Journal of Urban and Regional Studies on Contemporary India 7(1): 25–38 (2020)
Article

The Center for Contemporary India Studies, Hiroshima University
http://home.hiroshima-u.ac.jp/hindas/index.html

Spatial Patterns of COVID-19 Infection in India

KATSUMATA Yutaro* and TSUKIMORI Yoshiki**

*Assistant Professor, The Center for Contemporary India Studies at Hiroshima University; Research Fellow, National Institutes for the Humanities (NIHU), Japan
**Under Graduate Student, Hiroshima University, Japan
E-mail: ykatsumata18@gmail.com*, b175729@hiroshima-u.ac.jp**

Abstract This study clarified the spatial patterns of the COVID-19 infection in India up to August 1, 2020, through Geographical Information System (GIS) mapping. The research period was divided into four phases based on when the number of new infections reached 10,000, 50,000, and 100,000 per week, and analysis was conducted by state and district level based on this classification of phases. The trend of COVID-19 infection at the state level showed certain spatial patterns for each phase. The trend in the number of infections at the district level indicated that spatial patterns of infection spread had changed, even within the same district.

Key words COVID-19, spatial pattern, GIS, India

I. Introduction

This study aims to clarify spatial patterns of COVID-19 infection in India through mapping using a Geographical Information System (GIS). This paper drafted the manuscript for reporting as a prompt report.

COVID-19 has spread rapidly worldwide since the first outbreak in Wuhan, China, in December 2019. According to the World Health Organization (WHO), the total number of infections worldwide as of August 1, 2020, was 17,396,943, and the number of deaths was 675,060.[1] COVID-19 has a pressing influence on many areas of life, including society, economy, and culture, and each country is placed in an unprecedented situation.

Epidemics such as COVID-19 have been a matter of continuous interest in geography, and their study has made a constant contribution to the field (Kobayashi, 2008). In particular, the usefulness of the geographical approach for the study of epidemics was reevaluated by the spread of GIS (Nakaya et al., 2004; Nakaya, 2008; Mihara, 2019). Such research fields are referred to as “medical geography,” “spatial epidemiology,” and “spatial health research.” These research fields are interdisciplinary. Not only are the effective making and maintenance of the epidemic map enabled, but various spatial analyses using geographical information data are also conducted using GIS. Furthermore, GIS is applied to modeling, simulation, and monitoring of infection and diffusion of the epidemic. Thus, GIS is said to be a required technique in visualizing and analyzing an epidemic.

Some geographical studies have used GIS to analyze the trend of COVID-19 infection. Mollalo et al. (2020) analyzed the incidence of COVID-19 infection in the US at the national level. Using 35 variables about socioeconomic, behavioral, environmental, topographic, and demographic factors, they investigated the influence on the distribution of COVID-19 infection by spatial modeling using GIS. Kuebart and Stabler (2020) clarified the social-spatial process of COVID-19 diffusion in Germany. In addition, not only the probative studies at the national level but also studies at a more detailed spatial scale were obtained. Cordes and Castro (2020) identified a cluster of infections and the contributing factors based on zip code–level data[2] in New York. Subsequently, they noted that socioeconomic status and ethnic differences caused differences in COVID-19 infection. In addition, Zhou et al. (2020) described the effectiveness of big data in the analysis of COVID-19 infection by using GIS based on Chinese examples.

For studies on India, Murugesan et al. (2020) examined the spatial distribution of total infections, recovered cases, and deaths by state on April 21, 2020.[3] Here, the distribution of infections based on estimates using the inverse distance weighted (IDW) method[4] was depicted through a GIS map. On the other hand, Gupta et al. (2020) analyzed regional factors of distribution for total COVID-19 infections on April 27,[5] based on natural conditions and population density. They conducted a statistical analysis using air temperature, rainfall, actual evapotranspiration, solar radiation, specific humidity, wind speed, topographic altitude, and population density by state. In addition, this study suggested that relatively hot and dry Indian states at low altitudes are more susceptible to COVID-19.

In this manner, studies have also been conducted on COVID-19 infection in India. However, these studies considered the situation at one point in time; therefore, studies that analyze sequential infection trends do not exist. Further, the spatial level of analysis remains at the state level, and analysis at a more detailed spatial level remains a research problem.

This study examined the infection trend of COVID-19 from January 30, 2020, when the first case of infection was confirmed in India, to August 1 in time sequence. The spatial level of analysis in this study adopted the district level in addition to the state level because the district is a subordinate administrative division of the state. The method used in this study is simple: epidemic maps of the COVID-19 infection were made using GIS, and the infection trend was analyzed. Therefore, similar to previous studies, spatial analysis using GIS is not conducted. However, it is necessary to understand the spatial patterns of the COVID-19 infection to conduct exact spatial analysis. Furthermore, this study can contribute to providing basic material on COVID-19 infection in India.

The remainder of this paper is organized as follows. In Section II, the data used in this study are explained. Next, we conduct an outline of India, and changes in the pattern of COVID-19 infection in India are clarified in Section III. In Section IV, the trend of the COVID-19 infection at the state level is considered throughout India. The analysis of infection trends at the district level in two states is shown in Section V. Finally, in Section VI, the results of this study are summarized.

II. Data

In this study, we use data from covid19india.org collected and provided through a website.[6] Although covid19india.org is not public, it collects data and updates based on news from states or official information as a volunteer group. The website covid19india.org conveys information and provides the data it collects in API document and CSV file format. We can download those data and use them freely.

The type of data provided largely falls into two categories (Table 1). The first is "raw data," which this organization collected. Information after January 30, 2020, when the first case of COVID-19 infection was confirmed in India, is posted and is constantly updated. As of August 7, 2020, the CSV files were divided into 12 sheets.

The second category is data added according to purposes based on raw data. There are eight in total,[7] and these are divided into three categories. The first category relates to the overall infection trend of India, and the "case time-series" corresponds. The new infections, total infections, new recovered cases, total recovered cases, new deaths, and total deaths per day in India are shown for the time sequence in this seat. The second category included the infection situation in each state and district. Five seats—"state-wise," "district-wise," "state-wise daily," "states," and "districts"—correspond. The two seats for

Table 1. Data from covid19india.org

Sheet name	Description
raw data1	Till Apr 19th
raw data2	Apr 20th to Apr 26th
raw data3	Apr 27th to May 9th
raw data4	May 10th to May 23rd
raw data5	May 24th to June 4th
raw data6	June 5th to June 19th
raw data7	June 20th to June 30th
raw data8	July 1st to July 7th
raw data9	July 8th to July 13th
raw data10	July 14th to July 17th
raw data11	July 18th to July 22nd
raw data12	July 23th to Aug 7th
case time series	Time series of confirmed, recovered and deceased cases in India
state wise	The latest state-wise situation
district wise	The latest district-wise situation
state wise daily	Statewise timeseries of confirmed, recovered and deceased numbers
states	Statewise timeseries of confirmed, recovered and deceased numbers in long format
districts	Districtwise timeseries of confirmed, recovered and deceased numbers in long format
statewise tested numbers data	Number of tests conducted in each state, ventilators and hospital bed information reported in state bulletins
tested numbers icmr data	Number of tests reported by ICMR

Note: ICMR stands for the "Indian Council of Medical Research."
Source: Based on https://api.covid19india.org/documentation/csv/ (accessed August 7, 2020).

"state-wise" and "district-wise" placed the latest information as of the time the data were obtained. The seat for "state-wise daily" shows the new infections, recovered cases, and deaths per day for each state in time sequence. Furthermore, in "states" and "districts," the total infections, recovered cases, and deaths are added up per day for each state and district in time sequence. The third category concerns the enforcement situation of inspection and information about the number of beds in hospitals and respirators. The seat for "state-wise-tested numbers data" and "tested numbers icmr data" comes under this category.

In this study, we mainly use the seats for "case time-series," "states," and "districts" for analysis. In addition, the analysis object is limited to the case of infection. These data were collected based on multiple sources of information. However, the data from covid19india.org are useful for geographical studies because we can analyze at the district and state levels in the time sequence. In Gupta et al. (2020), the data of covid19india.org were used. In this manner, these data are used for academic studies.

