Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Multiple change point estimation of trends in Covid-19 infections and deaths in India as compared with WHO regions

Pavan Kumar S.T. a,*, Biswajit Lahiri b, Rafael Alvarado c

a College of Community Science, Central Agricultural University, Tura, Meghalaya 794005, India
b College of Fisheries, Central Agricultural University, Lembucherra, Tripura, India
c Carrera de Economía and Centro de Investigaciones Sociales y Económicas, Universidad Nacional de Loja, Loja 110150, Ecuador

A R T I C L E   I N F O

Article history:
Received 13 May 2021
Received in revised form 25 August 2021
Accepted 26 August 2021
Available online 3 September 2021

Keywords:
COVID-19
WHO regions
Multiple change point
Trend estimation
Pandemic

A B S T R A C T

The present study aims at estimating the multiple change points for the time series data of COVID-19 confirmed cases and deaths and trend estimation within the estimated multiple change points (MCP) in India as compared with WHO regions. The data were described using descriptive statistical measures, and for the estimation of change point’s E-divisive procedure was employed. Further, the trend within the estimated change points was tested using Sen’s slope and Mann Kendall tests. India, along with the African Region, American region, and South East Asia regions experienced a significant surge in the fresh cases up to the 5th Change point. Among the WHO regions, The American region was the worst hit by the pandemic in case of fresh cases and deaths. While the European region experienced an early negative trend of fresh cases during the 3rd and 4th change point, but later the situation reversed by the 5th (7th July 2020) and 6th (6th August 2020) change point. The trend of deaths in India and the South-East Asia Region was similar, and global deaths had a negative trend from the 4th (17th May 2020) Change point onwards. The change points were estimated with prefixed significance level $\alpha < 0.002$. Infections and deaths were positively significant for India and SEARO region across change points. Infection was significant at every 30 days interval across other WHO regions, and any delay in the infections was due to the interventions. The European region is expected to have a second wave of positive
infections during the 5th and 6th change points though the early two change points were negatively significant. The study highlights the efficacy of change point analysis in understanding the dynamics of covid-19 cases in India and across the world. It further helps to develop effective public health strategies.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

A decade after the outbreak of SARS (Severe Acute Respiratory Syndrome) and the Middle East Respiratory Syndrome Corona Virus (MERS-CoV) as a highly pathogenic virus in Middle Eastern countries (El Zowalaty and Järhult, 2020; Zaki et al., 2015), the world has been facing a serious problem of Novel Corona Virus Disease, which is well known as Covid-19. The Novel Corona Virus (2019 n-CoV) or Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) of zoonotic origin was first noticed in Wuhan City of Hubei Province of China (Guan et al., 2020; Lin et al., 2020; Zhu et al., 2019; Zu et al., 2020). It reaches fast across the world and became a public health emergency of international concern in January, declared by World Health Organization as pandemic on 11th March 2020. Corona virus can cause sickness in both human beings and animals. Patients with corona virus diseases show symptoms like dry cough, fever, fatigue, sore throat, and loss of sense of taste and smell etc. About 80 percent of people recover without treatment in hospitals, as per the (World Health Organization, 2020a) WHO Report (2020).

As of 4th March, 2021 globally 105 million cases of Covid-19 have been reported and caused 2.19 million deaths and India has reported 20.7 million cases and 0.263 million deaths. After the first reporting of a confirmed case of COVID-19 infection in Kerala, on January 30, 2020, India became the largest affected country in Asia by June 2020. Six major cities (Mumbai, Delhi, Ahmedabad, Chennai, Pune, and Kolkata) account for about half of all reported cases in the country (Ministry of Health and Family Welfare, 2020; Sahoo and Sapra, 2020). Thus, an effective model to track and predict the course of the epidemic is the need of the hour. In American region 38.7 million cases with 1.03 million deaths, Europe 30.1 million cases with 0.615 million deaths, South East Asia Region 24.3 million total cases with 0.334 million deaths have been reported. The United States of America ranked top in death rates with 1.03 million deaths. The death cases are highest in the American region (AMRO: 48%), followed by the European region (EURO: 34%), the South-East Asia Region (SEARO: 8%), the Eastern Mediterranean Region (EMRO: 6%), the African region (AFRO: 3%) and the West Pacific Region (WPRO-1%) (World Health Organization, 2020a).

