Real-time prediction for multi-wave COVID-19 outbreaks

Faihatuz Zuhairoh[ab], Dedi Rosadi[1, a]

[a]Department of Mathematics, Gadjah Mada University, Indonesia;
[b]Study Program of Mathematics Education, STKIP YPUP Makassar, Indonesia

Abstract

Intervention measures have been implemented worldwide to reduce the spread of the COVID-19 outbreak. The COVID-19 outbreak has occured in several waves of infection, so this paper is divided into three groups, namely those countries who have passed the pandemic period, those countries who are still experiencing a single-wave pandemic, and those countries who are experiencing a multi-wave pandemic. The purpose of this study is to develop a multi-wave Richards model with several changepoint detection methods so as to obtain more accurate prediction results, especially for the multi-wave case. We investigated epidemiological trends in different countries from January 2020 to October 2021 to determine the temporal changes during the epidemic with respect to the intervention strategy used. In this article, we adjust the daily cumulative epidemiological data for COVID-19 using the logistic growth model and the multi-wave Richards curve development model. The changepoint detection methods used include the interpolation method, the Pruned Exact Linear Time (PELT) method, and the Binary Segmentation (BS) method. The results of the analysis using 9 countries show that the Richards model development can be used to analyze multi-wave data using changepoint detection so that the initial data used for prediction on the last wave can be determined precisely. The changepoint used is the coincident changepoint generated by the PELT and BS methods. The interpolation method is only used to find out how many pandemic waves have occurred in given a country. Several waves have been identified and can better describe the data. Our results can find the peak of the pandemic and when it will end in each country, both for a single-wave pandemic and a multi-wave pandemic.

Keywords: prediction, changepoint, Richards model, multi-wave, PELT, binary segmentation, COVID-19

1. Introduction

The COVID-19 pandemic is still ongoing towards the end of 2021, and there are several countries that still have a high number of cases every day, while, at the same time, there are also several countries that have passed the pandemic period. In this paper, countries in the world will be divided into three major groups, namely countries that have finished going through a pandemic period, countries that are still experiencing one wave of the pandemic, and finally, countries that have experienced a multi-wave pandemic. Three countries were chosen to represent each group. The first group is represented by China, Paraguay, and Taiwan. The second group is represented by Vietnam, Cuba, and Thailand. The third group is represented by Indonesia, Australia, and the Netherlands.

In mid-October 2021, the cumulative cases of COVID-19 in the world reached 240 million cases, recovered cases reached 219 million and there were 4.9 million deaths. The highest number of cases

---

occurred in 3 countries, namely the USA, India, and Brazil with total cases reaching in the tens of millions (Worldometer, 2021). Prior to the COVID-19 pandemic, there were many illness outbreaks with significant fatality rates, including H1N1, HIV/AIDS, SARS, and others. The COVID-19 pandemic is the third coronavirus epidemic to emerge; previous outbreaks included SARS in 2002 and MERS in 2012, although neither outbreak approached the global scope of COVID-19. Both SARS and MERS had a greater case fatality rate (CFR) than COVID-19. This shows that more people died as a result of an infection (Feehan and Apostolopoulos, 2021).

The first case of COVID-19 was discovered in early December 2019 in Wuhan, China, and by January it had spread globally (Roberts *et al.*, 2021). Various strategies are used to reduce the daily increase in positive cases, such as the lockdown policy carried out by the Netherlands, Thailand, and China where human mobility has almost returned to normal after the lockdown was relaxed (Meier *et al.*, 2020; Haddawy *et al.*, 2021). Meanwhile, the Indonesian government has established the Implementation of Restrictions on Community Emergency Activities (PPKM Emergency) (Zuhairoh and Rosadi, 2020; Darmawan *et al.*, 2022). Several studies on vaccination as an effort to reduce the number of positive cases were also conducted in Australia by Edwards *et al.* (2021) which indicated that alternative policy measures may be needed to achieve sufficient vaccination coverage to end the pandemic. Vaccine policies were also implemented in Vietnam and Taiwan although communication programs to decrease risk perception and increase awareness about vaccines should be developed to facilitate vaccine acceptance (Nguyen *et al.*, 2021; Lo *et al.*, 2021). The impact of the COVID-19 pandemic has led the Paraguayan government to develop programs to provide mental health care and services for people facing self-isolation during a prolonged quarantine period (Kim *et al.*, 2021) while the Cuban government discusses the ongoing response to the COVID-19 pandemic in the context of the medical system in Cuba, its health tourism and diplomatic problems (Wylie, 2021). In addition, the previous COVID-19 research that we have done is about determining the basic reproduction number and the multi-state discrete-time Markov chain SVIRS model (Zuhairoh *et al.*, 2021, 2022EL).