On the other hand, this study used Arc GIS10.6.1 on mapping data about the COVID-19 infection. In addition, regarding geographical information data about the borders of states and districts, we use data purchased from ML Infomap (in Delhi, India). These map data are based on the Census of India 2011 and include data about population, social attributes, and employment for each district. Therefore, it is advantageous to add these data.

However, we faced some problems when using these geographical information data. The first problem is the handling of the states that gained separate status after 2011. Specifically, Telangana separated from Andhra Pradesh in 2014, and Jammu and Kashmir state was divided into Jammu and Kashmir union territory and Ladakh union territory in 2019. It is considered an independent state from Telangana, and its border data are reorganized. On the other hand, Jammu and Kashmir was treated as it was in 2011 because it is difficult for this state to segregate pre-2019 border data to correspond to the 2019 separation of Jammu and Kashmir state into the union territories of Jammu and Kashmir and Ladakh. The data on the COVID-19 infection are added based on the state constitution of 2020. Therefore, we added up the number of infections for the Jammu and Kashmir and Ladakh union territories and assumed the data of Jammu and Kashmir state.

The second problem is how to handle those districts that changed names or gained separate status after 2011. The data of the COVID-19 infection at the district level are added based on the name of the district as of 2020, similar to state-level data. Therefore, it is impossible to combine geographical information data for 2011. In addition, some districts with different names are included, even if they are the same district in 2011 and 2020. In combination with geographical information data, this causes complications. In this study, we conduct our analysis based on the border and name of the district in 2011. This is because population data by district included in geographical information data are used in addition to this paper. Therefore, the COVID-19 infection data at the level of district for which the name or border is different from 2011 are aggregated and corrected as referred to on the website of each state, map data of GADM,[8] and Google Earth. However, some districts cross district borders and have separate status, so these tabulations and corrections were not made for all districts.[9]

III. Overview of COVID-19 Infection in India

1. Study area

First, we explain Indian states and main urban agglomerations. As of August 2020, India had 28 states and 9 union territories. As noted above, Jammu and Kashmir and Ladakh union territories are treated as one state in this study. That is why this paper focuses on 29 states and 7 union territories (Figure 1). We subsequently refer to the state, including union territories.

According to the Census of India 2011, India's population in 2011 was 1,210,854,977. In terms of states, Uttar

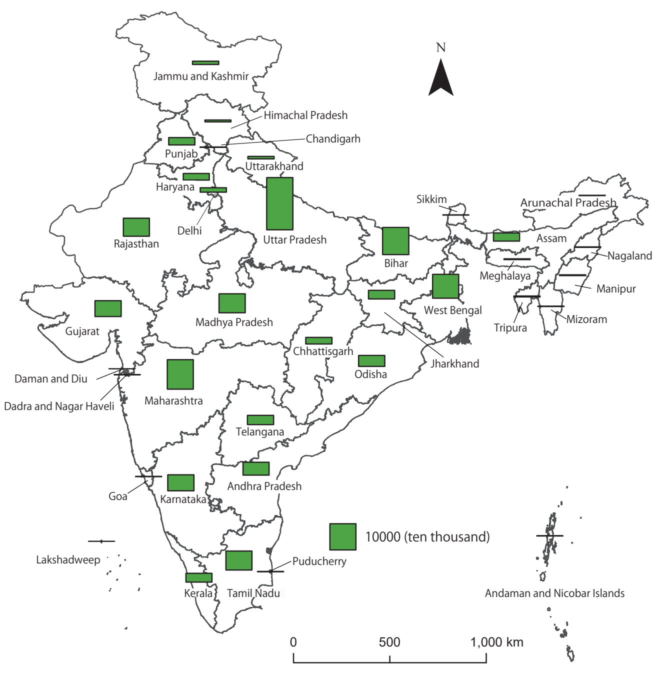

Figure 1. Population in India by state (2011)
Source: Based on the Census of India 2011.

Pradesh had the largest population (199,812,341), followed by Maharashtra (112,374,333) and Bihar (104,099,452). West Bengal came fourth, with a population of over 90 million. In addition, Madhya Pradesh (72,626,809), Tamil Nadu (72,147,030), Rajasthan (68,548,437), Karnataka 61,095,297, and Gujarat (60,439,692) all had populations of over 50 million.

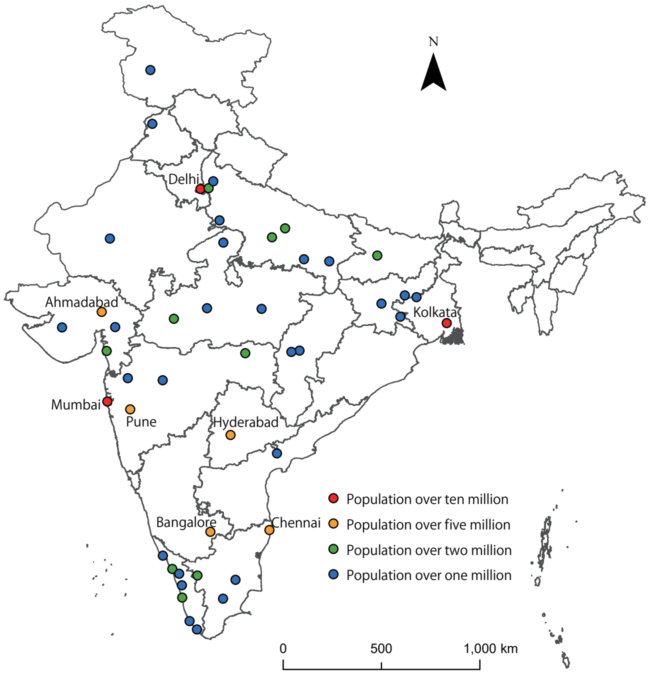

Figure 2. Distribution of urban agglomerations with populations over one million
Source: Based on the Census of India 2011.

The distribution of main urban agglomerations in India is shown in Figure 2. Urban agglomeration was set by the Census of India 2011, and the term as used in this study refers to main urban agglomerations with a population of one million and over. There were 46 main urban agglomerations in total. Of these, Mumbai (18,394,912), Delhi (16,349,831), and Kolkata (14,057,991) had populations of over ten million. Other urban agglomerations with more than five million inhabitants were Chennai (8,653,521), Bangalore (8,520,435), Hyderabad (7,677,018), Ahmedabad (6,357,693), and Pune (5,057,709). There were 10 urban agglomerations with more than two million people and 28 urban agglomerations with more than one million people.

2. Changes in COVID-19 infection

This section clarifies changes in the number of COVID-19 infections in India. New infections and total weekly cases of infections from January 30, 2020, to August 1 are shown in Figure 3.[10] Regarding the change in the number of infections, daily data are publicized but added up every week to assist us in finding characteristics in the time sequence more clearly.

