shown in Table 9. From Table 9, it is observed that the forecasted results of deaths are very near about to the official data of COVID-19. The results of Table 9 are also depicted graphically in Figures 7 and 8. This shows that the proposed model is well trained for the prediction of COVID-19 deaths during the upcoming third wave in India. The number of deaths during the third wave is approximated values, which may vary with the actual deaths in upcoming days. The predicted death during third wave for the upcoming days without vaccination impact, with different value of $R (\approx 1.5, 2.0, 2.5)$ is shown in Table 10. The results of Table 10 for deaths are also depicted graphically in Figure 9.

From Table 10, when the value of $R$ is 1.5, the total number of deaths as of after 15 days is around approx. 513 and as per the proposed model, it is expected to increase up to 4529 for deaths due to COVID-19 in India. Similarly, when the value of $R$ are 2.0 and 2.5, the maximum number of deaths during COVID-19 third wave are 4984 and 6235 respectively.

The predicted death during third wave for the upcoming days with vaccination impact, with different value of $R (\approx 1.5, 2.0, 2.5)$ is shown in Table 11. The results of Table 11 for deaths are also depicted graphically in Figure 10.

From Table 11, when the value of $R$ is 1.5, the total number of deaths as of after 15 days is around approx. 487 and as per the proposed model, it is expected to increase up to 4529 for deaths due to COVID-19 in India. Similarly, when the value of $R$ are 2.0 and 2.5, the maximum number of deaths during COVID-19 third wave are 4420 and 5380, respectively.

From both Figures 9 and 10, it can be concluded that as the value of $R$ increases, the number of deaths due to COVID-19. Thus, the maximum number of deaths starts increasing during the period of 29–78 days, without the impact of vaccination. With the impact of vaccination, the maximum number of deaths was observed after 36 days from the mid of December.

As the value of performance parameters are high, that’s why to check the accuracy of the proposed model, the error percentage is calculated between the official data and forecasted data of COVID-19 infected cases as well as deaths during the peak time of first wave and second wave in India. The results are shown in Table 12. By the analysis of this table, it can be concluded that the error between

Table 9: Forecasted COVID-19 deaths during first wave and second wave in India.

| Date       | Linguistic variable | Forecasted COVID-19 deaths | Date       | Linguistic variable | Forecasted COVID-19 deaths |
|------------|---------------------|----------------------------|------------|---------------------|----------------------------|
| 01/04/2020 | L(1)                | 5                          | 10/02/2021 | L(11)               | 95                          |
| 11/04/2020 | L(10)               | 34                         | 20/02/2021 | L(5)                | 74                          |
| 21/04/2020 | L(13)               | 46                         | 02/03/2021 | L(8)                | 83                          |
| 01/05/2020 | L(19)               | 70                         | 12/03/2021 | L(18)               | 169                         |
| 11/05/2020 | L(20)               | 66                         | 22/03/2021 | L(26)               | 226                         |
| 21/05/2020 | L(36)               | 133                        | 01/04/2021 | L(64)               | 613                         |
| 31/05/2020 | L(51)               | 207                        | 11/04/2021 | L(97)               | 1049                        |
| 10/06/2020 | L(75)               | 343                        | 21/04/2021 | L(116)              | 2360                        |
| 20/06/2020 | L(66)               | 293                        | 01/05/2021 | L(137)              | 3944                        |
| 30/06/2020 | L(107)              | 490                        | 11/05/2021 | L(142)              | 4457                        |
| 10/07/2020 | L(111)              | 505                        | 21/05/2021 | L(142)              | 4604                        |
| 20/07/2020 | L(122)              | 581                        | 31/05/2021 | L(124)              | 3193                        |
| 30/07/2020 | L(141)              | 769                        | 10/06/2021 | L(131)              | 3824                        |
| 09/08/2020 | L(158)              | 811                        | 20/06/2021 | L(109)              | 1834                        |
| 19/08/2020 | L(156)              | 777                        | 30/06/2021 | L(101)              | 1320                        |
| 29/08/2020 | L(154)              | 1017                       | 10/07/2021 | L(96)               | 1216                        |
| 08/09/2020 | L(163)              | 1181                       | 20/07/2021 | L(140)              | 4239                        |
| 18/09/2020 | L(169)              | 1295                       | 30/07/2021 | L(80)               | 839                         |
| 28/09/2020 | L(140)              | 1200                       | 09/08/2021 | L(51)               | 617                         |
| 08/10/2020 | L(155)              | 1392                       | 19/08/2021 | L(75)               | 784                         |
| 18/10/2020 | L(120)              | 807                        | 29/08/2021 | L(72)               | 768                         |
| 28/10/2020 | L(109)              | 936                        | 08/09/2021 | L(45)               | 580                         |
| 07/11/2020 | L(116)              | 978                        | 18/09/2021 | L(42)               | 547                         |
| 17/11/2020 | L(96)               | 682                        | 28/09/2021 | L(51)               | 478                         |
| 27/10/2020 | L(101)              | 696                        | 08/10/2021 | L(31)               | 249                         |
| 07/12/2020 | L(79)               | 596                        | 18/10/2021 | L(21)               | 149                         |
| 17/12/2020 | L(73)               | 552                        | 28/10/2021 | L(77)               | 534                         |
| 27/10/2020 | L(62)               | 383                        | 07/11/2021 | L(32)               | 267                         |
| 06/01/2021 | L(51)               | 324                        | 17/11/2021 | L(41)               | 297                         |
| 16/01/2021 | L(44)               | 283                        | 27/11/2021 | L(42)               | 358                         |
| 26/01/2021 | L(33)               | 242                        | 07/12/2021 | L(25)               | 234                         |