The logistic growth model and Richards curve model have been used previously by Wang *et al.* (2016) for studies of SARS and HIV, and real-time influenza pandemic forecasts by Nishiura and Chowell (2009). As for changepoint detection, there are many methods used including generalized likelihood ratio, Bayesian methods, Segment Neighbourhood and others (Eckley *et al.*, 2011). In addition there are Pruned Exact Linear Time (PELT) methods and Binary Segmentation (BS) by Killick and Eckley (2014) and James and Matteson (2014) whose focus is on applications where the number of chagepoints will increase as more data is collected. Research on multiple changepoints in previous time series data has also been carried out by Ma *et al.* (2020).

In our previous study on the prediction of COVID-19 in South Sulawesi, Indonesia and several countries in Southeast Asia (Zuhairoh and Rosadi, 2020, 2022), we were able to make predictions using the Richards curve model, although the Richards curve model results could not be used if a multi-wave infection occurs in a country or region. Thus, we use the interpolation method to determine how many pandemic waves occur in a country and then determine changepoint by comparing the PELT method and the BS method. After that, we used the parameter estimation results with the logistic growth model as the initial assumption to estimate the parameter values of the Richards curve model, and this method resulted resulted in better parameter estimates for predicting COVID-19 cases. The development of the multi-wave Richards curve model is expected to be a solution to obtain better predictive results for cases of multi-wave pandemics. The previous multistage Richards model was also further developed by Hsieh and Cheng (2006) whereby stages are distinguished by turning points (or inflection points) that denote acceleration after deceleration at the end of each S-shaped segment, which is the local minimum of the corresponding incidence curves. In this paper, however, we use a

different changepoint detection method.

If a country experiences a multi-wave pandemic, then data on the last wave is used to predict when the pandemic will end. The best model is then used, for predicting the countries that are still at the peak of the pandemic, to provide information on the peak of the pandemic and when the pandemic will end based on the maximum number of cases so that it can be used as input for dealing with the COVID-19 pandemic in their respective countries.

This research aims to predict the peak of the pandemic and when it will end. Our previous research used the Richards curve model to make predictions and obtain accurate results. However, using this model on data with multiple-waves produced less accurate results, so we developed a Richards curve model that first detected changepoints to determine the beginning of the last wave. We used the detection results to make predictions using the Richards curve model. This study used three changepoint detection methods, namely PELT, BS, and interpolation. The difference from previous studies, which is the contribution of this study, is to add a changepoint detection method in the Richards curve model to predict the spread of infectious diseases. This is necessary in order to overcome the weakness of the Richards curve model, which can only predict single-wave cases.

In Section 1 of this article, we describe several studies on COVID-19 predictions that have not yet been completed as for the end of 2021 and then explain the novelty of our research that includes the use of several changepoint detection methods to make predictions in the Richards curve model in multi-wave cases. In Section 2, we cite the data used and their sources and then arrange a Richards curve model algorithm for the multi-wave case. In Section 3, we present the detection results of changepoints and the results of predicting cases of COVID-19 in several countries using images. In Section 4, we describe the results by interpreting the images obtained in the results section, provide predictions of when the peak and the end of the COVID-19 pandemic will be reached in each of the sample countries, and compare the predicted results with the Richards and LGM curves. Finally, Section 5 contains the conclusions.

2. Materials and methods

2.1. Data

This study uses COVID-19 data sourced from Ritchie *et al.* (2021) which provides daily confirmed COVID-19 patient data from all countries in the world. The countries used in this study were divided into three groups, namely (1) countries that had finished going through a pandemic period represented by China, Paraguay, and Taiwan, (2) countries that were still experiencing a pandemic wave represented by Vietnam, Cuba, and Thailand, and (3) countries experiencing a multi-wave pandemic represented by Indonesia, Australia, and the Netherlands. We used data starting from the first case of COVID-19 entering each country and tracked the number of cases until October 12, 2021. In this paper, it is defined that a country that has passed the pandemic period is a country where in the last 1 month the number of positive confirmed cases of COVID-19 is less than 100 people.

