Applied Soft Computing 111 (2021) 107735

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Grey forecasting models based on internal optimization for Novel Corona virus (COVID-19)

Akash Saxena

Swami Keshvanand Institute of Technology, Management & Gramothan, Jaipur, Rajasthan, India

ARTICLE INFO

Article history:
Received 19 August 2020
Received in revised form 5 July 2021
Accepted 15 July 2021
Available online 25 July 2021

Keywords:
Corona
Grey forecasting models
Optimization
Mean Absolute Percentage Error (MAPE)

ABSTRACT

Pandemic forecasting has become an uphill task for the researchers on account of the paucity of sufficient data in the present times. The world is fighting with the Novel Coronavirus to save human life. In a bid to extend help to the concerned authorities, forecasting engines are invaluable assets. Considering this fact, the presented work is a proposal of two Internally Optimized Grey Prediction Models (IOGMs). These models are based on the modification of the conventional Grey Forecasting model (GM(1,1)). The IOGMs are formed by stacking infected case data with diverse overlap periods for forecasting pandemic spread at different locations in India. First, IOGM is tested using time series data. Its two models are then employed for forecasting the pandemic spread in three large Indian states namely, Rajasthan, Gujarat, Maharashtra and union territory Delhi.

Several test runs are carried out to evaluate the performance of proposed grey models and conventional grey models GM(1,1) and NGM(1,1,k). It is observed that the prediction accuracies of the proposed models are satisfactory and the forecasted results align with the mean infected cases. Investigations based on the evaluation of error indices indicate that the model with a higher overlap period provides better results.

1. Introduction

The month of December in 2019 was a watershed in the history of mankind with a series of cases reported suffering from inexplicable pneumonia at Wuhan, Hubei, China. The clinical investigation of the cases revealed their resemblance with viral pneumonia. Further, deep sequencing analysis from lower respiratory tract samples revealed that the root cause of this pneumonia is a new virus. People's Republic of China (Centre for Disease Control) attributed the cause of this unidentified pneumonia to a novel form of Coronavirus and later World Health Organization (WHO) declared it as COVID-19. Coronaviruses came from a family of Coronaviridae and the order Nidovirales. These are enveloped non-segmented positive-sense RNA viruses that are broadly distributed in humans and other mammals [1]. The WHO declared the deadly disease as a global pandemic on March 11, 2020 [2]. The initial cluster was developed in a local seafood market in Wuhan. An epidemiological alert was issued by WHO at the end of April 2020 which spread like a wildfire world over. It ultimately turn out to be the nemesis for the mankind in the form of pandemic. With the spread of this disease, implications of the pandemic appeared as loss of life and several health-related issues. For combating such situations, many scientists have come forward to help the community by forecasting certain parameters so that authorities concerned can frame preventive healthcare policies. The prediction models based on Data science and Machine Learning Methods (MLMs) have assumed importance in providing better understanding about growth and trend of pandemic with respect to time.

The forecasting of pandemic spread is quite challenging due to dependency of the pandemic spread on several factors. These factors are governed by the psychological behaviour of the community, combating strategies deployed by the authorities and of course,volatility and reliability of the data [3]. Moreover, the future does not repeat it in the same way as did in the past. Hence, considering the same policies which were employed in the past for forecasting the pandemic spread may not be applicable directly.

Evolving fear amongst the population due to the enhanced death rate and concerns of politicians to take adequate steps as preventive measures create a strong perception. Hence, this pandemic has emerged as an infodemic as well. However, cut-throat competition among the vaccine manufacturers is a welcome move and it is furthering the cause of protecting the lives of people against COVID. These facts set the strong foundation for discussion and research in the area of forecasting the pandemic spread and healthcare-related parameters.

Recently many prediction approaches have been reported by the researchers for fairly accurate forecasting about the spread

E-mail addresses: akash@skit.ac.in, aakash.saxena@hotmail.com.

https://doi.org/10.1016/j.asoc.2021.107735

Nomenclature

α	Background value coefficient
a	Grey development coefficient
b	Grey control parameter
$C^{(1)}_{(W,\mu)}$	First order accumulated sequence of mean infected cases for a week considering overlap period μ
$Z^{(1)}(m)$	Background value at mth instant
$\hat{C}^{(0)}_{(W,\mu)}(m)$	Forecasted value at mth instant
X_{Bi}	ith Benchmark time series
X^{in}_{Bi}(j)	jth In-sample for ith benchmark time series
X^{out}_{Bi}(j)	jth Out-sample for ith benchmark time series
ϵ	Error tolerance
$\hat{x}^{(0)}(m)$	Forecasted value of mth element of time series
$x^{(0)}(m)$	Actual value of mth element of time series
$\delta(m)$	Error in prediction of mth element of time series
μ	Period of overlap
$C^{(0)}_{(W,\mu)}$	Mean infected cases during time span of a week
$\hat{C}^{(0)}_{(W,6)}(m)$	Forecasted value at m^{th} instant considering 6 days of overlap
$\hat{C}^{(0)}_{(W,5)}(m)$	Forecasted value at m^{th} instant considering 5 days of overlap

of COVID, prediction of recovery rate and death rate. In Ref. [4], Maximum-Hasting (MH) parameter estimation method and the modified Susceptible Exposed Infectious Recovered (SEIR) model for prediction of COVID was developed. A study on the worst-hit states of India has been conducted in Ref. [5]. The study was based on system modelling and identification techniques. Time series-based approach in amalgamation with Long Short Term Associated Memory (LSTM) for prediction of COVID has been employed in Canada [6]. A similar approach based on LSTM has been reported in Ref. [7]. SIR model-based prediction has been developed in Ref. [8] for spread of Corona in Italy, China and France. Likewise, a prediction analysis of COVID based on deep learning models has also been presented in the study [9]. Real-time forecasting of COVID epidemic in China has been carried out in Ref. [10]. The Akaike Information Criterion (AIC) for model selection has been employed by the authors of Ref. [11] for comparing the SIR model and SEIR model. Along with this analysis authors also described and mentioned that forecasting a pandemic is not an easy task.

From this analysis it can be observed that forecasting of pandemic spread is a current research domain and new forecasting methods are welcomed to encounter this deadly disease.

In the past, various models based on grey prediction theories have been employed to predict various important parameters such as energy consumption [12], fuel consumption [13], load growth [14,15] and many more. Research in grey forecasting theory can be subdivided into four parts broadly:

1. Change in accumulation sequence operation.
2. Change in Grey equations.
3. Transformation of the original series with the use of some mathematical operators.
4. Parameter optimization by heuristic and metaheuristic methods.

Grey systems are defined as the systems that possess lacuna in information [16]. In other words, Grey systems contain some known information and a part of the information is unknown. Grey systems use accumulation operators to deal with randomness in data. Grey prediction Models (GPMs) have an inherent characteristic to transfer hidden original irregular data to strong regular data by using accumulation operation. The operator is known as Accumulating Generation Operator (AGO). For improving system performance, several research attempts have been made to formulate fractional-order accumulation operators. A concept of fractional order accumulation operator was put forward in researches [17] and [18]. An experiment of putting weights in accumulation sequence has also been conducted in the Ref. [19]. Another interesting approach based on time delay and multiple fractional-order grey system for forecasting the natural gas consumption in China was employed in Ref. [20].

Another interesting domain of research in GPMs is the transformation of the original data series into some other representative series. This transformation helps GPMs to achieve higher accuracy. In [21] a modified model of the Grey system has been presented and the transformation of the original time series was conducted with the help of logarithmic transformation. A technique based on background value optimization and data transformation was put forward in research for forecasting the energy consumption in Shanghai [22].

GPMs are based on Grey mathematical equations. Hence, the other aspect of conducted research in this area is related to change in the Grey system equations. A good example of this can be found in the development of Novel Grey Prediction Model (NGM) in Ref. [23]. An additional constant term has also been added in Grey system equation in order to overcome shortcomings of NGM and it has been named as NGM(1,1,k,c) [24]. Apparently, classical discrete model of Grey prediction can also be determined by changing the original Grey equations [25]. A novel discrete model was proposed for forecasting the CO_2 emission in China [26]. NGM method has been applied for forecasting the consumption of natural gas in China [27]. An alternative approach based on integrating heuristic time series and fuzzy theory for forecasting of the renewable energy in Taiwan was described in [28].

Optimization-based approaches and especially the approaches which are based on some nature-inspired algorithms have been employed with GPMs in recent years. These approaches are based on the parameter estimation of the Grey models. A novel time delay forecasting model based on a nature-inspired optimizer was performed in [29]. Further, Refs. [12] and [30] are fine examples of such approaches.

In addition to these published approaches, some experiments have been done to alter initial conditions of the Grey model in Ref. [26]. Authors of the paper employed alterable weighted coefficients in initial conditions. Another application of Grey prediction model for predicting sales in global integrated circuit industry is seen in Ref. [31]. Further, the Grey model has also been applied in power demand forecasting in Ref. [32]. The author employed residual modification with an artificial neural network for the modification of GM(1,1) model.

From the literature review of GPMs, the motivation for employing GPMs in pandemic forecasting is very clear and pragmatic. However, in some studies, it has been reported that GM(1,1) models fail to predict with required accuracy when data mutate swiftly or the associated variables are volatile.

In the past, it has been identified that forecasting accuracy of the grey models can be questionable when the initial conditions and starting points are not chosen correctly [15]. From this point of view, it is pragmatic that the involvement of optimization can enhance the forecasting accuracy. This involvement can be done at two levels during grey forecast. First, at the macro level by choosing the external optimization parameters such as data or selection of time series to develop different forecasting strategies. Secondly, by developing an internal optimization routine that integrates a few changes in forecast modelling and try to reduce the error between measured and sample data. Moreover, some researchers have emphasized in applying corrections in the internal parameter aggregation process by modifying the grey equation to achieve better results.

Considering a variable (infectious cases) as a grey variable that increases with every passing day and mutates swiftly, is difficult to forecast. Hence, the presented work primarily focuses on an internal optimization model that aims at enhancing the forecasting accuracy by choosing internal processing parameters through the optimization process. The research objectives of the paper are as follows:

1. To investigate the applicability of the GPMs on variety of benchmark time series data.
2. To propose the prediction model based on internal optimization and analyse the results based on average ranks obtained by stacking the forecaster's performance chronologically.
3. To employ the proposed internal optimization-based model and other grey models on the real data by taking two different overlap periods and construct two grey prediction models for forecasting pandemic spread in different hot-spots of India.
4. To evaluate the performance of these pandemic forecasting models based on error indices obtained for individual cities, average values of error for a particular city for both models.

Remaining part of the paper is organized as follows: In Section 2, proposal of IOGM is presented and explained. In Section 3, verification of proposed model is conducted on benchmark data series and comparative analysis is presented. Section 4, presents the simulation and results analysis of proposed grey models on different parts of India. Finally, Section 5, presents the concluding remarks of the work and throws light on the future directions of the research work.

2. Grey internal optimization prediction model

In this section, basics of GPM and its application for COVID-19 spread prediction are explained.

2.1. GM(1,1) model

Based on the above explanation, following mathematical expressions are considered for the proposed grey forecasting model. Following steps are followed for constructing forecasting engine.

The mean values of infected cases is considered in construction the initial time series. $C^{(0)}_{(W,\mu)}$ is representative denominator of time series. For the evolution of this series, successive elements are calculated based on overlap period of one week (7 days).

In general, consider a time series having an overlap period "μ" for the duration 'W'. The time series comprises of 'k' elements. The representation of this data series can be given as follows:

$$C^{(0)}_{(W,\mu)} = \left(C^{(0)}_{(W,\mu)}(1), C^{(0)}_{(W,\mu)}(2), C^{(0)}_{(W,\mu)}(3), \ldots\ldots\ldots C^{(0)}_{(W,\mu)}(k)\right) \quad (1)$$

By obtaining a one-time accumulating generation operation, the following series can be generated.

$$C^{(1)}_{(W,\mu)} = \left(C^{(1)}_{(W,\mu)}(1), C^{(1)}_{(W,\mu)}(2), C^{(1)}_{(W,\mu)}(3), \ldots\ldots\ldots C^{(1)}_{(W,\mu)}(k)\right) \quad (2)$$

$$C^{(1)}_{(W,\mu)}(m) = \sum_{i=1}^{m} C^{(0)}_{(W,\mu)}(i) \quad (3)$$

Where m=1, 2, 3.....k.

$$C^{(0)}_{(W,\mu)}(m) + aZ^{(1)}(m) = b \quad (4)$$

Where 'a' is Grey development coefficient and 'b' is Grey control parameter (driving coefficient). It is to be noted that the value of 'a' has a potential impact on background values (Z) of Grey derivatives, hence, the forecasted values get compromised due to the large value of 'a'.

$$Z^{(1)}(m) = (1-\alpha)C^{(1)}_{(W,\mu)}(m) + \alpha C^{(1)}_{(W,\mu)}(m-1) \quad (5)$$

Where m=2, 3.....k. Here, α is the background value production coefficient. The values of this coefficient should be optimized between an interval of [0, 1]. Further, the native Grey Model (1, 1) can be derived while keeping $\alpha = 0.5$. The following expression is for background values of grey derivatives for the native GM(1,1) model.

$$Z^{(1)}(m) = (0.5)C^{(1)}_{(W,\mu)}(m) + (0.5)C^{(1)}_{(W,\mu)}(m-1) \quad (6)$$

The expression (4) can be solved with the help of least square estimation method and expression for Grey development coefficient and driving coefficient can be expressed as follows:

