IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 17, Issue 2 Ser. I (Mar. – Apr. 2021), PP 39-44
www.iosrjournals.org

Predicting the Existing Confirmed Cases of the COVID-19 Epidemic in Hong Kong Based on Dynamic Transmission Rate Model

YANYU DING[1], JIAXING LI[2], WEILIANG SONG[2], DINGZHE XIE[2], TIAN SANG[3*]

[1]*School of Chemistry and Chemical Engineering, Shanghai University of Engineering Science, Shanghai, 201620, China*
[2]*School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai, 201620, China*
[3]*School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai, 201620, China*
** Corresponding author: TIAN SANG*

Abstract:
In this paper, we make an empirical analysis of the COVID-19 epidemic in Hong Kong based on the dynamic transmission rate model. First, through visual analysis, we find the three main stages of the development of the COVID-19 epidemic in Hong Kong. Furthermore, based on data driving, a suitable sliding window period is selected, and a three-parameter power function is used to perform nonlinear weighted fitting of the dynamic transmission rate. Finally, the inflection point of three rounds of the COVID-19 epidemic in Hong Kong and existing confirmed cases within 7 days after the date where selected data set ends are predicted. Experiments show that the empirical analysis results of this article are more accurate.

***Key Word**: COVID-19, Dynamic transmission rate, Nonlinear weighted fitting, Inflection point, Infectious diseases*

Date of Submission: 01-03-2021 Date of Acceptance: 14-03-2021

I. Introduction

At the beginning of the COVID-19 outbreak, Hong Kong's rapid response enabled the local epidemic to be controlled in a relatively short period and became a benchmark city for epidemic prevention and control. Unfortunately, in the context of the continuous outbreak of the global new crown pneumonia epidemic, with the outbreak in Europe and the United States, the number of imported cases from abroad has increased sharply, causing another outbreak of the epidemic that was about to be contained, and the number of existing confirmed cases continued to increase. The COVID-19 epidemic in Hong Kong rebounded in March 2020, July 2020, and November 2020. Figure 1 shows the trend of cumulative confirmed cases, cumulative deaths, cumulative cured cases, and existing confirmed cases since the outbreak of the COVID-19 epidemic in Hong Kong.

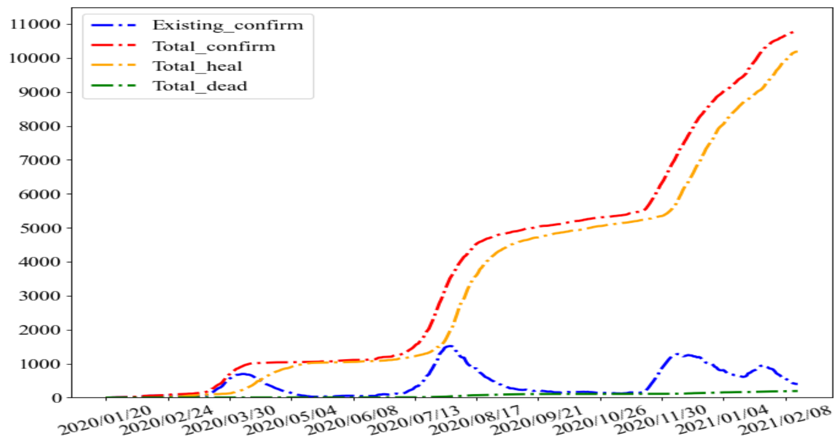


FIGURE 1: The development trend of the COVID-19 epidemic in Hong Kong

It can be seen from Figure 1 that since the outbreak of the COVID-19 epidemic in Hong Kong on January 20, 2020, the development of the epidemic has mainly gone through three rounds. They are from March 2020 to May 2020, July 2020 to October 2020, and November 2020 to January 2021.

During the spread of the COVID-19 epidemic, many researchers have researched the transmission mechanism and model prediction of the COVID-19 epidemic to find the best strategy for its prevention and control. Huang et al. [5] proposed a data-driven concise, and practical method for calculating the dynamic transmission rate of epidemics. Hu et al. [6], [7] developed the dynamic transmission rate model (DTRM) and dynamic growth rate model (DGRM) to predict the domestic and foreign COVID-19 epidemics, respectively. Xie et al. [8] constructed a nonlinear dynamic transmission rate model (NDTRM) based on support vector regression and predicted the domestic COVID-19 epidemic. Compared with the SEIR model [1], SIR model [2], and its extended models [3], [4], the data-driven DTRM has a wider application range, higher prediction accuracy, and robustness.

