Computers in Biology and Medicine 146 (2022) 105548

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Did the Tokyo Olympic Games enhance the transmission of COVID-19? An interpretation with machine learning

Akimasa Hirata [a,b,*], Sachiko Kodera [a], Yinliang Diao [a], Essam A. Rashed [c]

[a] Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Nagoya, 466-8555, Japan
[b] Center of Biomedical Physics and Information Technology, Nagoya Institute of Technology, Nagoya, 466-8555, Japan
[c] Graduate School of Information Science, University of Hyogo, Kobe, 650-0047, Japan

ARTICLE INFO

Keywords:
Machine learning
COVID-19
Olympic games
Viral transmission
Numerical modeling

ABSTRACT

Background: In the summer of 2021, the Olympic Games were held in Tokyo during the state of emergency due to the spread of COVID-19 pandemic. New daily positive cases (DPC) increased before the Olympic Games, and then decreased a few weeks after the Games. However, several cofactors influencing DPC exist; consequently, careful consideration is needed for future international events during an epidemic.
Methods: The impact of the Olympic Games on new DPC were evaluated in the Tokyo, Osaka, and Aichi Prefectures using a well-trained and -evaluated long short-term memory (LSTM) network. In addition, we proposed a compensation method based on effective reproduction number (ERN) to assess the effect of the national holidays on the DPC.
Results: During the spread phase, the estimated DPC with LSTM was 30%–60% lower than that of the observed value, but was consistent with the compensated value of the ERN for the three prefectures. During the decay phase, the estimated DPC was consistent with the observed values. The timing of the decay coincided with achievement of a fully-vaccinated rate of 10%–15% of people aged <65 years.
Conclusions: The up- and downsurge of the pandemic wave observed in July and September are likely attributable to high ERN during national holiday periods and to the vaccination effect, especially for people aged <65 years. The effect of national holidays in Tokyo was rather notable in Aichi and Osaka, which are distant from Tokyo. The effect of the Olympic Games on the spread and decay of the pandemic wave is neither dominant nor negligible due to the shifting of the national holiday dates to coincide with the Olympic Games.

1. Introduction

The Tokyo Olympic Games were held in Tokyo, Japan from July 23 to August 8, 2021, one year behind schedule (originally 2020) due to the COVID-19 pandemic caused by SARS-CoV-2. Just before the Games, the fifth wave of COVID-19, attributable to the new Delta viral variant (B.1.617.2), started to spread. Some dispute as to the possibly-significant contribution of the Olympic Games to the epidemic spread was noted before and during the event and the editorials was published on the safety issue of the Games [1]. Some newspapers reported “Tokyo fears that Games spread COVID-19” (August 8, 2021, The Japan Times). Similar articles have appeared worldwide and the topic has been debated scientifically [2–4]. The Games were held with limited physical public access and without spectators for the safety of participants and residents.

Although more than 320 cases have been linked to the Olympic Games [5], the event’s impact on the local community is unclear. In addition, new daily positive cases (DPC) in Tokyo have decreased significantly since the end of August 2021. After the peak DPC of 5773 cases was reached on August 13, 2021, new DPCs rapidly decreased to 219 at the end of September. The Ministry of Health Labor and Welfare of Japan mentioned in its national advisory board meeting (September 12, 2021) that “the reason for the abrupt decrease in DPC is unclear.”

It is not straightforward to resolve the open question of whether the DPC increase is attributable to the Olympic Games is complicated by several cofactors. It is crucial to better understand the effects of these pandemic cofactors on future international events, as well as their economic impacts [6], including medical resource allocation [7–9]. We have identified several parameters influencing new DPCs in our previous studies based on machine learning prediction (nonlinear regression)

* Corresponding author.Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, 466-8555, Japan.
E-mail address: ahirata@nitech.ac.jp (A. Hirata).

https://doi.org/10.1016/j.compbiomed.2022.105548
Received 8 February 2022; Received in revised form 2 April 2022; Accepted 18 April 2022
Available online 26 April 2022

[10,11]. In addition, the vaccination rate reached 40.7% and 60.2% (in Japan) at the end of August and September 2021, respectively [12].

Various computational or mathematical approaches have been used to evaluate viral transmission in the field of epidemiology [13–18]. Most studies have projected or forecasted COVID-19 status for a given scenario, or evaluated risk. In our previous study, we proposed a machine learning model based on long short-term memory (LSTM) that proved effective to consistently forecast the DPC using data from Japan. Using three cofactors (human mobility, metrological factors, and day of the week), we were able to estimate the DPC two weeks into the future with an 81.6% accuracy for six prefectures in Japan [11]. Human mobility was the dominant factor, especially at transit stations [11]. Some studies have applied mathematical or agent-based models for scenario-based analysis. Unlike previous studies, the aim of this study was to replicate past phenomena using a machine learning approach. The machine learning approach is more straightforward for such an application than the agent-based approach or the SEIR (Susceptible-Exposed-Infectious-Removed) model, in part because it is potentially case-sensitive for input parameters, e.g., sensitivity of initial conditions.

In this study, we numerically replicated new DPCs from July to September 2021 in three prefectures of Japan with a machine learning approach to provide insights concerning the effect of the Tokyo Olympic Games on COVID-19 spread. Dominant factors causing the up- and downsurge pandemic infections from July to September 2021 are discussed.

2. Related studies

Several groups have developed various forecasting models and analyzed different factors. We briefly reviewed studies based on machine learning, especially LSTM architecture. In Ref. [19], the error percentage of DPCs in India was less than 20% after five days. Several neural networks were considered for estimating the number of DPCs in different cities of India [20]; testing results showed an error of 3%–5% for one–three-day forecasting. That study compared different types of LSTM and found that bi-directional LSTM performed better than other models. A study conducted using the cumulative number of confirmed cases in Isfahan, Iran was used to test different machine learning forecasting models [21]. The input data included DPC and social determinants of health. The long-term prediction (i.e., more than a few months) had substantial error; the peak was predicted 1.5 months earlier. A rough estimate of the one-month prediction error is in the order of 20%–30% (from Fig. 6 in Ref. [21]). Data from Russia, Peru, and Iran obtained from January to July 2020 were used to validate a standard LSTM network [22]. The difference between the predicted and actual error was <12.8% for a three-day prediction. Another study used data from Canada to provide forecasts of different sets of future days and predicted the end of the COVID-19 outbreak to occur around June 2020 [23]; the accuracy of two-week estimation was 93.4%. In another study, data obtained from different European countries were compared using different models to forecast the two-week DPC [24]. Results indicate that LSTM provides more accurate estimations, and are consistent with those of another comparison study using data from 10 countries for a 48-day DPC prediction [25]. Different versions of LSTM were used to forecast different COVID-19 incidents for 30, 60, and 90 days using data from India [26]. Results indicated that stacked LSTM exhibited superior performance. Table S1 shows a brief comparison of related studies.

