

Original Article

Asian Pacific Journal of Tropical Medicine

doi: 10.4103/1995–7645.321613

5-Year Impact Factor: 2.285

EWMA control chart based on its first hitting time and coronavirus alert levels for monitoring symmetric COVID–19 cases

Areepong Yupaporn¹, Sunthornwat Rapin²✉¹Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand²Industrial Technology Program, Faculty of Science and Technology, Pathumwan Institute of Technology, Bangkok 10330, Thailand

ABSTRACT

Objective: To define the alert levels for the total number of COVID-19 cases derived by using quantile functions to monitor COVID-19 outbreaks *via* an exponentially weighted moving-average (EWMA) control chart based on the first hitting time of the total number of COVID-19 cases following a symmetric logistic growth curve.

Methods: The cumulative distribution function of the time for the total number of COVID-19 cases was used to construct a quantile function for classifying COVID-19 alert levels. The EWMA control chart control limits for monitoring a COVID-19 outbreak were formulated by applying the delta method and the sample mean and variance method. Samples were selected from countries and region including Thailand, Singapore, Vietnam, and Hong Kong to generate the total number of COVID-19 cases from February 15, 2020 to December 16, 2020, all of which followed symmetric patterns. A comparison of the two methods was made by applying them to a EWMA control chart based on the first hitting time for monitoring the COVID-19 outbreak in the sampled countries and region.

Results: The optimal first hitting times for the EWMA control chart for monitoring COVID-19 outbreaks in Thailand, Singapore, Vietnam, and Hong Kong were approximately 280, 208, 286, and 298 days, respectively.

Conclusions: The findings show that the sample mean and variance method can detect the first hitting time better than the delta method. Moreover, the COVID-19 alert levels can be defined into four stages for monitoring COVID-19 situation, which help the authorities to enact policies that monitor, control, and protect the population from a COVID-19 outbreak.

KEYWORDS: COVID-19 alert levels; Symmetric pattern of the total number of COVID-19 cases; Monitoring COVID-19 situation; EWMA control chart

1. Introduction

Coronavirus disease 2019 (COVID-19) is caused by a single-stranded RNA virus, SARS-COV-2, which was firstly reported in Wuhan, Hubei, China, in the end of 2019. Common symptoms of COVID-19 infection are fever, fatigue, body ache, dry cough, appetite loss, and loss of the sense of smell[1,2]. The COVID-19 outbreak has affected the interaction and activities of people abiding by the strict measures to control COVID-19 outbreaks, such as mask-wearing, washing hands often, physical and social distancing, quarantining after traveling both between and within countries. COVID-19 has caused a new norm in the human lifestyle and has affected the economic and social systems of most countries in the world.

Emerging technologies are being deployed to combat the COVID-19 pandemic. For example, smartphone applications are being used

Significance

COVID-19 has become a global concern. Monitoring the coronavirus alert levels is an important process. Control chart should be used to monitor the outbreak severity. This research applies statistical methods, EWMA control chart and quantile function, to monitor the total COVID-19 cases and classifies the coronavirus alert levels into 4 levels: very high, high, medium, and low, which help the authorities to enact policies that monitor, control, and protect the population from a COVID-19 outbreak.

✉To whom correspondence may be addressed. E-mail: rapin@pit.ac.th

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

©2021 Asian Pacific Journal of Tropical Medicine Produced by Wolters Kluwer-Medknow. All rights reserved.

How to cite this article: Yupaporn A, Rapin S. EWMA control chart based on its first hitting time and coronavirus alert levels for monitoring symmetric COVID-19 cases. Asian Pac J Trop Med 2021; 14(8): 364-374.

Article history: Received 20 April 2021 Revision 2 July 2021
Accepted 6 July 2021 Available online 23 July 2021

for connecting infected patients and the management information systems of hospitals. The rapid deployment of applications for self-screening, self-health check-ups, and COVID-19 hospital information systems, along with surveillance, testing, tracing, quarantining, and hospital management to combat the COVID-19 outbreak have been forthcoming[3]. Artificial intelligence is being used in systems to detect SARS-COV-2, fever, and other suspected symptoms of COVID-19. Telemedicine and mobile applications in response to the COVID-19 situation have been deployed to enable consultations on medical conditions *via* video calls with the physicians. In addition, drones and robotics are being used in hazardous areas infected with SARS-COV-2[4]. The Internet of Things is being used for COVID-19 patients to connect to hospital networks, while blockchain is being used for effective supply chain management. Moreover, cloud computing and big data have been deployed to store and provide data storage capacity for information about the COVID-19 situation[5].

Plans and policies have been laid out to handle, control, and monitor the COVID-19 outbreak. Recognizing the pattern of the total number of COVID-19 cases is important for acting on the policies to control COVID-19 outbreaks. The levels of COVID-19 severities which are called the COVID-19 alert levels have been defined by the authorities based on the COVID-19 situation. Moreover, the statistical control chart has been applied to monitor and control the COVID-19 outbreak. Thus, many researches have been conducted on the symmetric pattern of the total number of COVID-19 cases, COVID-19 alert levels, and statistical control charts for monitoring the COVID-19 outbreak. For example, the total number of COVID-19 cases in many countries, such as Thailand, Singapore, Sweden and Norway, follows a logistic growth curve, which is a symmetric pattern[6,7]. The number of cases relative to the maximum cumulative infected, the total confirmed, and the total recovered cases have been fitted to a logistic growth curve model for the COVID-19 suppression dynamics for each region in China[8]. Moreover, several countries and states have classified the COVID-19 alert levels for monitoring the COVID-19 situation and applied measures that people must abide by to protect the population from COVID-19 outbreak spikes. For example, New Zealand classified COVID-19 severity with a COVID-19 alert system[9]. The state of Alaska classified three statewide alert levels for evaluating the COVID-19 transmission risk[10]. The UK defined COVID-19 alert levels to communicate the current risk of COVID-19 transmission to the population[11].

Statistical control charts such as the exponentially weighted moving-average (EWMA) control chart and the proportional control chart have been adopted to monitor COVID-19 outbreaks. The EWMA control chart is being used for monitoring and producing

warnings of COVID-19 outbreaks in some major states in India[12], as well as to monitor and identify the control limits of the death variation of the COVID-19 situation in Pakistan[13]. The proportional control chart is being used to monitor the COVID-19 positivity rate, COVID-19 healing rate, and COVID-19 mortality rate over time[14]. Using quantitative methods to set the COVID-19 alert levels is a useful approach for supporting the authorities when making decisions about launching measures to protect the population during a COVID-19 outbreak. Therefore, the objective of the present research is to apply a quantitative and probability method (a quantile function) that is an inversion of the cumulative distribution function of the time over the total number of COVID-19 cases to set the COVID-19 alert levels when monitoring COVID-19 outbreaks by using a EWMA control chart. The two methods for estimating the expected value and variance of the COVID-19 cases as the population for the EWMA statistic and the results from the EWMA control chart based on the optimal first hitting time were evaluated and compared for COVID-19 outbreaks in Thailand, Singapore, Vietnam, and Hong Kong where the pattern of the total number of COVID-19 cases follows a logistic growth curve, which is a symmetric pattern.