MSE: $3.68E+04$, RMSE: $1.92E+04$
COVID-19 infected cases and deaths are minor. Hence, the proposed model can be applied for the prediction of COVID-19 infected cases and deaths in India during the upcoming third wave.

7. Analysis of Variance

Tables 6 and 9 shows that the proposed model forecasts the COVID-19 infected cases and deaths during the first and...
In order to verify the statistical validation of proposed model, one-way ANOVA test has been carried out in MINITAB-20, for the COVID-19 infected cases and deaths during peak time of first and second wave. Table 13 represents the results of ANOVA test. By the analysis of Table 13, it can be concluded that $P$ value is 0.8926 which is greater than the F-value 0.0188 and $P$ value is 0.8826 which is also greater than the F value 0.0224 at 95% confidence level for COVID-19 infected cases and deaths, respectively, during the peak time of first wave and second wave of COVID-19 in India.

**Table 10:** Forecasts COVID-19 deaths without vaccination impact during the upcoming third wave in India.

| Days | $R \approx 1.5$ | $R \approx 2.0$ | $R \approx 2.5$ |
|------|----------------|----------------|----------------|
| 1    | 278            | 291            | 318            |
| 8    | 318            | 378            | 413            |
| 15   | 513            | 582            | 631            |
| 22   | 680            | 624            | 952            |
| 29   | 755            | 738            | 980            |
| 36   | 784            | 1057           | 1098           |
| 43   | 1184           | 2026           | 1658           |
| 50   | 2257           | 3176           | 3160           |
| 57   | 3502           | 3952           | 4784           |
| 64   | 3749           | 4841           | 5249           |
| 71   | 4098           | 4419           | 5998           |
| 78   | 3509           | 3428           | 4913           |
| 85   | 2783           | 2382           | 3896           |
| 92   | 2012           | 1311           | 2817           |
| 99   | 1214           | 817            | 1700           |
| 106  | 1056           | 664            | 1476           |
| 113  | 707            | 508            | 990            |
| 120  | 413            | 388            | 578            |
| 127  | 356            | 331            | 498            |
| 134  | 244            | 227            | 342            |

**Figure 9:** Graphical representations of foreasted COVID-19 deaths during upcoming third wave in India without the impact of vaccination ($R \approx 1.5, 2.0, 2.5$).
Forecasted COVID-19 deaths during third wave with vaccination effect