In this paper, we use DTRM based on a three-parameter power function to predict the inflection point of the three rounds of the COVID-19 epidemic in Hong Kong and existing confirmed cases within 7 days after the date where the selected data set ends. The empirical results show that the prediction accuracy and robustness of this article are high, which can provide a valuable reference for the prevention and control of the COVID-19 epidemic in Hong Kong.

II. MODELS AND METHODS

In this section, we briefly introduce the core concepts of the DTRM, which plays an essential role in the prediction and analysis of the COVID-19 epidemic [5]-[8].

Let $N(t)$ be the existing confirmed cases at the time t. Then

$$N(t) = L(t) - K(t) - D(t), \tag{1}$$

where $L(t)$, $K(t)$, $D(t)$ are cumulative confirmed cases, cumulative deaths, and cumulative cures at the time t, respectively.

It is well-known the natural growth model is given by

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = q(t)N(t),\ q(t) \geq 0, \tag{2}$$

where $q(t)$ is the growth rate of the existing confirmed cases at the time t.

Without loss of generality, we can assume that

$$q(t) = g(t) - 1 \geq 0. \tag{3}$$

It follows from (2) and (3) that

$$\ln N(t) - \ln N(t_0) = \int_{t_0}^{t} g(x)\mathrm{d}x - (t - t_0). \tag{4}$$

Let

$$a_t = \int_{t_0}^{t} g(x)\mathrm{d}x / (t - t_0) - 1. \tag{5}$$

Then

$$N(t) = N(t_0)\exp\{a_t(t - t_0)\}. \tag{6}$$

In what follows, we introduce the so-called dynamic transmission rate, i.e.,

$$c_t = 1 + a_t. \tag{7}$$

It follows from (5) and (7) that

$$c_t = 1 + \frac{1}{t - t_0}\ln\frac{N(t)}{N(t_0)}. \tag{8}$$

Furthermore, we have

$$c_t = 1 + \frac{1}{k}\ln\frac{N(t)}{N(t-k)}. \tag{9}$$

where $k = t - t_0$ represents the sliding window period.

III. EMPIRICAL ANALYSIS

A. DATA SOURCES

This article selects the COVID-19 epidemic data in Hong Kong as the research object. The data comes from the Department of Health of the Hong Kong Special Administrative Region Government. The start and expiration date of the three rounds of Hong Kong epidemic data intercepted in this article are shown in Table 1.

TABLE 1: Experimental data sets

Stage	Start date	Expiration date
First-round	2020/03/10	2020/04/10
Second-round	2020/07/15	2020/08/15
Third-round	2020/11/12	2020/12/12

B. SLIDING WINDOW PERIOD SELECTION

Through observation and experimentation, this article comprehensively considers the generalization ability of the fitting function, the number of parameters, and whether there is a risk of overfitting, etc. Based on the [6], we choose the three-parameter power function, i.e.,

$$f(t) = ut^{v} + b \tag{10}$$

to fit the discrete value of dynamic transmission rate. Furthermore, we build a nonlinear weighted least squares model as follows

$$\arg\min \sum_{t_i=1}^{n} w_{t_i} [f(t_i) - \hat{c}_{t_i}]^2, \tag{11}$$

where

$$w_{t_i} = t_i / \sum_{t_j=1}^{n} t_j \tag{12}$$

is a nonlinear weighted function at a time t_i and n is the time series length of the fitted set. By solving the above model, we can obtain the values of the unknown parameters of $f(t)$ in (10) so that the concrete form of c_t can be obtained.

To avoid the phenomenon of inflection point prediction lag when the sliding window period is too large, this article stipulates that the value range of the sliding window period is an integer value from $\{1, 2, \ldots, 7\}$, i.e., the maximum sliding window period does not exceed one week.