Studies on various aspects of Covid-19 have already been done by many researchers globally. Many mathematical models worldwide are explored to contribute to the understanding of COVID-19 (Çakan, 2020; Ivorra et al., 2020; Khan and Atangana, 2020; Ndaïrou et al., 2020). A real-time data-driven model to track and predict the course of the epidemic is found effective in developing strategies for public health and economic crisis (Fanelli and Piazza, 2020; Gudbjartsson et al., 2020; Rafiq et al., 2020; Sahoo and Sapra, 2020; Salathé et al., 2020) The ARIMA model (Auto-Regressive Integrated Moving Average) has been widely used in modeling and forecasting the time series data in COVID-19 studies. It explains a given time series based on its past values, that is, its lags and the lagged forecast errors, so that equation can be used to forecast future values (Benvenuto et al., 2020; Dehesh et al., 2020; Gupta and Pal, 2020; Shi and Fang, 2020) The most frequently used models to study the epidemic evolution of disease with time are the SIR model (Susceptible, Infectious, or Recovered) and its variants (Kermack and McKendrick, 1927; Murray, 2002). The spatial distribution of the infection is however hardly be addressed by the SIR model, despite its usefulness in investigating the time evolution of an epidemic disease. In this regard, an epidemiological model that encompasses a fractal structure allows a more detailed description of the observed data about the virus in terms of geographical distribution (Abbasi et al., 2020). The local stability of equilibrium points with treatment, some sufficient conditions, and uniform asymptotic stability of
equilibria with general incidence rate were ensured by Fractional Susceptible–Exposed–Infectious–Removed (SEIR) model (Ricardo, 2018; Yang and Xu, 2020). The modified SEIQRP model synthesized from the generalized fractional-order SEIR model successfully captured the development process of COVID-19. It provides an important reference for understanding the trend of the outbreak in the USA (Xu et al., 2020). This epidemic spreading model on a network using concepts from percolation theory was tried out to describe the effects of lock-downs within a population. The network model performed well by using constant parameters, while more involvement of time-dependent parameters to be achieved with similar fitting accuracy in the SIR model (Crocco and Roman, 2020).

Deep Learning-based models were also used for predicting the number of novel coronavirus (COVID-19) positive reported cases by applying Recurrent Neural Network (RNN) based long-short term memory (LSTM) and its variants such as Deep LSTM, Convolutional LSTM, and Bidirectional LSTM. The models provided high accuracy for daily predictions and weekly predictions (Arora et al., 2020; Chimmula and Zhang, 2020). An evolutionary data analytics method called the Genetic Evolutionary Programming (GEP) on two datasets, e.g. confirmed cases (CC) and death cases (DC), were used to mathematically model the potential effect of corona virus in the fifteen most affected countries of the world. The model used simple linkage functions and provided highly reliable results for the time series prediction of COVID-19 in these countries (Salgotra et al., 2020).

Applications of the forecasting and time series analysis on the COVID-19 pandemic data series in Turkey indicate that the daily numbers of deaths and cases are expected to decrease in the short term (Kinaci et al., 2021). Similarly, the regression models for death cases in Iran showed an increasing trend but with some evidence of turning. The infection rate and the population density are having a polynomial relationship, signifying the importance of estimation of multiple change points and trends in a pandemic situation (Pourghasemi et al., 2020). The non-parametric Pettitt test (Pettitt, 1979) is an effective method of identifying the change in the temporal trend in any time series. Due to its sensitivity to breaks in the middle of temporal records, it has been extensively used (Gao et al., 2010; Gupta et al., 2020; Hänsel et al., 2016; Jaiswal et al., 2015; Mallakpour and Villarini, 2016; Wijngaard et al., 2003).

The change point analysis deciphers a complex pattern of change. It is a non-parametric technique and hence does not depend on any distributional assumptions. Due to this reason, non-parametric methods are being used more widely than parametric methods. A change point method mainly detects the point where the change has occurred significantly in the time-ordered observations (Matteson and James, 2014). The change point analysis is applied in a wide variety of fields like Financial Modeling (Teyssière and Kirmân, 2007), Bioinformatics (Muggeo and Adelfio, 2011; Vladimir et al., 2005), Engineering, Climatology (Li and Lund, 2012), Neuro Science (Koepcke et al., 2016) and other fields of science. The technique also applied to estimate the change point in the high dimensional time series data (Barigozzi et al., 2018) and to estimate the change points in the presence of outliers (Fearnhead and Rigail, 2019). In experimental and mathematical sciences, so-called retrospective AMOC (at most one change) change point problem arises in different issues like, epidemiological, quality control etc. But classical change point problems are considered in various bio-statistical and engineering applications (Alatawna et al., 2015). The Multiple change point for multivariate data using the e-divisive method estimates the change point hierarchically and tests the statistical significance of each estimated change points (Matteson and James, 2014). Often this method showed consistency due to the estimation procedure. In the field of epidemiology, the change point was estimated to identify the outbreak of certain diseases (Texier et al., 2016), and it helps to make appropriate decisions about the prevention of further outbreaks. The change points explain the pattern and infection cycle of an epidemic. Several studies have been conducted in India, Italy, China, and United States on the estimation of change points of the covid-19 epidemic by using the Mann–Kendall test and Sen’s slope to evaluate and verify the trends in corona virus disease (Jiang et al., 2020; Goswami et al., 2020; Gupta and Pradhan, 2020; Ison, 2020; Pedersen and Meneghini, 2020; Yang et al., 2020).

In this paper, we proposed to estimate the multiple change points and existing trends within the estimated change points of covid-19 cases and deaths in India as compared to the WHO region. The results from this study could be a yardstick in understanding the directional change of infection of
the virus in India and WHO regions as well. The detection of single or multiple change points alone may not bring clarity but, the dynamics within the change points would be extremely important. These data-oriented approaches help the researchers to further investigate the reasons for the varying trend of infection and deaths due to Covid-19.