2. Subjects and methods

This section contains the data collection protocol, mathematical and statistical methods, probabilistic analysis, COVID-19 alert levels, and the EWMA control chart for describing and monitoring a COVID-19 outbreak.

2.1. Data collection

The data for this research is the total number of COVID-19 cases over time. The total number of COVID-19 cases, which are secondary data, were collected and downloaded from the website Worldometer[15]. The total number of COVID-19 cases was specifically selected as the sample because the theories and analyses reported herein that comprise the main purpose for this research are based on the symmetric growth curve. The collected data of the total number of COVID-19 cases was complete with no missing values. The sampled countries or region for this research (Thailand, Singapore, Vietnam, and Hong Kong) were specifically selected because their total numbers of COVID-19 cases follow a symmetric growth curve (the logistic growth curve), and are suitable for estimation based on the symmetric pattern of the total number of COVID-19 cases[6] for the duration of February 15, 2020, to December 16, 2020.

2.2. The model for describing the behavior of a COVID-19 outbreak

The behavior of a COVID-19 outbreak can be represented by a logistic differential equation of the total number of COVID-19 cases[6,7]. The total number of COVID-19 cases can be described by using a logistic growth curve that is the solution for the logistic differential equation. Let $I(t)$ be the total cases at any time $t \geq 0$ and C be the carrying capacity of the environment. Thus, the logistic differential equation can be modeled as

$$\frac{dI(t)}{dt} = rI(t) \left(1 - \frac{I(t)}{C} \right) \tag{1}$$

where r is an intrinsic growth rate. One of the variables in Equation (1), $I(t)$, can be evaluated as a logistic growth curve (a symmetric growth curve) as follows:

$$I(t) = \frac{C}{1 + k \exp(-rt)}; k = \frac{C - I_0}{I_0}; C > I_0; I_0 > 0. \tag{2}$$

where I_0 is the initial total number of COVID-19 cases. The logistic growth curve in Equation (2) has an inflection point or median at $C/2$. Moreover, the parameters for the logistic growth curve consist of C , k and r . The least-squares-error method can be adopted to estimate these parameters as follows:

$$e(t) = \sum_{t=0}^n \left(I(t) - \hat{I}(t) \right)^2$$

where $e(t)$ is the error between actual and estimated values at time $t=0, 1, \dots, n$ and $\hat{I}(t)$ is the estimated logistic growth curve for time $t=0, 1, \dots, n$. Hence, the partial derivative with respect to parameters C , k and r can be applied to respectively estimate these parameters as:

$$\begin{aligned} \frac{\partial}{\partial C} e(t) &= 0 \\ \frac{\partial}{\partial k} e(t) &= 0 \\ \frac{\partial}{\partial r} e(t) &= 0 \end{aligned} \tag{3}$$

The parameters C , k and r can be solved from the system of Equation (3).

The accuracy of the estimation parameters and estimated logistic growth curve which represents the total number of COVID-19 cases is measured as the Root Mean Square Percentage Error (RMSPE).

$$RMSPE = \sqrt{\frac{\sum_{t=0}^n \left(\frac{I(t) - \hat{I}(t)}{I(t)} \right)^2}{n}}$$

The explanation of the independent variable (time) in relation to the dependent variable (estimated logistic growth) can be measured by using the coefficient of determination (R^2).

$$R^2 = \frac{\sum_{t=1}^n \left(\hat{I}(t) - \bar{I}(t) \right)^2}{\sum_{t=1}^n \left(I(t) - \bar{I}(t) \right)^2}$$

2.3. The probability density function, expected value, and variance of the time for the total number of COVID-19 cases

The cumulative distribution function (CDF: F) of the time for the total number of COVID-19 cases is defined as:

$$F(t) = \frac{1}{1 + k \exp(-rt)} = [1 + k \exp(-rt)]^{-1} \tag{4}$$

The probability density function (PDF: f) of the time for the total number of COVID-19 cases is a derivative of the CDF expressed as:

$$f(t) = \frac{d}{dt} F(t) = \frac{rk \exp(-rt)}{(1 + k \exp(-rt))^2} \tag{5}$$

The quantile function (QF: Q_p) with probability p is defined as

$$Q_p = \min \{ t \in \text{Real} : F(t) \geq p \}; p \in (0, 1)$$

In other words, Q_p gives the minimum value of time t where the cumulative distribution function exceeds the probability p .

$$Q_p = t = \frac{1}{r} \ln \left(\frac{pk}{1-p} \right); p \in (0, 1) \tag{6}$$

Especially,

$$\begin{aligned} Q_{0.01} &= \frac{1}{r} \ln \left(\frac{1}{99} k \right) \\ Q_{0.25} &= \frac{1}{r} \ln \left(\frac{1}{3} k \right) \\ Q_{0.50} &= \text{Median} = \frac{1}{r} \ln(k) \\ Q_{0.75} &= \frac{1}{r} \ln(3k) \\ Q_{0.99} &= \frac{1}{r} \ln(99k) \end{aligned}$$

The expected value $E(T)$ and variance $Var(T)$ of the time (T) for the total number of COVID-19 cases can be respectively calculated by using the moment generating function (m), Beta function (B), and Gamma function (Γ) as:

$$\begin{aligned} E(T) &= \frac{d}{ds} m(s) \Big|_{s=0} \\ &= k^{\frac{s}{r}} \ln(k) \frac{1}{r} B \left(1 + \frac{s}{r}, 1 - \frac{s}{r} \right) \Big|_{s=0} + k^{\frac{s}{r}} \left[\Gamma \left(1 + \frac{s}{r} \right) \frac{1}{r} \Gamma \left(1 - \frac{s}{r} \right) + \Gamma \left(1 + \frac{s}{r} \right) \Gamma \left(1 - \frac{s}{r} \right) \left(-\frac{1}{r} \right) \right] \Big|_{s=0} \end{aligned}$$

$$= \frac{\ln(k)}{r} \quad (7)$$

$$\begin{aligned} E(T^2) &= \left. \frac{d^2}{ds^2} m(s) \right|_{s=0} \\ &= \frac{\ln(k)}{r} \left[k^{\frac{s}{r}} \frac{\ln(k)}{r} B\left(1 + \frac{s}{r}, 1 - \frac{s}{r}\right) + k^{\frac{s}{r}} \frac{d}{ds} \left(B\left(1 + \frac{s}{r}, 1 - \frac{s}{r}\right) \right) \right] \Bigg|_{s=0} \\ &\quad + \left[\frac{d}{ds} \left(B\left(1 + \frac{s}{r}, 1 - \frac{s}{r}\right) \right) k^{\frac{s}{r}} \frac{\ln(k)}{r} + k^{\frac{s}{r}} \frac{d^2}{ds^2} \left(B\left(1 + \frac{s}{r}, 1 - \frac{s}{r}\right) \right) \right] \Bigg|_{s=0} \\ &= \left(\frac{\ln(k)}{r} \right)^2 + \frac{\pi^2}{3r^2} \end{aligned}$$

$$\begin{aligned} \text{Var}(T) &= \sigma_T^2 = E(T^2) - (E(T))^2 \\ &= \frac{\pi^2}{3r^2} \end{aligned} \quad (8)$$

Next, the estimate expected value and estimate variance of the total COVID-19 cases are evaluated.