In our previous study [11], the proposed framework provided more accurate and consistent estimations than those provided by Google Cloud for four-week estimation. The two-week projected DPC can be estimated with an 81.6% accuracy for six prefectures in Japan, which represent a range from 0.18 (best case) to 0.75 (worst case) of the average relative error of Google Cloud forecasting.

Overall, a straightforward comparison is not feasible because different studies use different metrics to compute the error; in some studies, errors were computed for cumulative data (not daily new cases).

3. Method and materials

3.1. Machine learning-based predictions and interpretation

Forecasting data of DPC is needed for comparison to confirm consistency. Recently, we developed a deep LSTM neural network that was proven to efficiently forecast DPC using a mixture of time series data including current status, meterological data, and mobility estimates [11]. The model consists of a multi-path LSTM neural network, followed by a fully-connected layer that is optimized to forecast future time-sequence datasets of DPC using time-sequence data. The multi-path LSTM connection considers bi-directional data alignment to empower data pattern extraction. A fully-connected linear layers of 600, 300, and 100 neurons flows into the LSTM network and ends with an output layer of 14 neurons (forecasting two weeks into the future). The training utilized cross-entropy cost function and the ADAM optimizer. This method is consistent with those of other studies [24,25], which demonstrated the superiority of the LSTM architecture in forecasting DPC using data from different countries.

We set input/output time frames to 14 days and included labels that define working days and holidays, state of emergency/normal conditions, and the dominant viral variant type. We used training data from August 1, 2020 to one day before the start date of forecasting (July 22, 2021 in Figs. 3 and 5, September 1, 2021 in Fig. 4). Review and comparison with other related methods can be found in Refs. [11,26]. The model's major principle is that current measurements with a short time period (two weeks) can be efficiently used to predict a future short time period (two weeks) using a set of LSTMs and a fully-connected layer. The design of the LSTM layer efficiently recognizes the time series data and the contribution of different inputs to the target output (the two-week predicted DPC). Different data paths with different sequence directions improve the prediction accuracy. Using a long training dataset period, spread- and decay-pandemic phases are efficiently recognized by the network. The fully-connected layer includes additional training parameters useful in handling training bias and data noise. An open-source software is available at https://github.com/erashed/CovidNet.

3.2. Vaccination rate and effectiveness at the population level

Vaccination effectiveness at the population level in each prefecture was used to approximate herd immunity, which is assumed as follows:

$$E(d) = \sum_{t=1}^{T} \sum_{i=0}^{d} (N_t(d-i) \cdot e_t(i)) \Big/ P, \tag{1}$$

$$e_t(i) = \begin{cases} a_t \cdot i/K (i \le 14), \\ a_t - s(i-14)(i > 14), \end{cases}$$

where d is day and N_t denotes the number of people who take newly-administered vaccination shot dose t. The variable e denotes the vaccination effect i days after inoculation; parameters a, b, and s are adjusted to reach a peak 14 days after inoculation, subsequently decreasing linearly. The variable P denotes the deemed population, expressed as the summation of the population of the entire prefecture and cumulative number of the second and third doses. The values of the parameters a_1 and a_2 for the Delta variant were defined as 0.605 and 0.756, respectively, based on a meta-analysis of systematic review (11 study groups) [27]. The parameter a_3 was defined as 0.956 based on information provided in the press release by Pfizer and BioNTech [28]. Note that the vaccination rate of BNT162b2 in Japan is approximately 90%. Furthermore, based on the waning immunity for infection protection, s was defined as 0.24 to match the model [29,30].

3.3. Estimation of compensated ERN with regression line

To validate the main factor triggering the increase in DPC during the

Olympic Games, we first studied the relationship between public mobility change and the effective reproduction number (ERN) in Tokyo, Osaka, and Aichi. We found a consistent linear correlation in most cases. Therefore, it becomes feasible to recompute ERN on specific days (national holidays) to eliminate the effect of public mobility change during holidays in DPC upsurges. The computation was performed for different regions to highlight the observed effect of the Olympic Games hosted in Tokyo but not observed in other regions.

3.4. Data

Data from three prefectures of Japan (Tokyo, Osaka, and Aichi) were used; these prefectures are the central districts in different areas, and are distant from each other. Thus, the mobility change attributable to the Olympic Games may be marginal in Osaka and Aichi. Data regarding the recorded COVID-19 DPCs were obtained by a dedicated online service provided by the Japanese Ministry of Health, Labor, and Welfare (https://www.mhlw.go.jp/stf/covid-19/open-data.html) (accessed on December 24, 2021) with local district websites.

From the viewpoint of mathematical representation of epidemics, ERN, which corresponds to the rate of change of new cases, is essential data rather than the new DPC itself. Different definitions of ERN, also defined as R_t, can be found [31–33]. ERN R_t was computed using the following equation:

$$R_t = \left(\sum_{i=1}^{s} DPC_{t-i} \Big/ \sum_{i=s+1}^{2s} DPC_{t-i} \right)^{\mu/s}, \tag{2}$$

where $s = 7$ means the number of days during a specified temporal period, and $\mu = 5$ (days) denotes the mean latency after COVID-19 infection.

Data on the weather conditions recorded at the main cities of interest within the target regions of this study were obtained from the Japan Meteorological Agency (https://www.jma.go.jp/jma/indexe.html) (accessed on December 27, 2021). Data from the largest cities (Tokyo, Osaka, and Aichi) were used as representative for the whole prefecture.

We also used data on public movements recorded by Google mobility reports from February 15, 2020 [34]. The mobility changes are labeled as retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residential areas, and provided as the percentage of change compared to day-of week mean values (baseline) from the five-week period comprising January 3 to February 6, 2020. The last data points corresponded to the baseline data in this study.

Fig. 1 shows the Google mobility data and DPC in Tokyo, Osaka, and Aichi for the summers of 2020 and 2021. The mobility effect was a prevalent factor in viral diffusion as a replacement for the degree of social distancing. In order to use the mobility as an indicator of viral diffusion, the impact of the incubation period must be considered. We defined the effective mobility values as an eight-day average, and considered mobility to have a five-day stride, i.e., 6–13 days before the corresponding day. Although not shown here, our results show that small shifts in this parameter do not significantly influence the correlation with R values [35].

