Journal of Current Science and Technology, May - August 2024 Vol. 14 No. 2, Article 29

ISSN 2630-0656 (Online)

Cite this article: Wongsathan, R., & Seedadan, I. (2024). Predictive analysis of COVID-19 epidemic in Thailand: evaluating control lockdown measures using LSTM networks. *Journal of Current Science and Technology*, *14*(2), Article 29. https://doi.org/10.59796/jcst.V14N2.2024.29

Journal of Current Science and Technology

Journal homepage: https://jcst.rsu.ac.th

Predictive Analysis of COVID-19 Epidemic in Thailand: Evaluating Control Lockdown Measures using LSTM Networks

Rati Wongsathan* and Isaravuth Seedadan

Department of Electrical Engineering, Faculty of Engineering and Technology, North-Chiang Mai University, Chiang Mai 50230, Thailand

*Corresponding author; E-mail: rati@northcm.ac.th

Received 27 November, 2023; Revised 14 December, 2023; Accepted 26 January, 2024
Published online 2 May, 2024

Abstract

This study addresses the critical objective of evaluating the effectiveness of non-pharmaceutical lockdown measures implemented during COVID-19 outbreaks in Thailand. Assessing the outcome of these measures provides valuable insight that can inform and guide response to future outbreaks. Utilizing a closed-loop forecasting model built on Long Short-Term Memory (LSTM) networks, the research focuses on achieving precise daily forecasts of COVID-19 cases. The methodology involves optimizing hyperparameters through grid-search and incorporating training data from other countries that implemented similar measures. The LSTM, configured with an optimal number of hidden processing units, utilizes past lagged data of daily infected cases as predictors to generate multi-step-ahead predicted values, which are subsequently used as predictors in a recursive approach. As a result, the predicted cases closely align with measured data, facilitating the estimation of the effective reproduction number (R_{eff}) to assess the performance of lockdown measures. The effectiveness of the lockdown measures is quantified at different time intervals: 51%, 41%, and 23% one day after implementation, increasing to 84%, 98%, and 34% after one week, and reaching 96%, 99%, and 73% at the endpoint of the first, second, and fourth waves of infection, respectively. Throughout these waves, the final R_{eff} remains below 1, indicating ongoing but controllable COVID-19, demonstrating the efficacy of the implemented lockdown measures. It is noted that these results are based on specific LSTM model, as the effectiveness of lockdown measures may vary with alternative modeling approaches. Therefore, the findings should be interpreted in the context of this LSTM-framework.

Keywords**:** *COVID-19 forecasting; Long Short-Term Memory* (*LSTM*)*; lockdown measures; hyperparameters optimization; effective reproduction number* (*R_{eff}*)*.*

1. Introduction

Since the COVID-19 outbreak originating in Wuhan, China (2019), countries worldwide have implemented a range of non-pharmaceutical interventions, including control measures such as lockdowns, face masking, self-quarantine, social distancing, and others. Lockdowns were initially found to be highly effective, reducing infection rates by 28% across 133 countries (Ge et al., 2022; Russell et al., 2021). Subsequently, face masking became a dominant measure during the next wave, with an average effectiveness of 30%. The rollout of vaccination strategies, albeit delayed, showed potential in curbing the epidemic and achieving herd immunity. However, individual countries adopted unique measures influenced by diverse factors, resulting in varying approaches across countries and epidemic waves. As a result, previous studies may not fully capture the effectiveness of these measures on a country-by-country basis. As of today (July 10, 2023), COVID-19 is on the decline, but it is still part of a seasonal cycle. Notably, there is no widespread testing for

infection; only confirmed cases are being reported (WHO).

In this study, Thailand, experiencing five COVID-19 waves, Figure 1 is examined in the context of lockdown implementations to evaluate performance efficiency. National lockdowns occurred twice: in the first wave (March 25, 2020) and the fourth wave (July 12, 2021). The second wave featured localized lockdowns in areas with a significant migrant worker population. In the first wave (January-May 2020), following a superspreading event in Bangkok, lockdown measures led to a significant reduction in COVID-19 transmission, Figure 1(a). According to the Susceptible-Exposed-Infected-Quarantined-Recovered (SEIQR) model (Patanarapeelert et al., 2022) there was a projected 91% decrease in the effective reproduction number (R_{eff}), a key epidemiological metric. In the second wave (December 2020-February 2021), Figure 1(b), localized lockdowns lowered R_{eff} by around 80%. Despite these measures, this wave characterized multiple peaks, deviating from the typical infection curve, posing prediction challenges. The SEIQR model failed to capture the evolving outbreak trend, consistently underestimating values.

In the fourth wave (June-December 2021), Figure 1(d), characterized by the delta variant, Thailand experienced a rapid surge in COVID-19 cases, surpassing one million by August 2021, despite a high vaccination rate. Managing over a hundred outbreak clusters posed a significant challenge. In response, a soft lockdown was reinstated after a six-month period, alongside ongoing vaccination efforts. Utilizing the Susceptible-Infected-Exposed-Recovered (SIER) model with a mobility index, lockdown impact, and vaccination rate, the lockdown reduced viral spread by approximately 15.5-18.2% (Polwiang, 2023). This model suggests that a two-month lockdown can decrease the effective reproduction number by over 60%, or one-fifth of the peak size without lockdown. Additionally, lockdown measures can shift the infection peak, aiding in medical supply allocation and future surge timing estimation. This highlights the effectiveness of lockdowns in slowing the virus spread.

Although lockdown measures help prevent new cases and mitigate the spread of the pandemic, they have a significant impact on the economy. These measures disrupt the supply chain, significantly restrict goods production from factories, and lead to delays in transactions (Nicola et al., 2020). Additionally, they affect consumption demand and result in service reductions, causing immediate liquidity shortages for firms and households. Household debt in Thailand is reported as the second-highest in East Asia by UNICEF-Thailand (2020). In 2020, the poverty rate increased to 9%, particularly among rural people, informal private employees, state employees, and own-account workers. Moreover, the deteriorating economy has an extreme impact on social issues, including increased unemployment rates, higher termination of employment, rising domestic violence, an increase in mental health issues, and reduced access to education. Contrastingly, with the restriction of movements, air pollution has improved, reducing PM-2.5 (particulate matter) levels by 20%. However, it poses challenges, particularly regarding mental health, due to a decrease in physical activity.

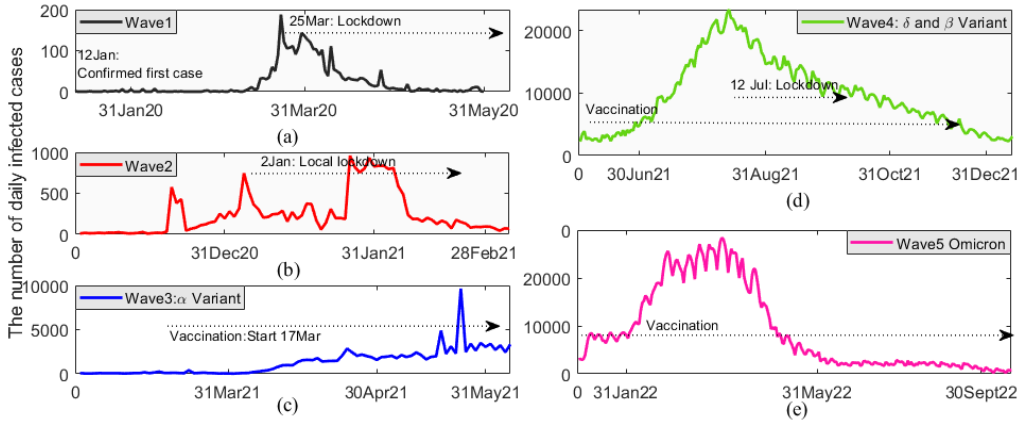


Figure 1 Daily infected COVID-19 cases against two main control measures; lockdowns and vaccinations, across the first to fifth waves of infection in Thailand

In Thailand, research on applying the LSTM model for predicting COVID-19 is quite limited. The research conducted by Kompunt et al. (2023) applied LSTM assisted with multilayered perceptron (MLP) techniques based on geometric information system (GIS) data to predict cumulative cases in the first three waves of infection all provinces in Thailand. The results showed superior accuracy (99.7%) compared to other state-of-the-art prediction models in different countries, such as support vector machine (SVM) (95%), radial basis function (RBF) (81.6%), nonlinear autoregressive with exogenous inputs (NARX), and decision tree (both 81.6%), and Bayesian (95.2%). However, differences in datasets, assumptions, and model complexity levels across study areas should be noted, potentially impacting the fairness of this comparison. Moreover, Vorathamthongdee, & Chongstitvatana (2023) suggest that incorporating data from countries paired with Thailand into the LSTM model improves the accuracy of predictions. In the study by Winalai et al. (2022), the delayed lag for predictors is set to 7, in accordance with the COVID-19 incubation period. Simultaneously, the number of steps for multi-step-ahead prediction is optimally determined as 4 for the LSTM (7,4) model used in the daily case prediction for Thailand.

