Sen's Innovative Method for Trend Analysis of Epidemic: A Case Study of Covid-19 Pandemic in India

Saha Dauji[1,2]

Abstract

Analysis of trend of epidemiological data helps to appreciate the progression of an epidemic and to develop monitoring and control strategies by the government agencies. Sen's Innovative Method suggests a graphical analysis, which can overcome many limitations of data such as short length, non-Gaussian nature, skewness or serial correlation. In this article, this method is applied for the first time on epidemiological data. For the case study, Covid-19 or SARS-CoV-2 data from India were employed. The results show that Sen's Innovative Method is capable of indicating the shift in epidemiological trend quite efficiently, before it is reflected in the time series or moving average plots. The graphical analysis worked particularly well in comparing the trends of monthly data. It is concluded that this method would be especially suitable for monitoring the epidemiological trend by breaking up the data into smaller segments, as was illustrated in the study.

Keywords Trend analysis · Covid-19 · SARS-CoV-2 · Novel Corona Virus · Sen's innovative method · Epidemiological trend

Introduction

Covid-19 or SARS-CoV-2, which is the latest in the family of corona viruses affecting humans, is similar to pneumonia and may have symptoms such as fever, breathing difficulty and lung infection. This virus originated reportedly from Wuhan province in China and spread over the entire world at a rapid pace. World Health Organization (WHO) declared 'public health emergency' on 30 January 2020, and named it as a pandemic on 11 March 2020 (Singh and Sharma 2020). In the initial days after the first case reported on 30 January 2020 (Singh and Sharma 2020), the cases reported in India were pretty few, and was not taken much seriously. Screening of passengers from China and other affected countries and quarantine was initiated. From middle of March 2020, the number of cases started picking up, as the secondary and tertiary infections increased in number and India faced a major health emergency. Government of India rose to the

occasion by calling a national lock-down from 25 March 2020 and initiated survey, monitoring, quarantine, and Covid-care facilities. The testing facilities in the beginning were few, reportedly 15 government laboratories (Singh and Sharma 2020)—but rapidly grew with private hospitals and laboratories being trained and equipped for Covid-19 tests. Later, the testing kits started getting produced indigenously and that made expanding the testing network more efficient and cost-effective. Detailed account of the initial development of Covid-19 in India can be found in Singh and Sharma (2020). Indian National Academy of Engineering highlighted the importance of participation of various research agencies in offering thoughts, analyses and technologies for addressing the Covid-19 scenario in India in a Special Issue (Mishra 2020).

As on 28 November 2020, using the data from the website: www.worldometers.info/coronavirus, the tally for the world was 62,275,649 infections (0.51% of population) and 1,454,203 deaths (2.34% of infections) whereas for India the tally stood at 9,309,871 infections (0.67% of population) and 135,752 deaths (1.46% of infections). India stood second only to USA in absolute numbers, which had 13,479,589 infections (4.06% of population) and 271,181 deaths (2.01%). The report of lower rate of infection could be due to the lower number of tests per million

✉ Saha Dauji
  acad.dauji@gmail.com; dauji_saha@yahoo.com

1 Nuclear Recycle Board, Bhabha Atomic Research Center, Mumbai 400094, India

2 Homi Bhabha National Institute, Mumbai 400094, India

population: 99,760 for India which is around one-sixth of USA (570,952). The death rate of India, lower by 0.55% from the developed country like USA, with advanced health care facilities is an interesting observation. This could be attributed to the general immunity of the Indian population to influenza type viruses, but it is not well explained in the literature. However, India is a vast country, and has a huge population of 1,385,530,551, which is second only to China (1,439,323,776). Hence for India, even a low percent infection, followed by low percent deaths would indicate a large threat to the lives of Indians. This makes the studies conducted for understanding and monitoring the Covid-19 scenario in India even more important.

There have been numerous studies on the Covid-19 spread this year, in which researchers have discussed epidemic models (Sahoo and Sapra 2020; Samui et al. 2020; Mahajan and Kaushal 2020; Mandal et al. 2020; Rafiq et al. 2020), *SIPHERD* (Susceptible-Infected or Symptomatic-Purely Asymptomatic-Hospitalized or Quarantined-Exposed-Recovered-Deceased) model (Mahajan et al. 2020), *SIR* (Susceptible-Infected-Recovered) model (Anand et al. 2020), *SUC* (Susceptible-Unidentified Infected-Confirmed) model (Lee et al. 2020), *SEIR* (Susceptible-Exposed-Infectious-Recovered) model (Pai et al. 2020), Gaussian mixture model (Singhal et al. 2020), growth models (power model: Asad et al. 2020, Verma et al. 2020; regression based model: Bhardwaj 2020), delay model (Contreras et al. 2020), numerical simulation (Diwan et al. 2020), and trend analysis of the infection (Gupta and Pal 2020; Singh and Sharma 2020). The different studies considered some factors directly whereas the other factors were based on certain assumptions. Researchers have tried to estimate the effectiveness of quarantine protocols (Dziugys 2020). Declining trend of Covid-19 in Italy and projection for the world was also reported (Martelloni and Martelloni 2020). Some researchers attempted to model the initial spread of Covid-19 in India (Ghosh 2020; Sarkar et al. 2020; Gupta and Pal 2020), some tried to evaluate the inter-state transmission (Ghosh et al. 2020), some examined the effect of meteorological parameters on Covid-19 spread in India (Goswami et al. 2020; Gupta et al. 2020; Kumar and Kumar 2020; Kumar 2020), lockdown effect (Sardar et al. 2020) or unlocking effect (Vaishnav and Vajpai 2020), and others predicted the decline timeline for Covid-19 in India (Bhattacharjee et al. 2020). In India, the mobile App: *Arogya Setu*, developed for smart phones as well as feature phones, has been widely publicized by the government and has been one of the monitoring mechanisms (Jhunjhunwala 2020). Data-driven techniques have also found some applications in modeling the Covid-19 spread: temporal dynamics (Ranjan 2020); real-time forecast and risk assessment (Chakraborty and Ghosh 2020); *ARIMA* (Gupta and Pal 2020; Sharma and Nigam 2020); *ARIMA* and *NAR* (Khan and Gupta 2020); machine learning model

(Kavadi et al. 2020; Swapnarekha et al. 2020); optimization of lock-down policies (Khadilkar et al. 2020); transmission model and correlation analysis (Kumar et al. 2020); big data and geospatial techniques (Shah and Patel 2020); deep learning (Shastri et al. 2020; Swapnarekha et al. 2020). The aforementioned discussion brings out that mathematical, statistical and data-driven techniques have found application in modeling the Covid-19 transmission, examination of the trend, or future projections.

Now let us consider the different variables involved in the Covid-19 spread and infection scenario in India. The Covid-19 pandemic data for the reported daily cases and deaths would be affected by several factors. In a country like India, where the health care system is not well established or documented, many cases may not get reported. There are social and religious barriers which result in secrecy about Covid-19 symptoms, tests, test results, and deaths. The migration of people during lock-down periods in non-conventional (not permitted) methods adds another variable, which would be difficult to model. The possible length of incubation period, reportedly 2 to 14 days (Singh and Sharma 2020), the often encountered mildness of symptoms (ignored or not reported by many), and asymptomatic transmissions (up to 80%: Singh and Sharma 2020) make documentation and tracking of infection more challenging. The testing laboratory and infrastructure has evolved over the pandemic progression and so has the number of tests per million population—this directly affects the number of confirmed cases. Then the probability of false negative tests was reportedly high for the rapid antigen Covid-19 test, and this would also contribute to variations in the number of reported cases.

Furthermore, there have been stages of Unlock in India starting in June 2020 as Unlock-1 subsequently followed by Unlock-2, Unlock-3, etc., on monthly basis. The mobility and public amenities have been opened up in these stages which would have direct and indirect bearing on the spread and reporting of Covid-19 cases. Last but not the least, the ground reality about safety precautions and quarantine practices in India might vary largely from the official rules and regulations particularly owing to the mental set-up and attitude of the population and modeling this factor could perhaps be the most challenging part of the problem. The imperfect lockdown scenario was examined in the literature for modeling Covid-19 transmission in India (Nadim and Chattopadhyay 2020). It is evident that consideration of all the influencing factors towards modeling of Covid-19 spread in India would be very difficult.

In India, initially, the patients were either hospitalized or quarantined, depending upon the severity of symptoms and were added to the number of new cases. When the patients tested Covid-19 negative, from quarantine or hospital as the case may be, then it was reported as 'recovery'. However, due to the limitations of the infrastructure and explosion

of the number of infections, later the patients were 'home quarantined' when having mild symptoms, and how or when they were added to the 'recovery' group is unclear. Suffice to say that the 'recovery' data reported in India were having different criterion over different periods of time. And consequently, the 'active' number of cases would also be affected. All other derived variables from these affected data would, in turn, be biased in some way or the other. Though studying the active cases, new cases and deaths using trend analysis or correlation analysis could reveal the transmission or fatality dynamics of Covid-19 in India, these could not be taken up due to the possible bias in the number of active cases, varying over time. Therefore, in this study, the data sets which are the most basic, namely, reported daily new cases and daily deaths, have been only considered.