In India, the first infection was confirmed on January 30, 2020, but further infections did not appear until March, other than two infections in the week of February 2. In March, the rate of infections showed a gradual increase. The number of new infections exceeded 10,000,

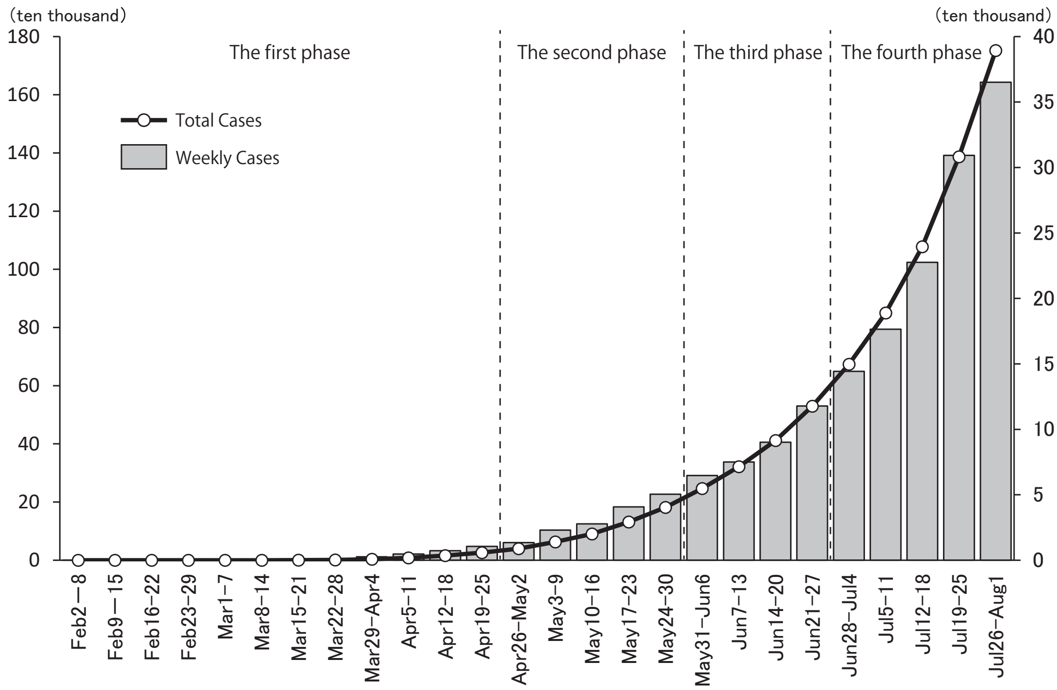

Figure 3. Trend in COVID-19 infection in India
Note: Data on January 30 included in the week of February 2.
Source: Based on data from covid19india.org.

and the number of total infections reached 25,000 in the week of April 19. Afterward, as the pace of infections increased, the number of new infections in the week of May 3 reached 23,039, an increase of nearly 10,000 from the 13,543 cases in the previous week. The number of new infections surpassed 50,000 cases in the week of May 24, and the number of total infections reached nearly 200,000. Subsequently, the number of new infections increased at a rate of more than 10,000 per week. Moreover, the number of new infections surpassed 100,000 cases in the week of July 21, and the number of total infections surpassed 500,000 cases. Afterward, an increase in the number of new infections became more pronounced, and more than 35,000 new infections were confirmed in the subsequent week. The total number of infections as of August 1 was 1,752,185, and the number of infections per 100,000 population was 144.7 cases.[11] In addition, the total number of deaths on the same day was 37,420, which is approximately 2.1% of the total number of infections.

Based on these trends, we divided the phases into four parts when the number of new infections reached 10,000, 50,000, and 100,000 in one week. Specifically, the first phase covered the period until the week of April 19 when the number of new infections reached 10,000 in one week. The second phase covered the week of May 24 when the number of new infections reached 50,000 in one week. The third phase covered the week of July 21, when the number of new infections reached 100,000 in one week, and the period after July 21 to August 1 was set as the fourth phase.

IV. Trends in COVID-19 Infection at the State Level

In this section, the trend for the number of COVID-19 infections at the state level is examined and divided into four phases. The data on the total number of COVID-19 infections by state include some data on infections in which the state was unknown. However, in comparison with the total number of infections for all of India, the capture rate is higher than 97% in all phases. Therefore, we judged that there was no problem with the use of state-level data. In addition, the number of new infections is similar, but only the total data by state for phase four has a slightly larger number than data for the whole of India.[12] As it is not a large numerical difference, this does not create a problem in this study. However, in the analysis in this section, we use data for the number of infections in which the state is known. Therefore, it is necessary to note that do not necessarily agree with the number shown in Figure 3 of the previous section when we mention the number of infections of India in this section.

1. The first phase

Figure 4-a shows the number of total infections until April 25, 2020, by state. The number of infections until this day was 26,283.

The greatest number of infections was confirmed in Maharashtra. The number of infections in this state was 7,628, and its share of all infections in India was approximately 29.0%. It was followed by Gujarat and Delhi, which had 3,071 (11.7%) and 2,625 (10.0%) cases, respectively. In addition, states accounting for more than 5% of the number of total infections included Rajasthan (2,083, 7.9%), Madhya Pradesh (1,945, 7.4%), Tamil Nadu (1,821, 6.9%), and Uttar Pradesh (1,793, 6.8%). These four states, combined with Maharashtra, Gujarat, and Delhi, account for approximately 80% of all infections in India. Thus, the number of COVID-19 infections in this phase tended to be concentrated in relatively populous states in western and northern India. On the other hand, no states in eastern and southern India were notably infected, except Tamil Nadu. Populous states such as Bihar (251, 1.0%), West Bengal (571, 2.2%), and Karnataka (500, 1.9%) did not have a large number of infections.

2. The second phase

Figure 4-b shows the number of total infections by state until May 30 divided into the number of new infections and that of non-new infections. The number of total infections as of this phase was 176,816, and there were 150,533 new infections. In addition, infections were confirmed in Nagaland, Sikkim, and Dadra and Nagar Haveli, where there were no cases of infection in the first phase; therefore, COVID-19 infection spread in all states except Daman and Diu and Lakshadweep.[13]

If we look at the number of new infections by state, Maharashtra had the highest number, as in the first phase. The number of new infections in this state was 57,540 (38.2%), and its share of all infections was higher than that in the previous phase. Maharashtra was followed by Tamil Nadu and Delhi. There were 19,363 cases in Tamil Nadu (12.9%) and 15,924 cases in Delhi (10.6%). Notably, the number of infections in Tamil Nadu has expanded significantly. The number of infections in this state was only 1,821 (6.9%) in the first phase, but it increased more than ten times in the second phase. The number of total infections in Tamil Nadu was 21,184 (12.0%), higher than that in Delhi (18,549, 10.5%). In addition, in Gujarat, which ranked highly for the number of infections in the first