![Graphical representations of forecasted COVID-19 deaths during upcoming third wave in India with the impact of vaccination](R≈1.5, 2.0, 2.5).

| Days | $R \approx 1.5$ | $R \approx 2.0$ | $R \approx 2.5$ |
|------|----------------|----------------|----------------|
| 1    | 258            | 276            | 293            |
| 8    | 296            | 354            | 391            |
| 15   | 487            | 517            | 622            |
| 22   | 458            | 566            | 789            |
| 29   | 541            | 777            | 939            |
| 36   | 596            | 877            | 1071           |
| 43   | 642            | 880            | 1310           |
| 50   | 806            | 1076           | 1334           |
| 57   | 952            | 1274           | 1698           |
| 64   | 1023           | 2021           | 2941           |
| 71   | 1498           | 3176           | 5380           |
| 78   | 2761           | 3200           | 4057           |
| 85   | 3463           | 2423           | 2782           |
| 92   | 2682           | 1291           | 1530           |
| 99   | 1245           | 1019           | 1348           |
| 106  | 937            | 817            | 949            |
| 113  | 921            | 677            | 563            |
| 120  | 678            | 450            | 479            |
| 127  | 402            | 374            | 381            |
| 134  | 342            | 318            | 276            |
8. Conclusion

In India, the number of infected cases and deaths due to COVID-19 is increasing daily. To determine the plan and policies, the study of the spreading of COVID-19 in India becomes crucial. As the second wave is not entirely ruled out, there is a possibility for the third wave. The peak during the upcoming third wave and determination of the approximate maximum number of COVID-19 infected cases and deaths at a particular day is crucial for India. The main outcomes of the proposed model are discussed as follows:

(i) The present study presents a hybrid FTS forecasted model based on the PSO and EFCM algorithm, named as FTSPSOEFCM.

(ii) Established fitness function for PSO to increase its convergence speed and basic FCM is extended by using an exponential function to tolerate the effect of outliers, named as EFCM.

(iii) The effectiveness of the proposed model has been illustrated by testing it on enrolment data of Alabama University, TAIFEX data, first and second wave COVID-19 data and the obtained results are very close to the existing data with less error rate.

(iv) Developed an efficient model to predict the duration and approximate COVID-19 infected cases and deaths in India during the upcoming third wave because there is a possibility of a third wave in India. This study demonstrates that third wave of COVID-19 could occur in India, while also illustrating that it is unlikely for any such resurgence to be as large as the second wave.

(v) Proposed model predicts that the peak of third wave will occur approximately after 40–70 days from the mid of December.

(vi) Furthermore, the impact of vaccination on infected cases and deaths during the upcoming third wave in India is also studied. The peak of COVID-19 infected during the third wave is observed between 52 and 65 days, 45–60 days and 40–55 days for different values of $R$, without the impact of vaccination. With the implementation of the vaccine on the Indian people, the peak of COVID-19 infected during third wave will be shifted in forward direction and observed in between 70 and 85 days, 62–79 days and 52–70 days. The results obtained in this article for Indian data can be beneficial for the health authorities and the decision-making process.

(vii) Furthermore, the COVID-19 deaths will also be estimated during the upcoming wave. The results of the predicted infected cases and deaths during the upcoming third wave in India are shown graphically.

Thus, the proposed model is essential for government and health care decision-makers to make protection plans during the impending pandemic. But the duration and COVID-19 infected cases and deaths during the third wave

\[\text{Table 12: Comparison of official and forecasted data of COVID-19 infected cases and deaths during peak time of first wave and second wave in India.}\]