A review by Kotwal et al. (2020) revealed that due to the multiple evolving factors and related uncertainties, the short term projection of Covid-19 in India could be the preferred option as long term projections would involve possible higher inaccuracies. Consequently, for the purpose of monitoring the trend (India) and modifying the government and health care policies according to the current trend of Covid-19, it becomes imperative to employ a trend test, which would have good detection power for short length data. Trend analysis for Covid-19 pandemic in India has been predominantly reported with numbers plot and cumulative plots (natural and logarithmic) of daily cases and daily deaths (Gupta and Pal 2020), or numbers plot and cumulative plot of daily cases and daily deaths, tests performed and positive results, virus reproduction rate, case fatality rate, doubling period (Singh and Sharma 2020). For monitoring the pandemic scenario, rapid evaluation of the short-term trend would be essential. The epidemiological progression data would be serially correlated; as the number of new cases being reported over the transmission horizon (typically 14 days was suggested for this case) would be correlated to the present total infection—depending on the transmission rate. It is observed that application of statistical trend tests for examining Covid-19 pandemic in India is missing in readily available literature. This could be possibly due to the rapidly evolving pandemic scenario coupled with the limitations of the traditional trend tests such as Man-Kendall (Mann 1945; Kendall 1975) and Spearman's *Rho* (Sen, 1978), when applied for short length data or serially correlated data. Sen's innovative method proposed by Sen (2012) specifically overcame many restrictive assumptions of the aforementioned traditional methods: less power for small sample size, less power for data with serial correlation, non-normal probability distributions, skewness of data, and requirement of pre-whitening of data before application, among others. This method has been established by Sen (2012) with extensive simulation studies, and has been later applied for hydro-meteorological time series including:

rainfall and runoff (Sen 2012; Dabanli et al. 2016), temperature (Sonali and Kumar 2013; Sen 2014; Dabanli et al. 2016), relative humidity (Dabanli et al. 2016) and droughts (Malik et al. 2020). To the knowledge of the author, evaluation of epidemiological trend has not been previously attempted with this approach. However, considering the strengths of this approach, particularly for the short length data, as well as the complex and rapidly evolving nature of epidemic or pandemic progression, this could prove to be extremely handy for the government agencies in handling evolving epidemic scenarios. At this backdrop, the objective of this study is to examine whether this approach can be useful for epidemiological cases, and this would be performed using the Covid-19 pandemic data for the country: India.

Presently, with sufficiently long data from last 8–9 months, it is hereby assumed that all the variety of factors influencing the Covid-19 spread in India (discussed earlier) have been reflected in the chosen data, namely, the daily cases and daily deaths reported. Therefore, the trend analysis of the Covid-19 data from India could be informative and instructive for the government and the health care authorities for making informed decisions. Further, as the trend detection by Sen's innovative method is not limited by any assumption regarding the data length or distribution, it would be particularly suitable for evaluation of the variations in trend of the data. Of course, the reasons for the variation in trend, as discerned from the data, would be open to interpretation.

Research Significance

The article examines whether Sen's innovative method of trend analysis can be useful in evaluation of epidemiological trend, with a case study of Covid-19 pandemic outbreak in India. The advantages of Sen's method are many-fold (discussed earlier), and if found suitable, application of this method would help the government and health workers to monitor the pandemic progression and formulate monitoring and control policies in a more efficient manner. A method of presenting normalized scatter plot for depicting Sen's (2012) innovative trend over different time spans is proposed in this study, for ease of comparison and appreciation of the variations of trend over the respective time spans.

Data and Methodology

Data

The data for daily new cases and daily deaths have been extracted from the website: www.worldometers.info/coronavirus/country/india/ from time to time. Frequently during the entire duration, the total reported cases and the total
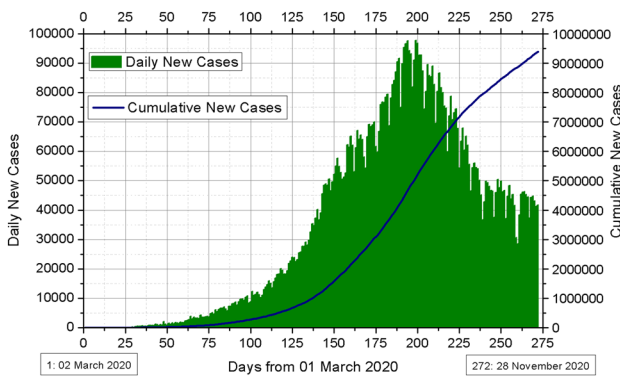
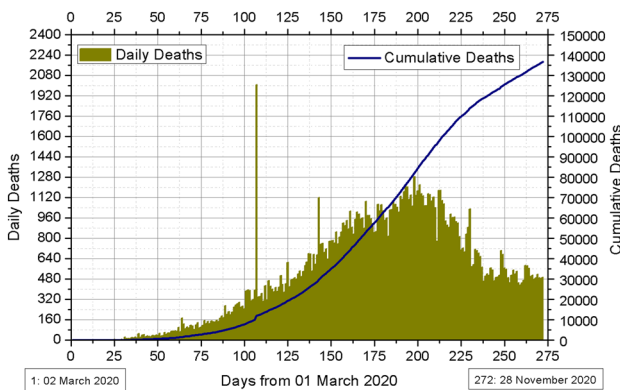**Fig. 1** New cases of Covid-19 in India: daily and cumulative plot

Fig. 2 Death due to Covid-19 in India: daily and cumulative plot

death have been corroborated with the data from the website: www.mohfw.gov.in, which is maintained by Ministry of Health and Family Welfare, Government of India. As recently as 28 November 2020, the variations between the two websites have been found to be 0.05% for total cases and 0.02% for total deaths—indicating that the data are accurate enough for a study of epidemiological trend. There is one entry (2006) on 16 June 2020 for daily death on the website: www.worldometers.info which has been explained as follows: there was additional entry of 1328 backlog deaths from Maharashtra on that day. Thus, this entry of 2006 has been reduced by that number to 678 for trend analysis presented in this study. However, for the purposes of reporting the total deaths and other statistics, this entry is included. The daily new cases and cumulative new cases are plotted in Fig. 1; daily deaths and cumulative deaths are plotted in Fig. 2. In these figures, the time series have been plotted with time starting from 01 March 2020. Therefore, entry for time 1 corresponds to that of 02 March 2020 and the last entry at time 272 corresponds to 28 November 2020.

From Figs. 1 and 2, it can be noted that the number of daily new cases as well as daily deaths due to Covid-19 in

India is presently on a declining mode (as on 28 November 2020), with the change happening after time: 200, i.e., 17 September 2020. The rate of decline reduced after around time 235, which would be 22 October 2020.

Sen's Innovative Trend Analysis Methodology

The methodology proposed by Sen (2012) for detection of trend in a time series depends on a graphical approach and interpretation of the scatter plots developed from the time series in the following manner. The time series is divided into two equal parts, each of which is arranged in ascending order. The first half is plotted along the abscissa (or x-axis) and the second along the ordinate (or y-axis) to obtain a scatter plot, from which interpretations about the trend would be possible. Care should be taken to keep the range same in both axes. A diagonal from the minima to the maxima (45° or 1:1 line) divides the plot into two triangles, upper and lower. The interpretation provided by Sen (2012) and further elaborated by Dabanli et al. (2016) is as follows:

(1) If the points fall in a straight line almost parallel to the diagonal (45° or 1:1 line) in the upper (lower) triangle, then the trend present in the data is increasing (decreasing).
(2) The distance between the diagonal (45° or 1:1 line) and the straight line of the scatter points indicates the strength of the trend, higher distance signifies stronger trend.
(3) In all cases where the points does not fall in a straight line parallel to the diagonal, the data exhibit a non-monotonic trend, which could be increasing (decreasing) depending on the location of the points with respect to the diagonal—in upper (lower) triangle on the diagram.

In traditional trend analysis, the slope (and intercept) of the trend line is generally estimated using least square regression or by Sen's non-parametric slope estimator (Sen 1968). In the present study employing Sen's Innovative Method, the slope (s) for the trend over different time durations have been evaluated using the equation proposed by Sen (2015), reproduced as Eq. 1:

$$s = \frac{2(\bar{y}_2 - \bar{y}_1)}{n} \tag{1}$$

where, $\bar{y}_1$ and $\bar{y}_2$ are the arithmetic averages of the first and the second halves of the dependent variable, y, sequence, and n is the total number of data.

The intercept (a) may be evaluated using Eq. 2, reproduced from Sen (2015):

$$a = \bar{y} - \frac{2(\bar{y}_2 - \bar{y}_1)}{n} \times \bar{t} \qquad (2)$$

where, $\bar{y}$ and $\bar{t}$ are the arithmetic average of dependent variable and time sequence, respectively.

A brief explanation for interpreting the scatter plots on the basis of which Sen's (2012) Innovative Trend analysis is interpreted, is reiterated here. The abscissa (or x-axis) signifies the values of the variable over the first half of the time series arranged in ascending order, whereas the ordinate (or y-axis) depicts the same for the second half of the time series under consideration. The ranges of the abscissa and ordinate are maintained same and this is a necessary condition for correct interpretation. The diagonal line connecting minima and maxima (1:1 line or 45° line) signifies the 'no trend' condition. The points falling in upper triangle signify increasing trend and those in the lower triangle indicate decreasing trend. The information about the monotonic or non-monotonic trend or the strength of the trend is interpreted as described earlier in the section. Readers are directed to the original articles (Sen 2012, 2015; Dabanli et al. 2016) for further details about construction and interpretation of the graphical method of Sen's Innovative Trend analysis.

