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.
COVID-19 second wave: District level study of concentration of confirmed cases and fatality in India

Dr. Debarshi Ghosh a, Apurba Sarkar b,*, Dr. Pradip Chouhan c

a Department of Geography, Dhuggari Girl’s College, PO. Dhuggari, Jalpaiguri, PIN-735210, WB, India
b Department of Geography, University of Gorur Banga, Mokdumpur, Malda, West Bengal, India PIN-732103
c Department of Geography, University of Gorur Banga, Malda, University of Gorur Banga, Mokdumpur, Malda, West Bengal, India PIN-732103

A R T I C L E   I N F O

Keyword: Pandemic Fatality Ratio Infection mortality ratio Spatial autocorrelation Location quotient Moran’s I statistics

A B S T R A C T

The paper aims to reveal the spatial pattern of the concentration of COVID-19 confirmed cases and the spread of the pandemic from the Case Fatality Ratio. The study has been accomplished with district-level data. The analysis of the spatial pattern decoding has been done considering the Global and Local Moran’s I statistics comprising the linear trend of spatial autocorrelation for the whole India. The timeframe has been divided considering the surge of the second wave in March, 2021 and the peak of the wave in May 2021. The spatial clustering technique presents both the concentration of confirmed cases using Location Quotient analysis and the pattern of spread of the infection-related fatality throughout the country. The high Location Quotient of the confirmed cases strongly clustered around the Mumbai-Pune region, Kerala-Karnataka region, Garhwal Himachal, NCT of Delhi and Ladakh-Kashmir-Himachal Pradesh region during the period of the study. In May, the concentration has randomly clustered around the middle part of India. The Case Fatality Ratio was high in Maharshtara, Madhya Pradesh, Punjab and Haryana at the surge of the second wave. During the peak (May), two significant clusters of high Case Fatality Ratio are observed in and around the Mumbai urban (Maharashtra) and NCT of Delhi (including Punjab-Haryana).

1. Introduction

COVID-19 pandemic has been impacting the life and economy across the globe since December 2019 and has caused major disruptions (Walker et al. 2020). The COVID-19 pandemic has resurfaced in India in the form of a hard-hitting second wave. The COVID-19 has brought a threatening challenge to Indian society and the economy (Sarkar and Chouhan, 2021). India’s devastating second wave of COVID-19 has overwhelmed its health system and the country (Ranjana, 2020, Ghosh et al., 2020). The second wave of COVID-19, caused by severe acute respiratory syndrome (SARS-CoV-2), has struck India severely, with a significant case fatality rate (Tomar and Gupta, 2020). The situation in India is more critical as it has a huge population, poor medical infrastructure and complex socio-economic structure, where self-isolation, social distancing and quality treatment are the key controlling factors to neutralise the impact of the disease (Kaliya-Perumal et al., 2020, Bhyuan, 2021).

The growth of towns and the consequent need for more supplies have damaged the delicate environment of India, where there are high levels of smog, fine dust, and water pollution. Sulfur dioxide (SO2), Nitrogen dioxide (NO2), and particulate matter (PM) contribute in part to the toxins causing environmental contamination (Sarkar and Chouhan, 2020, Huang and Brown, 2021, Bherwani et al., 2021). Many Indian urban communities, including Mumbai, Kolkata, and Pune, are at the risk of air contamination (Conibear et al., 2018). One out of eight (about 12.5%) deaths in 2017 in the country were attributable to high rates of respiratory disease, stroke, heart disease, diabetes, and lung cancer, all conditions for which a certain percentage of cases result from severe air pollution (Gurjar et al., 2016). Some relevant scientific literature highlights that exposure to air pollution may be relevant to virus infection spread, and more recent literature focuses on COVID-19 diffusion (Cheng et al., 2020, Report et al., 2020, Saha and Chouhan, 2021). On January 30 2020, the Director-General WHO declared that the outbreak of novel coronavirus (2019-nCoV) constitutes a Public Health Emergence of International Concern (PHEIC) as per the advice of the International Health Regulations (IHR) Emergency Committee (Black et al., 2020). In the first surge in 2020, COVID-19 has infected nearly 20 million people across the globe, with 90 countries in the community transmission stage (Bherwani et al., 2021). The daily reported confirmed cases started to rise from February 2021 in India (Sengupta et al., 2021). The mid of April 2021, registered sudden hike over thousands of daily death was observed around the country (Khanna, 2020). Multiple fac-

