OSIR, September 2024, Volume 17, Issue 3, p.146-154

Outbreak, Surveillance, Investigation & Response (OSIR) Journal

Field Epidemiology Training Program, Division of Epidemiology
Department of Disease Control, Ministry of Public Health, Thailand
Tel: +6625903894, Fax: +6625903845, Email: osireditor@osirjournal.net, https://he02.tci-thaijo.org/index.php/OSIR

Accuracy of COVID-19 Prediction Modeling Techniques

Panithee Thammawijaya[1*], Rapeepong Suphanchaimat[2,3]

1 Department of Disease Control, Ministry of Public Health, Thailand

2 Division of Epidemiology, Department of Disease Control, Ministry of Public Health, Thailand

3 International Health Policy Program, Ministry of Public Health, Thailand

*Corresponding author, email address: viewfetp@gmail.com

Received: 30 Aug 2024; Revised: 10 Sep 2024; Accepted: 12 Sep 2024
https://doi.org/10.59096/osir.v17i3.270877

Abstract

The unprecedented impact of the COVID-19 pandemic has revealed that forecasting capability is critically needed in making strategic decisions and formulating reasonable countermeasures. This study aimed to assess the predictive accuracy in forecasting the numbers of COVID-19 cases using Thailand's national COVID-19 surveillance database from January 2020–June 2021 based on three analytical models: a susceptible-exposed-infectious-recovery compartmental model, an auto-regressive integrated moving average model, and a long short-term memory (LSTM) network model. All forecasting methods had model parameters adjusted weekly according to the most recent situation and predictive accuracy measures, including the mean absolute percentage error (MAPE). We found that the MAPE values ranged from 19.65%–22.54%, 28.95%–32.35%, 47.78%–53.55%, and 75.03–84.91% for forecasting one, two, four, and eight weeks ahead, respectively. Among the three models, the LSTM model had slightly higher accuracy than the other two models within the same forecasting range. These prediction models can be used for short-range forecasts in other similar settings while long-range forecasting requires monitoring and updating periodically.

Keywords: COVID-19, forecast, predictive accuracy, compartmental model, ARIMA, LSTM

Introduction

The coronavirus disease 2019 (COVID-19) pandemic, caused by the severe acute respiratory syndrome coronavirus 2 during 2019–2023 has had an unprecedented impact around the world.[1,2] The pandemic was declared a public health emergency of international concern by the World Health Organization in January 2020, and lasted until May 2023.[3] At least 765 million cases and 6.9 million deaths associated with COVID-19 infection were reported worldwide.[4] The speed and scale of the outbreak required countries to mobilize enormous resources including vaccines, medicines, equipment, and medical supplies to prevent the spread of disease.[4,5]

To manage resources more effectively and efficiently, many disease forecasting techniques have been applied in making strategic decisions and formulating reasonable countermeasures. For example, traditional regression models and time series analysis methods, such as auto-regressive integrated moving average (ARIMA), exponential smoothing techniques, and decomposition, were used in forecasting various diseases such as influenza, dengue hemorrhagic fever, and COVID-19.[6–8] The susceptible-exposed-infectious-recovered (SEIR) model, which is a well-known mathematical compartmental model, is another commonly used technique, especially for estimating the effects of interventions.[9,10] Recent advances in digital technology, big data management, and computational capability have also enabled the applicability of data-intensive methods such as machine learning and deep learning. Many techniques, such as random forest, convoluted neural networks, and long short-term memory (LSTM) networks have been used for predicting various kinds of diseases, including COVID-19.[11–13] Machine learning is arguably more flexible than traditional statistical forecasting models, particularly for identifying complex and non-linear relationships. LSTM is a type of machine learning technique under the framework of recurrent neural network models, and is commonly used in disease forecasting.[14–16]

Thailand was the first country after China to detect a confirmed case of COVID-19, which was reported on 12 Jan 2020.[17] The first large disease cluster started in March 2020, and thereafter cases were continuously reported across the country.[18] The most severe period in terms of cases and deaths occurred in 2021, and it was not until 2022 that the severity started to abate.[19] In October 2022, the Thai Ministry of Public Health (MOPH) announced the end of the epidemic in Thailand and classified COVID-19 as an endemic disease.[20]

The Department of Disease Control (DDC), Thailand's national agency for disease prevention and control under the MOPH, established the National COVID-19 Surveillance (NCS) system in January 2020. The NCS is an electronic surveillance database that stores data of individual COVID-19 cases reported from all hospitals in the country. The data are analyzed and disseminated to all stakeholders via the DDC COVID-19 situation dashboard.[19] During the pandemic, efforts were made to forecast COVID-19 cases based on the NCS data using various modeling techniques.[21,22] However, there was a lack of systematic evaluation of the performance of these techniques when applied to Thailand's context. Since the pattern of COVID-19 cases and the quality of surveillance systems vary among countries, this study aimed to assess the accuracy of forecasting weekly COVID-19 cases at one, two, four, and eight weeks ahead using three models, namely SEIR, ARIMA, and LSTM models, based on data from the Thai NCS database during 2020–2021. Results from this study should provide a rationale on the selection of appropriate forecasting techniques, execution, and interpretation for better responses during future pandemics.