Methodology

The various time periods considered for different analyses are: total duration (TD) from 02 March 2020 to 28 November 2020; two halves of TD; four quarters of TD; and individual months. To start with, the data statistics and the box plots for the various time durations have been examined. The moving averages are calculated for 2 days to 14 days and plotted. Subsequently, using Sen's Innovative method (Sen 2012), the trends of the number of daily new cases and number of daily deaths have been explored. This study is focused primarily on identifying the changes in the trend of the epidemic data (daily new cases / deaths) and this would be useful for dynamic formulation of management policies. As the trend is expected to vary over time, the slope would definitely be of interest for comparison. On the other hand, the intercept would keep changing over time, and would have been important if the objective was to identify a single trend and project it over some time period in future. As would be shown in the analysis, a single trend over entire time period fails to identify the short term variations, which would be critical for public epidemiological policies. Therefore, in this article, the slopes evaluated according to the Eq. 1 for different temporal durations have been reported, and as the intercepts are not pertinent for this study, they have not been reported.

In an attempt to succinctly present the variations in trends over different time scales, the scatter plot as suggested by Sen (2012) has been presented in non-dimensional form, by dividing all entries by the maximum value in that time span, to arrive at values between 0 and 1 for all the scatter points. Thereby, for different time scales, say for halves, for quarters or for monthly data, the trends of the new cases or deaths have been represented in a single plot. Normalization for presentation of Sen's Innovative trend over different time spans for comparing the variation of trend over time is proposed in this study. To the best knowledge of the author, such normalization has not been earlier reported in the literature in this context.

Results and Discussions

Statistical Features of the Data

The data statistics for daily cases are presented for the various time durations in Table 1, and on monthly basis in Table 2, respectively. Similarly, the daily death statistics are listed in Table 3 for the various durations, and in Table 4 for

Table 1 Data statistics for daily Covid-19 cases in India over different time durations (02 March 2020 to 28 November 2020)

Statistic	March–november 2020	First half	Second half	First quarter	Second quarter	Third quarter	Fourth quarter
Number	272	136	136	68	68	68	68
Mean	34,517.70	7132.39	61,903.01	875.46	13,389.32	67,440.24	56,365.79
Median	30,279	3556	59,795	537	11,132	65,383	49,882
Standard Deviation	30,801.47	8449.55	17,946.37	1030.58	7957.78	18,066.80	16,133.99
Coefficient of Variation	0.89	1.18	0.29	1.18	0.59	0.27	0.29
Maximum	97,859	32,682	97,859	3932	32,682	97,859	89,688
90 percentile	79,922	21,040	89,349	2408	25,637	92,917	80,858
10 percentile	147	17	41,506	7	4547	46,219	40,507
Minimum	0	0	28,555	0	3113	34,820	28,555

Table 2 Data statistics for daily Covid-19 cases in India on monthly basis: 02 march 2020 to 28 november 2020

Statistic	March	April	May	June	July	August	September	October	November
Number	30	30	31	30	31	31	30	31	28
Mean	46.47	1110.17	5024.06	13,172.77	35,847.16	64,093.06	87,410.93	60,406.90	43,217.79
Median	20	979	4630	12,199	34,820	64,142	88,196	59,893	44,689
Standard Deviation	56.88	490.26	1863.49	3683.87	11,537.77	8256.65	7331.90	12,271.33	5175.31
Coefficient of Variation	1.22	0.44	0.37	0.28	0.32	0.13	0.08	0.20	0.12
Maximum	227	2013	8782	20,131	57,704	79,457	97,859	81,693	50,465
90 percentile	124	1805	7300	18,282	50,525	76,665	96,763	75,479	47,821
10 percentile	1	528	2963	8849	22,721	52,783	77,854	46,498	37,350
Minimum	0	440	2394	7761	19,428	50,629	69,671	36,838	28,555

monthly basis data. It is mentioned here that for the statistics of Covid-19 deaths which included the high entry of 2006 (dated 16 June 2020) the standard deviation and coefficient of variation are reported for both 2006 and 678 (removing the 1328 backlog cases from Maharashtra). The box and whisker plots are depicted in Fig. 3 for the daily new cases and Fig. 4 for the daily deaths, for which entry for 16 June 2020 has been taken as 678. These box and whisker plots (Figs. 3 and Fig. 4) represent the data statistics over different time intervals in pictorial form for visual appreciation.

Moving average plots help to depict the variation in time series by smoothing out the local fluctuations, for the particular time period chosen for calculating the moving average. Therefore, inference about the trend in a time series can be attempted from such plots. The moving averages are plotted in Fig. 5 for the daily new cases and Fig. 6 for the daily deaths from 2-day to 14-day blocks. Peaks are predominant in the period before September 2020 (day 200: 17 September 2020 in Fig. 5) whereas valleys feature mostly after that. This leads to the conclusion that the trend of daily new cases started declining from that date. For Fig. 6, the entry for 16 June 2020 has been taken as 2006, and that explains the spike for daily death and the plateau in moving averages around that time. However, the other peaks and valleys in the plot indicate fluctuations in the numbers over short durations, typically less than 14 days, as the 14-day plot is quite smooth except around 16 June 2020 (which has been explained before). It is noted from Fig. 5 that the moving average of daily cases over 7-day and 14-day are quite smooth (shown with thick lines), when compared to the other durations. This periodicity of the data could not be explained with the available information, but could be related to the mean and maximum incubation period of the virus in population before appearance of symptoms. However, no such behavior has been observed for the daily death data (Fig. 6). The moving average plots depict a declining trend of daily new cases (Fig. 5) as well as the daily deaths (Fig. 6), typically after day 200, i.e., 17 September 2020. After day 235, i.e. 22 October 2020, the declining trend

reduced markedly for both new cases (Fig. 5) and daily deaths (Fig. 6). These shifts in trend, which are similar to observations in Figs. 1 and 2, would be compared with the results of trend analysis by Sen's Innovative Method in Sect. "Results and Discussions".

The percent increase in daily new cases, calculated over the cumulative cases till the previous day, as well as the percent increase in daily deaths, calculated over the cumulative deaths till the previous day, could be informative for monitoring the trends. These have been plotted in Fig. 7, where it can be observed that both the lines tend to get asymptotic with progression of time. In Fig. 7a, where the entire period has been plotted, the spike for daily deaths on time 107 (16 June 2020) is corresponding to the backlog clearance entry of 2006, as discussed earlier. After day 50 (20 April 2020), the rate of increase for both daily new cases and daily deaths are below 10% except for the aforementioned spike in daily deaths. Figure 7b depicts the same plot, but after day 200 (17 September 2020), after which the declining phase appeared to have started according to the earlier observations. During the period (from 200: 17 September 2020 to 240: 27 October 2020) the rate of increase for both daily new cases and daily deaths reduced from around 2% to around 0.6%, but subsequently the rates are fluctuating between 0.4% and 0.6% for both the variables. It would be difficult to get any projection after this period from the available inputs.

Graphical Evaluation of Trend of Covid-19 Progression in India by Sen's Innovative Method

Duration: Total (02 March to 28 November, 2020)

To start with, the trend of daily new cases and daily deaths for the entire duration (272 data) by Sen's Innovative Method is presented in Fig. 8a, b, respectively. As explained in Sect. "Sen's Innovative Trend Analysis Methodology", the interpretation of these plots would be that both daily new cases and daily deaths exhibit non-monotonic increasing trend when the entire duration is considered together.

Table 3 Data statistics for daily Covid-19 deaths in India over different time durations (02 March 2020 to 28 November 2020)

Statistic	March–november 2020	March–november 2020 #	First half	First half #	Second half	First quarter	Second quarter	Second quarter #	Third quarter	Fourth quarter
Number	272	272	136	136	136	68	68	68	68	68
Mean	502.66	497.78	183.30	173.54	822.02	29.19	337.41	317.88	941.71	702.34
Median	490	490	104	104	848	18	311	311	957	586
Standard deviation	397.99	387.48	237.14	182.60	237.13	36.15	252.58	153.58	171.21	234.32
Coefficient of variation	0.79	0.78	1.29	1.05	0.29	1.24	0.75	0.48	0.18	0.33
Maximum	2006	1283	2006	678	1283	175	2006	678	1283	1179
90perc	1064	1056	456	456	1132	72	511	511	1143	1078
10perc	4	4	0	0	496	0	128	128	687	478
Minimum	0	0	0	0	435	0	82	82	543	435

#These statistics are calculated using the value of 678 deaths in place of 2006 deaths for June 16, 2020, removing the backlog entries of 1328 from Maharashtra on the day

This analysis would be constrained in its inability to offer information regarding the evolution of the Covid-19 trend in India, and for that purpose, the data are repeatedly segmented into smaller time periods—in the following sub-sections.