Countries in the first group, namely the group of countries that passed the pandemic period, were analyzed using data from the start of the last wave to just before the peak of the pandemic to predict when the pandemic would end, and then this data was compared with the real data available until October 12, 2021. Countries in the second group, namely the countries that were still experiencing one wave of the pandemic, were analyzed using data collected until October 12, 2021, then predictions were made on when the pandemic will peak and when the pandemic will end in each country. Meanwhile, countries in the third group, namely countries experiencing multi-wave pandemics, were analyzed using an improvised version of the Richards curve model to predict when the pandemic

would end. The data used starts from the beginning of the last wave until October 12, 2021.

2.2. Models

The Logistic Growth Model (LGM) and the Richards curve model were used in this work as phenomenological models. These two models have been frequently employed to simulate epidemics of different infectious diseases. The Richards curve model is a more detailed version of the LGM (Maleki, 2020a).

Verhulst initially used the LGM to estimate population increase in 1838 (Hsieh, 2009). The differential equation of the LGM model is

$$C'(t) = rC(t)\left[1 - \frac{C(t)}{K}\right] \tag{2.1}$$

of which the solution is:

$$C(t) = \frac{K}{1 + be^{-rt}}, \tag{2.2}$$

where $C(t)$ denotes the number of cumulative cases at time t, K denotes the expected final epidemic size (total number of cases), t denotes time, b is a constant and r is the per capita growth rate.

To obtain better prediction results, the LGM model equation given in equation (2.1) needs to be modified by adding parameter a, which measures the deviation from the simple symmetric logistic curve (Roosa *et al.*, 2020) so that equation (2.1) can be written as equation (2.3) known as the Richards curve model,

$$C'(t) = rC(t)\left[1 - \left(\frac{C(t)}{K}\right)^a\right], \tag{2.3}$$

where the analytic solution of the Richards curve model is (Hsieh, 2009),

$$C(t) = \frac{K}{\left[1 + e^{-r(t-t_m)}\right]^{\frac{1}{a}}}, \qquad t_m = t_i + \frac{\ln a}{r}, \tag{2.4}$$

where $C(t)$ is the number of cumulative cases at time t, K is the total case number of the outbreak, t is time, r is the growth rate of the infected population, a is the exponent of deviation from the standard logistic curve, and t_i is the inflection point. The Richards curve parameter can be determined using the data of patients who tested positive for COVID-19. The value of K in each country was the cumulative number of cases in that country.

Countries experiencing a multi-wave pandemic can use the multi-wave Richards curve model. The steps for developing the multi-wave Richards curve model are given as follows,

1. Make a data plot of positive confirmed COVID-19 cases for each country. For outbreaks with a single-wave pandemic, it is possible to estimate parameters directly (K, r, a, and t_m).

2. Use the interpolation method to see how many waves of the pandemic occurred. The interpolation method uses a gradient between adjacent interpolation points. To detect a turning point, we first define a gradient (G_i) between adjacent interpolated points. This gradient is then used to assess the presence or absence of a turning point. This paper uses the two criteria defined in Guven's study (Guven *et al.*, 2014). The first criterion determines whether the gradient changes sign such

that there is a local/absolute minimum or maximum value. While the second criterion is needed to ensure that the location of this turn is correctly recognized even when the slope does not change sign. The two criteria can be written as follows,

$$\text{Condition } 1 \rightarrow G_{i-1} > 0 \ \& \ G_i < 0 \ \| \ G_{i-1} < 0 \ \& \ G_i > 0 \tag{2.5}$$

$$\text{Condition } 2 \rightarrow \frac{3}{4} * |G_{i-1}| > |G_i| \ \| \ \frac{3}{4} * |G_i| > |G_{i-1}| \tag{2.6}$$

3. Find the inflection point, t_{min} that separates the two epidemic waves using the PELT and BS methods. The PELT method's main assumption governing computational time is that the number of inflection points increases linearly as the data set increases, which indicates that the changepoints are spread across the data rather than being confined to a single region. The PELT method algorithm developed by Killick *et al.* (2012) is as **Algorithm 1**.

Algorithm 1: Pruned Exact Linear Time (PELT) method.

Input   :

A set of data of the form, $(x_1, x_2, \dots, x_n)$ where $x_i \in \mathbb{R}$.

A measure of fit $C(\cdot)$ dependent on the data.