Vaccination rates were obtained from the Government of the Chief Information Officers Portal, Japan (https://cio.go.jp/c19vaccine_opendata).

The time course of new DPCs is plotted in Fig. 1, along with the relationship between the time-averaged mobility of different zoning categories in the aforementioned prefectures.

3.5. Scenarios

The replication of DPC was conducted using two scenarios. First, the DPCs was estimated through the adjustment of the ERN such that the

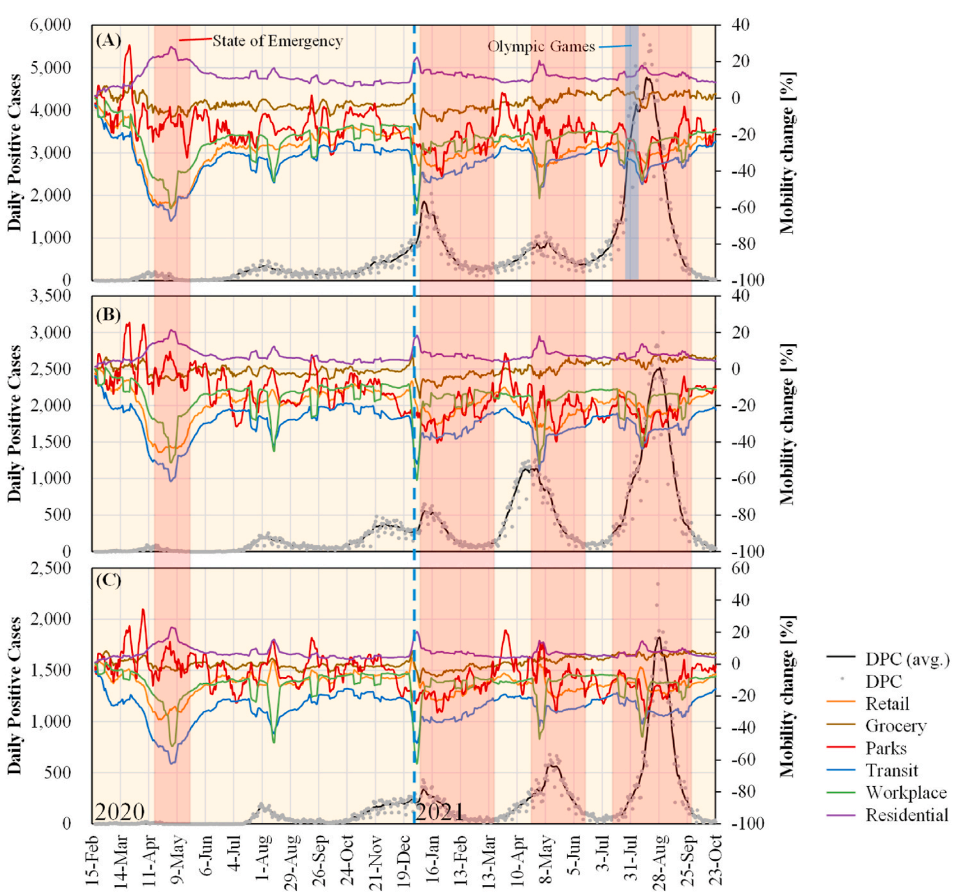

Fig. 1. Daily positive cases (DPC) and Google mobility data of different categories for (A) Tokyo, (B) Osaka, and (C) Aichi from February 15, 2020 to October 23, 2021. Black lines represent a seven-day average of DPC for raw data (dots). The peak values of DPC (seven-day average) in the 3rd, 4th, and 5th waves in Tokyo were 1861 cases, 934 cases, and 4774 cases, respectively, on January 8, May 10, and August 16, 2021. The peak DPCs in Osaka were 553 cases, 1134 cases, and 2518 cases on January 8, April 29, and August 29, for the 3rd, 4th, and 5th waves, respectively, whereas they were 339 cases, 575 cases, and 1822 cases on January 8, May 13, and August 29 for Aichi, respectively. A time lag of up to two weeks can be observed in different prefectures.

influence of public movements on national holidays was excluded. The adjusted ERN estimated from the regression line represents the relationship between the mobility change rate at transit stations and their corresponding ERN. Second, the machine learning model based on LSTM was trained using meteorological data, mobility estimates, day labels, and corresponding DPC number. Training data comprised August 1, 2020 to July 22, 2021, and validation was considered from July 23, 2021.

4. Results

4.1. Relationship between mobility and ERN

Fig. 2 shows the correlation between time-averaged mobility at transit stations and the ERN in the three prefectures. The determination coefficient and parameters of the regression lines of the three areas are listed in Table 1. Statistical significance was accepted at $p < 0.05$. From Fig. 2, similar singular behaviors were observed from August 1–4, 2021. During this period, the ERN increased despite a decrease in the mobility change rate. When considering the virus incubation period of 5–6 days [36,37], including a 7–10 day lag before going to hospital, this singular behavior approximately corresponded to the period of July 22–27, 2021, and matches the national holidays (July 22–23, 2021) and weekend (July 24–25, 2021), the dates of which were adjusted for the opening of the Olympic Games. However, this ERN upsurge becomes modest around August 5, 2021 and buried in the plots on the regression line of the fifth wave. In addition, the "Obon" holiday took place from August 13–15, 2021, and was somewhat extendable from August 7, 2021 (Saturday) to match the period of the Olympic Games; August 8, 2021 was changed to a national holiday to accommodate the Olympic Games. A similar tendency of ERN upsurge was observed even for this period. The ERN during the above holidays was 30%–50% higher than the regression line to characterize the relationship between ERN and mobility at the transit stations.

4.2. Effect of national holidays and vaccination on new DPC

The aforementioned observation highlighted that national holidays may have contributed to the DPC upsurge. To investigate the impact of the abovementioned holidays, the ERN was adjusted in these periods. As shown in Fig. 3 A, the ERN of spread duration from August 1–4 and August 16–22, 2021 were determined using the regression line. To maintain consistency, the ERN were also adjusted for the three days before and after these periods. Fig. 3 A–C shows that in all three prefectures, new DPCs decreased after adjusting for the effect of holidays. Peak new DPCs decreased by 53.4%, 33.8%, and 60.4% in Tokyo, Osaka, and Aichi, respectively. In the same figure, the estimation using machine learning was also plotted. The estimated values were consistent with the two adjusted curves, whereas the observed values were higher.