Methods

This is a predictive modeling study based on retrospective data on the number of weekly COVID-19 cases in Thailand, using three modeling techniques as described below.

Data

This study used weekly COVID-19 cases during 1 Jan 2020 to 29 Jun 2021. During the study period, all COVID-19 cases were diagnosed in hospital and every positive case was confirmed either by reverse transcription polymerase chain reaction or rapid antigen test. As of the end of June 2021, COVID-19 vaccination coverage for two doses was less than five percent of the total Thai population.[23] To minimize the influence of immunization coverage on the assessment of predictive accuracy, this study excluded cases diagnosed after 29 Jun 2021.

Prediction Models

Susceptible-exposed-infectious-recovered (SEIR) model

To set up the compartmental model, the population in Thailand was divided into four compartments, namely susceptible (S), exposed (E), infectious (I), and recovered (R), representing those without infection, non-immune and therefore susceptible for getting infection; those contracting the infection but still not-transmissible; those contracting the infection and transmissible; and those recovering from the disease and not transmissible, respectively. Since dying from COVID-19 should only occur among those who developed symptoms, the mortality rate from the disease was added to the "I" compartment. The dynamics of the SEIR model are defined in the following differential equations:

$$\frac{dS}{dt} = -\frac{\beta S\, I}{N}$$

$$\frac{dE}{dt} = \frac{\beta S\, I}{N} - \sigma\, E$$

$$\frac{dE}{dt} = \sigma\, E - \gamma\, I - \mu\, I$$

$$\frac{dR}{dt} = \gamma\, I$$

$$\frac{dD}{dt} = \mu\, I$$

where S, E, I, R stand respectively for the numbers of cases in the population in the S, E, I, and R compartments at each time step; N and D are the estimated population and COVID-19-associated deaths each time step, respectively; β is the per capita transmission coefficient (or effective contact rate); σ is the reciprocal of the latency period; γ is the recovery rate; and μ is the COVID-19-associated mortality rate among cases. The total population of Thailand in 2020 was 66,534 thousand and the initial values of S, E, I, and R were estimated and revised weekly based on the numbers of reported cases and deaths.[19,24] The β parameter was estimated using the effective reproductive number (R_t) calculated by the team from the Robert Koch Institute.[25] Parameters σ and γ were calculated using the reciprocal of the latent period of 5.5 days and infectious period of 10 days, respectively.[26,27] Parameter μ is the estimated case-fatality ratio of COVID-19 and equal to 1.37%.[28] The model's initial values were reviewed and adjusted weekly when the mean absolute percentage error of the most recent four weeks was larger than 30%. The adjusted values were approved by an epidemiologist and an infectious disease expert. The model transmission dynamics were run using the R language and environment.[29]

Auto-regressive integrated moving average (ARIMA) model

The general form for the ARIMA (p, d, q) model with auto-regressive parameter (p), moving average parameter (q), and differencing parameter (d) was applied in this study. Due to the seasonal pattern of COVID-19, we also included a seasonal component in the model. The structure of the basic ARIMA (p, d, q) model is given by:

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where ϕ_a (a=1, 2,…, p) and θ_b (b=0, 1, 2,…, q) are the auto-regressive and moving average parameters, i.e., regression coefficients, of the model, respectively. Y_t and ε_t represent the value to be predicted and its error at time week t. Parameter θ_0 is the intercept of the regression line. We used the "auto.arima" function in the "forecast" package in R to identify the appropriate and non-stationary data adjustment and ARIMA model structure, including seasonality components.[29,30] The function algorithm uses conditional-sum-of-squares to find starting values, then fits the model using maximum likelihood. We used the AIC, AICc and BIC values to choose the best model.

Long short-term memory (LSTM) model

We employed a two-layer LSTM network structure. We normalized COVID-19 cases using the Min-Max scaling method to scale the data between 0–1. We used a 4-week moving historical data window to predict future cases. We set the 50-neuron unit for both the first and the second layers. The model's weights were adjusted using the backpropagation through time method. We compiled the model using the loss function of the mean square error (MSE), optimizer (Adam), epoch number of 300 and learning rate of 0.005, as suggested by a previous study.[31] All LSTM network models were conducted in R.[29]

Model Training, Testing and Predictive Accuracy Assessment

The number of COVID-19 cases from 1 Jan 2020–29 Jun 2021 was analyzed. Cases reported between 1 Jan 2020–29 Dec 2020 (the first 52 weeks in the dataset) were used as the training set in all three models for predicting the number of cases at weeks 53, 54, 56, and 60. Cases diagnosed in weeks 1–53 were then used to predict the number of cases in weeks 54, 55, 57, and 61. The model fitting and parameter adjustment were conducted weekly and this process continued until the end of the study. After obtaining the predicted values of the targeted future weeks, each of the predicted values was compared with the actual data, (the testing set), for each week. To compare the predictive accuracy of the three models quantitatively, the following metrics were used: mean absolute error (MAE), mean absolute percentage error (MAPE) and the coefficient of determination (R^2) as given by the following equations:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|Y_i - \hat{Y}_i|$$

$$\text{MAPE} = \left(\frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i-\hat{Y}_i|}{Y_i}\right)x100$$

$$R^2 = 1-\frac{\sum_{i=1}^{N}(Y_i-\hat{Y}_i)^2}{\sum_{i=1}^{N}(Y_i-\bar{Y})^2}$$

where Y_i is the actual number of cases, $\hat{Y}_i$ is the predicted number of cases, and $\bar{Y}$ is the average number of actual cases during the study period.

