

RESEARCH

Open Access



# Simple mathematical model for predicting COVID-19 outbreaks in Japan based on epidemic waves with a cyclical trend

Hiroki Manabe<sup>1\*</sup>, Toshie Manabe<sup>2,3</sup>, Yuki Honda<sup>1</sup>, Yoshihiro Kawade<sup>3</sup>, Dan Kambayashi<sup>3,4</sup>, Yoshiki Manabe<sup>5</sup> and Koichiro Kudo<sup>6,7</sup>

## Abstract

**Background** Several models have been used to predict outbreaks during the COVID-19 pandemic, with limited success. We developed a simple mathematical model to accurately predict future epidemic waves.

**Methods** We used data from the Ministry of Health, Labour and Welfare of Japan for newly confirmed COVID-19 cases. COVID-19 case data were summarized as weekly data, and epidemic waves were visualized and identified. The periodicity of COVID-19 in each prefecture of Japan was confirmed using time-series analysis and the autocorrelation coefficient, which was used to investigate the longer-term pattern of COVID-19 cases. Outcomes using the autocorrelation coefficient were visualized via a correlogram to capture the periodicity of the data. An algorithm for a simple prediction model of the seventh COVID-19 wave in Japan comprised three steps. Step 1: machine learning techniques were used to depict the regression lines for each epidemic wave, denoting the “rising trend line”; Step 2: an exponential function with good fit was identified from data of rising straight lines up to the sixth wave, and the timing of the rise of the seventh wave and speed of its spread were calculated; Step 3: a logistic function was created using the values calculated in Step 2 as coefficients to predict the seventh wave. The accuracy of the model in predicting the seventh wave was confirmed using data up to the sixth wave.

**Results** Up to March 31, 2023, the correlation coefficient value was approximately 0.5, indicating significant periodicity. The spread of COVID-19 in Japan was repeated in a cycle of approximately 140 days. Although there was a slight lag in the starting and peak times in our predicted seventh wave compared with the actual epidemic, our developed prediction model had a fairly high degree of accuracy.

**Conclusion** Our newly developed prediction model based on the rising trend line could predict COVID-19 outbreaks up to a few months in advance with high accuracy. The findings of the present study warrant further investigation regarding application to emerging infectious diseases other than COVID-19 in which the epidemic wave has high periodicity.

**Keywords** COVID-19, SARS-CoV-2, Prediction model, Mathematical model, Machine learning technique, Periodicity, Rising trend line, Epidemic wave, Autocorrelation coefficient, Correlogram

\*Correspondence:

Hiroki Manabe  
manabe@shitennoji.ac.jp

Full list of author information is available at the end of the article



© The Author(s) 2024. **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/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Following reports of the first patients with viral pneumonia caused by COVID-19 in December 2019 from Wuhan City in Hubei Province, China, the disease rapidly spread throughout the world, leading to the COVID-19 pandemic [1, 2]. In Japan, the first patient with COVID-19 was reported on January 15, 2020. This was the second imported case from China, following those from Thailand [3, 4]. Since then, people in Japan have experienced eight epidemic waves of COVID-19, with approximately 33.5 million laboratory-confirmed cases and 74,000 deaths as of March 31, 2023 [5]. Each time an epidemic wave occurs, the number of laboratory-confirmed COVID-19 cases increases rapidly. This rapid elevation in the number of cases caused increased fears of health system collapse, with hospitals experiencing difficulty treating patients with diseases other than COVID-19, such as myocardial infarction and heart failure [6], as well as an increase in the number of patients with out-of-hospital cardiac arrest [7]. However, if the timing and shape of an epidemic wave could be predicted in advance with high accuracy, hospitals and health care systems could better prepare according to the predicted number of cases and expected time frame.

Several prediction models have been proposed for infectious diseases, including the historical mathematical model based on differential equations such as the SIR (susceptible, infectious, recovered) model [8]; models involving time-series analysis with use of past data to predict the future, such as the autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA) models [9–11]; and models that learn data patterns for prediction and classification, such as the Prophet model [12, 13]. Additionally, the effective application of outbreak prediction or forecasting models is crucial for obtaining insightful information regarding the transmission dynamics of a disease and its consequences. However, standard prediction models that deliver accurate results have not yet been established [14]. Various factors increase the uncertainty of prediction models, including known and unknown variables, differences in population/behavioral complexity in different geopolitical areas, people's vaccination status, the evolution of new strains, medical measurements, and variations in containment strategies [15–17]. Therefore, developing a standard prediction model for COVID-19 outbreaks that matches the real-world data with adjustment for these risk factors is challenging. Under these conditions, a highly accurate prediction model for COVID-19 in Japan is still lacking. In a comparison among the 47 prefectures of Japan, the speed of increase in the number of cases differs according to prefecture [5]. A prediction model of COVID-19 for Japan must be adjusted to the conditions in each prefecture; therefore, a highly accurate,

tailor-made prediction model must incorporate the local conditions. However, it can be hypothesized that if medical organizations and local governments in each location can produce accurate epidemic predictions, this would contribute to the preparation of countermeasures against COVID-19 throughout the health system that match the conditions in each region.

The aim of the present study was to predict COVID-19 outbreaks in Japan using a simple mathematical model.

## Methods

### Dataset

We used data that were made openly available by the Ministry of Health, Labour and Welfare of Japan from the first confirmed COVID-19 case on January 16, 2020 up to May 8, 2023 [18]. During this period in Japan, successive variations in the progression and containment of the epidemic, or waves, were observed eight times. Data for daily counts of new COVID-19 cases at both national and prefectural levels were collected and used for developing the model.