Given the impact of lockdown measures on the socio-economy and the limited research on their effectiveness in Thailand, there is a need for performance evaluation. Current assessments primarily rely on predictive models in epidemiology, including mathematical model such as exponential and non-linear growth models, statistical model such as an autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA), and regression models (Wongsathan, 2021a; Sengsri, & khunratchasana, 2023). Ordinary differential equation-based compartmental models, such as SIR, SIER, and their variants, have also been employed (Mahikul et al., 2021; Teekasap et al., 2022). These models are utilized in long-term COVID-19 forecasting and require assumptions and parameter estimations that are likely burdened with errors due to a lack of adaptability. They are suitable for predicting smooth curves, like cumulative infection cases that follow the standard S-shape curve. However, when interventions occur, this curve may deform. Additionally, capturing daily infection cases characterized by high fluctuations due to the nature of infection and control interventions may render these models less effective. Specifically, for compartmental models, the results consistently present underestimated values on average. In contrast, short-term forecasts based on long short-term memory (LSTM) recurrent neural networks, which have memory units for past data, can adapt themselves through the flow of information units, called gates, making them suitable for time-sequential COVID-19 data. Due to the complex and nonlinear nature of disease transmission under control measures, accurate adaptive models are essential. To address this gap, this study utilizes LSTM neural networks, renowned for their ability to capture long-term dependencies in infectious time-series data (Ao, & Fayek, 2023; Absar et al., 2022), and in other fields, for precise daily infection predictions. The details of the LSTM-based prediction model for daily infected cases are provided in Section 3.1.

Using raw fluctuation data poses a prediction challenge to LSTM, as opposed to cumulative case predictions (Sunthornwat, & Areepong, 2021; Wongsathan, & Puangmanee, 2023), which smooth out these fluctuations providing predictions based on averages may not accurately reflect the real situation. To enhance LSTM's learning, hyperparameters are optimized through grid search to improve LSTM training in terms of reducing the overfitting problem. Furthermore, training data from countries that implemented similar lockdown measures before Thailand, such as China, Hong Kong, and Vietnam, are incorporated to improve prediction accuracy. With partial real data from the first, second, and fourth waves of ongoing infection in Thailand, LSTM adopts a closed-loop forecasting approach, utilizing prior short-term predictions as input for subsequent time intervals (long-term predictions). Through this approach, error propagation from previous forecasts is mitigated by selecting an optimized time lags for autoregressive terms of predictors and the response length for predictions. Based on the predictions, the effectiveness of lockdown measures of each wave of infection is analyzed, assessed, and quantified by estimating effective reproduction number (R_{eff}), a main epidemic criterion. This is done before and after implementation through a mathematical demographic approach, as detailed in Section 3.2.

This study significantly contributes by utilizing a forecasting model, LSTM networks, to provide precise daily COVID-19 forecasts. The optimized methodology and international data integration demonstrate the effectiveness of non-pharmaceutical lockdown measures, offering valuable insights for future outbreak responses.

2. Objectives

This paper aims to evaluate the effectiveness of non-pharmaceutical lockdown measures during COVID-19 outbreaks in Thailand, informing future response strategies. It develops a closed-loop forecasting model using LSTM networks to precisely forecast daily COVID-19 cases, incorporating training data from multiple countries. By assessing lockdown measures' performance and quantifying their effectiveness at different intervals, the study demonstrates the efficacy of the implemented measures while acknowledging potential variability in results based on alternative modeling approaches.

3. Materials and Methods

3.1 LSTM-based Daily COVID-19 Cases Prediction Model

The LSTM-based COVID-19 cases prediction model, Figure 2, comprises three layers, including input layer, LSTM layer, and output layer. The input layer presents an input window of the sequential autocorrelated data of daily infected COVID-19 cases of prior *lag* days for autoregressive terms, $\mathbf{Covid}_t = [Covid_{t-lag+1}, \ldots, Covid_{t-1}, Covid_t]^T \in R^{lag\times 1}$. Within the LSTM layer, LSTM units are employed to capture and remember the complex relationships among these lagged terms. The LSTM generates forecasts to responses in subsequent *res* days, $\mathbf{Covid}^{predict} = [Covid_{t+1}, Covid_{t+2}, \ldots, Covid_{t+res}] \in R^{res\times 1}$, within the output layer.

In the first stage called forget gate (grey box in Figure 2), previous short-term memories ($\mathbf{H}_{t-1}$) and current inputs ($\mathbf{Covid}_t$) are weighted through $\mathbf{W}_F$, combined with biases ($\mathbf{b}_F$), and subjected to a sigmoid function, $\sigma(x) = (1-e^{-x})^{-1} \in [0, 1]$. The resulting output ($\mathbf{F}_t$), serving as the percentage to remember/forget information, is then multiplied by the previous long-term memories ($\mathbf{C}_{t-1}$). This stage interacts between input, short, and long-term memories. The second stage called input gate (violet box in Figure 2) involves weighted $\mathbf{Covid}_t$ and $\mathbf{H}_{t-1}$ through $\mathbf{W}_C$ combined with biases ($\mathbf{b}_C$), passing through a hyperbolic tangent function, $tanh(x) = 2\sigma(2x)-1 \in [-1, 1]$, to yield potential long-term memories in the intermediate cell state ($\tilde{C}_t$). Simultaneously, the other part of this stage, the weighted $\mathbf{H}_{t-1}$ and $\mathbf{Covid}_t$ through $\mathbf{W}_I$ combined with biases ($\mathbf{b}_I$) are passed through σ function to determine the percentage of these potential memories for updating the preceding ones. In the final stage called output gate (green box, Figure 2), the new long-term memories pass through *tanh* function, yielding potential short-term memory. This value is then multiplied by the output of a sigmoid function, which incorporates the weighted $\mathbf{H}_{t-1}$ and $\mathbf{Covid}_t$ through $\mathbf{W}_O$ along with biases ($\mathbf{b}_O$). The resulting new short-term memory represents the LSTM network's output which is passed through the hidden layer to generate the prediction outputs.

The gate (or dense neural network) in LSTM unit of the forget gate, input gate, and output gate can be mathematically expressed, respectively, as

$$F_t=\sigma(W_F^{Covid}\times Covid_t+W_F^H\times H_{t-1}+b_F), \quad (1)$$

$$I_t=\sigma(W_I^{Covid}\times Covid_t+W_I^H\times H_{t-1}+b_I), \quad (2)$$

and

$$O_t=\sigma(W_O^{Covid}\times Covid_t+W_O^H\times H_{t-1}+b_O). \quad (3)$$

As a result, the hidden state output ($\mathbf{H}_t$), or the prediction output at time *t* derived from the sub-outputs of the three aforementioned gates, can be expressed as follows:

$$H_t=\tanh(C_t)\otimes O_t, \quad (4)$$

where the current memory cell and intermediate cell state $\in R^{hid\times 1}$ are respectively as:

$$C_t=F_t\otimes C_{t-1}+I_t\otimes \tilde{C}_t, \quad (5)$$

$$\tilde{C}_t=\tanh(W_C^{Cov}\times Covid_t+W_C^H\times H_{t-1}+b_C). \quad (6)$$

Here, $\mathbf{H}_{t-1} \in R^{hid\times 1}$ represents the prior hidden layers, $W_F^{Covid}, W_I^{Covid}, W_O^{Covid}$, and $W_C^{Covid} \in R^{hid\times lag}$ denote the forward weight matrices, W_F^H, W_I^H, W_O^H and $W_C^H \in R^{hid\times hid}$ denote the recurrent weight matrices of the forget, input, output, and candidate gates, respectively. The parameter *hid* represents the number of hidden LSTM units, determining the extent to which information is learned by the layer, and b_F, b_I, b_O, and $b_C \in R^{hid\times 1}$ signify bias vectors. The operator $\otimes$ denotes element-wise multiplication. The initial values are

set as follows: $\mathbf{C}_0 = \underline{\mathbf{0}}$ and $\mathbf{H}_0 = \underline{\mathbf{0}}$, or they can be initialized randomly.

In training LSTM, the most widely used method, an extended version of the familiar backpropagation algorithm, is backpropagation through time (BPTT) (Werbos, 1988). By using BPTT, unrolling in time, and treating it as a multilayer feedforward neural network, the parameters, $\Theta = \{\mathbf{W}^{covid}, \mathbf{W}^{H}, \mathbf{b}\}$ are updated through the dataset size N to minimize the cost function J, $\underset{\Theta}{argmin}J$, where

$$J=(1/N)\sum_{i=1}^{N} L_i(Covid^{actual},Covid^{predict}), \quad (7)$$

with the loss function,

$$L=(1/2)(Covid^{actual}-Covid^{predict})^2. \quad (8)$$

3.2 Effective Reproduction Number Under Lockdown Measure

The epidemic's evolution can be indexed by R_0, the basic reproduction number, representing the ratio of new infections generated by one infected person. When $R_0 > 1$ (unstable), it indicates an epidemic; when $R_0 < 1$ (stable), the virus may die out. R_0 varies with mathematical models and assumptions. For instance, in the SIR model by using the dominant eigenvalue of the next-generation matrix, $R_0 = \beta/(\gamma+\mu)$, while in the SEIR model, $R_0 = \beta\times\varepsilon/(\gamma+\mu)(\varepsilon+\mu)$, and in the exponential growth model, $R_0 = e^{\kappa\tau}$, where β represents the effective contact rate, γ is the recovery rate, ε is the progression rate from exposed to infected, μ is the natural death rate, κ represents the growth rate, and τ denotes the duration time. However, utilizing R_0 assumes a fully susceptible population without immunization, implying the absence of immunization.