| Date       | Official data Infected cases | Forecasted value | Error percentage | Official data Deaths | Forecasted value | Error percentage |
|------------|------------------------------|------------------|------------------|----------------------|------------------|------------------|
|            | First wave 2020             |                  |                  |                      |                  |                  |
| 02/09/2020 | 82865                      | 88654            | 6.9861           | 1026                 | 1100             | 7.2125           |
| 09/09/2020 | 95536                      | 101325           | 6.0595           | 1168                 | 1242             | 6.3356           |
| 16/09/2020 | 97860                      | 103649           | 5.9156           | 1140                 | 1214             | 6.4912           |
| 23/09/2020 | 86703                      | 92492            | 6.6768           | 1123                 | 1197             | 6.5895           |
| 30/09/2020 | 86748                      | 92537            | 6.6734           | 1179                 | 1258             | 6.7006           |
|            | Second wave 2021            |                  |                  |                      |                  |                  |
| 22/04/2021 | 332531                     | 342425           | 2.9754           | 2257                 | 2301             | 1.9495           |
| 29/04/2021 | 386773                     | 396667           | 2.5581           | 3502                 | 3561             | 1.6848           |
| 06/05/2021 | 414280                     | 427174           | 3.1124           | 3923                 | 4182             | 6.6021           |
| 13/05/2021 | 343005                     | 355899           | 3.7591           | 4000                 | 4101             | 2.5250           |
| 20/05/2021 | 259242                     | 272136           | 4.9737           | 4209                 | 4315             | 2.5184           |

\[\text{Table 13: ANOVA analysis of COVID-19 infected cases and deaths during peak time of first wave and second wave in India.}\]

| Source | DF | Adj. SS       | Adj. MS       | F-value | P-value |
|--------|----|---------------|---------------|---------|---------|
| Proposed model versus official data for infected cases in India during first and second wave | Between-group | 1 | 3.82E + 08  | 3.82E + 08 | 0.0188  | 0.8926 |
|        | Within-group | 18 | 3.67E + 11   | 2.04E + 10 |         |         |
|        | Total       | 19 | 3.67E + 11   | 3.67E + 11 |         |         |
| Proposed model versus official data for deaths in India during first and second wave | Between-group | 1 | 4.46E + 04   | 4.46E + 04 | 0.0224  | 0.8826 |
|        | Within-group | 18 | 3.57E + 07   | 1.99E + 06  |         |         |
|        | Total       | 19 | 3.58E + 07   | 3.58E + 07  |         |         |
may eventually decrease or increase. It may be varied from the actual trend of third wave in India. This model is of urgent significance to combat the crisis effectively as it will help decision makers for estimating the number of COVID-19 infected cases and deaths. The proposed model may be enabling for calculating important parameter such as infection rates and deaths rates, which will help decision maker to have a more accurate grasp of the transmission trained for other disease, if occurs in future. As a future direction, other meta-heuristic algorithms can be used to optimize the proposed forecasting model.

Data Availability

The numerical data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

On behalf of all authors, the corresponding author states that there are no conflicts of interest.

References

[1] S. Deb and M. Majumdar, “A Time Series Method to Analyze Incidence Pattern and Estimate Reproduction Number of COVID-19,” 2020, https://arxiv.org/abs/2003.10655.

[2] S. Mandal, T. Bhatnagar, N. Arinaminpathy et al., “Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: a mathematical model-based approach,” Indian Journal of Medical Research, vol. 151, no. 1-2, pp. 190–199, 2020.

[3] J. Bhola, V. R. Venkateswaran, and M. Kou, Corona Epidemic in Indian Context: Predictive Mathematical Modelling, MedRxiv, New York, NY, USA, 2020.

[4] S. Mondal and S. Ghosh, Fear of Exponential Growth in COVID-19 Data of India and Future Sketching, MedRxiv, New York, NY, USA, 2020.

[5] M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak, and T. K. Kar, “A model based study on the dynamics of COVID-19: prediction and control,” Chaos, Solitons, and Fractals, vol. 136, Article ID 109889, 2020.

[6] M. Maleki, M. R. Mahmoudi, M. H. Heydari, and K. Pho, “Modeling and forecasting the spread and death rate of coronavirus (COVID-19) in the world using time series models,” Chaos, Solitons & Fractals, vol. 140, Article ID 110151, 2020.

[7] M. R. Mahmoudi, D. Baleanu, Z. Mansor, B. A. Tuan, and K. H. Pho, “Fuzzy clustering method to compare the spread rate of COVID-19 in the high risks countries,” Chaos, Solitons, and Fractals, vol. 140, Article ID 110230, 2020.