https://doi.org/10.1016/j.envc.2021.100221
Received 22 May 2021; Received in revised form 18 July 2021; Accepted 21 July 2021
2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
tors are involved in driving the second wave of COVID-19 in India, such as the complex interplay of mutant strains, violation of COVID appropriate behaviour, and government and public complacency on initiation of the vaccination drive (Kar et al., 2021). The situation turned into a bleak one when the country witnessed the daily deaths of over three thousand at the end of April 2021. On April 26, 2021, India saw the highest daily tally of new SARS-CoV-2 infections ever recorded globally, 360,960, taking its pandemic total to 16 million cases, second only to the USA with more than 200,000 deaths (Thiagarajan, 2021). As of May 18, India had reported more than 26.4 million confirmed cases and over 274,000 deaths from COVID-19 (Balasri et al., 2021). 29.27 million cases have been reported in India during the pandemic, with a case fatality rate of 1.24% (363,079 deaths) up to June 11, 2021 (Kar et al., 2021). At the beginning of the second wave, the country’s Case Fatality Ratio (CFR) has hovered around 1.35% to 1.40%. At the onset of the COVID-19 pandemic, India imposed the world’s strictest nationwide lockdown beginning from March 25, 2020 (The Lancet, 2021). But, the situation during the onset of the second wave was aweary in the whole country. WHO confirmed 15,510 new cases in India on March 1, 2021, and the peak had been formed with 414,188 confirmed cases on May 7 (Saha and Chouhan, 2021). The curve of new cases indicated a sharp rise from April onwards in India. Surprisingly, the present peak pandemic situation addresses a difference of 3,14,692 new confirmed cases than the previous peak (September 18, 2020) (Nishirya, 2010). It indicates the severity of the second outbreak in India. On April 17, 2021, the country peaked, considering confirmed cases that accounted for 10.6% of the country’s total population. On May 11, 2021, the WHO reported about 16,411 confirmed cases and 178 deaths per million (COVID-19: STATUS ACROSS STATES 2021). On the same day, the country’s CFR was 1.09%. The spread of the infection is more belligerent in the states of North-East India. From the beginning of the second wave, the outbreak was centred around the megacities of India, especially in NCT of Delhi, Mumbai and Bengaluru urban (The Lancet, 2021). On January 4, 2021, a total of 7,57% people of the country were infected by the coronavirus (Adviser et al., 2021). But, by May 11, the figure reached up to 16.67% (Thakur et al., 2012). At this stage, the main challenge is to organise the basic life saving treatment facilities at the grass-root level and especially assure the supply of oxygen and vaccines (Joannidis, 2021). In addition to this, the legislative assembly election in Assam, West Bengal, Tamil Nadu and Kerala with mass political gatherings and rallies has made the situation out of control. India is the largest democratic country globally; however, looking at the current COVID-19 situation, the state elections (Assam, Puducherry, Kerala, Tamil Nadu, and West Bengal) could have been postponed till normalization (Samarasekera, 2021). Many states did not go for the full lockdown this time and relied on night curfew to keep the local economy alive. As well the largest vaccine drive in the world was started in India in 1st week of April. At the same time, 6.06% of the total population was vaccinated (1st and 2nd). On May 11, roughly 13.3% people have got vaccinated (COVID-19: STATUS ACROSS STATES 2021). The increase is not so promising for such a country with a billion-plus population. The changing nature of the virus and countrywide high oxygen demand worsened (Timilisina et al., 2020). A rapid increase in the daily incidence of serious cases creates a shortage of medical instruments, oxygen, hospital beds and lifesaving drugs (Gupta et al., 2021). In a knee jerk reaction, the central government and its expert team started constructing the badly needed medical facilities to overcome the shortages. However, this frenetic activity of augmenting facilities comes in the middle of an ongoing and exponential rise in cases. In contrast, it should have started way ahead of these doldrums, say experts (Black et al., 2020). The country’s average CFR falls significantly down from 1.45% (in January 2021) to 1.09% (in the first week of May) (Grech and Cuschieri, 2020). Considering the second outbreak, the study incorporates a new angle to find out the district wise spatial pattern of the concentration of COVID-19 confirmed cases and Case Fatality Ratio (CFR) in India (Dhillon et al., 2020).

The paper aims to know the spatial clustering pattern of both the concentration of confirmed cases and CFR. Spatial Cluster analysis (SCA) has two general purposes. The first is data preprocessing, where clustering is used as an automated, unsupervised step for organising data, increasing the efficiency of search or query algorithms, and improving the system’s overall performance. The second is exploratory, where clustering is used for knowledge discovery and for motivating new hypotheses based on discovered patterns (Sengupta et al., 2021). In this study, the SCA technique well presented the pattern of spread of the infection in the second wave. The work will surely help the policymakers of the country-firstly to understand the spread of the second wave from the concentration of positive cases and its spatial clustering, secondly to identify the hot-spots of the concentration and thirdly to implement COVID-19 pandemic combat strategies like demarcation of containment zones, lockdown strategies, vaccine drive with paramount importance, blocking intra-state movements and most importantly thorough preparation at the medical level (district, block and village level). This study incorporates a completely new angle to reveal the hidden geospatial pattern of the second wave at the district level with statistical statistics.