In this study, R_0 estimation, often challenging with indirect methods, is achieved directly. predicted daily COVID-19 cases ($Covid^{predict}$) generated by the LSTM model are used, with just one epidemic parameter (γ), through a mathematical demographic approach. R_0, with a lag of t_0 days between the infection date and reported cases, is defined as the ratio of the number of new infected cases on day t to the proportion of active infected cases during the period $[t_{init}, t_{final}]$, expressed as (Rosero-Bixby, & Rosero-Bixby, 2022),

$$R_0(t\text{-}t_0)=\frac{Covid^{predict}(t)}{\sum_{a=t_{init}}^{a=t_{final}}\left(\frac{\gamma e^{-\gamma a}}{e^{-\gamma t_{init}}-e^{-\gamma t_{final}}}\right)\times Covid^{predict}(t\text{-}a)}, \quad (9)$$

where the first term in the denominator represents the survival distribution function for infectious individuals remaining contagious (t–a) days after infection. $R_0(t)$ is simulated with $\gamma = 0.1$, following WHO (2020), subject to possible future changes.

To estimate infection dynamic in a partially susceptible population while accounting for control measures like lockdown measures starting at time t_c, the effective reproduction number as a function of time, $R_{eff}(t)$. This function takes into account the impact of these measures on R_0 and can be expressed as,

$$R_{eff}(t)=\begin{cases} R_0(t) & 0\le t<t_c \\ R_0(t)(1-l(t)) & t\ge t_c \end{cases} \quad (10)$$

where $l(t)$ represents the effectiveness of the lockdown measures.

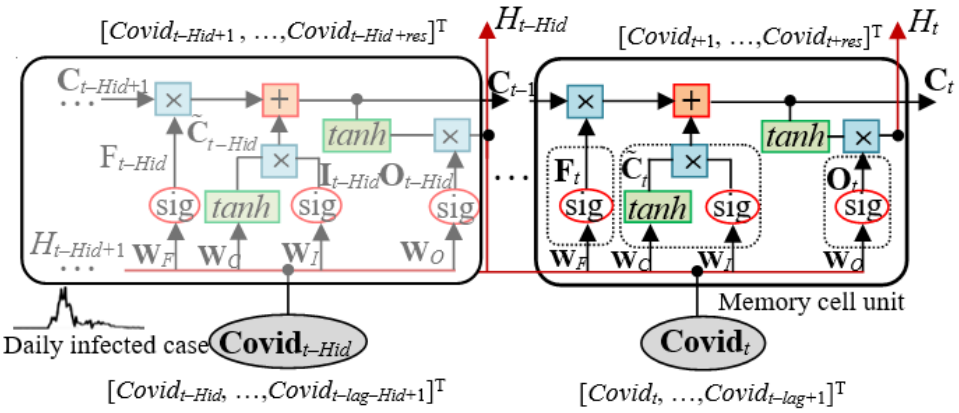


Figure 2 LSTM-based daily COVID-19 cases prediction model with a single hidden layer, incorporating *lag* autoregressive terms and *Res* projection terms

3.3 COVID-19 Dataset

In the pre-training phase, daily COVID-19 case data from Thailand and other Asian countries, which had experienced similar outbreaks and implemented lockdown measures before Europe and America, were combined. The acquisition of secondary data on active COVID-19 cases in Thailand involved retrieving information from the Ministry of Public Health, Thailand (https:// public.tableau.com/views/SATCOVID ashboard/1-dash-tiles?: showVizHome=no), while data from other countries were sourced from the Worldometer website (https://www.worldometers.info/coronavirus/) and the World Health Organization (WHO). The countries selected for data training, along with the initial data from each wave of infection in Thailand that applied lockdown measures before Thailand, were chosen based on information obtained from the Oxford COVID-19 Government Response Tracker (OxCGRT) (https://extranet.who.int/countryplanning cycles/reportsportal/oxford-covid-19-government-response-tracker-oxcgrt). This dataset included COVID-19 data with the initial lockdown periods from China (starting on January 23, 2020, with 60 data points), Hong Kong (on February 7, 2020, with 50 data points), and Vietnam (on February 12, 2020, with 57 data points).

Meanwhile, the number of infected cases used for LSTM model training and testing during the first, second, and fourth waves in Thailand were 163 (from January 1, 2020, to May 31, 2020), 82 (from December 1, 2020, to February 20, 2021), and 214 (from June 1, 2021, to December 31, 2021), respectively. Therefore, a total of 626 sample data points (459 from Thailand's cases and 167 from outside countries) were gathered for use in this work, as shown in Table 1. This raw daily data does not need smoothing, unlike other research, as we aim to test the robustness of the LSTM model to capture the relationship between data under high fluctuations.

For splitting training and testing data, the training data consisted of 167 data from other countries, with additional training data from the first wave (30 data) out of 163 data, resulting in a ratio of 60%:40%, for the first wave of infection. Similarly, for the second wave of infection, the training data included 330 data, with some from the second wave (30 data) out of 82 data, creating a ratio of 80%:20%. Additionally, for the fourth wave of infection, the training data comprised 412 data from other countries, the first and second waves of infection, with some from the fourth wave (60 data) out of 214 data, resulting in a ratio of 75%:25%.

3.4 Data Preprocessing

The dataset utilized in this study is carefully selected to include data for estimating the reproduction numbers observed in Thailand. During the data cleaning process, it is determined that there are few missing values from the Thailand official website database. Since official records are checked daily, and as a result, few missing values have been identified. They are replaced by interpolation using simple moving average (SMA) values. To fill the missing values at time t, a 3-period SMA is used as $SMA_t = (Covid_{t-1} + Covid_{t-2} + Covid_{t-3})/3$. Regarding data from countries outside Thailand retrieved from the Worldometer website and the World Health Organization (WHO), any missing values found in these countries have already been replaced through interpolated values.

Since LSTMs are sensitive to the scale of input data, the maximum and minimum values outside the training dataset are identified to ensure data consistency and support the predicted data. During the scaling process, [*Min*, *Max*] is set as [0, 1×10^3], [0, 2×10^3], and [0, 3×10^4] for the 1st, 2nd, and 4th waves of infection, respectively. Subsequently, a normalization process scales the dataset into the range [-1, 1], by a linear formula: $2(Covid - Min)/(Max - Min) - 1$. This helps mitigate differences in sizes among different data sources.

For forecasting COVID-19 cases, the predictors comprise past data with the time lag (*lag*), while the response parameter (*res*), representing the step-ahead forecast, is prepared for use in training, testing, and forecasting the LSTM model. The preparation of sequential data for training, testing, and forecasting phases is illustrated in Figure 3. It can be divided into two cases, including *lag* > *res* (Figure 3(a)) with examples using 16 training samples to forecast 9 steps ahead, and *lag* < *res* (Figure 3(b)) with examples using 20 training samples to forecast 15 steps ahead.

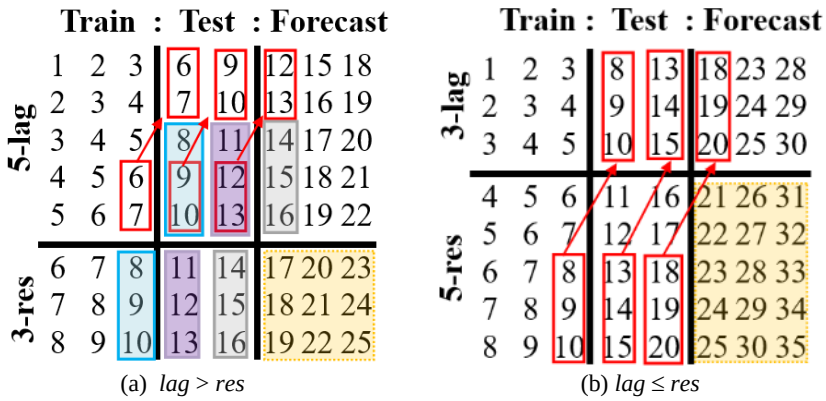
3.5 Hyperparameter selection

For forecasting COVID-19 cases, selecting *lag* and *res* are crucial. This is performed through grid search. Given that the incubation period from exposure to symptom onset spans about 2-14 days (averaging 5.1 days in early 2020) and varies based on individuals and virus variants (shorter in Omicron, approximately 3 days compared to the

other variants), exploring different values for *lag* predictors of 2, 3, 5, 7, 10, and 14 days with *res* (2, 3, 5, and 7 days) is essential. Consequently, the grid search spans the domain *lag* × *res* ∈ [2,3,5,7,10,14]×[2,3,5,7]. So, *lag* and *res* parameters are selected after other parameters are set. For hyperparameter selecting, the number of hidden units (*hid*) and the number of iterations (*Epoch*) are selected through the grid search of *hid*×*Epoch* ∈ [1, 500]×[50, 2000], while using a pilot *lag* of 5 and *res* of 1 to reduce the dimension of the search. To enhance training, ADAM optimizer, with its adaptive learning rates, is suitable for LSTM in addressing complex, nonlinear problems like epidemics. A small initial learning rate (0.001) is employed. Moreover, the dropout rate in LSTM is crucial for preventing overfitting. It represents the fraction of input units to randomly drop during training, typically set between [0.2, 0.5].