Duration: Split in Halves (02 March to 28 November, 2020)

Next, the total period is divided into two parts (136 data in each) and the trends of daily new cases and daily deaths over these halves by Sen's Innovative Method are depicted respectively in Fig. 9a, b. Whereas the first half period shows non-monotonic increasing trend, the fall of points very steep for daily new cases (Fig. 9a) and steep for daily deaths (Fig. 9b), the second half (18 July 2020 to 28 November 2020) clearly indicates that the decreasing trend is monotonic for daily new cases (Fig. 9a) whereas for daily deaths (Fig. 9b), the decreasing trend is non-monotonic. The clarity in the variation of trend would be improved by considering smaller segments of time, as would be explained in subsequent sub-sections.

Duration: Split into Quarters (02 March to 28 November, 2020)

In a similar manner, the trends for the four quarters (68 data in each) are presented in Fig. 10a, b, respectively, for the daily new cases and daily deaths. The increasing trend is non-monotonic for daily new cases for the first two quarters, followed by a monotonic increasing trend in the third and then changing over to a non-monotonic decreasing trend for the fourth quarter. In case of daily deaths the first quarter is marked by non-monotonic increasing trend, followed by monotonic increasing trend over the next two quarters, and subsequently changing over to non-monotonic decreasing trend for the fourth quarter (22 September 2020 to 28 November 2020). Thus, disaggregation of the entire period into smaller segments paid dividends in narrowing down the period of change-over of the trend of Covid-19 pandemic progression in India. This aspect would be further examined with the monthly trends.

Duration: Monthly (March 2020 to November 2020)

The monthly trend analysis depicted in Fig. 11a, b, respectively, for new cases and deaths, helps to bring further clarity in the variation of trend over time for the case. The non-monotonic increasing trend for new cases (Fig. 11a) from March to August is marked by the scatter line becoming less steep and at the same time comes closer to the 1:1 diagonal progressively and in September depicts decreasing trend. In the month of October, the trend becomes non-monotonic decreasing, and falls to a weak monotonic decreasing trend

Table 4 Data Statistics for Daily Covid-19 Deaths in India on Monthly basis: 02 March 2020 to 28 November 2020

Statistic	March	April	May	June	June #	July	August	September	October	November
Number	30	30	31	30	30	31	31	30	31	28
Mean	1.17	37.30	137.23	400.07	355.80	617.45	931.74	1109.10	756.16	520.54
Median	0	36	132	354	354	610	944	1127	714	505
Standard deviation	1.90	17.62	44.97	312.32	96.18	152.37	82.03	91.40	194.29	62.42
Coefficient of variation	1.62	0.47	0.33	0.78	0.27	0.25	0.09	0.08	0.26	0.12
Max	8	75	269	2006	678	1120	1089	1283	1096	704
90perc	3	60	190	430	430	776	1019	1203	991	585
10perc	0	18	92	258	258	444	835	1026	509	451
Min	0	13	68	200	200	377	758	777	463	435

[#]These statistics are calculated using the value of 678 deaths in place of 2006 deaths for June 16, 2020, removing the backlog entries of 1328 from Maharashtra on the day

Fig. 3 Daily new cases of Covid-19 in India: box and whiskers plot
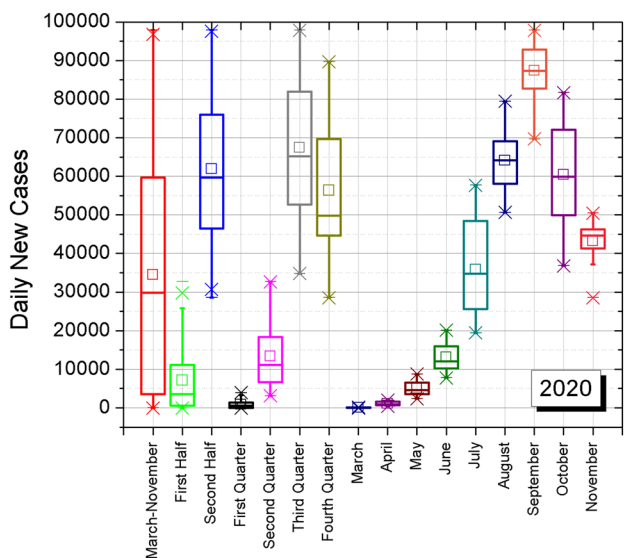
Fig. 4 Daily deaths due to Covid-19 in India: box and whiskers plot

in November. The variation of the trend of deaths reported daily is a little different over the months, when compared to the daily new cases. The trend is non-monotonic increasing for the months of March to May, and thereafter it becomes monotonic increasing trend for June and July. In August, the monotonic increasing trend becomes weaker, and changes over to weak decreasing trend in September. October is marked by strong non-monotonic decreasing trend of daily deaths, similar to the new cases, and becomes weak decreasing trend in November.

The daily new cases and daily deaths scatter plots were very steep for March and April (and this can be correlated with the high percent increase over previous days total for first 50 days: Fig. 7a) and gradually flattened out to moderately increasing trend towards August. Looking more closely at the figures would bring out the fact that the distance of
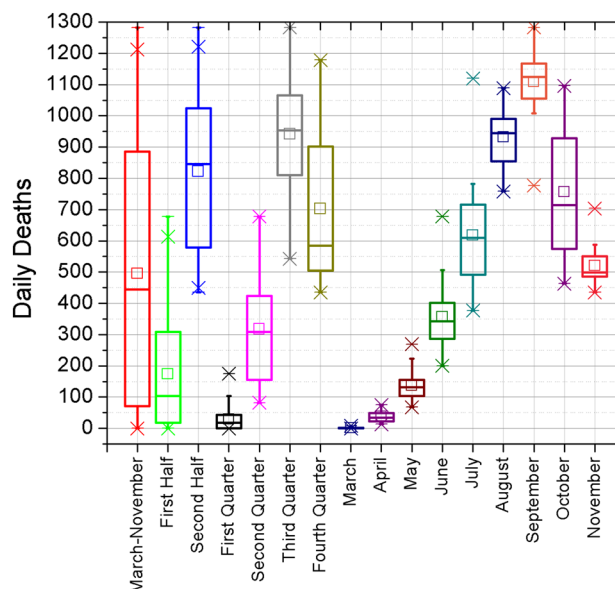
the trend line from the 1:1 line in August is much less of that of July, though it was positive (i.e., increasing trend). This fact signifies that the plateau had been reached with the trend line coming closer to the 1:1 line in August. The plateau in the trend, which was identified in the month of September from the time series (Figs. 1 and 2) and moving average plots (Figs. 5 and 6), could be identified 1 month earlier: in August, when using Sen's Innovative Method for trend analysis (Fig. 11a, b). This would be further corroborated with the numerical values of slope of trend line in Sect. "Progression of Slope of Trend Over Various Time Scales (March 2020 to November 2020)" (Fig. 12). September is marked with weak decreasing trend for both new cases and daily deaths, and for October, both the trends

Fig. 5 Daily new cases of Covid-19 in India: moving average plot
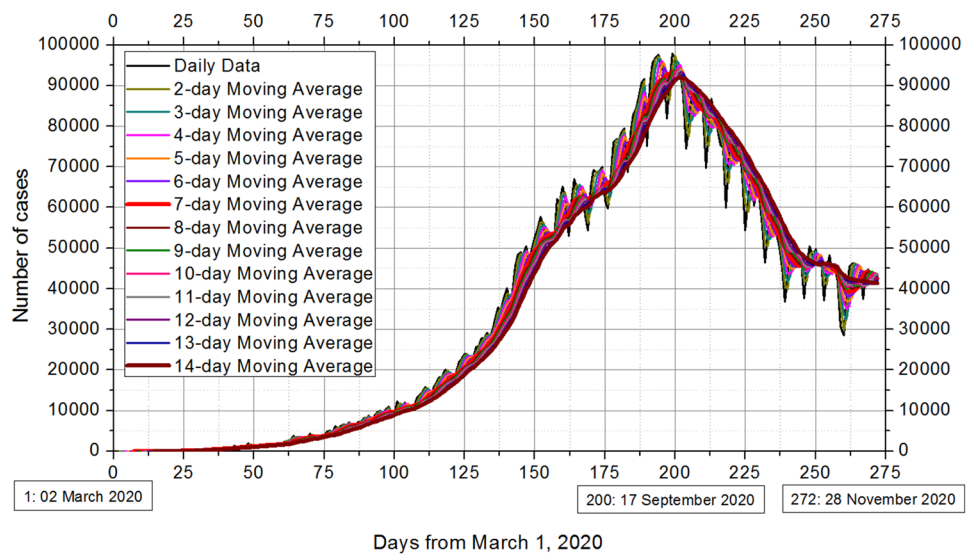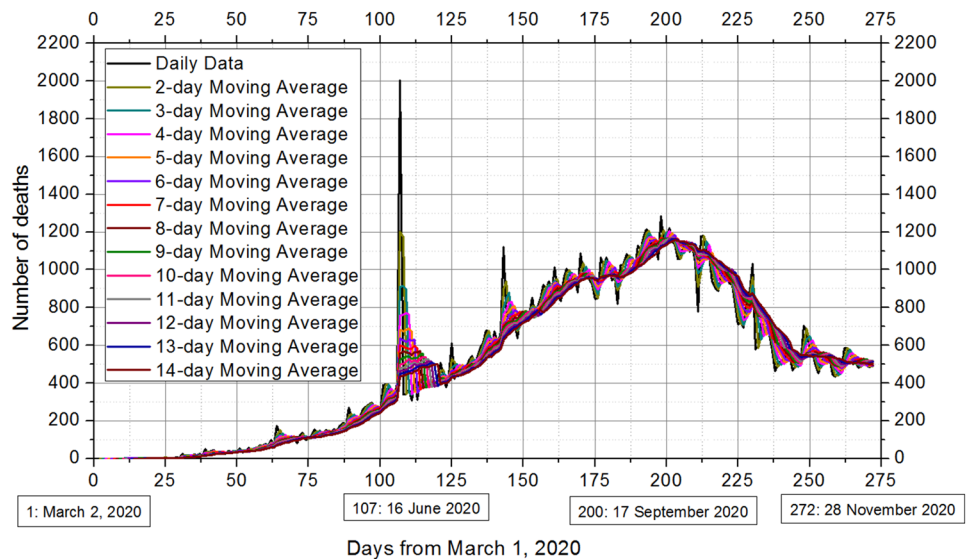

