

Estimating the Basic Reproductive Number for Omicron Outbreak in Hangzhou, China

Jiong-Hang Wu, Hao-Zhang Wu, Yu-Quan Xie, Hao-Bo Xu, Yu-Jie Shi and Juan Li*

School of Computer Science and Technology (School of Artificial Intelligence), Zhejiang Sci-Tech University, Hangzhou, 310018, P.R. China.

*Corresponding author email id: lijuanjuan0120@163.com

Date of publication (dd/mm/yyyy): 22/04/2023

Abstract – Since the first case of COVID-19 was reported in 2019, the epidemic is still raging around the world, and human life safety and health are facing a major threat. To quantify the transmission rate and severity of the Omicron epidemic in the early stage of Hangzhou, based on the epidemic data of Hangzhou in China from March 25, 2022 to April 26, 2022, we applied multiple mathematical formulas to directly calculate the basic reproductive number (R_0) and effective reproductive number (R_e). This is of great significance for retrospectively evaluating the transmission ability of Omicron virus, observing the development of the epidemic situation, as well as analyzing the effectiveness of prevention and control of Omicron virus in Hangzhou.

Keywords – COVID-19, Basic (Effective) Reproductive Number, Exponential Growth Method, Maximum Likelihood Method, Sequential Bayesian Method, Prevention and Control.

I. INTRODUCTION

Since the outbreak of COVID-19, human health has been constantly under threat. Even if countries adopt strict epidemic prevention measures, it is difficult to resist the epidemic. According to real-time statistics from Johns Hopkins University (CSSE), as of 2 April 2023, the total number of laboratory-confirmed cases during the pandemic has reached more than 67,000,000, 6,000,000 deaths, and 4,000,000 new cases have been reported, including 28,000 deaths in the last 28 days [1]. The global fight against the COVID-19 pandemic is far from over.

In the past two years, SARS-COV-2 variants of concern (VOC) have been evolving in countries or regions with high virus transmission rates and low vaccine coverage [2]. When assessing the infectivity of different variants, a parameter in the epidemiological concept has attracted much attention, the basic reproductive number (Reproductive R_0), which refers to the average number of secondary cases that an infected person can infect during his or her infection in the absence of any intervention [3]. In the process of epidemic prevention and control, R_0 can be used to describe the infectibility of a certain disease, evaluate the difficulty of disease control, help predict the speed and scale of disease transmission, as well as the level of population immunity required for disease control, etc. R_0 has become a very important threshold parameter used to quantify the spread of diseases [4]. According to R_0 estimation, from Alpha to Delta, and then to Omicron, the transmission capacity of VOCs (Delta: 3.2-8, Omicron: 3.47-10) has exceeded the original strain (2.5-2.79) by 3-5 times [5] [6]. According to the definition of R_0 , the best time for its value that is likely to be accurately estimated is at the beginning of the outbreak. As various interventions are implemented, such as testing [7], isolation [8], [9], tracing [7], vaccines [10], and other non-drug interventions [11], or if there are many infections already, the calculation of R_0 will deviate significantly from reality. Therefore, another important epidemiological index, effective reproductive number (Reproductive R_e), is introduced, which is time-changing and usually defined to quantify the number of secondary cases that an average infected person can infect during the infection period in a population where the

target population is not all susceptible at time t [12]. When evaluating the effectiveness of vaccines or other public health measures, the effective reproductive number (R_e) is a more appropriate indicator than R_0 . Effective assessment of epidemic transmission situation has become a hot issue of widespread concern in global public health, R_e estimated value of provides an indispensable reference for different countries and regions to respond to the epidemic.

This study will focus on the Omicron epidemic in the initial stage of Hangzhou, China. Driven by the actual epidemic data, we will use three kinds of mathematical calculation methods (Exponential growth method, Maximum likelihood method, Sequential Bayesian method) to estimate R_e . The results quantify the transmission rate of epidemic and analyze the impact of different calculation methods, providing scientific reference and foundation to better understand the epidemic development and evaluate the potential effectiveness of prevention and control measures in Hangzhou.