Optimizing hyperparameters for LSTM models involves experimenting with momentum and batch size. The momentum value significantly influences convergence, with a small momentum delicately navigating noisy gradients but potentially slowing convergence. Conversely, a high momentum accelerates convergence, but excessive values may lead to overshooting. Experimentation with different momentum values is crucial for striking a balance between convergence speed and stability. Simultaneously, batch size in LSTM-based predictions is pivotal for performance. A larger batch size may result in faster training convergence but could require more memory, while a smaller batch size may introduce more noise but potentially converge to better solutions. The choice of batch size requires experimentation to find the right balance for the specific dataset, considering the trade-off between training speed and accuracy. The test optimizes momentum values (0.1, 0.5, and 0.9) and batch powers of 2 (32, 64, 128) for the LSTM model. However, commonly used momentum values for LSTM models are around 0.9 or 0.99, chosen based on faster convergence in scenarios with noisy gradients.

Two hyperparameters, the number of hidden units (*hid*) set at 50, 50, and 20, and the number of iterations (*Epoch*) at 200, 200, and 100 for the first, second, and fourth waves infection, respectively, were selected through a grid search within the search of *hid*×*Epoch*∈[1, 500]×[50, 2000], as depicted in Figure 4(a)-4(c) and Figure 4(d)-4(f). These selections were made based on the lowest RMSE criteria not only in the training phase but also in the testing phase, taking into account the cost of computation. The dropout rate was set at 20% after observing multiple experiments, and adaptive learning was improved through the use of the Adam optimizer with an initial learning rate of 0.001. This decision was based on the selection criteria, indicating that further incrementally increasing these parameters from small values did not lead to improved error performance. Furthermore, it was determined that using 5-*lag*, 5-*lag*, and 7-*lag* predictors corresponding to short-term forecasts of 2-*res*, 5-*res*, and 2-*res* for the first, second, and fourth waves of infection, respectively, provided optimal results for training and validation of the LSTM model, as illustrated in Figure 4(g)-4(i). The hyperparameter's selecting range and the best choices for the LSTM model used in the first, second, and fourth wave of infection in Thailand are summarized in Table 2.



(a) *lag* > *res* (b) *lag* ≤ *res*

Figure 3 Data configuration set up for training, testing, and forecasting using actual data, and generating forecast data for two scenarios: (a) 16:9 and (b) 20:15

Table 1 Daily cases of COVID-19 data (*C*) in Thailand during the first, second, and fourth waves and data from other countries, including China, Hong Kong, and Vietnam, used in training and testing LSTM-based prediction model

Date	*C*	Date	*C*	Date	*C*	Date	*C*	Date	*C*	Date	*C*	Date	*C*	Date	*C*	Date	*C*	Date	*C*
1st Wave		7/3/20	2	29/4/20	9	5/1/21	527	11/6/21	2290	18/8/21	20515	25/10/21	8675	**China**		13/2/20	2	20/4/20	1
1/1/20	0	8/3/20	0	30/4/20	7	6/1/21	365	12/6/21	3277	19/8/21	20902	26/10/21	7706	23/1/20	93	14/2/20	0	21/4/20	3
2/1/20	0	9/3/20	0	1/5/20	6	7/1/21	305	13/6/21	2804	20/8/21	19851	27/10/21	8452	24/1/20	277	15/2/20	2	22/4/20	2
3/1/20	0	10/3/20	3	2/5/20	6	8/1/21	205	14/6/21	3355	21/8/21	20571	28/10/21	9658	25/1/20	483	16/2/20	1	23/4/20	1
4/1/20	0	11/3/20	6	3/5/20	3	9/1/21	212	15/6/21	3000	22/8/21	19014	29/10/21	8968	26/1/20	666	17/2/20	2	24/4/20	5
5/1/20	0	12/3/20	11	4/5/20	18	10/1/21	245	16/6/21	2331	23/8/21	17491	30/10/21	9224	27/1/20	802	18/2/20	0	25/4/20	0
6/1/20	0	13/3/20	5	5/5/20	1	11/1/21	249	17/6/21	3129	24/8/21	17165	31/10/21	8859	28/1/20	2632	19/2/20	2	26/4/20	1
7/1/20	0	14/3/20	7	6/5/20	1	12/1/21	287	18/6/21	3058	25/8/21	18417	1/11/21	8165	29/1/20	576	20/2/20	4	27/4/20	5
8/1/20	0	15/3/20	32	7/5/20	3	13/1/21	157	19/6/21	3667	26/8/21	18501	2/11/21	7574	30/1/20	2054	21/2/20	3	28/4/20	5
9/1/20	0	16/3/20	33	8/5/20	8	14/1/21	271	20/6/21	3682	27/8/21	18702	3/11/21	7679	31/1/20	1659	22/2/20	1	29/4/20	5
10/1/20	0	17/3/20	30	9/5/20	4	15/1/21	188	21/6/21	3175	28/8/21	17984	4/11/21	7982	1/2/20	2088	23/2/20	1	**Vietnam**	
11/1/20	0	18/3/20	35	10/5/20	5	16/1/21	230	22/6/21	4059	29/8/21	16536	5/11/21	8148	2/2/20	4737	24/2/20	3	12/2/20	1
12/1/20	1	19/3/20	60	11/5/20	6	17/1/21	374	23/6/21	3174	30/8/21	15972	6/11/21	8467	3/2/20	3086	25/2/20	9	13/2/20	0
13/1/20	0	20/3/20	50	13/5/20	1	18/1/21	369	24/6/21	4108	31/8/21	14666	7/11/21	7960	4/2/20	3989	26/2/20	11	14/2/20	1
14/1/20	0	21/3/20	89	14/5/20	7	19/1/21	171	25/6/21	3644	1/9/21	14802	8/11/21	7592	5/2/20	3729	27/2/20	1	15/2/20	0
15/1/20	0	22/3/20	188	15/5/20	0	20/1/21	59	26/6/21	4161	2/9/21	14956	9/11/21	6904	6/2/20	3144	28/2/20	3	16/2/20	0
16/1/20	0	23/3/20	122	16/5/20	3	21/1/21	142	27/6/21	3995	3/9/21	14653	10/11/21	6978	7/2/20	3522	29/2/20	3	17/2/20	0
17/1/20	1	24/3/20	106	17/5/20	3	22/1/21	309	28/6/21	5406	4/9/21	15942	11/11/21	7496	8/2/20	2703	1/3/20	0	18/2/20	0
18/1/20	0	25/3/20	107	18/5/20	2	23/1/21	198	29/6/21	4662	5/9/21	15452	12/11/21	7305	9/2/20	3012	12/5/20	1	19/2/20	0
19/1/20	0	26/3/20	111	19/5/20	1	24/1/21	198	30/6/21	4786	6/9/21	13988	13/11/21	7057	10/2/20	2516	2/3/20	3	20/2/20	0
20/1/20	0	27/3/20	91	20/5/20	3	25/1/21	187	1/7/21	5533	7/9/21	13821	14/11/21	7079	11/2/20	2021	3/3/20	2	21/2/20	0
21/1/20	0	28/3/20	109	21/5/20	0	26/1/21	959	2/7/21	6087	8/9/21	14176	15/11/21	6343	12/2/20	372	4/3/20	1	22/2/20	0
22/1/20	2	29/3/20	143	22/5/20	3	27/1/21	819	3/7/21	6230	9/9/21	16031	16/11/21	5947	13/2/20	5133	5/3/20	5	23/2/20	0
23/1/20	0	30/3/20	136	23/5/20	0	28/1/21	756	4/7/21	5916	10/9/21	14403	17/11/21	6524	14/2/20	6460	6/3/20	0	24/2/20	0
24/1/20	1	31/3/20	127	24/5/20	2	29/1/21	802	5/7/21	6166	11/9/21	15191	18/11/21	6901	15/2/20	2055	7/3/20	1	25/2/20	0
25/1/20	1	1/4/20	120	25/5/20	3	30/1/21	930	6/7/21	5420	12/9/21	14028	19/11/21	6855	16/2/20	2099	8/3/20	5	26/2/20	0
26/1/20	2	2/4/20	104	26/5/20	9	31/1/21	829	7/7/21	6519	13/9/21	12583	20/11/21	6595	17/2/20	1918	9/3/20	5	27/2/20	0
27/1/20	0	3/4/20	103	27/5/20	11	1/2/21	836	8/7/21	7058	14/9/21	11786	21/11/21	7006	18/2/20	1775	10/3/20	5	28/2/20	0
28/1/20	6	4/4/20	89	28/5/20	11	2/2/21	836	9/7/21	9276	15/9/21	13798	22/11/21	6428	19/2/20	407	11/3/20	7	29/2/20	0
29/1/20	0	5/4/20	102	29/5/20	1	3/2/21	795	10/7/21	9326	16/9/21	13897	23/11/21	5126	20/2/20	453	12/3/20	1	1/3/20	0
30/1/20	0	6/4/20	51	30/5/20	0	4/2/21	809	11/7/21	9539	17/9/21	14555	24/11/21	5857	21/2/20	473	13/3/20	2	12/5/20	0
31/1/20	5	7/4/20	38	31/5/20	4	5/2/21	586	12/7/21	8656	18/9/21	14109	25/11/21	6335	22/2/20	1450	14/3/20	1	2/3/20	0
1/2/20	0	8/4/20	111	**2nd Wave**		6/2/21	490	13/7/21	8656	19/9/21	13576	26/11/21	6559	23/2/20	16	15/3/20	1	3/3/20	0
2/2/20	0	9/4/20	54	1/12/20	10	7/2/21	237	14/7/21	9317	20/9/21	12709	27/11/21	6073	24/2/20	214	16/3/20	4	4/3/20	0
3/2/20	0	10/4/20	50	2/12/20	18	8/2/21	186	15/7/21	9186	21/9/21	10919	28/11/21	5854	25/2/20	508	17/3/20	0	5/3/20	0
4/2/20	6	11/4/20	45	3/12/20	13	9/2/21	189	16/7/21	9692	22/9/21	11252	29/11/21	4753	26/2/20	405	18/3/20	5	6/3/20	2
5/2/20	0	12/4/20	33	4/12/20	14	10/2/21	157	17/7/21	10082	23/9/21	13576	30/11/21	4306	27/2/20	433	19/3/20	0	7/3/20	2
6/2/20	0	13/4/20	28	5/12/20	19	11/2/21	201	18/7/21	11397	24/9/21	12697	1/12/21	4886	28/2/20	326	20/3/20	2	8/3/20	10
7/2/20	0	14/4/20	34	6/12/20	14	12/2/21	175	19/7/21	11784	25/9/21	11975	2/12/21	4971	29/2/20	427	21/3/20	1	9/3/20	1
8/2/20	7	15/4/20	30	7/12/20	21	13/2/21	126	20/7/21	11305	26/9/21	12353	3/12/21	4912	1/3/20	575	22/3/20	6	10/3/20	3
9/2/20	0	16/4/20	29	8/12/20	19	14/2/21	166	21/7/21	13002	27/9/21	10288	4/12/21	5896	12/5/20	200	23/3/20	1	11/3/20	4
10/2/20	0	17/4/20	28	9/12/20	25	15/2/21	143	22/7/21	13655	28/9/21	9489	5/12/21	4704	2/3/20	125	24/3/20	5	12/3/20	6
11/2/20	1	18/4/20	33	10/12/20	18	16/2/21	72	23/7/21	14575	29/9/21	10414	6/12/21	4000	3/3/20	120	25/3/20	6	13/3/20	3
12/2/20	0	19/4/20	32	11/12/20	11	17/2/21	175	24/7/21	14260	30/9/21	11646	7/12/21	3525	4/3/20	151	26/3/20	3	14/3/20	6
13/2/20	0	20/4/20	27	12/12/20	12	18/2/21	150	25/7/21	15335	1/10/21	11754	8/12/21	3618	5/3/20	151	27/3/20	2	15/3/20	4
14/2/20	0	21/4/20	19	13/12/20	17	19/2/21	130	26/7/21	15376	2/10/21	11375	9/12/21	4203	6/3/20	79	28/3/20	0	16/3/20	4
15/2/20	1	22/4/20	15	14/12/20	28	20/2/21	82	27/7/21	14150	3/10/21	10828	10/12/21	4193	7/3/20	47	29/3/20	3	17/3/20	5
16/2/20	0	23/4/20	13	15/12/20	9	**4th Wave**		28/7/21	16533	4/10/21	9930	11/12/21	4079	8/3/20	36	30/3/20	3	18/3/20	10
17/2/20	1	24/4/20	15	16/12/20	15	1/6/21	2230	29/7/21	17669	5/10/21	9869	12/12/21	3787	9/3/20	22	31/3/20	0	19/3/20	9
18/2/20	0	25/4/20	53	17/12/20	20	2/6/21	3440	30/7/21	17345	6/10/21	9866	13/12/21	3398	10/3/20	28	1/4/20	0	20/3/20	6
19/2/20	0	26/4/20	15	18/12/20	16	3/6/21	3886	31/7/21	18912	7/10/21	11200	14/12/21	2862	11/3/20	8	2/4/20	2	21/3/20	3
20/2/20	0	27/4/20	9	19/12/20	34	4/6/21	2631	1/8/21	18027	8/10/21	11140	15/12/21	3370	12/3/20	8	3/4/20	0	22/3/20	19
21/2/20	0	28/4/20	7	20/12/20	576	5/6/21	2817	2/8/21	17970	9/10/21	10630	16/12/21	3684	13/3/20	26	4/4/20	2	23/3/20	10
22/2/20	0	29/4/20	9	21/12/20	382	6/6/21	2671	3/8/21	18901	10/10/21	10817	17/12/21	3537	14/3/20	21	5/4/20	1	24/3/20	11
23/2/20	0	30/4/20	7	22/12/20	427	7/6/21	2419	4/8/21	20200	11/10/21	10035	18/12/21	3132	15/3/20	20	6/4/20	2	25/3/20	7
24/2/20	0	1/5/20	6	23/12/20	46	8/6/21	2662	5/8/21	20920	12/10/21	9445	19/12/21	2899	16/3/20	18	7/4/20	0	26/3/20	12
25/2/20	2	2/5/20	6	24/12/20	67	9/6/21	2680	6/8/21	21379	13/10/21	10064	20/12/21	2525	17/3/20	25	8/4/20	2	27/3/20	10
26/2/20	3	28/4/20	7	25/12/20	81	10/6/21	2310	7/8/21	21838	14/10/21	11276	21/12/21	2476	18/3/20	27	9/4/20	4	28/3/20	11
27/2/20	0	29/4/20	9	26/12/20	110	11/6/21	2290	8/8/21	19983	15/10/21	10486	22/12/21	2532	19/3/20	46	10/4/20	3	29/3/20	14
28/2/20	1	30/4/20	7	27/12/20	121	12/6/21	3277	9/8/21	19603	16/10/21	10648	23/12/21	2940	20/3/20	38	11/4/20	1	30/3/20	15
29/2/20	1	1/5/20	6	28/12/20	144	13/6/21	2804	10/8/21	19843	17/10/21	10863	24/12/21	2671	21/3/20	86	12/4/20	1	31/3/20	4
1/3/20	0	2/5/20	6	29/12/20	155	14/6/21	3355	11/8/21	21038	18/10/21	10111	25/12/21	2766	**Hong Kong**		13/4/20	3	1/4/20	11
12/5/20	2	3/5/20	3	30/12/20	250	5/6/21	2817	12/8/21	22782	19/10/21	9122	26/12/21	2532	7/2/20	2	14/4/20	9	2/4/20	15
2/3/20	1	4/5/20	18	31/12/20	194	6/6/21	2671	13/8/21	23418	20/10/21	8918	27/12/21	2437	8/2/20	0	15/4/20	11	3/4/20	6
3/3/20	0	5/5/20	1	1/1/21	279	7/6/21	2419	14/8/21	22086	21/10/21	9727	28/12/21	2305	9/2/20	3	16/4/20	1	4/4/20	1
4/3/20	0	6/5/20	1	2/1/21	216	8/6/21	2662	15/8/21	21882	22/10/21	9810	29/12/21	2575	10/2/20	3	17/4/20	3	5/4/20	1
5/3/20	4	7/5/20	3	3/1/21	315	9/6/21	2680	16/8/21	21157	23/10/21	9742	30/12/21	3037	11/2/20	0	18/4/20	3	6/4/20	4
6/3/20	1	28/4/20	7	4/1/21	745	10/6/21	2310	17/8/21	20128	24/10/21	9351	31/12/21	3111	12/2/20	0	19/4/20	0	7/4/20	4