The model development and the evaluation of predictive accuracy were calculated for the whole country and for two sub-national areas: (1) Greater Bangkok Region, consisting of Bangkok and three surrounding provinces: Nonthaburi, Samut Prakan and Pathum Thani, and (2) the remaining provinces of Thailand.

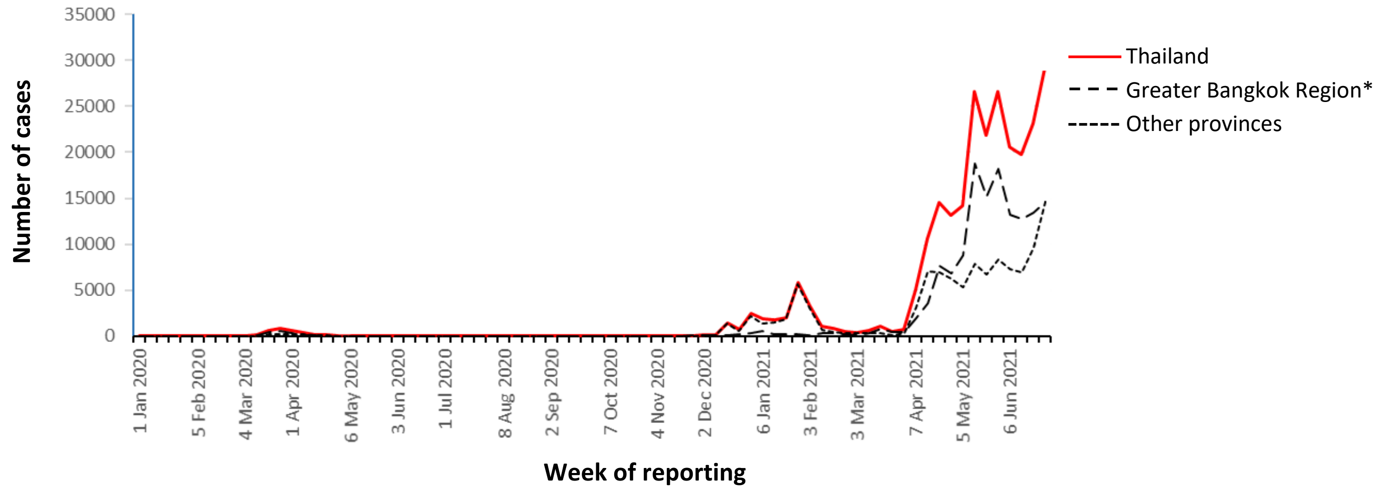
Ethics

This study used aggregated COVID-19 cases from the NCS database of the DDC with no individual or personal identifiable information. This analysis is one of DDC's public health mandates in preparing for a better response to the next pandemic and was therefore exempted from ethical review for research in humans.

Results

COVID-19 Epidemic Pattern in Thailand, January 2020–June 2021

The first COVID-19 case in Thailand was reported in the second week of 2020 (Figure 1). Subsequently, 1–12 cases per week were reported until the first large-scale outbreak occurred in mid-March, lasting until the end of April. In this first outbreak period, over 100 cases were reported weekly, with the peak week having more than 800 cases. After April 2020, the number of weekly cases decreased to fewer than 100 and this continued until December 2020 when a second large-scale outbreak occurred. The number of cases increased from two to three figures and in some weeks reached four figures. The number of cases in this wave of the outbreak peaked in the last week of January 2021, with more than five thousand cases reported. Subsequently, the number of cases decreased to less than 1,000 per week. However, in early April 2021, another outbreak occurred, with cases increasing to over ten thousand per week. By the end of June, the outbreak was still present. The epidemic from the beginning until June 2021 revealed cases that were mostly (56%) from the Greater Bangkok Region, with the remainder scattered in other provinces throughout the country. However, the outbreak patterns in both areas were similar.