[8] N. Kumar and S. Susan, “Particle swarm optimization of partitions and fuzzy order for fuzzy time series forecasting of COVID-19,” Chaos, Solitons & Fractals, vol. 110, 2021.

[9] M. A. Elleuch, A. B. Hassena, M. Abdelhedi, and F. S. Pinto, “Real-time prediction of COVID-19 patients health situations using artificial neural networks and fuzzy interval mathematical modeling,” Applied Soft Computing, vol. 110, 2021.

[10] T. Chakraborty and I. Ghosh, “Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis,” Chaos, Solitons, and Fractals, vol. 135, Article ID 109850, 2020.

[11] R. Salgotra, M. Gandomi, and A. H. Gandomi, “Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming,” Chaos, Solitons, and Fractals, vol. 138, Article ID 109945, 2020.

[12] C. Kavitha, A. Gowrisankar, and S. Banerjee, “The second and third waves in India: when will the pandemic be culminated?” The European Physical Journal Plus, vol. 136, pp. 1–12, 2021.

[13] S. Shringi, H. Sharma, P. N. Rath, J. C. Bansal, and A. Nagar, “Modified SIR-V model for COVID-19 spread prediction for northern and southern states in India,” Chaos, Solitons & Fractals, vol. 148, Article ID 111039, 2021.

[14] N. V. Tinh, “Forecasting of COVID-19 confirmed cases in Vietnam using fuzzy time series model combined with particle swarm optimization,” Computational Research Progress in Applied Science and Engineering, vol. 6, no. 2, pp. 114–120, 2020.

[15] A. A. Husain, B. Surarso, Farikhin, and B. Irawanto, “Forecasting model of COVID-19 cases using fuzzy time series using percentage change,” Journal of Physics: Conference Series, vol. 1943, 2021.

[16] M. Amouch and N. Karim, “Modeling the dynamic of COVID-19 with different types of transmissions,” Chaos, Solitons & Fractals, vol. 150, 2021.

[17] N. Kumar and H. Kumar, “A Novel Hybrid Fuzzy Time Series Model for Prediction of COVID-19 Infected Cases and Deaths in India,” ISA Transactions, vol. 21, 2021.

[18] A. H. Gandomi, X.-S. Yang, S. Talatahari, and A. H. Alavi, “Metaheuristic algorithms in modelling and optimization,” in Metaheuristic Applications In Structures And Infrastructures, pp. 1–24, Elsevier, Amsterdam, Netherlands, 1st edition, 2013.

[19] K. Hussain, M. N. Mohd Salleh, S. Cheng, and Y. Shi, “Metaheuristic research: a comprehensive survey,” Artificial Intelligence Review, vol. 52, no. 4, pp. 2191–2233, 2019.

[20] B. Abdollahzadeh, F. S. Gharehchopogh, and S. Mirjalili, “African vultures optimization algorithm: a new nature-inspired metaheuristic algorithm for global optimization problems,” Computers & Industrial Engineering, vol. 158, pp. 1–37, 2021.

[21] B. Abdollahzadeh, F. Soleimanian Gharehchopogh, and S. Mirjalili, “Artificial gorilla troops optimizer: a new nature-inspired metaheuristic algorithm for global optimization problems,” International Journal of Intelligent Systems, vol. 36, no. 10, pp. 5887–5938, 2021.

[22] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in Proceedings of the IEEE International Conference on Neural Networks, pp. 1942–1948, Perth, Australia, December 1995.

[23] D. Tian and Z. Shi, “MPSO: modified particle swarm optimization and its applications,” Swarm and Evolutionary Computation, vol. 41, pp. 49–68, 2018.

[24] Q. Song and B. S. Chissom, “Forecasting enrollments with particle swarm optimization,” in Metaheuristical algorithms in modelling and optimization,” in Metaheuristic Applications In Structures And Infrastructures, pp. 1–24, Elsevier, Amsterdam, Netherlands, 1st edition, 2013.