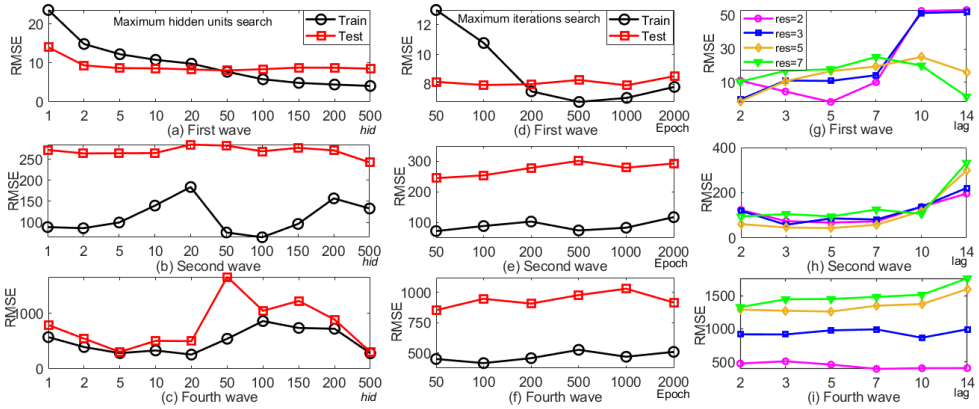


Figure 4 Selection on LSTM hyperparameters (a)-(c) *hid*, (d)-(f) Epoch, and (g)-(i) input-output parameters, *lag* and *res*, for the first, second, and fourth waves of infection in Thailand

Table 2 Hyperparameter selection for LSTM(5,2), LSTM(5,5), and LSTM(7, 2) configurations in predicting daily COVID-19 cases during the first, second, and fourth waves of infection in Thailand, respectively

Parameter	Range	Selected value			Parameter	Range	Selected value		
		1st wave	2nd wave	4th wave			1st wave	2nd wave	4th wave
Time lag (*lag*)	[2,14]	5	5	7	Optimizer	-	ADAM	ADAM	ADAM
Multi-step-ahead forecast (*res*)	[2, 7]	2	5	2	Initial learning rate	[0.0005, 0.1]	0.001	0.001	0.001
					Dropout ratio	[0, 0.8]	0.2	0.2	0.2
Hidden units (*hid*)	[1,500]	50	50	20	Batch size	[2, 128]	32	64	64
Epoch	[50, 2000]	200	200	100	Momentum	[0.1, 0.99]	0.95	0.95	0.95

The optimal LSTM(*lag*, *res*) configurations for predicting the first, second, and fourth waves of infection, as derived from Table 2, are denoted as LSTM(5,2), LSTM(5,5), and LSTM(7,2), respectively.

