



Delay Time Parameter and Its Confidence Interval of Predictive Time Series of COVID-19 Outbreak Between the First and the Second Wave

Rapin Sunthornwat¹, Sirikanlaya Sookkhee^{2*}

¹Industrial Technology Program, Faculty of Science and Technology, Pathumwan Institute of Technology, Bangkok, 10330, Thailand

²Department of Mathematics, Faculty of Education, Sisaket Rajabhat University, Sisaket, 33000, Thailand

*E-mail: sirikanlaya.s@sskru.ac.th

Abstract. The outbreak of coronavirus disease 2019 (COVID-19) has become a major problem facing humans all around the world. For governments, in order to deal with the outbreak and protect the population, it is important to predict the number of infectious cases in the future to monitor the COVID-19 situation. This research aimed to compare the effectiveness of the logistic and the delay logistic time series in predicting the total number of infectious cases by using actual data from four countries, i.e. Thailand, South Korea, Egypt, and Nigeria. The total number of COVID-19 cases was collected during the first and the second wave of the COVID-19 outbreak. The validation and accuracy of the predictive growth curve time series were determined based on statistical values, i.e. the coefficient of determination and the root mean squared percentage error. It was found that the logistic time series was more appropriate for predicting the first wave in the four countries. For the second wave, the delay logistic time series was preferable. Moreover, the confidence interval based on Chebyshev's inequality of delay time between the first and the second wave of the COVID-19 outbreak is also proposed.

Keywords: *confidence interval; Coronavirus disease 2019; delay logistic time series; logistic time series.*

1 Introduction

The world has been facing an epidemic of COVID-19, a new strain of coronavirus that first appeared in December 2019 in the capital of Hubei Province, China, where the virus rapidly spread among the residents. The general symptoms after getting COVID-2019 consist of fever, coughing, shortness of breath, weakness, fatigue, headache, and diarrhea [1]-[3]. Measures for alleviating the COVID-19 outbreak have been applied in all countries in the world. During the first wave of the COVID-19 outbreak, many countries imposed lockdowns to suppress the spread of the virus. However, these lockdowns have negatively affected the economy, especially, the industrial,

service, and tourism sectors. After the first wave, the lockdowns were lifted in order to let the economy recover. Movement and migration of people were allowed again. This increased the risk of COVID-19 disease transmission among people [4][5]. Subsequently, a second wave of the COVID-19 outbreak emerged after relaxation of the preventive measures. This caused an increase in the number of COVID-19 cases from the first wave to the second wave. For example, according to the WHO Thailand Situation Report [6], Thailand had approximately 1,300 infected people during the first wave. During the second wave, the total number of cases was 15,465 people.

A model for predicting the number of infected people is an important tool to help governments develop policies to control severe epidemics. A lot of research has been done related to COVID-19 forecasting, using models such as the logistic growth model, the generalized logistic growth model, the Richards model, the simple Gaussian model, the Ratkowsky model, and compartment models for the first as well as for the second wave. Studying the logistic growth model, the generalized logistic growth model and the generalized Richards model for forecasting the number of infected cases in 29 provinces in China and other regions in the world, the findings revealed that different outbreak levels can be classified into three groups. The authors suggest that the forecasts in some countries are largely misleading due to several factors, such as case definition, testing capacity, testing protocols, and reporting system and time [7]. The logistic growth model, the generalized logistic growth model, and the generalized growth model were used for daily forecasting of confirmed cases in India [8]. The compartmental model (generalized SEIR model), the logistic growth model, and the simple Gaussian model for predicting the spread of COVID-19 were studied in Iraq and Egypt [9]. The Richards model, the Gompertz model, the logistic model, the Ratkowsky model, the compartmental model (SIRD model), and the SIR model were investigated for making projections of the COVID-19 pandemic dynamics in Iran [10][11]. The logistics model for estimating the number of COVID-19 cases in Sweden and other Nordic countries, the USA, Brazil, and India has been studied and validated for COVID-19 situation control policies, such as strict lockdown and herd immunity policies [12][13]. A comparison was made between utilizing the logistic growth model and the Gompertz growth model for estimating COVID-19 cumulative cases in Southeast Asian countries [14]. Meanwhile, a second wave has occurred in many countries in Europe, the United States, New York, and Asian countries [15]. The researchers compared an estimate of the second wave using the logistic model with the final size of the second wave. They found that the data followed the logistics curve during the first wave and then started to deviate from it, indicating the beginning of the second wave of the epidemic [16].

A question that challenges statisticians is how to identify when a second wave will occur and how long it will take to enter the so-called second wave phase and how long the delay time between the first and the second wave is, which is important to know when making decisions on policy-level planning to cope with severe pandemic outbreaks.

The present research focused on the delay time parameter and its confidence interval in relation to the first and the second wave of the COVID-19 outbreak based on the total number of COVID-19 cumulative cases in the four sampled countries of this research: Thailand, South Korea, Egypt, and Nigeria. By comparing the parameters of the logistic and the delay logistic growth model, their performance was measured in terms of the coefficient of determination and the root mean squared percentage error.

2 Materials and Methods

This section presents the mathematical and statistical background and the material and methods used in this research.

2.1 Data collection

The data used in this research were gathered from the Worldometers website [17]. This website provides data about the COVID-19 outbreak worldwide. The provided data consists of total coronavirus cases, daily new cases, active cases, total coronavirus deaths, daily new deaths, newly infected, newly recovered, recovery rate, and death rate, etc. However, the total numbers of coronavirus cases, i.e. the cumulative number of COVID-19 cases during the first and the second wave of the COVID-19 outbreak, were collected for this research. The sampled countries for this research were Thailand, South Korea, Egypt, and Nigeria. The period of collected data was from February 15, 2020 ($t = 0$) to January 10, 2021 ($t = 330$).