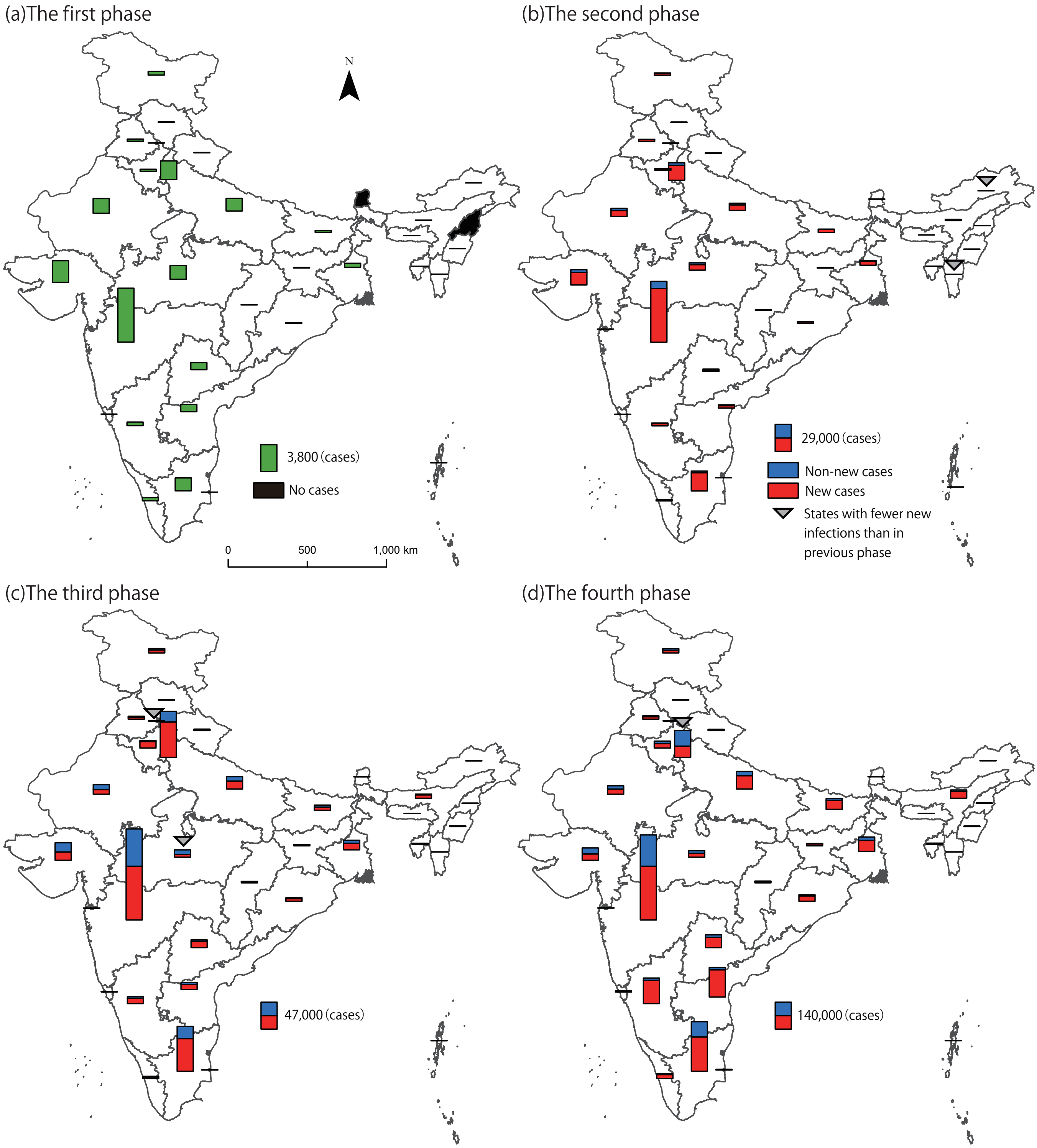

Figure 4. Number of infections by state in India
Source: Based on data from covid19india.org.

phase, the number of new infections was 13,285 (8.8%), which was the fourth-highest in India. The total number of infections (16,356, 9.3%) was still comparable to that of Delhi.

On the other hand, apart from these four states, no other state had a share of more than 5% of new and total infections. Further, the share of the total number of new infections for these four states reached 70.5%. Therefore, during this phase, there was a significant increase in the number of new infections, but they tended to be concentrated among the top-four states.

3. The third phase

Figure 4-c shows the number of total infections by state until June 27. The total number of infections was 521,564, and the number of new infections was 344,748. The num-

ber of new infections more than doubled compared to the previous phase.

Maharashtra remained the state with the largest number of new infections, with 93,965 cases. However, its share of the number of new infections was 27.3%, a decline of more than 10% from the second phase. Delhi and Tamil Nadu had the highest number of new infections after Maharashtra. There was no change in the composition of the top-three states in terms of the number of new infections compared to the second phase. The number of new infections was 61,639 in Delhi and 57,151 in Tamil Nadu; both states had over 50,000 cases. Delhi represented 17.9% of all new infections, an increase of 7.3% from the second phase, and Tamil Nadu represented 16.6%. Thus, the number of new infections in Delhi exceeded that in Tamil Nadu, which was the characteristic of the third phase.

Further, the number of new infections in Gujarat, which had the fourth-largest number of infections in the second phase, was confirmed as 14,417 (4.2%) in this phase. In the third phase, although Gujarat had the fourth-highest number of new infections in India, the difference between Delhi and Tamil Nadu significantly increased in comparison with the second phase. In addition, the increase in the number of new infections was confined to approximately 1,000, as the number in the second phase was 13,285 (8.8%). Gujarat was not hit by a large increase in the number of new infections, as were the three states mentioned above. Consequently, both Gujarat's share of the number of new infections and that of the total number of infections decreased to 5.9% from 9.3%.

Finally, we mention states other than these four states. The share of the number of new infections by the four highest states was 70.5% in the second phase, but it decreased to 65.9% in this phase. However, this depended on the decline of Gujarat's share; the share for the three-highest states was unchanged at 61.7%. On the other hand, the third-phase share of the number of new/total infections over 5% was also not observed other than in these four states. The number of new infections in Bihar, West Bengal, and Karnataka mentioned in the first phase was 5,415 (1.6%), 111,581 (3.4%), and 9,001 (2.3%), respectively. Although the number of new infections increased compared to the first and second phases, the number of infections remained low relative to the population scale.

4. The fourth phase

Figure 4-d shows the number of total infections until August 1 by state. The number of total infections until this day was 1,752,171, and the number of new infections in the fourth phase was 1,230,607. The number of new infections in the fourth phase increased approximately 3.5 times from the third phase.

If we look at the number of new infections by state, Maharashtra retained the largest share. However, its share of the number of new infections decreased by 22.2% from the third phase. The next three states with the highest number of new infections were Tamil Nadu, Andhra Pradesh, and Karnataka. Of these three states, Tamil Nadu had the highest number of new infections in the second and third phases. In Tamil Nadu, 173,403 (14.1%) new infections were confirmed during this phase. On the other hand, in the other two states, many infections did not appear until the pre-phase. The number of new infections in the third phase in Andhra Pradesh was 8,824 (2.6%), but it rapidly increased to 137,924 cases (11.2%) in the fourth phase. Thus, the number of new infections in this state rivaled that in Tamil Nadu. Similarly, the number of new infections in Karnataka largely increased from 9,001 (2.6%) in the third phase to 117,364 (9.5%) in the fourth phase. Thus, some states had seen a significant increase in the number of new infections in this phase, even states where the number of new infections had been low in the previous phase. Uttar Pradesh, West Bengal, Telangana, and Bihar also recorded prominent rates of infection, but the number of new infections was not as large as in Andhra Pradesh and Karnataka.

Increasingly important is the trend of new infections in Delhi. As we have discussed, this state was one of the central states of the COVID-19 outbreak in India since the first phase. The number of total infections in Delhi in the third phase was 80,188 (15.4%), second to Maharashtra. However, the number of new infections in Delhi, which was 61,639 in the third phase, declined in the fourth phase, amounting to 56,528. The number of new infections in Delhi by week increased almost consistently throughout the week of June 21, peaking at 23,442 new infections in one week (Figure 5). On the other hand, the number of new infections had been decreasing since then, with 7,185 new cases during the week of July 26. In addition to Delhi, in some states, the number of infections decreased compared to the previous phase; all these states had a low incidence of infections (Figure 4-b~d). That is why the trend of such infections in Delhi is noteworthy.

Thus, a new trend in the spatial pattern of COVID-19 infections was observed in the fourth phase.

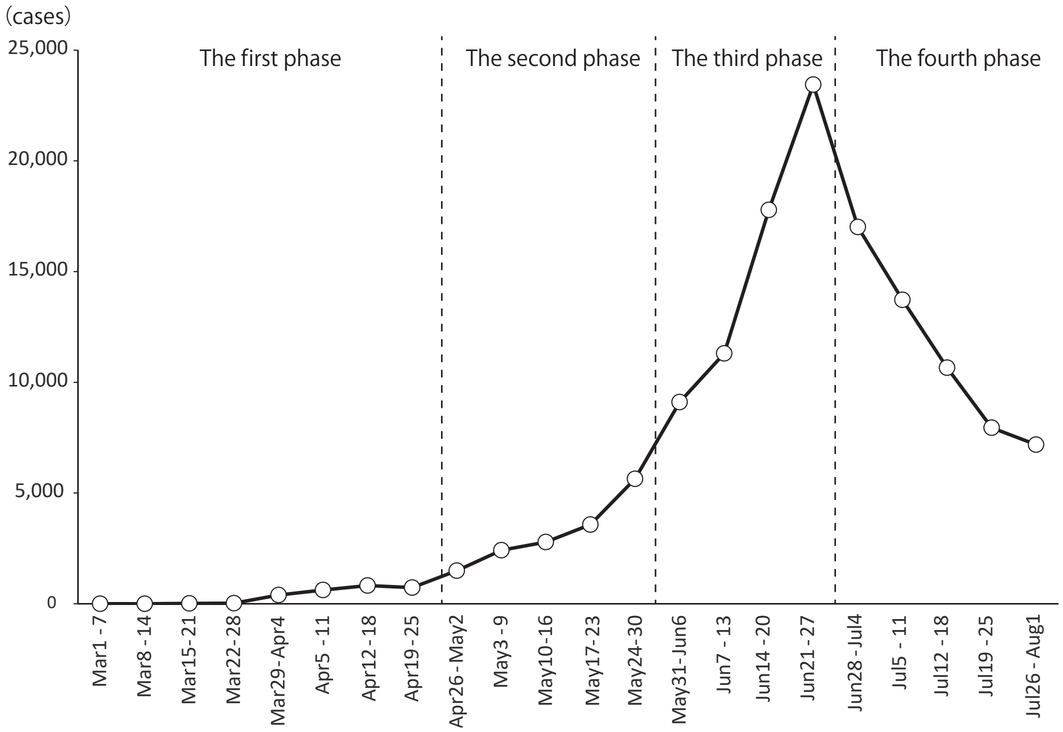

Figure 5. Number of new infections in Delhi by week
Note: Prior to February 29, there were no infections.
Source: Based on data from covid19india.org.