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y \quad (7)$$

$$B = \begin{bmatrix} -1\times((1-\alpha)\times C^{(1)}_{(W,\mu)}(2) + \alpha\times C^{(1)}_{(W,\mu)}(1)) & 1 \\ -1\times((1-\alpha)\times C^{(1)}_{(W,\mu)}(3) + \alpha\times C^{(1)}_{(W,\mu)}(2)) & 1 \\ \vdots & \vdots \\ -1\times((1-\alpha)\times C^{(1)}_{(W,\mu)}(m) + \alpha\times C^{(1)}_{(W,\mu)}(m-1)) & 1 \\ \vdots & \vdots \\ -1\times((1-\alpha)\times C^{(1)}_{(W,\mu)}(k) + \alpha\times C^{(1)}_{(W,\mu)}(k-1)) & 1 \end{bmatrix} \quad (8)$$

in simplified form it can be written as

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ \vdots & \vdots \\ -Z^{(1)}(k) & 1 \end{bmatrix} \quad (9)$$

$$Y = \begin{bmatrix} C^{(0)}_{(W,\mu)}(2) \\ C^{(0)}_{(W,\mu)}(3) \\ \vdots \\ C^{(0)}_{(W,\mu)}(k) \end{bmatrix} \quad (10)$$

The solution of Eq. (4) can be written as

$$\hat{C}^{(1)}_{(W,\mu)}(m) = \left(C^{(0)}_{(W,\mu)}(1) - \frac{b}{a}\right) \times e^{-a\times(m-1)} + \frac{b}{a} \quad (11)$$

Where $\hat{C}^{(1)}_{(W,\mu)}(m)$ is the m^{th} associated value. For obtaining predicted values of original time series Inverse Accumulation generation operation is required and can be represented as per following equations.

$$\hat{C}^{(0)}_{(W,\mu)}(1) = C^{(0)}_{(W,\mu)}(1) \quad (12)$$

This expression holds good for m=1. Generalized equation can be written as

$$\hat{C}^{(0)}_{(W,\mu)}(m) = \hat{C}^{(1)}_{(W,\mu)}(m) - \hat{C}^{(1)}_{(W,\mu)}(m-1) \quad (13)$$

Eq. (13) is the generalized expression for m = 2, 3, ……k. after rearranging the expressions one can get a generalized expression for forecasted values at m^{th} instance.

$$\hat{C}^{(0)}_{(W,\mu)}(m) = \left(C^{(0)}_{(W,\mu)}(1) - \frac{b}{a}\right) \times e^{-a\times(k-1)} \times (1 - e^{a}) \tag{14}$$

From these expressions, it can be concluded that the internal optimization of tunable parameters can have a potential impact on the forecast accuracy. Hence, these parameters should be tuned properly. The following subsection presents a discussion on the need of this optimization.

2.2. Discussion

After considering the facts involved in the development of a forecasting engine it appears that there is a potential impact of internal parameters on the accuracy of the forecast. In references [33,34], it has been pointed out that near accurate estimation of the background values can be expressed as per Eq. (5). The relationship between background value production coefficient α and development coefficient can be defined as follows:

$$\alpha = \frac{1}{a} - \frac{1}{(e^{a} - 1)} \tag{15}$$

Further, Chang et al. [35] proved that proper optimization of background value production coefficient can enhance the forecasting accuracy. A detailed explanation regarding this can be found in Ref. [36].

By using the L-Hopital rule as applied in [15] it can be concluded that for diverse values of ‘a’ the parameter α revolves around 0.5 value. Moreover, it can be said that higher values of ‘a’ can lead to erroneous results. As ‘a’ approaches to zero, α approaches to 0.5. Fig. 1 exhibits this relationship where 16 samples of ‘a’ are considered between span $[-1,1]$ and α is calculated as per expression (15). As indicated in different researches, larger values of ‘a’ yield erroneous forecast because of greater difference between α and a. Hence, in a defined search (objective) space the error function is employed to bridge this difference through optimization.

2.3. Proposed internal optimization model

Based on the discussion in previous subsection, an optimization routine is formulated for predicting COVID-19 infected cases. Following are the steps involved in constructing the model.

Step 1: Start the iterative search by taking $\alpha(0) = 0.5$, while taking $\alpha = 0.5$ the model becomes conventional grey model. Calculate the background values as per Eq. (5) and further calculate values of a and b.

Step 2: Now by substituting the values of ‘a’, in expression (15) new value of α that is designated as $\alpha(1)$ is obtained. Calculate the absolute error between obtained $\alpha(1)$ and initial value i.e. (0.5), if the value of error is greater than tolerance then stop the loop otherwise perturb the value of α. The absolute error can be defined for two successive iterations by $\Delta(t)$, where ‘t’ denotes current iteration.

$$\Delta(t) = |\alpha(t) - \alpha(t-1)| \tag{16}$$

Now, update α as

$$\widetilde{\alpha} = \alpha(t) + \Delta(t) \tag{17}$$

Where $\widetilde{\alpha}$ is perturbed value, again calculate the values of a and b from the expression (5)–(7) and compute the absolute error between $\widetilde{\alpha}$ and $\alpha(t)$.

Step 3: If the error reduces then increase the loop counter by 1 and accept the perturbation vector $\Delta(t)$ and append $\alpha(t+1)$ as $\widetilde{\alpha}$ as in same direction. Otherwise reject the perturbation value and assign opposite perturbation $(-\Delta(t))$. Now, update the alpha as

$$\widetilde{\alpha} = \alpha(t) - \Delta(t) \tag{18}$$

and $\alpha(t+1) = \widetilde{\alpha}$. Repeat the process, till the error between the successive iterations of the α becomes less than tolerance value.

Step 4: Now the optimized model can be realized with the modified α as represented by Eq. (14). For simulating the time series and prediction, ϵ is taken 10E-8 in this work.

Algorithm 1: Pseudo code of Proposed IOGM

Initialize α and construct the original GM(1,1) i.e. by setting $\alpha = 0.5$
Calculate the background values, a and b as per Eq. (5).
Calculate updated α by substituting the values of a, in Eq. (15).
Calculate the absolute error between $\alpha(1)$ and initial value i.e. ($\alpha(0)$=0.5)
while $\Delta(t) < \epsilon$ **do**
    Perturb α by $\Delta(t)$, where $\Delta(t) = \alpha(t) - \alpha(t-1)$.
    Perturb $\alpha(t)$ to obtain $\alpha(t+1)$ and apply condition for choosing perturbation value by calculating absolute error between perturbed α and previous value.
    $\alpha(t+1) = \begin{cases} \alpha(t) + \Delta(t) & \text{if } \Delta(t+1) < \Delta(t) \\ \alpha(t) - \Delta(t) & \text{Otherwise} \end{cases}$
    Check feasibility of new positions and evaluate these positions
end while
Print the values of a, b and α, and construct IOGM for forecast.

In this work, an internal optimization scheme is employed for forecasting COVID-19 cases. The flow of algorithm along with data stacking process are shown in Figs. 2 and 3 respectively. Following steps are considered for framing GPMs for forecasting the pandemic:

- For staking data into Model-I and Model-II, two different overlap periods (μ=5 and 6) are considered. The data of three different states and Delhi are depicted in the result Section 4.
- Further, the data stacked in model array are segregated into two parts simulation and validation (forecasted) parts.
- On the basis of tolerance value ’ϵ’ obtained from simulated data, grey models have been constructed and coefficients ‘a’ and ‘b’ are calculated.

However, it is quite necessary to judge the forecasting performance of the IOGM in comparison to classical GM model and Novel Grey Model (NGM) on benchmark time series. Following section depicts this analysis in depth.

3. Verification of proposed optimization model with benchmark time series data

For understanding the impact of internal optimization, let us consider a homogeneous Geometric Progression data series. This series is defined as follows:

$$A = (1, 2, 4, 8, 16, 32, 64, 128) \tag{19}$$

While forecasting the next value of series from GM(1,1) model, α is set to 0.5. From this value of α, undermentioned series is

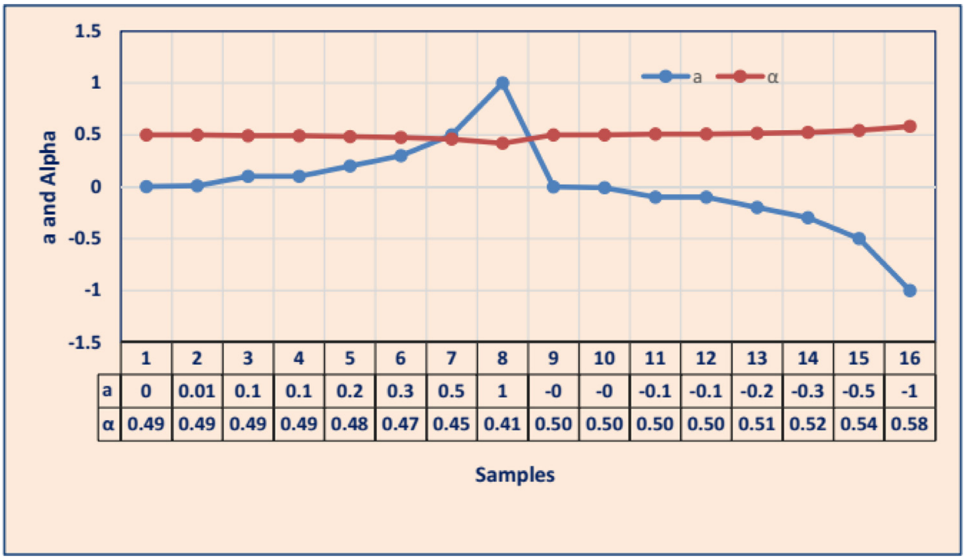

Fig. 1. Relationship between a and α.

obtained.

$$A_{\hat{GM(1,1)}} = (0.97, 1.89, 3.69, 7.19, 14.00, 27.27, 53.13, 103.48, \mathbf{201.56}) \quad (20)$$

For this experiment obtained value of a is -0.667. The forecasted value of the series is written in bold face. From this forecasted value (201.56), it can be observed that the error is very high. A huge difference exists between actual value of the series (256) and the forecasted value of GM(1,1) model. Further, using optimized model, following series is obtained.

$$A_{\hat{IOGM(1,1)}} = (0.99, 1.99, 3.99, 7.99, 15.99, 31.99, 63.99, 127.99, \mathbf{255.9963}) \quad (21)$$

For this model, value of α is 0.5573 and value of a is -0.6931. It is observed that forecasted value is 255.9963. This value is quite close to the actual value as compared to non-optimized model. This fact validates the necessity of internal optimization for improving the forecasting accuracy of the conventional GM(1,1). Further, for verification of the internal optimized model of grey forecasting, certain benchmarks time series are used here from [37] and same are defined as under:

1. **Homogeneous Exponential Sequence (B1):** The series can be identified with the help of following formula:

$$X_{B1} : x_{B1}^{(0)}(i) = 0.8 \times 1.5^i, i = 1, 2, 3.....15 \quad (22)$$

2. **Non-homogeneous Exponential Sequence (B2):** The series can be identified with the help of following formula:

$$X_{B2} : x_{B2}^{(0)}(i) = 1.3 \times 1.8^i + 3.5, i = 1, 2, 3.....15 \quad (23)$$

3. **Approximate Non-homogeneous Exponential Sequence (B3):** The series can be identified with the help of following formula:

$$X_{B3} : x_{B3}^{(0)}(i) = 1.3 \times 1.4^i + 1.6, i = 1, 2, 3.....15 \quad (24)$$

4. **Random Number Sequence (B4):** The series can be identified as per following sequence:

$$X_{B4} : x_{B4}^{(0)} = [78.3571, 35.0894, 40.8045, 48.9120, 58.0352, 66.4530, 77.8831, 68.2308, 73.0020, 79.3541, 79.4816, 105.474, 95.2562, 119.923, 133.4256] \quad (25)$$

For evaluation of the performance of the proposed IOGM, simulations are carried out on B1–B4. For a better understanding of the computed forecasting accuracy, the whole process is subdivided into two parts. In the first part, 10 out of 15 samples are considered for building the grey architecture and for calculation of internal parameters of the forecaster. In the second part, remaining five samples of each series are employed as Out-Samples for evaluating the performance of the forecaster on the basis of error and Mean Absolute Percentage Error (MAPE).

Both of these indices are defined as under:

$$\delta(m) = \frac{\left|\hat{x}^{(0)}(m) - x^{(0)}(m)\right|}{x^{(0)}(m)} \times 100 \quad (26)$$

$$MAPE = \frac{1}{n}\sum_{n}^{k=1}\left|\frac{\hat{x}^{(0)}(m) - x^{(0)}(m)}{x^{(0)}(m)}\right| \times 100 \quad (27)$$

Table 1 shows the results of B1. Similar to previously reported results on Geometric Progression Series, it is observed that the IOGM model is better as far as simulated value error analysis is concerned. It is further observed that NGM is not suitable for this kind of time series. Grey models are developed with current series having 10 data points. Values of 'a 'and 'b' are provided along with the name of models in respective rows. Further, as per the analysis conducted on forecasted values, MAPE of models have been calculated and it is observed that the IOGM model gives the best results as MAPE values are optimal for this model.

Further, Table 2 shows results of B2. High errors are there for the non-homogeneous exponential model. The coefficients and internal parameters have been calculated based on In-Samples and the rest five Out-Samples are simulated with the grey equations of associated models. It is observed that MAPE for simulated values is optimal for IOGM. MAPE for forecasted values is also very competitive. However, NGM model possesses the optimal value of MAPE for forecasted values.