**Greater Bangkok Region consists of Bangkok Metropolitan, Nonthaburi, Samut Prakan, and Pathum Thani provinces*

Figure 1. Number of reported COVID-19 cases in Thailand, Greater Bangkok Region, and other provinces by week of reporting, 1 Jan 2020–29 Jun 2021

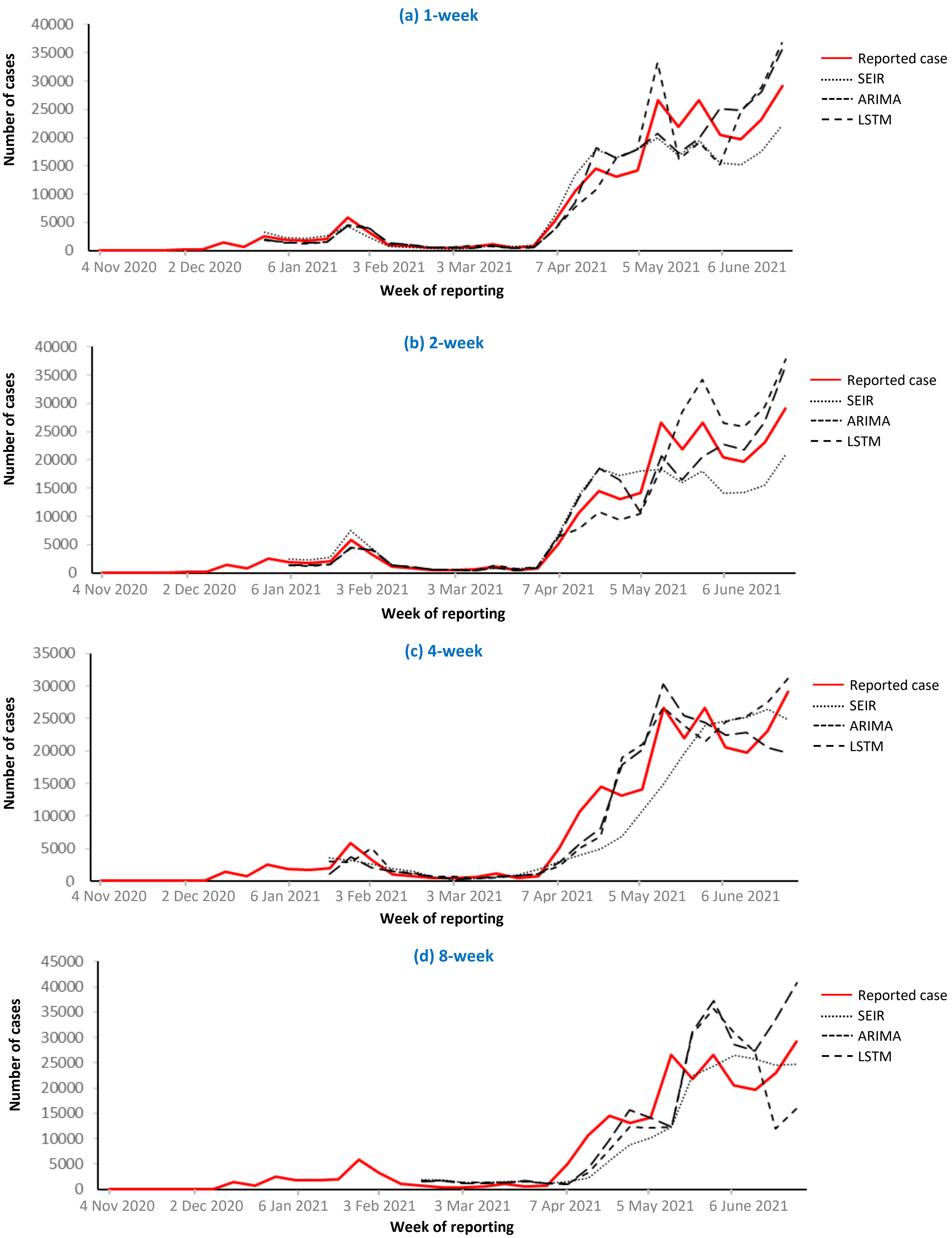
Predictive Accuracy

Figure 2 illustrates the predicted weekly numbers of COVID-19 cases using the SEIR, ARIMA, and LSTM models at one, two, four, and eight weeks ahead (Figure 2a, 2b, 2c, and 2d, respectively) compared to the reported numbers during 30 Dec 2020–29 Jun 2021. The differences between the predicted and reported numbers, i.e., prediction error, varied by forecasting model and time. During periods of severe outbreaks, the prediction error was higher and longer forecasts were associated with larger prediction errors. The relationships between the actual and predicted number of cases are visualized in Figure 3. The pattern was similar to that shown in Figure 2; longer forecasts were associated with weaker correlations, i.e., lower values of R^2. Table 1 quantitatively characterizes the prediction errors for each model and forecast week. The MAE and MAPE values for all models increased with increasing forecasting period. The MAPE values of the three models ranged from 19.65%–22.54%, 28.95%–32.35%, 47.78%–53.55%, and 75.03–84.91% for the forecasts at one, two, four, and eight weeks ahead, respectively. Among the three models, the LSTM model had consistently lower MAE and MAPE values compared to the other two models within the same forecasting range.