2.2 Predictive Time Series and Its Parameter Estimation for the COVID-19 Outbreak

The growth curve time series as a predictive time series for describing the COVID-19 outbreak is a solution to the logistic differential equations. The solution is called the logistic growth curve time series, which is a flattened curve after passing its inflection point. This property corresponds to the behavior of the COVID-19 outbreak [13][14]. Let $T(t)$ be a logistic growth curve time series of the total COVID-19 cases at any time t . The logistic differential equation was developed in [18] as:

$$\frac{d}{dt}T(t)=rT(t)\left[1-\frac{T(t)}{C}\right]; T(t=0)=T_0, \tag{1}$$

where T_0 is the initial condition for an infectious COVID-19 case, r is the intrinsic growth rate, t is the time, and C is the carrying capacity.

The Eq. (1) can be solved by a partial fraction and separable method [13][14]. Therefore, the logistic growth curve time series is carried out as follows:

$$T(t)=\frac{C}{(1+K\exp(-rt))}, \tag{2}$$

where $K=\frac{C-T_0}{T_0}$.

Asymptotic behavior of the logistic growth curve time series will converge to the carrying capacity.

The inflection point of logistic growth curve time series is at $\frac{C}{2}$, which is the maximum outbreak point, and it is equivalent to the peak time at $t = \frac{\ln(C)}{r}$. The parameters K, r, C of the logistic growth curve time series can be estimated by the least square error method. Let $e(t)$ be an error function of the difference between the actual and the estimated value at any time t . The sum square error S can be evaluated as follows:

$$S = \sum_{t=1}^n [e(t)]^2 = \sum_{t=1}^n [\text{actual}(t) - \text{estimate}(t)]^2.$$

To minimize the sum square error, the partial derivative and setting to zero are computed by the calculus concept:

$$\begin{aligned} \frac{\partial S}{\partial K} &= 0 \\ \frac{\partial S}{\partial r} &= 0 \\ \frac{\partial S}{\partial C} &= 0 \end{aligned} \tag{3}$$

The optimal parameters or estimated parameters K, r, C can be solved by the least square error method. The validation and accuracy of the predictive growth curve time series are based on statistic the coefficient of determination (R^2) and the root mean squared percentage error (RMSPE) [13][14].

2.3 Logistic Growth Curve Time Series of the Total COVID-19 Outbreak During the First and the Second Wave

The variables related to the analysis of the predictive time series of the first and the second wave of the COVID-19 outbreak are defined as shown in Figure 1. T_0 denotes the initial total number of COVID-19 cases at starting time t_1 of the first wave. C_1 denotes the carrying capacity of the total number of COVID-19 cases at ending time t_{C1} of the first wave. Also, C_1 will become the initial total number of COVID-19 cases at starting time t_2 of the second wave. C_2 denotes the carrying capacity of the total number of COVID-19 cases at ending time t_{C2} of the second wave. I_1 denotes the inflection point of the first wave of the COVID-19 outbreak at time t_{I1} , while I_2 denotes the inflection point of the second wave of the COVID-19 outbreak at time t_{I2} . Moreover, the delay time or lag time between the first and the second wave, denoted by τ , appears before the second wave. Based on the symmetric property of the logistic growth time series, the ending time of the first and the second wave can be evaluated as:

$$t_{C1} = \sqrt{(t_{I1} - t_1)^2 + \left(\frac{C_1}{2} - T_0\right)^2 - \left(C_1 - \frac{C_1}{2}\right)^2} + t_{I1}$$

$$t_{C2} = \sqrt{(t_{I2} - t_2)^2 + \left(\frac{C_2}{2} - C_1\right)^2 - \left(C_2 - \frac{C_2}{2}\right)^2} + t_{I2}$$

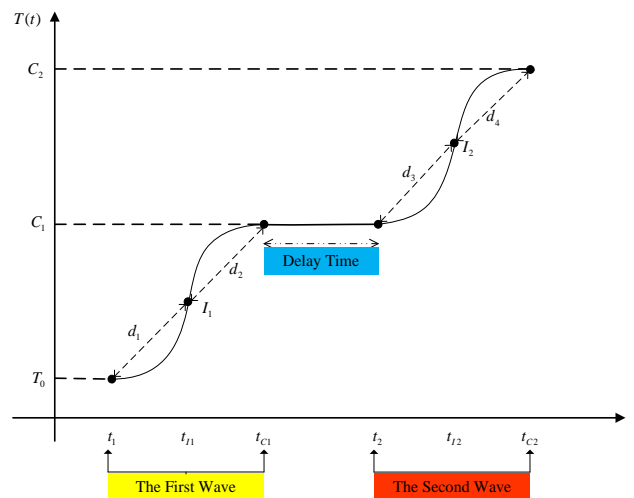


Figure 1 The structure of the predictive time series of the COVID-19 outbreak between the first and the second wave.