The end date of each epidemic wave was defined as the week with the minimum number of weekly cases after the peak of the epidemic (in which the number of weekly cases was greater than that in the following weeks). Using this definition, the duration of each epidemic wave in Japan was as follows. Wave 1 was from January 16, 2020 to May 24, 2020; Wave 2: May 25, 2020 to September 27, 2020; Wave 3: September 28, 2020 to February 28, 2021; Wave 4: February, March 1, 2021 to June 20, 2021; Wave 5: June 21, 2021 to October 3, 2021; Wave 6: November 29, 2021 to June 19, 2022; Wave 7: June 20, 2022 to October 9, 2022; and Wave 8 was from October 10, 2022 to April 2, 2023.

### Assessing the periodicity of epidemic waves using time-series analysis and the autocorrelation coefficient

We used time-series analysis and the autocorrelation coefficient (ACF) to examine the periodicity of COVID-19 epidemic waves in Japan. A time-series analysis necessitates decomposing the data into fundamental elements: trends that capture global trends and irregular, short-term fluctuations (called noise). Given the pronounced short-term periodicity influenced by weekdays among COVID-19 cases in Japan, a moving average method was adopted. This involved calculating a 7-day average centered around each day to smooth out variations, treating it as the daily count of newly confirmed cases.

The ACF was applied to investigate longer-term patterns of COVID-19 cases [19]. The ACF is used to calculate the correlation coefficient between the original data and data shifted by specific time lags (referred to as lags), aiming to ascertain the presence of periodicity. The ACF outcomes were visualized using a correlogram that

captures the periodicity of the data, representing the correlation coefficients between the original data  $x$  and the time-shifted data  $y$ . This is a plot with lags on the horizontal axis and autocorrelation on the vertical axis, illustrating the relationship. The graph initiates from lag 0, where the correlation coefficient is 1 from a comparison with the identical dataset. Generally, as the lag increases, the correlation tends to attenuate, although instances arise where the strength of the correlation re-emerges. This phenomenon signifies autocorrelation, serving as a guide for exploring periodicity. We conducted time-series analysis of changes in the ACFs after January 2020. Then, the collected data were used to derive a predictive framework for epidemic trends.

**Algorithm for a simple prediction model of the seventh COVID-19 wave in Japan**

**Step 1.** A time series of COVID-19 case datasets (i.e., daily new confirmed cases) from January 15, 2020 to April 23, 2023 was generated in Python programming language (Ver. 3.0.14) using pandas (ver. 1.2.4) to conduct the tabular data analysis. We used a function of a linear regression model from the machine learning library scikit-learn (ver. 0.24.1) to fit the regression lines for each

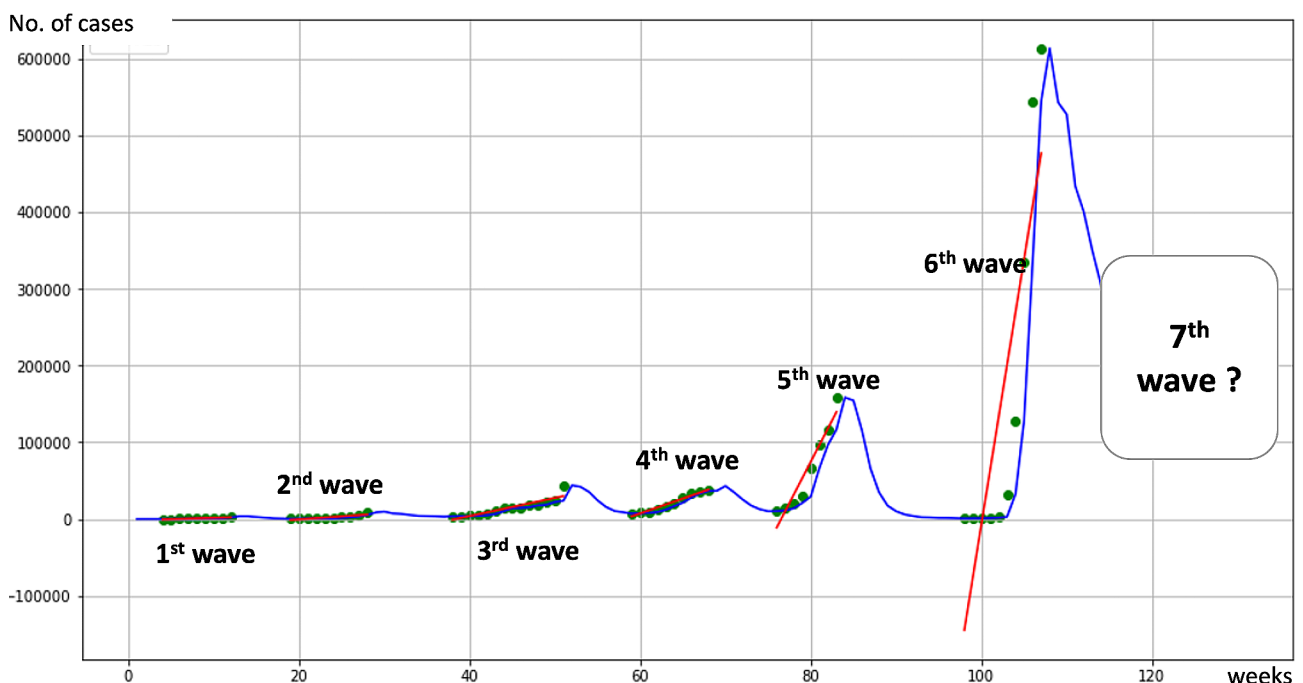
epidemic wave, which we denoted the “rising trend line.” (Fig. 1). The data were visualized using matplotlib (ver. 3.3.4) and seaborn (0.11.1).