Table 1. Accuracy measures of COVID-19 predictions at one, two, four, and eight weeks ahead using SEIR, ARIMA, and LSTM models compared with weekly reported COVID-19 cases in Thailand, 30 Dec 2020–29 Jun 2021

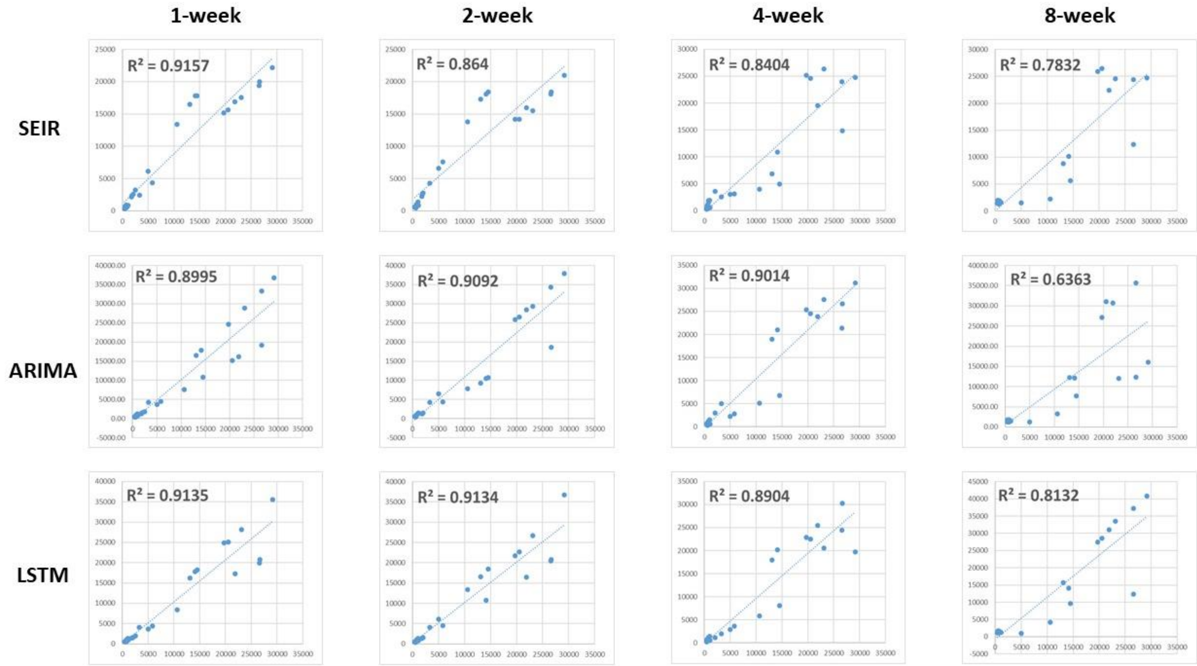
Model	Accuracy measure	Forecasting period (weeks ahead)			
		1	2	4	8
SEIR	MAE	2,349	2,854	3,373	4,169
	MAPE	21.12	32.35	49.84	78.70
	R^2	0.92	0.86	0.84	0.78
ARIMA	MAE	2,466	2,805	3,630	5,002
	MAPE	22.54	30.92	53.55	84.91
	R^2	0.90	0.91	0.90	0.64
LSTM	MAE	2,224	2,111	2,897	3,985
	MAPE	19.65	28.95	47.78	75.03
	R^2	0.91	0.91	0.89	0.81

SEIR: susceptible-exposed-infectious-recovered compartmental model. ARIMA: auto-regressive integrated moving average model. LSTM: long short-term memory network model. MAE: mean absolute error. MAPE: mean absolute percentage error. R^2: coefficient of determination.



SEIR: susceptible-exposed-infectious-recovered compartmental model. ARIMA: auto-regressive integrated moving average model. LSTM: long short-term memory network model

Figure 2. Number of reported COVID-19 cases of Thailand by week of reporting and predicted numbers at one (2a), two (2b), four (2c), and eight (2d) weeks ahead using SEIR, ARIMA, and LSTM models, 30 Dec 2020–29 Jun 2021



SEIR: susceptible-exposed-infectious-recovered compartmental model. ARIMA: auto-regressive integrated moving average model. LSTM: long short-term memory network model.

Figure 3. Scatterplot and coefficient of determination (R2) between numbers of reported COVID-19 cases in Thailand (x-axis) by week of reporting, 30 Dec 2020–29 Jun 2021, and predicted numbers (y-axis) at one, two, four, and eight weeks ahead using SEIR, ARIMA, and LSTM models

Subgroup Analysis for Predictive Accuracy by Area

The predictive accuracy based on the MAE of the two sub-national areas, Greater Bangkok Region and other provinces, are shown in Table 2. For both areas, the predictive accuracy had a similar pattern to that of the whole country. However, based on the forecasting performance, the accuracy was higher in the Greater Bangkok Region.