The B3 benchmark consists of an approximate non-homogeneous model, the simulated results are shown in Table 3 . From the Table, it can be observed that MAPE value is optimal for IOGM (7.63). However, it is worth mentioning here that other competitive models NGM and GM possess high simulation errors for this sequence. This fact indicates that the IOGM model

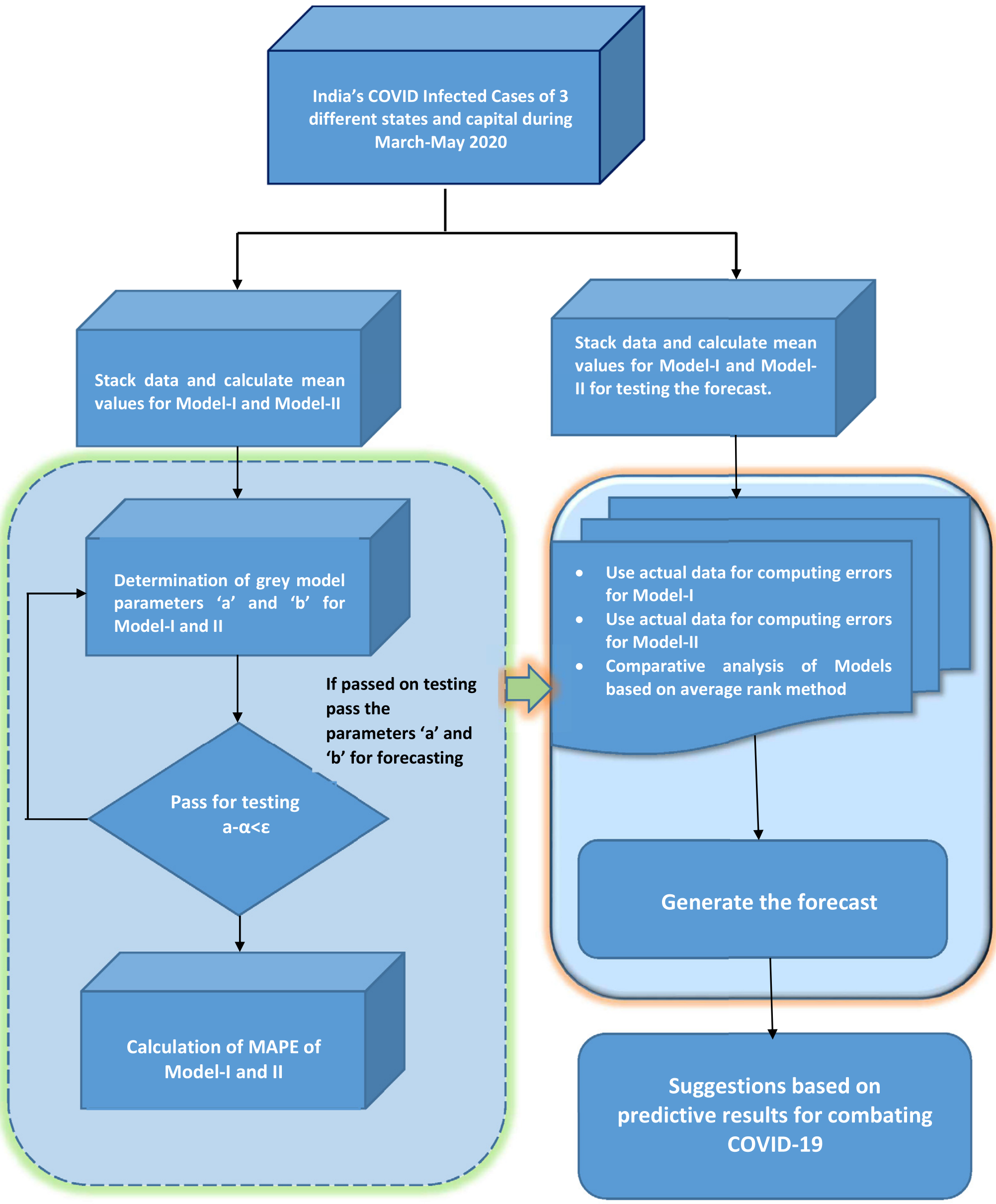

Fig. 2. Flow Diagram of Proposed Internal Optimization Grey Models.

suits well for this kind of time series. Further, inspecting the forecasted results of Out-Samples, it can be concluded that the IOGM model possesses least MAPE. On the other hand, the NGM model possesses highest simulated and forecasted errors.

To test the applicability of the IOGM model further, a random sequence time series is considered. Comparative analysis of the performance of different grey models is depicted through errors in simulated and forecasted values of the series in Table 4. In-Samples are employed for building the grey prediction models, by observing the individual error component and accumulated MAPE for the simulation model. It is observed that the NGM model is not suitable for forecasting the random samples. Further, the GM model also possesses higher MAPE as compared to the IOGM model. By comparing the MAPEs of these models it can be concluded that the IOGM model possesses optimal MAPE.

3.1. Discussion

By comparing the results, on all benchmarks, on the basis of MAPE and arranging ranks on the basis of performance, one can calculate the average rank obtained from models. It can be concluded that IOGM secured 1.25 average rank as the performance of IOGM is better than other models on three out of four benchmarks. NGM performance is very weak on these benchmarks as the average rank possessed by this model is 2.5. Except, non-homogeneous model, performance of NGM is comparatively weak. Further, the average rank obtained by the GM model is 2.25. The same analysis is depicted in Fig. 4, where (a) segment shows the average rank and the rank obtained on the basis of MAPE of simulated and forecasted data of benchmark time series. Segment (b) shows the graphical representation of rank-based comparison of forecasters for simulated data. Likewise, segment (c) and (d) show rank-based analysis of forecasted data and MAPEs of forecasting engines respectively. From this analysis, it can be concluded that for all types of time series, the performance of IOGM is satisfactory. Following points support the argument for choosing IOGM for COVID-19 forecasting:

1. It is observed that the overall performance of IOGM is competitive with conventional GM and NGM models. This

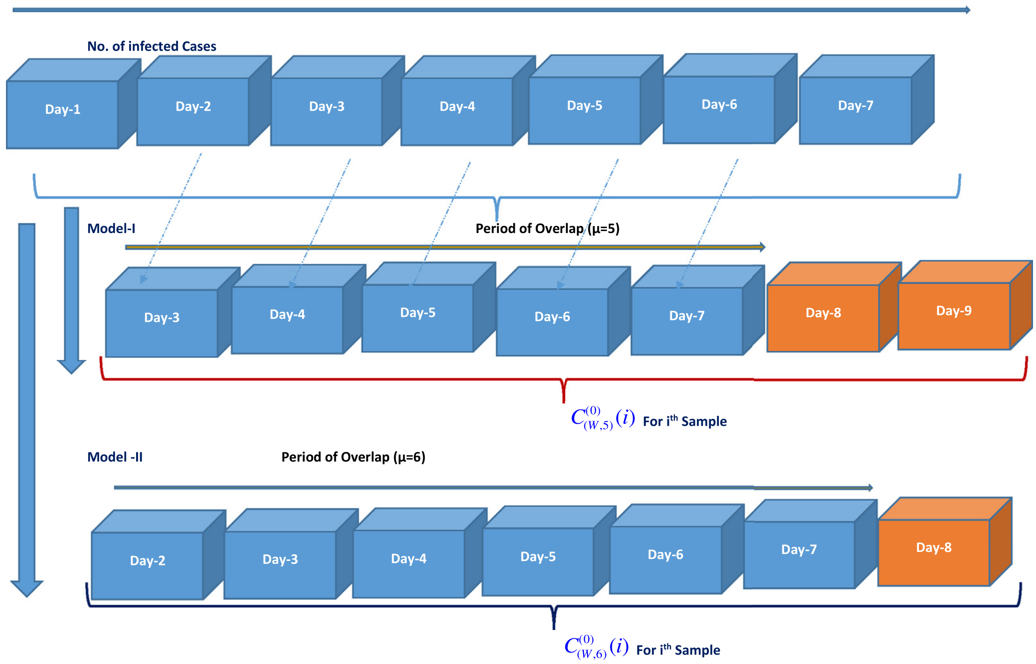

Fig. 3. Proposed Grey Forecasting Models based on Internal Optimization.

Table 1
Simulated and Forecasted results of different Grey Models on B1.

Simulated results							
Sample	In-sample	GM(1,1) [14]	Error	NGM(1,1,k) [23]	Error	IOGM	Error
		a = −0.400, b = 0.9600		a = −0.3867, b = 0.2266		a = −0.4055, b = 0.9731	
		Simulated value		Simulated value		Simulated value	
$X^{in}_{B1}(1)$	1.2	1.2	0	1.2	0	1.2	0
$X^{in}_{B1}(2)$	1.8	1.770568912	1.63506	0.972690805	45.96162	1.8	0
$X^{in}_{B1}(3)$	2.7	2.641378431	2.171169	1.70861309	36.71803	2.7	0
$X^{in}_{B1}(4)$	4.05	3.940473579	2.704356	2.791985048	31.0621	4.05	0
$X^{in}_{B1}(5)$	6.075	5.878495806	3.234637	4.386847474	27.78852	6.075	0
$X^{in}_{B1}(6)$	9.1125	8.769685228	3.762028	6.734689447	26.09394	9.1125	0
$X^{in}_{B1}(7)$	13.66875	13.08283301	4.286544	10.19101386	25.44297	13.66875	0
$X^{in}_{B1}(8)$	20.503125	19.51729341	4.808202	15.27916653	25.47884	20.503125	0
$X^{in}_{B1}(9)$	30.7546875	29.11638033	5.327016	22.76957965	25.96387	30.7546875	0
$X^{in}_{B1}(10)$	46.13203125	43.43653529	5.843003	33.79642812	26.73978	46.13203125	0
MAPE		3.752446129		30.13885312		**0**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$X^{out}_{B1}(1)$	69.19804688	64.7998	6.356027	50.02936286	27.70119	69.19720902	0.001211
$X^{out}_{B1}(2)$	103.7970703	96.6699	6.866447	73.92632408	28.77802	103.7957088	0.001312
$X^{out}_{B1}(3)$	155.6956055	144.2145	7.374072	109.1057149	29.9237	155.6934061	0.001413
$X^{out}_{B1}(4)$	233.5434082	215.1427	7.878924	160.8942887	31.10733	233.5398735	0.001514
$X^{out}_{B1}(5)$	350.3151123	320.9551	8.381029	237.1337093	32.30846	350.3094568	0.001614
MAPE		7.37		29.96374116		**0.001412603**	

is on the basis of average rank obtained in simulated data (In-Samples) and forecasted data (Out-Samples).

2. Application of IOGM on previously reported approaches, motivated the author to employ this prediction theory for forecasting the pandemic growth in terms of reported infected cases. As it is indicated in the results of the benchmark data series that IOGM is compatible for all types of data and it can give fruitful results.
3. Further, it is apparent from the results that the average rank method is a suitable criterion for evaluating the performance of different prediction models.

Based on the results on known time-series data, the following section presents an application of conventional GM, NGM and proposed IOGMs for forecasting the spread of pandemic at different locations in India.

Table 2
Simulated and Forecasted results of different Grey Models on B2.

Simulated results							
Sample	In-sample	GM(1,1) [14]	Error	NGM(1,1,k) [23]	Error	IOGM	Error
		$a = -0.5544$, $b = -2.1190$		$a = -0.5622$, $b = -0.7212$		$a = -0.5693$, $b = -2.1749$	
		Simulated value		Simulated value		Simulated value	
$X_{B2}^{in}(1)$	5.84	5.84	0	5.84	0	5.84	0
$X_{B2}^{in}(2)$	7.712	1.495361505	80.60994	2.999639095	61.10426	1.549460351	79.90845
$X_{B2}^{in}(3)$	11.0816	2.603375408	76.50722	4.295022134	61.24186	2.738029543	75.29211
$X_{B2}^{in}(4)$	17.1469	4.532391325	73.56728	6.567789434	61.69693	4.838333405	71.78304
$X_{B2}^{in}(5)$	28.0644	7.890744863	71.88344	10.55539082	62.38868	8.549750751	69.53524
$X_{B2}^{in}(6)$	47.7159	13.73752839	71.20975	17.55169212	63.21626	15.10814402	68.3373
$X_{B2}^{in}(7)$	83.0886	23.91658702	71.21556	29.82679865	64.10242	26.69738831	67.86877
$X_{B2}^{in}(8)$	146.7595	41.63799475	71.62842	51.36364135	65.00149	47.1765785	67.8545
$X_{B2}^{in}(9)$	261.3671	72.49038524	72.26492	89.1503271	65.89076	83.36506679	68.10422
$X_{B2}^{in}(10)$	467.66	126.2033867	73.01386	155.4475853	66.76056	147.3132343	0
MAPE		73.54448706		63.48924641		**63.18707065**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$X_{B2}^{out}(1)$	838.989331	219.7159631	73.81183	271.7670336	67.60781	260.3151395	68.97277
$X_{B2}^{out}(2)$	1507.380796	382.5182961	74.62365	475.8511003	68.43192	459.9992131	69.48354
$X_{B2}^{out}(3)$	2710.485433	665.9518262	75.43053	833.9193931	69.23358	812.8581246	70.01061
$X_{B2}^{out}(4)$	4876.073779	1159.40032	76.22267	1462.155121	70.01368	1436.390133	70.54208
$X_{B2}^{out}(5)$	8774.132801	2018.477987	76.99513	2564.403319	70.77314	2538.224753	71.0715
MAPE		75.4167604		**69.21202647**		**70.01609993**	

Table 3
Simulated and Forecasted results of different Grey Models on B3.