II. MATERIALS AND METHODS

2.1. Data Sources

The study collected data on daily confirmed cases reported by Hangzhou Municipal Health Commission from March 25 to April 26, 2022, including infection date, newly infected case number, the infection severity (asymptomatic/mild/severe), and the epidemic source [13].

2.1. Methods

2.2.1. Exponential Growth Method (EGM)

The exponential growth method is a calculation method of R_0 based on mathematical derivation [14]. This method first assumes that the number of newly infected people per day is still increasing exponentially, so the R_0 value used to calculate the early stage of an epidemic is more appropriate. The opposite is another reproductive number R_e , which has a similar meaning to R_0 , but “all susceptible people” is replaced by “some immune people”, and it is obviously more widely used in practice. The following algorithm is compatible with both reproductive numbers, and since the derivation assumes that the population is not entirely composed of susceptible populations, it is used R_0 here directly.

$$R_0 = (1 + r\bar{T}_L)(1 + r\bar{T}_I), \quad (1)$$

where r denotes the growth rate of exponential growth, $\bar{T}_L$ is the average length of incubation period of the infected person, and $\bar{T}_I$ represents the length of recovery. If $b(t)$ denotes the number of new infections on day t . Δt is the time length since the beginning of the epidemic on day t , then

$$b(t) = b(t - \Delta t)e^{r\Delta t}. \quad (2)$$

Arranging (2),

$$r = \ln(Yt) / t, \quad (3)$$

with

$$Yt = \frac{b(t)}{b(t-\Delta t)}. \quad (4)$$

According to (1) and (2), we obtain,

$$R_0 = 1 + rTG + p * (1 - p) * (rTG * rTG), \quad (5)$$

where

$$TG = TI + TL, p = TL/TG. \quad (6)$$

2.2.2. Maximum Likelihood Estimation (MLE)

Maximum likelihood estimation is a statistical method based on the principle of maximum likelihood, which is the application of probability theory in statistics [14]. MLE provides a method to evaluate the parameters of a model given the observed data, that is, “the model is determined, the parameters are unknown”. It is called MLE by observing the results of a number of experiments and using the test results to obtain a parameter value that maximizes the probability of the sample occurrence. The estimated value can be compared with the estimated value of the branching process of the offspring mean. The branching process assumes either that the sequence intervals are single-point distributed with mean 1 or that we can cluster the data into generations on the basis of prior knowledge of the sequence intervals or clusters evident in the data (reasonable for diseases with long-latency distributions). For example, if the mean value of consecutive intervals is μ , the vector of daily case denoted by N_T would be:

$$N = \{N_1, N_2, \dots, N_T\}. \quad (7)$$

According to (7), N can also be divided into several generations, for example:

$$N^* = \{N_0, \sum_{t=1}^{\mu} N_t, \sum_{t=1+\mu}^{2\mu} N_t, \dots, \sum_{t=(m-1)\mu+1}^{m\mu} N_t\}, \quad (8)$$

where

$$T/\mu = m. \quad (9)$$

In this case, using N^* to estimate R_0 , we have:

$$\tilde{R}_0 = \frac{\sum_{i=1}^m N_i^*}{\sum_{i=1}^m N_{i-1}^*}, \quad (10)$$

where N_i^* corresponds one-to-one with the vector N^* in (8).

2.2.3. Sequential Bayesian Method (SBM)

The SBM method estimates the R_0 value by the average duration of the infection period, and the prior distribution of R_0 at each time point is the posterior distribution at the previous time point [15]. The incidence number N_{t+1} at time $t + 1$ is approximately the Poisson distribution of mean $N_t e^{(\gamma(R-1))}$, where $\frac{1}{\gamma}$ describes the average duration of the infection period, and the calculation formula is given by [16]:

$$P(R|N_0, \dots, N_{t+1}) = \frac{P(N_{t+1}|R, N_0, \dots, N_t)P(R|N_0, \dots, N_t)}{P(N_0, \dots, N_{t+1})}. \quad (11)$$

III. NUMERICAL RESULTS

According to the available data, we plotted the regional distribution and time changes of the epidemic in Han-

-gzhou, as shown in Figure 1 and Figure 2. From Figure 1, one can be intuitively observed that the epidemic in Hangzhou is geographically concentrated, mainly occurring in three districts: Yuhang, Binjiang, and Gongshu. Figure 2 shows that the daily changes of new cases in Hangzhou over time. Hangzhou's Omicron outbreak has shown two peaks in a short period of time, with the previous peak being smaller than the latter. And the infected person is almost asymptomatic.