Table 2. Mean absolute percentage error (MAPE) of COVID-19 forecasts at one, two, four, and eight weeks ahead using SEIR, ARIMA, and LSTM models, compared with weekly reported COVID-19 cases of Greater Bangkok Region* and other provinces of Thailand, 30 Dec 2020–29 Jun 2021

Model	Area	MAPE at forecasting period (weeks ahead)			
		1	2	4	8
SEIR	Greater Bangkok Region	19.10	28.98	47.55	76.55
	Other Provinces	24.55	33.59	55.08	87.39
ARIMA	Greater Bangkok Region	20.22	29.11	50.12	81.44
	Other Provinces	25.78	31.47	56.05	86.67
LSTM	Greater Bangkok Region	18.15	25.02	46.89	72.50
	Other Provinces	21.31	30.40	50.02	78.89

*SEIR: susceptible-exposed-infectious-recovered compartmental model. ARIMA: auto-regressive integrated moving average model. LSTM: long short-term memory network model. *Greater Bangkok Region consists of Bangkok Metropolitan, Nonthaburi, Samut Prakan, and Pathum Thani provinces.*

Discussion

This study demonstrates the real-world performance of three predictive models commonly used for COVID-19 within the context of Thailand's epidemic situation and disease surveillance system. Overall, our study results indicate that longer forecasts were less accurate regardless of the model used. This is consistent with well-known concepts of forecasting, particularly during a highly dynamic situation such as an epidemic of a newly emerging disease.[32,33] We have shown quantitatively the level of forecasting inaccuracy.

The MAPE values of the 1-, 2-, and 4-week ahead forecasts for all three models were approximately 20, 30 and 50%, respectively. These findings are consistent with a study demonstrating accuracy of various models conducted by the U.S. Centers for Disease Control and Prevention for the same forecast horizons while other studies reported either lower or higher values.[34–36] For the 8-week ahead forecasts, all three models gave errors exceeding 70%. Although there are very limited studies on long-range COVID-19 forecasts using the models presented in this study, the marked increase in inaccuracy over longer time horizons (e.g., four weeks or more) found in this study highlights the fact that these models may only be suitable for short-term forecasts, especially during times of high volatility such as during a pandemic.[32,33]

Among the three models, the LSTM model slightly outperformed the others, consistent with other studies.[37,38] One possible explanation is that LSTM model has a unique algorithm that allows it to "remember" past values (and errors) to inform future predictions longer than many machine learning techniques.[15,16,31,37] For the SEIR and ARIMA models, which were not very different in term of short-term forecast accuracy, what matters may be practical issues. One can conduct ARIMA using an automated tool, e.g., "auto.arima" function in R, that facilitates identifying the best model (as tested in this study), whereas the SEIR model requires modelers to design the structure and set parameters by themselves based on a literature review and data from multiple sources.[30,39,40] In contrast to an SEIR model, traditional ARIMA models require a time series to forecast its future values with or without the use of other exogenous information.[32,33] Therefore, in practice, using ARIMA for general short-range forecasting is likely to be a better option. SEIR will play an important role in forecasting different scenarios, such as comparing the effects of alternative measures of different types.[39,40] Another observation found in this study (Figure 2) is that the predicted values changed more slowly than the actual values. This reflects lags in forecasts usually found in models that rely on calculating forecasts from actual values from the recent past.[33]

Our models could forecast COVID-19 cases in Bangkok and surrounding areas more accurately than in other provinces, although the accuracy trajectories of the two areas were similar to the whole country. This phenomenon may be partly explained by the fact that the population in Bangkok and nearby provinces are denser and have similar characteristics to large cities, i.e., they are more homogeneous, when compared to other provinces, which are divided into urban and rural areas. The pattern of an epidemic spreading in the former were therefore more clear and less diverse than in the latter.

Limitations

One important limitation of this study is that the reported disease information could be incomplete, which is commonly found in disease surveillance systems. This issue might cause the forecasts to be less accurate due to the inaccuracy of the amount and pattern of the input data in the model. However, this limitation should be acceptable as the aim of this study was to assess the predictive performance under real circumstances. Another point to consider is that we chose to analyze the data at a time when most of the country's population were not vaccinated. This was to prevent the rapid increase in vaccine coverage from interfering with the accuracy assessment of our forecasts. This may result in differences in accuracy if these models are deployed at a time when a high proportion of the population has been vaccinated.

Conclusion and Recommendations

This study presented three models for forecasting emerging disease situations whose accuracy could be arguably acceptable for short-range forecasts when knowledge about the disease is limited. We found that the SEIR, ARIMA, and LSTM models had a similar accuracy for short-range (less than two weeks) forecasts and the LSTM model was slightly more accurate than the others. However, long-range forecasts were less accurate. Therefore, researchers using these models should monitor and update the forecasts periodically. Although this study was conducted in the context of Thailand, the results of the study are likely to reflect characteristics of models that can be applied in other countries experiencing similar epidemics.

Acknowledgements

The authors would like to thank the Division of Epidemiology, Department of Disease Control, Ministry of Public Health, Thailand for providing the data.

Funding

This study has no funding support.

Conflicts of Interests

The authors of this study have no conflicts of interest.

Suggested Citation

Thammawijaya P, Suphanchaimat R. Accuracy of COVID-19 prediction modeling techniques. OSIR. 2024 Sep;17(3):146–54. doi:10.59096/osir.v17i3.270877.