In this paper, we use mean absolute error (MAE) and root mean square error (RMSE) as the evaluation indicators for selecting the best sliding window period, and they are defined by

$$\mathrm{MAE} = \frac{1}{N} \sum_{t=1}^{N} |c_t - \hat{c}_t| \tag{13}$$

and

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (c_t - \hat{c}_t)^2}, \tag{14}$$

respectively, where $\hat{c}_t$ is the predicted value of dynamic transmission rate at the time t, N is the length of the time series of the predictive value. For selecting the optimal sliding window period, we refer to [6]-[8].

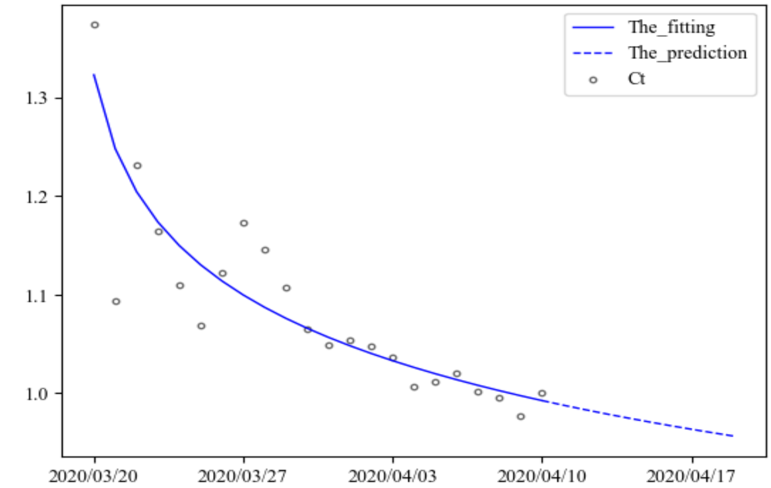
In Table 2, we provide the fitting parameters, the best sliding window period, and the fitting evaluation indexes of the three-parameter power function for each round of the COVID-19 epidemic in Hong Kong.

TABLE 2: Fitting parameters, optimal sliding window period, and evaluation indicators

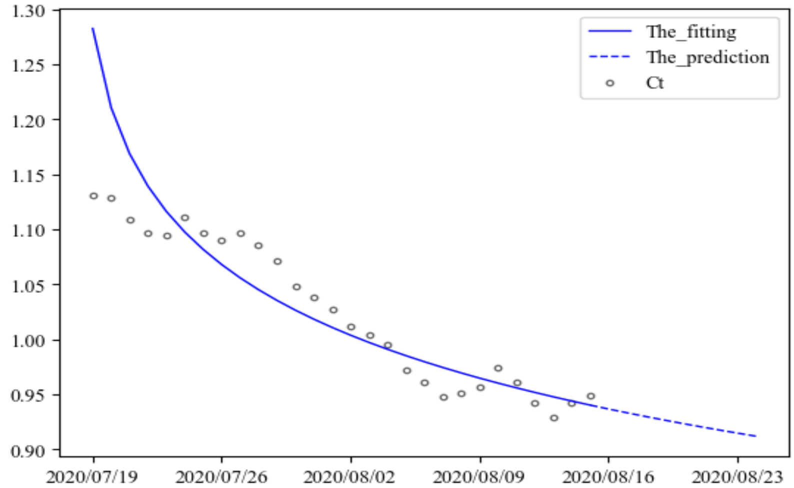
Stage	u	v	b	k	MAE	RMSE
First-round	12.0021	-0.0090	-10.6791	1	0.0272	0.0290
Second-round	14.7654	-0.0070	-13.4828	3	0.0138	0.0167
Third-round	11.5873	-0.0074	-10.3338	7	0.0181	0.0195