2.4. Estimating the expected value and variance of the total number of COVID-19 cases

These can be calculated *via* two methods.

2.4.1. The delta method^[16]

Obviously, the total number of COVID-19 cases I is a function of time t , i.e., $I=I(t)$. The respective estimated expected value and variance of random variable I based on a Taylor series expansion about $t=E(T)$ are:

$$\begin{aligned} E(I) &= I(E(T)) + \frac{1}{2} I''(E(T)) \text{Var}(T) + \frac{1}{6} I'''(E(T)) \text{Var}(T) E[(t-E(T))^3] + \dots \\ \text{Var}(I) &= \sigma_I^2 = \text{Var}[IE(T)] + (I-E(T))I'(E(T)) = [I'(E(T))]^2 \text{Var}(T) + \dots \end{aligned}$$

Therefore, the second order approximation for the expected value of I is:

$$E(I) \approx I(E(T)) + \frac{1}{2} \frac{d^2}{dt^2} (I(E(T))) \text{Var}(T) \quad (9)$$

The first order approximation for the variance of I is:

$$\text{Var}(I) = \sigma_I^2 \approx \left(\frac{dI}{dt} I(E(T)) \right)^2 \text{Var}(T) \quad (10)$$

2.4.2. The sample mean and variance method

The expected value of I can be estimated by using the sample mean as:

$$E(I) \approx \bar{I} = \frac{\sum_{t=1}^n I(t)}{n} \quad (11)$$

The variance of I can be estimated by using the sample variance as:

$$\text{Var}(I) = \sigma_I^2 \approx \frac{\sum_{i=1}^n (I(i) - \bar{I})^2}{n-1} \quad (12)$$

2.5. The exponentially weighted moving average control chart for monitoring a COVID-19 outbreak

The EWMA control chart^[17-19] is based on EWMA statistic $z(t)$ for smoothing parameter $\lambda \in (0, 1]$, which can be written as:

$$Z(t) = \lambda I(t) + (1-\lambda)z(t-1); \quad t=1, 2, 3, \dots, n; \quad Z(0) = E(I) \quad (13)$$

In this research, the EWMA control chart was applied for monitoring the total number of COVID-19 cases. Thus, the assumption is that the $I(t)$'s are independent identically distributed random variables with expected value $E(I(t)) = E(I)$ and variance

$$\sigma_{I(t)}^2 = \sigma_I^2.$$

Therefore, expected value $E(z)$ of the EWMA statistic is:

$$E(z) = \lambda E(I) + (1-\lambda)E(z(t-1))$$

$$E(z) = E(I)$$

and variance $\sigma_{z(t)}^2 = \sigma_z^2$ of the EWMA statistic is:

$$\sigma_z^2 = \frac{\lambda}{2-\lambda} [1 - (1-\lambda)^2] \sigma_I^2$$

The control limits of the EWMA control chart consist of:

$$\text{Upper control limit (UCL)} = E(z) + d\sigma_z$$

$$\text{Center line (CL)} = E(z)$$

$$\text{Lower control limit (LCL)} = E(z) - d\sigma_z$$

where d is the width of the control limit.

Moreover, the total number of COVID-19 cases is estimated *via* logistic growth, which is a gradual increase in the first period, exponential growth in the middle period, and slow growth in the end period. To monitor the total number of COVID-19 cases, the first hitting time (FHT), or first passage time^[20-22], is the run length of the number of samples at which point the EWMA control chart gives the first signal over the upper control limit. The first hitting time of the EWMA control chart is defined as:

$$FHT = \inf\{t > 0 \mid z(t) > UCL\}$$

The first hitting time for monitoring a COVID-19 outbreak through the EWMA control chart should be as large as possible because it reveals the efficiency of controlling the COVID-19 outbreak.

2.6. The COVID-19 alert levels

Here, we present the COVID-19 alert levels for monitoring a COVID-19 outbreak. The COVID-19 alert levels are important for the authorities to make decisions on monitoring, controlling, and protecting the population during a COVID-19 outbreak. In this research, the COVID-19 alert levels were classified into four levels. However, the total number of COVID-19 cases for a logistic growth curve can be described by a symmetric growth curve.

Figure 1 shows a plot of the quantile functions corresponding to the COVID-19 alert levels. The total number of COVID-19 cases starts increasing over the range $(Q_{0.01}, Q_{0.25})$ in alert level 4 (L_4) with very high outbreak severity. Subsequently, the total number of COVID-

19 cases exponentially increases over the range ($Q_{0.25}, Q_{0.50}$) in alert level 3 (L_3) with high outbreak severity. After the inflection point (median), the total number of COVID-19 cases starts decreasing over the range ($Q_{0.50}, Q_{0.75}$) in alert level 2 (L_2) with medium outbreak severity. Following this, the total number of COVID-19 cases significantly decreases and then becomes stable over the range ($Q_{0.75}, Q_{0.99}$) in alert level 1 (L_1) with low outbreak severity.

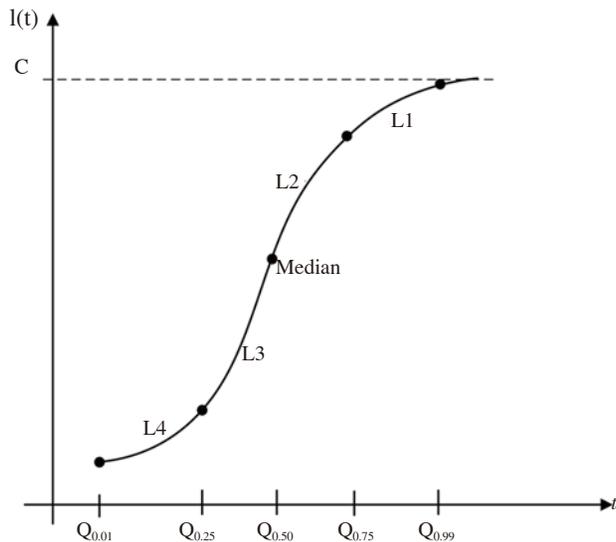


Figure 1. The COVID-19 alert levels and a symmetric growth curve (logistic growth curve) of the total number of COVID-19 cases.

3. Results

3.1. The COVID-19 outbreak in Thailand

As shown in Figure 2A, the estimated logistic growth curve for the total number of COVID-19 cases in Thailand from February 15, 2020, to December 16, 2020, is:

$$I(t) = \frac{3462.31}{1 + 108.59 \exp(-0.09t)}$$

and the quantile functions of the total number of COVID-19 cases are 0.96, 37.44, 48.91, 60.36 and 96.84.