2. Materials and methods

2.1. Data for the study

The World Health Organization (WHO), as a global organization manages, and maintains a wide range of data related to global public health and wellbeing. The WHO created a separate dashboard for the purpose of displaying the real-time global database of COVID-19 and also provides a downloadable database for global researchers. This database is comprised of published records by the Ministries of various countries across the globe. The present study is based on the secondary data of the daily fresh cases and deaths of COVID-19, collected from the WHO website (World Health Organization, 2020a). The time-series data across WHO regions comprises of the time frame starting from 4th January 2020 up to 9th September 2020 (250 Days), and for India, new cases and deaths data collected in the time window 30th January 2020 (Since first cases detected on 30th January 2020 in Kerala state) to 9th September 2020 (224 Days). The data were summarized using descriptive measures.

2.2. E-divisive method

The change points in this method are estimated with a data splitting approach. The first step in detecting the change point is to divide the data series into candidate phases, $X_{1:t}$ and $X_{t+1:n}$ for which the characteristic function differ maximally.

$$\hat{Q}(\tau) = \frac{\tau(n-\tau)}{n} \left[ \frac{2}{\tau(n-\tau)} \sum_{i=1}^{r} \sum_{j=r+1}^{n} \|X_i - X_j\| \right.$$  

$$- \left( \frac{\tau}{2} \right)^{-1} \sum_{i=1}^{r} \sum_{k=i+1}^{r} \|X_i - X_k\| - \left( \frac{n-\tau}{2} \right)^{-1} \sum_{j=r+1}^{n-1} \sum_{k=j+1}^{n} \|X_j - X_k\| \right]$$

where $\| \cdot \|$ denotes the Euclidean distance. $n$ denotes the size of time series. The optimum change point is estimated by considering, which $\tau^*$ value maximizes $\hat{Q}$. The second step is to estimate if the change point is significant through the permutation test. After the change point is estimated, then the significance is tested at pre-specified significance level $\alpha < 0.002$ level, a point which maximizes the value of $\hat{Q}$. The method is conducted by generating $R$ permuted time series obtained by randomly changing the time order of the sequence.

At the third step, after the change point obtained in the first step found significant, the series is further divided into one more phase to find any additional change points. Those points are hierarchically structured. This process continues till the optimal (Non-significance of the change point) change points obtained for the data series, and no further bisection of data into phases is showing significance. The e-divisive method is consistent with the estimation of change points (Matteson and James, 2014). The data has been analyzed using the R package (RStudio Team, 2020) esp. for non-parametric multiple change point analysis of multivariate data.

2.3. Mann–Kendal’s Z

Mann–Kendal (Kendall, 1975) test is mainly used to test whether univariate time series data has a monotonic trend (Upward or Downward). It assumes that the data do not show any autocorrelation. The hypothesis under this test is as follows;
\( H_0 = \text{No trend} \)
\( H_1 = \text{There exists a trend (Upward or Downward)} \)

The test statistic is given by

\[ S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}(x_j - x_i) \]

where \( x_j \) and \( x_i \) are sequential data values. The variance of \( S \) is given as

\[ \text{Var} = \frac{1}{18} \left[ n(n-1)(2n+5) - \sum_t f_t (f_t - 1)(2f_t + 5) \right] \]

where ‘\( t \)’ varies over set of tied ranks and \( f_t \) is the frequency that rank \( t \) appears. The test statistic will be

\[ z = \begin{cases} 
\frac{(s - 1)}{se}, & S > 0 \\
0, & S = 0 \\
\frac{(s + 1)S}{se}, & S < 0 
\end{cases} \]

where \( se = \text{Standard deviation: if there is no monotonic trend. The significance of the trend is tested at 5% (} \alpha = 0.05 \text{) level.} \)

2.4. Sen’s slope \( Q \)

The magnitude of the trend is estimated by Sen’s slope. The magnitude of the trend estimated using Sen’s Slope \( Q \) (Sen, 1968) is based on the median values of variables \( (X_{ij}) \). The test statistics is given by

\[ Q = \begin{cases} 
\frac{\beta_{\frac{N+1}{2}}}{N} & \text{where } N \text{ is odd} \\
\frac{1}{2} \left( \beta_{\frac{N}{2}} + \beta_{\frac{N+2}{2}} \right) & \text{where } N \text{ is even} 
\end{cases} \]

where \( N \) represents the length of the sample and \( \beta \)-Slope estimator. A positive \( Q \) value indicates an upward trend and a negative value represents a downward trend. The magnitude of the trend is tested at 5% (\( \alpha = 0.05 \)) significance level is considered.

3. Results

3.1. Descriptive statistics

Descriptive statistics of the Covid-19 data presented in Tables 1 and 2 revealed that on average 20,000 new fresh cases and 330 deaths were reported daily till 9th September in India. The major portion of new fresh cases and deaths in the SEARO region was reported from India. Total death cases in India were more than the total average deaths from EURO and AFRO region together. As such, there was no significant peak was observed in the case of fresh cases in India but, the fresh cases continued to increase steadily. The daily average covid-19 confirmed cases were estimated to be 110,010 across the world, and the major portion of the confirmed cases globally was contributed from the AMRO region (More than 50 percent of the confirmed cases) followed by SEARO, EURO, EMRO, AFRO, and WPRO region.