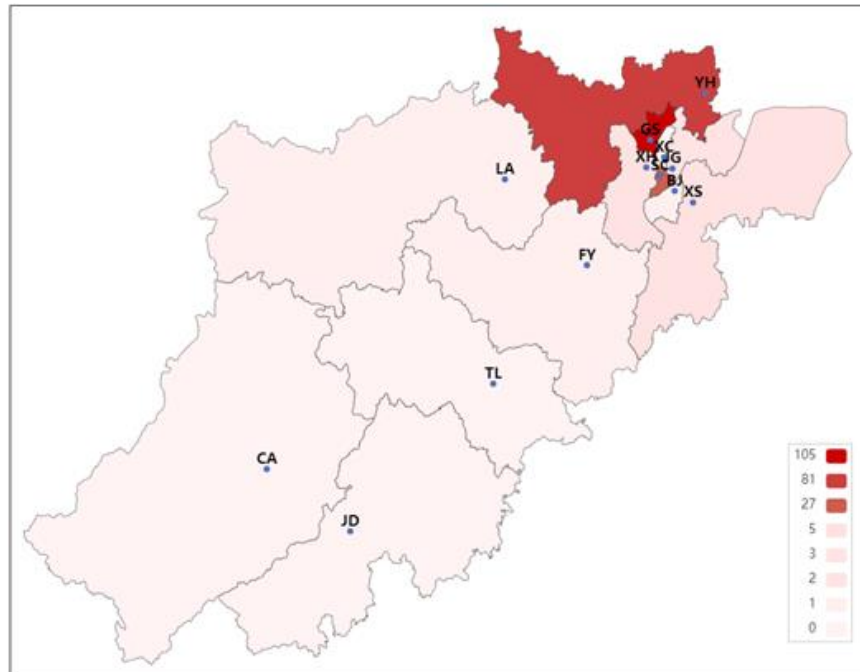


Fig. 1. The regional distribution of COVID-19 cases in Hangzhou, China, from March 25 to April 26, 2022, where the letter symbols are shorthand for the initials in the Chinese names of different regions of Hangzhou. Data source from Hangzhou Municipal People's Government Public Emergency website [13].

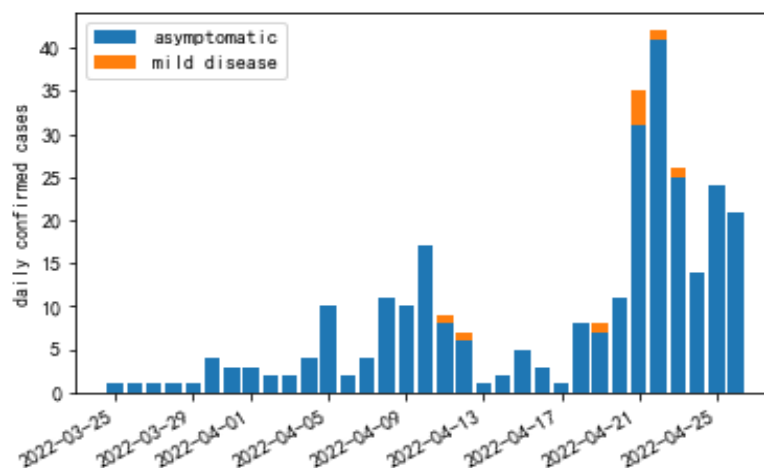


Fig. 2. Omicron epidemic change of Hangzhou, China from March 25 to April 26, 2022. Data source from Hangzhou Municipal People's Government Public Emergency website [13].