2. Date source & measures

2.1. Understanding the concentration of confirmed cases

The purpose of the Location Quotient (LQ) technique is to yield a coefficient to understand how well industry can concentrate in a region (Mo et al., 2020, Mack and Jacobson, 1996). The LQ measures the ratio between the local and national share of productive activities of a particular industry in a region, usually using employment to represent productive activities (Tian, 2013). It is measured on a simple numerical scale. The value of LQ can be presented in several ways: a) LQ less than one indicates that industry is under-presented or the output is not sufficient to meet the local demand etc., b) LQ 1 means the study region’s share of the industry is identical to the reference region’s share or the employment is equal in the sector for the national and regional economy and c) LQ greater than 1 suggests that the study region has more than its share of an industry or the output is more than sufficient (Lu, 2000). The LQ can be stated as a ratio of ratios. The technique is frequently used specially in studying industrial concentration, agglomeration and in regional economics. It can also be considered as agglomeration index like Gini Index, Maurel-Sedillot (MS) index etc (Crawley et al., 2013). The theoretical foundation of LQ was constructed by Ellison and Glaeser (1997) and Guimaraes et al. (2009) (Ellison and Glaeser, 1997, Guimaraes et al., 2009). Further, Billings and Johnson (2012) examined the statistical accuracy of LQ (Billings and Johnson, 2012). In this study, which district-wise concentration of the confirmed cases of covid-19 has been conceptualised with LQ technique. We took the objective to know which part of the country, the second wave of the Covid-19 pandemic has concentrated intensely. In addition to this, the district-wise concentration of the confirmed cases can also be checked compared to the whole country. The formula of LQ has been reshaped for this study:

$$LQ_j = \frac{X_{ij}}{X_{ij}/X_{jk}}$$

(1)

Where, $X_{ij}$ represents confirmed cases in a region j, $X_i$ is the total population of an area j, $X_{ij}$ is total confirmed cases of the whole country, and $X_{jk}$ is the total population of the entire country.

This approach will also lead to identify the spatial pattern of the concentration of confirmed cases in the country. Thus, the mobility of the spatial concentration (confirmed cases) from one region to another has been presented with maps considering a time lag.

2.2. Measurement of Case Fatality Ratio (CFR)

The Case Fatality Ratio (CFR) is a measure of the severity of the condition as it corresponds to the proportion of the population with a cer-
tain condition who die from that condition during the reference period (Nishiura, 2010, Aguiar and Stollenwerk, 2020). But the measurement is crude in sense that it does not account for the changes in the demography of positive cases and deaths during the different stages of an epidemic (Vigod et al., 2012, Barouki et al., 2021). Considering the lacuna, the CFR can be used as a basis for estimating and monitoring the number of infected individuals in a population, which may be subsequently used to inform policy decisions relating to public health interventions and lockdown strategies (Russell et al., 2020, Roques et al., 2020). In epidemiology, the ratio stands for finding the proportion of the people who die from specific diseases among all the individuals diagnosed with the diseases over a while. As many cases of previous covid-19 were asymptomatic, generalised data on the true number of persons infected lack in India. Mortality rates, therefore, are calculated from confirmed cases, which overestimate the Case Fatality Ratio (CFR). The ratio can also be used to evaluate the effect of new treatments (Tian, 2013). It is conventionally expressed as a percentage and represents a measure of disease’s severity. The calculation comprises the proportion of deaths among all infected individuals. The CFR can be measured as the proportion of individuals diagnosed with a disease who die from that disease and is therefore a measure of severity among detected cases. In this study, we tried to capture the fatality risk among the confirmed covid-19 cases. The calculation technique is:

\[
\text{CFR} = \frac{\text{Total number of deaths from diseases}}{\text{Total number of confirmed cases of diseases}} \times 100
\]

The CFR clustering pattern has been cross checked with the dot density mapping considering the deaths on the selected dates. It tells us about the spread of the severity of the infection in the country.