Simulated results							
Sample	In-sample	GM(1,1) [14]	Error	NGM(1,1,k) [23]	Error	IOGM	Error
		$a = -0.3026$, $b = 1.7313$		$a = -0.2810$, $b = 0.4217$		$a = -0.3049$, $b = 1.7449$	
		Simulated value		Simulated value		Simulated value	
$X_{B3}^{in}(1)$	3.32	3.32	0	3.32	0	3.32	0
$X_{B3}^{in}(2)$	4.248	3.194908005	24.7903	1.795999681	57.72129	3.223590794	24.11509
XB3in(3)	5.1672	4.323835724	16.32149	2.865544851	44.54357	4.372860104	15.37273
$X_{B3}^{in}(4)$	5.6041	5.851672516	4.417703	4.282038509	23.59097	5.931865026	5.848665
$X_{B3}^{in}(5)$	8.6217	7.919373775	8.146029	6.158026863	28.57526	8.04668383	6.669406
$X_{B3}^{in}(6)$	11.4084	10.71770179	6.054295	8.642564853	24.24385	10.91547437	4.320725
$X_{B3}^{in}(7)$	15.4138	14.50482511	5.89715	11.93305917	22.58198	14.80704142	3.936463
$X_{B3}^{in}(8)$	20.88583	19.63013673	6.012178	16.29095304	21.99997	20.08602357	3.829421
$X_{B3}^{in}(9)$	28.5594	26.5664884	6.978128	22.06249893	22.74873	27.2470598	4.595125
$X_{B3}^{in}(10)$	39.3031	35.95381507	8.521681	29.70626973	24.41749	36.96113693	0
MAPE		9.682106842		30.04701125		**7.631959339**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$X_{B3}^{out}(1)$	55.24434	48.65817412	11.92188	39.82959428	27.90285	50.14124257	9.237322
$X_{B3}^{out}(2)$	75.90208	65.85164617	13.24132	53.23681206	29.86119	68.01751625	10.38781
$X_{B3}^{out}(3)$	106.7829	89.12046911	16.5405	70.9931813	33.51634	92.26700975	13.59383
$X_{B3}^{out}(4)$	146.3561	120.6113814	17.59047	94.50951808	35.42495	125.161893	14.48126
$X_{B3}^{out}(5)$	203.8385	163.2296764	19.92206	125.6542914	38.35596	169.784406	16.70641
MAPE		15.84324534		33.01225931		**12.88132694**	

4. Simulation results

The proposed internal optimization based model is implemented to predict the infected cases in different states (Gujarat, Rajasthan and Maharashtra) and union territory Delhi. The data for this study has been taken from [38],[39]. For better understanding, it is to be noted here that variable $C_{(W,6)}^{(0)}(1)$ indicates the mean values of infected cases of COVID-19 for the duration of the first week of April to the Second week of May 2020. Time series is constructed by excluding two values for Model-1 and excluding only one entry for Model-2. For considering the uncertainty and unavailability of the data in a few cases, the mean of available data is considered for the forecast. A few points may be noted here:

Table 4
Simulated and Forecasted results of different Grey Models on B4.

Simulated results							
Sample	In-sample	GM(1,1) [14]	Error	NGM(1,1,k) [23]	Error	IOGM	Error
		a = −0.0854, b = 33.9309		a = 0.3757, b = 29.7504		a = −0.0854, b = 33.9549	
		Simulated value		Simulated value		Simulated value	
$X_{B4}^{in}(1)$	78.3571	78.3571	0	78.3571	0	78.3571	0
$X_{B4}^{in}(2)$	35.0894	42.40830908	20.85789	13.43485632	61.71249	42.43763141	20.94146
$X_{B4}^{in}(3)$	40.8045	46.18944512	13.19694	34.02779312	16.60774	46.22328777	13.27988
$X_{B4}^{in}(4)$	48.912	50.30770825	2.853509	48.17121952	1.514517	50.34664428	2.933113
$X_{B4}^{in}(5)$	58.0352	54.79315681	5.586339	57.88506036	0.258704	54.83782553	5.509371
$X_{B4}^{in}(6)$	66.453	59.67852915	10.19438	64.55661937	2.853717	59.72964339	10.11746
$X_{B4}^{in}(7)$	77.8831	64.99948257	16.54225	69.13871006	11.22758	65.05783672	16.46733
$X_{B4}^{in}(8)$	68.2308	70.79485361	3.757912	72.28573387	5.942967	70.86133247	3.855345
$X_{B4}^{in}(9)$	73.002	77.10694146	5.623053	74.44713999	1.97959	77.18253008	5.726597
$X_{B4}^{in}(10)$	79.3541	83.98181674	5.83173	75.93161443	4.312928	84.06761124	0
MAPE		9.382667169		11.82336045		**8.758949572**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$X_{B4}^{out}(1)$	79.4816	91.4696577	15.08281	76.95116571	3.183673	91.56687727	15.20513
$X_{B4}^{out}(2)$	105.7574	99.62511655	5.798444	77.6514033	26.57591	99.73511664	5.694432
$X_{B4}^{out}(3)$	95.2562	108.5077183	13.91145	78.13233319	17.97664	108.6320052	14.04193
$X_{B4}^{out}(4)$	119.923	118.1822951	1.451519	78.46264045	34.57248	118.3225423	1.334571
$X_{B4}^{out}(5)$	133.4256	128.7194598	3.527164	78.68949865	41.02369	128.8775255	3.408697
MAPE		7.954277077		24.66648042		**7.93695043**	

Fig. 4. Comparative analysis of Forecasting engines on the basis of rank.

- Forecasting of such time series is quite difficult due to the unavailability of data. Also, at the initial stage, the value of the variable is quite small and it takes large abrupt changes at a later stage. This change sometimes is quite higher than the previous value. Also, it is to be noted here that the forecasting of the day ahead cases are merely meaningful as this short time forecast will give very less time to authorities for taking any preventive action. On the basis of this fact, the work reported in this paper addresses two representative models that can forecast weekly mean infected cases.
- Forecast, on the basis of mean values of the infected cases, can be helpful for authorities to cater to the needs of the patients and to think on needed health care and infrastructure. The advantage associated with these models is that they can give predictions of infected cases for an upcoming week.

4.1. Results of Model-I

On the basis of the discussion and steps presented in Section 2.3, two models are constructed. These are named as Model-I and Model-II. As shown in Fig. 3, the period of overlap is 5 days, which indicates that new time series element $C^{(0)}_{(W,\mu)}(m)$ consists of 5 same values and last two are replaced by new values of infected case. The results of Model-I are shown in Tables 5–8 for Rajasthan, Maharashtra, Delhi and Gujarat respectively.

4.1.1. Forecasted results of Rajasthan

Forecasted results of model-I for the state of Rajasthan are depicted in Table 5. The data of reported infected cases have been segregated into two parts. First, 11 samples are taken as simulated data and the remaining five samples are considered for validation of grey models and for generating forecasts. The data of 6th April 2020–12th April 2020 is considered as $\hat{C}^{(0)}_{(W,5)}(1)$ and the data of mean infected cases during 26th April 2020 to 5th May 2020 is denoted as $\hat{C}^{(0)}_{(W,5)}(11)$. Simulation process with defined time series with 11 data points provides the coefficient 'a' and 'b' for representative models of GM, NGM and IOGM. These values are depicted with corresponding grey models.

From the obtained results, it can be concluded that proposed model gives competitive performance as per Lewis' criterion for model evaluation [30]. This criterion has been used for evaluating the performance of the grey models on the basis of calculated MAPE. Simulated and forecasted results of these grey models are depicted in Fig. 5. It is observed that the NGM model gives better results as the increment in the forecasted values are at a comparatively low exponential rate. Further, errors in the forecasting and simulation process have been encapsulated in Fig. 6.

4.1.2. Forecasted results of Maharashtra

Table 6 shows the comparative analysis of different grey forecasters with the proposed IOGM. The analysis is being depicted through the calculation of MAPE for simulated data and forecasted data. Data of the infected cases for the state of Maharashtra are employed for the construction of IOGM and other forecasters. The first 11 samples are taken for constructing grey architectures and the internal parameters (a, b) of these, are depicted with respective grey forecasters. For simplification, it is to be noted that the data sample from 6th April 2020 to 12th April 2020 is considered as $\hat{C}^{(0)}_{(W,5)}(1)$ while $\hat{C}^{(0)}_{(W,5)}(11)$ represents the data of mean infected cases during 26th April 2020 to 5th May 2020. After careful evaluation of the results, it can be concluded that for this particular data, IOGM's performance indicator i.e. MAPE is competitive (4.134597266). However, it can be concluded from this result that further improvement in the performance of IOGM may be possible, if the overlap period is varied. It is observed that the NGM model gives better results in this case, as the forecasted values increase at a comparatively lower exponential rate. Graphical representation of simulated and forecasted results along with errors in forecasting and simulation process of this model for Maharashtra state have been encapsulated in Figs. 5 and 6 respectively.

4.1.3. Forecasted results of Delhi

Results of grey models along with proposed IOGM (model-I) for Delhi are depicted in Table 7. Data of infected cases have been segregated into two parts, the first 10 samples are taken as simulated data and the remaining 6 samples have been considered for validation and forecasting purposes. The data from 7th April 2020 to 13th April 2020 are depicted in Table 7 as $\hat{C}^{(0)}_{(W,5)}(1)$ and the data of mean infected cases during 25th April 2020 to 1st May 2020 are denoted as $\hat{C}^{(0)}_{(W,5)}(10)$. Coefficients of constructed grey models are depicted along with the models. In addition to that error in the prediction of each sample is also shown in this analysis. Careful observation of Table 7 yields the fact that the proposed model exhibits competitive performance as values of MAPEs are competitive (3.278769158) and (2.160679097) for simulated and forecasted data. Depiction of forecasting performance and errors is exhibited in Figs. 5 and 6. It is observed that the IOGM model gives better results in this case, as the forecasted values increase with a comparatively higher exponential rate.

4.1.4. Forecasted results of Gujarat

Forecasted results of model-I for Gujarat state are depicted in Table 8. Similar to the results reported in the previous subsection for different states, the data of infected cases have been subdivided into two parts. First 5 samples are taken as simulated data and the remaining 5 samples are considered for validation and forecasting purpose. The time series of this model can be identified as $\hat{C}^{(0)}_{(W,5)}(1)$ - $\hat{C}^{(0)}_{(W,5)}(5)$. The mean value of infected cases from 18th April 2020 to 24th April 2020 is considered as the first element of the time series and the mean value of infected cases from 21st April 2020 to 27th April 2020 is considered as the last element of the time series. Likewise, the data employed for simulation, yield the coefficients 'a' and 'b' for representative models of GM, NGM and IOGM respectively. The entries of these parameters are also depicted. Inspecting the obtained results, it is easily predictable that the proposed model gives a competitive performance as values of MAPEs are quite competitive. Proposed IOGM exhibits superior performance with simulated MAPE (0.4525292188769158) and forecasted MAPE (4.831822784). Simulated and forecasted results of these models are depicted in Fig. 5. It is observed that the IOGM model gives competitive results in this case as the forecasted values increase with a comparatively higher exponential rate. Errors in the forecasting and simulation process of this model for the state of Gujarat have been exhibited in Fig. 6.

4.2. Results of Model-II

The results of proposed Model-II are presented in this section. The difference between this model and the first model is that it employs an extended overlap period. The compilation of forecasted results is presented in Tables 9–12 for Delhi, Maharashtra, Rajasthan and Gujarat respectively.

4.2.1. Forecasted results of Delhi

Table 9 depicts results of forecasters in terms of In-Samples and Out-Samples for three grey forecasting models as described in the previous section also. Internal parameters of grey models ('a' and 'b') have also been depicted along with the errors. For constructing the time series, data from 6th April 2020 to 12th April 2020 is considered as $\hat{C}^{(0)}_{(W,6)}(1)$ and $\hat{C}^{(0)}_{(W,6)}(27)$ represents the data of mean infected cases from 2nd May 2020 to 8th May 2020. Time series is divided into two parts. First 27 samples are considered for constructing the grey model and for obtaining parameters ('a', 'b'). The remaining, 5 samples are employed to evaluate the performance of constructed model.

As observed from the Table 9, IOGM outperforms other opponents, when MAPE of Out-Sample is considered. It is observed that values of MAPEs are quite competitive for the proposed IOGM. The MAPE values are quite competitive for simulated (4.441512436) and forecasted data (2.316478228). Pictorial representation of all these models is depicted in Fig. 7. It is also worth mentioning here that IOGM model gives competitive results in this case, as the forecasted value increases with every sample at a higher exponential rate. In addition to that, the graphical representation of simulation and forecasting errors for Delhi Model-II is presented in Fig. 8.

Table 5
Simulated and Forecasted results for Rajasthan on Model-I.

Simulated results							
Sample	In-sample	GM(1,1) [14] a = −0.1411, b = 596.2662 Simulated value	Error	NGM(1,1,k) [23] a = 0.0154, b = 227.0934 Simulated value	Error	IOGM a = −0.14135, b = 596.3051 Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(1)$	355.5714286	355.5714286	0	355.5714286	0	355.5714286	0
$\hat{C}^{(0)}_{(W,5)}(2)$	503.8571429	694.262049	37.78946	332.8709539	33.93545	695.3219623	37.99982
$\hat{C}^{(0)}_{(W,5)}(3)$	685	799.465716	16.71032	553.1245434	19.25189	800.8675726	16.91497
$\hat{C}^{(0)}_{(W,5)}(4)$	871.8571429	920.6112188	5.59198	770.0057058	11.68212	922.4343593	5.80109
$\hat{C}^{(0)}_{(W,5)}(5)$	1061.428571	1060.114273	0.123824	983.5660783	7.335632	1062.454239	0.096631
$\hat{C}^{(0)}_{(W,5)}(6)$	1256.142857	1220.756654	2.817052	1193.856507	4.95854	1223.728277	2.580485
$\hat{C}^{(0)}_{(W,5)}(7)$	1493.714286	1405.741671	5.889521	1400.927061	6.211846	1409.482726	5.639068
$\hat{C}^{(0)}_{(W,5)}(8)$	1727.714286	1618.758037	6.306381	1604.82704	7.112706	1623.433561	6.035762
$\hat{C}^{(0)}_{(W,5)}(9)$	1933.285714	1864.053429	3.581068	1805.604991	6.604338	1869.860821	3.280679
$\hat{C}^{(0)}_{(W,5)}(10)$	2111.714286	2146.519187	1.648182	2003.308718	5.133534	2153.694228	1.987956
$\hat{C}^{(0)}_{(W,5)}(11)$	2278.571429	2471.787853	8.479718	2197.985291	3.536696	2480.611805	8.866976
MAPE		8.085228394		9.614796169		**8.109403844**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(12)$	2467.285714	2846.34548	15.36343	2389.681062	3.145345	2853.655642	15.65972
$\hat{C}^{(0)}_{(W,5)}(13)$	2681.571429	3277.66098	22.22911	2578.44167	3.84587	3286.82307	22.57078
$\hat{C}^{(0)}_{(W,5)}(14)$	2920.571429	3774.335045	29.23276	2764.312058	5.350301	3785.742658	29.62335
$\hat{C}^{(0)}_{(W,5)}(15)$	3171.428571	4346.271662	37.0446	2947.336479	7.065967	4360.395179	37.48994
$\hat{C}^{(0)}_{(W,5)}(16)$	3437.714286	5004.875596	45.58731	3127.558511	9.022151	5022.27643	46.09348
MAPE		29.89144242		5.685926932		**30.28745404**	

Table 6
Simulated and Forecasted results for Maharashtra on Model-I.