Nearly 55 percent of the total confirmed death cases across the world were reported from the American region, followed by EURO (25%), SEARO (9.7%), EMRO (6%), and the rest were from the AFRO and WPRO regions. With reference to the death cases, WPRO, EURO, and AFRO regions were highly inconsistent, and the WPRO region reported a peaked number of death cases in a single day after 100 days of the COVID-19 outbreak (Table 2).
Table 1
Descriptive statistics new Covid-19 cases in India, world and across WHO regions.

| Parameters | AFRO | AMRO | EMRO | EURO | SEARO | WPRO | India | World |
|------------|------|------|------|------|-------|------|-------|-------|
| Mean       | 4390 | 57000| 8160 | 18500| 19900 | 2110 | 19500 | 11000 |
| Median     | 1990 | 44200| 7990 | 19700| 4610  | 1310 | 5610  | 85800 |
| SD         | 5420 | 53800| 6880 | 13100| 2750  | 2050 | 26200 | 98000 |
| Range      | 20600| 173000|21800| 44100| 97100 | 15200| 90800 | 307000|

Where SD-Standard Deviation.

Table 2
Descriptive statistics new Covid-19 death cases in the world and across WHO regions.

| Parameters | AFRO | AMRO | EMRO | EURO | SEARO | WPRO | India | World |
|------------|------|------|------|------|-------|------|-------|-------|
| Mean       | 93.5 | 1980 | 215  | 895  | 348   | 45.6 | 330   | 3580  |
| Median     | 42.5 | 2200 | 159  | 422  | 125   | 29.5 | 143   | 4240  |
| SD         | 118  | 1710 | 179  | 1170 | 422   | 89.0 | 382   | 2600  |
| Range      | 632  | 7750 | 576  | 5170 | 2090  | 1310 | 2000  | 12400 |

Where, SD-Standard Deviation.

Table 3
Multiple change point detection for new Covid-19 cases in the world and India.

| India       | Change Point | Day       | p-value  | 31     | 61     | 91     | 121    | 169   | \( \hat{K} \) = 6 |
|-------------|--------------|-----------|----------|--------|--------|--------|--------|-------|------------------|
| Day         | (30-Jan-20)  | (29-Feb-20)| (30-Mar-20)| (29-Apr-20)| (29-May-20)| (08-Jul-20)|        |       |                  |
| Day         | (04-Jan-20)  | (03-Feb-20)| (24-Mar-20)| (17-May-20)| (18-Jun-20)| (18-Jul-20)|        |       |                  |

Values in the parenthesis are p-value, T = 250 for WHO regions, T = 224 for India, Alpha = 1.

3.2. Multiple change point for new cases and deaths: India’s scenario

Results from Table 3, 4, S1 and S2 reveal that there was a total of 7 change points (\( \hat{K} \)) estimated for the AFRO and WPRO region whereas, 6 change points were estimated for the rest of the region and also for India. The significance level (\( \alpha \)) < 0.002 was considered for estimation of each change points, and R = 499 iterations used performing the permutation tests. The e-divisive procedure was executed with \( \alpha = 1 \). In the case of the India and World scenario, the first significant change point with respect to change in mean (\( \mu \)) was estimated on the 31st day after a very first case was reported (Fig. 1 and S4 Fig.). The significant change points were estimated at every 30 days interval for India and WPRO region up to 121 days of covid-19 with a significance level \( \alpha < 0.002 \). After that, for India, it almost took approximately 50 (8th July-20) days to reach the next significant change point, and it was clear that the spread of the virus narrowed its gap in India (S1 Table). It was clear that the number of daily fresh cases in India were declined between the 5th and 6th change points while fresh cases around the world were continued in 30 days cycle. The number of fresh deaths was declined from the 4th change point onwards in India (Delayed by a week between the 4th and 5th change point). Around the globe, the fresh cases declined from the 3rd change point onwards. The gap between the change points was continued to widen till the 5th change point.

3.3. WHO scenario: New cases and deaths

In the case of AMRO (S1 Fig.), EMRO (S2 Fig.), EURO (S3 Fig.), and SEARO (S5 Fig.) regions, the virus transmission between the 3rd to 4th change points (March–May, 2020) were a little delayed. Later, subsequent change points were detected at 30 days intervals AMRO, EURO, and SEARO except for the EMRO region (5th and 6th change points). Similarly, the AFRO region (S6 Fig.) experienced that the spread of the virus was delayed between 4th and 5th CP (April–May, 20), but after this point, it had fallen into 30 days cycle. The results indicate that the spread of virus infection was
Table 4

| India  | Change point | Day       | p-value | \(\hat{K}\) | \(p\)-value | Day       |
|--------|--------------|-----------|---------|-------------|-------------|-----------|
|        | 1            | (30-Jan-20) | (0.002) |              | (0.002)     | (04-Jan-20) |
|        | 37           | (06-Mar-20) | (0.002) |              | (0.002)     | (25-Feb-20) |
|        | 67           | (05-Apr-20) | (0.002) |              | (0.002)     | (26-Mar-20) |
|        | 97           | (05-May-20) | (0.002) |              | (0.004)     | (04-May-20) |
|        | 134          | (11-Jun-20) | (0.002) |              | (0.118)     | (16-Jul-20) |
|        | 176          | (23-Jul-20) | (0.002) |              |             |           |

Values in the parenthesis are \(p\)-value, \(T = 250\) for WHO regions, \(T = 224\) for India, Alpha = 1.