The next epidemic wave might be predicted on the basis of regular changes in the slope of this line.

**Step 2.** We then identified an exponential function with good fit from the data for rising straight lines up to the sixth epidemic wave and calculated the timing for the rise of the seventh wave and its speed of spread using the GeoGebra’s function for applying an exponential model of bivariate regression analysis (GeoGebra Classic 5.2.826.0-d, International GeoGebra Institute, Linz, Austria). Regression analysis was performed after plotting the data calculated in step 1 on a coordinate plane to fit the logistic growth curves. The obtained model is shown in Fig. 2.

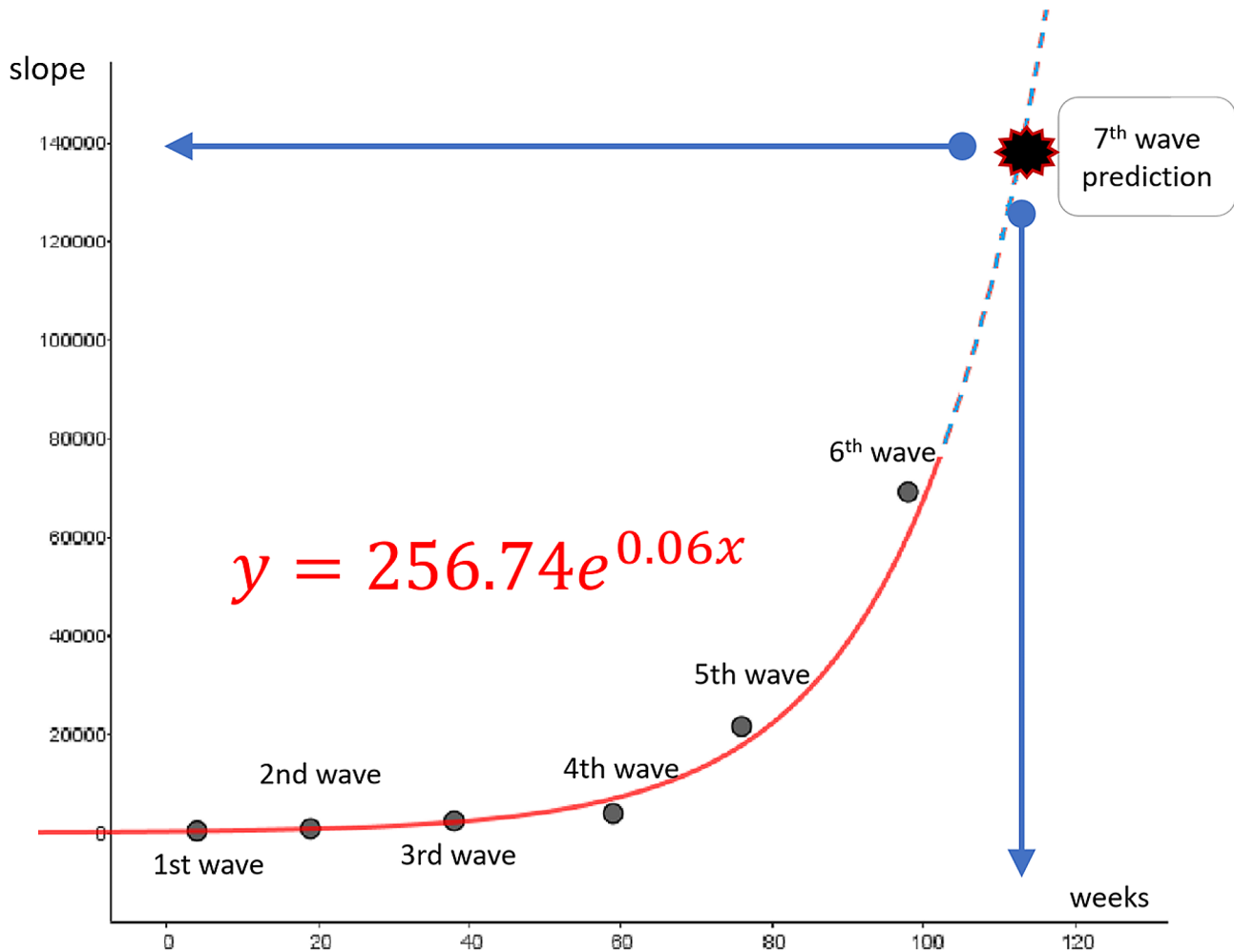
The surge timing was predicted using the average interval between waves up to the sixth wave, and the exponential function was used to calculate its expansion speed. After the calculated data were plotted on a coordinate plane, regression analysis was performed, and the function with a good fit was adopted as model (1).

$$y=256.74e^{0.06x} \tag{1}$$



	1 <sup>st</sup> wave	2 <sup>nd</sup> wave	3 <sup>rd</sup> wave	4 <sup>th</sup> wave	5 <sup>th</sup> wave	6 <sup>th</sup> wave	7 <sup>th</sup> wave
No. of weeks	4	19	38	59	76	98	Prediction
slope of rising line	345	787	2366	3879	21538	69103	

**Fig. 1** Epidemic waves of COVID-19 in Japan (blue line) with the rising trend line (red line)



**Fig. 2** Prediction diagram of speed of increase in number of COVID-19 cases and timing of outbreak

The average interval between each epidemic wave was 19 weeks (95% confidence interval, 13.1–25.0). We divided the duration from the time of occurrence of an epidemic to the time of its peak into two periods, a rising period and peak period. We then calculated the slope of the rising trend line. We predicted the seventh wave during the convergence period of the sixth wave. This period was longer than those in the previous epidemic period, so we added 10 weeks to the start time of the seventh wave. The coefficients used to predict the slope of the rising trend line are shown in Additional file 1.