2.3. Understanding the spatial pattern of LQ & CFR

The spatial pattern of the CFR has been studied using Global Moran-I statistics (spatial statistics) considering Moran’s I index, Z-score, P-value and Expected Index (Lee and Li, 2017, Chen, 2013). The Global Moran’s I determine the probability that the observed value could be found in the random distribution or not (Haque et al., 2020). The spatial distribution pattern of the Fatality Ratio has been scaled considering the contiguity scale of distribution which varies from dispersed to clustered pattern. The statistics conceptualises the spatial pattern based on contiguity edge only function. It only finds neighboring polygon features that share a common boundary. The spatial distance among the neighbour polygons has been measured using the Euclidean distance method. The Global Moran I summary indicates z-score value in the standardised normal.

Fig. 1. District-wise spatial variation and clustered pattern of confirmed cases; a) LQ of confirmed cases (01-03-2021), b) LQ of confirmed cases (10-05-2021), c) spatial cluster of the concentration of confirmed cases (01-03-2021) and d) spatial cluster of the concentration of confirmed cases (10-05-2021).
probability distribution and determines the likelihood of spatial pattern (dispersed, random and clustered) (Zou et al., 2014).

The Anselin Local Moran’s I (Zou et al., 2014) statistics along with the Local Indicators of Spatial Associations (LISA) cluster maps were prepared to assess the univariate local and global patterns of spatial dependence for the neighbourhood states considering the Fatality Ratio (Anselin, 1995, Scott and Janikas, 2010). It is useful in determining spatial outliers. The approach produces Local Moran’s I Index, z-score, pseudo-value and cluster-outlier type (COType) (Table 1). The analysis wants to check whether the apparent similarity (spatial clustering of high or low values) or dissimilarity (an outlier) is more pronounced than one would expect in a random distribution. The COType indicates statistically significant clusters and outliers for a 0.05 level of significance. The authors used the spatial weight matrix that yielded the highest Moran’s I values for understanding the spatial association. The matrix weights are a key component in measuring spatial dependence.

Here, the authors generated contiguity weights following the Queen’s contiguity method. The method considers the common boundary as well as the vertices of the neighbour polygons. The Queen’s criteria consider eight neighbours. The matrix weights have been generated carefully considering the topology of the spatial structure produced as 'Connectivity Graph' in the GIS environment. Finally, to identify the clusters of the districts (Fatality Ratio), the Local Moran cluster map or Local Indicator of Spatial Association (LISA) clustering has been adopted. This clustering approach is designed to reject the null hypothesis of spatial randomness favouring an alternative clustering. LISA approach functions in two ways: a) it brings out an assessment of significance for each location attributes and b) it reveals a proportional relationship between the local and global statistics. The LISA clustering has been done with permutations of 999 where the smallest possible pseudo-p-value is 0.001 and all other pseudo values will be even multiples of this value. The LISA clustering on FR in India produces a good and the rationale fit to compare the India’s district-level data with the world’s reality. Finally, Moran’s scatter plot helps to determine the linear fit of the regression slope exactly following the Moran’s I value (Anselin, 2003). The plot classified the nature of spatial auto-correlation into four categories that correspond with the LISA CoTypes. The four quadrants in the scatter plot connect between global and local spatial autocorrelation.

3. Results & discussion

Numerous production sectors have grown and in addition to the steel and metallurgical sectors, the textile and oil refining industries have also expanded. This economic growth has increased the number of jobs in cities and the populations of large urban areas, such as Delhi and Mumbai (Gopalan and Misra, 2020). These megacities have high
population concentration, and both the cities are over crowded with migrating populations from all over India. The population density of Mumbai Urban and NCT of Delhi are 19,652 per sq. km and 11,320 per sq. km. Respectively. Social distancing has considered a preventive approach worldwide which has ensured that Covid-19 can be contained. But, in the Dharavi slum of Mumbai, the population density is more than 277,136 per sq. km. In NCT of Delhi, Munirka (urban village) and Kirti Nagar are highly congested urban slums.

The farmer protest in November 2020 at the New Delhi border areas with approximately 150 to 300 thousand people made the situation highly vulnerable for the infection to spread. The prescribed ‘Social distancing’ was at stake considering the entire situation. The household density of the urban slum areas is one of the most vital causes of infection spread. The intra-urban mobility from slum to town central Business Districts (CBD) and other workplaces increases the chance of such infection spread. The international airports of NCT of Delhi, Mumbai, Bengaluru and Kolkata acted as the gateways to allow the dangerous mutating Covid-19 strains to spread across the main urban hubs of the country. The migrant labours borrow the new strains to the different corners of the country. In addition, the relaxation phase of domestic transportation (especially railways) from November 2020 up to almost mid of April made the situation favourable for rapid infection spread. A recent study revealed the composite analysis based on socio-economic-political parameters at the district level to construct a new Socio-Economic Vulnerability Index (Sarkar and Chouhan, 2021). The index indicated the high vulnerability of North-West Delhi, Central Delhi, Kolkata, Mumbai and Ghaziabad. In our study, the spatial auto-correlation technique for LQ and CFR indices also indicates the same hot spots during the second wave. Previously, many printed discussions and analyses focused on the socio-economic impact of the Covid-19 pandemic in India (Bashir et al., 2020). We also have much literature dealing with finding the cause of infection spread from clinical epidemiological (Laxminarayan et al., 2020, Dehning et al., 2020, Xu et al., 2020) and societal viewpoints (Sharifi and Khavari-Garmsir, 2020, Somani et al., 2020). In this paper, the infection spread phenomenon and its concentration have been identified with keeping a relation between geographical association (spatial auto-correlation), logic of Geographical Information system (GIS), and statistical foundation.

