



Regional trends in the number of COVID-19 cases

Keisuke Chujo¹ · Tatsunori Seki² · Toshiki Murata¹ · Yu Kimura¹ · Tomoaki Sakurai² · Satoshi Miyata² · Hiroyasu Inoue^{3,4} · Nobuyasu Ito⁴

Received: 29 April 2023 / Accepted: 19 January 2024 / Published online: 12 March 2024
© The Author(s) 2024

Abstract

In this study, we analysed the novel coronavirus disease (COVID-19) cases data to investigate the regional infection trends in Japan. There had been seven outbreaks by October 2022 in Japan. In each outbreak, the number of COVID-19 cases has increased at different rates in different regions. The prefectural infection ratio is defined using COVID-19 cases data. We calculate the prefectural infection ratio and study the characteristic of each pandemic wave. The prefectural order of infection progression is estimated in each past wave of the COVID-19 pandemic. This study shows that the infection spread from the Kanto region in the fourth pandemic wave and the infection spread simultaneously from four regions in the sixth wave. It is also found that the infection situation trend in Okinawa differs from that in the other regions.

Keywords COVID-19 · Infection progression · Regional trend · Spatio-temporal analysis

This work was presented in part at the joint symposium of the 28th International Symposium on Artificial Life and Robotics, the 8th International Symposium on BioComplexity, and the 6th International Symposium on Swarm Behavior and Bio-Inspired Robotics (Beppu, Oita and Online, January 25–27, 2023).

Keisuke Chujo is the presenter of this paper.

✉ Keisuke Chujo
keisuke.chujo@g.softbank.co.jp

Satoshi Miyata
satoshi.miyata@g.softbank.co.jp

Hiroyasu Inoue
inoue@gsis.u-hyogo.ac.jp

Nobuyasu Ito
nobuyasu.ito@riken.jp

- ¹ Solution Engineering Division, SoftBank Corp., Tokyo, Japan
- ² Data Infrastructure Strategy Division, SoftBank Corp., Tokyo, Japan
- ³ Graduate School of Information Science, University of Hyogo, Hyogo, Japan
- ⁴ RIKEN Center for Computational Science, Tokyo, Japan

1 Introduction

The COVID-19 pandemic caused by the SARS-CoV-2, first reported at the end of 2019, has spread rapidly worldwide. Countries around the world have taken various measures including quarantines, to contain its spread. However, 3 years later, COVID-19 has not been eliminated yet and resulted in several outbreaks. In Japan, the disease has also spread quickly throughout the country, since the first case was detected and new cases were still occurring by the end of 2022. Studying the spread of infections can be useful for combating both current and future infectious diseases. Because COVID-19 is a highly contagious infectious disease, once it starts to spread, it spreads almost simultaneously, but never evenly across space, with some areas having more infected people than others at the same time.

Previous studies have suggested that population fluidity may significantly affect COVID-19 rate. For example, a reduction in travel has been reported to affect three epidemiological outcome measures, the number of export cases, the probability of an outbreak and the time delay to an outbreak [1, 2]. Noda reported a correlation between the number of effective reproductions estimated from the epi-curves of the target prefectures (Tokyo, Osaka, Aichi, Hokkaido, Kanagawa and Kyoto) and the behaviour patterns classified from the mesh data using non-negative tensor factorization [3]. It has also been reported that there is a correlation

between the influx of population data for Tokyo during the pandemic and the number of effective reproductions, considering time delays [4].

However, a few studies have investigated the variations in infection status trends between regions during multiple waves of the pandemic. This study investigates inter-regional trends in COVID-19 pandemic. It is necessary to prepare for pandemics in general, not only for COVID-19, as they are likely to occur in the future as well. The data obtained could be helpful for the current and future control of infectious diseases.

2 Methodology

2.1 Definition of the indicator used in this study

In this study, we use open data on the novel coronavirus infections provided by the Japanese Ministry of Health, Labour and Welfare [5]. First, the population-adjusted infection rate for each prefecture is calculated by dividing the infection number in the prefecture by its population. The prefectural infection ratio is defined as T_r in Eq. (1)

$$T_r(t) = \frac{C_r(t)/P_r}{\sum_r C_r(t) / \sum_r P_r}, \quad (1)$$

where the $C_r(t)$ and P_r are the number of positive of COVID-19 tests at a specific time in region r and the population number in region r , respectively. As this study focuses on the prefectures in Japan, r indicates one of the 47 prefectures. The parameter t refers to time. The prefectural infection ratio is abbreviated as the PIR in the following text in this study.