**Step 3.** A logistic growth model is often used to fit the time series analysis in studies of infectious diseases [20–22] (2).

$$f(t) = \frac{1}{1 + e^{-t}} \quad (2)$$

Using the logistic growth model, the variable  $t$  represents the number of days (duration) in each epidemic wave;  $a_i$  is the number of infected people in each epidemic;  $b_i$  is

the inflection point of the curve representing the transition (number of days at the peak in an epidemic wave); and the speed of infection. Using  $c_i$  as a coefficient to adjust, the total number of infected people in that epidemic wave is expressed by the following formula.

$$f_i(t) = \frac{a_i}{1 + e^{-\frac{b_i-t}{c_i}}} \quad (3)$$

The following equation, which differentiates this with respect to the time variable  $t$ , can be expressed as the daily number of cases. Then, the variables  $a_i$ ,  $b_i$ , and  $c_i$  in each epidemic wave are calculated as follows:

$$\frac{df_i(t)}{dt} = \frac{a_i e^{-\frac{b_i-t}{c_i}}}{c_i (1 + e^{-\frac{b_i-t}{c_i}})^2} \quad (4)$$

We used Python programming language (Ver. 3.0.14) for the analytical process throughout the present study.

For all analyses, significance levels were two-tailed, and  $p < 0.05$  was considered statistically significant.

## Results

### Periodicity of COVID-19 in Japan using time-series analysis and the autocorrelation coefficient (ACF)

The temporal distribution of COVID-19 cases during the whole observational period (January 2020 to March 2023) is shown in Fig. 3.

We used advanced machine learning techniques to ascertain and visually represent the epidemic waves with a rising trend line (Fig. 1). We also assessed the distribution of COVID-19 cases with a rising trend line in Japan's 47 prefectures; different trends among epidemic waves in each prefecture can be observed (Additional file 2).

Figure 4 shows the correlogram using the whole observational period (up to March 31, 2023).

Between January 16, 2020 and May 8, 2023, the correlation coefficient was approximately 0.5, indicating significant periodicity. The peak of lag was approximately 140 (red arrow). Therefore, the spread of infection was repeated in a cycle of approximately 140 days, with high periodicity.

### Prediction of the seventh COVID-19 wave and its accuracy

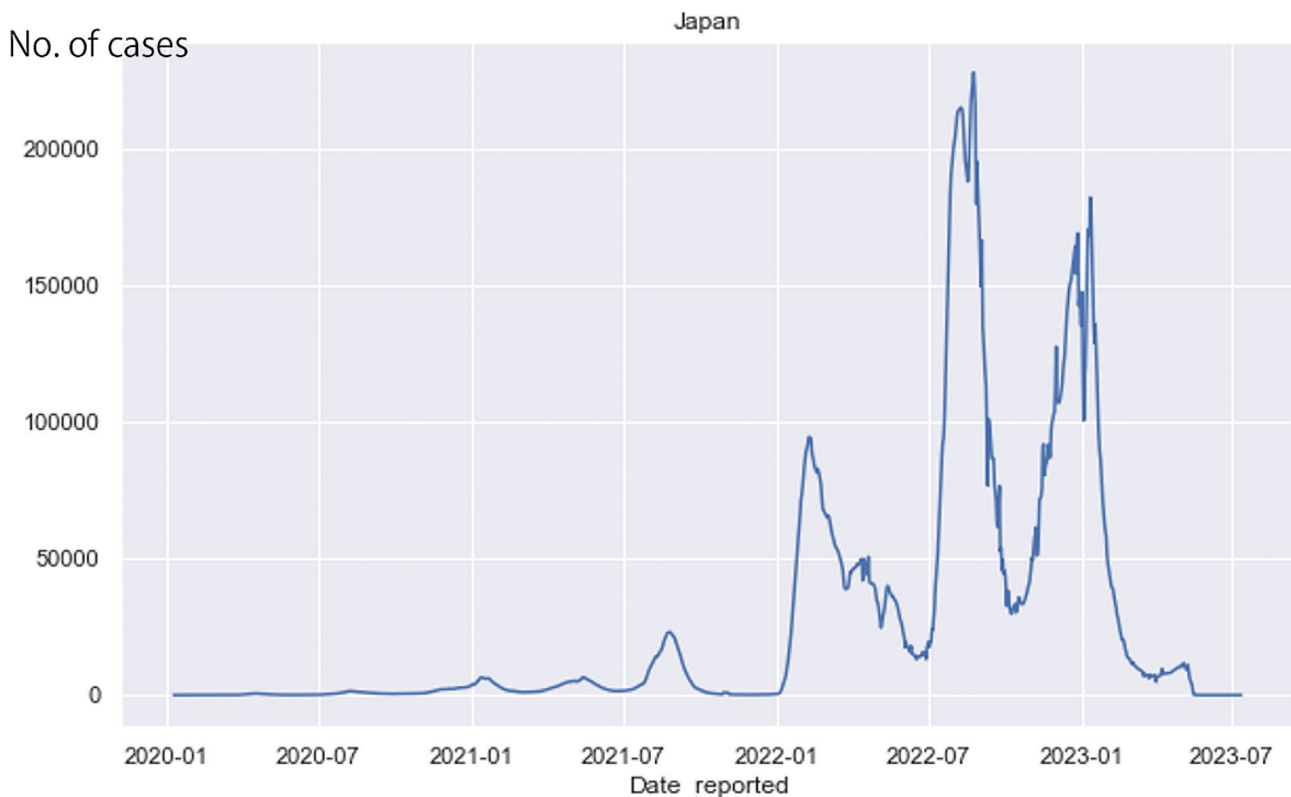
The epidemic curves using actual data for the number of infected cases and the approximate model are shown

in Fig. 5, including the time of prediction of the seventh wave (March 15, 2022) and the simulated wave using the prediction model, as well as the actual number of cases after the time of prediction.