A penalty constant β which does not depend on the number or location of changepoints.

A constant K that satisfies equation $C(x_{(t+1):s}) + C(x_{(s+1):T}) + K \leq C(x_{(t+1):T})$.

Initialise: Let n = length of data and set $F(0) = -\beta$, $cp(0) = NULL$. $R_1 = \{0\}$.
Iterate for $\tau^* = 1, 2, \dots, n$

1. Calculate $F(\tau^*) = \min_{\tau \in R_{\tau^*}} [F(\tau) + C(x_{(\tau+1):\tau^*}) + \beta]$.

2. Let $\tau^1 = \arg\{\min_{\tau \in R_{\tau^*}} [F(\tau) + C(x_{(\tau+1):\tau^*}) + \beta]\}$.

3. Set $cp(\tau^*) = [cp(\tau^1), \tau^1]$.

4. Set $R_{\tau^*+1} = \{\tau \in R_{\tau^*} \cup \{\tau^*\} : F(\tau) + C(x_{(\tau+1):\tau^*}) + K \leq F(\tau^*)\}$.

Output : the changepoints recorded in $cp(n)$.

Where $C(\cdot)$ is the cost function for each data segment, β is a penalty constant that does not depend on the number and position of the inflection points, t is time, x_i is the number of positive cases on day i, n is the amount of data, τ is the place of the changepoint in the data set, and τ^* is iterations that occur at each data point.

While the BS method can be used to extend the single changepoint method to multiple changepoints, we start by applying the detection method to the entire data set. If no changepoint is detected, then the detection process stops, if not, then we divide the data into two segments, i.e., before and after the changepoint, and apply the detection method to each segment. This procedure is repeated until no further changepoints are detected. The BS method algorithm developed by Eckley *et al.* (2011) is as **Algorithm 2**.

Where $\Lambda(\cdot)$ is test statistics, $\hat{\tau}(\cdot)$ is the location estimator from the changepoint, C is the set of the detected changepoints, s is the set of data segments that need to be tested for the point of change, r is the position of the changepoint in the original data set, and x_i is the number of positive cases on day i.

4. Use data starting from $t_{min} + 1$, i.e., a day after the start of the last wave, then use the data to determine the initial assumptions in estimate the parameters of the Richards curve model, where these parameters are used to predict when the pandemic will peak and when the pandemic will end.

Algorithm 2: Binary Segmentation (BS) method.

Input :

    A set of data of the form, $(x_1, x_2, \ldots, x_n)$.

    A test statistic $\Lambda(\cdot)$ dependent on the data.

    An estimator of the changepoint position $\hat{\tau}(\cdot)$.

    A rejection threshold C.

Initialise: Let $C = \emptyset$, and $S = \{[1, n]\}$.

Iterate while $S \neq \emptyset$

1. Choose an element of S; denote this element as $[s, t]$.

2. If $\Lambda(x_{s:t}) < C$, remove $[s, t]$ from S.

3. If $\Lambda(x_{s:t}) \geq C$ then:

    (a) remove $[s, t]$ from S;

    (b) calculate $r = \hat{\tau}(x_{s:t}) + s - 1$, and add r to C;

    (c) if $r \neq s$ add $[s, r]$ to S;

    (d) if $r \neq t - 1$ add $[r + 1, t]$ to S.

Output : the set of changepoints recorded C.

---

3. Results

3.1. Changepoint detection

The first step is to draw the number of confirmed COVID-19 cases in each group of sampled countries to see the number of pandemic waves that occur. Figure 1 is based on a plot of data for each country of the number of positive confirmed cases. Figure 1 also shows which countries have gone through a pandemic period, which countries are still experiencing one pandemic wave, and which countries have experienced a multi-wave pandemic.

Figure 1 shows that China and Taiwan are countries that have gone through a pandemic period and have had only one pandemic wave, while Paraguay has also gone through a multi-wave pandemic. To verify this, Figure 2 is based on a data plot equipped with the PELT, BS, and interpolation methods. Similarly, the curves generated by the interpolation method for Vietnam, Cuba, and Thailand show that they are still experiencing one pandemic wave. Meanwhile, Indonesia, Australia and the Netherlands have experienced a multi-wave pandemic. After it is known that a country is experiencing a multi-wave pandemic, the next step is to determine the changepoint to find out the initial data used to make the predictions.