The LQ analysis identified the Covid-19 infection concentration district-wise (Fig. 1 a & b). The concentration has been checked based on confirmed cases on a particular day. On the first day of March 2021, the spatial pattern of high concentration (greater than 10) was seen, especially in Delhi, Bengaluru, and Mumbai. These urban centres were the pivot of infection at the beginning of March. The National Capi-
tal of Delhi indicated a large concentration gap considering the confirmed cases and even outmatched with Bengaluru Urban (second highest). The spatial scaling of the LQ has been modified to new categories: a) less than 1 (poor concentration), b) 1.0 to 5.0 (moderate concentration), c) 5.0 to 12.0 (High concentration), and d) greater than 12 (very high concentration). In the march, a high concentration of confirmed cases prevailed in South India, and Bengaluru urban was the hot spot. In the whole, middle, middle-east and middle-west part of the country was under moderate to poor concentration region except NCT Delhi as the central hot spot of North India. The whole eastern part, except some districts of Arunachal Pradesh, falls within the moderate concentration category. Surprisingly, some districts in Haryana, Himachal and
Table 2
Global Moran’s I summery of CFR

| Date          | Global Moran’s I Index | Z-score | p-value | Spatial pattern                                      |
|---------------|------------------------|---------|---------|-----------------------------------------------------|
| 01-03-2021(b) | 0.594024               | 26.5107 | 0.000   | Given the z-score of 26.5107 there is a less than 1% likelihood that is clustered pattern could be the result of random chance. |
| 10-05-2021(P) | 0.466944               | 20.7097 | 0.000   | Given the z-score of 20.7097 there is a less than 1% likelihood that is clustered pattern could be the result of random chance. |

Uttaranchal, which are adjacent to the NCT of Delhi, indicated high concentration of confirmed cases. On May 10, 2021, a very high concentration can be observed at Bengaluru urban, Lakshadweep and Mumbai metropolitan regions. The districts of the state of Telangana vary from poorly to moderately concentrated in comparison to the whole country. Districts of Nagpur and Pune of Maharashtra indicate the increasing concentration of the confirmed cases. In the North of India, the districts of Haryana and Himachal show the same increasing concentration (high to very high). In the whole eastern part, some isolated patches of moderate to high concentration can be observed viz. Darjeeling, North and South 24 Parganas in West Bengal and Kamrup metropolitan in Assam. Except the Chennai urban, the whole Tamil Nadu is showing less concentration.

The moderate to high concentration in the middle part is quite random in nature. The LISA cluster map (Fig. 1c) of LQ (01-03-2021) indicates a High-High (HH) concentration of confirmed cases at Mumbai-Puna region, Kerala-Karnataka region, Ghawhal Himachal, NCT of Delhi, Ladakh-Kashmir-Himachal Pradesh region at the beginning of the second wave in India. The pattern has been confirmed by the z-score distribution map (Fig. 3c&d), which checks the likelihood of the cluster pattern comes from random chance. The regions have statistically significant clusters of high values. The Low-Low (LL) concentration is somehow fragmented in Uttar Pradesh, Madhya Pradesh, Telengana, Sikkim, Assam, Meghalaya, Manipur and Nagaland. A statistically significant low concentration of the confirmed cases stretches along a linear belt form middle to east India. The probability map tells the high value of cluster patterns at 0.05 significance level could be the result of random chance (999 permutations). The low-value clusters (LL) can be presented significantly at a 0.001 significance level in the LISA map with greater accuracy in rejecting the null hypothesis (the spatial clustering pattern is a random one). The LH (Low-High outlier) cluster in Goa state indicates some high confirmed cases surrounding it. In May, the HH clusters indicate a new concentration at the North-Eastern part of Maharashtra (Nagpur district) and the districts of Chhattisgarh (especially Bametara district) (Fig. 1d). The high concentration of confirmed cases in (HH type) Leh (Ladakh), Gurugram (Haryana), Lahaul and Spiti (Himachal) districts can be observed during this peak time.