2.4 Analysis of Delay Time Between the First and the Second Wave of the COVID-19 Outbreak

2.4.1 Delay logistic growth time series for delay time between the first and the second wave

The delay logistic growth time series, a solution to the delay logistic differential equation, is more generalized than the logistic growth time series in Eq. (2). The logistic growth time series is under the assumption that the process relies on the growth rate of the relative number of individuals. On the other hand, the delay logistic growth time series is under the assumption that the time series process is not instantaneous. Thus, the delay logistic differential equation extended from the logistic differential equation in Eq. (4) can be defined as follows:

$$\frac{d}{dt}T(t) = rT(t-\tau)\left[1 - \frac{T(t)}{C}\right]; T(t = 0) = T_0, \tag{4}$$

where $\tau > 0$ is the delay time or lag time. An analytical solution of the delay logistic differential equation cannot be computed. A numerical solution, the method of steps, was adopted to solve this problem. The method of steps is conducted to transform the delayed logistic differential equations in a given interval to ordinary differential equations over that interval by using the iteration steps of the next interval [19]. Let $T_0(t)$ be the initial condition function for the delay logistic differential equations. The process of the method of steps is given as:

1. 1st step: On the interval $t \in [-\tau, 0]$, then $T(t) = T_0(t)$.
2. 2nd step: On the interval $t \in [0, \tau]$, then $T(t-\tau) = T_0(t-\tau)$. The solution of the delay logistic differential equation $\frac{d}{dt}T(t) = rT_0(t-\tau)\left[1 - \frac{T(t)}{C}\right]$ as given $T_1(t)$.
3. 3rd step: On the interval $t \in [\tau, 2\tau]$, then $T(t-\tau) = T_1(t-\tau)$. The solution of the delay logistic differential equation $\frac{d}{dt}T(t) = rT_1(t-\tau)\left[1 - \frac{T(t)}{C}\right]$ is solved as given $T_1(t)$. The steps are continuously repeated until the desired time subsequent interval is reached.

2.4.2 Cross correlation for estimate delay time

The cross correlation [20] between two time series $x(t)$ and $y(t)$ can be given by:

$X_{\text{corr}_\tau}(x,y) = \int_{-\infty}^{\infty} x(t)y(t-\tau)dt$ for continuous time series $x(t)$ and $y(t)$

$X_{\text{corr}_\tau}(x,y) = \sum_{t=-\infty}^{\infty} x(t)y(t-\tau)$ for discrete time series $x(t)$ and $y(t)$,

where $\tau > 0$ is the delay time or lag time. This is useful for measuring the similarity between two time series and detecting the lag or delay of the two time series.

2.4.3 Confidence interval for delay time

In this research, the delay time between the first and the second wave of the COVID-19 outbreak was assumed to follow a uniform distribution on the interval $[t_{C1}, t_2]$. Let t follow a uniform distribution on interval $[t_{C1}, t_2]$ where t_{C1} is the minimum time and t_2 is the maximum time. The probability density function (pdf) of the uniform distribution is defined as:

$$\text{pdf} : f(t; t_{C1}, t_2) = \begin{cases} \frac{1}{t_2 - t_{C1}} & ; t \in [t_{C1}, t_2] \\ 0 & ; t \notin [t_{C1}, t_2] \end{cases}$$

The cumulative distribution function (cdf) of the uniform distribution is defined as:

$$\text{cdf} : F(t; t_{C1}, t_2) = \begin{cases} 0 & ; t < t_{C1} \\ \frac{t - t_{C1}}{t_2 - t_{C1}} & ; t \in [t_{C1}, t_2] \\ 1 & ; t > t_2 \end{cases}$$

Let M_n be a statistic or an order statistics estimator $\text{Max}(\{t_i \in [t_{C1}, t_2]\})$ of independent and identically distributed random variables $t_i \in [t_{C1}, t_2]$. The cumulative distribution function of M_n can be derived as:

$$\mathbf{Pr}(M_n \leq x) = \mathbf{Pr}([T_1 \leq x] \cap [T_2 \leq x] \cap \dots \cap [T_n \leq x]) = \prod_{i=1}^n \mathbf{Pr}(T_i \leq x).$$

Thus, the cumulative distribution function of the maximum value M_n between t_{C1} and t_2 is evaluated as:

$$\text{cdf} : F_{M_n}(t; t_{C1}, t_2) = \prod_{i=1}^n F_{T_i}(t) = \begin{cases} 0 & ; t < t_{C1} \\ \left(\frac{t - t_{C1}}{t_2 - t_{C1}}\right)^n & ; t \in [t_{C1}, t_2] \\ 1 & ; t > t_2 \end{cases}$$

The probability density function of the maximum value M_n between t_{C1} and t_2 is evaluated as:

$$\text{pdf} : f_{M_n}(t) = \frac{d}{dx} F_{M_n}(t) = \frac{n}{(t_2 - t_{C1})^n} (t - t_{C1})^{n-1} \text{ for } t_{C1} \leq t \leq t_2.$$

The expected value of M_n can be computed as:

$$\begin{aligned} \mathbf{E}(M_n) &= \int_{t_{C1}}^{t_2} t f_{M_n}(t) dt \\ &= \int_{t_{C1}}^{t_2} t \frac{n}{(t_2-t_{C1})^n} (t-t_{C1})^{n-1} dt \\ &= \frac{n}{(t_2-t_{C1})^n} \int_{t_{C1}}^{t_2} t(t-t_{C1})^{n-1} dt \\ &= \frac{n}{(t_2-t_{C1})^n} \left[\frac{(t_2-t_{C1})^{n+1}}{n+1} + \frac{t_{C1}(t_2-t_{C1})^n}{n} \right] \\ &= \frac{n(t_2-t_{C1})}{n+1} + t_{C1}. \end{aligned}$$

The expected value of $(M_n)^2$ can be computed as:

$$\begin{aligned} \mathbf{E}(M_n^2) &= \int_{t_{C1}}^{t_2} t^2 f_{M_n}(t) dt \\ &= \int_{t_{C1}}^{t_2} t^2 \frac{n}{(t_2-t_{C1})^n} (t-t_{C1})^{n-1} dt \\ &= \frac{n}{(t_2-t_{C1})^n} \int_{t_{C1}}^{t_2} t^2 (t-t_{C1})^{n-1} dt \\ &= \frac{n}{(t_2-t_{C1})^n} \left[\frac{(t_2-t_{C1})^{n+2}}{n+2} + \frac{2t_{C1}(t_2-t_{C1})^{n+1}}{n+1} + \frac{(t_{C1})^2(t_2-t_{C1})^n}{n} \right] \\ &= \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2. \end{aligned}$$