V. Trends in COVID-19 Infection at the District Level

In this section, we provide spatial resolution and clarify changes in the number of the COVID-19 infections at the district level. However, as described in Section II, it is not possible to handle data on the number of infections for all districts using GIS. Further, the capture ratio of infections for district data may be less sufficient than data on the number of infections by state. Hence, unlike the analysis by state level in Section IV, it is impossible to analyze the number of infections by district level for the whole of India. We discuss the trend in COVID-19 infections by district using some case studies of specific districts in Section V. [14]

The states that serve as an example are Maharashtra and Karnataka. Maharashtra had been the state with the highest number of infections in India since the first phase. On the other hand, in Karnataka, it was in the fourth phase that the number of new infections increased rapidly. Thus, some differences are accepted in both states to increase the number of COVID-19 infections. Therefore, this state serves as a significant case study for analysis at the district level. Like state-level data, data on the number of infections by district also include data in which the infected district was unknown. Here, an unidentified number is excluded from data to be used for analysis in this paper. In addition, the total number of infections by district may be larger than the number of infections by state. For these, it was judged that there were no major problems as with the analysis in the previous section.[15] Hence, when referring to the total number of infections by state, which is the sum of data by district, the figure for infections is different from the data by state in Section III.

1. Maharashtra

First, as a state overview of Maharashtra, population by district is shown. Figure 6 shows the population by

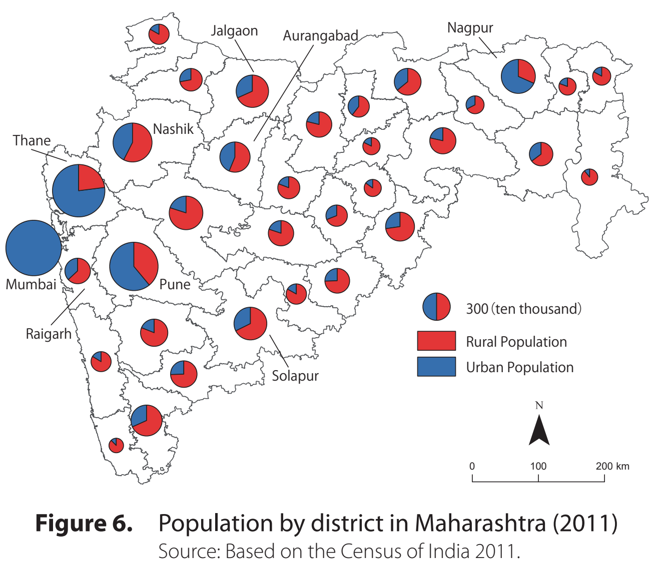

Figure 6. Population by district in Maharashtra (2011)
Source: Based on the Census of India 2011.

district in Maharashtra divided into rural and urban populations. As per the Census of India 2011, the population of this state was 112,374,333, with a rural population of 61,556,074 (54.8%) and an urban population of 50,818,259 (45.2%).[16] Thane had the largest population—11,060,148. This district has the character of the satellite city of Mumbai, and nearly 80% of the population was urban. Thane had the largest population in this state because Mumbai was divided into the Mumbai district and Mumbai suburban district. However, in data about the COVID-19 infection by district, these two districts are collectively treated as the Mumbai district. Therefore, we conveniently integrated the Mumbai suburban district with the Mumbai district and analyzed it. The total population of these two districts was 12,442,373[17] and was classified as an urban population. Pune forms an urban agglomeration with a population of over five million (Figure 2). This district had a population of 9,429,408, and its urban population ratio was over 60%. There were three other urban agglomerations in Maharashtra, with a population of more than one million each. The districts forming the core of these urban agglomerations were Nashik, Nagpur, and Aurangabad, with populations of 6,107,187, 4,653,570, and 3,701,282, respectively. Of these, only Nagpur (68.3%) had an urban population ratio higher than 50%, but the remaining two districts were approximately 40%.

Next, we examined the trend of COVID-19 infection by district. Initially, the number of infections in the whole state was 8,043, of which Mumbai accounted for 5,407 and 67.2% of all cases in the first phase (Figure 7-a). This was followed by Pune (1,052 cases) and Thane (879 cases), with these three districts accounting for more than 90% of all cases.

The number of new infections was 57,066 in the second phase. Mumbai (33,035 cases, 57.9%), Thane (9,212 cases, 16.1%), and Pune (6,485 cases, 11.4%) accounted for more than 10% of new infections, with no change in the states with the highest number of infections (Figure 7-b). These three states accounted for 85.4% of new infections, which represented a decrease compared to the first phase but remained highly concentrated. The number of total