As shown in Figure 2B, the PDF of the time for the total number of COVID-19 cases in Thailand is positively skewed. The probability of the total number of COVID-19 cases in Thailand from February 15, 2020, to December 16, 2020, reached the maximum value at around day 50 and approached zero at around day 120. The expected value and variance of the times for the COVID-19 outbreak in Thailand are 319.08 and 1.52×10^4 , respectively. The sample value and variance of the COVID-19 outbreak in Thailand are 2.89×10^3 and 1.38×10^6 , respectively.

As shown in Figure 2, the EWMA control chart was applied for monitoring the total number of COVID-19 cases in Thailand from February 15, 2020, to December 16, 2020. The results from the

experiment, the first hitting time based on the delta method (method 1) was 0 for all values of the smoothing parameter (Figure 2C) whereas the sample mean and variance method (method 2) (Figure 2D) for smoothing parameter $\lambda=0.3$ and $d=2$ more effectively detected the first hitting time (280 days) than with the other smoothing parameter values. Hence, the optimal first hitting time on the EWMA control chart for monitoring a COVID-19 outbreak in Thailand is approximately 280 days for $\lambda=0.3$ and $d=2$.

3.2. The COVID-19 outbreak in Singapore

As shown in Figure 3A, the estimated logistic growth curve for the total number of COVID-19 cases in Singapore from February 15, 2020, to December 16, 2020, is:

$$I(t) = \frac{57314.95}{1 + 68.59 \exp(-0.04t)}$$

and the quantile functions of the total number of COVID-19 cases are: -8.71, 74.25, 100.32, 126.39 and 209.35.

As shown in Figure 3B, the PDF of the total number of COVID-19 cases in Singapore is positively skewed. The probability of the total number of COVID-19 cases in Singapore from February 15, 2020, to December 16, 2020, reached the maximum value at around day 100 and approached zero at around day 290. The expected value and variance of the times for the COVID-19 outbreak in Singapore are 200.57, and 7.40×10^3 , respectively. The sample value and variance for the COVID-19 outbreak in Singapore are 3.79×10^4 , and 5.13×10^8 , respectively.

As shown in Figure 3, the EWMA control chart was applied for monitoring the total number of COVID-19 cases in Singapore from February 15, 2020, to December 16, 2020. The results from the experiment, the first hitting time based on the delta method (method 1) was 0 for all values of the smoothing parameter (Figure 3C) whereas the sample mean and variance method (method 2) (Figure 3D) for smoothing parameter $\lambda=0.3$ and $d=2$ more effectively detected the first hitting time (208 days) than for the other smoothing parameter values. Thus, the optimal first hitting time on the EWMA control chart for monitoring the COVID-19 outbreak in Singapore is approximately 208 days for $\lambda=0.3$ and $d=2$.

3.3. The COVID-19 outbreak in Vietnam

As shown in Figure 4A, the estimated logistic growth curve for the total number of COVID-19 cases in Vietnam from February 15, 2020 to December 16, 2020, is:

$$I(t) = \frac{1509.36}{1 + 28.43 \exp(-0.02t)}$$

and the quantile functions of the total number of COVID-19 cases are: -66.42, 119.69, 178.16, 236.63 and 422.74.

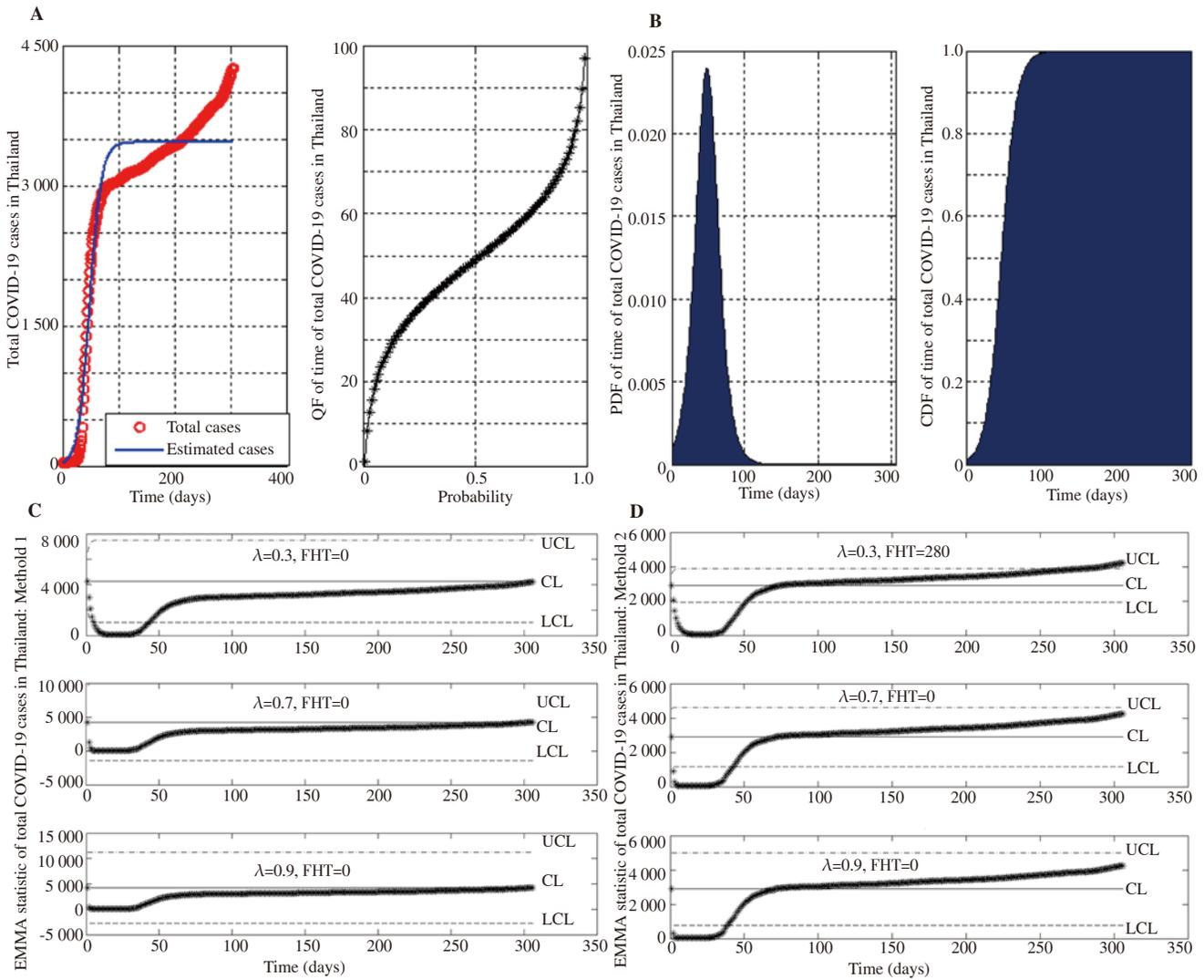


Figure 2. (A) Estimated total number of COVID-19 cases and quantile function of the time for the total number of COVID-19 cases in Thailand and (B) the probability density function and cumulative density function of the time for the total number of COVID-19 cases in Thailand. The EWMA control chart for monitoring a COVID-19 outbreak in Thailand: The expected value and variance estimated *via* (C) the delta method and (D) the sample mean and variance method. FHT: first hitting time; λ : Smoothing parameter. UCL: Upper control limit; CL: Center line; LCL: Lower control limit.