Fig. 6 Daily deaths due to Covid-19 in India: moving average plot

become strong and non-monotonic decreasing. The strength of the trends reduced from October to November, and they changed over to monotonic decreasing in November. Thus, analysis of trend of monthly trend of Covid-19 data using Sen's Innovative Method further facilitated the pinpointing of the change-over of the pandemic progression in India. These findings would be further corroborated with the slopes (Eq. 1) evaluated in the following sub-section.

Progression of Slope of Trend Over Various Time Scales (March 2020 to November 2020)

The slopes of the trend, evaluated according to the Eq. 1, are presented in Fig. 12a for daily new cases and Fig. 12b for daily deaths occurring due to Covid-19 in India. These figures are particularly informative about the shifting trend

of daily new cases as well as the daily deaths in India. The slope value becomes negative (decreasing trend) in second half, fourth quarter, and month of September when considering the half, quarter and monthly data. However, the interesting fact is that the plateau in slopes is identified in July, after which in August it reduces drastically to almost half for daily new cases (July: 1365; August: 678) or one-fourth for daily deaths (July: 15.21; August: 3.8) and eventually becomes negative in September. This observation is valid for both the variables considered in this study, though the numerical values are different. The negative slope decreases in the month of November for both new cases as well as daily deaths. Thus it is concluded that early identification of plateau and start of the decreasing trend can be possible when using Sen's Innovative Method for estimation of slope, along with the graphical trend analysis. Such
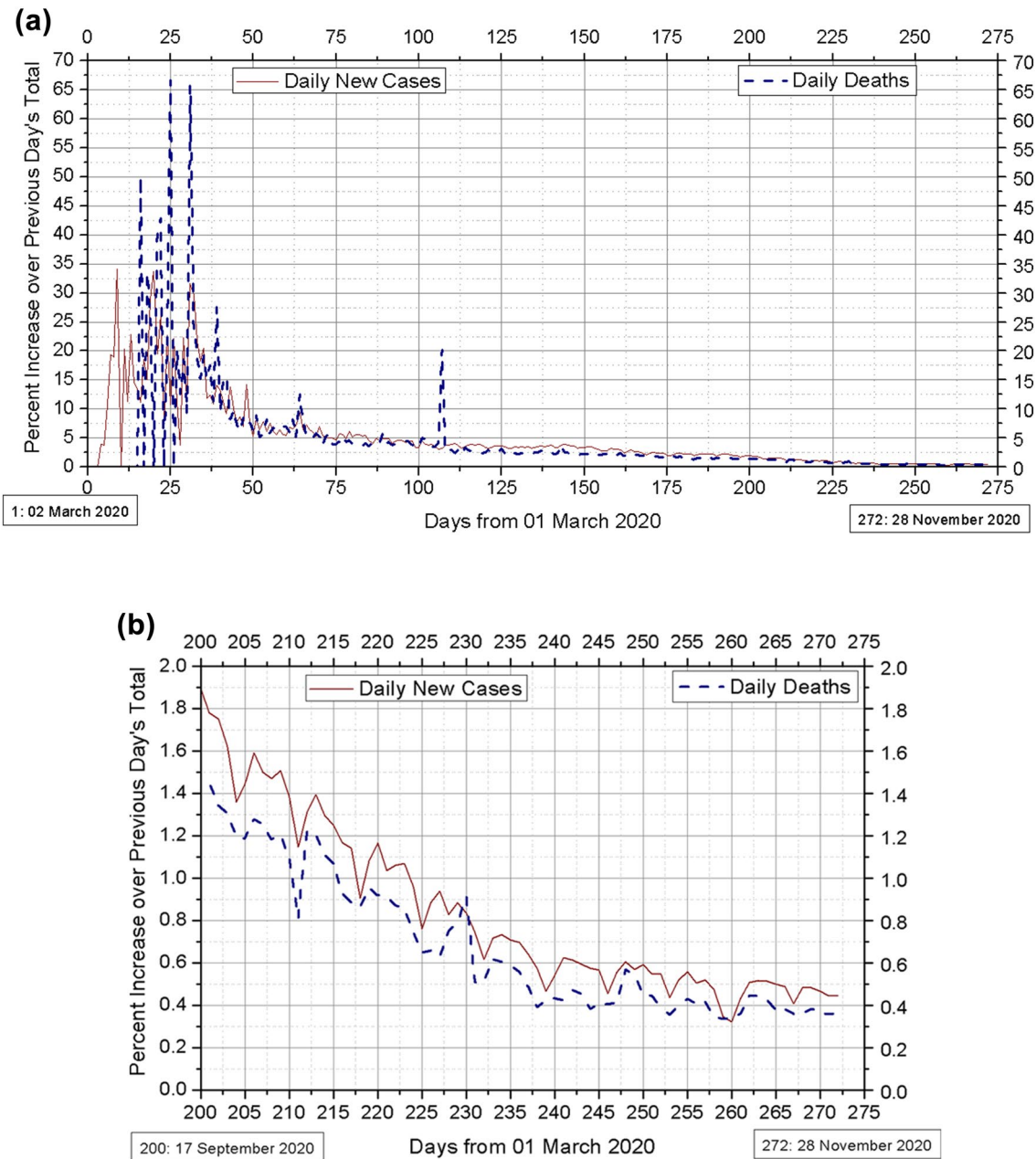
Fig. 7 Percent increase over previous day's total: **a** 02 March 2020 to 28 November 2020 **b** 17 September to 28 November 2020

information could be vital in formulating policies for control of pandemic and de-regulating the lock-down conditions as implemented for India during Covid-19 pandemic. The turn-around of new cases and deaths happened in the month of July, after Unlock-1 being implemented in June and thereby it is highlighted that this method could provide crucial inputs for government planning. However, a note of caution: this analysis presented here is performed by taking India as a whole. Similar analysis at state and district level would be advocated to critically monitor the pandemic progression in different states or districts.