Simulated results							
Sample	In-sample	GM(1,1) [14] a = −0.1884, b = 1316.9084 Simulated value	Error	NGM(1,1,k) [23] a = −0.0955, b = 431.4991 Simulated value	Error	IOGM a = −0.1890, b = 1320.8668 Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(1)$	1028.142857	1028.142857	0	1028.142857	0	1028.142857	0
$\hat{C}^{(0)}_{(W,5)}(2)$	1386.428571	1662.275573	19.89623	778.5358378	43.84595	1667.741783	20.29049
$\hat{C}^{(0)}_{(W,5)}(3)$	1834.714286	2006.884178	9.384017	1309.287013	28.6381	2014.606608	9.804923
$\hat{C}^{(0)}_{(W,5)}(4)$	2352.571429	2422.93406	2.990882	1893.202632	19.52624	2433.614022	3.444852
$\hat{C}^{(0)}_{(W,5)}(5)$	2872.428571	2925.235808	1.838418	2535.608085	11.72598	2939.768579	2.344358
$\hat{C}^{(0)}_{(W,5)}(6)$	3522.428571	3531.670413	0.262371	3242.3622	7.950945	3551.195556	0.81668
$\hat{C}^{(0)}_{(W,5)}(7)$	4274.857143	4263.825798	0.258052	4019.910673	5.963859	4289.790008	0.349318
$\hat{C}^{(0)}_{(W,5)}(8)$	5234.714286	5147.765311	1.661007	4875.344855	6.86512	5182.000828	1.006998
$\hat{C}^{(0)}_{(W,5)}(9)$	6355	6214.955523	2.20369	5816.466424	8.474171	6259.777876	1.498381
$\hat{C}^{(0)}_{(W,5)}(10)$	7500.428571	7503.386386	0.039435	6851.858541	8.647106	7561.716093	0.81712
$\hat{C}^{(0)}_{(W,5)}(11)$	8690.571429	9058.923599	4.238526	7990.964126	8.050188	9134.437579	5.107445
MAPE		3.888420598		13.60796909		**4.134597266**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(12)$	10027.28571	10936.94134	9.071803	9244.17198	7.809828	11034.26112	10.04235
$\hat{C}^{(0)}_{(W,5)}(13)$	11578.28571	13204.29349	14.0436	10622.91153	8.25143	13329.21895	15.12256
$\hat{C}^{(0)}_{(W,5)}(14)$	13442.57143	15941.6935	18.5911	12139.75709	9.691705	16101.49297	19.77986
$\hat{C}^{(0)}_{(W,5)}(15)$	15590.14286	19246.58761	23.45357	13808.54248	11.42774	19450.35765	24.76061
$\hat{C}^{(0)}_{(W,5)}(16)$	18037.14286	23236.6238	28.82652	15644.48727	13.26516	23495.73505	30.26306
MAPE		18.79731828		10.08917147		**19.99368955**	

Table 7
Simulated and Forecasted results for Delhi on Model-I.

Simulated results							
Sample	In-sample	GM(1,1) [14] a = −0.1199, b = 965.4484 Simulated value	Error	NGM(1,1,k) [23] a = 0.1151, b = 461.3211 Simulated value	Error	IOGM a = −0.1201, b = 966.620 Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(1)$	744.8571429	744.8571429	0	744.8571429	0	744.8571429	0
$\hat{C}^{(0)}_{(W,5)}(2)$	968.4285714	1120.648553	15.71825	576.8271099	40.43679	1122.089215	15.86701
$\hat{C}^{(0)}_{(W,5)}(3)$	1239	1263.412352	1.970327	949.8658806	23.33609	1265.216358	2.115929
$\hat{C}^{(0)}_{(W,5)}(4)$	1459.857143	1424.363388	2.431317	1282.342684	12.15971	1426.599962	2.278112
$\hat{C}^{(0)}_{(W,5)}(5)$	1698.857143	1605.818606	5.476537	1578.667982	7.074707	1608.568716	5.314657
$\hat{C}^{(0)}_{(W,5)}(6)$	1865.428571	1810.390114	2.950446	1842.77267	1.214515	1813.748341	2.770421
$\hat{C}^{(0)}_{(W,5)}(7)$	2066.285714	2041.022785	1.222625	2078.16022	0.574679	2045.09948	1.025329
$\hat{C}^{(0)}_{(W,5)}(8)$	2286.142857	2301.036653	0.651481	2287.953159	0.079186	2305.960418	0.866856
$\hat{C}^{(0)}_{(W,5)}(9)$	2563.571429	2594.174704	1.193775	2474.93449	3.457557	2600.095253	1.424724
$\hat{C}^{(0)}_{(W,5)}(10)$	2899.142857	2924.656757	0.88005	2641.584609	8.883945	2931.748209	1.124655
MAPE		3.249480357		9.721718244		**3.278769158**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(11)$	3236.714286	3297.24021	1.86998	2790.114207	13.79795	3305.704877	2.131501
$\hat{C}^{(0)}_{(W,5)}(12)$	3683.571429	3717.288526	0.915337	2922.4936	20.66141	3727.361273	1.188788
$\hat{C}^{(0)}_{(W,5)}(13)$	4195	4190.848437	0.098965	3040.478863	27.52136	4202.801694	0.185976
$\hat{C}^{(0)}_{(W,5)}(14)$	4846.142857	4724.736996	2.505206	3145.635126	35.08992	4738.886517	2.213231
$\hat{C}^{(0)}_{(W,5)}(15)$	5560.428571	5326.639705	4.204512	3239.357336	41.74267	5343.351187	3.903969
$\hat{C}^{(0)}_{(W,5)}(16)$	6233.142857	6005.221152	3.65661	3322.888763	46.69	6024.917838	3.34061
MAPE		2.2084349		30.91721722		**2.160679097**	

Table 8
Simulated and Forecasted results for Gujarat on Model-I.

Simulated results							
Sample	In-sample	GM(1,1) [14] a = −0.1538, b = 1806.6475 Simulated value	Error	NGM(1,1,k) [23] a = 0.4067, b = 1634.4552 Simulated value	Error	IOGM a = −0.1540, b = 1810.2249 Simulated value	Error
$\hat{C}^{(0)}_{(W,5)}(1)$	1784.714286	1784.714286	0	1784.714286	0	1784.714286	0
$\hat{C}^{(0)}_{(W,5)}(2)$	2234.142857	2249.788215	0.700285	1463.378541	34.49933	2254.19028	0.897321
$\hat{C}^{(0)}_{(W,5)}(3)$	2650.857143	2623.902535	1.016826	2317.289441	12.58339	2629.492419	0.805955
$\hat{C}^{(0)}_{(W,5)}(4)$	3077.142857	3060.227833	0.549699	2885.858555	6.216296	3067.278943	0.320554
$\hat{C}^{(0)}_{(W,5)}(5)$	3569.428571	3569.109091	0.00895	3264.435335	8.544596	3577.952935	0.238816
MAPE		0.455152029		12.36872234		**0.452529218**	
Forecasted results							
	Out-sample	Simulated Value	Error	Simulated Value	Error	Simulated Value	Error
$\hat{C}^{(0)}_{(W,5)}(6)$	4125.142857	4162.611543	0.9083	3516.507377	14.75429	4173.649492	1.175878
$\hat{C}^{(0)}_{(W,5)}(7)$	4751.285714	4854.806737	2.1788	3684.347344	22.45578	4868.524097	2.467509
$\hat{C}^{(0)}_{(W,5)}(8)$	5467.571429	5662.106157	3.557973	3796.102121	30.5706	5679.088991	3.868583
$\hat{C}^{(0)}_{(W,5)}(9)$	6224.428571	6603.650336	6.092475	3870.513063	37.81738	6624.605553	6.429136
$\hat{C}^{(0)}_{(W,5)}(10)$	7011.142857	7701.762657	9.850317	3920.058939	44.08816	7727.542006	10.21801
MAPE		4.517573198		29.93724099		**4.831822784**	

4.2.2. *Forecasted results of Maharashtra*

Comparative analysis of the forecasting results for the state of Maharashtra is presented in Table 10. For constructing the time series for grey models like previous case studies, 27 such indicators have been considered for simulation and the remaining 5 samples are considered as Out-Samples for evaluating the constructed grey models. The internal parameters of grey forecasting systems such as ('a' and 'b') are shown in the respective columns. The data from 5th April 2020 to 11th April 2020 is considered as $\hat{C}^{(0)}_{(W,6)}(1)$ and $\hat{C}^{(0)}_{(W,6)}(27)$ represents the data of mean infected cases from 2nd May 2020 to 8th May 2020.

A careful inspection of obtained MAPEs for simulated data In-Samples and Out-Samples indicate that MAPEs are quite competitive for proposed IOGM for simulated results (9.881179166)

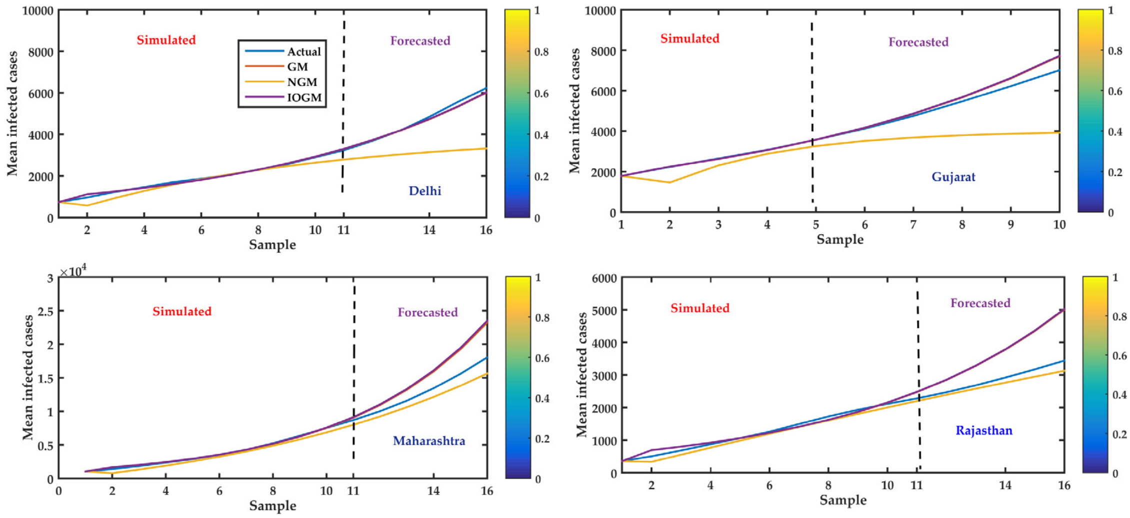

Fig. 5. Forecasted Results of Proposed Grey Model-I for different states and Delhi.

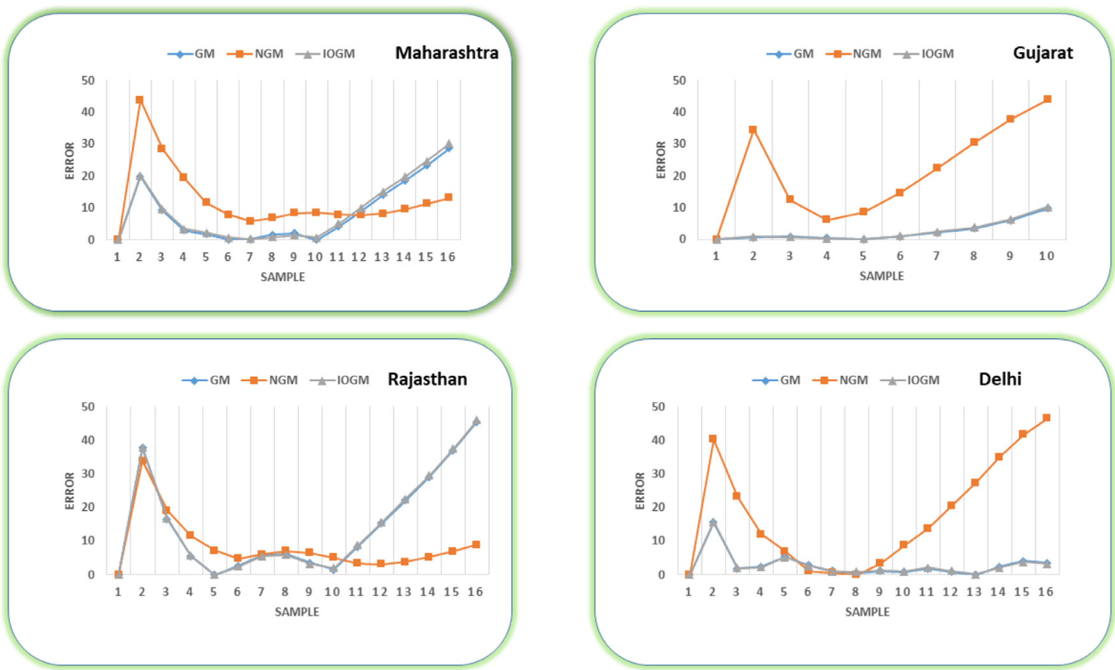

Fig. 6. Error of Proposed Grey Model-I for different states and Delhi.

and competitive for forecasted results (12.36321992). Graphical representation of the forecasting performance along with errors for each model are depicted in Figs. 7 and 8 respectively. It is concluded that Model-II provides quite competitive results with the proposed IOGM.

4.2.3. Forecasted results of Rajasthan

Forecasting results of Rajasthan state for all three grey models are depicted in Table 11 and graphical representation of the forecasting results is presented in Fig. 7. For construction of the time series, from 6th April 2020 to 12th April 2020 is considered as $\hat{C}^{(0)}_{(W,6)}(1)$ and data of mean infected cases from 3rd May 2020 to 9th May 2020 is considered as $\hat{C}^{(0)}_{(W,6)}(27)$. Inspecting the forecasting results from Table 11 and Fig. 7, it is observed that for this particular case NGM model provides better results. Similar to previously reported results all internal parameters of the grey system are shown in Table 11.