References

1. World Health Organization. The impact of COVID-19 on global health goals [Internet]. Geneva: World Health Organization; 2021 May 21 [cited 2024 Jul 21]. <https://www.who.int/news-room/spotlight/the-impact-of-covid-19-on-global-health-goals>

2. World Health Organization. The impact of COVID-19 on health and care workers: a closer look at deaths [Internet]. Gevena: World Health Organization; 2021 [cited 2024 Jul 21]. 18 p. <https://iris.who.int/handle/10665/345300>

3. World Health Organization. Statement on the fifteenth meeting of the IHR (2005) emergency committee on the COVID-19 pandemic [Internet]. Geneva: World Health Organization; 2023 May 5 [cited 2024 Aug 30]. <https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic>

4. World Health Organization. WHO COVID-19 dashboard [Internet]. Geneva: World Health Organization; 2024 [cited 2024 Jul 21]. <https://data.who.int/dashboards/covid19/cases?n=c>

5. Mathematica. Surging the public health workforce: lessons learned from the COVID-19 response at state, tribal, local, and territorial public health agencies [Internet]. Washington (DC): Office of the Assistant Secretary for Planning and Evaluation; 2023 Sep [cited 2024 Jul 21]. <https://aspe.hhs.gov/sites/default/files/documents/da3419d752f6f8c0359a126a9045fefb/surging-the-ph-workforce.pdf>

6. Song X, Xiao J, Deng J, Kang Q, Zhang Y, Xu J. Time series analysis of influenza incidence in Chinese provinces from 2004 to 2011. Medicine. 2016 Jun;95(26):e3929. doi:10.1097/MD.0000000000003929.

7. Karasinghe N, Peiris S, Jayathilaka R, Dharmasena T. Forecasting weekly dengue incidence in Sri Lanka: Modified Autoregressive Integrated Moving Average modeling approach. PLOS ONE. 2024 Mar 8;19(3):e0299953. doi:10.1371/journal.pone.0299953.

8. Abdelaziz M, Ahmed A, Riad A, Abderrezak G, Djida A. Forecasting daily confirmed COVID-19 cases in Algeria using ARIMA models. East Mediterr Health J. 2023 Jul 31;29(7):515–9. doi:10.26719/emhj.23.054.

9. Newton EA, Reiter P. A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics. Am J Trop Med Hyg. 1992 Dec;47(6):709–20. doi:10.4269/ajtmh.1992.47.709.

10. Abernethy GM, Glass DH. Optimal COVID-19 lockdown strategies in an age-structured SEIR model of Northern Ireland. J R Soc Interface. 2022 Mar;19(188):20210896. doi:10.1098/rsif.2021.0896.

11. Li G, Li Y, Han G, Jiang C, Geng M, Guo N, et al. Forecasting and analyzing influenza activity in Hebei Province, China, using a CNN-LSTM hybrid model. BMC Public Health. 2024 Aug 12;24(1):2171. doi:10.1186/s12889-024-19590-8.

12. Xu J, Xu K, Li Z, Meng F, Tu T, Xu L, et al. Forecast of Dengue Cases in 20 Chinese Cities Based on the Deep Learning Method. Int J Environ Res Public Health. 2020 Jan 10;17(2):453. doi: 10.3390/ijerph17020453.

13. Iqbal M, Al-Obeidat F, Maqbool F, Razzaq S, Anwar S, Tubaishat A, Khan MS, Shah B. COVID-19 Patient Count Prediction Using LSTM. IEEE Trans Comput Soc Syst. 2021 Feb 19;8(4):974–81. doi:10.1109/TCSS.2021.3056769.

14. Santangelo OE, Gentile V, Pizzo S, Giordano D, Cedrone F. Machine learning and prediction of infectious diseases: a systematic review. Mach Learn Knowl Extr. 2023 Feb 1;5(1):175–98. doi:10.3390/make5010013.

15. Absar N, Uddin N, Khandaker MU, Ullah H. The efficacy of deep learning based LSTM model in forecasting the outbreak of contagious diseases. Infect Dis Model. 2022 Mar;7(1):170–83. doi:10.1016/j.idm.2021.12.005.

16. Sembiring I, Wahyuni SN, Sediyono E. LSTM algorithm optimization for COVID-19 prediction model. Heliyon. 2024 Feb 16;10(4):e26158. doi:10.1016/j.heliyon.2024.e26158.

17. World Health Organization. Thailand responding to the novel coronavirus [Internet]. Geneva: World Health Organization; 2020 Jan 13 [cited 2024 Aug 12]. <https://www.who.int/thailand/news/detail/13-01-2020-thailand-responding-to-the-novel-coronavirus>

18. Jindahra P, Wongboonsin K, Wongboonsin P. Demographic and initial outbreak patterns of COVID-19 in Thailand. J Pop Research. 2022;39(4):567–88. doi:10.1007/s12546-021-09276-y.

19. Thailand Department of Disease Control. DDC COVID-19 Interactive Dashboard [Internet]. Nonthaburi: Department of Disease Control; [cited 2024 Aug 12] <https://ddc.moph.go.th/covid19-dashboard/?dashboard=main>. Thai.