The infection fatality rate, the probability of dying of a person who is infected, is one of the most important features of the COVID-19 pandemic. The CFR is directly related to the expected total mortality burden of a country. Here, the asymptomatic confirmed cases are also being taken into consideration. Firstly, the CFR has been checked from Global Moran’s I statistics (Table 2). The CFR of March and May both produces statistically significant cluster patterns rejecting the null hypothesis (p-value 0.000). The CFR greater than 2.8% (very high) is observed at Mumbai-Pune-Nask and Punjab-Haryana region in March. North-East India did not show the severity of CFR in March except Kolkata urban. During May (peak time), the spatial pattern indicates the same concentration of very high CFR (greater than 2.045) surrounding the Mumbai urban and interestingly Goa has also come under this category. The districts surrounding the NCT of Delhi show the worst situation indeed. Although, the data reveals a slight decrease in the CFR than March. Fig. 2a indicates the spread of CFR to East India. Using local Moran I statistics, the LISA map (Fig. 2c) reveals the HH clusters (significant) in four parts of the country, i.e. Mumbai-Pune-Nask districts (Maharashtra), Bhopal-Indore-Sagar districts (Madhyaaprades), Delhi-Gurugam-Chandigarh districts (Punjab-Haryana-NCT of Delhi), Kolkata-North 24 Parganas-East Medinipur districts (West Bengal). The z-score distribution map (Fig. 3c) supports the clustering pattern strongly. The significance map (Fig. 3a) indicates that the Delhi-Gurugam-Chandigarh cluster rejects the null hypothesis of random clustering pattern at 0.001 significance level. In Eastern India, the LL pattern indicates a statistically significant cluster of low CFR in March. But, in May, the CFR spatial distribution (Fig. 2b) and LISA map (Fig. 2d) indicates only two clusters with statistically significant high values of CFR, i.e. Mumbai-Pune-Nask districts (Maharashtra) and Delhi-Gurugam-Chandigarh districts (Punjab-Haryana-NCT of Delhi). The CFR in other regions is statistically insignificant even in Kolkata, Chennai and Bengaluru urban districts. Now, the East Indian districts indicate a fragmented LL category of clustering. The statistical like-lihood of rejection of the null hypothesis is observed at a precision level of 0.01 at the Delhi-Gurugam-Chandigarh cluster, indicates a statistically reliable clustered pattern of high CFR (Fig. 3b&d). The dot-density maps (Fig. 4a & b) also confirmed the LISA spatial clustering pattern. The Global Moran I values (Moran Index) (Fig. 4c&d) for both the time frame indicates a linear expansion of the CFR values to the HH quadrant, and relatively more data weightage falls on the right half of the standard normal distribution curve.

4. Conclusion

In this study, the concentration of confirmed cases comprising the asymptomatic patients reveals the pattern of statistically significant clustering around the two megacities of India i.e. NCT of and Mumbai. Surprisingly, in eastern Kashmir and Leh-Ladakh, the LQ revealed statistically significant HH cluster in March. But, in May, it becomes a non-significant cluster. In Eastern India comprising the districts of Ganga and Brahmaputra valley in Bihar, West Bengal and Assam, Nagaland, Manipur and Meghalaya have a relatively low concentration of LQ. In South India, whole Kerala and districts of Karnataka surrounding Bengaluru urban show the severity of the concentration of confirmed cases. The NCT of Delhi and Mumbai remain the main hot spot of concentration till May. In this peak, the confirmed cases gradually spreads over the whole Maharashtra and the neighbouring state of Chhattisgarh starts affected rapidly. The CFR indicates wide dissimilarity in comparison to the LQ index in LISA maps. In March, the very high concentration of confirmed cases in Kerala and districts surrounding Bengaluru urban did not report high CFR. The CFR cluster bridged up the districts of Maharashtra with some districts of middle-North India and shows a continuous directional progression (N-E). CFR is insignificant in the districts of eastern Kashmir and Leh-Ladakh. The LQ and CFR pattern matches by chance only in case of Eastern India (LL category). In May, the HH clusters of CFR clearly indicates two pivots of the infection i.e. Mumbai urban-Pune-Nask-Kolhapur region (comprising Goa) and NCT of Delhi (comprising Punjab & Haryana). The districts of the far west (Rajasthan & Gujrat), middle India (Madhyaaprades, Chhattisgarh, Bihar & Uttar Pradesh), West Bengal, Sikkim and Tamil Nadu do not show any significant trend of clustering of CFR in this peak time of Covid-19
pandemic. This analysis decodes the spatial pattern of the concentration and spread of the infection from March to May (second wave). Finally, the pattern would give us some idea how to combat with the present situation and to get prepared for the predicted third wave.