The PIR compares the prefectural population-adjusted infection rate with the national population-adjusted infection rate as a baseline. In other words, a PIR value greater than 1.0 for a prefecture indicates a more active state of infection in the region than the national average. A PIR value closer to 1.0 also indicates that the infection in the prefecture is closer to the national trend, while a PIR of a prefecture less than 1.0 indicates that the infection situation in a region is more moderate than that at the national level. The number of infected individuals was calculated as the cumulative value within the target period of each pandemic wave. This indicates that the national cumulative number of infected persons was also small during the early part of the target period. The value of the national cumulative number of infections is included in the denominator of the PIR; therefore, the value of the population-adjusted infection rate by prefecture tends to be higher in the early part of each target period.

2.2 Target period

Prior to commencing the study, the following seven periods are defined as COVID-19 pandemic waves. Although there is no clear delimitation of each COVID-19 pandemic wave period, the boundaries of the pandemic waves are determined referring to daily data regarding the number of patients nationwide.

Regarding the fifth and seventh waves, the possible interference of previous waves cannot be ruled out, because the fifth and seventh waves are start before the number of national daily COVID-19 cases drop overall. The changeover periods between waves 4 and 5 and waves 6 and 7 occur at similar times of the year. The SARS-CoV-2, which causes COVID-19, is characterised by its continued spread through a vast number of mutations. First-wave data were excluded from the scope of this study, because the number of infected persons in this wave was relatively small (Table 1).

3 Results and discussion

For the total of seven COVID-19 pandemic waves until the end of October 2022, the PIR is calculated daily and presented as time-series data. These time-series data are sorted in order of peak values for each prefecture and the top eight regions are extracted (Table 2) based on the PIR. Each wave of change in PIR has different characteristics and no general rule applicable to all pandemic waves has been determined yet. In this study, we share the results of our investigation regarding the fourth and sixth waves, because they have different and distinct characteristics.

Eight prefectures are extracted in the order of the peak values from the PIR time-series (Table 2) and hierarchically clustered against the two temporal features. Ward's method, a type of hierarchical agglomerative clustering, was used for the clustering. The Euclidean distance was adopted as the distance between PIR values of each prefectures.

Table 1 Periods of the COVID-19 pandemic waves

	Start date	End date
1st wave	1st February, 2020	30th June, 2020
2nd wave	1st July, 2020	15th November, 2020
3rd wave	1st December, 2020	28th February, 2021
4th wave	1st March, 2021	30th June, 2021
5th wave	1st July, 2021	30th September, 2021
6th wave	1st January, 2022	14th June, 2022
7th wave	15th June, 2022	11th October, 2022

Table 2 The top eight prefectures in descending order of the PIR value of each COVID-19 pandemic wave

	2nd wave		3rd wave	
	Pref.	Peak value	Pref.	Peak value
1	Tokyo	9.68	Osaka	2.69
2	Kagoshima	7.55	Tokyo	2.68
3	Okinawa	3.62	Hokkaido	2.62
4	Saitama	2.27	Aichi	1.70
5	Osaka	2.14	Kanagawa	1.69
6	Fukuoka	1.95	Okinawa	1.58
7	Tochigi	1.58	Hyogo	1.43
8	Aichi	1.56	Chiba	1.41
	4th wave		5th wave	
	Pref.	Peak value	Pref.	Peak value
1	Chiba	4.37	Tokyo	5.09
2	Ibaraki	3.62	Okinawa	3.32
3	Osaka	3.62	Kanagawa	2.15
4	Miyagi	3.25	Chiba	1.87
5	Okinawa	3.06	Osaka	1.58
6	Tokyo	2.69	Saitama	1.52
7	Hyogo	2.06	Fukui	1.25
8	Saitama	1.88	Akita	1.12
	6th wave		7th wave	
	Pref.	Peak value	Pref.	Peak value
1	Okinawa	22.63	Okinawa	8.47
2	Yamaguchi	5.46	Shimane	2.44
3	Hiroshima	4.27	Kumamoto	2.29
4	Shimane	2.43	Saga	2.07
5	Osaka	2.35	Kagoshima	1.90
6	Shiga	2.32	Aomori	1.70
7	Tochigi	2.22	Tokyo	1.59
8	Gunma	1.98	Fukui	1.57