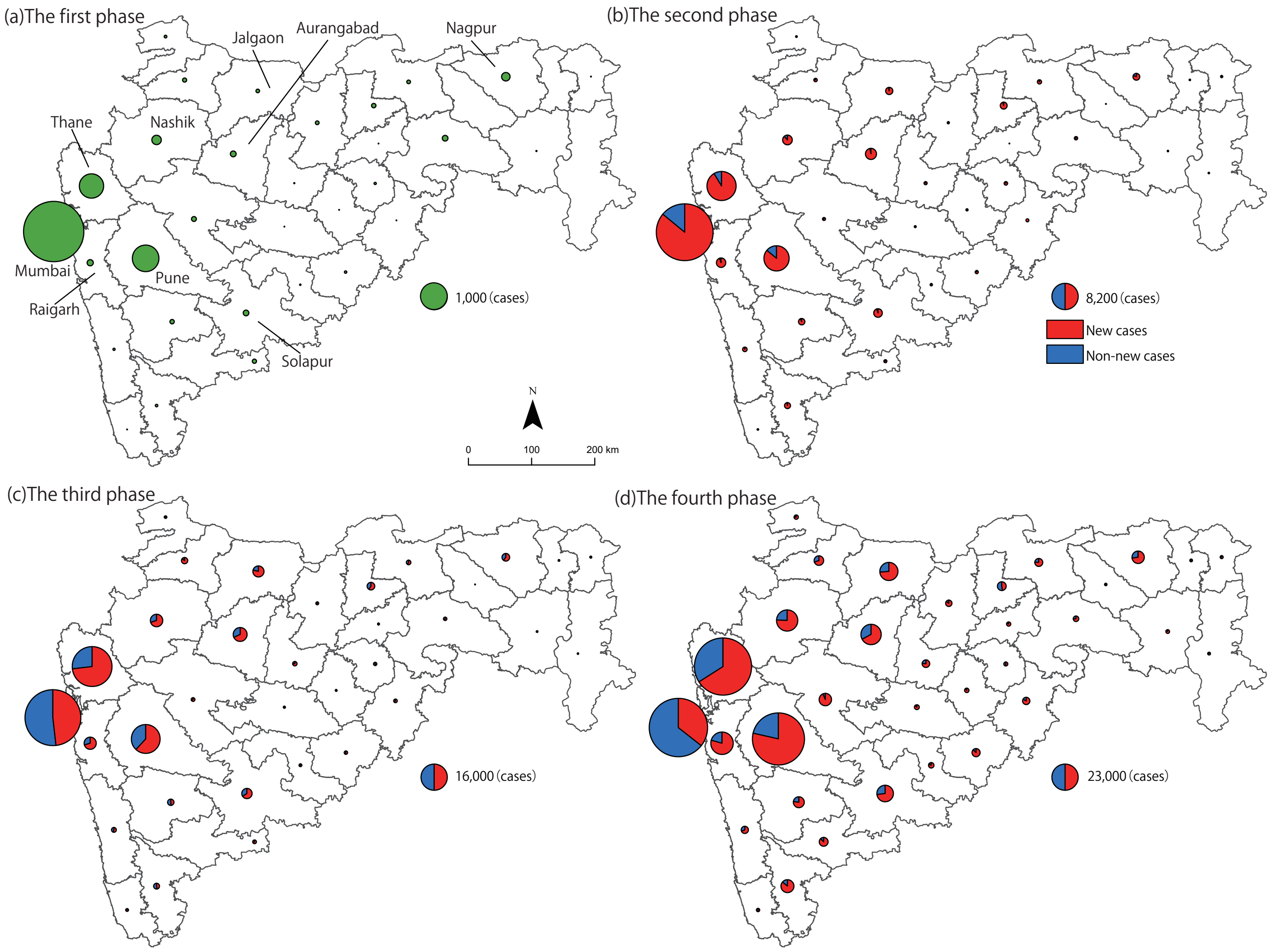

Figure 7. Number of infections by district in Maharashtra
Source: Based on data from covid19india.org.

infections was similar. Other than these three districts, districts with over 1,000 infections included Aurangabad (1,462 cases), Nashik (1,111 cases), and Raigarh (1,042 cases). These districts are located in the western part of Maharashtra. Aurangabad and Nashik are at the core of urban agglomerations with populations of at least a million. On the other hand, the total number of infections in Nagpur, an urban agglomeration of one million, was confined to 556 cases.

In the third phase, the number of new infections reached 93,952, and the total number of infections exceeded 150,000 (Figure 7-c). There was no change in the top-three districts in terms of the number of new infections. Mumbai, Thane, and Pune had 35,810, 27,524, and 12,224 new infections, respectively, in this phase. Mumbai still had the largest number of new infections, but the number itself was not considerably different from the second phase. On the other hand, the number of new infections in Thane and Pune significantly increased. In particular, the number of new infections in Thane increased approximately three times compared to the second phase and accounted for approximately 30% of all new infections in Maharashtra. Mumbai's share of new infections in Maharashtra thereby decreased from 57.9% to 38.1%. Thus, the number of new infections in the third phase was similar to that in the second phase in Mumbai, while it continued to increase significantly in Thane and Pune in this phase. Regarding the other districts, Aurangabad (3,124 cases), Nashik (2,592 cases), Raigarh (2,476 cases), Jalgaon (2,292 cases), and Solapur (1,664 cases) followed Pune, with more than 1,000 new infections each. As indicated in the second phase, it turns out that all these districts are located in the western part of the state.

Finally, the infection trend in the fourth phase is discussed (Figure 7-d). The number of new infections was 272,253, and that of total infections accounted for over 400,000. The top-three new districts for infections were Thane (72,877 cases), Pune (72,169 cases), and Mumbai (41,079 cases). There was no change in the composition of the top-three districts in terms of the number of cases. Notably, the number of new infections in Thane and Pune increased significantly following the third phase and was higher than that in Mumbai in this phase. On the other hand, the number of new infections in Mumbai had been stable compared to these two districts since the second phase, although it increased compared to the third phase. Consequently, Mumbai, Thane, and Pune accounted for 26.7% (115,331 cases), 25.6% (110,492 cases), and 21.3% (91,930 cases), respectively, of total infections. Further, the number of new infections exceeded 10,000 in some districts in the fourth phase, except for these three districts. The districts with a large number of new infections following Mumbai were Raigarh (13,437 cases), Nashik (11,524 cases), Aurangabad (9,425 cases), Jalgaon (8,200 cases), and Solapur (6,818 cases). The number of new infections surpassed 5,000 in all five states. As in the second and third phases, all districts are located in the western part of the state. In the case of Nagpur, the number of new infections increased from 844 to 3,787 from the third to the fourth phase, but the spread of infection had not been confirmed to be as large as in the five districts mentioned above.

2. Karnataka

The population of Karnataka in 2011 was 61,095,297, with a rural population of 37,469,335 (61.3%) and an urban population of 23,625,962 (38.7%). The largest district in terms of population was Bangalore (9,621,551), with an urban population of over 90% (Figure 8). Bangalore was the only urban agglomeration in Karnataka with a population of one million or more. The second-most populous district after Bangalore was Belgaum. The population of this district was 4,779,661, but the urban

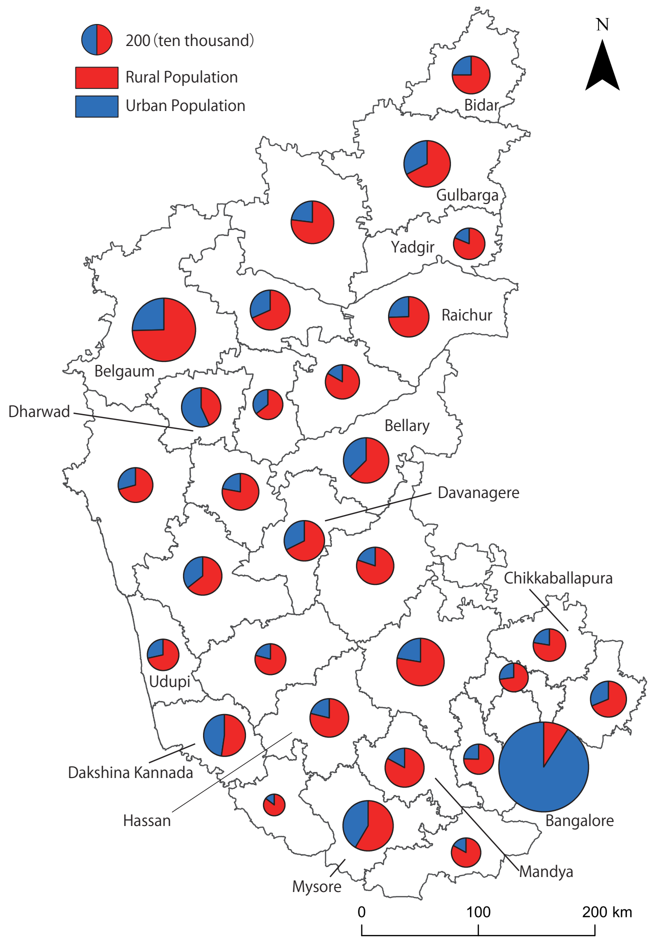

Figure 8. Population by district in Karnataka (2011)
Source: Based on the Census of India 2011.

population ratio was 25.3% lower than the ratio for the whole state. Apart from Bangalore, only Dharwad in this state had an urban population ratio of more than 50%. Only two districts had an urban population ratio of more than 40%—Mysore and Dakshina Kannada.