3.6 Performance Metrics

The performance metrics used to assess the LSTM-based prediction model for daily COVID-19 cases in Thailand include the mean absolute error (MSE) (11) and the root mean square error (RMSE) (12). Additionally, the effectiveness of the lockdown measures is evaluate using $R_{eff}(t)$ criterion (10) in three scenarios: $R_{eff}(1)$ after one day of implementing lockdown, $R_{eff}(7)$ after one week of implementing lockdown, and $R_{eff}(\infty)$ at the end of each wave. As a linear metric, MSE treats errors equally on average, while the quadratic metric RMSE assigns relatively higher wieght to large errors, thereby detecting additional errors. Both MAE and RMSE, negative-oriented metrics where lower scores indicate better performance, are typically used in conjuction to explain the variation in forecast errors. Meanwhile, $R_{eff}(1)$ assesses the immeadiat impact of the control measure, $R_{eff}(7)$ evaluates the adaptiation after the enforcement of control measures, and $R_{eff}(\infty)$ measures the final impact of this measure.

$$\text{MAE}=\frac{1}{N}\sum_{i=1}^{N}\left|\text{Covid}^{\text{actual}}(i)-\text{Covid}^{\text{predict}}(i)\right|. \quad (11)$$

$$\text{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\text{Covid}^{\text{actual}}(i)-\text{Covid}^{\text{predict}}(i)\right)^2}. \quad (12)$$

3.7 Ethics

Because this study solely utilized publicly available secondary data and did not involve human participation, formal ethics approval was not required. Nevertheless, maintaining ethical research standards was a priority, ensuring that the results were not related to the privacy of individual information.

4. Results and Discussion

4.1 Prediction Results of LSTM Model

The predicted daily COVID-19 cases, both before, during, and after the implementation of lockdown measures, as well as $R_{eff}(t)$ and cumulative cases based on these predictions, are depicted in Figures 5, 6, and 7 for the first, second, and fourth waves, respectively. As depicted in Figures 5(a), 6(a), and 7(a), the prediction results of daily cases are consistent with the measured data,

even under high fluctuations in the daily infected case data, which indicates the LSTM's ability to capture long-term dependencies in infectious data. In addition, $R_{eff}(t)$ depicted in Figures 5(b), 6(b), and 7(b), as well as cumulative cases shown in Figures 5(c), 6(c), and 7(c), evaluated using the predicted values obtained from the LSTM is consistent with that evaluated from the measured value.

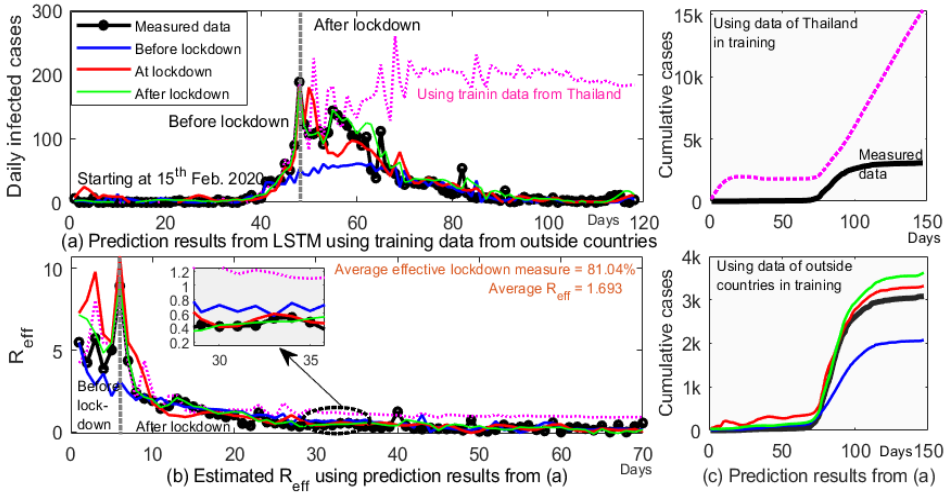


Figure 5 (a) Prediction results of daily infected cases using the LSTM network with different training datasets, (b) R_{eff} based on the prediction results from (a), and (c) Cumulative cases derived from predicted daily cases from (a) for the first wave of COVID-19 in Thailand

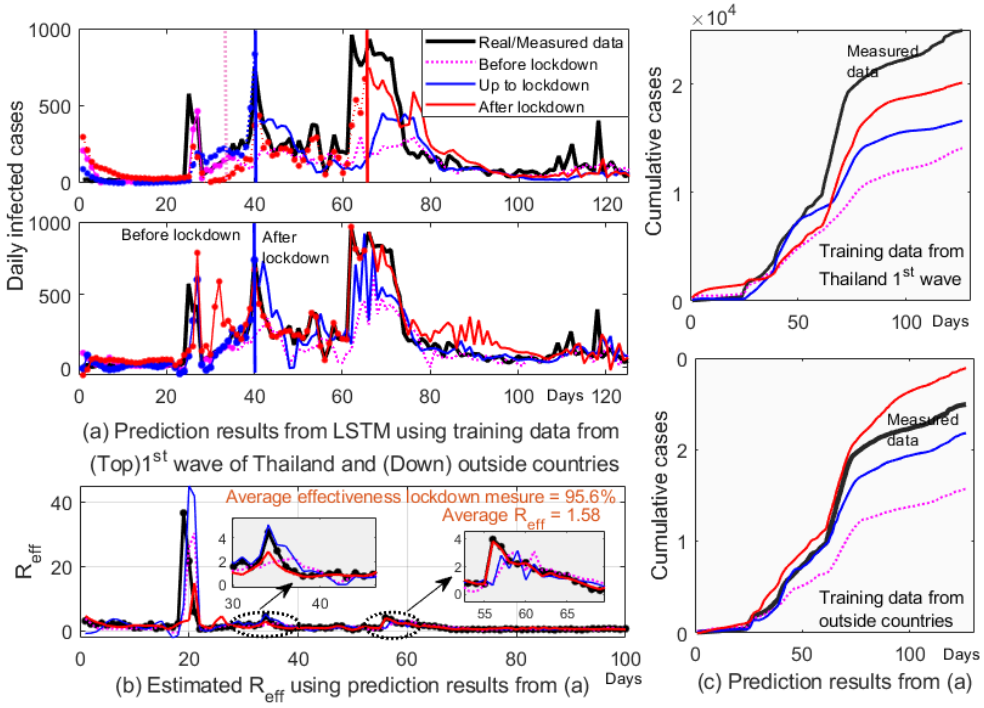


Figure 6 (a) Prediction results of daily infected cases using the LSTM network with different training datasets, (b) R_{eff} based on the prediction results from (a), and (c) Cumulative cases derived from predicted daily cases from (a) for the second wave of COVID-19 in Thailand

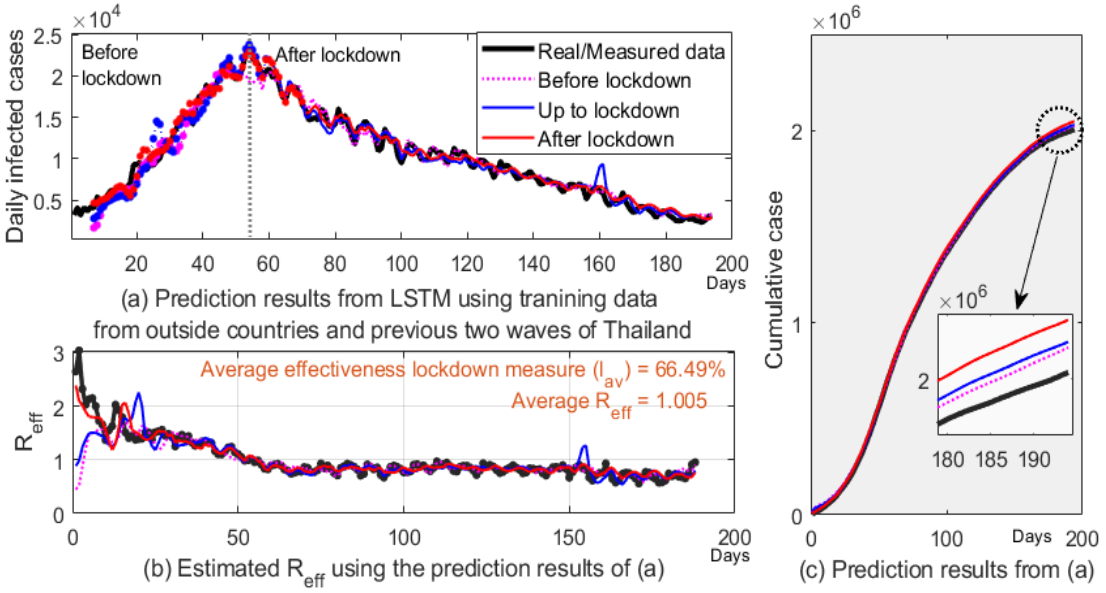


Figure 7 (a) Prediction results of daily infected cases using the LSTM network with different training datasets, (b) R_{eff} based on the prediction results from (a), and (c) Cumulative cases derived from predicted daily cases from (a) for the fourth wave of COVID-19 in Thailand