Countries that have gone through the pandemic period, shown in Figure 2, are known to have similarities with Paraguay and Taiwan, namely that they have experienced multi-waves, while China has only experienced a single-wave pandemic. For China, data from the beginning of the pandemic until February 22, 2020 was used. Meanwhile, for Paraguay and Taiwan, data starting from the detection of the first changepoints on the lines generated by the PELT and BS methods was used. From Figure 2, it can be seen that some changepoints generated from the PELT and BS methods are coincident and some are not. The PELT method detects more changepoints than the BS method, however, for predictions using the Richards curve model, better prediction results were obtained using the BS method. Figure 3 shows that the prediction results with the Richards curve model are closer to the actual data than the prediction results using the LGM.

Predictions for countries in the second group, namely countries that are still experiencing a single-
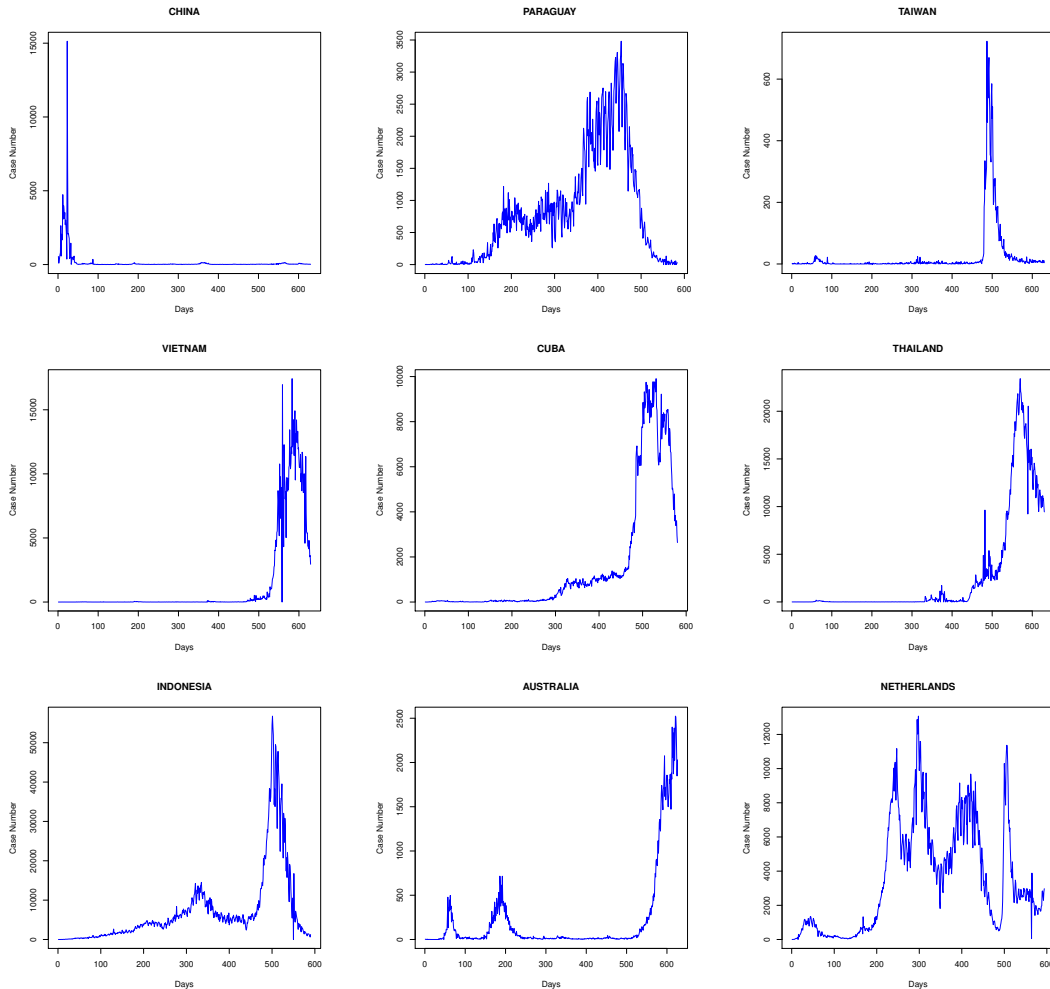
Figure 1: *Plot of confirmed COVID-19 cases.*

wave, used data from the beginning of the pandemic to October 12, 2021. As for countries experiencing a multi-wave pandemic, data on the last wave will be used. Australia and the Netherlands still show significant additions of positive cases every day while Indonesia is at the end of the second wave with additional cases of less than 1,000 cases per day.