Fig. 4 shows that some discrepancies between the observed and estimated values were observed in September. In this figure, the observed mobility was used, whereas the remaining parameters were the observed values (before August 31, 2021). The estimated values in Fig. 4 were consistent with measured values, in contrast to those from mid-July.

4.3. Effect of vaccination effectiveness and mobility on ERN

Fig. 5 shows the estimated new DPCs when the vaccination effectiveness at the population level was maintained at a constant level after

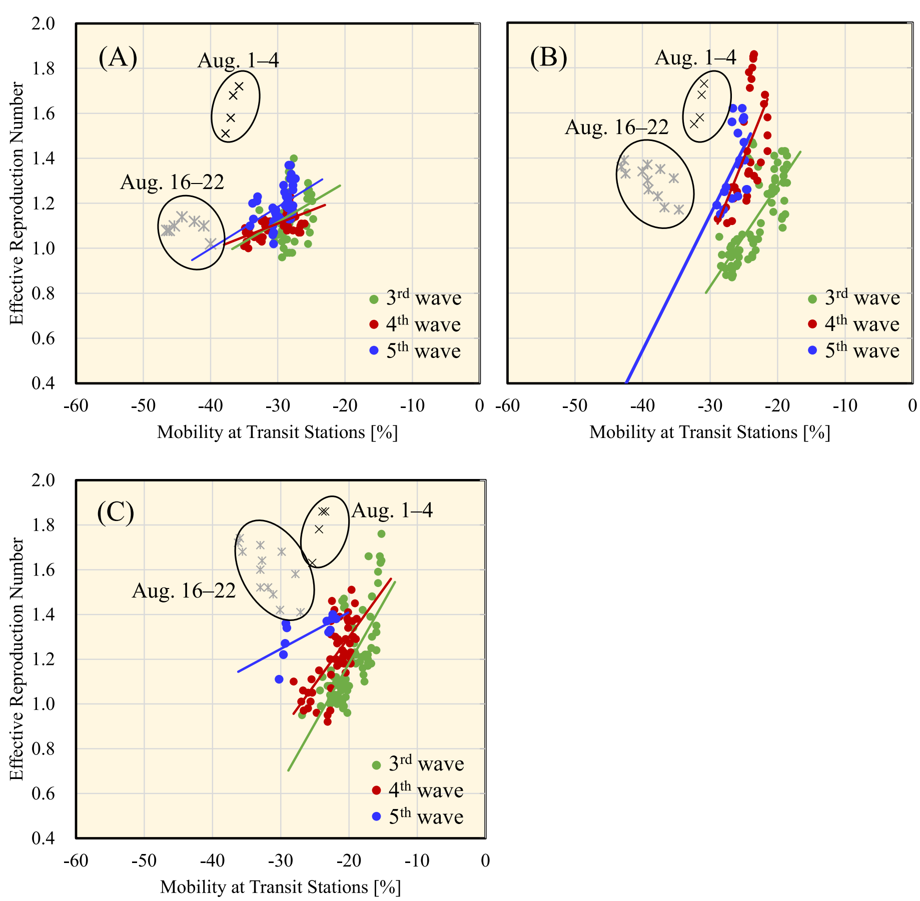

Fig. 2. Relationship between the seven-day averaged mobility (with latency of seven days) at transit stations and effective reproduction number in (A) Tokyo, (B) Osaka, and (C) Aichi. To derive the regression line for 5th wave, days corresponding of the national holidays were excluded.

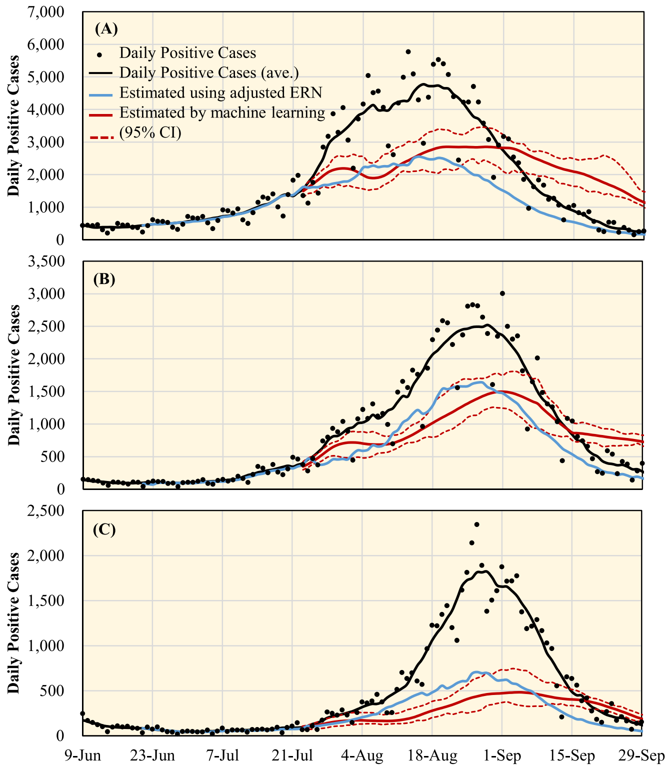

Fig. 3. Daily positive cases (seven-day average) estimated using the adjusted effective reproduction number (ERN) during national holidays and machine learning in (A) Tokyo, (B) Osaka, and (C) Aichi. The ERN during national holidays was adjusted by substituting the observed mobility into the regression line defined in Fig. 2 and Table 1.

July 20, 2021. In this figure, new DPCs increased and reached 9400 cases on September 4, 2021 (95% CI; 6000–13,400 cases). Compared with the observed values, this estimation is approximately twice the observed DPCs, and its peak was two weeks behind the observed value.

Similarly, when the mobility was set to 0% after July 23, 2021 (i.e., the same as pre-pandemic mobility), the DPC was also replicated with machine learning. The new DPC reached 3700 cases on August 19, 2021 (95% CI; 2800–3900 cases). This peak value was 38% higher than the estimation with adjusted ERN excluding the effect of holidays.

5. Discussion

In this study, we analyzed DPC data in three major prefectures in Japan from July to September 2021. A major motivation was the Olympic Games, held from July 23 to August 8, 2021. It was crucial to analyze this situation for future international events, and to consider its impact on economics in making policy decisions [38,39]. Similarly to a previous study, we conducted a scenario-based analysis [14] with a LSTM model.