As shown in Figure 4B, the PDF of the total number of COVID-19 cases in Vietnam is quite symmetric. The probability of the total number of COVID-19 cases in Vietnam from February 15, 2020, to December 16, 2020, reached the maximum value at around day 170. The expected value and variance of the time for the total number of COVID-19 cases in Vietnam are 240.17 and 1.69×10^4 , respectively. The sample value and variance of the COVID-19 outbreak in Vietnam are 646.02, and 2.11×10^5 , respectively.

As shown in Figure 4, the EWMA control chart was applied for monitoring the total number of COVID-19 cases in Vietnam from February 15, 2020, to December 16, 2020. The results from the experiment, the first hitting time based on the delta method (method 1) was 0 for all values of the smoothing parameter (Figure 4C) whereas the sample mean and variance method (method 2) (Figure 4D) detected first hitting times of 198, 286, and 0 for $\lambda=0.7$ and $d=2$,

respectively. For smoothing parameter $\lambda=0.3, 0.7$ and 0.9 , the sample mean and variance method more effectively detected the first hitting time (286 days) than for the other smoothing parameter values. Hence, the optimal first hitting time on the EWMA control chart for monitoring the COVID-19 outbreak in Vietnam is approximately 286 days for $\lambda=0.7$ and $d=2$.

3.4. The COVID-19 outbreak in Hong Kong

As shown in Figure 5A, the estimated logistic growth curve for the total number of COVID-19 cases in Hong Kong from February 15, 2020, to December 16, 2020, is:

$$I(t) = \frac{6618.04}{1 + 59.99 \exp(-0.02t)}$$

and the quantile functions of the total number of COVID-19 cases

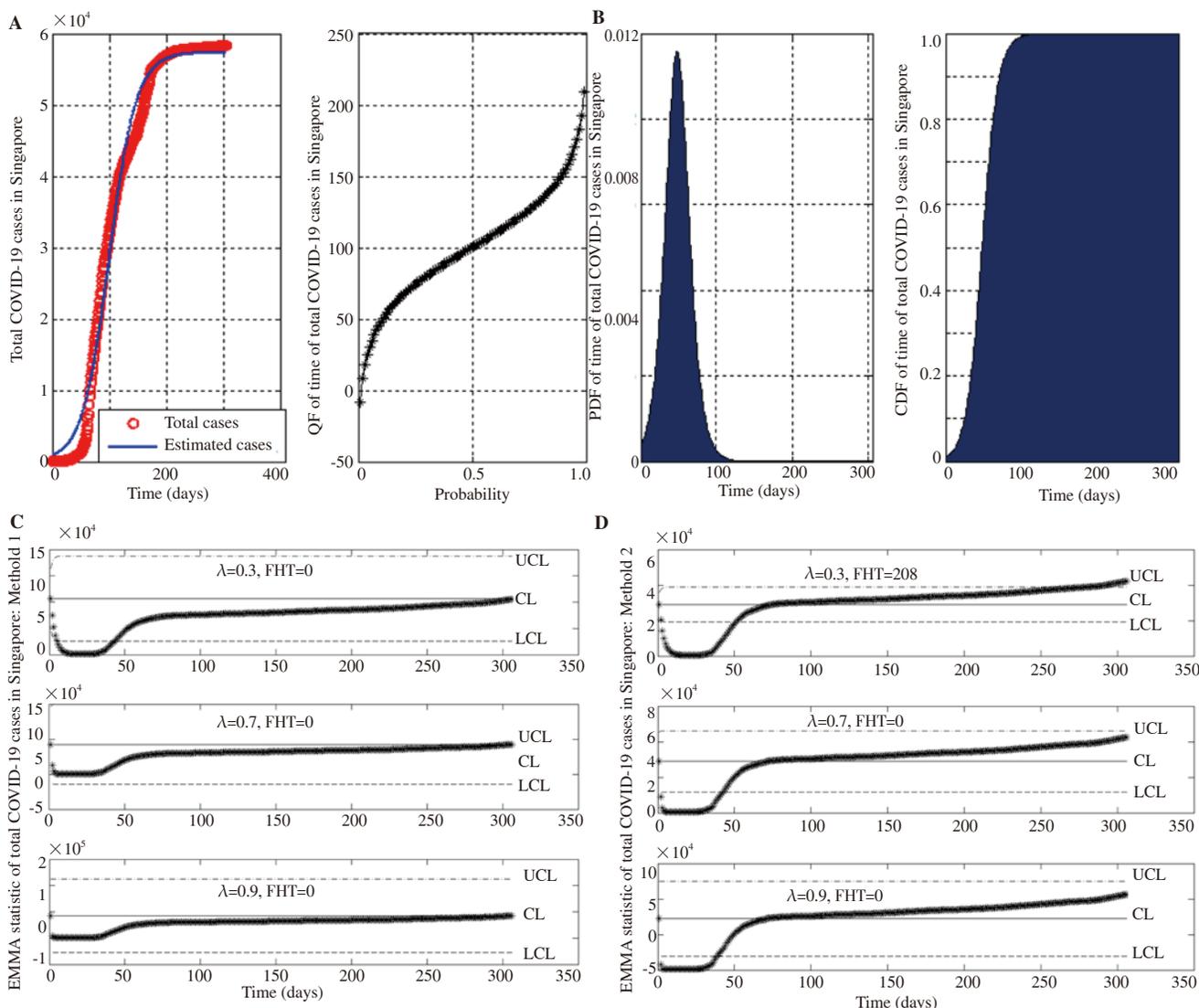


Figure 3. (A) Estimated total number of COVID-19 cases and quantile function of the time for the total number of COVID-19 cases in Singapore and (B) the probability density function and cumulative density function of the time for the total number of COVID-19 cases in Singapore. The EWMA control chart for monitoring a COVID-19 outbreak in Singapore: The expected value and variance estimated *via* (C) the delta method and (D) the sample mean and variance method. FHT: First hitting time; λ : Smoothing parameter. UCL: Upper control limit; CL: Center line; LCL: Lower control limit.

are: -21.24, 127.04, 173.63, 220.22 and 368.50.

As shown in Figure 5B, the probability density function of the total number of COVID-19 cases in Hong Kong is quite symmetric. The probability of the total number of COVID-19 cases in Hong Kong from February 15, 2020, to December 16, 2020, reached the maximum value at around day 160. The expected value and variance of the time for the COVID-19 outbreak in Hong Kong are 319.08, and $1.52e+04$, respectively. The sample value and variance of the COVID-19 outbreak in Hong Kong are $2.91e+03$, and $5.12e+06$, respectively.