[25] A. H. Gandomi, X.-S. Yang, S. Talatahari, and A. H. Alavi, “Metaheuristic algorithms in modelling and optimization,” in Metaheuristic Applications In Structures And Infrastructures, pp. 1–24, Elsevier, Amsterdam, Netherlands, 1st edition, 2013.

[26] S.-T. Li, Y.-C. Cheng, and S.-Y. Lin, “A FCM-based deterministic forecasting model for fuzzy time series forecasting based on improved fuzzy c-means clustering algorithm,” in Proceedings of the 2018 IEEE International Conference on Progress in Informatics and Computing (PIC), pp. 80–84, IEEE, Suzhou, China, December 2018.
[28] N. Kumar, H. Kumar, and K. Sharma, “Extension of FCM by introducing new distance metric,” SN Applied Sciences, vol. 2, no. 4, pp. 1–21, 2020.

[29] Y. Zhang, H. Qu, W. Wang, and J. Zhao, “A novel fuzzy time series forecasting model based on multiple linear regression and time series clustering,” Mathematical Problems in Engineering, vol. 2020, Article ID 9546792, 17 pages, 2020.

[30] I.-H. Kuo, S.-J. Horng, T.-W. Kao, T.-L. Lin, C.-L. Lee, and Y. Pan, “An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization,” Expert Systems with Applications, vol. 36, no. 3, pp. 6108–6117, 2009.

[31] I.-H. Kuo, S.-J. Horng, Y.-H. Chen et al., “Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization,” Expert Systems with Applications, vol. 37, no. 2, pp. 1494–1502, 2010.

[32] Y.-L. Huang, S.-J. Horng, M. He et al., “A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization,” Expert Systems with Applications, vol. 38, no. 7, pp. 8014–8023, 2011.

[33] P. Singh and B. Borah, “Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization,” International Journal of Approximate Reasoning, vol. 55, no. 3, pp. 812–833, 2014.

[34] Y.-L. Huang, W.-K. Hsieh, Shih-Wei, and Lin, "A trend based forecasting model using fuzzy time series and PSO algorithm," in Proceedings of the 2019 International Conference on Computation, Communication and Engineering, pp. 21–24, Fujian, China, November 2019.

[35] N. V. Tinh, “Enhanced forecasting accuracy of fuzzy time series model based on combined fuzzy c-mean clustering with particle swarm optimization,” International Journal of Computational Intelligence and Applications, vol. 19, no. 2, pp. 1–26, 2020.

[36] L. Basnarkov, “SEAIR epidemic spreading model of COVID-19,” Chaos, Solitons & Fractals, vol. 142, pp. 1–12, 2020.

[37] S. Kumar and S. S. Gangwar, “A fuzzy time series forecasting method induced by intuitionistic fuzzy sets,” International Journal of Modelling, Simulation, and Scientific Computing, vol. 6, no. 4, pp. 1–23, 2015.

[38] Y. N. Wang, Y. Lei, X. Fan, and Y. Wang, “Intuitionistic fuzzy time series forecasting model based on intuitionistic fuzzy reasoning,” Mathematical Problems in Engineering, vol. 2016, Article ID 5035160, 12 pages, 2016.

[39] S. Karthick and N. Gomathi, “Galactic swarm-improved whale optimization algorithm-based resource management in internet of things,” International Journal of Communication Systems, vol. 35, no. 3, 2021.

[40] C. Kocak, E. Egrioglu, and E. Bas, “A new deep intuitionistic fuzzy time series forecasting method based on long short-term memory,” The Journal of Supercomputing, vol. 77, no. 6, pp. 6178–6196, 2021.

[41] E. H. Ruspini, “A new approach to clustering,” Information and Control, vol. 15, no. 1, pp. 22–32, 1969.

[42] E. H. Ruspini, “Numerical methods for fuzzy clustering,” Information Sciences, vol. 2, no. 3, pp. 319–350, 1970.

[43] J. C. Bezdek, “Objective function clustering,” in Pattern Recognition with Fuzzy Objective Function Algorithms, pp. 43–93, Springer, Boston, MA, USA, 1981.

[44] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” Journal of Cybernetics, vol. 3, no. 3, pp. 32–57, 1973.