Fig. 1. Multiple change point detection for Daily New confirmed Covid-19 cases in the World, and India.

Fig. 2. Change in a Univariate Gaussian Sequence for World New Covid Cases

Fig. 3. Change in a Univariate Gaussian Sequence for India New Covid Cases

Fig. 4. Change in a Univariate Gaussian Sequence for WHO regions

Fig. 5. Change in a Univariate Gaussian Sequence for EMRO region

Scores of fresh cases and deaths for the SEARO region (includes India) was on the higher side as compared to other regions (Fig. 2). The trend of death cases was slightly slower than that of the number of new confirmed cases on daily basis in India (Figs. 3 & 4) and across the WHO region (Fig. S7) as well. The estimated change points on 12th July 2020 and 9th August 2020 for EMRO and SEARO respectively were non-significant (S2Table).

New death cases due to covid-19 also reported similarly as new confirmed cases across regions (Fig. 5, S8–S13 Fig.). The death cases of covid-19 were delayed between the 4th and 5th change point. On average, it took approximately 44 days otherwise, the death cycle was at 30 days interval. In the case of the AMRO region (S8 Fig.), the first significant change point was observed in 93 days \((p < 0.002)\) after the first death reported in the region but later change was found in 30 days (Table 4 & S2 Table).

3.4. Trend estimation within the estimated multiple change points: India’s and world scenario

The trend within estimated change points of fresh covid-19 cases is presented in Table 5. The trend of new cases in India continued to increase exponentially. The magnitude of the trend during the 2nd change point was 2.429 \((p < 0.05)\), while the same was 893.474 \((p < 0.05)\) during the 6th change point. The rate of increase in trend was significant across change points. The rate of increase between 2nd and 3rd change point was at 95 percent followed by 72 percent between 3rd and 4th, 61 percent between 4th and 5th and 47 percent between 5th and 6th change point. The
Fig. 2. Trend of Covid-19 new cases & Deaths in India and world.

Fig. 3. Comparative Trend of Covid-19 new confirmed and death cases in India.

Fig. 4. Multiple change point detection for Daily New Death Covid-19 cases in India.
growth rate between change points continued to fall but, the magnitude of the trend of the fresh cases kept on increasing. The magnitude of the trend in case of the world fresh cases was also in the increasing trend till 5th change point. The highest growth rate has been observed between the 1st and 2nd change point i.e., 80 percent, and between the 3rd and 4th change point whereas in the final change point (6th) the magnitude declined from 3070 to 256 (−1098%).

For India, the magnitude of the fresh deaths increased up to the 5th change point (11th June 2020), while in the final change point, there was a slight decline in the trend of fresh deaths. The rate of increase was similar to that of fresh cases. The rate was decreased by negative 5 percent between the 5th and 6th change points, and it has been observed that the spread of the virus was widened its gap during the 5th and 6th change points. In the case of world death cases, the trend was positive significant up to 3rd change point later the trend found non-significant (4th and 5th change point). The growth rate was found positive during the 1st and 2nd change points, and it was negative from 3rd change point onwards. The negative rate due to fewer death cases reported from the WHO regions ie., AMRO, EURO, EMRO, and WPRO.

3.5. WHO region scenario

The Sen’s slope values indicate that the African region had experienced a significant monotonic positive trend in the covid-19 cases till the 5th change point. After that, the region’s confirmed cases declined at a significant rate (Slope = −263 at p < 0.05), and a similar trend was observed in the American region (6th Change point). The negative trend was observed in the East Mediterranean region after 125 days of Covid-19 (5th Change point = −75.0 at p < 0.05 and 6th Change point = −14.4 at p < 0.05) (S3 Table).

The European region experienced an early significant negative trend after 45 days (17th February), and it prevailed till 129 days (4th CP). But the trend of virus infection reversed after 129 days
(Second week of May). There was a significant surge in the fresh cases and deaths due to covid-19 for South East Asia and in India throughout the study period. In the WPRO region, the cases had an abrupt negative trend during the 2nd CP, while it was turned to positive in the 3rd CP, and in the latter, there was no monotonic trend observed in the new COVID-19 infections. The surge in the number of fresh deaths was significantly higher in the American region during the initial 100 days, while later, the trend remained stagnant (Fig. S1). In the EURO region, the fresh deaths significantly decreased after 50 days (22nd February, 3rd Change point Slope \(-38.6\), 4th Change point slope\(-25.1\)) consequently for 2 change points. In the WPRO region, the death cases decreased after 31 days till the 5th change point. Later the trend was found positive during the 6th change point but found non-significant. The death cases reported across the world were on the decreasing trend from the 83rd day onward (26th March), but the trend was not significant (S4 Table).