In next, to assess the transmission capacity of this epidemic, three methods (SGM, MLE, and SBM) were used for quantitative analysis of R_e based on the number of daily confirmed cases reported by the Health Commission of Hangzhou in China from March 25 to April 26, 2022. In this study, SGM and MLE were carried out using VScode software and python language, and SBM was performed using R language.

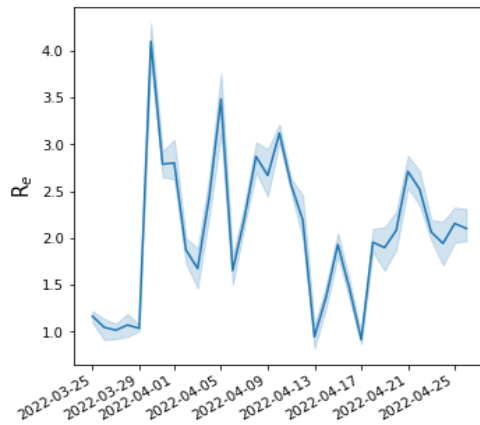


Fig. 3. EGM Result.

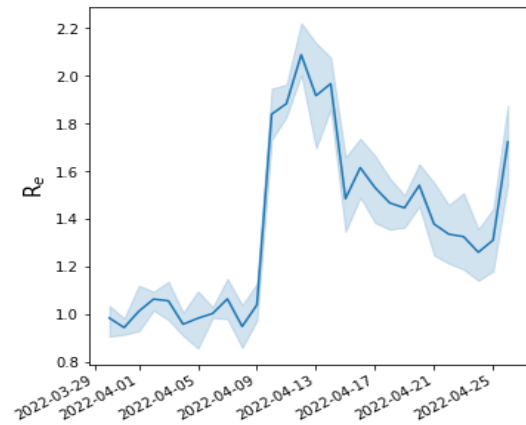


Fig. 4..MLE Result.

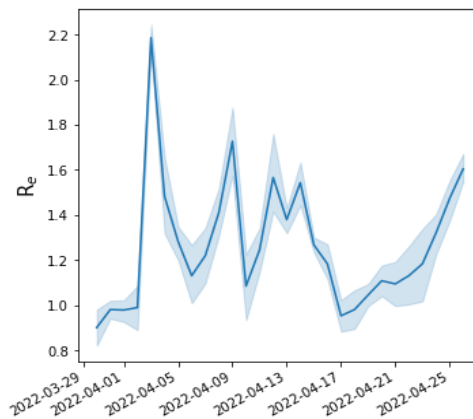


Fig. 5. SBM Result.

3.1. EGM

According to Equation (5), we plotted the situation of R_e over time, where the recovery period was $TI = 4.40$ days [17], and the incubation period was $TL = 2.52$ days [18]. Depending on the range of data available, the initial value is set by the data on March 25, and t is the number of days. It can be seen from Figure 3 that the R_e value estimated by EGM is in a stable state at about 1 at the beginning, and then it increases greatly to about 4.5 on March 30, and then declines rapidly until April 9, which is in the fluctuation state from 2.0 to 3.0, and falls to the range of 1.0 to 2.0 on April 10, and rises on April 20. It reaches more than 2.0. From the overall point of view, the change of R_e is relatively large.

3.2. MLE

R_e here is calculated according to Formula (10), where the sequence interval is $\mu = 2.75$ days [19]. Since this method needs the data of the first five days to calculate the R_e of the sixth day, so the plot of R_e over time starts on March 29, as shown in Figure 4. It can be intuitively seen from the figure that R_e is always in the state of around 1.0 before April 9, rises to above 2.0 on April 9, and then quickly falls on April 13. After April 20, it gradually rises to more than 2.0. In general, the fluctuation of R_e is larger than that of the EGM.

3.3. SBM

The numerical implementation of SBM method is based on Equation (11), where the mean of gamma distribution is 2.13 and the standard deviation is 1.69 [18]. As can be seen from Figure 5, before April 2, it is always in a stable state of about 1.0, rises to more than 2.20 on April 2, and then rapidly falls to the range of 2.0 to 1.0, reaches about 1.0 on April 17, and begins to rise by 2.0 on April 20. In general, the fluctuation trend of R_e is similar to the result obtained by EGM.