Consent for publication

As the corresponding author, I give you the publisher’s consent on behalf of all the authors.

Ethical approval

On behalf of all the authors, we approve that we have gone through all the journal’s ethical standards and confirm that all the ethics were taken into consideration during this research.

Human and animals’ rights

This article does not contain any studies with human participants or animals performed by the authors.

Availability of data

The data are collected from the data repository of https://www.covid19india.org which is publicly available and could be accessed upon a request subject to non-commercial and academic interest only.

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.

References

Adviser, P.S., Ministry, H., Minister, U.H., Minister, U.H., General, D.C., India situation, 2021.

Aguiar, K., Stollenwerk, N., 2020. Condition-specific mortality risk can explain differences in COVID-19 case fatality ratios around the globe. Public Health 188, 18–20. doi:10.1016/j.puhe.2020.08.021.

Anselin, L., 1995. Local Indicators of Spatial Association—LISA. Geogr. Anal. 27, 93–115. doi:10.1111/j.0169-5884.1995.tb00389.x.

Anselin, L., 2003. An Introduction to Spatial Autocorrelation Analysis with GeoDa. Spatial Analysis Laboratory (SAL), pp. 1–20.

Balsari, S., Udawadia, Z., Shaikh, A., Ghafrar, A., Kataria, S., 2021. Contextualising evidence-based recommendations for the second wave of the COVID-19 pandemic in India. Lancet Infect. Dis. 21, 905–907. doi:10.1016/S1473-3099(21)00397-2.

Barouki, R., Kogevinas, M., Audouze, K., Beleosva, K., Bergman, A., Bimbbaum, L., Boekhold, S., Denys, S., Deseille, C., Dravkik, E., Frumkin, H., Garric, J., Destoumieux-Garon, D., Haines, A., Huss, A., Jensen, G., Karkates, S., Klasova, J., Koskela, M., Laden, F., Marano, F., Franziska Matthes-Wiener, E., Morris, G., Nowacki, J., Paloniemi, R., Pearce, N., Peters, A., Rekola, A., Sarigiannis, D., Šebková, K., Slama, R., Staehelin, B., Tonne, C., Vermeulen, R., Vines, P., 2021. The COVID-19 pandemic and global environmental change: emerging research needs. Environ. Int. 146. doi:10.1016/j.envint.2021.106722.

Bashir, M.F., Ma, B., Shahzad, L., 2020. A brief review of socio-economic and environmental impact of Covid-19. Air Quality. Atmos. Health 13, 1403–1409. doi:10.1007/s13361-020-00894-8.

Bherwani, H., Gautam, S., Gupta, A., 2021. Qualitative and quantitative analyses of impact of COVID-19 on sustainable development goals (SDGs) in Indian subcontinent with a focus on air quality. Int. J. Environ. Sci. Technol. 18, 1019–1028. doi:10.1007/s13762-020-03122-z.

Bherwani, H., Gautam, S., Gupta, A., 2021. Qualitative and quantitative analyses of impact of COVID-19 on sustainable development goals (SDGs) in Indian subcontinent with a focus on air quality. Int. J. Environ. Sci. Technol. 18, 1019–1028. doi:10.1007/s13762-020-03122-z.

Bhuian, A., 2021. India begins COVID-19 vaccination amid trial allegations. Lancet (London, England) 397, 284. doi:10.1016/S0140-6736(21)00145-8.

Billings, S.B., Johnson, E.R., 2012. The location quotient as an estimator of industrial concentration. Reg. Sci. Urban. Econ. 42, 642–647. doi:10.1016/j.regsciurbeco.2012.03.003.

Black, J.R.M., Bailey, C., Przeworska, J., Djeleta, K.K., Swanton, C., 2020. COVID-19: the case for health-care worker screening to prevent hospital transmission. Lancet North Am. Ed. 395, 1418–1420. doi:10.1016/S0140-6736(20)30917-X.

Chen, Y., 2013. New approaches for calculating Moran’s index of spatial auto-correlation. PLoS One 8. doi:10.1371/journal.pone.0065836.
Flasche, S., Eggo, R.M., Edmunds, W.J., Kucharski, A.J., 2020. Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. Euro-surveillance 25, 6–10. doi:10.2807/1560-7917.ES.2020.25.12.2000256.

Saha, Jay, et al., 2021. Spatial-temporal variations in community mobility during lockdown, unlock, and the second wave of COVID-19 in India: A data-based analysis using Google's Community Mobility Reports. Spatial and Spatio-temporal Epidemiology doi:10.1016/j.sste.2021.100442, In press.