As shown in Figure 5, the EWMA control chart was applied for monitoring the total number of COVID-19 cases in Hong Kong from February 15, 2020, to December 16, 2020. The results from the experiment, the first hitting time based on the delta method (method 1) was 0 for all values of the smoothing parameter (Figure 5C) whereas the

sample mean and variance method (method 2) (Figure 5D) detected first hitting times of 202, 290, and 298 for $\lambda=0.3$, 0.7 and 0.9, respectively. For smoothing parameter $\lambda=0.9$ and $d=2$, the sample mean and variance method more effectively detected the first hitting time (298 days) than for the other smoothing parameter values. Thus, the optimal first hitting time on the EWMA control chart for monitoring the COVID-19 outbreak in Hong Kong is approximately 298 days for $\lambda=0.9$ and $d=2$.

Here, we compare, discuss, and suggest applying the EWMA control chart and parameter estimation for monitoring symmetric patterns of the total number of COVID-19 cases for the sampled countries and region: Thailand, Singapore, Vietnam, and Hong Kong.

Table 1 reports the estimated parameters, EWMA control chart, and the COVID-19 alert levels for monitoring COVID-19 outbreaks to

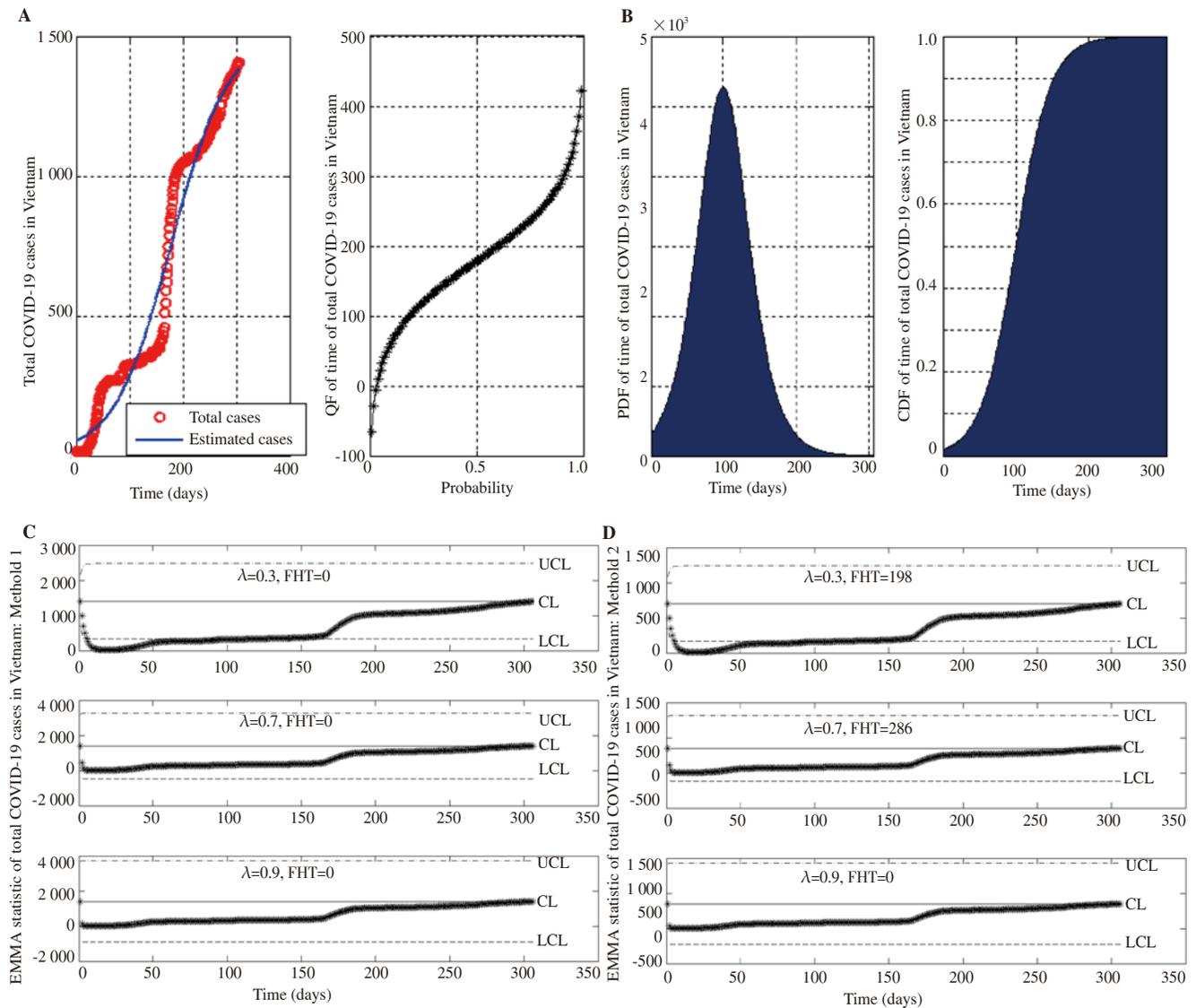


Figure 4. (A) Estimated total number of COVID-19 cases and its quantile function in Vietnam and (B) The probability density function and cumulative density function of the time for the total number of COVID-19 cases in Vietnam. The EWMA control chart for monitoring a COVID-19 outbreak in Vietnam: The expected value and variance estimated *via* (C) the delta method and (D) the sample mean and variance method. FHT: First hitting time; λ : Smoothing parameter. UCL: Upper control limit; CL: Center line; LCL: Lower control limit.

illustrate the levels of severity in each country. The delta model for estimating the expected value and the variance of the total number of COVID-19 cases could not detect the first hitting time on the EWMA control chart. However, the mean and variance model with the optimal smoothing parameter for the EWMA control chart of the total number of COVID-19 cases could detect the following first hitting times: 280 days with $\lambda=0.3$ for the COVID-19 outbreak in Thailand, 208 with $\lambda=0.3$ for the COVID-19 outbreak in Singapore, 286 days with $\lambda=0.7$ for the COVID-19 outbreak in Vietnam, and 298 days with $\lambda=0.9$ for the COVID-19 outbreak in Hong Kong. Furthermore, when using smoothing parameter $\lambda=0.3$ with the EWMA control chart for monitoring the COVID-19 outbreaks, the maximum first hitting times were 280, 208, 202, and 198 days in Thailand, Singapore, Hong Kong, and Vietnam, respectively. These

results imply that Thailand can more effectively hold the COVID-19 outbreak time before breaking the upper control limit of the EWMA control chart and thus better protect its population from COVID-19 outbreaks compared to the other countries or region. Moreover, the total number of COVID-19 cases escalate rapidly in COVID-19 alert level 4, thus signifying that policies for controlling a COVID-19 outbreak, such as lockdown, social distancing, *etc.*, should be strictly encouraged. COVID-19 total cases are exponentially increasing for alert level 3, which suggests that the policies mentioned for level 4 should be strictly maintained to suppress the total number of COVID-19 cases. For alert level 2, the COVID-19 total cases are starting to gradually decrease after having passed the inflection point at $Q_{0.05}$. However, applying the policies mentioned previously should still be strictly applied to suppress the total number of COVID-19