5.1. Spread phase

As shown in Fig. 1, a similar trend of time course of DPC and mobility can be found in the three different prefectures of Japan. Moreover, the relative (normalized) peak for DPC demonstrates a different trend among different prefectures, e.g., the peak of the fourth wave in Tokyo is relatively lower than that of other prefectures.

From Fig. 2, the ERN was 20%–80% higher than the regression line during two holiday periods, which is attributable to human behavior. Note that this enhancement was not significant during the Obon holiday (August 13–15, 2021) in Tokyo, whereas it was notable in Osaka and Aichi, suggesting that social behavior differs between prefectures.

To confirm its impact on the upsurge, DPC was numerically replicated in terms of compensated ERN, assuming that there were no national holidays during the period. As shown in Fig. 3, the estimated new DPC was smaller than the actual DPC value. Once the difference in ERN was compensated for, the peak DPCs were suppressed by 30%–60%. The point to be emphasized here is that an upshift of the ERN was observed not only in Tokyo but also in Osaka and Aichi, and only during the holidays, suggesting that this upsurge may not be fully-attributable to the Olympic Games. Instead, some holiday dates were adjusted to the Olympic Games; thus, the contribution may not be neglected from the viewpoint of human behavior. Note that the effect of imported new cases related the Games may not be significant. The number of positive cases linked to the Olympics was 320 cases (from July 1, 2021), which is smaller than the new DPC (a few thousand cases daily) [40].

5.2. Superspreading during the spread phase

A similar upsurge was observed in the fourth wave, which appeared towards the end of December and New Year Day; at that time the ERN was 20%–60% higher than the regression line. This upsurge was not observed in the holidays during the earlier (1st to 3rd) waves, as people were more self-constraining [41]. Model accuracy marginally improved (less than 10%) for the training period after the fourth wave (i.e., early 2021); only one period for national holiday. If the training period was from the third wave, this improvement was not observed. This data insufficiency regarding the holidays prevents the improvement of accuracy of the LSTM method. One potential reason for this upsurge during the holidays after the fourth wave may be attributable to the frequency of human contact in different communities. In Ref. [42], it has been modified to represent the social behaviors of people where the generated communities are restricted and reflect spatiotemporal constraints in real life. They concluded that the bridge between infected and fresh clusters may trigger new virus spread. This may happen especially during holidays of Japan because unlike in other countries, holidays are usually simultaneously taken all at once by the general population [43]. Even in other countries, the ERN has been influenced by policy, holidays, etc., and consistent with that in Japan [44,45], which changed time by time, and thus this topic remains still open question to be resolved.

5.3. Decay phase

New DPCs suddenly decreased in September, sparking debate as to its cause. Our machine learning estimation suggests that this decrease was mainly due to vaccination, which had been increasing linearly (Fig. 4 A), and is consistent with forecasting in the end of August (Fig. 4 B–D). Thus, the abrupt decay, reported by newspapers as "mysterious" [46], was mainly attributable to two factors: (1) vaccination, and (2) the decrease of the ERN, which had temporarily increased due to the holidays. The latter factor vanished with latency immediately post-holidays. However, some discrepancy still exists, which may be attributable to the age category (i.e., vaccination strategy first focused on the elderly), the age-dependent effect, and cluster infections [47].

Asymptomatic infected people may be included in the machine learning model without defining them explicitly. The number of asymptomatic infected people is approximately four times larger than those with symptoms [48]. The symptomatic infected people is 1% of the total population, whereas 4% population who obtained antibodies from asymptomatic infections. This percentage is lower than the vaccination rate of 30%. Considering the incubation lag period, the decay timing also coincided with a 10%–15% fully-vaccinated rate in people aged <65 years. Vaccination prevalence among the more behaviorally-active generation (aged <65 years) may have influenced the decrease in DPC (see Figure A1).

From Fig. 5, the decrease of the ERN due to mobility and vaccination in the downsurge is associated by 38% and 253%, respectively. The