20. TAT Newsroom. Thailand ends COVID-19 emergency decree on 30 September 2022 [Internet]. Bangkok: Tourism Authority of Thailand; 2022 Sep 23 [cited 2024 Aug 12]. <https://www.tatnews.org/2022/09/thailand-ends-covid-19-emergency-decree-on-30-september-2022/>

21. Suphanchaimat R, Teekasap P, Nittayasoot N, Phaiyarom M, Cetthakrikul N. Forecasted Trends of the new COVID-19 epidemic due to the Omicron Variant in Thailand, 2022. Vaccines. 2022 Jun 27;10(7):1024. doi:10.3390/vaccines 10071024.

22. Winalai C, Anupong S, Modchang C, Chadsuthi S. LSTM-Powered COVID-19 prediction in central Thailand incorporating meteorological and particulate matter data with a multi-feature selection approach. Heliyon. 2024 Apr 26; 10(9):e30319. doi:10.1016/j.heliyon.2024.e30319.

23. Immunization Center, Thailand Ministry of Public Health. Covid-19 vaccination infographic [Internet]. San Francisco (CA): Internet Archive; [cited 2024 Aug 30]. <https://web.archive.org/web/20210802001320 /https://dashboard-vaccine.moph.go.th/>

24. Institute for Population and Social Research. Mahidol population gazette: population of Thailand, 2020 (Vol.29: January 2020). Nakhon Pathom: Institute for Population and Social Research, Mahidol University; 2020 Mar 17 [cited 2024 Aug 12]. <https://ipsr.mahidol. ac.th/en/population-of-thailand-2020-vol-29-january-2020/>

25. Annuziato A, Asikainen T. Effective reproduction number estimation from data series. Luxembourg: Publications Office of the European Union; 2020 [cited 2024 Aug 30]; <https://publications.jrc.ec.europa.eu/repositor y/bitstream/JRC121343/r0_technical_note_v3. 4_1.pdf>

26. Xin H, Li Y, Wu P, Li Z, Lau EHY, Qin Y, et al. Estimating the latent period of coronavirus disease 2019 (COVID-19). Clin Infect Dis. 2022 May 3;74(9):1678–81. doi:10.1093/cid/ciab746.

27. Gregory CO, Hall A. CDC Yellow Book 2024: COVID-19 [Internet]. Atlanta: U.S. Centers for Disease Control and Prevention; [cited 2024 Aug 12]. <https://wwwnc.cdc.gov/travel/yellow book/2024/infections-diseases/covid-19#:~:text =People%20with%20COVID%2D19%20can>

28. Luo G, Zhang X, Zheng H, He D. Infection fatality ratio and case fatality ratio of COVID-19. Int J Infect Dis. 2021 Dec;113:43–6. doi:10.1016/j.ijid.2021.10.004.

29. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2016 [cited 2024 Aug 12]. <https://www.R-project.org/>

30. R Core Team. R: fit best ARIMA model to univariate time series [Internet]. Vienna: R Foundation for Statistical Computing; [cited 2024 Aug 12]. <https://search.r-project.org/CR AN/refmans/forecast/html/auto.arima.html>

31. Xu L, Magar R, Barati Farimani A. Forecasting COVID-19 new cases using deep learning methods. Computers in Biology and Medicine. 2022 May;144:105342. doi:10.1016/j.compbio med.2022.105342.

32. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting: methods and applications. Journal of the Operational Research Society. 1984;35(1):79. doi:10.1057/jors.1984.11.

33. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. 3rd ed. Melbourne: Otexts; 2021.

34. Chharia A, Jeevan G, Jha RA, Liu M, Berman JM, Glorioso C. Accuracy of US CDC COVID-19 forecasting models. Front Public Health. 2024 Jun 26;12:1359368. doi:10.3389/fpubh.2024.1359368.

35. Ahmad G, Ahmed F, Rizwan MS, Muhammad J, Fatima SH, Ikram A, et al. Evaluating data-driven methods for short-term forecasts of cumulative SARS-CoV2 cases. PLoS One. 2021 May 21;16(5):e0252147. doi:10.1371/journal. pone.0252147.

36. Rahman MS, Chowdhury AH, Amrin M. Accuracy comparison of ARIMA and XGBoost forecasting models in predicting the incidence of COVID-19 in Bangladesh. PLOS Glob Public Health. 2022 May 18;2(5):e0000495. doi:10.1371/journal.pgph.0000495.

37. Rguibi MA, Moussa N, Madani A, Aaroud A, Zine-dine K. Forecasting Covid-19 Transmission with ARIMA and LSTM Techniques in Morocco. SN Comput Sci. 2022;3(2):133. doi:10.1007/s42979-022-01019-x.

38. Zhao D, Zhang H, Cao Q, Wang Z, He S, Zhou M, et al. The research of ARIMA, GM(1,1), and LSTM models for prediction of TB cases in China. PLoS One. 2022 Feb 23;17(2):e0262734. doi:10.1371/journal.pone.0262734.

39. Vynnycky E, White RG, Fine PEM. An introduction to infectious disease modelling. New York: Oxford University Press; 2011.

40. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton (NJ): Princeton University Press; 2008.