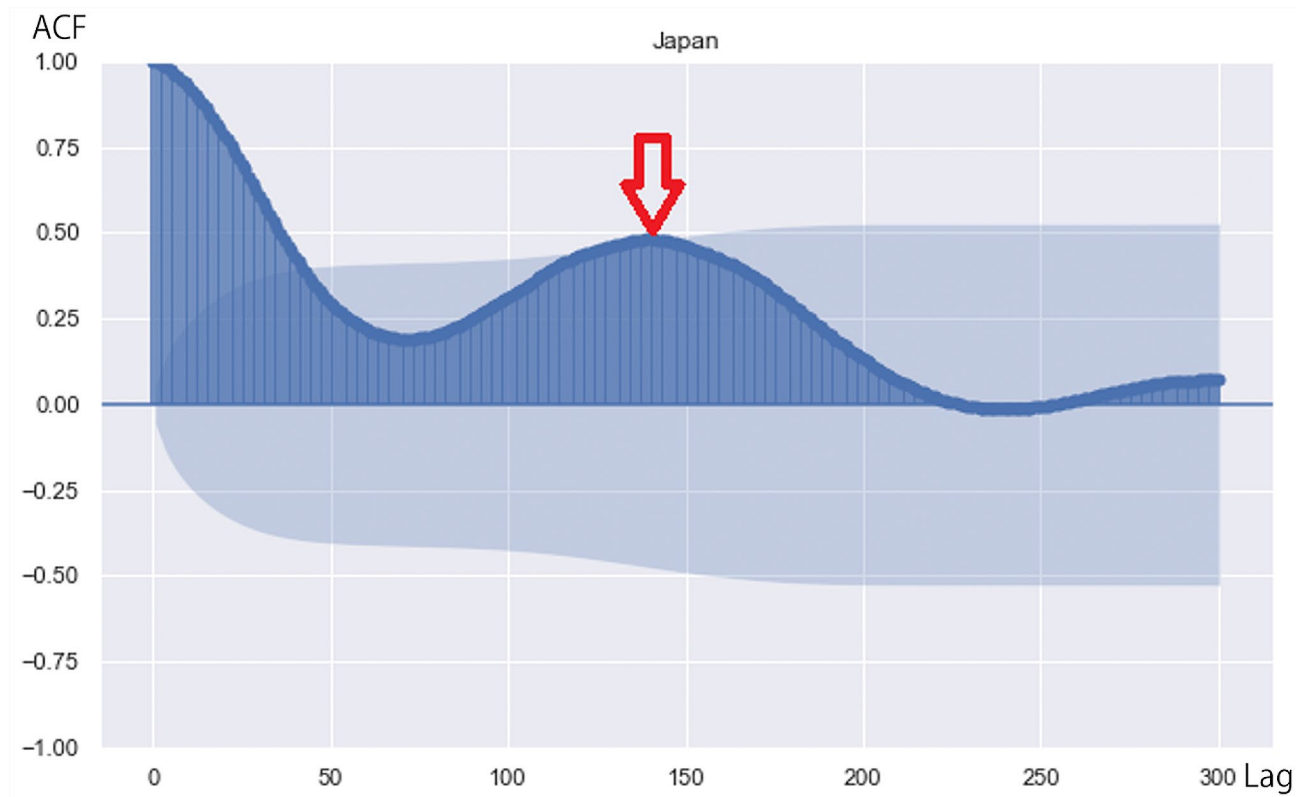
The actual distribution of COVID-19 cases and the distribution using the approximated model, which were used as the basic tool for developing the prediction model, presented similar distributions by mid-March 2022 (Fig. 5). In the predicted seventh wave, although the starting time and peak time of the epidemic were slightly behind, the mean absolute percent error (MAPE) during March 13, 2022 to October 12, 2022 indicated 53.5% but 20.0% for the 3 months from July 1, 2022 to October 1, 2022 during the outbreak and 13.9% for the 1-month period from July 1, 2022 to August 1, 2022 during the expansion period. Therefore, although this provides the rationale for setting it to 200 days, if there had been a 15-day difference, it would have been possible to predict the outbreak and spread with a high degree of accuracy.