The variance of M_n can be computed as:

$$\begin{aligned} \mathbf{Var}(M_n) &= \mathbf{E}(M_n^2) - (\mathbf{E}(M_n))^2 \\ &= \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2 - \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]^2. \end{aligned}$$

The estimator M_n can be transformed to an unbiased estimator by $B_n = \frac{n+1}{n} M_n$, with the expected value and the variance are:

$$\mathbf{E}(B_n) = \frac{n+1}{n} \mathbf{E}(M_n)$$

$$= \frac{n+1}{n} \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]$$

$$\text{Var}(B_n) = \left(\frac{n+1}{n}\right)^2 \text{Var}(M_n)$$

$$= \left(\frac{n+1}{n}\right)^2 \left\{ \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2 - \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]^2 \right\}$$

To determine the confidence interval of the starting time t_2 of the second wave of the COVID-19 outbreak, Chebyshev’s inequality is applied.

$\Pr(|B_n - t_2| \geq \varepsilon) \leq \frac{\text{Var}(B_n)}{\varepsilon^2}$, where ε is any positive real number.

$$\Pr(B_n + \varepsilon \geq t_2 \geq B_n - \varepsilon) \leq \frac{\left(\frac{n+1}{n}\right)^2 \left\{ \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2 - \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]^2 \right\}}{\varepsilon^2}$$

$$\Pr \left(\begin{array}{c} B_n - \sqrt{\frac{\left(\frac{n+1}{n}\right)^2 \left\{ \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2 - \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]^2 \right\}}{\alpha}} \\ \leq t_2 \leq \\ \sqrt{\frac{\left(\frac{n+1}{n}\right)^2 \left\{ \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2 - \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]^2 \right\}}{\alpha}} \end{array} \right) \geq 1 - \alpha$$

where $\varepsilon = \sqrt{\frac{\left(\frac{n+1}{n}\right)^2 \left\{ \frac{n(t_2-t_{C1})^2}{n+2} + \frac{2nt_{C1}}{n+1} (t_2-t_{C1}) + (t_{C1})^2 - \left[\frac{n(t_2-t_{C1})}{n+1} + t_{C1} \right]^2 \right\}}{\alpha}}$.

Then, the confidence interval $(1-\alpha)100\%$ for starting time t_2 of the second wave of the COVID-19 outbreak is $B_n - \varepsilon$ for the lower confidence limit (LCL) and $B_n + \varepsilon$ for the upper confidence limit (UCL). Therefore, the confidence interval $(1-\alpha)100\%$ for the delay time between the first and the second wave of the COVID-19 outbreak is $t_{C1} - \varepsilon$ for LCL and $\frac{n+1}{n} B_n + \varepsilon$ for the UCL.

3 Results and Discussion

In this section, the results of this research are demonstrated and interpreted.

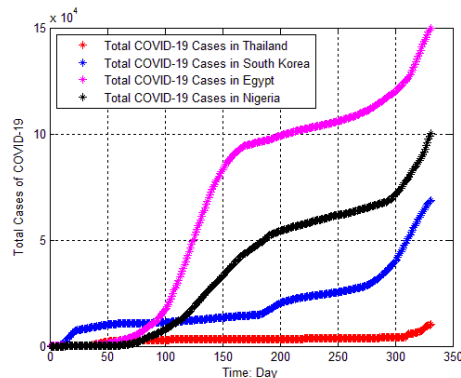


Figure 2 The total numbers of COVID cases for the four sampled countries.

The total numbers of COVID-19 cases for the four countries are shown in Figure 2. Thailand had a relatively low and the flattest infection trend, followed by South Korea, Nigeria, and Egypt, respectively. According to Figure 2, the total number of COVID-19 cases in the four countries from starting point to around day 150 was flat. Then, the rate of increase in the cumulative infections was relatively steady or became slightly higher from day 151 to day 300. Especially the total number of COVID-19 cases in Egypt was quite high compared with the other countries.

3.1.1 Comparison and Analysis of Predictive Time Series on COVID-19 Outbreak Between the First and the Second Wave in Thailand

The logistic and delay logistic time series for the estimated total number of COVID-19 cases during the first wave in Thailand are shown in Figure 3(a) and 3(b), respectively. The real total number of COVID-19 cases is indicated by the circle, while the solid line indicates the estimated total number of COVID-19 cases. It can be seen that the number of infections estimated by the logistic time series was a better match than that of the delay logistic time series for the first wave of the COVID-19 outbreak. The cross correlations showed that the estimated delay time for the second wave was approximately 90 days after the first wave, as can be seen in Figure 4(a). Namely, from day 0 to 330, the first date of the second wave was day 91. The prediction of the time series of the total number of COVID-19 cases during the second wave showed that the delay logistic time series is preferable. It showed that the logistic time series is not suitable for predicting the total number of COVID-19 cases, as shown in Figure 4(b).