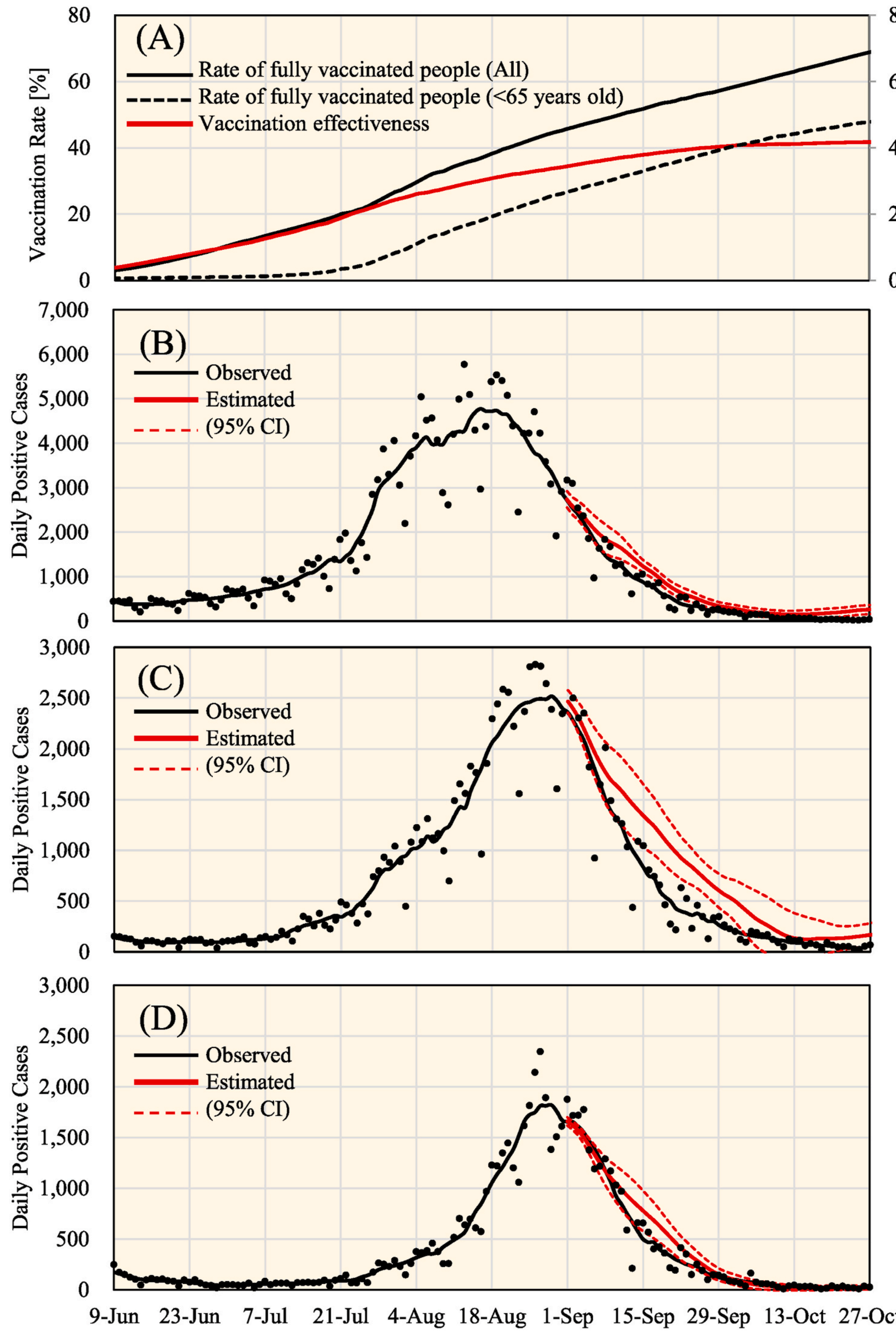

Fig. 4. (A) Vaccination rate and vaccination effectiveness in Japan. For vaccination effectiveness, daily positive cases were estimated using machine learning in (B) Tokyo, (C) Osaka, and (D) Aichi. The dashed lines show the variation in each estimation.

decrease of the ERN from 1.2 to 0.8 (averaged from August 20 to September 20, 2021, respectively) can be mainly explained by these two factors (Figure A1). The impact of weather on COVID-19 up- and downsurges from July to September 2021 in Japan (not shown here to avoid repetition) was <2%. This is because that the change in weather during the up- and downsurge was marginal due to its sufficiently short duration.

6. Summary

We applied a machine learning approach to replicate the DPC in three prefectures of Japan from July to September 2021, during which the Olympic Games were held. Our computational data suggests no clear evidence that the Games directly influenced the upsurge of the fifth COVID-19 infection wave because a similar tendency was observed in Osaka and Aichi, which are distant from Tokyo. Instead, the shift in holiday dates, designated to match the Olympic Games, likely contributed to the upsurge. Such superspreading events or phenomena cannot be resolved because the tendencies were different from those of earlier waves. Except for this factor, the machine learning approach would be a promising tool for estimation and management of future events during a pandemic.

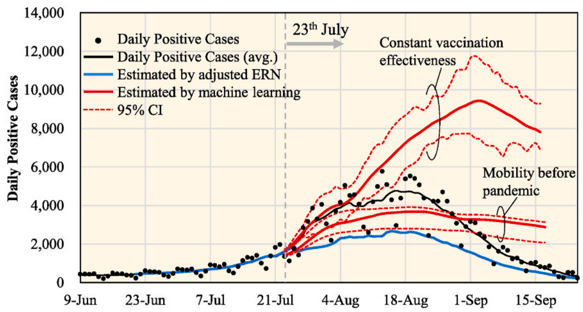

Fig. 5. Daily positive cases were replicated with machine learning in Tokyo after July 23, 2021for cases where vaccination was not conducted and mobility was assumed to be identical to that before the pandemic. The impact of vaccination is larger than that of the mobility.

Table 1
Parameters of regression line ($y = ax + b$) in Fig. 2 and corresponding coefficients of determination for different COVID-19 waves.

	Waves	a	b	R^2
Tokyo	3rd	0.0177	1.65	0.0774 ($p = 0{\cdot}035$)
	4th	0.0117	1.46	0.377 ($p < 0{\cdot}0001$)
	5th	0.0185	1.74	0.177 ($p = 0{\cdot}006$)
Osaka	3rd	0.0448	2.17	0.587 ($p < 0.0001$)
	4th	0.0751	3.28	0.353 ($p = 0.0004$)
	5th	0.0601	2.95	0.258 ($p = 0.026$)
Aichi	3rd	0.0539	2.26	0.513 ($p < 0.0001$)
	4th	0.0423	2.15	0.451 ($p < 0.0001$)
	5th	0.0164	1.74	0.473 ($p = 0.019$)

Contributors

AH, ER, SK, and YD conceived this study. YD and SK contributed data curation. ER and SK analyzed the data. AH and YD verified all the data. AH led the writing of the paper. All authors commented on the paper, and edited the final manuscript. All authors gave final approval of the version for submission.

Data statement

The main part of the code (LSTM) used in this study is made publicly available (https://github.com/erashed/CovidNet). All data processed in this study will be available with reasonable request, ending 5 years following article publication.

Declaration of competing interest

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

Acknowledgments

This research was conducted as part of “Covid-19 AI & Simulation Project” run by Mitsubishi Research Institute commissioned by Cabinet Secretariat; AH received general funding from Ministry of Education, Sports, Arts, Science and Technology, Japan, as a director of Center of Biomedical Physics and Information Technology, Nagoya Institute of Technology, to conduct research related to public health engineering. The funders of the study and the membership of Japanese national committees had no role in study design; data collection, analysis, and interpretation; preparation of the manuscript.

Preliminary results in this study have been presented in the meeting of COVID-19 AI & Simulation Project under Cabinet Secretariat, Japan (October 19, 2021) and Advisory Board Meeting of Ministry of Health and Welfare, Japan (November 9, 2021).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2022.105548.

Appendix

To clarify how much vaccination effectiveness was needed to reduce infection, we investigated the correlation between the population vaccination effectiveness and the ERN. It is difficult to make direct comparisons because of the different day-to-day mobilities. Therefore, the ERN was adjusted to match −25% based on the slope of the regression line obtained in Fig. 2.

Figure A1 shows the correlation between vaccination effectiveness and adjusted ERN. The correlation during the national holiday is different from that on other days, as shown in Fig. 2. The adjusted ERN without the national holidays starts to decrease when the population vaccination effectiveness $<0{\cdot}3$. The rate of fully vaccine people at that time was 40%–45% for the population of all age and 20%–30% for those aged <65 years. In addition, the vaccination effectiveness was 0.33–0.35 when the ERN is less than 1, i.e., the number of new positive cases began to decrease.