4. Discussion

The e-divisive method estimates the change points in the hierarchical fashion with prefixed significance level $\alpha < 0.002$. The findings of the study clearly indicate that the infection of the virus was significantly changing every 30 days interval in India and most of the regions as well. The significant change points were estimated in a similar fashion for WPRO and India. In India, the infection of the virus was delayed after 29th May i.e., the 5th change point. It took 48 days to find the next significant change point. Delay in the infection was mainly attributed to government interventions like imposition of lockdown, public awareness with regard to social distancing, wearing masks in public, and other public health measures (Flaxman et al., 2020). The nationwide lockdown in India was announced in the month of March (Wikipedia, 2021), and its effect was observed after 30 days in case of new confirmed cases and 60 days in case of death cases. The trend of fresh cases and deaths was on the increasing trend in India, as compared to other regions. The trend was almost similar to that of the SEARO region in both fresh and death cases, but the trends of fresh cases have a wider range (Fig. 2). The trend of new cases exceeds a wider range than that of deaths in India (Fig. 3).

The trend of India’s covid-19 infections was on the increasing mode in all the estimated change points, while the world cases were on declining trend during 6th change point but the trend in both cases was significant. The box whisker plots revealed that the trend of fresh cases in India surpassed the world’s trend i.e., the fresh infections were significantly recorded daily in India (Gupta et al., 2020). The early lockdown process yielded a significant downtrend of fresh cases in the rest of the world during the last change point. Increased testing and better medical facility in the developed countries were also resulting in reduced infections in the world cases (Stephen, 2021; WHO, 2020; World Health Organization, 2020a,b; World Health Organization, 2020b). The trend of fresh deaths was increasing in all the estimated change points for India. Whereas, the world deaths declined from the 4th change point (4th May 2020) onwards and found a negative trend during the 5th change point. There was a huge improvement in the trend of fresh deaths across the world over the initial lockdown; the countries were able to contain the number of deaths by imposing strict covid protocols (Gatto et al., 2020; Leung et al., 2020; Wurtzer et al., 2020). Many countries in the WHO region started to unlock the process during May–June 2020 in phases to boost the economy by providing relief to people living in the lower economic stratum, job losses and unemployed population (Gopalan and Misra, 2020; Agoramorthy and Hsu, 2021). Lifting the restrictions on various public activities resulted in a surge in cases and narrowed the infection gap between 5th and 6th change points across the WHO regions. Better health infrastructure, public attitude towards health warnings helped reduced infections and deaths cases in most of the world regions. Any changes in fresh cases or deaths could be inferred as the result of interventions and mitigation strategies (Delhing et al., 2020).

5. Conclusion

The COVID-19 pandemic came out a big challenge for human civilization. The containment of this virus has also become a challenge to the scientists’ community across the world. Different
data-driven approaches are needed presently to evaluate the infection and re-infection cycle of the virus to develop a uniform mitigating strategy and standardize the global protocol. Data-oriented approaches are important because the level of infection and death rates are varying from region to region. Since the infections and deaths are significantly changing every 30 days, the country should be well prepared to tackle the health crisis. The efforts should be made to contain the next level of possible infections. In the case of the European region, it is clear that there could be another wave of infections by looking at the latest trends during the 5th and 6th change points, and for the WPRO region also the same trend is observed. The new infections and deaths were still significant for the SEARO region and India. Therefore, estimation of change points using e-divisive at different time frames would yield a significant breakthrough in understanding the change in infection at a particular time frame, and it helps to know the days to reach the next significant mean infections. Along with change point analysis, the trend analysis would yield a better result in forming a global health strategy. The combination of trend estimation and change point analysis plays a crucial role in reflecting the magnitude of infections across estimated change points, and also helps to understand the infection cycle.

CRediT authorship contribution statement

**Pavan Kumar S.T.**: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Visualization, Data curation, Resources. **Biswajit Lahiri**: Conceptualization, Writing – review & editing, Formatting, Supervision. **Rafael Alvarado**: Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Authors are thankful to the data management teams of the WHO for the quality database. Authors are also thankful to the Central Agricultural University, Imphal, Manipur, India, for providing necessary infrastructural support for conducting the study.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical approval

Ethical approval is not required for the study because it does not include any human subject.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spasta.2021.100538.
Xu, C., Yu, Y., Chen, Y., Lu, Z., 2020. Forecast analysis of the epidemics trend of COVID-19 in the USA by a generalized fractional-order SEIR model. Nonlinear Dynam. 101, 1621–1634. http://dx.doi.org/10.1007/s11071-020-05946-3.

Croccolo, F., Roman, H.E., 2020. Spreading of infections on random graphs: A percolation-type model for COVID-19. Chaos Solitons Fractals 13 (110077). http://dx.doi.org/10.1016/j.chaos.2020.110077.

Arora, P., Kumar, H., Panigrahi, B.K., 2020. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. Chaos Solitons Fractals 139, 110017. http://dx.doi.org/10.1016/j.chaos.2020.110017.

Chinnula, V.K.R., Zhang, L., 2020. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos Solitons Fractals 135, 109864. http://dx.doi.org/10.1016/j.chaos.2020.109864.