Discussion

The trend analysis using Sen's Innovative Method was found to be quite effective in detecting the variations in the trend of Covid-19 pandemic in India. Particularly when smaller segments of the time series were separately analyzed, the variation of trend over time was clearly discernable. The slope estimator in Sen's Innovative Method could provide valuable information about the variable reaching the plateau, in advance of other reported methods of trend detection for Covid-19 in India. The plateau of the daily new
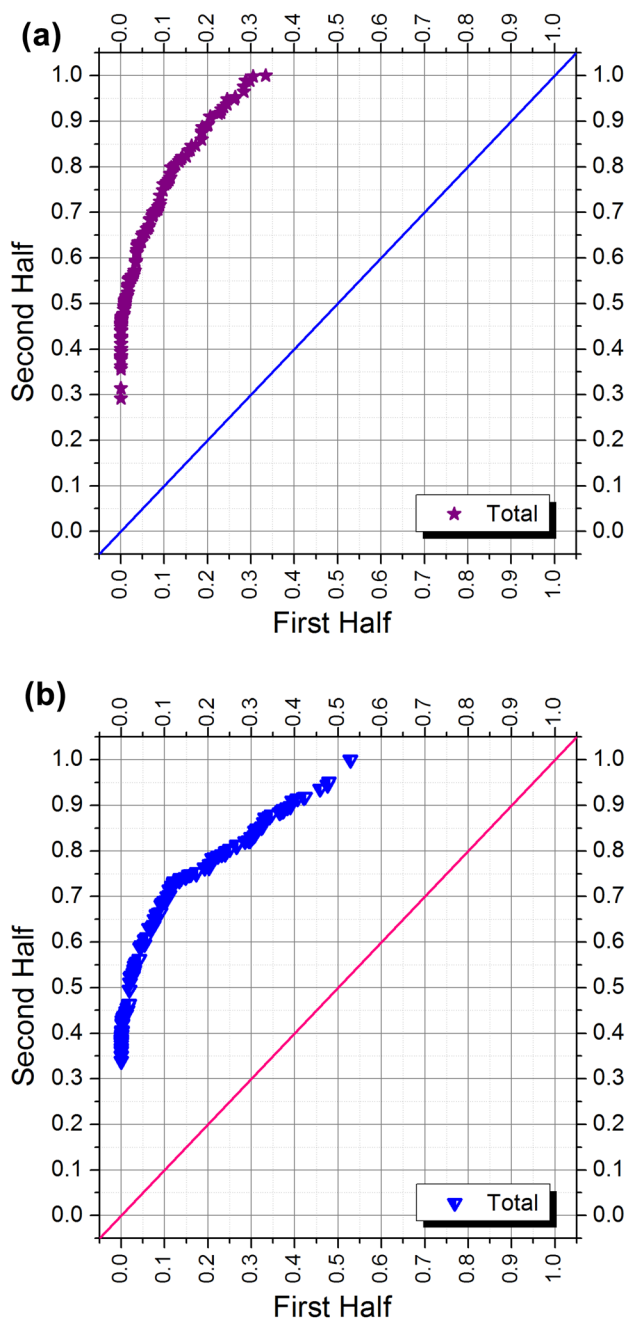
Fig. 8 Trend of Covid-19 in India using Sen's Innovative Method: 02 March to 28 November 2020 **a** Daily new cases **b** Daily deaths
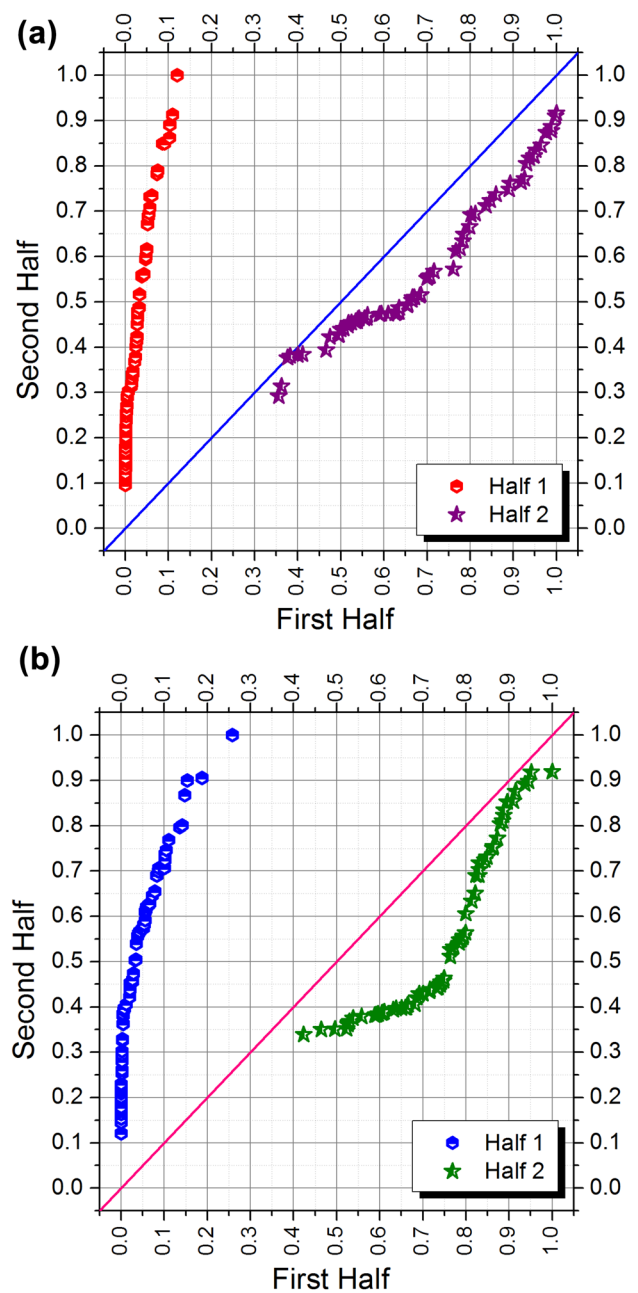

Fig. 9 Trend of Covid-19 in India using Sen's Innovative Method: 02 March to 28 November 2020 in two halves **a** Daily new cases; **b** Daily deaths

cases and the daily deaths could be identified in the month of July—2 months before it was reflected in the time series plot or moving average plot. Hence, this method would be extremely helpful for monitoring pandemic situation for India similar to Covid-19, and other epidemiological cases. When applied at the state or district data, this method could highlight the state-specific or district-specific trends and any adjustments to the national policies required at local levels would be easy to formulate and implement. In this study,

the analysis was limited to the two basic data, the daily new cases and the daily deaths reported for Covid-19 in India. However, it might be necessary to monitor other basic variables (say, recoveries or active cases) or derived variables (tests performed and positive results, virus reproduction rate, case fatality rate, doubling period, etc.) for better appreciation of epidemic situation. Sen's Innovative trend analysis can be performed for any other variable, basic or derived, and useful information would be obtained in a similar
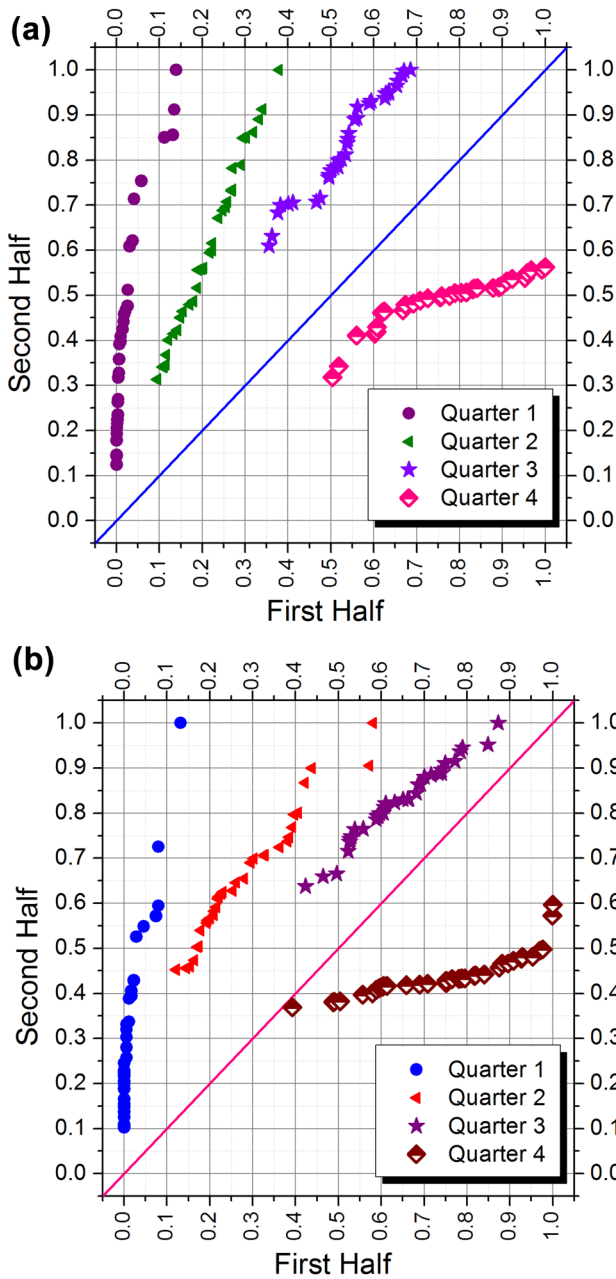
Fig. 10 Trend of Covid-19 in India using Sen's Innovative Method: 02 March to 28 November 2020 in four quarters **a** Daily new cases; **b** Daily deaths
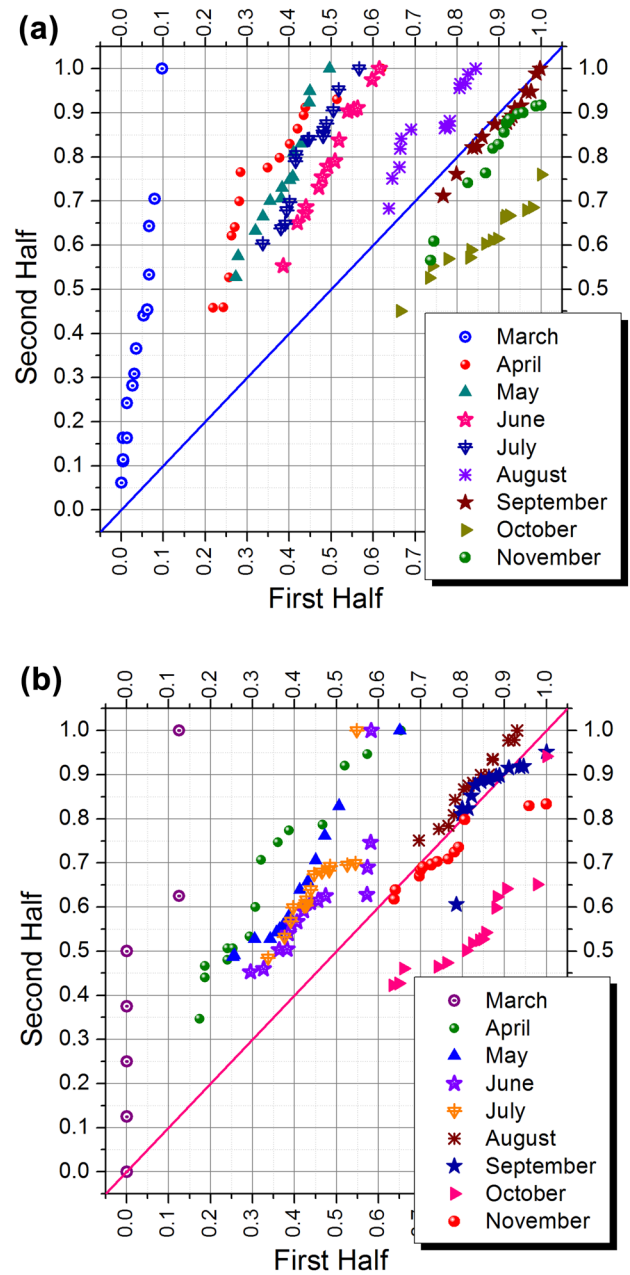

Fig. 11 Monthly trend of Covid-19 in India using Sen's Innovative Method: march to november 2020 **a** daily new cases; **b** daily deaths

manner, as discussed for the daily new cases and daily deaths in this study. It is reiterated here that for the rapidly evolving pandemic scenario segmenting of the entire time period was necessary to correctly appreciate the variation of the trend of transmission or fatality due to Covid-19 in India.