Table 3 Performance metrics for evaluating daily infected cases prediction and assessing lockdown measure effectiveness using $R_{eff}(t)$ for the first, second, and fourth wave of infection in Thailand

Wave	Training data	Training		Forecasting		$R_{eff}(t)$			
		MAE	RMSE	MAE	RMSE	$R_{eff}(0)$	$R_{eff}(1)$	$R_{eff}(7)$	$R_{eff}(\infty)$
1st	Measured data	-	-	-	-	8.93	4.37	1.45	0.4
	Thailand's data	38.4	43.2	105.3	118.7	10.8	5.5	1.45	0.4
	Outside countries data prior to TL's lockdown	2.3	3.3	18.91	32.79	3.13	1.99	1.99	0.6
	Outside countries and TL data during lockdown	9.2	13.2	12.99	21.55	8.93	4.37	1.2	0.4
	Outside countries and TL data after TL lockdow	2.7	3.9	10.56	17.48	8.93	4.36	1.44	0.4
2nd	Measured data	-	-	-	-	36.57	21.64	0.62	0.28
	1st wave TL data prior to lockdown	59.4	113.1	84.70	110.9	24.93	0.51	1.25	0.49
	1st wave TL data up to lockdown	73.5	155.4	131.7	233.4	56.14	12.0	3.20	0.44
	1st wave TL data after lockdown	101.1	154.9	82.65	124.2	6.43	4.49	0.23	0.27
	Outside countries before lockdown	38.2	93.1	119.4	194.3	30.33	0.26	1.25	0.29
	Outside countries data during lockdown	37.9	88.6	99.88	165.3	45.03	41.79	1.03	0.37
	Outside countries after lockdown	59.4	113.1	84.70	110.9	14.79	1.83	4.45	0.45
4th	Measured data	-	-	-	-	3.00	2.31	1.99	0.81
	Outside countries before lockdown	1680	2930	670	910	1.68	1.54	1.19	0.77
	Outside countries during lockdown	1480	2670	640	840	2.25	2.00	1.17	0.84
	Outside countries after lockdown	790	1090	630	780	2.38	2.09	1.79	0.81

The performance of the LSTM-based daily COVID-19 cases prediction using forecasting errors both training and testing, as well as the assessment of lockdown measure effectiveness using $R_{eff}(t)$, are reported in Table 3. For both the first and second waves of infection, utilizing only Thailand's COVID-19 data in training LSTM model results in higher MAE and RMSE compared to training with data from other countries. Specifically, during the first wave and the associated lockdown period, there is an increase in MAE and RMSE when contrasted with the periods before and after lockdown measure. This elevation is attributed to heightened variations during the lockdown, reaching the maximum number of infected cases. The rapid interruption by this measure causes the infection trend to deviate, leading to higher prediction errors. However, across all cases with different training data, the indifferent between RMSE and MAE indicates minimal variation errors and unlikely occurrence of large errors. Conversely, using training data after the outbreak in the second wave results in higher MAE and RMSE. This is because of the multiple peaks observed after implementing lockdown measures (see Figure 1). Notably, the fourth wave displays very high MAE and RMSE due to a substantial increase in the number of infected cases, exceeding those of the second wave by over 10 times in daily cases or 100 times in cumulative cases. When comparing percentage errors relative to the final epidemic size of 3×10^3, 3×10^4, and 2×10^6 for the first, second, and fourth waves, respectively, Figure 5-7(c), the average percentage errors in testing are determined as 3×10^{-3}, 2.8×10^{-3}, and 3×10^{-4}. These values indicate that LSTM forecasting in the fourth wave is more effective than in the other waves, because of the distribution of daily cases in the fourth wave, which exhibits a Gaussian shape. For both the second and fourth waves, a significant difference between RMSE and MAE, with RMSE > MAE, is observed, indicating a high variation of errors and the presence of large errors. For the $R_{eff}(t)$ criterion, except for the second wave of infection, the rest exhibit good alignment with the measured value corresponding to (8). Furthermore, it is observed that $R_{eff}(t)$ is dynamic and changes over time due to the evolving nature of the epidemic and the implemented interventions. A comparison, including the prediction errors and the assessment of the effectiveness of lockdown measures in Thailand, is presented in Table 4 with the findings from previous research and our proposed LSTM model.

Table 4 Comparison of outcome results between this work with previous research

Research	Model	Prediction performance			Research	Model	Lockdown measure efficacy		
		1st wave	2nd wave	4th wave			1st wave	2nd wave	4th wave
Our research	LSTM	RMSE (17.5) MAE (10.6)	RMSE (84.7) MAE (110.9)	RMSE (630) MAE (780)	Our research	LSTM	96%	98%	66.5%
Tantrakarnapa, & Bhopdhornangkul (2020)	SEIR	RMSE (12.8)	-	-	Patanarapeelert et al. (2022)	SEIQR	91%	80%	-
Wongsathan (2021b)	Gaussian function	RMSE (14.4)	-	-	Polwiang (2023)	SEIR	-	-	60%
	MLR	RMSE (13.8-40.7)	-	-					
Vorathamthongdee, & Chongstitvatana (2023)	LSTM	RMSE (507-1,017) MAE (380-886)			Uansri et al. (2021)	SEIR	-	-	60%
Winalai et al. (2022)	LSTM	RMSE (400-800)							

4.2 Predictions Analysis and Performance Efficiency of Lockdown Measures

In the first wave of infection, utilizing only local data from Thailand before implementing lockdown measures, the predicted results (dotted pink line in Figure 5(a)) deviate significantly from observed data. Daily cases increase and eventually stabilize, resulting in an overestimation of cumulative cases (Figure 5(c) top). This is also evident in Figure 5(b), where $R_{eff}(t) > 1$, indicating an ongoing and uncontrollable outbreak, not matching the actual $R_{eff}(t)$. While using training data from other countries, the prediction results for daily infected cases align well but show underestimation (blue line) and overestimation (green line) of the cumulative cases before and after implementing lockdown (Figure 5(c) down). In contrast, during the implementation of the lockdown, both daily and cumulative predicted cases (red lines) are consistent with measured data. In all cases, $R_{eff}(\infty) < 1$, indicating effective containment with lockdown. To achieve accurate predictions, it is advisable to use training data from countries with similar control measures and make predictions during lockdown implementation.

According to Figure 5(b) and Table 3, $R_{eff}(0)$ is 8.93 before lockdown measures, indicating a severe outbreak situation where one infected person can transmit the virus to about 9 others. However, $R_{eff}(1)$ drops to 4.37 within a day after implementation, with an estimated lockdown effectiveness of 51% ($l(1)$) using (8). After a week, $R_{eff}(7)$ is 1.45, estimating a lockdown effectiveness of 84% ($l(7)$). Notably, $R_{eff}(t)$ remains below 1 for $t > 14$, effectively comprehensive control the outbreak for two weeks, matching the virus's incubation period. At the end of this wave, $R_{eff}(\infty) < 1$ and $l(\infty)$ reaches 96%, matching previous findings (91%) (Patanarapeelert et al., 2022). This suggests that COVID-19 is ongoing but controllable, thereby demonstrating the effectiveness of the lockdown measures for this wave.

Questions arise regarding the high effectiveness of lockdown measures during the first wave of infection in 2020. Was it due to the strict containment measures with few infected cases? In terms of disease control, this measure can be considered successful, but it resulted in widespread social and economic impacts. In early 2021, the second wave of infection experienced a rise in active cases, leading to a more relaxed approach, which, compared to the initial outbreak, resulted in a serious resurgence of the virus. The second wave, marked by multiple peaks (triple peaks), poses a challenging prediction task. Figure 6(a) shows that LSTM predictions using training data from Thailand's first wave do not effectively track epidemic trends. Cumulative cases are underestimated, largely confined to the first and second peaks (Figure 6(c) Top). Utilizing datasets from other countries for LSTM training moderately improves predictions for both daily and cumulative cases (Figure 6(a) and 6(c) Down).

As previously mentioned, the relaxation of lockdown measures resulted in a notably high $R_{eff}(0)$ of approximately 36.5 (see Table 3), mainly due to the presence of numerous clusters. This outbreak primarily affected vulnerable areas, especially in provinces with a significant population of immigrant workers, resulting in a limited number of susceptible individuals and a relatively small number of infections. In response, localized lockdowns were enforced, leading to a reduction in $R_{eff}(1)$ to about 21.6, indicating an effectiveness of approximately 41% ($l(1)$). After one week, $R_{eff}(7)$ significantly dropped to 0.62, with $l(7)$ increasing to approximately 98%, rendering the outbreak controllable. The localized lockdown's effectiveness more than doubled after one week, outperforming other measures, largely attributed to identifying specific vulnerable areas and streamlining lockdown control. However, this effectiveness differs significantly from the 80% reported by Patanarapeelert et al. (2022).