Salge, R., Gandomi, M., Gandomi, A.H., 2020. Evolutionary modelling of the COVID-19 pandemic in fifteen most affected countries. Chaos Solitons Fractals 140 (110118), 1–45. http://dx.doi.org/10.1016/j.chaos.2020.110118.

Kinaci, H., Unsal, G.M., Kasap, R., 2021. A close look at 2019 novel coronavirus (COVID-19) infections in Turkey using time series analysis and efficiency analysis. Chaos Solitons Fractals 143, 1–11. http://dx.doi.org/10.1016/j.chaos.2020.110583.

Pourghasemi, H.R., Poutian, S., Heidari, B., Farajzadeh, Z., Shamsi, S.R.F., Babaie, S., Khoosavi, R., Etemadi, M., Ghanbarian, G., Farhadi, A., Safaeian, R., Heidari, Z., Tarazkar, M.H., Tiefenbacher, J.P., Azmi, A., Sadeghian, F., 2020. Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). Int. J. Infect. Dis. 98, 90–108. http://dx.doi.org/10.1016/j.ijid.2020.06.058.

Pettitt, A.N., 1979. A non-parametric approach to the change point problem. J. R. Stat. Soc. C 28, 126–135. http://dx.doi.org/10.1111/j.1467-9868.1979.tb00237.x.

Gao, P., Mu, X.-M., Wang, F., Li, R., 2010. Changes in stream flow and sediment discharge and the response to human activities in the middle reaches of the Yellow River. Hydrofl. Earth Syst. Sci. Discuss. 7, 6793–6822. http://dx.doi.org/10.5194/hessd-7-6793-2010.

Gupta, A., Pradhan, B., Nizam, K., Maulud, A., 2020. Estimating the impact of daily weather on the temporal pattern of COVID-19 outbreak in India. Earth Syst. Environ. 4, 523–534. http://dx.doi.org/10.1007/s41748-020-00179-1.

Hänsel, S., Medeiros, D.M., Matschullat, J., Petta, R.A., de Mendonça, S.J., 2016. Assessing homogeneity and climate variability of temperature and precipitation series in the capitals of north-eastern Brazil. Front. Earth Sci. 4, 1–21. http://dx.doi.org/10.3389/feart.2016.00029.

Jaiswal, R., Lohani, A., Tiwari, H., 2015. Statistical analysis for change detection and trend assessment in climatological parameters. Environ. Process. 2, 729–749. http://dx.doi.org/10.1007/s40710-015-0105-3.

Mallakpour, I., Villarini, G., 2016. A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. Hydrofl. Sci. J. 61 (2), 245–254. http://dx.doi.org/10.1002/hyss.2015.1008482.

Wijngaard, J.B., Klein Tank, A.M.G., Können, G.P., 2003. Homogeneity of 20th century European daily temperature and precipitation series. Int. J. Climatol. 23 (6), 679–692. http://dx.doi.org/10.1002/joc.906.

Matteon, David S., James, Nicholas A., 2014. A nonparametric approach for multiple change point analysis of multivariate data. J. Amer. Statist. Assoc. 109 (505), 334–345. http://dx.doi.org/10.1080/01621459.2013.849605.

Teyssièr, G., Kirman, 2007. In: Teyssière, G., Kirman, A.P. (Eds.), Long Memory in Economics. http://dx.doi.org/10.1007/978-3-540-34625-8.

Muggeo, V.M.R., Adelfio, G., 2011. Efficient change point detection for genomic sequences of continuous measurements. Bioinformatics 27 (2), 161–166. http://dx.doi.org/10.1093/bioinformatics/btq647.

Vladimir, N.M., Dorman, K.S., Fang, F., Suchard, M.A., 2005. Dual multiple change-point model leads to more accurate recombination detection. Bioinformatics 21 (13), 3034–3042. http://dx.doi.org/10.1093/bioinformatics/bti459.

Li, S., Lund, R., 2012. Multiple change point detection via genetic algorithms. J. Clim. 25 (2), 674–686. http://dx.doi.org/10.1175/2011JCLI4553.1.

Koeckele, L., Ashida, G., Kretzberg, J., 2016. Single and multiple change point detection in spike trains: Comparison of different CUSUM methods. Front. Syst. Neurosci. 10, 51. http://dx.doi.org/10.3389/fnsys.2016.00051.

Barigozzi, M., Cho, H., Fryzlewicz, P., 2018. Simultaneous multiple change-point and factor analysis for high-dimensional time series. J. Econometrics 206, 187–225. http://dx.doi.org/10.1016/j.jeconom.2018.05.003.

Fearnhead, P., Rigail, G., 2019. Changepoint detection in the presence of outliers. J. Amer. Statist. Assoc. 114 (525), 169–183. http://dx.doi.org/10.1080/01621459.2017.138546.

Alatawna, L., Yancu, Y., Gurevich, G., 2019. Change point trend analysis of GNI per capita in selected European countries and Israel. 2015. In: Proceedings of the 9th International Days of Statistics and Economics, Prague, September 10–12, 2015.