C. WEIGHTED FITTING RESULTS

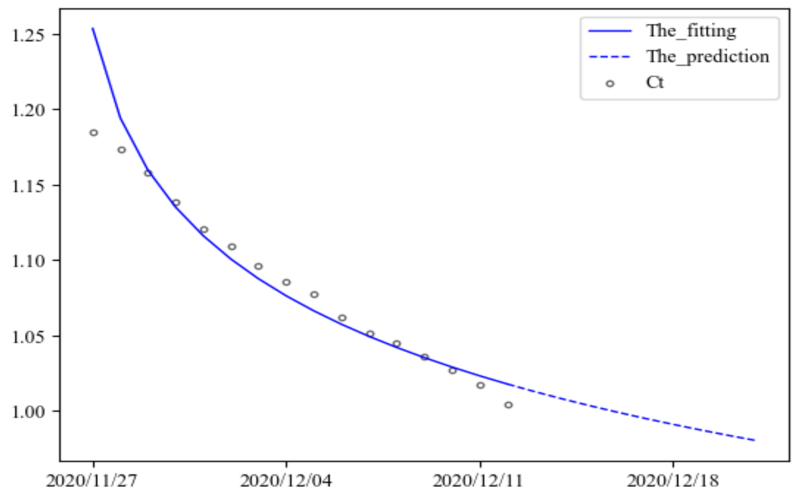
Based on the DTRM, we obtained the fitting results of the three rounds of the COVID-19 epidemic in Hong Kong, as shown in Figure 2.



(a) First round



(b) Second round



(c) Third round

FIGURE 2: The prediction results based on DTRM

It can be seen from Figure 2 that after each outbreak of the epidemic, the Hong Kong government has attached great importance to it and adopted timely control measures. Although the peak of the dynamic transmission rate of the severe epidemic at first was high, with the continuous strengthening of the government's control measures, the dynamic transmission rate declined rapidly, and each round of the epidemic was controlled about half a month after the outbreak. However, due to the continuous rebound of overseas epidemics and imported cases triggering local associations, Hong Kong has a relatively large number of epidemic outbreaks, and a new wave of the rebound has occurred after being contained.

Let $\hat{f}(t) = 1$, we can get the inflection point dates t^* for each of the three rounds of the COVID-19 epidemic in Hong Kong, as shown in Table 3.

TABLE 3: Actual and estimated inflection points

Stage	Actual inflection point	Estimated inflection point
First-round	2020/04/07	2020/04/08
Second-round	2020/08/02	2020/08/02
Third-round	2020/12/09	2020/12/15

From Table 3, we can get the predicted values of the inflection point dates of the three rounds of epidemics as 2020/04/08, 2020/08/02, and 2020/12/15. The error between the prediction result and the actual inflection point date is within a sliding window period, indicating that the model has a good prediction effect.

At the same time, we predict the existing confirmed cases within 7 days after the date where the selected data set ends of the three rounds of the epidemic in Hong Kong from

$$\hat{N}(t) = \frac{1}{k}\sum_{i=1}^{k} \hat{N}(t-i)\exp\{i(\hat{c}_t - 1)\}, \tag{13}$$

where $N(t)$ represents the estimated value of the existing confirmed cases at the time t. When $N(t-i)$ is unknown, $\hat{N}(t-i)$ can be used instead [7].

The forecast number, real number, and deviation rates of the existing confirmed cases in the three rounds of epidemics are shown in Tables 4, 5, and 6, respectively.

TABLE 4: Existing confirmed cases within 7 days of the first round of epidemic

Date	Forecast number	Real number	Deviation rate
2020/04/11	667	660	1.06%
2020/04/12	661	640	3.28%
2020/04/13	656	608	7.95%
2020/04/14	652	574	13.59%
2020/04/15	648	527	23.00%
2020/04/16	645	528	22.10%
2020/04/17	641	485	32.22%

TABLE 5: Existing confirmed cases within 7 days of the second round of epidemic

Date	Forecast number	Real number	Deviation rate
2020/08/16	784	862	9.10%
2020/08/17	756	856	11.67%
2020/08/18	729	812	10.20%
2020/08/19	705	735	4.12%
2020/08/20	683	702	2.76%
2020/08/21	662	656	0.95%
2020/08/22	643	607	6.00%

TABLE 6: Existing confirmed cases within 7 days of the third round of epidemic

Date	Forecast number	Real number	Deviation rate
2020/12/13	1317	1222	7.76%
2020/12/14	1312	1237	6.02%
2020/12/15	1305	1253	4.18%
2020/12/16	1298	1241	4.61%
2020/12/17	1290	1240	4.06%
2020/12/18	1282	1212	5.78%

2020/12/19	1274	1211	5.19%

To evaluate the accuracy of the model, we use mean absolute percentage error (MAPE) as an evaluation index for the accuracy of the number of people prediction, which is given by

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|N(t) - \hat{N}(t)|}{N(t)} \times 100\% . \tag{15}$$

In what follows, the real existing confirmed cases accumulated in 7 days (SUM-RECCA), the predicted existing confirmed cases accumulated in 7 days (SUM-PECCA), and their corresponding MAPEs are shown in Table 7.