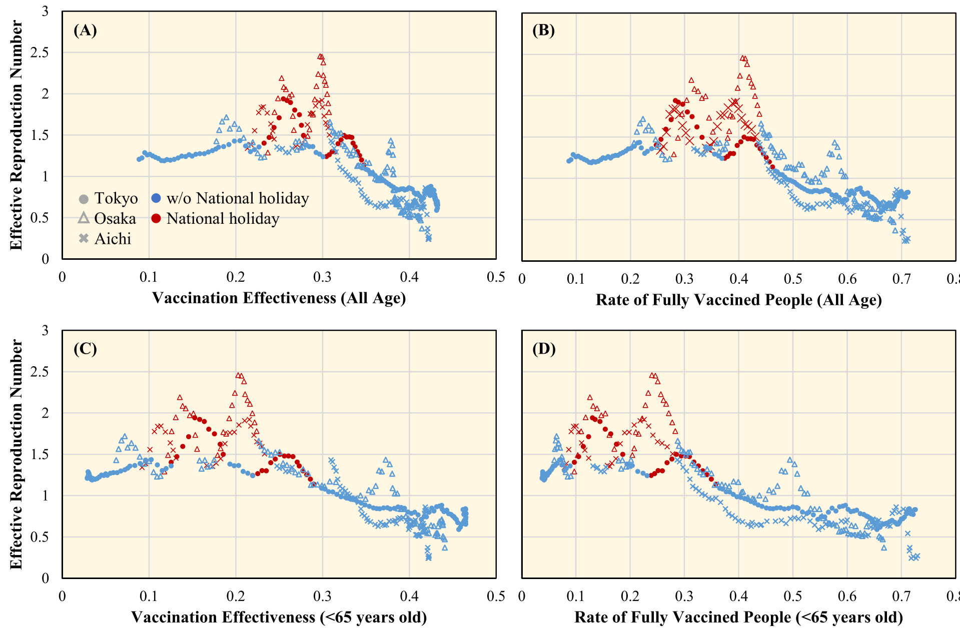

Fig. A1. Correlation between effective reproduction number and (A, C) vaccination effectiveness and (B, D) rate of fully vaccine people in three prefectures. Vaccine effectiveness and rate of fully vaccinated people are shown for (A, B) all ages and (C, D) < 65 years. Similar threshold of population vaccination effectiveness to suppress the effective reproduction number was observed for three prefectures.

References

[1] K. Shimizu, D. Sridhar, K. Taniguchi, K. Shibuya, Reconsider this summer's Olympic and Paralympic games, Br. Med. J. (2021) n962.

[2] V.T. Hoang, J.A. Al-Tawfiq, P. Gautret, The Tokyo olympic games and the risk of COVID-19, Curr. Trpic. Med. Rep. 7 (2020) 126–132.

[3] K. Taku, H. Arai, Impact of COVID-19 on athletes and coaches, and their values in Japan: repercussions of postponing the Tokyo 2020 olympic and paralympic games, J. Loss Trauma 25 (2020) 623–630.

[4] V. Gallego, H. Nishiura, R. Sah, A.J. Rodriguez-Morales, The COVID-19 outbreak and implications for the Tokyo 2020 summer olympic games, Trav. Med. Infect. Dis. 34 (2020), 101604.

[5] Reuter, Factbox: Coronavirus Cases at the Tokyo Olympics, 2021.

[6] M. Komarov, Tokyo, Olympic Games: the Economic Impact of the COVID-19 Pandemic, Problemy Dalnego Vostoka, 2020, pp. 49–59 (2021).

[7] Y. Miyata, N. Hasegawa, H. Nagase, N. Kutsumura, R. Nagatomi, K. Nakamura, T. Majima, Medical and social countermeasures for Tokyo olympic in 2021, J. Phys. Fitness. Sport. Med. 9 (2020) 344.

[8] K. Shimizu, S. Gilmour, H. Mase, P.M. Le, A. Teshima, H. Sakamoto, S. Nomura, COVID-19 and heat illness in Tokyo, Japan: implications for the summer olympic and paralympic games in 2021, Int. J. Environ. Res. Publ. Health 18 (2021) 3620.

[9] E.J. Emanuel, G. Persad, R. Upshur, B. Thome, M. Parker, A. Glickman, C. Zhang, C. Boyle, M. Smith, J.P. Phillips, Fair Allocation of Scarce Medical Resources in the Time of Covid-19, Mass Medical Soc, 2020.

[10] E.A. Rashed, A. Hirata, Infectivity Upsurge by COVID-19 viral variants in Japan: evidence from deep learning modeling, Int. J. Environ. Res. Publ. Health 18 (2021) 7799.

[11] E.A. Rashed, A. Hirata, One-year lesson: machine learning prediction of COVID-19 positive cases with meteorological data and mobility estimate in Japan, Int. J. Environ. Res. Publ. Health 18 (2021) 5736.

[12] Our World in Data, Coronavirus (COVID-19) Vaccinations, Availabe online: https://ourworldindata.org/covid-vaccinations.

[13] E. Cuevas, An agent-based model to evaluate the COVID-19 transmission risks in facilities, Comput. Biol. Med. 121 (2020), 103827.

[14] A.M. Salman, I. Ahmed, M.H. Mohd, M.S. Jamiluddin, M.A. Dheyab, Scenario analysis of COVID-19 transmission dynamics in Malaysia with the possibility of reinfection and limited medical resources scenarios, Comput. Biol. Med. 133 (2021), 104372.

[15] R. Zafarnejad, P.M. Griffin, Assessing school-based policy actions for COVID-19: an agent-based analysis of incremental infection risk, Comput. Biol. Med. 134 (2021), 104518.

[16] B.M. Castro, Y.d.A. de Melo, N.F. Dos Santos, A.L. da Costa Barcellos, R. Choren, R. M. Salles, Multi-agent simulation model for the evaluation of COVID-19 transmission, Comput. Biol. Med. 136 (2021), 104645.

[17] S. Chowdhury, M. Forkan, S.F. Ahmed, P. Agarwal, A.S. Ali, S. Muyeen, Modeling the SARS-CoV-2 parallel transmission dynamics: asymptomatic and symptomatic pathways, Comput. Biol. Med. (2022), 105264.

[18] Y. Quintero, D. Ardila, E. Camargo, F. Rivas, J. Aguilar, Machine learning models for the prediction of the SEIRD variables for the COVID-19 pandemic based on a deep dependence analysis of variables, Comput. Biol. Med. 134 (2021), 104500.