## Discussion

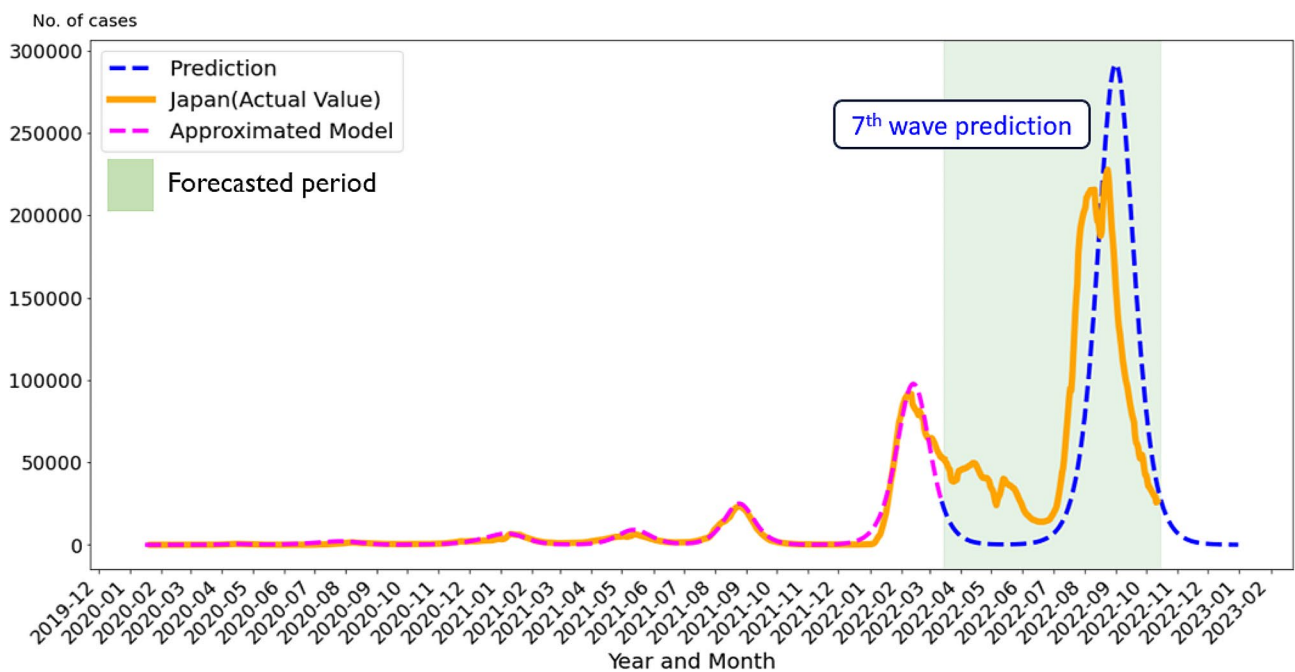
We constructed a simple prediction model by combining the depicted rising trend line. Our model showed a high degree of accuracy, especially when the distribution of COVID-19 cases had substantial periodicity. The spread of COVID-19 in Japan was repeated in a cycle of approximately 140 days, with high periodicity.



**Fig. 3** Temporal distribution of COVID-19 cases, January 2020–March 2023



**Fig. 4** Correlogram using data from January 2020 to March 2023 for COVID-19 cases in Japan. Red arrow indicates the peak of lag. Light blue band indicates 95% confidence interval. Values outside the band (dark blue), which are the autocorrelation coefficient, are statistically significant



**Fig. 5** Distribution of COVID-19 cases and simulated wave developed using a prediction model. Orange line: actual distribution of COVID-19 cases. Pink line: distribution using the approximated model. Blue line: simulated wave developed using prediction. Gray shade: Forecasted period

After the emergence of the COVID-19 pandemic, SARS-CoV-2 spread rapidly across countries worldwide, including Japan. It threatened people's daily lives and caused medical care challenges such as health system collapse, which would make it impossible for patients with COVID-19 to be treated in a hospital. However, if it were possible to predict the timing of a future outbreak with high accuracy, as well as the trends in the number of cases over time, targeted infection control measures could be efficiently planned and the medical system could adequately prepare. We constructed a model to predict the surge timing of an epidemic wave using an exponential function, which yielded empirical evidence to support this model up to the seventh COVID-19 wave in Japan. The periodicity in SARS-CoV-2 transmission may stem primarily from changes in the implemented public health and social measures, adherence levels, antigenic drift, and seasonality influenced by environmental factors. However, because COVID-19 is an emerging infectious disease, risk factors that affect the spread of infection have not been clearly elucidated. Additionally, COVID-19 infection may be transmitted from asymptomatic or pre-symptomatic individuals [23] and many cases of reinfection have been reported [24]. Owing to these issues, it is considered difficult to predict COVID-19 outbreaks with high accuracy and in a timely fashion using conventional SIR models and models using the basic reproduction number or effective reproduction number.

From our previous studies, we have seen that some emerging and re-emerging infectious diseases have a high cyclic trend [25, 26]. We hypothesized that if periodicity exists for COVID-19 cases in Japan, we can predict the timing of a future outbreak. Although the periodicity varied slightly among prefectures, the spread of COVID-19 across Japan exhibited a cycle of approximately 140 days, with high periodicity. This number is reasonable within the context of the COVID-19 pandemic in Japan, where epidemic waves occurred mainly during the summer and winter. We then calculated and plotted the rising trend line for the next epidemic wave, after the peak of the previous epidemic wave had ended. The timing from the calculation differed from that of a graphical model using a line rising with the increase in the number of cases per day [27]. Parag et al. derived a novel method that can estimate the probability for the end of an epidemic [28]. Incorporating this method into our prediction model confers the possibility of analyzing the timing of convergence of the epidemic wave after the peak and potentially analyzing the timing of convergence of post-peak epidemic waves, leading to enhanced analysis by the prediction model. Thus, we would be able to predict the number of COVID-19 cases over time during the next epidemic wave months in advance. These models are simple; however, by combining the calculation methods for the cyclic

trend and the rising trend line. This method could predict the starting time, peak, and number of cases in the seventh COVID-19 wave in Japan, during the previous wave. However, more work is needed to verify that our findings apply to other locations and other time frames.