3.1 Fourth pandemic wave

From the time-series data of the PIR, two time-related features are extracted: the number of days from the start date of the fourth wave until the date of the peak value and the number of days from the start date of the fourth wave until the first day that the PIR exceeded ad hoc 75% of the peak. For the purpose of examining the order of infection by prefecture, it is necessary not only to clarify the peak period of infection but also to define and indicate the time when infection is almost prevalent. Because if the PIR remains high near the peak, infection activity may have begun well before the peak. There are various ways to define the almost prevalent date, but since the peak PIR value varies depending on the epidemic wave and the prefecture, we define it as ad hoc 75% of the peak PIR value for each prefecture. Figure 1 shows the results of

extracting the quantities and plotting them on a scatterplot. Figure 2 shows the time trend of the fourth wave of the infection using the Japanese national map.

In the fourth wave, only Okinawa has two peaks; therefore, it is divided into an earlier and later peaks, as shown in Fig. 1. Hierarchical clustering confirms the presence of four groups which are (1) Ibaraki, Chiba, Tokyo, and Saitama; (2) Miyagi and Okinawa1; (3) Hyogo and Osaka; and (4) Okinawa2. Accordingly, it can be assumed that the general infection order the pandemic followed this trend. The elements within the earliest group in the fourth wave are regionally close to each other. The map in Fig. 2 shows that the lower the lightness of the colour is, the greater the PIR becomes. The area around Kanto was more prominent in March 2021. Miyagi and Okinawa in April 2021 and Osaka, Hyogo, and Nara in May 2021 exhibited larger PIRs. The PIR values during the fourth wave are shown in Fig. 3.

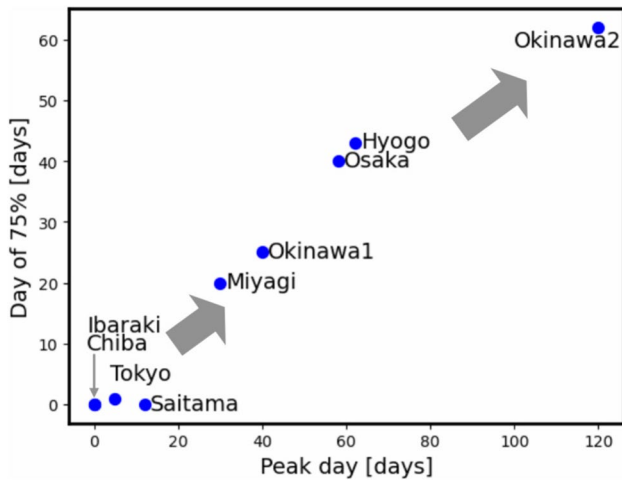


Fig. 1 The peak period of infection in the fourth wave for each prefecture. The vertical axis indicates the number of days taken from the start date of the wave to reach 75% of the peak value indicating the time of spread of infection. The horizontal axis indicates the number of days taken from the start date of the wave to reach the peak value. Time flows from the bottom left to the top right

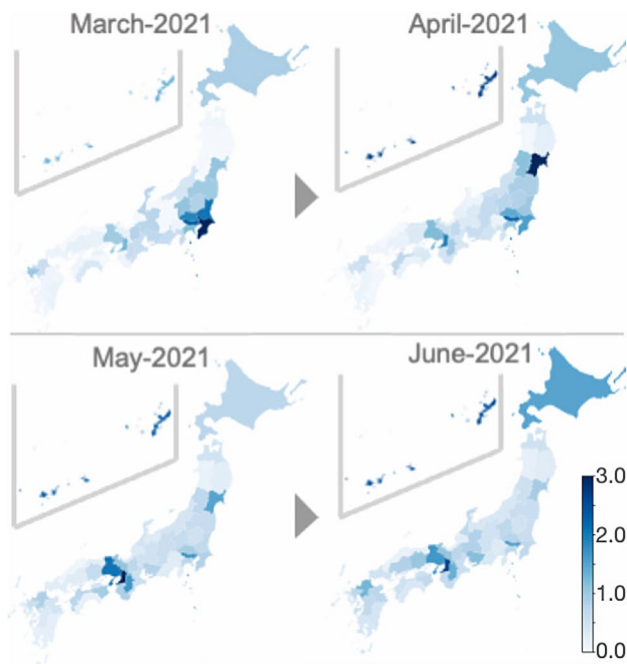


Fig. 2 Choropleth map of the PIR in the fourth wave. The spatial spread over time of COVID-19 is plotted from March to June 2022