It should be noted that whereas in this study, the segmentation of the entire period was performed up to monthly data, further segmenting or different segmenting could also be performed to examine the effects of lockdown and other control measures, or Unlock and other de-regulating measures, introduction of vaccine to the populations, or other significant controlling factor pertaining to disease transmission, progression, or fatality. Therefore, analysis of trend of different variables of interest, coupled with a study of the estimated slope of trend line adopting Sen's Innovative Method—performed over judiciously chosen time segments would be suggested for monitoring and control of the epidemic scenarios in future. This small case study has indicated immense potential of this approach for understanding
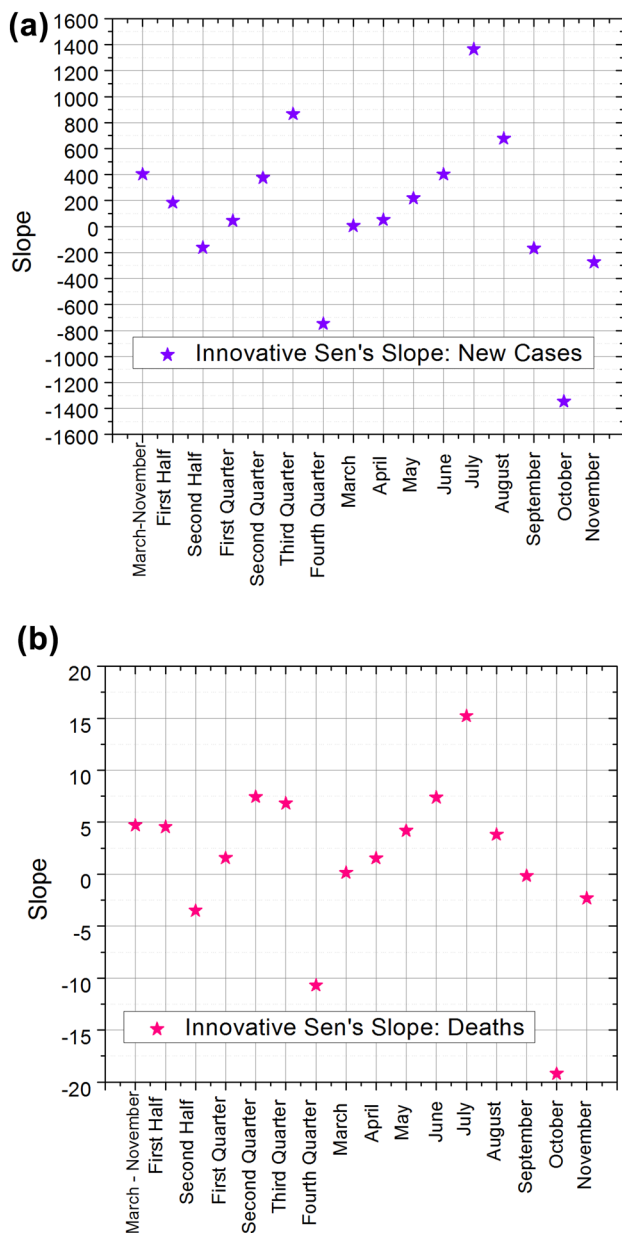
Fig. 12 Slope of trend of Covid-19 in India using Sen's Innovative Method over different temporal durations, 02 March to 28 November 2020 **a** daily new cases **b** daily deaths

the epidemic progression trend and this in turn would be valuable for formulation of government policies for control of spread of the epidemic.

Conclusions

In this article, Sen's Innovative method of trend analysis was applied for evaluating the epidemiological trend, with the case study of Covid-19 pandemic data (daily new cases and

daily deaths) from India. The following conclusions can be drawn from the limitations of the study:

- Sen's Innovative Method is effective in capturing the epidemiological trend, and its progression over time, as demonstrated for Covid-19 pandemic in India.
- The test provided more insights on the progression of pandemic in India when the entire data were disaggregated into smaller segments, up to monthly time interval.
- The evaluation of slope by Sen (2015) formulation, which is suggested for analysis of trend by Sen's Innovative Method, used along with the graphical trend detection by Sen (2012) could provide crucial inputs for formulating and modifying government policies of regulation or de-regulation in rapidly evolving pandemic scenarios.
- The turn-around for India as a whole could be identified in month of July 2020, using Sen's Innovative Method, for both new cases and daily deaths. This could be identified two months earlier than the time series plots or moving average plots.
- The normalized scatter plots for depicting the trend by Sen's (2012) innovative method, as proposed in this study, helped efficient representation of the variation of trend over different time intervals.

This study was limited to analysis of two variables, namely, daily new cases and daily deaths. For both these variables, the evolution of trend over time remained mostly similar. Other variables of interest such as recoveries or active cases, and derived variables such as tests performed and positive results, virus reproduction rate, case fatality rate, doubling period, etc., could provide further insight into the evolution of the epidemic. Effect of significant events such as specific regulating or de-regulating policy, or healthcare landmarks such as vaccination campaign can be examined by analysis of trend in this approach, before and after the event. A detailed analysis of the data from individual states along with that of entire country, would be more informative and the same is advocated for better insights in the Covid-19 trends of different states or regions in India. The information would provide important inputs for adjusting the national policies according to the state-specific trend and progression of the pandemic. Corroboration of the findings of the trend analysis with the other mathematical or statistical models could be interesting. Validation of Sen's Innovative Method for epidemiological trend analysis might be performed with data from other countries / epidemic in future for generalizing the conclusions drawn from this study.

and suggestions from the anonymous reviewers have contributed immensely towards improvement in the manuscript pertaining to technical content and presentation and the author is sincerely grateful for the same.

Availability of Data and Material The data used this study were extracted from the websites: www.worldometers.info/coronavirus/country/india/ and www.mohfw.gov.in. The listing of data can be shared by the author on request.

Code Availability The author declares that no specific code was developed for the study.

Declarations

Conflict of interest The author declares that there is no conflict of interest.

References

Anand N, Sabarinath A, Geetha S, Somanath S (2020) Predicting the Spread of COVID-19 Using SIR Model Augmented to Incorporate Quarantine and Testing. Trans Indian Natl Acad Eng 5:141–148. https://doi.org/10.1007/s41403-020-00151-5

Asad A, Srivastava S, Verma MK (2020) Evolution of COVID-19 Pandemic in India. Trans Indian Natl Acad Eng. https://doi.org/10.1007/s41403-020-00166-y

Bhardwaj R (2020) A Predictive Model for the Evolution of COVID-19. Trans Indian Natl Acad Eng 5:133–140. https://doi.org/10.1007/s41403-020-00130-w

Bhattacharjee A, Kumar M, Patel KK (2020) When COVID-19 will decline in India? Prediction by combination of recovery and case load rate. Clin Epidemiol Glob Health. https://doi.org/10.1016/j.cegh.2020.06.004

Chakraborty T, Ghosh I (2020) Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis. Chaos Solitons Fract 135:109850. https://doi.org/10.1016/j.chaos.2020.109850

Contreras S, Biron-Lattes JP, Villavicencio HA, Medina-Ortiz D, Llanovarced-Kawles N, Olivera-Nappa A (2020) Statistically-based methodology for revealing real contagion trends and correcting delay-induced errors in the assessment of COVID-19 pandemic. Chaos Solitons Fract 139:110087. https://doi.org/10.1016/j.chaos.2020.110087

Dabanli I, Sen Z, Yeleğen MO, Şişman E, Selek B, Güçlü YS (2016) Trend Assessment by the Innovative-Şen Method. Water Resour Manage. https://doi.org/10.1007/s11269-016-1478-4

Diwan SS, Ravichandran S, Govindarajan R, Narasimha R (2020) Understanding transmission dynamics of COVID-19-type infectionsby direct numerical simulations of cough/sneeze flows. Transactions of the Indian National Academy of Engineering 5:255–261. https://doi.org/10.1007/s41403-020-00106-w