We confirmed that the shape and rising trend line of the epidemic wave differed depending on each prefecture in Step 2 of the algorithm. However, in Japan, the same COVID-19 countermeasures were taken nationwide under government initiatives. Therefore, in Step 3 of the algorithm, we used national data and analyzed data throughout Japan to build an infection prediction model, in preparation for building different models for each prefecture and region in the future. We also found that each prefecture in Japan had a different periodicity. However, periodicity was analyzed by integrating national data, and analysis that considered periodicity in each prefecture was not carried out. When examining periodicity using a correlogram, the periodicity may change depending on the time of the analysis owing to changes in the amount of data used; it was not possible to reliably confirm the periodicity of 140 days using national data for Japan. Additionally, during a prolonged epidemic period, it is unclear whether it can be assumed that the speed of increase will consistently follow the proposed exponentially increasing pattern in the future, especially after society transitions to the normal endemic phase of SARS-CoV-2 circulation. Although there are various known and unknown variables and factors related to the conditions leading to infection, we developed our prediction model using solely case incidence data. It is unclear which of these factors are true influencing factors, and their weight. Additionally, mutations of the virus, the availability of vaccines and people's vaccination status, and individual knowledge about prevention measures are related to the spread of COVID-19. However, it is difficult to obtain such data in a timely manner, because these data change constantly over time. Moreover, there are limits to available data related to the characteristics of COVID-19, such as who an infected person infects, how many people they infect, and the number of pre-symptomatic and asymptomatic cases for each variant. In the future, we are considering using available data from medical institutions and public health centers, calculating predicted values for unavailable data, and developing a predictive model using these data.

Despite these limitations, it is crucial to consider the various patterns of prediction models in preparation for future emerging infectious diseases to protect the health of populations worldwide.

## Conclusions

Our simple model, which uses periodicity and the rising trend line, showed that if past outbreaks have periodicity, the spread of COVID-19 can be predicted up to a few months in advance. The study findings suggest the possibility of predicting the starting point of a future infectious disease outbreak and the number of infected individuals, contributing to early policy decision-making and advanced health system preparation. The results suggest that our developed simple mathematical prediction model can facilitate tailor-made epidemic prediction of COVID-19 outbreaks in the future.

## Abbreviations

COVID-19	Coronavirus disease 2019
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
WHO	World Health Organization
ACF	Autocorrelation coefficient

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12879-024-09354-5>.

Supplementary Material 1

Supplementary Material 2

## Acknowledgements

We thank Analisa Avila, MPH, ELS for editing a draft of this manuscript.

## Author contributions

Conceived and designed the study: HM; performed the experiments: HM, TM; Data collection: HM, YH, YK, DK, YM; analyzed the data: HM; interpreted the study results: HM, TM, YH, YK, DK, YM, KK; wrote the first draft of the manuscript: HM, TM. All authors read and approved the final manuscript.

## Funding

This study was supported by a grant from the Japan Science and Technology (JST) Mirai Program (#20345310), and the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI, Promotion of Joint International Research B, #20KK0218). These funders had no role in the design, methods, participant recruitment, data collection and analysis, or manuscript writing.

## Data availability

Data is provided within the manuscript or supplementary information files.

## Declarations

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Ethics approval and consent to participate

The requirement for ethical approval was waived because the study used open-access, anonymized information.

### Author details

<sup>1</sup>Shitennoji University, 3-2-1 Gakuenmae, Habikino City 583-8501, Osaka, Japan

<sup>2</sup>Nagoya City University School of Data Science, Nagoya City, Aichi, Japan

<sup>3</sup>Nagoya City University Graduate School of Medicine, Nagoya City, Aichi, Japan

<sup>4</sup>Showa Pharmaceutical University, Machida, Tokyo, Japan

<sup>5</sup>Tokyo University Graduate School of Engineering, Tokyo, Japan

<sup>6</sup>Waseda University Organization Regional and inter-regional Studies, Tokyo, Japan