Inspecting the values of MAPEs for IOGM, it has been observed that the performance of the IOGM is acceptable and competitive as per Lewis's criterion. It is observed that values of MAPEs are (12.63328453) for proposed IOGM simulated data (27 data points) and (16.00238203) for the remaining 5 Out-Samples. Simulated and forecasted results of these models are depicted in Fig. 7. It is observed that the IOGM model gives competitive results in this case as the increment in infected cases and forecasted values increase with a higher exponential rate. Further, the error analysis for Rajasthan (Model-II) is depicted in Fig. 8.

4.2.4. Forecasted results of Gujarat

Table 12 shows the comparative analysis of different grey models along with the proposed IOGM (Model-II). The analysis is depicted through calculated error for each sample and the same is shown along with the sample. All parameters obtained for an internal grey mechanism such as ‘a’ and ‘b’ are shown in Table 12.

Table 9
Simulated and Forecasted results for Delhi on Model-II.

Simulated results							
Sample	In-sample	GM(1,1) [14] a = −0.0611, b = 904.781 Simulated value	Error	NGM(1,1,k) [23] a = 0.0031, b = 159.4292 Simulated value	Error	IOGM a = −0.0611, b = 904.9924 Simulated value	Error
$\hat{C}^{(0)}_{(W,6)}(1)$	664	664	0	664	0	664	0
$\hat{C}^{(0)}_{(W,6)}(2)$	744.8571429	974.7735719	30.86718	236.7697971	68.21272	975.0890441	30.90954
$\hat{C}^{(0)}_{(W,6)}(3)$	835	1036.206602	24.0966	395.2235387	52.66784	1036.561559	24.13911
$\hat{C}^{(0)}_{(W,6)}(4)$	968.4285714	1101.511318	13.74213	553.1889791	42.87767	1101.909484	13.78325
$\hat{C}^{(0)}_{(W,6)}(5)$	1109.142857	1170.931724	5.570866	710.667623	35.92641	1171.377137	5.611025
$\hat{C}^{(0)}_{(W,6)}(6)$	1239	1244.727204	0.462244	867.6609707	29.97087	1245.224238	0.50236
$\hat{C}^{(0)}_{(W,6)}(7)$	1345	1323.173487	1.622789	1024.170518	23.85349	1323.726879	1.581645
$\hat{C}^{(0)}_{(W,6)}(8)$	1459.857143	1406.56368	3.650594	1180.197754	19.15663	1407.17856	3.608475
$\hat{C}^{(0)}_{(W,6)}(9)$	1577.571429	1495.209363	5.220814	1335.744168	15.32908	1495.891284	5.177588
$\hat{C}^{(0)}_{(W,6)}(10)$	1698.857143	1589.441751	6.440529	1490.811239	12.24623	1590.196722	6.396089
$\hat{C}^{(0)}_{(W,6)}(11)$	1780.428571	1689.612935	5.100774	1645.400446	7.584024	1690.447455	5.053902
$\hat{C}^{(0)}_{(W,6)}(12)$	1865.428571	1796.097194	3.716646	1799.513261	3.533521	1797.018293	3.667269
$\hat{C}^{(0)}_{(W,6)}(13)$	1961.142857	1909.292397	2.64389	1953.151152	0.407502	1910.307674	2.59212
$\hat{C}^{(0)}_{(W,6)}(14)$	2066.285714	2029.621487	1.774403	2106.315583	1.937286	2030.739154	1.720312
$\hat{C}^{(0)}_{(W,6)}(15)$	2181.571429	2157.534062	1.101837	2259.008012	3.549578	2158.762994	1.045505
$\hat{C}^{(0)}_{(W,6)}(16)$	2286.142857	2293.508055	0.322167	2411.229895	5.471532	2294.857837	0.381209
$\hat{C}^{(0)}_{(W,6)}(17)$	2416.857143	2438.051519	0.87694	2562.98268	6.046097	2439.532503	0.938217
$\hat{C}^{(0)}_{(W,6)}(18)$	2563.571429	2591.704527	1.097418	2714.267815	5.878377	2593.327892	1.160742
$\hat{C}^{(0)}_{(W,6)}(19)$	2729	2755.041189	0.954239	2865.086739	4.986689	2756.818999	1.019384
$\hat{C}^{(0)}_{(W,6)}(20)$	2899.142857	2928.671797	1.01854	3015.440891	4.011463	2930.617072	1.085639
$\hat{C}^{(0)}_{(W,6)}(21)$	3061.857143	3113.245104	1.678327	3165.331701	3.37947	3115.371893	1.747787
$\hat{C}^{(0)}_{(W,6)}(22)$	3236.714286	3309.450751	2.247232	3314.760598	2.411282	3311.774208	2.319016
$\hat{C}^{(0)}_{(W,6)}(23)$	3450.571429	3518.021842	1.954761	3463.729006	0.381316	3520.558309	2.028269
$\hat{C}^{(0)}_{(W,6)}(24)$	3683.571429	3739.737681	1.524777	3612.238342	1.93652	3742.504781	1.599897
$\hat{C}^{(0)}_{(W,6)}(25)$	3939.285714	3975.42669	0.91745	3760.290023	4.543862	3978.443419	0.994031
$\hat{C}^{(0)}_{(W,6)}(26)$	4195	4225.969497	0.738248	3907.885459	6.844208	4229.256331	0.816599
$\hat{C}^{(0)}_{(W,6)}(27)$	4494	4492.302231	0.037779	4055.026055	9.768001	4495.881235	0.041861
MAPE		4.421451124		13.81154333		4.441512436	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,6)}(28)$	4846.142857	4775.420018	1.459363	4201.713212	13.29778	4779.314967	1.378991
$\hat{C}^{(0)}_{(W,6)}(29)$	5214.714286	5076.380702	2.652755	4347.94833	16.62154	5080.617204	2.571514
$\hat{C}^{(0)}_{(W,6)}(30)$	5560.428571	5396.308792	2.951567	4493.7328	19.1837	5400.914431	2.868738
$\hat{C}^{(0)}_{(W,6)}(31)$	5899.571429	5736.39967	2.765824	4639.068011	21.36602	5741.404148	2.680996
$\hat{C}^{(0)}_{(W,6)}(32)$	6233.142857	6097.924051	2.169352	4783.955347	23.24971	6103.359351	2.082152
MAPE		2.399772252		18.74374981		2.316478228	

For constructing the time series for the state of Gujarat infected cases, 15 samples are taken for constructing the model and the remaining five samples are kept for testing the efficacy of the proposed IOGM.

The data from 18th April 2020 to 24th April 2020 represent $\hat{C}^{(0)}_{(W,6)}(1)$, and $\hat{C}^{(0)}_{(W,6)}(15)$ represent the data of mean infected cases from 2nd May 2020 to 8th May 2020. After assessment of results, it can be concluded that for this particular data IOGM's performance is better than competitors as the MAPE values obtained for simulated data is optimal (**0.889304425**) and competitive for forecasted results (4.338970394). It is also worth mentioning here that the rise in infected cases in this particular state is swift comparatively, hence the NGM model provides pessimistic results. The same fact can be observed from higher MAPEs obtained from the NGM model (11.25642155) for simulated data. The MAPE is very high (20.02449253) for forecasted data. Graphical representation of the forecasted results is presented in Fig. 7. Further, the analysis depicting simulation and forecasting errors for Gujarat Model-II is shown in Fig. 8. From this analysis, it can be concluded that proposed IOGM performs satisfactorily.

Table 10
Simulated and Forecasted results for Maharashtra on Model-II.

Simulated results							
Sample	In-sample	GM(1,1) [14] $a = -0.0845$, $b = 1604.26$ Simulated value	Error	NGM(1,1,k) [23] $a = -0.0468$, $b = 236.57$ Simulated value	Error	IOGM $a = -0.0845$, $b = 1604.921$ Simulated value	Error
$\hat{C}^{(0)}_{(W,6)}(1)$	1028.142857	1028.142857	0	1028.142857	0	1028.142857	0
$\hat{C}^{(0)}_{(W,6)}(2)$	1209.714286	1764.746184	45.88124	411.5522969	65.97938	1765.352263	45.93134
$\hat{C}^{(0)}_{(W,6)}(3)$	1386.428571	1920.443046	38.51727	673.440986	51.4262	1921.025603	38.55929
$\hat{C}^{(0)}_{(W,6)}(4)$	1596.285714	2089.876453	30.9212	947.8644603	40.62063	2090.426622	30.95567
$\hat{C}^{(0)}_{(W,6)}(5)$	1834.714286	2274.258326	23.95708	1235.422673	32.66403	2274.765863	23.98475
$\hat{C}^{(0)}_{(W,6)}(6)$	2089.571429	2474.907513	18.44091	1536.744291	26.45648	2475.360615	18.4626
$\hat{C}^{(0)}_{(W,6)}(7)$	2352.571429	2693.259216	14.48151	1852.488075	21.25688	2693.644333	14.49788
$\hat{C}^{(0)}_{(W,6)}(8)$	2602.428571	2930.875262	12.62078	2183.344311	16.10358	2931.176875	12.63237
$\hat{C}^{(0)}_{(W,6)}(9)$	2872.428571	3189.45527	11.03689	2530.036328	11.91996	3189.655652	11.04386
$\hat{C}^{(0)}_{(W,6)}(10)$	3189.285714	3470.848812	8.828406	2893.322075	9.279935	3470.927759	8.830882
$\hat{C}^{(0)}_{(W,6)}(11)$	3522.428571	3777.068639	7.229105	3273.995776	7.052884	3777.003169	7.227247
$\hat{C}^{(0)}_{(W,6)}(12)$	3884.428571	4110.30508	5.814922	3672.889674	5.445818	4110.069102	5.808847
$\hat{C}^{(0)}_{(W,6)}(13)$	4274.857143	4472.941708	4.633712	4090.87584	4.3038	4472.505653	4.623511
$\hat{C}^{(0)}_{(W,6)}(14)$	4735.571429	4867.572391	2.787435	4528.86809	4.364908	4866.902799	2.773295
$\hat{C}^{(0)}_{(W,6)}(15)$	5234.714286	5297.019841	1.190238	4987.823974	4.716405	5296.078908	1.172263
$\hat{C}^{(0)}_{(W,6)}(16)$	5802.857143	5764.355812	0.663489	5468.746875	5.757686	5763.100879	0.685115
$\hat{C}^{(0)}_{(W,6)}(17)$	6355	6272.923063	1.291533	5972.688203	6.015921	6271.306059	1.316978
$\hat{C}^{(0)}_{(W,6)}(18)$	6915.142857	6826.359274	1.283901	6500.749687	5.992547	6824.326089	1.313303
$\hat{C}^{(0)}_{(W,6)}(19)$	7500.428571	7428.623062	0.957352	7054.085792	5.950897	7426.112859	0.99082
$\hat{C}^{(0)}_{(W,6)}(20)$	8109.428571	8084.022301	0.313293	7633.906238	5.86382	8080.966747	0.350972
$\hat{C}^{(0)}_{(W,6)}(21)$	8690.571429	8797.244928	1.227462	8241.478646	5.167586	8793.567349	1.185146
$\hat{C}^{(0)}_{(W,6)}(22)$	9360.428571	9573.392483	2.275151	8878.131307	5.152513	9569.00692	2.228299
$\hat{C}^{(0)}_{(W,6)}(23)$	10027.28571	10418.01659	3.896676	9545.256092	4.807179	10412.82676	3.844919
$\hat{C}^{(0)}_{(W,6)}(24)$	10728.14286	11337.15869	5.676806	10244.31149	4.509927	11331.05683	5.619929
$\hat{C}^{(0)}_{(W,6)}(25)$	11578.28571	12337.3932	6.556303	10976.82579	5.194723	12330.2588	6.494684
$\hat{C}^{(0)}_{(W,6)}(26)$	12465	13425.87461	7.708581	11744.40045	5.780983	13417.573	7.641982
$\hat{C}^{(0)}_{(W,6)}(27)$	13442.57143	14610.3886	8.687454	12548.71355	6.649456	14600.76937	8.615896
MAPE		9.884396406		13.64570872		**9.881179166**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,6)}(28)$	14510.57143	15899.40778	9.571204	13391.52351	7.711949	15888.30308	9.494675
$\hat{C}^{(0)}_{(W,6)}(29)$	15590.14286	17302.15223	10.98136	14274.6729	8.437831	17289.37484	10.8994
$\hat{C}^{(0)}_{(W,6)}(30)$	16723.28571	18828.65549	12.58945	15200.0925	9.108217	18813.99675	12.5018
$\hat{C}^{(0)}_{(W,6)}(31)$	18037.14286	20489.83636	13.59802	16169.80548	10.35273	20473.06377	13.50503
$\hat{C}^{(0)}_{(W,6)}(32)$	19302.85714	22297.57691	15.51439	17185.93186	10.9669	22278.43162	15.4152
MAPE		12.45088273		9.31552661		**12.36321992**	

4.3. Comparative analysis of the proposed IOGM models

To showcase the efficacy of the proposed approach, the analysis based on errors reported in simulated and forecasted data have already been discussed in the previous section. On the basis of MAPE values of forecasting models, it can be concluded that proposed models based on different overlap periods and mean infected cases in the duration of a span of 6–7 days can be a potential tool for alignment of medical facilities and policy decisions. Further, to have a clear insight, comparative analysis of these proposed grey models are depicted in Fig. 9. The following points can be concluded from this analysis:

1. Fig. 9(a) and (c) show the MAPE of forecasted results of IOGM models. It can be easily concluded that the proposed IOGM (Model-II) provides competitive results as compared to the results obtained by conventional GM and NGM models.
2. Fig. 9(b) depicts the average rank analysis. For conducting this analysis, developed models have given rank as per the performance. The evaluation of performance is based on the calculated MAPE for simulated and forecasted samples. After taking the mean of the MAPE obtained from forecasted and simulated models, it has been observed that the average rank of IOGM (Model-II) is I **(1.5)** as compared to

Table 11
Simulated and Forecasted results for Rajasthan on Model-II.