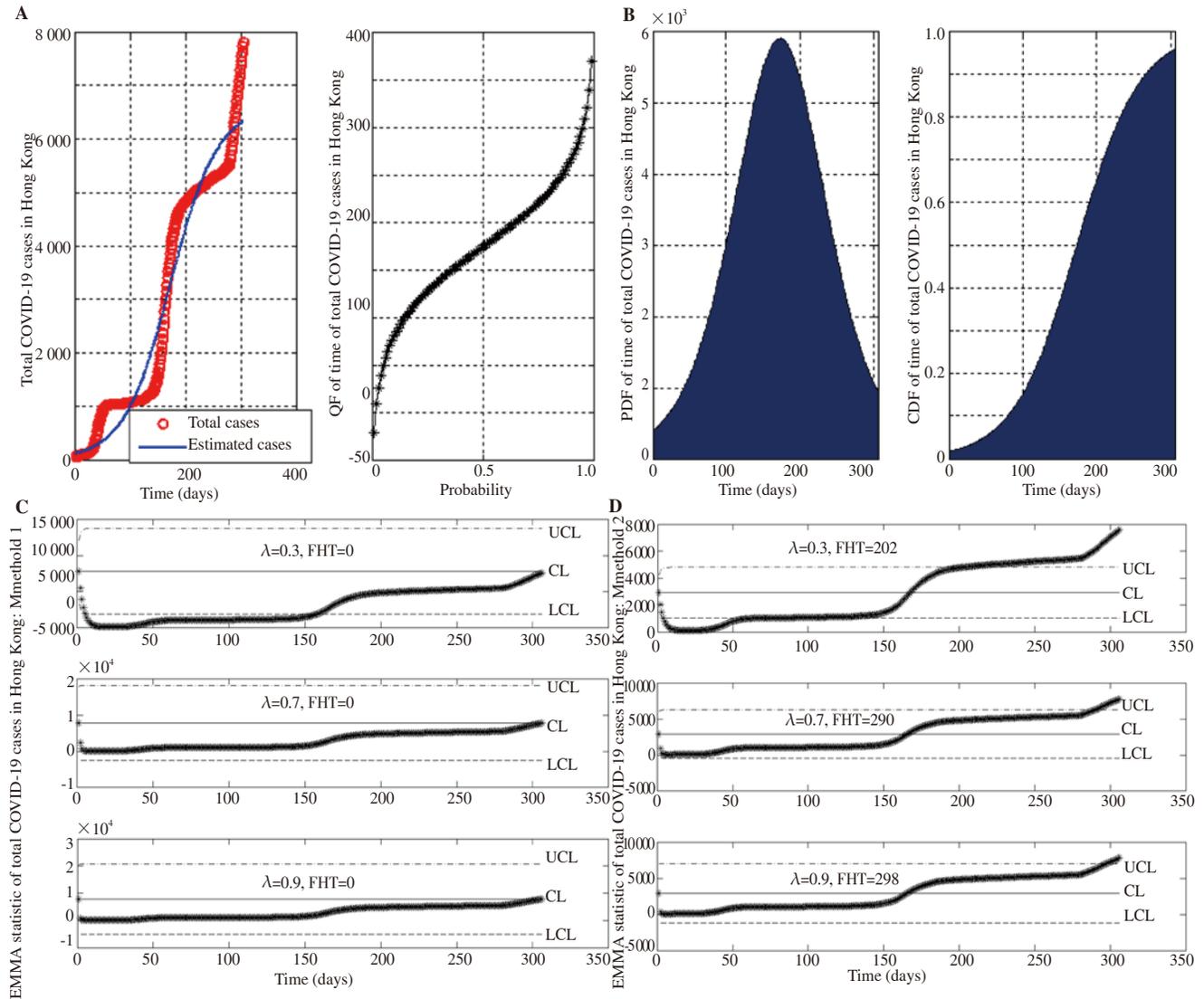


Figure 5. (A) Estimated total number of COVID-19 cases and its quantile function in Hong Kong and (B) the probability density function (PDF) and cumulative density function (CDF) of the time for the total number of COVID-19 cases in Hong Kong. The EWMA control chart for monitoring a COVID-19 outbreak in Hong Kong: The expected value and variance estimated *via* (C) the delta method and (D) the sample mean and variance method. FHT: First hitting time; λ : Smoothing parameter. UCL: Upper control limit; CL: Center line; LCL: Lower control limit.

Table 1. Comparison of the COVID-19 outbreaks in sampled countries and region.

Country	Logistic growth curve ($C/r/k/R^2/RMSPE$)	COVID-19 alert levels ($Q_{0.01}/Q_{0.25}/Q_{0.50}/Q_{0.75}/Q_{0.99}$)	EWMA control chart (λ/FHT)	
			Delta method	Mean and variance method
Thailand	3 462.31/0.09/108.59/0.94/0.71	0.96/37.44/48.91/60.36/96.84	0.3/0	0.3/280
			0.7/0	0.7/0
			0.9/0	0.9/0
Singapore	57 314.95/0.04/68.59/0.99/0.67	-8.71/74.25/100.32/126.39/209.35	0.3/0	0.3/208
			0.7/0	0.7/0
			0.9/0	0.9/0
Vietnam	1 509.36/0.02/28.43/0.96/0.68	-66.42/119.69/178.16/236.63/422.74	0.3/0	0.3/198
			0.7/0	0.7/286
			0.9/0	0.9/0
Hong Kong	6618.04/0.02/59.99/0.95/0.13	-21.24/127.04/173.63/220.22/368.50	0.3/0	0.3/202
			0.7/0	0.7/290
			0.9/0	0.9/298

cases. For alert level 1, the total number of COVID-19 cases is low or almost nonexistent, and so considerations in adjusting public health and social measures are recommended.

4. Discussion

The scope of this research was the study of probabilistic methods for the symmetric pattern of the total number of COVID-19 cases. The findings of this research are restricted to investigating the PDF of the symmetric pattern of the total number of COVID-19 cases. The samples are from only four countries or region: Thailand, Singapore, Vietnam, and Hong Kong. The outbreak time range was restricted to the period of February 15, 2020, to December 16, 2020. The approaches outlined in this research could be applied to symmetric patterns of the total number of COVID-19 cases in other countries. Moreover, we only used the EWMA control chart for monitoring the COVID-19 outbreaks. The application of the approach in this research on COVID-19 outbreaks in other countries has been left for the future. Asymmetric patterns for the total number of COVID-19 cases could be forecasted and estimated by using the PDF. Furthermore, other control charts such as the modified EWMA and cumulative sum, *etc.*, for monitoring COVID-19 outbreaks could be applied to classify the COVID-19 alert levels.