TABLE 7: Existing confirmed cases accumulated in 7 days

Stage	SUM-RECCA	SUM-PECCA	MAPE
First-round	4570	4022	14.74%
Second-round	5230	4962	6.40%
Third-round	8616	9078	5.37%

Table 7 shows that all MAPEs are achieved suitable prediction. The MAPE of the three rounds of the epidemic in Hong Kong is 14.74%, 6.40%, and 5.37%, respectively, indicating that the empirical analysis results are satisfactory. Furthermore, we can see that the number of existing confirmed cases in Hong Kong is still not optimistic. Based on the actual situation, it is not difficult to see that the main reasons for the ups and downs of the Hong Kong epidemic are: 1. The confirmed cases involve a wide range of people. 2. There are many loopholes, and prevention and control work still needs to be strengthened. 3. New coronavirus detection capabilities need to be further improved. Therefore, in this context, individuals must take protective measures, and the government should also learn from the prevention and control experience of the Chinese mainland and strive to prevent imports and internal rebounds.

IV. Conclusion

In this paper, we used the DTRM based on nonlinear weighted fitting to predict the existing confirmed cases of the three rounds of the COVID-19 epidemic in Hong Kong. The fitting results are highly consistent with the empirical data, which can provide a scientific reference for the prevention and control of the COVID-19 epidemic in Hong Kong.

The use of machine learning and data assimilation methods to optimize and improve the DTRM is a topic worthy of continued research.

Acknowledgments

This work was supported by the National Statistical Science Research Project of China (No. 2020LY067), and the Shanghai University Student Innovation Training Project (No. CS2021002).

References

[1]. J. Wu, K. Leung, G. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, Lancet, 2020, 395(10225): 689–697.

[2]. Y. Zhang, J. Li, Prediction and analysis of propagation of novel coronavirus pneumonia epidemic based on SIR model, Journal of Anhui University of Technology (Natural Science), 2020, 37(1):94–101 (In Chinese)

[3]. X. Wang, S. Tang, Y. Chen, et al., When will be the resumption of work in Wuhan and its surrounding areas during COVID-19 epidemic? A data-driven network modeling analysis, Scientia Sinica Mathematica, 2020, 50(7):969–978 (In Chinese)

[4]. Z. Yang, Z. Zeng, K. Wang, et al., Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions, Journal of Thoracic Disease, 2020, 12(2):165–174.

[5]. N. Huang, F. Qiao, A data driven time-dependent transmission rate for tracking an epidemic: a case study of 2019-nCoV, Science Bulletin, 2020, 65(06):425–427.

[6]. Y. Hu, J. Kong, L. Yang, et al., A dynamic growth rate model and its application in global COVID-19 epidemic analysis, Acta Mathematicae Applicatae Sinica, 2020, 43(2):452–467 (In Chinese).

[7]. Y. Hu, Y. Liu, L. Wu, et al., A dynamic transmission rate model and its application in epidemic analysis, Operations Research Transactions, 2020, 24(3):27–42 (In Chinese).

[8]. X. Xie, K. Luo, Y. Zhang, et al., Nonlinear combinational dynamic trans-mission rate model and COVID-19 epidemic analysis and prediction in China, Operations Research Transactions, 2021, 25(1):17–30 (In Chinese)

[9]. G.A. McPherson, P. Molenaar, C. Raper, et al, Analysis of dose-response curves and calculation of agonist dissociation constants using a weighted nonlinear curve fitting program, 1983, 10(4):231-41.

TIAN SANG, et. al. " Predicting the Existing Confirmed Cases of the COVID-19 Epidemic in Hong Kong Based on Dynamic Transmission Rate Model." *IOSR Journal of Mathematics (IOSR-JM)*, 17(2), (2021): pp. 39-44.