3.2 Sixth pandemic wave

For the sixth pandemic wave, the highest PIR value in Okinawa is 22.63, which is more than four times the peak value in Yamaguchi, which had the second highest peak in Japan. Okinawa had shown an upward trend since May 2022. After declining slightly in May, the values in Osaka

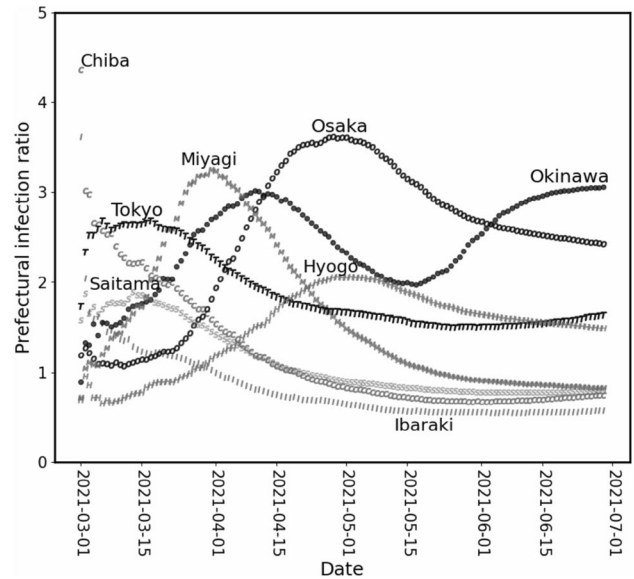


Fig. 3 Regional trend feature of the fourth wave. The top eight prefectures in descending order of the PIR value in the fourth pandemic wave are plotted on this graph. The vertical axis indicates the PIR which is defined in Eq. (1). The PIR value greater than 1.0 indicates a more active state of infection in the region than the national average. A matter that should be paid attention is only Okinawa has two peaks

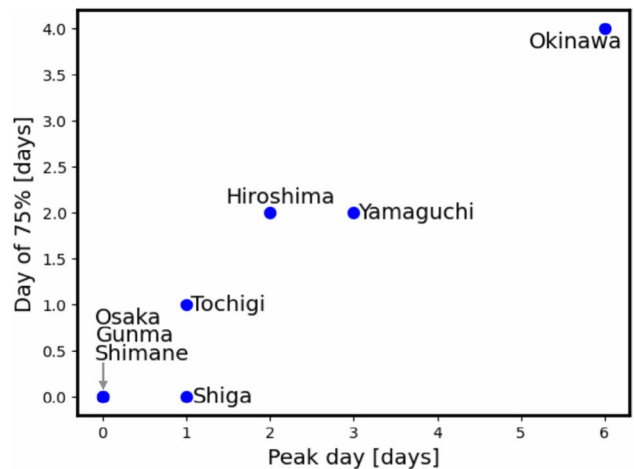


Fig. 4 The infection peak in the sixth wave for each prefecture. The vertical axis indicates the number of days taken from the wave start date to reach 75% of the peak value. The horizontal axis indicates the number of days taken from the wave start date to reach the peak value

remained largely unchanged and were greater than 1, which were higher than the values obtained at the national level. For the sixth wave, the scatter plot of the temporal features and the map of infection trends were similar to those of the fourth wave.

By comparing Fig. 1 with Fig. 4, observe that the period on the horizontal axis in Fig. 4 is shorter. This implies that both the peak and the 75% of the peak values in the sixth

wave were concentrated at the beginning of the sixth pandemic wave. Moreover, the period of the sixth wave was longer. After some initial regional differences, the trend has been gradual without much change. This indicates that the infection spread at an almost constant level relative to the national trend.

The initial situation of the sixth wave differs from that of the fourth wave, in which the infection appears to have started in four different areas simultaneously. The four areas are Okinawa, the Chugoku District, the Kinki District, and the Kanto District. The map shows that the four areas appear to have been infected early on in the wave, followed by the neighbouring prefectures (Fig. 5).