Probabilistic methods such as the PDF, quantile functions, and optimal first hitting times were used to set COVID-19 alert levels for the total number of COVID-19 cases with a symmetric pattern. The EWMA control chart was adopted to monitor and set the alert levels for COVID-19 outbreaks. Symmetric patterns of the total number of COVID-19 cases in Thailand, Singapore, Vietnam, and Hong Kong, were investigated. The sample mean and variance method and the delta method were used for estimating the mean and variance to detect the first hitting time on the EWMA control chart, of which the former method detected the first hitting time better than the latter. Four COVID-19 alert levels were defined for monitoring COVID-19 outbreaks based on quantile functions: $(Q_{0.01}, Q_{0.25})$ for level 4, $(Q_{0.25}, Q_{0.50})$ for level 3, $(Q_{0.50}, Q_{0.75})$ for level 2, and $(Q_{0.75}, Q_{0.99})$ for level 1. For example, the quantile functions for Thailand were $Q_{0.01}=0.96$, $Q_{0.25}=37.44$, $Q_{0.50}=48.91$, $Q_{0.75}=60.36$, and $Q_{0.99}=96.84$. However, for this research, the delta method which was used for estimating the mean and variance to compare the sample mean and variance method. The sample countries for this research: Thailand, Singapore, Vietnam, and Hong Kong, the delta method could not detect the first hitting time on the EWMA control chart but the delta method could detect the first hitting time on the EWMA control chart for the other countries.

The COVID-19 alert levels identified by using the probabilistic methods mentioned earlier are to help the authorities to enact

policies that monitor, control, and protect the population from a COVID-19 outbreak. For example, a strict lockdown policy should be employed for a COVID-19 alert level 4 situation to flatten the total number of COVID-19 cases following a symmetric logistic growth curve while relaxation of the policy should be applied for a COVID-19 alert level 1 situation. Moreover, a new COVID-19 wave can be identified by the EWMA control chart when the monitoring shows that the infection rate is out-of-control, *i.e.*, exceeds the upper control limit.

Conflict of interest statement

We declare that we have no conflict of interest.

Funding

The research was funding by King Mongkut's University of Technology North Bangkok Contract no. KMUTNB-61-KNOW-014.

Authors' contributions

All authors substantially contributed to drafting, revising the article as well as the final version of the manuscript. R.S. contributed to the conception and design of the study and the theoretical proof. Y.A. provided data and performed the data analysis. Both R.S. and Y.A. authors contributed to analyze, interpret, and summarize the results of the study.

References

- [1] Huynh G, Nguyen TN, Tran VK, Vo KN, Vo VT, Pham LA. Knowledge and attitude toward COVID-19 among healthcare workers at District 2 Hospital, Ho Chi Minh City. *Asian Pac J Trop Med* 2020; **13**: 260-265.
- [2] Chakraborty C, Sharma AR, Bhattacharya M, Sharma G, Lee SS. The 2019 novel coronavirus disease (COVID-19) pandemic: A zoonotic prospective. *Asian Pac J Trop Med* 2020; **13**: 242-246.
- [3] Intawong K, Olson D, Chariyalertsak S. Application technology to fight the COVID-19 pandemic: Lessons learned in Thailand. *Biochem Biophys Res Commun* 2021; **534**: 830-836.
- [4] Khan H, Kushwah KK, Singh S, Urkude H, Maurya MR, Sadasivuni KK. Smart technologies driven approaches to tackle COVID-19 pandemic: A review. *3 Biotech* 2021; **11**: 1-22.
- [5] Amtul W, Jana S. *Successful role of smart technology to combat COVID-*

19. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). 2020; 772-777. doi: 10.1109/I-SMAC49090.2020.
- [6] Areepong Y, Sunthornwat R. Predictive models for cumulative confirmed COVID-19 cases by day in Southeast Asia. *Comput Model Eng Sci* 2020; **125**(3): 927-942.
- [7] Sunthornwat R, Areepong Y. Predictive model for the total daily COVID-19 cases with herd immunity policy. *Chiang Mai Univ J Nat Sci* 2021; **20**(1): e2021016.
- [8] Zou Y, Pan S, Zhao P, Han L, Wang X, Hemerik L, et al. Outbreak analysis with a logistic growth model shows COVID-19 suppression dynamics in China. *PLoS One* 2020; **15**(6): e0235247.
- [9] New Zealand Government. *COVID-19 alert system*. 2021. [Online]. Available from: <https://covid19.govt.nz/alert-system>. [Accessed on 7th April 2021].
- [10] Alaskan Department of Health and Social Services. *Alaska COVID-19 alert levels*. 2021. [Online]. Available from: <http://dhss.alaska.gov/dph/Epi/id/Pages/COVID-19/alertlevels.aspx#current>. [Accessed on 7th April 2021].
- [11] United Kingdom Government. *Guidance UK COVID-19 alert level methodology: An overview*. 2021. [Online]. Available from: <https://www.gov.uk/government/publications/uk-covid-19-alert-level-methodology-an-overview/uk-covid-19-alert-level-methodology-an-overview>. [Accessed on 7 April 2021].
- [12] Singh BP, Madhusudan JV, Tiwari AK, Singh S, Das UD. Evaluation of EWMA control charts for monitoring spread of transformed observations of COVID-19 in India. *Asian J Res Infect Dis* 2020; **5**: 25-36.
- [13] Mahmood Y, Ishtiaq S, Khoo MBC, Teh SY, Khan H. Monitoring of three-phase variations in the mortality of COVID-19 pandemic using control charts: Where does Pakistan stand? *Int J Qual Health Care* 2021; **33**(2): mzab062. doi: 10.1093/intqhc/mzab062.
- [14] Mbaye M, Sarr N, Ngom B. Construction of control charts for monitoring various parameters related to the management of the COVID-19 pandemic. *J Biosci Med* 2021; **9**: 9-19.
- [15] Worldometer. *COVID-19 coronavirus pandemic*. 2021. [Online]. Available from: <https://www.worldometers.info/coronavirus/#countries>. [Accessed on 7 April 2021].
- [16] Haym B, Seon MH. *Probability models in engineering and science*. Boca Raton: Taylor & Francis; 2005.
- [17] Sunthornwat R, Areepong Y. Average run length on CUSUM control chart for seasonal and non-seasonal moving average processes with exogenous variables. *Symmetry* 2020; **12**(1): 173. doi: 10.3390/sym12010173.
- [18] Supharakonsakun Y, Areepong Y, Sukparungsee S. The exact solution of the average run length on a modified EWMA control chart for the first-order moving-average process. *ScienceAsia* 2020; **46**(1): 109-118.
- [19] Phanthuna P, Areepong Y, Sukparungsee S. Exact run length evaluation on a two-sided modified exponentially weighted moving average chart for monitoring process mean. *Comput Model Eng Sci* 2021; **127**(1): 23-41.
- [20] Phanyaem S. Estimating the average run length of CUSUM control chart for seasonal autoregressive integrated moving average of order (P,D,Q) L Model. *J App Sci* 2020; **19**(2): 52-60.
- [21] Chananet C, Areepong Y, Sukparungsee S. An approximate formula for ARL in moving average chart with ZINB Data. *Thai Stat* 2015; **13**(2): 209-222.
- [22] Peerajit W, Areepong Y, Sukparungsee S. Explicit analytical solutions for ARL of CUSUM chart for a long-memory SARFIMA model. *Commun Stat Simul Comput* 2019; **48**(4): 1176-1190.