<sup>7</sup>Kawakita General Hospital, Tokyo, Japan

Received: 30 December 2023 / Accepted: 26 April 2024

Published online: 09 May 2024

## References

- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;15(10223):497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (Accessed March 15, 2020).
- Ministry of Health, Labour and Welfare. Incidence of pneumonia patients relating to new corona virus infection (The 1st case). [https://www.mhlw.go.jp/stf/newpage\\_08906.html](https://www.mhlw.go.jp/stf/newpage_08906.html). (Accessed March 15, 2023).
- World Health Organization. Novel Coronavirus (2019-nCoV) Situation Report – 1. [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10_4) (Accessed April 5, 2020).
- Ministry of Health, Labour and Welfare. Situation report of COVID-19 (March 31, 2023). [https://www.mhlw.go.jp/stf/newpage\\_32443.html](https://www.mhlw.go.jp/stf/newpage_32443.html). (Accessed April 3, 2023).
- Baum A, Schwartz MD. Admissions to Veterans affairs hospitals for Emergency conditions during the COVID-19 pandemic. *JAMA*. 2020;324(1):96–9. <https://doi.org/10.1001/jama.2020.9972>
- Baldi E, Sechi GM, Mare C, Canevari F, Brancaglione A, Primi R, et al. Lombardia CARE Researchers. Out-of-hospital cardiac arrest during the Covid-19 outbreak in Italy. *N Engl J Med*. 2020;383(5):496–8. <https://doi.org/10.1056/NEJMc2010418>
- Kermack William Ogilvy and McKendrick A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* 1927;A115700–721. <https://doi.org/10.1098/rspa.1927.0118>
- Helpfenstein U. The use of transfer function models, intervention analysis and related time series methods in epidemiology. *Int J Epidemiol*. 1991;20(3):808–15. <https://doi.org/10.1093/ije/20.3.808>. PMID: 1955267.
- Wang Y, Yan Z, Wang D, Yang M, Li Z, Gong X, et al. Prediction and analysis of COVID-19 daily new cases and cumulative cases: times series forecasting and machine learning models. *BMC Infect Dis*. 2022;22(1):495. <https://doi.org/10.1186/s12879-022-07472-6>
- Helpfenstein U. Box-Jenkins modelling in medical research. *Stat Methods Med Res*. 1996;5(1):3–22. <https://doi.org/10.1177/096228029600500102>
- Zhao D, Zhang R, Zhang H, He S. Prediction of global omicron pandemic using ARIMA, MLR, and Prophet models. *Sci Rep*. 2022;12(1):18138. <https://doi.org/10.1038/s41598-022-23154-4>
- Tulshyan V, Sharma D, Mittal M. An Eye on the future of COVID-19: prediction of likely positive cases and fatality in India over a 30-Day Horizon using the Prophet Model. *Disaster Med Public Health Prep*. 2022;16(3):980–6. <https://doi.org/10.1017/dmp.2020.444>
- Naeem M, Yu J, Aamir M, Khan SA, Adeleye O, Khan Z. Comparative analysis of machine learning approaches to analyze and predict the COVID-19 outbreak. *PeerJ Comput Sci*. 2021;7:e746. <https://doi.org/10.7717/peerj-cs.746>
- Ardabili SF, Mosavi A, Ghamisi P, Ferdinand F, Varkonyi-Koczy AR, Reuter U, Rabczuk T, Atkinson PM, Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S et al. P Clinical Trial Group. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med*. 2020;383(27):2603–2615. <https://doi.org/10.1056/NEJMoa2034577>
- Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. H BNT162b2 mRNA Covid-19 vaccine in a Nationwide Mass Vaccination setting. *N Engl J Med*. 2021;384(15):1412–23. <https://doi.org/10.1056/NEJMoa2101765>
- Magazine N, Zhang T, Wu Y, McGee MC, Veggiani G, Huang W. Mutations and evolution of the SARS-CoV-2 Spike Protein. *Viruses*. 2022;14(3):640. <https://doi.org/10.3390/v14030640>
- Ministry of Health, Labour and Welfare. Open data. <https://www.mhlw.go.jp/stf/covid-19/open-data.html> (Accessed August 1, 2023).



19. Chen Y. An analytical process of spatial autocorrelation functions based on Moran's index. *PLoS ONE*. 2021;16(4):e0249589. <https://doi.org/10.1371/journal.pone.0249589>
20. Zou Y, Pan S, Zhao P, Han L, Wang X, Hemerik L, Knops J, van der Werf W. Outbreak analysis with a logistic growth model shows COVID-19 suppression dynamics in China. *PLoS ONE*. 2020;15(6):e0235247. <https://doi.org/10.1371/journal.pone.0235247>. PMID: 32598342; PMCID: PMC7323941.
21. Tao J, Ma Y, Luo C, Huang J, Zhang T, Yin F. Summary of the COVID-19 epidemic and estimating the effects of emergency responses in China. *Sci Rep*. 2021;11(1):717. <https://doi.org/10.1038/s41598-020-80201-8>. PMID: 33436848; PMCID: PMC7803947.
22. Chen DG, Chen X, Chen JK. Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. *Glob Health Res Policy*. 2020;5:25. <https://doi.org/10.1186/s41256-020-00152-5>. PMID: 32435695; PMCID: PMC7225094.
23. Xu C, Zhang Z, Huang X, Cheng K, Guo S, Wang X et al. A study on the transmission dynamics of COVID-19 considering the impact of asymptomatic infection. *J Biol Dyn*. 2023;17(1):2244980. <https://doi.org/10.1080/17513758.2023.2244980>. Erratum in: *J Biol Dyn*. 2023;17(1):2259761. PMID: 37656780.
24. Ciuffreda L, Lorenzo-Salazar JM, García-Martínez de Artola D, Gil-Campesino H, Alcoba-Florez J, Rodríguez-Pérez H, et al. Reinfection rate and disease severity of the BA.5 Omicron SARS-CoV-2 lineage compared with previously circulating variants of concern in the Canary Islands (Spain). *Emerg Microbes Infect*. 2023;12(1):2202281. <https://doi.org/10.1080/22221751.2023.2202281>
25. Manabe T, Takasaki J, Kudo K. Seasonality of newly notified pulmonary tuberculosis in Japan, 2007–2015. *BMC Infect Dis*. 2019;19(1):497. <https://doi.org/10.1186/s12879-019-3957-8>. PMID: 31170932; PMCID: PMC6555020.
26. Manabe T, Yamaoka K, Tango T, Binh NG, Co DX, Tuan ND, et al. Chronological, geographical, and seasonal trends of human cases of avian influenza A (H5N1) in Vietnam, 2003–2014: a spatial analysis. *BMC Infect Dis*. 2016;16:64. <https://doi.org/10.1186/s12879-016-1391-8>
27. Perneger T, Kevorkian A, Grenet T, Gallée H, Gayet-Ageron A. Alternative graphical displays for the monitoring of epidemic outbreaks, with application to COVID-19 mortality. *BMC Med Res Methodol*. 2020;20(1):248. <https://doi.org/10.1186/s12874-020-01122-8>. Erratum in: *BMC Med Res Methodol*. 2020;20(1):265. PMID: 33023505; PMCID: PMC7537983.
28. Parag KV, Donnelly CA, Jha R, Thompson RN. An exact method for quantifying the reliability of end-of-epidemic declarations in real time. *PLoS Comput Biol*. 2020;16(11):e1008478. <https://doi.org/10.1371/journal.pcbi.1008478>

### Publisher's Note

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