Texier, Gaëtan, Farhadi, Magnim, Pellegrin, Liliane, Jackson, Michael L., Meynard, Jean-Baptiste, Deparis, Xavier, Chaudet, Hervé, 2016. Outbreak definition by change point analysis: a tool for public health decision? BMC Med. Inform. Decis. Mak. 16, 33. http://dx.doi.org/10.1186/s12911-016-0271-x.

Jiang, F., Zhao, Z., Shao, X., 2020. Time series analysis by series analysis of COVID-19 infection curve: A change-point perspective. J. Econometrics Retrieved from https://dx.doi.org/10.1016/j.jeconom.2020.07.039.

Goswami, K., Bharali, S., Hazarika, J., 2020. Projections for COVID-19 pandemic in India and effect of temperature and humidity. Diabetes Metabol. Syndr.: Clin. Res. Rev. 14 (5), 801–805. http://dx.doi.org/10.1016/j.dsx.2020.05.045.

Gupta, A., Pradhan, B., 2020. Assessment of temporal trend of Covid-19 outbreak in India. http://dx.doi.org/10.31219/osf.io/qyre6.

Ison, D., 2020. Statistical procedures for evaluating trends in coronavirus disease-19 cases in the United States. Int. J. Health Sci. 14 (5), 23–31. PMID: 32952502.

Pedersen, M.G., Meneghini, M., 2020. Data-driven estimation of change points reveal correlation between face mask use and accelerated curtailment of the COVID-19 epidemic in Italy. Retrieved from https://dx.doi.org/10.1101/2020.06.29.2014152.

Yang, W., Deng, M., Li, C., Huang, J., 2020. Spatio-temporal patterns of the 2019-nCov epidemic at the country level in Hubei Province, China. Int. J. Environ. Res. Public Health 17 (2563), 1–11. http://dx.doi.org/10.3390/ijerph17072563.
RStudio Team, RStudio Team, 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, URL http://www.rstudio.com/.

Kendall, M.G., 1975. Rank Correlation Methods. Oxford University Press, New York, NY.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall’s Tau. J. Amer. Statist. Assoc. 63 (324), 1379–1389. http://dx.doi.org/10.1080/01621459.1968.10480934.

Flaxman, S., Mishra, S., Gandy, A.H., Juliette, T.U., Thomas, A.M., Helen, C., Charles, W., Harrison, Z., Tresnia, B., Jeffrey, W.E., Mélodie, M., Azra, C.G., Christl, A.D., Steven, M.R., Michaela, A.C.V., Neil, M.F., Lucy, C.O., Samir, B., Imperial College COVID-19 Response Team, 2020. Estimating the effects of non-pharmaceutical interventions on COVID–19 in Europe. Nature http://dx.doi.org/10.1038/s41586-020-2405-7.

Wikipedia, 2021. Covid-19 lockdown in India. https://en.wikipedia.org/wiki/COVID-19_lockdown_in_India. Last Accessed on 4th 2021.

Stephen, Mary, 2021. England in Lockdown: On rapid spread of Covid-19 variant. 6th January 2021. https://www.thehindu.com/opinion/editorial/england-in-lockdown-on-rapid-spread-ofcovid-19-variant/article33504233.ece. Last Accessed on 4th 2021.

WHO, WHO, 2020. Situation Report. World Health Organization, Available online: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.

Word Health Organization, 2020a. Weekly update on Corona Virus Disease (COVID-19), 27th September 2020.

Word Health Organization, 2020b. Weekly update on Corona Virus Disease (COVID-19), 11th October, 2020.

World Health Organization, 2020b. WHO Western Pacific regional action plan for response to large-scale community outbreaks of COVID-19.

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., Rinaldo, A., 2020. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. Proc. Natl. Acad. Sci. USA 117 (19), 10484–10491. http://dx.doi.org/10.1073/pnas.2004978117.

Leung, K., Wu, J.T., Liu, D., Leung, G.M., 2020. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet 395, 1382–1393. http://dx.doi.org/10.1016/S0140-6736(20)30746-7.

Wurtzer, S., Marechal, V., Mouchel, J.M., Maday, Y., Teyssou, R., Richard, E., Almayrac, J.L., Moulin, L., 2020. Evaluation of lockdown impact on SARS-CoV-2 dynamics through viral genome quantification in Paris wastewaters. Eurosurveillance 25 (50), http://dx.doi.org/10.2807/1560-7917. pii=2000776.

Gopalani, H.S., Misra, A., 2020. COVID-19 pandemic and challenges for socio-economic issues, healthcare and National Health Programs in India. Diabetes Metabol. Syndr.: Clin. Res. Rev. 14, 757–759. http://dx.doi.org/10.1016/j.dsx.2020.05.041.

Agoramoorthy, Govindasamy, Hsu, Minna J., 2021. How the coronavirus lockdown impacts the impoverished in India. J. Racial Ethnic Health Disparities 8, 1–6. http://dx.doi.org/10.1007/s40615-020-00905-5.

Delhing, J., Zierenberg, J., Spitzner, F.P., Wibral, M., Neto, J.P., Wilczek, M., Priesemann, V., 2020. Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. Science http://dx.doi.org/10.1126/science.abb9789.