Džiugys A, Bieliunas M, Skarbalius G, Misiulis E, Navakas R (2020) Simplified model of Covid-19 epidemic prognosis under quarantine and estimation of quarantine effectiveness. Chaos, Solitons Fractals 140:110162. https://doi.org/10.1016/j.chaos.2020.110162

Ghosh K, Sengupta N, Manna D, De SK (2020) Inter-state transmission potential and vulnerability of COVID-19 in India. Progress

in Disaster Science 7:100114. https://doi.org/10.1016/j.pdisas.2020.100114

Ghosh S (2020) Predictive model with analysis of the initial spread of COVID-19 in India. Int J Med Informatics 143:104262. https://doi.org/10.1016/j.ijmedinf.2020.104262

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

Gupta R, Pal SK (2020). Trend Analysis and Forecasting of COVID-19 outbreak in India. medRxiv preprint. https://doi.org/https://doi.org/10.1101/2020.03.26.20044511

Gupta S, Raghuwanshi GS, Chanda A (2020) Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020. Sci Total Environ 728:138860. https://doi.org/10.1016/j.scitotenv.2020.138860

Jhunjhunwala A (2020) Role of telecom network to manage COVID-19 in India: Aarogya Setu. Trans Indian Natl Acad Eng 5:157–161. https://doi.org/10.1007/s41403-020-00109-7

Kavadi DP, Patan R, Ramachandran M, Gandomi AH (2020) Partial derivative nonlinear global pandemic machine learning prediction of COVID 19. Chaos Solitons Fract 139:110056. https://doi.org/10.1016/j.chaos.2020.110056

Kendall MG (1975) Rank correlation methods. Oxford University Press, New York

Khalidkar H, Ganu T, Seetharam DP (2020) Optimising lockdown policies for epidemic control using reinforcement learning: an AI-driven control approach compatible with existing disease and network models. Transactions of the Indian National Academy of Engineering 5:129–132. https://doi.org/10.1007/s41403-020-00129-3

Khan FA, Gupta R (2020) ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. J Saf Sci Resilience 1:12–18. https://doi.org/10.1016/j.jnlssr.2020.06.007

Kotwal A, Yadav AK, Yadav J, Kotwal J, Khune S (2020) Predictive models of COVID-19 in India: a rapid review. Med J Armed Forces India. https://doi.org/10.1016/j.mjafi.2020.06.001

Kumar A, Rani P, Kumar R, Sharma V, Purohit SR (2020) Data-driven modelling and prediction of COVID-19 infection in India and correlation analysis of the virus transmission with socio-economic factors. Diabet Metab Syndr 14:1231–1240. https://doi.org/10.1016/j.dsx.2020.07.008

Kumar G, Kumar RR (2020) A correlation study between meteorological parameters and COVID-19 pandemic in Mumbai, India. Diabet Metab Syndr 14:1735–1742. https://doi.org/10.1016/j.dsx.2020.09.002

Kumar S (2020) Effect of meteorological parameters on spread of COVID-19 in India and air quality during lockdown. Sci Total Environ 745:141021. https://doi.org/10.1016/j.scitotenv.2020.141021

Lee C, Li Y, Kim J (2020) The susceptible-unidentified infected-confirmed (SUC) epidemic model for estimating unidentified infected population for COVID-19. Chaos Solitons Fract 139:110090. https://doi.org/10.1016/j.chaos.2020.110090

Mahajan A, Sivadas NA, Solanki R (2020) An epidemic model SIPHERD and its application for prediction of the spread of COVID-19 infection in India. Chaos Solitons Fract 140:110156. https://doi.org/10.1016/j.chaos.2020.110156

Mahajan P, Kaushal J (2020) Epidemic trend of COVID-19 transmission in India during lockdown-1 phase. J Community Health. https://doi.org/10.1007/s10900-020-00863-3

Malik A, Kumar A, Pham QB, Zhu S, Linh NTT, Tri DQ (2020) Identification of EDI trend using Mann-Kendall and Şen-innovative trend methods (Uttarakhand, India). Arab J Geosci 13:951. https://doi.org/10.1007/s12517-020-05926-2

Mandal M, Jana S, Nandi SK, Khatua A, Adak S, Kar TK (2020) A model based study on the dynamics of COVID-19: prediction and control. Chaos Solitons Fract 136:109889. https://doi.org/10.1016/j.chaos.2020.109889

Mann HB (1945) Nonparametric tests against trend. Econometrica 13(3):245–259

Martelloni G, Martelloni G (2020) Modelling the downhill of the Sars-Cov-2 in Italy and a universal forecast of the epidemic in the world. Chaos Solitons Fract 139:110064. https://doi.org/10.1016/j.chaos.2020.110064

Mishra S (2020) Foreword by President of INAE. Trans Indian Natl Acad Eng 5:89. https://doi.org/10.1007/s41403-020-00154-2

Nadim SS, Chattopadhyay J (2020) Occurrence of backward bifurcation and prediction of disease transmission with imperfect lockdown: a case study on COVID-19. Chaos Solitons Fract 140:110163. https://doi.org/10.1016/j.chaos.2020.110163

Pai C, Bhaskar A, Rawoot V (2020) Investigating the dynamics of COVID-19 pandemic in India under lockdown. Chaos Solitons Fract 138:109988. https://doi.org/10.1016/j.chaos.2020.109988

Rafiq D, Suhail SA, Bazaz MA (2020) Evaluation and prediction of COVID-19 in India: a case study of worst hit states. Chaos Solitons Fract 139:110014. https://doi.org/10.1016/j.chaos.2020.110014

Ranjan R (2020) Temporal dynamics of COVID-19 outbreak and future projections: a data-driven approach. Trans Indian Natl Acad Eng 5:109–115. https://doi.org/10.1007/s41403-020-00112-y

Sahoo BK, Sapra SK (2020) A data driven epidemic model to analyse the lockdown effect and predict the course of COVID-19 progress in India. Chaos Solitons Fract 139:110034. https://doi.org/10.1016/j.chaos.2020.110034

Samui P, Mondal J, Khajanchi S (2020) A mathematical model for COVID-19 transmission dynamics with a case study of India. Chaos Solitons Fract 140:110173. https://doi.org/10.1016/j.chaos.2020.110173

Sardar T, Nadim SS, Rana S, Chattopadhyay, (2020) Assessment of lockdown effect in some states and overall India: a predictive mathematical study on COVID-19 outbreak. Chaos Solitons Fract 139:110078. https://doi.org/10.1016/j.chaos.2020.110078

Sarkar K, Khajanchi S, Nieto JJ (2020) Modeling and forecasting the COVID-19 pandemic in India. Chaos Solitons Fract 139:110049. https://doi.org/10.1016/j.chaos.2020.110049

Sen PK (1968) Estimates of the regression coefficient based on Kendall's tau. J Am Stat Assoc 63:1379–1389

Sen Z (2012) Innovative trend analysis methodology. ASCE J Hydrol Eng 17(9):1042–1046. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000556

Sen Z (2014) Trend Identification Simulation and Application. ASCE Journal of Hydrologic Engineering. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000811

Sen Z (2015) Innovative trend significance test and applications. Theoret Appl Climatol. https://doi.org/10.1007/s00704-015-1681-x

Shah P, Patel CR (2020) Prevention is better than cure: an application of big data and geospatial technology in mitigating pandemic. Transactions of the Indian National Academy of Engineering 5:187–192. https://doi.org/10.1007/s41403-020-00120-y

Sharma VK, Nigam U (2020) Modeling and forecasting of COVID-19 growth curve in India. Trans Indian Natl Acad Eng. https://doi.org/10.1007/s41403-020-00165-z

Shastri S, Singh K, Kumar S, Kour P, Mansotra V (2020) Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. Chaos Solitons Fract 140:110227. https://doi.org/10.1016/j.chaos.2020.110227

Singh SK, Sharma SN (2020) Situational Analysis and Trend of COVID-19 Virus in India. J Commun Dis 50(2):38–45. https://doi.org/10.4321/0019.5138.202017

Singhal A, Singh P, Lall B, Joshi SD (2020) Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. Chaos Solitons Fract 138:110023. https://doi.org/10.1016/j.chaos.2020.110023

Swapnarekha H, Behera SK, Nayak J, Naik B (2020) Role of intelligent computing in COVID-19 prognosis: a state-of-the-art review. Chaos Solitons Fract 138:109947. https://doi.org/10.1016/j.chaos.2020.109947

Sonali P, Kumar DN (2013) Review of trend detection methods and their application to detect temperature changes in India. J Hydrol 476:212–227. https://doi.org/10.1016/j.jhydrol.2012.10.034

Vaishnav V, Vajpai J (2020) Assessment of impact of relaxation in lockdown and forecast of preparation for combating COVID-19 pandemic in India using Group Method of Data Handling. Chaos Solitons Fract 140:110191. https://doi.org/10.1016/j.chaos.2020.110191

Verma MK, Asad A, Chatterjee S (2020) COVID-19 pandemic: power law spread and flattening of the curve. Trans Indian Natl Acad Eng 5:103–108. https://doi.org/10.1007/s41403-020-00104-y

www.worldometers.info/coronavirus/country/india/. Accessed on 28 November 2020 for Covid19 data of India and world.

www.mohfw.gov.in. Accessed on 28 November 2020 for Covid19 data of India.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.