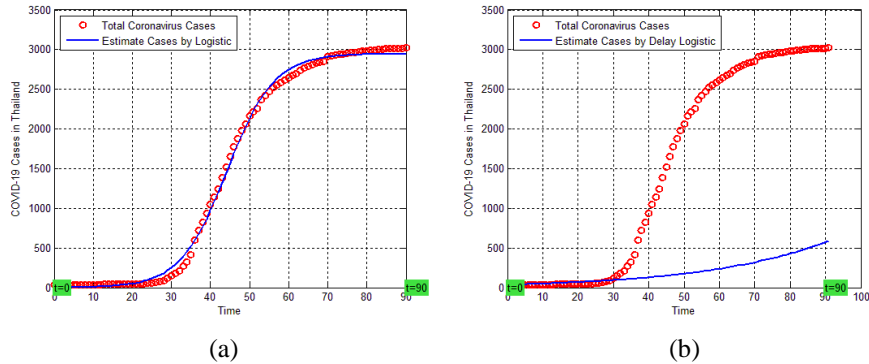


Figure 3 Logistic time series (a) and delay logistic time series (b) for the total number of COVID-19 cases during the first wave in Thailand.

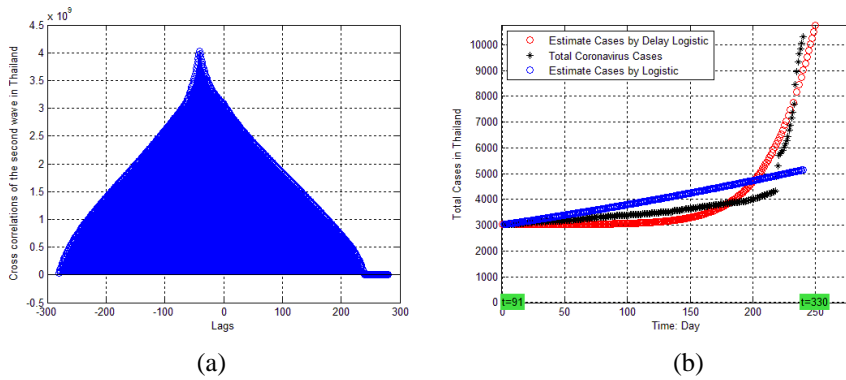


Figure 4 Cross correlations for estimated delay time (a) and predictive time series for the total number of COVID-19 cases (b) during the second wave in Thailand.

3.1.2 Comparison and Analysis of Predictive Time Series on COVID-19 Outbreak Between the First and the Second Wave in South Korea

The logistic and delay logistic time series for the predicted total number of COVID-19 cases for the first wave in South Korea are shown in Figure 5(a) and 5(b), respectively. The real total number of COVID-19 cases is indicated by the circle while the solid line indicates the estimated total number of COVID-19 cases. It can be seen that the estimated total number of COVID-19 cases of the logistic time series was a better match than that of the delay logistic time series for the first wave. The cross correlations showed that the estimated delay time of the second wave was about 85 days after the first wave, as can be seen in Figure 6(a), from day 0 to 330, the first date of the second wave was day 86. The prediction of the time series of the total number of COVID-19 cases for the second wave showed that the logistic time series is preferable. It showed that

the delay logistic time series was not suitable for explaining the total number of COVID-19 cases, as can be seen in Figure 6(b).

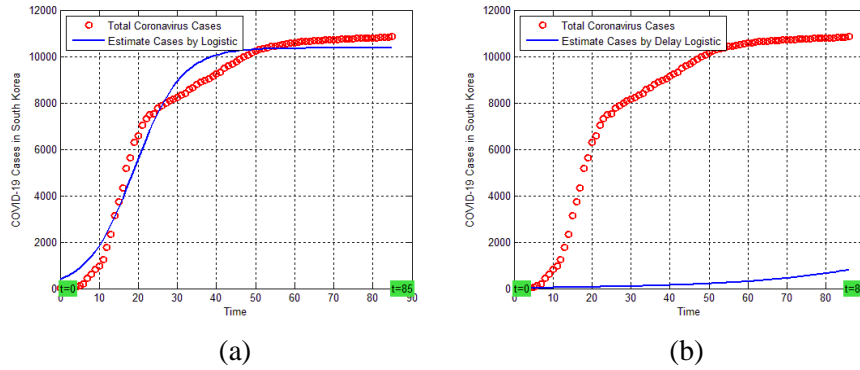


Figure 5 Logistic time series (a) and delay logistic time series for the total number of COVID-19 cases during the first wave in South Korea.

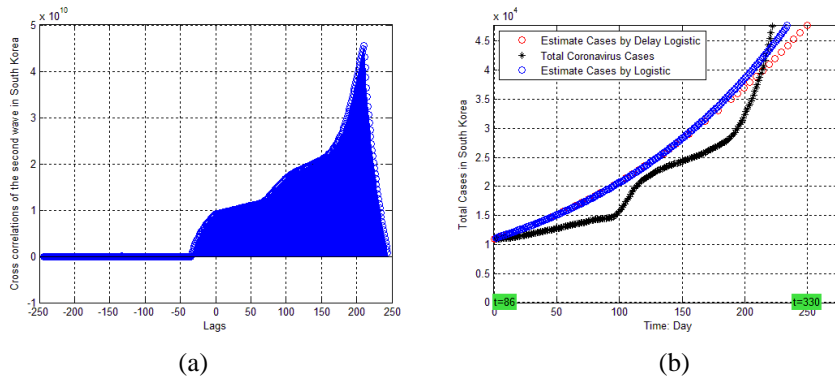


Figure 6 Cross correlations for estimated delay time (a) and predictive time series of the total number of COVID-19 cases (b) during the second wave in South Korea.

3.1.3 Comparison and Analysis of Predictive Time Series on COVID-19 Outbreak Between the First and the Second Wave in Egypt

The logistic and delay logistic time series for the predicted total number of COVID-19 cases during the first wave in Egypt are shown in Figure 7(a) and 7(b), respectively. The real total number of COVID-19 cases is indicated by the circle while the solid line indicates the estimated total COVID-19 cases. It was shown that the estimated total number of cases of infections from the logistic time series was a better match than that of the delay logistic time series for the first wave. The cross correlations showed that the estimated delay time for the second wave was about 220 days after the first wave, as can be seen in Figure

8(a). Namely, from day 0 to 330, the first date of the second wave was day 221. The prediction of the time series of the total number of COVID-19 cases for the second wave showed that the logistic time series and delay logistic time series performed equally well in this case, as shown in Figure 8(b).