3.3 Discussion

Regarding the fourth wave, we examine the sequence of infection more locally. The order of infection outbreaks in the prefectures adjacent to Miyagi and in Chiba and

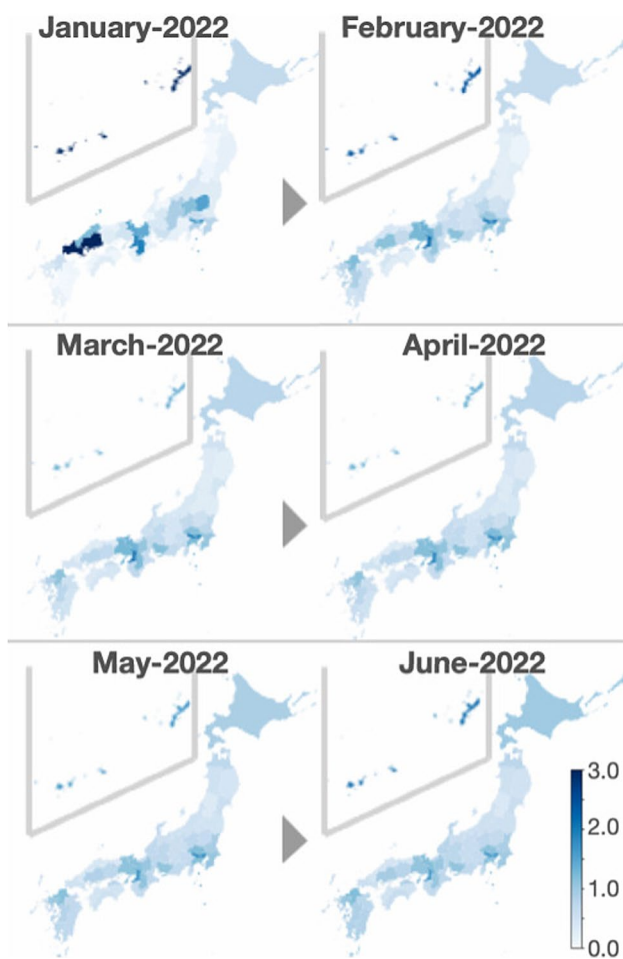


Fig. 5 Choropleth map of the prefectural infection ratio in the sixth wave. The spatial spread over time of COVID-19 is plotted from January to June, 2022

Ibaraki shows that infection outbreaks occur in ascending order of latitude as Fig. 6. The results suggest that the infection propagates overland from the Kanto area to Miyagi along neighbouring prefectures in the fourth wave.

On the other hand, the timing of the pandemic in Okinawa is earlier than that in Kagoshima, which is the prefecture nearest to Okinawa. Kagoshima later experienced an infection sequence in the Kyushu region. This result indicates that it is unlikely that the infection was transmitted southwards from the Kyushu region to Okinawa and unlikely that it is transmitted northwards from Okinawa to the Kyushu region. Regarding the access from other prefecture to Okinawa, there are only two modes of transportation, by boat or by plane, because Okinawa is surrounded by sea. It takes approximately 1 day by boat from Kagoshima to the neighbouring prefecture, whereas it takes 1 h and 20 min by plane. Naha City, the capital of Okinawa Prefecture, is approximately 660 km away from Kagoshima City, capital of Kagoshima Prefecture. Most people should visit Okinawa by plane. An air passenger movement survey conducted at Naha Airport showed that passengers arriving at Naha Airport came from airports in major cities, such as Tokyo, Osaka, and Fukuoka [7]. This suggests the need to consider linkages with prefectures including large cities rather than geographically closer prefectures when analysing the infection influence regarding Okinawa. However, the significantly highest value in Okinawa in the early part of the sixth wave is likely due to other factors. This is because infections in other large cities are less prevalent (Fig. 7).