Subsequently, we confirmed the trend of COVID-19 infections by district in Karnataka. Figure 9-a shows the number of infections by district in the first phase. The

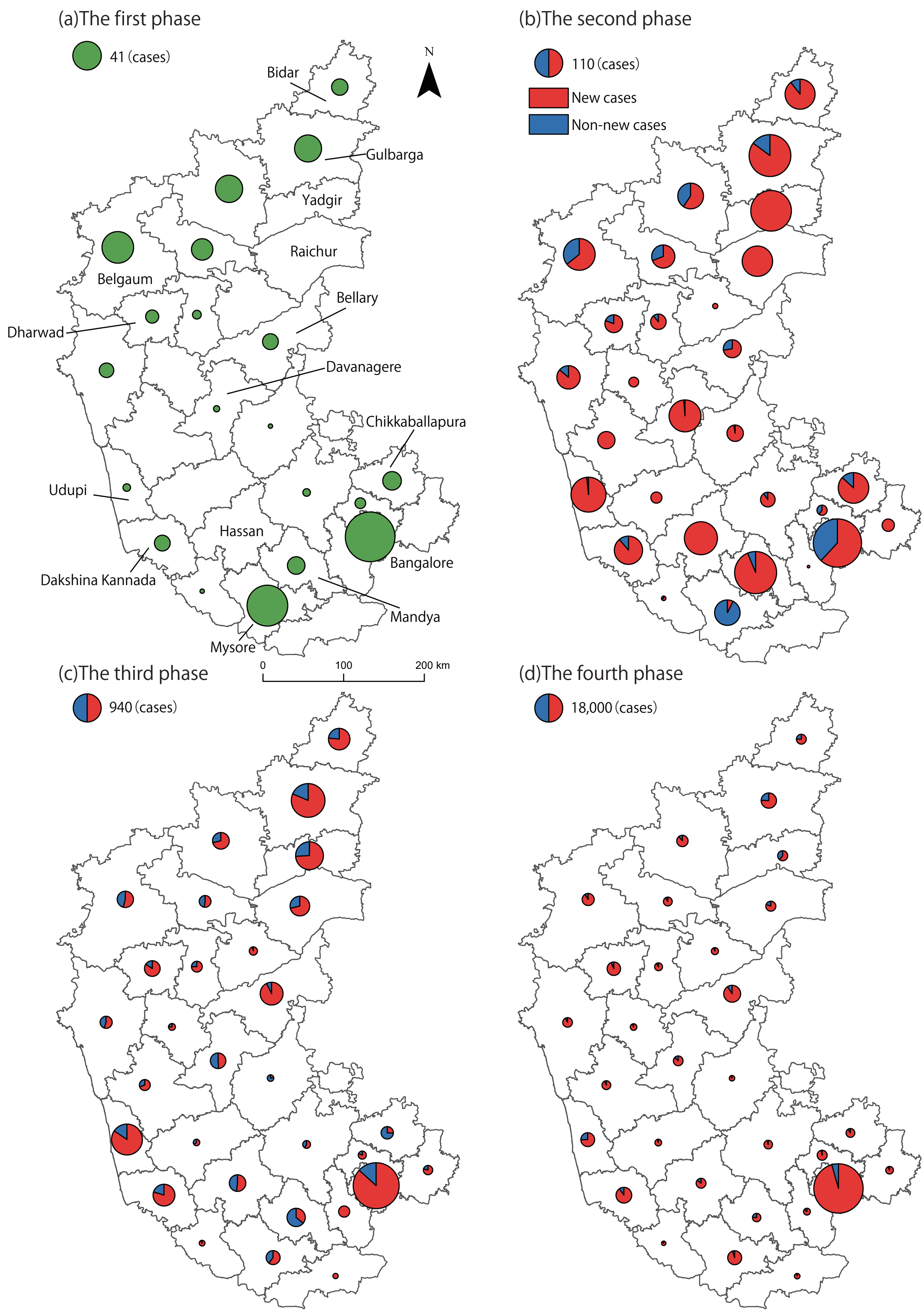

Figure 9. Number of infections by district in Karnataka
Source: Based on data from covid19india.org.

number of infections in this state was 483. The district with the largest number of infections was Bangalore (129 cases, 26.7%). Mysore and Belgaum were next with 87 cases and 52 cases, respectively, and only these three districts had more than 50 cases. These three districts accounted for 55.5% of all infections, but the concentration of cases in certain districts was smaller than that in Maharashtra.

The number of new infections was 2,417 in the second phase. As for the number of new infections by district, more than 200 infections were counted in Mandya (241, 10.0%), Yadgir (240, 9.9%), Gulbarga (216, 8.9%), and Bangalore (209, 8.6%) (Figure 9-b). Notably, some districts had more new cases than Bangalore, albeit with some differences. In addition, seven districts with 100 or more new infections were seen: Udupi (174), Hassan (158), Davanagere (146), Raichur (134), Chikkaballapura (118), Bidar (117), and Dakshina Kannada (103). On the other hand, the number of new infections in Mysore and Belgaum, which were located next to Bangalore in the first phase, did not reach 100 cases.[18] Thus, in the second phase, the trend of new infections concentrated in specific districts was weaker than that in the first phase. The distribution of these districts was dispersed throughout the state and was not regionally uneven.

By the third phase, the number of new infections in this state was 8,987, and the total number of infections exceeded 10,000. The number of new infections by district was found to be the highest in Bangalore, with 2,193 cases (24.4%), followed by Gulbarga (1,110 cases, 12.4%), Udupi (962 cases, 10.7%), Yadgir (689 cases, 7.7%), Bellary (579 cases, 6.4%), and Dakshina Kannada (452 cases, 5.0%) (Figure 9-c). The concentration of infections in specific districts was not strong, similar to the second phase, compared to Maharashtra. However, considering that the proportion of new infections in Bangalore reached nearly 25%, and the top-three districts accounted for nearly 50% of all new infections, these trends were weaker than those in the second phase. Focusing on the distribution of these five districts, we could find three regional patterns: 1) Bangalore, 2) the southwestern part of the state, and 3) the northeastern part of the state. This distribution pattern accorded with the distribution of the total number of infections.

Finally, we examined the trend of infection in the fourth phase (Figure 9-d). The number of new infections increased rapidly to 117,364 in this phase. The total number of infections reached nearly 130,000, making it the fifth-highest in India. As for the number of new infections by district, Bangalore accounted for 54,865 cases, 46.7% of all cases. The total number of infections in this district was 57,396, accounting for 44.4% of infections for this state. The number of new infections in Bangalore was extremely high, as shown by the fact that Bellary had the second-highest number of new infections after Bangalore, with 6,046 cases (5.2%). However, districts other than Bangalore were also hit by a significant increase in the number of new infections compared to the third phase. The number of new infections surpassed 1,000 in 24 of 29 districts, except Bangalore. Except for a few districts such as Gulbarga, Udupi, and Yadgir, where the number of total infections was relatively high in the third phase, most of the total infections were accounted for by new infections in the fourth phase.[19] As described above, Karnataka experienced a rapid increase in the number of new infections in this phase. Despite a significant increase in the number of infections in many districts, the sharp increase in Bangalore is particularly noteworthy.

VI. Conclusion

This study reveals spatial patterns of COVID-19 infection in India up to August 1, 2020, through GIS mapping.

First, we examined changes in the COVID-19 infection in India. Trends since January 30, when the first infection was confirmed in India, were examined based on data compiled weekly to understand better the characteristics of the infection trend in a time sequence. On this basis, the relevant period was divided into four phases based on the number of new infections reaching 10,000, 50,000, and 100,000 per week, and analysis was conducted by state and district based on this classification of phases.

If examine the trend of infection by state in India, infections in the first phase tended to be concentrated in states with relatively large populations located in western and northern India. In particular, Maharashtra, Gujarat, and Delhi were at the center of the wave of infections in the first phase. On the other hand, regarding states located in eastern and southern India, only Tamil Nadu stood out, and the number of infections was not considerably large. In the second phase, the spread of infections in Tamil Nadu was noted. In addition, four states (Maharashtra, Tamil Nadu, Delhi, and Gujarat) accounted for over 70% of new infections. Although the number of new infections increased significantly in the second phase, it tended to be concentrated in these four states. In the third phase, Maharashtra continued to have the largest number of new infections, followed by Delhi and Tamil Nadu. In this phase, Delhi stood out for the expansion of infections. On the other hand, Gujarat, which had the second-most infections after these three states, did not experience a significant increase in the number of new cases in the third phase. For the remaining states, the number of new infections remained low compared to the above three states, although it increased in this phase, even in states with a certain population size, such as Bihar, West Bengal, and Karnataka. The major changes in spatial patterns of the COVID-19 infection were accepted in the fourth phase. Maharashtra remained the state with the largest number of new infections, followed by Tamil Nadu. Importantly, Andhra Pradesh and Karnataka had the second-highest number of new infections after these two states. In both Andhra Pradesh and Karnataka, the number of new infections increased rapidly in the fourth phase, although there were not so many cases until the third phase. Further, the trend of new infections in Delhi was noteworthy. Delhi had been one of the states most affected by the COVID-19 infection in India since the first phase. However, the number of new infections in Delhi declined in the fourth phase.