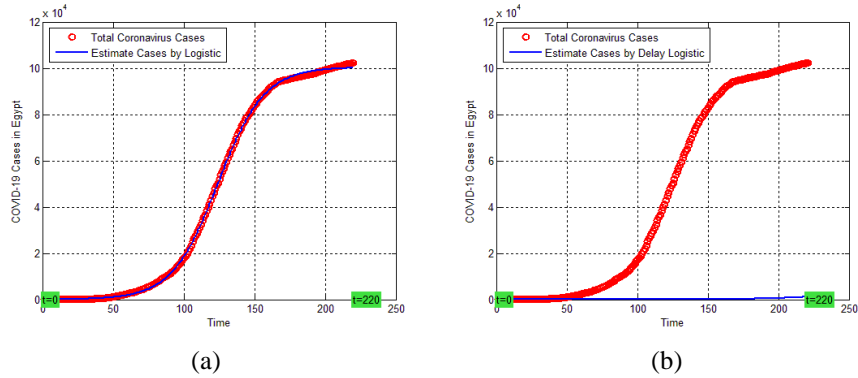


Figure 7 Logistic time series (a) and delay logistic time series (b) for the total number of COVID-19 cases during the first wave in Egypt.

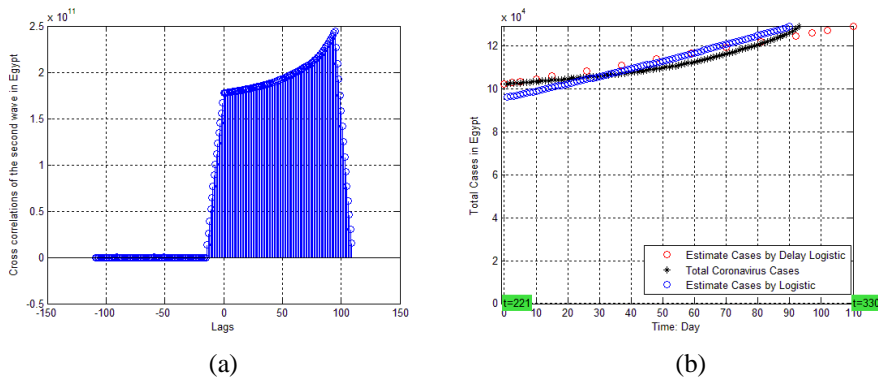


Figure 8 Cross correlations for estimated delay time (a) and predictive time series of the total number of COVID-19 cases (b) during the second wave in Egypt.

3.1.4 Comparison and Analysis of Predictive Time Series on COVID-19 Outbreak between the First and the Second Wave in Nigeria

The logistic and the delay logistic time series for the predicted total number of COVID-19 cases during the first wave in Nigeria are shown in Figure 9(a) and 9(b), respectively. The real total number of COVID-19 cases is indicated by the circle, while the solid line indicates the estimated total number of COVID-19 cases. It can be seen that the estimated total number of cases of infections by the

logistic time series was a better match than that of the delay logistic time series for the first wave. The cross correlations showed that the estimated delay time for the second wave was about 210 days after the first wave, as shown in Figure 10(a). Namely, from day 0 to 330, the first date of the second wave was day 211. The prediction of the time series of the total number of COVID-19 cases for the second wave showed that the logistic time series and the delay logistic time series performed equally well, as can be seen in Figure 10(b).

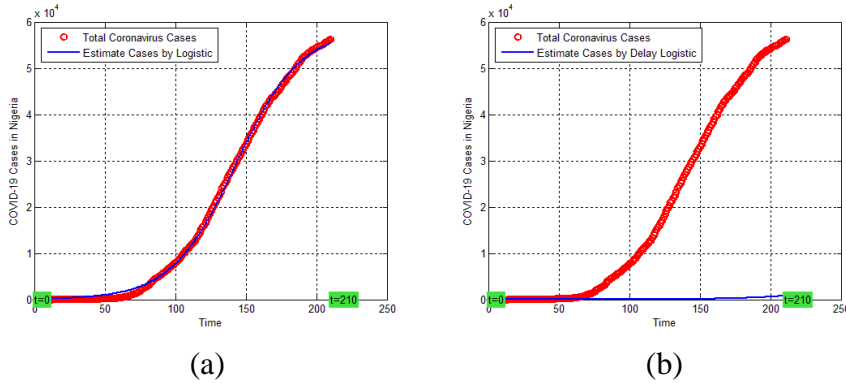


Figure 9 Logistic time series (a) and delay logistic time series (b) for the total number of COVID-19 cases during the first wave in Nigeria.

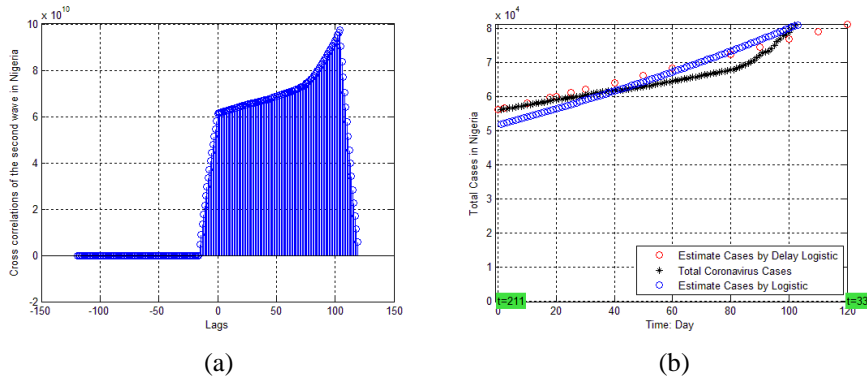


Figure 10 Cross correlations for estimated delay time (a) and predictive time series of the total number of COVID-19 cases (b) during the second wave in Nigeria.