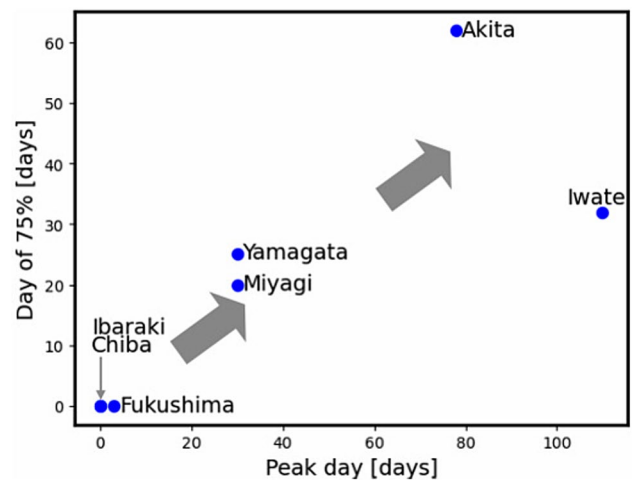


Fig. 6 The peak period of infection in the fourth wave for the prefectures around Miyagi. The vertical axis indicates the number of days taken from the start date of the wave to reach 75% of the peak value. The horizontal axis indicates the number of days taken from the start date of the wave to reach the peak value

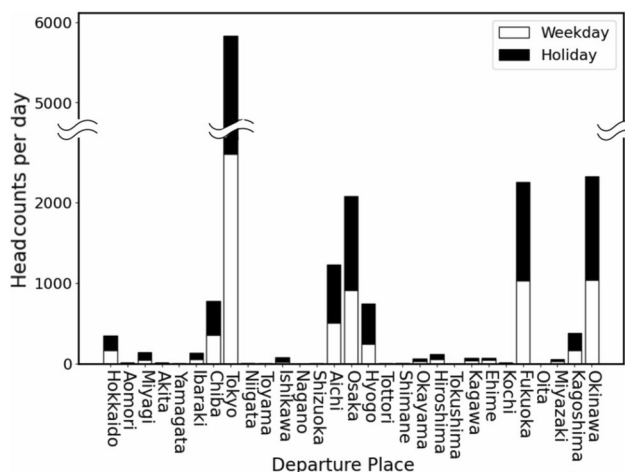


Fig. 7 Number of passengers per day arriving at Naha Airport. The figures are for 1 weekday and one holiday in October or November 2021. The horizontal axis indicates the prefecture of departure, where the headcount is at least 1

4 Summary and perspectives

This study examined the regional characteristics of the transmission of past COVID-19 during past pandemic waves in Japan. The prefectural infection ratio (PIR) was defined to chronologically identify regions with active infection. We further analysed the time-series data using agglomerative hierarchical clustering. This allowed us to determine the temporal order of the regions in which the infection was active. It was also found that the infection trends in Okinawa differed from those in other regions.

In the near future, population influx data will also be used to investigate the relationship between wide-area human flows and infection trends [8].

Data availability Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source,

provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Anzai A, Kobayashi T, Linton NM, Kinoshita R, Hayashi K, Suzuki A, Yang Y, Jung S, Miyama T, Akhmetzhanov AR, Nishiura H (2020) Assessing the impact of reduced travel on exportation dynamics of novel coronavirus infection (COVID-19). *J Clin Med* 9(2):601
2. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, Pastore y Piontti A, Mu K, Rossi L, Sun K, Viboud C, Xiong X, Yu H, Elizabeth Halloran M, Longini IM Jr, Vespignani A (2020) The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368(6489):395
3. COVID-19 AI & Simulation Project Website, I. Noda (2022) Correlation between movement pattern classification and spread of infection based on movement trajectory data" [in Japanese]. https://www.covid19-ai.jp/ja-jp/presentation/2020_rq4_count_emeasures_analysis/articles/article038/. Accessed 20th Dec 2022
4. Kimura Y, Seki T, Miyata S, Arai Y, Murata T, Inoue H, Ito N (2022) Hotspot analysis of COVID-19 infection using mobile-phone location data. *Artif Life Robot*. <https://doi.org/10.1007/s10015-022-00830-2>
5. Open data for novel corona-virus disease of the Japanese Ministry of Health, Labour and Welfare Website, <https://www.mhlw.go.jp/stf/covid-19/open-data.html>. Accessed 20 Dec 2022
6. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, Rinaldo A (2020) Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *PNAS* 117(19):10484
7. Air Passenger Movement Survey of the Japanese Ministry of Land, Infrastructure, Transport and Tourism Website, https://www.mlit.go.jp/koku/koku_tk6_000001.html. Accessed 20th Dec 2022
8. SoftBank Corporation Website (Zenkoku-Ugoki-Tokei) <https://www.softbank.jp/biz/services/analytics/ugoki>. Accessed 20 Dec 2022

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