IV. DISCUSSION AND SUMMARY

This study focused on the hot issues of public health in Hangzhou of China, and conduct a retrospective quantitative analysis of the important epidemiological indicator R_e based on the actual epidemic data in Hangzhou. Driven by the Omicron epidemic data in Hangzhou of China, this study used EGM, MLE and SBM to evaluate the R_e value of Omicron, and quantified the transmission ability of Omicron in Hangzhou from multiple perspectives, making the results more scientific. The numerical analysis showed that the interval of the basic reproductive number of Omicron epidemic in Hangzhou was about 2.055 (95%CI 1.78, 2.33). However, different models would lead to different results, where the results obtained by EGM and SBM were very close. The MLE is expected the larger sequence interval to be the better when estimating R_e , but the sequence interval of omicron is short, which may result in an unprecise result.

Significantly, the R_e values in this study were estimated with the use of mathematical or statistical models based on retrospective data. In fact, its estimation results are also related to many factors, such as the accuracy of data, the distribution of data, and the structure of the model, etc. Therefore, the R_e values calculated in the past epidemics of the same infectious disease may not be applicable to the future outbreaks of the same infectious disease [20]. In addition, due to the delay of data collection, their practicability is further limited [21].

In a conclusion, when analyzing the basic reproductive number of an epidemic, different methods have a great impact on the obtained results, a reasonable estimation result can provide not only an intuitive understanding of the development situation and prevention and control effect of COVID-19 in Hangzhou, but also a useful reference for the evaluation of the transmission capacity of the epidemic in the short term. Therefore, how to choose an appropriate method to realize the quantification of the basic reproductive number of an epidemic is still a topic worthy of in-depth research.

REFERENCES

- [1] "COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)". [Online]. Available: <https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6>
- [2] "World Health Organization, Tracking SARS-CoV-2 variants". [Online]. Available: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>
- [3] P. Dreessche, J. Watmough, "Reproductive numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, Issue. 1. 2002, pp. 29–48, Doi: 10.1016/S0025-5564(02)00108-6.
- [4] Gulenkin VM, Korennoy FI, Karaulov AK, et al, "Cartographical analysis of African swine fever outbreaks in the territory of the Russian Federation and computer modeling of the basic reproductive ratio," *Preventive Veterinary Medicine*, vol. 102, Issue. 3. 2011, pp. 167-174, Doi: 10.1016/j.prevetmed.2011.07.004.
- [5] "Omicron variant and booster COVID-19 vaccines". [Online]. Available: [https://www.thelancet.com/journals/lanres/article/PIIS2213-2600\(21\)00559-2/fulltext](https://www.thelancet.com/journals/lanres/article/PIIS2213-2600(21)00559-2/fulltext)
- [6] Y Liu and R. Joacim, "The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus," *Journal of Travel Medicine*, Vol. 28, Issue. 7. 2021, Doi: 10.1093/JTM/TAAB124.
- [7] Ludwig A, Berthiaume P, Orpana H, et al, "Assessing the impact of varying levels of case detection and contact tracing on COVID-19 transmission in Canada during lifting of restrictive closures using a dynamic compartmental model," *Can Commun Dis Rep*, Vol. 46, Issue. 1112.2020, pp. 409-421, Doi: 10.14745/CCDR.V46I1112A08.
- [8] Shanshan Feng, Juping Zhang, Juan Li, et al, "The impact of quarantine and medical resources on the control of COVID-19 in Wuhan based on a household model," *Bulletin of Mathematical Biology*, Volume. 84, Issue. 4. 2022, pp. 47-47, Doi: 10.1007/S11538-021-00989-Y.
- [9] Tang B, Xia F, Tang SY, et al, "The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemics in the final phase of the current outbreak in China," *International Journal of Infectious Diseases*, Vol. 95, Issue. Epub.2020, pp. 288-293,

Doi: 10.1016/j.ijid.2020.03.018.