3.1.5 Discussion and Comparison

A comparison of the parameter estimations of the logistic and the delay logistic time series for the first wave and the second wave of the COVID-19 outbreak is provided in Tables 1 and 2, respectively. The accuracy of the predictive time series of the total number of COVID-19 cases from the logistic and delay

logistic time series is based on the maximum of the coefficient of determination, which should approach one, and the minimum of the root mean squared percentage error, which should approach zero.

Table 1 Parameter Estimation of Logistic and Delay Logistic Time Series for COVID-19 Outbreak During the First Wave

Country	T_0	Logistic Time Series			Delay Logistic Time Series			
		r	C	R^2 RMSPE	r	C	τ	R^2 RMSPE
Thailand	34	0.167	2946	0.998* 0.206*	0.191	3025	14	0.862 0.634
South Korea	28	0.167	10358	0.977* 0.668*	0.185	10874	8	0.512 240
Egypt	1	0.061	100385	1.000* 0.968*	0.044	102254	6	0.456 560
Nigeria	1	0.043	58314	0.999* 0.988*	0.0539	112354	12	0.857 615

Note: *appropriate value

Table 1 shows the estimated parameters of the logistic and the delay logistic time series for the the first wave of the COVID-19 outbreak. The validation and accuracy of the logistic time series for the first wave were based on R^2 and RMSPE: Thailand (0.998, 0.206), South Korea (0.977, 0.668), Egypt (1.000, 0.968), and Nigeria (0.999, 0.988). Meanwhile, the validation and accuracy of the delay logistic time series for the first wave were Thailand (0.862, 0.634), South Korea (0.512, 240), Egypt (0.456, 560), and Nigeria (0.857, 615).

Table 2 shows the estimated parameters of the logistic and the delay logistic time series for the second wave of the COVID-19 outbreak. The validation and accuracy of the logistic time series for the second wave were based on R^2 and RMSPE: Thailand (0.473, 0.027), South Korea (0.895, 1.992), Egypt (0.883, 0.019), and Nigeria (0.863, 0.042). Meanwhile, the validation and accuracy of the delay logistic time series for the second wave were: Thailand (0.874, 0.024), South Korea (0.984, 0.038), Egypt (0.983, 0.001), and Nigeria (0.973, 0.003).

Comparing Thailand, South Korea, Egypt, and Nigeria, the longest delay time for the first wave of the COVID-19 outbreak was in Thailand, i.e. 14 days. The longest delay time for the second wave of the COVID-19 outbreak was in South Korea, i.e. 198 days. Based on the accuracy and validation of the predictive time series among the four countries, the logistic time series provided an R^2 larger than and an RMSPE smaller than the delay logistic time series for the first wave. In contrast, the delay logistic time series provided an R^2 larger than and an RMSPE smaller than the logistic time series for the second wave. This implies that the logistic time series was more suitable for estimating the total

number for the first wave of the COVID-19 cases than the delay logistic time series; on the other hand, the delay logistic time series was more suitable for estimating the total number for the second wave of COVID-19 cases than the logistic time series.

Table 2 Parameter Estimation of Logistic and Delay Logistic Time Series for COVID-19 Outbreak During the Second Wave

Country	T ₀	Logistic Time Series			Delay Logistic Time Series			
		r	C	R ² RMSPE	r	C	τ with α = 0.05 (LCL-UCL)	R ² RMSPE
Thailand	3025	0.0028	20587	0.473 0.027	0.0045	20596	41 (86-140)	0.874* 0.024*
South Korea	10874	0.0063	2073313	0.895 1.992	0.0075	137328	198 (81-290)	0.984* 0.038*
Egypt	102254	0.0034	4518732	0.883 0.019	0.0035	299584	95 (216-323)	0.983* 0.001*
Nigeria	56256	0.0044	4746755	0.863 0.042	0.0048	200174	104 (206-322)	0.973* 0.003*

Note: *appropriate value

4 Conclusion

This research focused on the delay time parameters and estimating the confidence of the predictive time series, the logistic and the delay logistic time series for the first and the second wave of the COVID-19 outbreak using the total number of infections from four countries: Thailand, South Korea, Egypt, and Nigeria.

The findings showed that the logistic time series was more suitable for estimating the total number of COVID-19 cases for the first wave in these four countries and the delay logistic time series was more suitable for estimating the total number of COVID-19 cases for the second wave. The maximum delay time shows a slow outbreak; the minimum delay time shows a fast outbreak. For example, Thailand had the longest maximum delay time for the first wave, 14 days, compared to the other countries. This means that there was a relatively slow outbreak in Thailand. Thus, Thailand could effectively control the COVID-19 outbreak during the first wave. However, South Korea could effectively control the COVID-19 outbreak during the second wave, because a maximum delay time of 198 days occurred in South Korea. Egypt had a shorter minimum delay time for the first wave, 6 days, compared to Thailand, South Korea, and Nigeria. This shows that there was a fast outbreak in Egypt during the first wave. During the second wave, Thailand had a minimum delay time of 41 days. This shows that there was a fast outbreak in Thailand during the

second wave. As discussed previously, the logistic time series was more appropriate to forecast the total number of COVID-19 cases for the first wave and the delay logistic time series was more appropriate for predicting the total number of COVID-19 cases during the second wave, but this is based on the numbers from only four countries: Thailand, South Korea, Egypt, and Nigeria. Future research could focus on the total number of COVID-19 cases from other countries in the world.