Subsequently, we used spatial resolution and examined the infection trend by district. The cases of Maharashtra and Karnataka, where there are differences in timing and increases in the number of infections, were used as case studies. To summarize the trend of COVID-19 infection in Maharashtra, it was found that COVID-19 spread mainly in Mumbai but also in the western part of this state, especially in districts that form urban agglomerations and have a certain population size. By the second phase, the majority of infections were in Mumbai. However, the number of new infections in Mumbai remained relatively stable subsequently. Conversely, Thane and Pune saw a remarkable expansion of the number of new infections after the third phase. In particular, these two districts had significantly more new infections than Mumbai in the fourth phase.

On the other hand, if we turn to Karnataka, Bangalore had the highest number of infections in the first phase. However, these cases were not as concentrated as in Maharashtra. In the second phase, the number of new infections in some districts exceeded that in Bangalore, and such cases were less concentrated in certain districts than in the first phase. Bangalore was exposed to new infections again when it entered the third phase. In contrast to Maharashtra, the number of new infections in the top-three districts was nearly 50% of those in the state as a whole, although infections were still not strongly concentrated in particular districts. Focusing on the distribution of the top-five districts in terms of the number of new infections, we found three regional patterns: 1) Bangalore, 2) the southwestern part of the state, and 3) the northeastern part of the state. This pattern was consistent with the distribution of the total number of infections. The number of new infections in this state increased rapidly in the fourth phase. This led to a significant increase in the number of new infections in many districts. However, the number of new infections in Bangalore was extremely high in Karnataka, with nearly half of the cases concentrated in this district.

In summary, the trend of COVID-19 infections in India by state showed certain spatial patterns for each phase. In terms of the trend for the number of infections at the district level, it was clear that spatial patterns of infection spread had changed, even within the same district.

However, this study does not examine the factors that define such spatial patterns of COVID-19 infection in India. Gupta et al. (2020) explained the distribution of the total number of COVID-19 cases based on natural conditions. However, the results of this study suggest that, in addition to population and city size, infections might be linked to urban systems, such as urban hierarchy and inter-urban connectivity. Such an association between epidemics and urban systems is pointed out in Kobayashi (2008) and is one of the important factors in analyzing the diffusion of an epidemic. However, these are not merely hypotheses, and as stated in previous studies, it is necessary to conduct spatial analysis using GIS that considers various factors. It is also necessary to continue analyzing the trend of COVID-19 infections in India. These points remain research issues.

Acknowledgment

This study used data from the website covid19india.org. I would like to thank everyone who collected and updated these data. This study is part of the results of a project for the National Institutes for the Humanities (NIHU).

Notes

1. According to “Coronavirus disease(COVID-19) Situation Report 194” https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200801-covid-19-sitrep-194.pdf?sfvrsn=401287f3_2 (accessed August 21, 2020).
2. It is the postal code system used by the United States Postal Service (USPS).
3. Data from the Ministry of Health and Family Welfare were used.
4. It is a method of interpolation that estimates an unknown value based on a set of known numerical data (according to the ESRI Japan website). https://www.esrij.com/gis-guide/spatial/interpolation/ (accessed August 21, 2020).

5. As in this study, data from covid19india.org were used.
6. https://www.covid19india.org/ (accessed August 7, 2020).
7. In addition, a sheet summarizing the data sources collected by covid19india.org has been updated.
8. https://gadm.org/index.html (accessed August 7, 2020).
9. In particular, a significant change in the composition of the district was observed in Telangana.
10. As there was only one case on January 30, we included it in the number of cases for the week of February 2.
11. Based on population from the Census of India 2011.
12. The capture rate of data on the number of new infections exceeds 96% in both the second and third phases. In the fourth phase, the total number of new infections in the national-level data was 1,222,595, compared to 1,230,607 in the state-level data, with a difference of approximately 8,000.
13. In "states" sheet used for analysis in this chapter, Dadra and Nagar Haveli and Daman and Diu are aggregated together. However, if we check the "state-wise daily" sheet, we can understand the composition of Dadra and Nagar Haveli and Daman and Diu.
14. As April 26 is the first date for the "district" sheet to be tabulated, the first phase's data up to April 26 are used for convenience in this chapter.
15. The capture rate of Maharashtra's data on new and total infections exceeded 99% for all phases but the first phase. Maharashtra's data on district-level infections in the first phase was 8,043, which was over-counted compared to the state-level data (7,628 cases). Karnataka's data capture rate for new and total infected cases exceeded 96% for all phases.
16. For reference, the total rural and urban population in India in 2011 was 833,748,852 (68.9%) and 377,106,125 (31.1%), respectively (according to the Census of India 2011).
17. In 2011, the population of Mumbai and Mumbai suburban districts was 3,085,411 and 9,356,962, respectively (according to the Census of India 2011).
18. Belgaum had 94 new infections in this phase, while Mysore had only 7 cases.
19. In Bangalore, new infections in the fourth plase accounted for 95.6%.

References

Cordes, J. and Castro, M.C. (2020): Spatial Analysis of COVID-19 Clusters and Contextual Factors in New York City. *Spatial and Spatio-temporal Epidemiology*, 34, 1–8. DOI:https://doi.org/10.1016/j.sste.2020.100355 (E)

Gupta, A., Banerjee, S. and Das, S. (2020): Significance of Geographical Factors to the COVID-19 Outbreak in India. *Modeling Earth Systems and Environment*, 17, 1–9. DOI:https://doi.org/10.1007/s40808-020-00838-2 (E)

Kobayashi, S. (2008): Hendo suru Sekai heno Apurochi: Jimbunchirigaku toha Donna Gakumon ka. Kobayashi, S. and Sugiura, Y. (eds.): *Kaitei ban Jimbunchirigaku*. Hosodaigakukyoikushinkokyokai, 11–33. (J)

Kuebart, A. and Stabler, M. (2020): Infectious Diseases as Socio-spatial Processes: The COVID-19 Outbreak in Germany. *Journal of Economic and Social Geography*, 111, 482–496. DOI:https://doi.org/10.1111/tesg.12429 (E)

Mihara, A. (2019): Exploring the Research Trends in Medical Geography in Japan while Focusing on Research for Medical Services. *Jimbun Kenkyu* (Kanagawa University), 197, 55・57–72. (JE)

Mollalo, A., Vahedi, B. and Rivera, K.M. (2020): GIS-based Spatial Modeling of COVID-19 Incidence Rate in the Continental United States. *Science of the Total Environment*, 728, 1–8. DOI:https://doi.org/10.1016/j.scitotenv.2020.138884 (E)

Murugesan, B., Karuppannan, S., Mengistie, A.T., Ranganathan, M. and Gopalakrishnan, G. (2020): Distribution and Trend Analysis of COVID-19 in India: Geospatial Approach. *Journal of Geographical Studies*, 4(1), 1–9. DOI:https://doi.org/10.21523/gcj5.20040101 (E)

Nakaya, T. (2008): Spatial Epidemiology in Geographic Information System Environments. *Journal of the National Institute of Public Health (Hoken Iryo Kagaku)*, 57, 99–116. (JE)

Nakaya, T., Tanimura, S., Nihei, N. and Horikoshi, Y. (eds.) (2004): *Hoken Iryo no tame no GIS. Kokonshoin*. (J)

Zhou, C., Su, F., Pei, T., Zhanga, A., Du, Y., Luo, B., Cao, Z., Wang, J., Yuan, W., Zhu, Y., Song, C., Chen, J., Xu, J., Li, F., Ma, T., Jiang, L., Yan, F., Yi, J., Hu, Y., Liao, Y. and Xiao, H. (2020): COVID-19: Challenges to GIS with Big Data. *Geography and Sustainability*, 1, 77–87. DOI:https://doi.org/10.1016/j.geosus.2020.03.005 (E)