- [10] Elena Aruffo, Pei Yuan, Yi Tan, et al, "Mathematical modelling of vaccination rollout and NPIs lifting on COVID-19 transmission with VOC: a case study in Toronto, Canada," BMC Public Health, Vol. 22, Issue. 1. 2022, pp. 1349-1349, Doi: 10.1186/S12889-022-13597-9.
- [11] Pei Yuan, Juan Li, Elena Aruffo, et al, "Efficacy of 'stay-at-home' policy and transmission of COVID-19 in Toronto, Canada: a mathematical modeling study," CMAJ open, Volume. 10, Issue. 2. 2022, pp. E367-E378, Doi: 10.9778/cmajo.20200242.
- [12] A Cintrón-Arias, C Castillo-Chávez, L. Bettencourt, et al, "The estimation of the effective reproductive number from disease outbreak data," Mathematical Biosciences and Engineering, Vol. 6, Issue. 2. 2009, pp. 261-82, Doi: 10.3934/mbe.2009.6.261.
- [13] Hangzhou Municipal People's Government, Special topic of public emergencies, epidemic prevention and control. [Online]. Available: <http://www.hangzhou.gov.cn/col/col1229285516/index.html?uid=6350825&pageNum=1>
- [14] Wallinga J, Lipsitch M, "How generation intervals shape the relationship between growth rates and reproductive numbers," Proceedings of the Royal Society B: Biological Sciences, Vol. 274, Issue. 1609. 2007, pp. 599-604, Doi: 10.1098/rspb.2006.3754.
- [15] White L Forsberg, Pagano M, "A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic," Statistics in Medicine, Vol. 27, Issue. 16. 2008, pp. 2999-3016, Doi: 10.1002/sim.3136.
- [16] Bettencourt Luís M A, Ribeiro Ruy M, "Real time bayesian estimation of the epidemic potential of emerging infectious diseases," PLoS ONE, Vol. 3, Issue. 5. 2008, pp. e2185, Doi: 10.1371/journal.pone.0002185.
- [17] Anderson Roy M, Heesterbeek Hans, Klinkenberg Don, et al, "How will country-based mitigation measures influence the course of the COVID-19 epidemic," The Lancet, Vol. 395, Issue. Epub. 2020, pp. 931-934, Doi: 10.1016/S0140-6736(20)30567-5.
- [18] Liu Tian, Xu Qinwen, He Shiqi, et al, "Estimation of transmission dynamics parameters for OmicronBA.5.2 variant—based on real-world data," Disease Surveillance, accepted, 2023.
- [19] Hu Shixiong, Wang Wei, Wang Yan, et al, "Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China," Nature Communications, Vol. 12, Issue. 1. 2021, pp. 1533-1533, Doi: 10.1038/S41467-021-21710-6.
- [20] Delamater PL, Street EJ, Leslie TF, et al, "Complexity of the Basic Reproduction Number (R_0)," Emerging Infectious Diseases, Vol. 25, Issue. 1. 2019, pp. 1-4, Doi: 10.3201/eid2501.171901.
- [21] Heffernan JM, Smith RJ, Wahl LM, "Perspectives on the basic reproductive ratio," Journal of the Royal Society Interface, Vol. 2, Issue. 4. 2005, pp. 281-93, Doi: 10.3201/eid2501.171901.

AUTHOR'S PROFILE

First Author

Jiong-Hang Wu, Department of Computer Science and Technology, Zhejiang Sci-Tech University, China.

Second Author

Hao-Zhang Wu, Department of Computer Science and Technology, Zhejiang Sci-Tech University, China.

Third Author

Yu-Quan Xie, Department of Computer Science and Technology, Zhejiang Sci-Tech University, China.

Fourth Author

Hao-Bo Xu, Department of Computer Science and Technology, Zhejiang Sci-Tech University, China.

Fifth Author

Yu-Jie Shi, Department of Computer Science and Technology, Zhejiang Sci-Tech University, China.

Sixth Author

Juan Li, Department of Computer Science and Technology, Zhejiang Sci-Tech University, China.