Sample	In-sample	GM(1,1) [14]	Error	NGM(1,1,k) [23]	Error	IOGM	Error
Simulated results							
		a = −0.0586, b = 700.4246		a = 0.0160, b = 129.94		a = −0.0585, b = 700.6304	
		Simulated value		Simulated value		Simulated value	
$\hat{C}^{(0)}_{(W,6)}(1)$	355.5714286	355.5714286	0	355.5714286	0	355.5714286	0
$\hat{C}^{(0)}_{(W,6)}(2)$	427	742.8266132	73.96408	187.9169088	55.99136	742.9783216	73.99961
$\hat{C}^{(0)}_{(W,6)}(3)$	503.8571429	787.6736058	56.32876	313.8585409	37.70882	787.7683486	56.34756
$\hat{C}^{(0)}_{(W,6)}(4)$	588.2857143	835.2281654	41.97662	437.8058071	25.57939	835.258517	41.98178
$\hat{C}^{(0)}_{(W,6)}(5)$	685	885.6537569	29.29252	559.7902895	18.27879	885.6116033	29.28637
$\hat{C}^{(0)}_{(W,6)}(6)$	776.4285714	939.1237145	20.9543	679.84307	12.43971	939.0001969	20.93839
$\hat{C}^{(0)}_{(W,6)}(7)$	871.8571429	995.821837	14.21846	797.9947383	8.471847	995.6072915	14.19386
$\hat{C}^{(0)}_{(W,6)}(8)$	968.4285714	1055.94302	9.036748	914.2753998	5.59186	1055.626913	9.004107
$\hat{C}^{(0)}_{(W,6)}(9)$	1061.428571	1119.693925	5.489333	1028.714683	3.082062	1119.264783	5.448903
$\hat{C}^{(0)}_{(W,6)}(10)$	1156.571429	1187.293691	2.656322	1141.341747	1.316796	1186.739026	2.608364
$\hat{C}^{(0)}_{(W,6)}(11)$	1256.142857	1258.974687	0.225439	1252.18529	0.315057	1258.280916	0.170208
$\hat{C}^{(0)}_{(W,6)}(12)$	1369.857143	1334.98331	2.545801	1361.273556	0.626605	1334.135667	2.607679
$\hat{C}^{(0)}_{(W,6)}(13)$	1493.714286	1415.580836	5.230816	1468.634339	1.679032	1414.56328	5.298939
$\hat{C}^{(0)}_{(W,6)}(14)$	1612.714286	1501.04431	6.92435	1574.294995	2.382275	1499.839426	6.999061
$\hat{C}^{(0)}_{(W,6)}(15)$	1727.714286	1591.667508	7.874379	1678.282448	2.861112	1590.256395	7.956055
$\hat{C}^{(0)}_{(W,6)}(16)$	1832.285714	1687.761939	7.887622	1780.623194	2.819567	1686.124101	7.97701
$\hat{C}^{(0)}_{(W,6)}(17)$	1933.285714	1789.657921	7.429207	1881.343308	2.686742	1787.771137	7.526801
$\hat{C}^{(0)}_{(W,6)}(18)$	2031.285714	1897.705714	6.576131	1980.468456	2.501729	1895.545906	6.682458
$\hat{C}^{(0)}_{(W,6)}(19)$	2111.714286	2012.276722	4.708855	2078.023894	1.595405	2009.817817	4.825296
$\hat{C}^{(0)}_{(W,6)}(20)$	2190	2133.764776	2.567818	2174.034479	0.729019	2130.978545	2.695044
$\hat{C}^{(0)}_{(W,6)}(21)$	2278.571429	2262.58748	0.70149	2268.524676	0.440923	2259.443379	0.839476
$\hat{C}^{(0)}_{(W,6)}(22)$	2368.857143	2399.187653	1.280386	2361.51856	0.309794	2395.652642	1.131157
$\hat{C}^{(0)}_{(W,6)}(23)$	2467.285714	2544.034848	3.110671	2453.039827	0.577391	2540.073204	2.950104
$\hat{C}^{(0)}_{(W,6)}(24)$	2567.428571	2697.626965	5.071159	2543.111796	0.947126	2693.200077	4.898734
$\hat{C}^{(0)}_{(W,6)}(25)$	2681.571429	2860.491965	6.672227	2631.757419	1.857642	2855.558118	6.488236
$\hat{C}^{(0)}_{(W,6)}(26)$	2795	3033.189685	8.521992	2718.999281	2.719167	3027.70382	8.325718
$\hat{C}^{(0)}_{(W,6)}(27)$	2920.571429	3216.313758	10.12618	2804.859614	3.961958	3210.227229	9.917778
MAPE		12.64339478		7.313747646		**12.63328453**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,6)}(28)$	3041	3410.493661	12.1504	2889.360293	4.986508	3403.753958	11.92877
$\hat{C}^{(0)}_{(W,6)}(29)$	3171.428571	3616.396871	14.03053	2972.52285	6.271802	3608.947336	13.79564
$\hat{C}^{(0)}_{(W,6)}(30)$	3305.142857	3834.731165	16.02316	3054.368475	7.587399	3826.510681	15.77444
$\hat{C}^{(0)}_{(W,6)}(31)$	3437.714286	4066.247049	18.28345	3134.918023	8.808069	4057.189709	18.01998
$\hat{C}^{(0)}_{(W,6)}(32)$	3570.142857	4311.740342	20.77221	3214.192017	9.970213	4301.775091	20.49308
MAPE		16.25194971		7.524798279		**16.00238203**	

other grey models i.e. **2 and 2.5**. However, the results for Model-I are quite comparable with the original GM model. Hence, it is to be noted that for Model-II proposed IOGM based methodology provides very competitive results. This method is suitable for forecasting as it produces meaningful results without knowing the pattern of variables. It can generate a reliable forecast for planning combating strategies.

3. Fig. 9(d) depicts the average MAPE analysis obtained by IOGM models (state-wise). As shown, the average MAPE obtained by models I and II are (11.76 and 7.82), (16.29, 11.37), (13.09, 9.54) and (21.95, 13.25) for Delhi, Maharashtra, Gujarat and Rajasthan respectively. It can be concluded that the proposed IOGM model-II exhibits superior performance as the average MAPE calculated for forecasting is optimal.

4.4. Discussion

From the results reported in this section, it can be concluded that the estimated results are always higher than the actual infected cases. This indicator is sufficient enough to spark an

Table 12
Simulated and Forecasted results for Gujarat on Model-II.

Sample	In-sample	GM(1,1) [14] a = −0.0738, b = 1904.2076 Simulated value	Error	NGM(1,1,k) [23] a = −0.1212, b = 730.2219 Simulated value	Error	IOGM a = −0.0738, b = 1905.07252622308 Simulated value	Error
Simulated results							
$\hat{C}^{(0)}_{(W,6)}(1)$	1784.714286	1784.714286	0	1784.714286	0	1784.714286	0
$\hat{C}^{(0)}_{(W,6)}(2)$	2013.714286	2112.805827	4.920834	834.7845948	58.54503	2113.801078	4.970258
$\hat{C}^{(0)}_{(W,6)}(3)$	2234.142857	2274.533853	1.807897	1427.202548	36.11856	2275.681355	1.859259
$\hat{C}^{(0)}_{(W,6)}(4)$	2443.714286	2448.641604	0.201632	1951.998846	20.12164	2449.958835	0.255535
$\hat{C}^{(0)}_{(W,6)}(5)$	2650.857143	2636.076705	0.557572	2416.892175	8.826012	2637.582929	0.500752
$\hat{C}^{(0)}_{(W,6)}(6)$	2862.571429	2837.85932	0.863284	2828.72017	1.182547	2839.575755	0.803322
$\hat{C}^{(0)}_{(W,6)}(7)$	3077.142857	3055.0877	0.716741	3193.539984	3.782636	3057.037707	0.653371
$\hat{C}^{(0)}_{(W,6)}(8)$	3316.428571	3288.944166	0.828735	3516.717373	6.039292	3291.153449	0.762119
$\hat{C}^{(0)}_{(W,6)}(9)$	3569.428571	3540.701541	0.804808	3803.00562	6.543822	3543.198372	0.734857
$\hat{C}^{(0)}_{(W,6)}(10)$	3841.714286	3811.730078	0.78049	4056.615444	5.593887	3814.545537	0.707204
$\hat{C}^{(0)}_{(W,6)}(11)$	4125.142857	4103.504918	0.524538	4281.276927	3.784937	4106.67316	0.447735
$\hat{C}^{(0)}_{(W,6)}(12)$	4429	4417.61412	0.257076	4480.294385	1.158148	4421.172662	0.176729
$\hat{C}^{(0)}_{(W,6)}(13)$	4751.285714	4755.767302	0.094324	4656.594959	1.99295	4759.757337	0.178302
$\hat{C}^{(0)}_{(W,6)}(14)$	5104.285714	5119.804948	0.304043	4812.771671	5.711162	5124.27169	0.391553
$\hat{C}^{(0)}_{(W,6)}(15)$	5467.571429	5511.708425	0.80725	4951.121564	9.44569	5516.701481	0.898572
MAPE		0.897948273		11.25642155		**0.889304425**	
Forecasted results							
	Out-sample	Simulated value	Error	Simulated value	Error	Simulated value	Error
$\hat{C}^{(0)}_{(W,6)}(16)$	5841.428571	5933.610767	1.578076	5073.679484	13.14317	5936.501843	1.627569
$\hat{C}^{(0)}_{(W,6)}(17)$	6224.428571	6387.808283	2.624815	5182.248007	16.74339	6391.134291	2.678249
$\hat{C}^{(0)}_{(W,6)}(18)$	6616	6876.773058	3.941552	5278.423957	20.21729	6880.58365	3.999148
$\hat{C}^{(0)}_{(W,6)}(19)$	7011.142857	7403.166407	5.591436	5363.621885	23.49861	7407.51629	5.653478
$\hat{C}^{(0)}_{(W,6)}(20)$	7402.142857	7969.853358	7.669543	5439.09488	26.52	7974.802774	7.736407
MAPE		4.281084185		20.02449253		**4.338970394**	

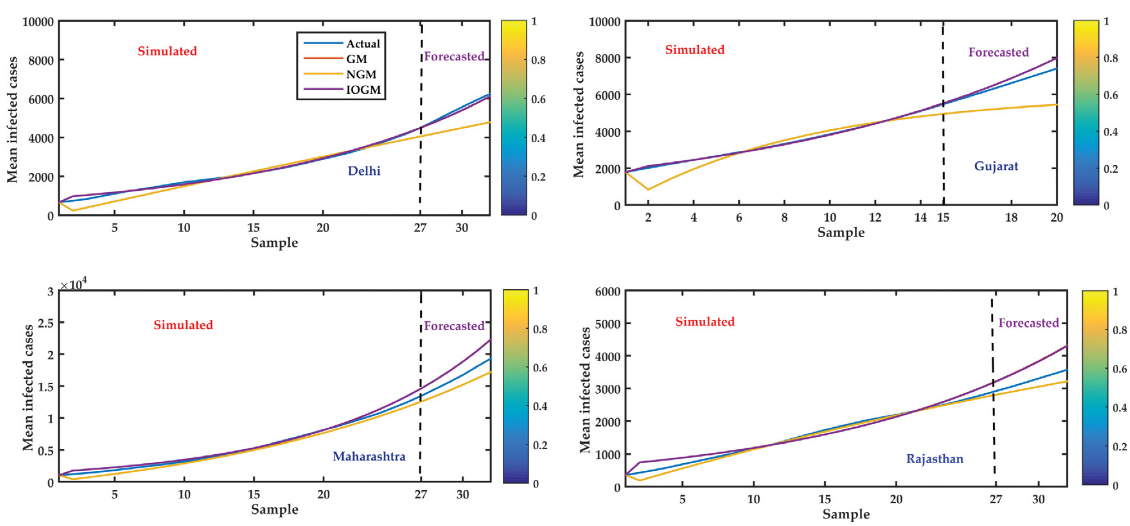

Fig. 7. Forecasted Results of Proposed Grey Model-II for different states and Delhi.

alarm to the authorities. However, the authenticity of the forecast largely depends upon the removal of potential uncertainties in the data. Another problem with the forecasting of epidemic and pandemic is that the data of confirmed cases multi-folds with time. Apart from these issues, forecasting is immensely valuable as it allows us to foresee many preventive and corrective measures in health care. Model-I and II give a sufficient amount of accuracy in the prediction of mean weekly infected cases. Higher values of MAPE can be justified with the larger population in three states and Delhi. Following recommendations can be drawn from this forecast:

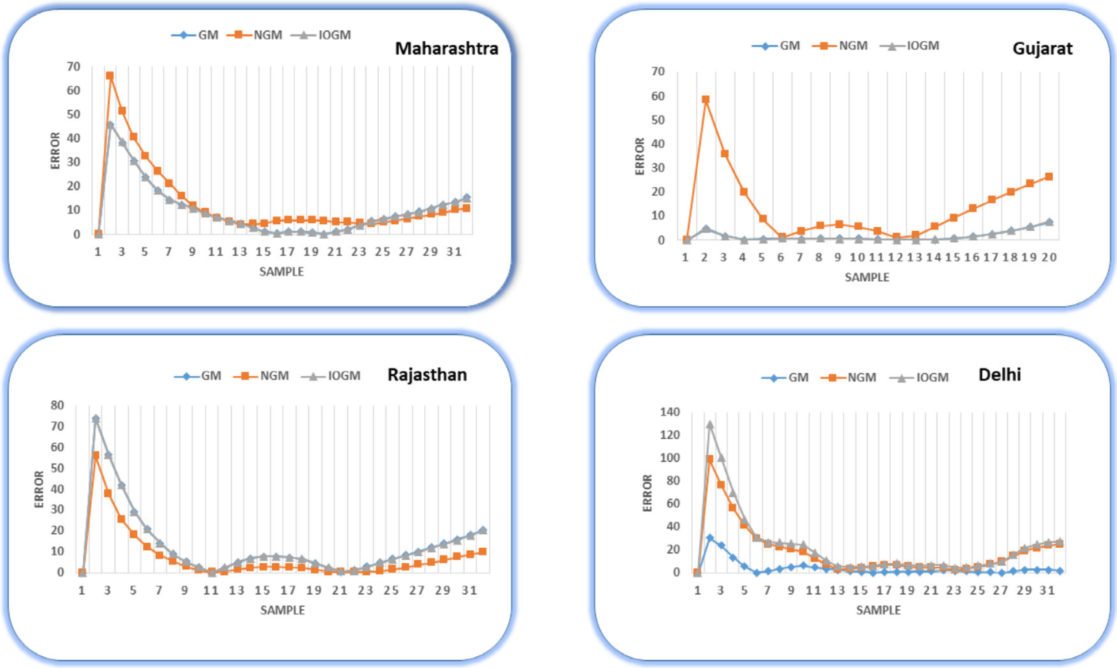

Fig. 8. Error of Proposed Grey Model-II for different states and Delhi.