[19] A. Tomar, N. Gupta, Prediction for the spread of COVID-19 in India and effectiveness of preventive measures, Sci. Total Environ. (2020), 138762.

[20] P. Arora, H. Kumar, B.K. Panigrahi, Prediction and analysis of COVID-19 positive cases using deep learning models: a descriptive case study of India, Chaos, Soliton. Fractal. 139 (2020), 110017.

[21] R. Kafieh, N. Saeedizadeh, R. Arian, Z. Amini, N.D. Serej, A. Vaezi, S.H. Javanmard, Isfahan and COVID-19: deep spatiotemporal representation, chaos, Soliton. Fractal. 141 (2020), 110339.

[22] P. Wang, X. Zheng, G. Ai, D. Liu, B. Zhu, Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: case studies in Russia, Peru and Iran, Chaos, Soliton. Fractal. 140 (2020), 110214.

[23] V.K.R. Chimmula, L. Zhang, Time series forecasting of COVID-19 transmission in Canada using LSTM networks, Chaos, Soliton. Fractal. 135 (2020), 109864.

[24] İ. Kırbaş, A. Sözen, A.D. Tuncer, F.Ş. Kazancıoğlu, Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches, Chaos, Soliton. Fractal. 138 (2020), 110015.

[25] F. Shahid, A. Zameer, M. Muneeb, Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM, Chaos, Soliton. Fractal. 140 (2020), 110212.

[26] J. Devaraj, R.M. Elavarasan, R. Pugazhendhi, G. Shafiullah, S. Ganesan, A. K. Jeysree, I.A. Khan, E. Hossain, Forecasting of COVID-19 cases using deep learning models: is it reliable and practically significant? Results Phys. 21 (2021), 103817.

[27] B. Zeng, L. Gao, Q. Zhou, K. Yu, F. Sun, Effectiveness of COVID-19 Vaccines against SARS-CoV-2 Variants of Concern: a Systematic Review and Meta-Analysis, medRxiv, 2021.

[28] Pfizer and BioNTech, Pfizer and BioNTech Announce Phase 3 Trial Data Showing High Efficacy of a Booster Dose of Their COVID-19 Vaccine, 2021.

[29] S.Y. Tartof, J.M. Slezak, H. Fischer, V. Hong, B.K. Ackerson, O.N. Ranasinghe, T. B. Frankland, O.A. Ogun, J.M. Zamparo, S. Gray, Six-month Effectiveness of BNT162B2 mRNA COVID-19 Vaccine in a Large US Integrated Health System: a Retrospective Cohort Study, 2021. NCT04848584.

[30] S. Kodera, E.A. Rashed, A. Hirata, Estimation of real-world vaccination effectiveness of mRNA COVID-19 vaccines against Delta and omicron variants in Japan, Vaccines 10 (2022) 430.

[31] J. Wallinga, M. Lipsitch, How generation intervals shape the relationship between growth rates and reproductive numbers, Proc. Biol. Sci. 274 (2007) 599–604.

[32] T. Obadia, R. Haneef, P.-Y. Boëlle, The R0 package: a toolbox to estimate reproduction numbers for epidemic outbreaks, BMC Med. Inf. Decis. Making 12 (2012) 1–9.

[33] J. Ma, Estimating epidemic exponential growth rate and basic reproduction number, Infect. Dis. Model. 5 (2020) 129–141.

[34] Google, Google Mobility, Availabe online: https://www.google.com/covid19/mobility/.

[35] S. Kodera, E.A. Rashed, A. Hirata, Mobility-dependent and Mobility-Compensated Effective Reproduction Number of COVID-19 Viral Variants: New Metric for Infectivity Evaluation, arXiv, 2021, 2111.02961.
[36] S.A. Lauer, K.H. Grantz, Q. Bi, F.K. Jones, Q. Zheng, H.R. Meredith, A.S. Azman, N. G. Reich, J. Lessler, The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application, Ann. Intern. Med. 172 (2020) 577–582.
[37] C. McAloon, Á. Collins, K. Hunt, A. Barber, A.W. Byrne, F. Butler, M. Casey, J. Griffin, E. Lane, D. McEvoy, Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research, BMJ Open 10 (2020), e039652.
[38] D. Fujii, T. Nakata, COVID-19 and Output in Japan, Center for Advanced Research in Finance, Faculty of Economi, University of Tokyo, 2021. CARF-F-505.
[39] H. Inoue, Y. Todo, The propagation of economic impacts through supply chains: the case of a mega-city lockdown to prevent the spread of COVID-19, PLoS One 15 (2020), e0239251.
[40] BBC News, Tokyo Olympic Games: when Are They and How Will Covid Affect Them?, 23 July 2021.
[41] T. Kuniya, Evaluation of the effect of the state of emergency for the first wave of COVID-19 in Japan, Infect. Dis. Model. 5 (2020) 580–587.
[42] Y. Ohsawa, M. Tsubokura, Stay with your community: bridges between clusters trigger expansion of COVID-19, PLoS One 15 (2020), e0242766.
[43] Y. Ohe, Characteristics and issues of rural tourism in Japan, WIT Trans. Ecol. Environ. (2008) 115.
[44] L. Casini, M. Roccetti, A cross-regional analysis of the COVID-19 spread during the 2020 Italian vacation period: results from three computational models are compared, Sensors 20 (2020) 7319.
[45] L. Santamaría, J. Hortal, COVID-19 effective reproduction number dropped during Spain's nationwide dropdown, then spiked at lower-incidence regions, Sci. Total Environ. 751 (2021), 142257.
[46] Mainichi Newspaper, Why Have New COVID Cases Declined So Quickly in Japan, and Why Is Caution Needed?, 2021.
[47] Y. Furuse, E. Sando, N. Tsuchiya, R. Miyahara, I. Yasuda, Y.K. Ko, M. Saito, K. Morimoto, T. Imamura, Y. Shobugawa, Clusters of coronavirus disease in communities, Japan, January–April 2020, Emerg. Infect. Dis. 26 (2020) 2176.
[48] T. Sanada, T. Honda, F. Yasui, K. Yamaji, T. Munakata, N. Yamamoto, M. Kurano, Y. Matsumoto, R. Kohno, S. Toyama, Serologic survey of IgG against SARS-CoV-2 among hospital visitors without a history of SARS-CoV-2 infection in Tokyo, 2020-2021, J. Epidemiol. (2021), JE20210324.