References

- [1] Jaber, R.M., Mafrachi, B., Al-Ani, A. & Shkara, M., *Awareness and Perception of COVID-19 Among the General Population: A Middle Eastern Survey*, PLOS ONE **16**(4), 2021. DOI: [10.1371/journal.pone.0250461](https://doi.org/10.1371/journal.pone.0250461).
- [2] Duque, M., Lucaccioni, H., Costa, C., Marques, R., Antunes, D., Hansen, L. & Sá Machado, R., *COVID-19 Symptoms: A Case-control Study, Portugal, March-April 2020*, *Epidemiology and Infection*, **149**, 2021. DOI: [10.1017/S095026882100042X](https://doi.org/10.1017/S095026882100042X).
- [3] Massachusetts General Hospital, *Common Symptoms of COVID-19*, <https://www.massgeneral.org/news/coronavirus/symptoms>. (10 February 2021).
- [4] Piccoli, L., Dzankic, J. & Ruedin, D., *Citizenship, Migration and Mobility in a Pandemic (CMMP): A Global Dataset of COVID-19 Restrictions on Human Movement*, PLOS ONE, **16**(3), 2021. DOI: [10.1371/journal.pone.0248066](https://doi.org/10.1371/journal.pone.0248066).
- [5] Simona B. *Pandemic Overview: Migration Problems in the European Union*, *CIFILE Journal of International Law*, **2**(3), pp. 21-31, 2021.
- [6] World Health Organization, *Coronavirus Disease (COVID-19) Pandemic*, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. (10 February 2021)
- [7] Wu, K., Darcet, D., Wang, Q. & Sornette, D., *Generalized Logistic Growth Modeling of the COVID-19 Outbreak: Comparing the Dynamics in the 29 provinces in China and in the Rest of the World*, *Nonlinear Dyn*, **101**, pp. 1561-1581, 2020. DOI: [10.1007/s11071-020-05862-6](https://doi.org/10.1007/s11071-020-05862-6).
- [8] Jain, M., Bhati, P.K., Kataria, P. & Kumar, R., *Modelling Logistic Growth Model for COVID-19 Pandemic in India*, *Proceedings of the 5th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, pp. 784-789, 2020.
- [9] Ibrahim, M.A. & Al-Najafi, A., *Modeling, Control, and Prediction of the Spread of COVID-19 Using Compartmental, Logistic, and Gauss Models: A Case Study in Iraq and Egypt*, *Processes*, **8**(11), p. 1400, 2020. DOI: [10.3390/pr8111400](https://doi.org/10.3390/pr8111400).

- [10] Sedaghat, A., Oloomi Abbas, S.A., Ashtian Malayer, M., Rezaei N. & Mosavi, A., *Coronavirus (COVID-19) Outbreak Prediction Using Epidemiological Models of Richards Gompertz Logistic Ratkowsky and SIRD*, medRxiv, 2020. DOI: [10.1101/2020.03.11.20024901](https://doi.org/10.1101/2020.03.11.20024901)
- [11] Wijaya, K.P., Ganegoda, N., Jayathunga, Y. *et al.* *An Epidemic Model Integrating Direct and Fomite Transmission as well as Household Structure Applied to COVID-19*, J. Math. Industry, **11**(1), 2021. DOI: 10.1186/s13362-020-00097-x.
- [12] Bhardwaj, R., *A Predictive Model for the Evolution of COVID-19*, Trans Indian Natl Acad Eng, **5**, pp. 133-140, 2020. DOI: [10.1007/s41403-020-00130-w](https://doi.org/10.1007/s41403-020-00130-w).
- [13] Sunthornwat, R., & Areepong, Y., *Predictive Model for the Total Daily COVID-19 Cases with Herd Immunity Policy*, Chiang Mai University Journal of Natural Sciences, **20**(1), pp. 1-13, 2021.
- [14] Areepong, Y. & Sunthornwat, R., *Predictive Models for Cumulative Confirmed COVID-19 Cases by Day in Southeast Asia*, CMES-Computer Modeling in Engineering & Sciences, **125**, pp. 927-942, 2020.
- [15] Ganegoda, N.C., Wijaya, K.P., Amadi, M., Erandi, K.K.W.H. & Aldila, D., *Interrelationship Between Daily COVID-19 Cases and Average Temperature as well as Relative Humidity in Germany*, Scientific Reports, **11**(1), p. 11302, 2021. DOI: [10.1038/s41598-021-90873-5](https://doi.org/10.1038/s41598-021-90873-5).
- [16] Batista, M., *Estimation of the Final Size of the Second Phase of the Coronavirus COVID 19 Epidemic by the Logistic Model*, Cold Spring Harbor Laboratory Press, pp. 1-13, 2020. DOI: [10.1101/2020.03.11.20024901](https://doi.org/10.1101/2020.03.11.20024901).
- [17] Worldometer, *Covid-19 Coronavirus Pandemic*, <https://www.worldometers.info/coronavirus/#countries>. (1 March 2021)
- [18] Verhulst, P.F., *Notice Sur La Loi Que LPopulation Suit Dans Son Accroissement*, Correspondence Mathématique et Physique, **1838**, pp. 113-121, 1975.
- [19] Bani-Yaghoub, M., *Analysis and Applications of Delay Differential Equations in Biology and Medicine*, arXiv, 1701.04173, pp. 1-23, 2017.
- [20] Jonathan, Y.S., *Digital Signal Processing: A Computer Science Perspective*, USA: John Wiley & Sons Inc, 2000.