Saha, J., Chouhan, P., 2021. Lockdown and unlock for the COVID-19 pandemic and associated residential mobility in India. Int. J. Infect. Dis. 104, 382-389. doi:10.1016/j.ijid.2020.11.187.

Samarasekera, U., 2021. India grapples with second wave of COVID-19. The Lancet Microbe, 2, e238. doi:10.1016/S2666-5247(21)00123-3.

Sarkar, A., Chouhan, P., 2020. Modeling spatial determinants of urban expansion of Siliguri a metropolitan city of India using logistic regression. Model. Earth Syst. Environ. 6, 2317–2331. doi:10.1007/s40808-020-00815-9.

Sarkar, A., Chouhan, P., 2021. COVID-19: District level vulnerability assessment in India. Clin. Epidemiol. Glob. Health 9, 204–215. doi:10.1016/jcej.2020.08.017.

Scott, L.M., Janikas, M.V., 2010. Handbook of applied spatial analysis. Handb. Appl. Spatial Anal. 27–41. doi:10.1007/978-3-642-03647-7.

Sengupta, P., Ganguli, B., SenRoy, S., Chatterjee, A., 2021. An analysis of COVID-19 clusters in India: Two case studies on Nizamuddin and Dharavi. BMC Public Health 21, 1–21. doi:10.1186/s12889-021-10491-8.

Sengupta, P., Ganguli, B., SenRoy, S., Chatterjee, A., 2021. An analysis of COVID-19 clusters in India: Two case studies on Nizamuddin and Dharavi. BMC Public Health 21, 1–21. doi:10.1186/s12889-021-10491-8.

Sharifi, A., Khavarian-Garmir, A.R., 2020. The COVID-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. Sci. Total Environ. 749, 1–3. doi:10.1016/j.scitotenv.2020.142391.

Somani, M., Srivastava, A.N., Gummadidavalli, S.K., Sharma, A., 2020. Indirect implications of COVID-19 towards sustainable environment: an investigation in Indian context. Bioresour. Technol. Rep. 11, 100491. doi:10.1016/j.biotech.2020.100491.

Thakur, P.K., Laha, C., Aggarwal, S.P., 2012. River bank erosion hazard study of river Ganga, upstream of Farakka barrage using remote sensing and GIS. Nat. Hazards 61, 967–987. doi:10.1007/s11069-011-9944-z.

The Lancet, 2021. India’s COVID-19 emergency. Lancet (London, England) 397, 1683. doi:10.1016/S0140-6736(21)01052-7.

Thiagarajan, K., 2021. Why is India having a covid-19 surge? BMJ n1124. doi:10.1136/bmj.n1124.

Tian, Z., 2013. Measuring agglomeration using the standardized location quotient with a bootstrap method. J. Reg. Anal. Policy 43, 186–197.

Timilsina, B., Adhikari, N., Kafle, S., Paudel, S., Poudel, S., Gautam, D., 2020. Addressing impact of COVID-19 post pandemic on farming and agricultural deeds. Asian J. Adv. Res. Rep. 11, 28–35. doi:10.9734/ajarr/2020/v11i430272.

Tomar, A., Gupta, N., 2020. Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. Sci. Total Environ. 728, 138762. doi:10.1016/j.scitotenv.2020.138762.

Vigod, S.N., Seeman, M.V., Ray, J.G., Anderson, G.M., Dennis, C.L., Grigoriadis, S., Gruneir, A., Kurdyak, P.A., Rochon, P.A., 2012. Temporal trends in general and age-specific fertility rates among women with schizophrenia (1996-2009): A population-based study in Ontario, Canada. Schizophr. Res. 139, 169–175. doi:10.1016/j.schres.2012.05.010.

Xu, B., Gutierrez, B., Mekaru, S., Sewalk, K., Goodwin, L, Loskill, A, Cohn, E.L., Hawen, Y., Hill, S.C., Cobo, M.M., Zarebski, A.E., Li, S., Wu, C.H., Hulland, E., Morgan, J.D., Wang, L., O’Brien, K., Scarpino, S.V.V., Brownstein, J.S., Pylun, O.G., Pigott, D.M., Kraemer, M.U.G., 2020. Epidemiological data from the COVID-19 outbreak, real-time case information. Sci. Data 7, 1–6. doi:10.1038/s41597-020-0448-0.

Zou, B., Peng, F., Wan, N., Mamady, K., Wilson, G.J., 2014. Spatial cluster detection of air pollution exposure inequities across the United States. PLoS One 9. doi:10.1371/journal.pone.0091917.