	GM(1,1)	NGM(1,1,K)	IOGM(1,1)
Delhi	2.2084349	30.91721722	2.160679097
Maharashtra	18.79731828	10.08917147	19.99368955
Gujarat	4.517573198	29.93724099	4.831822784
Rajasthan	29.89144242	5.685926932	30.28745404

a. MAPE of Forecasted Results by Model-I

Model	Case	Place	GM	NGM	IOGM
Model-1	Simulation	Delhi	2	3	1
		Maharashtra	1	3	2
		Gujarat	2	3	1
		Rajasthan	1	3	2
	Case	Place	GM	NGM	IOGM
	Forecasted	Delhi	2	3	1
		Maharashtra	2	1	3
		Gujarat	1	3	2
		Rajasthan	3	1	2
	Average Rank		1.75	2.5	1.75
Model-II	Simulation	Delhi	1	3	2
		Maharashtra	2	3	1
		Gujarat	2	3	1
		Rajasthan	2	3	1
	Case	Place	GM	NGM	IOGM
	Forecasted	Delhi	2	3	1
		Maharashtra	3	1	2
		Gujarat	1	3	2
		Rajasthan	3	1	2
	Average Rank		2	2.5	1.5

b. Rank Based Analysis

	GM(1,1)	NGM(1,1,K)	IOGM(1,1)
Delhi	2.399772252	18.74374981	2.316478228
Maharashtra	12.45088273	9.31552661	12.36321992
Gujarat	4.281084185	20.02449253	4.338970394
Rajasthan	16.25194971	7.524798279	16.00238203

c. MAPE of Forecasted Results (Model-II)

d. Average MAPE for forecasting model-I and II for different states

Fig. 9. Comparative Analysis of proposed grey models on the basis of average ranks..

1. It can be seen that the performance of the models relies on the mean infected cases of a duration of more than five days. It is also a known fact that by taking the average of infected cases, the forecaster can easily deal with the randomness in data. This randomness is due to environment, policies, strategic decisions, sentiments and medical conditions.
2. Considering a large population and the density of population in the major states of India, it may be concluded that based on the predictions of pandemic spread in these states, authorities can take decisions on the availability of Intensive Care Units and for severe cases, more ventilators can be procured.
3. It is also empirical to spread awareness of this deadly disease in rural areas. Here, the authorities can plan online/offline campaigns to educate people before it hits the masses. Also, the strategies can be framed to impose lockdown in certain states and the period of lockdown can also be calculated based on this forecast.
4. In addition to the above-cited recommendations, special care is to be taken of those patients who are already suffering from other diseases and taking regular treatment from hospitals. Based on the infected cases forecast, local hospitals, schools and some unused official buildings can be converted into COVID relief and cure centres. Based on the prediction results, the supply of required first aid

treatment equipment and medicines can be foreseen. An awareness programme for the first line of medication can be developed. The knowledge of these programmes can be disseminated at different levels.

5. Conclusion

Novel Coronavirus poses a threat to human beings. This has significantly changed the way of thinking towards life. As the pandemic hit masses, the prediction methods offer help to medical practitioners, policymakers, and leaders of the states and countries to combat this disease effectively. The work reported in this paper has discussed difficulties in the forecasting of spread of pandemic and it has offered a probable solution in the form of the proposed grey mathematics-based optimization model. Following are the major conclusions of this work:

- This paper has presented theoretical aspects of GPMs. Also, it has discussed how the accuracy of these models is compromised due to the inherent nature of the models. An Internal Optimization-Based Model has been proposed for addressing these issues. This model has been validated on benchmark time series data. After the validation of this model, two sub-models have been developed with the help of different overlap periods to conduct the forecast of pandemic spread.
- These models are based on the mean values of the infected cases in the three major states and Delhi consisting of different overlap patterns i.e. 5 and 6 days respectively.
- The proposed prediction models are based on internal optimization and also on the hypothesis that performance can substantially be enhanced with the help of a careful selection of the grey model's internal parameters. Further, both models have been tested on three major states and Delhi and forecasting of the infected cases has been done. It is observed that the values of error indices are optimal as compared with non-optimized models.
- The comparison of optimized models and non-optimized conventional models such as GM and NGM has been done in terms of evaluation of error indices. Further, this analysis has been extended to the evaluation of the average rank associated with these models. It has been observed that the proposed models perform satisfactorily as ranks obtained by these models are optimal in comparison to other grey models. Further, it is stated that the results of the proposed models are closely aligned with the actual data.
- For extending the analysis,the average MAPE of proposed IOGM models (place wise) have been evaluated. Moreover, it is observed that the proposed model-II (with a higher overlap period) yields satisfactory results. Based on prediction results, certain suggestions and recommendations have been framed. These recommendations can be further utilized for framing the policies and preventive strategies for the COVID-19 by the Government of India.

The comparative analysis of performance of other grey models for forecasting Corona spread can be a future research direction. It will be interesting to develop new grey models with the application of nature-inspired optimizers in future.

CRediT authorship contribution statement

Akash Saxena: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Visualization, Investigation, Supervision, Software, Validation, Writing – review & editing.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influecne the work reported in this paper.

References

[1] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.

[2] D. Cucinotta, M. Vanelli, WHO Declares COVID-19 a pandemic, Acta Bio-Medica: Atenei Parmensis 91 (1) (2020) 157–160.

[3] F. Petropoulos, S. Makridakis, Forecasting the novel coronavirus COVID-19, PLoS One 15 (3) (2020) e0231236.

[4] Z. Zhao, X. Li, F. Liu, G. Zhu, C. Ma, L. Wang, Prediction of the COVID-19 spread in African countries and implications for prevention and controls: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya, Sci. Total Environ. (2020) 138959.

[5] D. Rafiq, S.A. Suhail, M.A. Bazaz, Evaluation and prediction of COVID-19 in India: A case study of worst hit states, Chaos Solitons Fractals (2020) 110014.

[6] V.K.R. Chimmula, L. Zhang, Time series forecasting of COVID-19 transmission in Canada using LSTM networks, Chaos Solitons Fractals (2020) 109864.

[7] A. Tomar, N. Gupta, Prediction for the spread of COVID-19 in India and effectiveness of preventive measures, Sci. Total Environ. (2020) 138762.

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

[9] P. Arora, H. Kumar, B.K. Panigrahi, Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India, Chaos Solitons Fractals (2020) 110017.

[10] K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. Hyman, P. Yan, G. Chowell, Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, Infect. Disease Model. 5 (2020) 256–263.

[11] W.C. Roda, M.B. Varughese, D. Han, M.Y. Li, Why is it difficult to accurately predict the COVID-19 epidemic? Infectious Disease Modelling (2020).

[12] P. Jiang, Q. Zhou, H. Jiang, Y. Dong, An optimized forecasting approach based on grey theory and Cuckoo search algorithm: a case study for electricity consumption in New South Wales, in: Abstract Appl. Anal., 2014, Hindawi, 2014.

[13] W. Zhou, J.-M. He, Generalized GM (1, 1) model and its application in forecasting of fuel production, Appl. Math. Model. 37 (9) (2013) 6234–6243.

[14] D. Akay, M. Atak, Grey prediction with rolling mechanism for electricity demand forecasting of Turkey, Energy 32 (9) (2007) 1670–1675.

[15] M. Jin, X. Zhou, Z.M. Zhang, M.M. Tentzeris, Short-term power load forecasting using grey correlation contest modeling, Expert Syst. Appl. 39 (1) (2012) 773–779.

[16] D. Ju-Long, Control problems of grey systems, Systems Control Lett. 1 (5) (1982) 288–294.

[17] L. Wu, S. Liu, L. Yao, S. Yan, D. Liu, Grey system model with the fractional order accumulation, Commun. Nonlinear Sci. Numer. Simul. 18 (7) (2013) 1775–1785.

[18] Y. Yang, D. Xue, Continuous fractional-order grey model and electricity prediction research based on the observation error feedback, Energy 115 (2016) 722–733.

[19] L. Wu, H. Zhao, Discrete grey model with the weighted accumulation, Soft Comput. 23 (23) (2019) 12873–12881.

[20] Y. Hu, X. Ma, W. Li, W. Wu, D. Tu, Forecasting manufacturing industrial natural gas consumption of China using a novel time-delayed fractional grey model with multiple fractional order, Comput. Appl. Math. 39 (4) (2020) 1–30.

[21] C.-C. Hsu, C.-Y. Chen, A modified Grey forecasting model for long-term prediction, J. Chin. Inst. Eng. 26 (3) (2003) 301–308.

[22] K. Li, T. Zhang, A novel grey forecasting model and its application in forecasting the energy consumption in Shanghai, Energy Syst. (2019) 1–16.

[23] J. Cui, S.-f. Liu, B. Zeng, N.-m. Xie, A novel grey forecasting model and its optimization, Appl. Math. Model. 37 (6) (2013) 4399–4406.

[24] P.-Y. Chen, H.-M. Yu, Foundation settlement prediction based on a novel NGM model, Math. Probl. Eng. 2014 (2014).

[25] N.-m. Xie, S.-f. Liu, Discrete grey forecasting model and its optimization, Appl. Math. Model. 33 (2) (2009) 1173–1186.

[26] S. Ding, K.W. Hipel, Y.-g. Dang, Forecasting China's electricity consumption using a new grey prediction model, Energy 149 (2018) 314–328.

[27] C. Zheng, W.-Z. Wu, J. Jiang, Q. Li, Forecasting natural gas consumption of china using a novel grey model, Complexity 2020 (2020).

[28] S.-L. Lu, Integrating heuristic time series with modified grey forecasting for renewable energy in Taiwan, Renew. Energy 133 (2019) 1436–1444.

[29] X. Ma, X. Mei, W. Wu, X. Wu, B. Zeng, A novel fractional time delayed grey model with Grey Wolf Optimizer and its applications in forecasting the natural gas and coal consumption in Chongqing China, Energy 178 (2019) 487–507.
[30] P. Zhang, X. Ma, K. She, A novel power-driven grey model with whale optimization algorithm and its application in forecasting the residential energy consumption in China, Complexity 2019 (2019).
[31] L.-C. Hsu, Applying the grey prediction model to the global integrated circuit industry, Technol. Forecast. Soc. Change 70 (6) (2003) 563–574.
[32] C.-C. Hsu, C.-Y. Chen, Applications of improved grey prediction model for power demand forecasting, Energy Convers. Manage. 44 (14) (2003) 2241–2249.
[33] H. Zhuan, Mechanism and application of GM (1, 1) and improvement methods, Syst. Eng. Theory Practice 2 (1993) 56–62.
[34] Z. Hengyang, Modeling mechanism and prerequisites for GM (1, 1) and the revised methods, Syst. Eng.-Theory Methodol. Appl. 2 (3) (1993) 56–62.
[35] S.-C. Chang, H.-C. Lai, H.-C. Yu, A variable P value rolling grey forecasting model for Taiwan semiconductor industry production, Technol. Forecast. Soc. Change 72 (5) (2005) 623–640.
[36] N. Xu, Y. Dang, S. Ding, Optimization method of background value in GM (1, 1) model based on least error, Control Decis. 30 (2) (2015) 283–288.
[37] B. Zeng, C. Li, Forecasting the natural gas demand in China using a self-adapting intelligent grey model, Energy 112 (2016) 810–825.
[38] kaggle data, COVID-19 corona virus India dataset, 2020, https://www.kaggle.com/imdevskp/covid19-corona-virus-india-dataset.
[39] M. of Health, F. Welfare, MHFQ COVID-19 India dataset, 2020, https://www.mohfw.gov.in/.

Dr. Akash Saxena received his Ph.D. and Master of Technology Degree (with Honours) in Electrical Engineering from Malviya National Institute of Technology, Jaipur in 2015 and 2008 respectively and a Bachelor of Engineering Degree (with Honours) in Electrical Engineering from Engineering College Kota in 2001. Dr. Akash is the proud recipient of Chartered Management Institute (Level 3) Diploma in Management which represents a global benchmark for excellence. Also, he is a senior member of IEEE. Dr. Akash Saxena has been working as a Professor in Electrical Engineering Department of SKIT, Jaipur, India since 2014 with almost two decades of teaching.
As an academician, he has been associated with numerous institutions as Member, Doctoral Research Committee and Member, Board of Studies. He has presented his research work at various platforms to showcase his findings to the fellow researchers and providing a pathway to the future researchers. His intellectual work has been published in leading journals in the form of short communication/letters /articles/research papers. A number of prestigious professional organizations have taken him on board as an editor, reviewer and adviser. Dr. Akash Saxena has to his credit a number of accomplished sponsored projects from the prestigious government organizations. Dr. Akash Saxena who has been a passionate researcher is currently immersed into extensive research in the areas of Computational Intelligence, Application of Artificial Intelligence in the Power System, and Smart Grid. He has published more than 70 quality research papers in international refereed journals and conferences and supervised 22 Master theses hitherto.