In the fourth wave, with a better understanding of the epidemic, super-spreader events were notably absent. Moreover, people increased awareness and adherence to safety measures after experiencing multiple waves may have led to a reduction in secondary infections, including the efficacy of vaccines, coupled with a higher proportion of the population having received full vaccination, likely contributed to an increase in herd immunity. These correspond to the observed small R_{eff} shown in Table 3, further diminishing the potential for widespread transmission. However, the COVID-19 outbreak may endure longer than expected, mainly due to the increased transmission capability of the Delta variant. Delayed lockdown initiation, as depicted in Figure 1(d), and inadequate case detection have led to an underestimation of the situation. This reflects the least stringent lockdown measure when compared to those implemented during the first and second waves. Challenges in vaccine availability and distribution have hindered the buildup of population immunity and effectiveness against the Delta variant. Until now, the LSTM model effectively tracks epidemic trends in both daily and cumulative cases, owing to

adequate training data from other countries and Thailand's previous three waves. In Figure 7(a) and (c), the model performs well. Notably, in Figure 7(c), predictions before implementing lockdown measures (blue line) outperform those made afterward (red line).

While strict lockdown measures were effective in controlling the first and second waves, led to R_{eff} rapidly reaching near zero within a few weeks, the fourth wave faced challenges due to the implementation of softer measures. The Delta variant's heightened transmission rate, combined with the partial implementation of vaccines, contributed to a reduced effectiveness as reflected in the slower decline of $R_{eff}(t)$ (as seen in Figure 6(b)). $R_{eff}(1)$ decreased from 3.0 to 2.31 (Table 3), indicating an effectiveness of approximately 23% ($l(1)$). After one week, $l(7)$ increased to 34%. These results are consistent with the expectations of the COVID Situation Management Center of Thailand regarding a 20% effective lockdown. At the end of this wave, $R_{eff}(\infty)$ was 0.81, implying an effectiveness of 73% ($l(\infty)$). The average effectiveness of the lockdown measures during this wave was approximately 66.5%, which closely aligns with the 60% figure reported by Polwiang (2023). However, R_{eff} slowly declined over a few months. Given the high daily case count, peaking at 25,000 cases, a targeted vaccination plan was concurrently implemented to reduce new infections in time, a critical step for managing hospital bed occupancy and reducing mortality.

4.3 Cross-Domain Applications of LSTM Models: Beyond COVID-19 Prediction

The application of LSTM models in predicting COVID-19 cases extends beyond public health, offering valuable contributions to various domains. In non-health sectors, LSTM models are pivotal for optimizing resource allocation. This includes applications in supply chain management, where demand forecasting and efficient inventory management lead to cost savings. Additionally, the energy sector benefits from LSTM predictions by optimizing consumption patterns for sustainable practices. Workforce planning can leverage LSTM insights for anticipating staffing needs and optimizing personnel allocation.

The research on lockdown control measures adds applicability, enhancing decision-making across sectors. Integrating LSTM predictions with lockdown effectiveness studies aids workforce planning and informs decision-making in financial markets, climate change impact modeling, transportation planning, and traffic flow predictions. LSTM insights also contribute to educational institutions, assisting in enrollment predictions and assessing the impact of disruptions on educational patterns.

5. Conclusions

In this study, the effectiveness of lockdown measures in controlling the spread of COVID-19 during Thailand's first, second, and fourth waves of infection was evaluated using R_{eff} in conjunction with LSTM-based predictions. Accurate predictions of daily infected cases, along with the utilization of data from countries with prior experience in implementing these measures, consistently demonstrated the success of lockdowns, with R_{eff} consistently remaining below 1. The effectiveness of lockdowns exhibited variation, influenced by factors such as the extent of measures, viral transmission rates, and unique interventions in each wave. Future work should optimize lockdown strategies, account for virus evolution and vaccination efforts, and improve predictive models with broader data sources to better manage COVID-19 and emerging diseases.

6. References

Absar, N., Uddin, N., Khandaker, M. U., & Ullah, H. (2022). The efficacy of deep learning based LSTM model in forecasting the outbreak of contagious diseases. *Infectious Disease Modelling*, *7*(1), 170-183. https://doi.org/10.1016/j.idm.2021.12.005

Ao, S. I., & Fayek, H. (2023). Continual deep learning for time series modeling. *Sensors*, *23*(16), Article 7167. https://doi.org/10.3390/s23167167

Ge, Y., Zhang, W. B., Liu, H., Ruktanonchai, C. W., Hu, M., Wu, X., ... & Lai, S. (2022). Impacts of worldwide individual non-pharmaceutical interventions on COVID-19 transmission across waves and space. *International Journal of Applied Earth Observation and Geoinformation*, *106*, Article 102649. https://doi.org/10.1016/j.jag.2021.102649

Kompunt, P., Yongjoh, S., Aimtongkham, P., Muneesawang, P., Faksri, K., & So-In, C. (2023). A Hybrid LSTM and MLP Scheme for COVID-19 Prediction: A Case Study in Thailand. *Trends in Sciences*, *20*(10), 6884-6884. https://doi.org/10.48048/tis.2023.6884

Mahikul, W., Chotsiri, P., Ploddi, K., & Pan-Ngum, W. (2021). Evaluating the impact of intervention strategies on the first wave and predicting the second wave of COVID-19 in Thailand: a mathematical modeling study. *Biology*, *10*(2),

Article 80. https://doi.org/10.3390/biology10020080

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., ... & Agha, R. (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, *78*, 185-193. https://doi.org/10.1016/j.ijsu.2020.04.018

Patanarapeelert, K., Songprasert, W., & Patanarapeelert, N. (2022). Modeling dynamic responses to COVID-19 epidemics: a case study in Thailand. *Tropical Medicine and Infectious Disease*, *7*(10), Article 303. https://doi.org/10.3390/tropicalmed7100303

Polwiang, S. (2023). The lockdown and vaccination distribution in Thailand's COVID-19 epidemic: A model study. *Infectious Disease Modelling*, *8*(2), 551-561. https://doi.org/10.1016/j.idm.2023.05.002

Rosero-Bixby, L., & Miller, T. (2022). The mathematics of the reproduction number R for Covid-19: A primer for demographers. *Vienna Yearbook of Population Research*, *20*, 1-24. https://doi.org/10.1553/populationyearbook2022.res1.3

Russell, T. W., Wu, J. T., Clifford, S., Edmunds, W. J., Kucharski, A. J., & Jit, M. (2021). Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study. *The Lancet Public Health*, *6*(1), e12-e20. https://doi.org/10.1016/S2468-2667(20)30263-2

Sengsri, S., & Khunratchasana, K. (2023). Comparison of machine learning algorithms with regression analysis to predict the COVID-19 outbreak in Thailand. *Indonesian Journal of Electrical Engineering and Computer Science*, *31*(1), 299-304. https://doi.org/10.11591/ijeecs.v31.i1.pp299-304

Sunthornwat, R., & Areepong, Y. (2021). Predictive models for the number of cumulative cases for spreading coronavirus disease 2019 in the world. *Engineering & Applied Science Research*, *48*(4). 432-435. https://ph01.tci-thaijo.org/index.php/easr/ article/view/241789

Tantrakarnapa, K., & Bhopdhornangkul, B. (2020). Challenging the spread of COVID-19 in Thailand. *One Health*, *11*, Article 100173. https://doi.org/10.1016/j.onehlt.2020.100173

Teekasap, P., Tuangratananon, T., Phaiyarom, M., & Suphanchaimat, R. (2022). Misperceptions about the Impact of Lockdown on the Number of Newly Reported COVID-19 Cases. *Outbreak, Surveillance, Investigation & Response (OSIR) Journal*, *15*(1), 28-32. https://doi.org/10.59096/osir.v15i1.262499

Uansri, S., Tuangratananon, T., Phaiyarom, M., Rajatanavin, N., Suphanchaimat, R., & Jaruwanno, W. (2021). Predicted impact of the lockdown measure in response to coronavirus disease 2019 (COVID-19) in greater Bangkok, Thailand, 2021. *International Journal of Environmental Research and Public Health*, *18*(23), Article 12816. https://doi: 10.3390/ijerph182312816

Vorathamthongdee, S., & Chongstitvatana, P. (2023). Predictive Analysis of COVID-19 Patients in Thailand using Multiple Countries Data. *ECTI Transaction on Application Research and Development*, *3*(1), 1-8. https://doi.org/10.37936/ectiard.2023-3-1.248647

Werbos, P. J. (1988). Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, *1*(4), 339-356. https://doi.org/10.1016/0893-6080(88)90007-X

Winalai, C., Chadsuthi, S., Anupong, S., Pattanasiri, B., Suttirat, P., & Modchang, C. (2022, June 8-11). *Long short-term memory model for the COVID-19 pandemic in Thailand* [Conference presentation]. Proc. for the 25th International Annual Symposium on Computational Science and Engineering (ANSCSE25), Khon Kaen, Thailand. https://www.researchgate.net/publication/361616232_Long Short-Term_Memory_Model_for_the_COVID-19 Pandemic in Thailand

Wongsathan, R. (2021a). Real-time prediction of the COVID-19 epidemic in Thailand using simple model-free method and time series regression model. *Walailak Journal of Science and Technology (WJST)*, *18*(14), Article 10028. https://doi.org/10.48048/wjst.2021.10028

Wongsathan, R. (2021b). Real-time prediction of the COVID-19 epidemic inThailand using simple model-free method and time series regression model. *Walailak Journal of Science & Technology*, *18*(41), 1-11. https://doi.org/10.48048/wjst.2021.10028

Wongsathan, R., & Puangmanee, W. (2023). DNN and RNN Models derived by PSO for Predicting COVID-19 and R t under Control Measure. *International Journal of Computer Information Systems & Industrial Management Applications*, *15*, 418-428. https://doi.org/10.11113/aej.v14.2008