3.2. Determination of parameters to make predictions with the Richards curve model

In the next step, the initial assumptions used in the LGM can be seen in Table 1. The initial assumptions were obtained from the data on COVID-19 cases in the countries in Figure 1 that have not passed the pandemic period, where K was the total number of COVID-19 19 cases, r was the ratio of the per capita growth rate of the population infected, and t_m was the reflection point of t_i obtained from equation (2.4). In determining the initial assumption value, only countries in Group 2 and 3 will be used because countries in Group 1 have passed the pandemic period. Countries in Group 2 use all the data
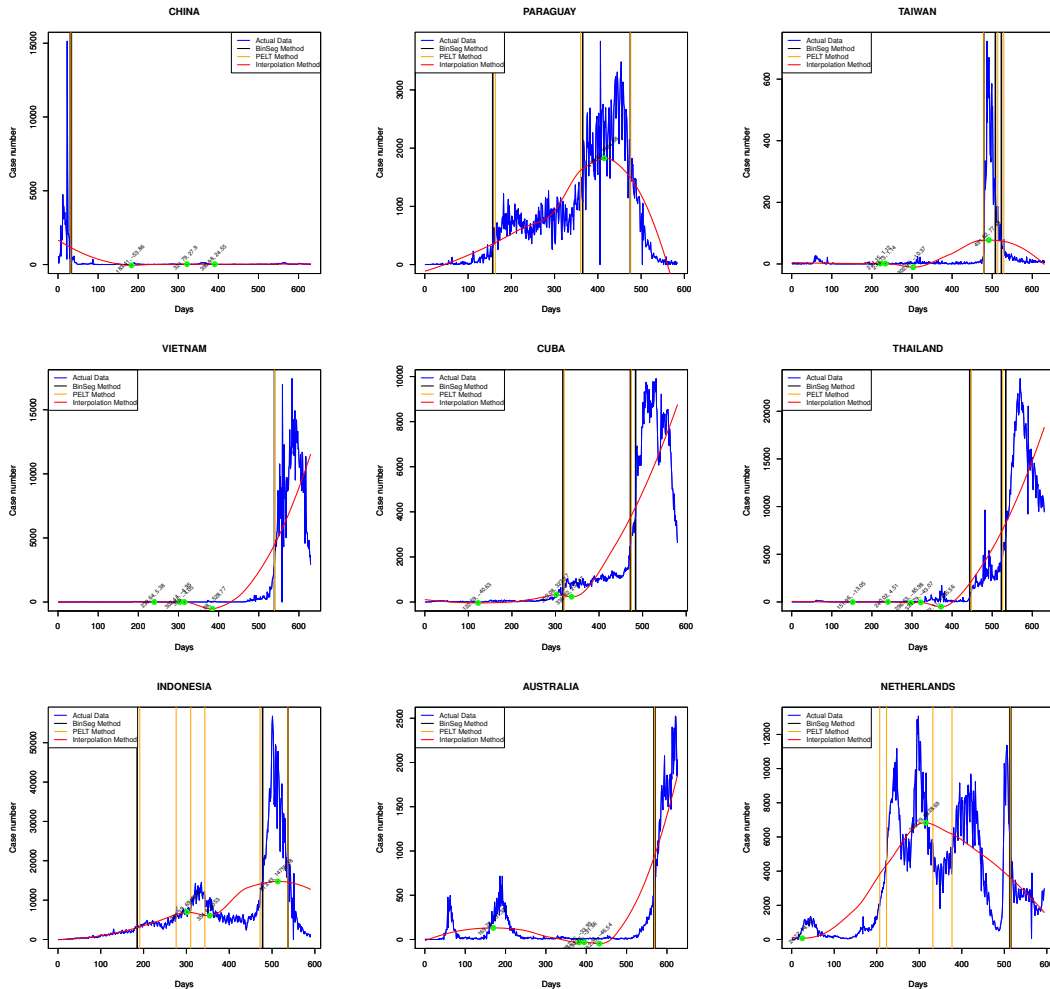
Figure 2: *Changepoint detection in each country.*

from the start of the pandemic to October 12, 2021, and countries in Group 3 use the data from the beginning of the last wave until October 2021.

Using these initial parameters from Table 1, we estimated the parameters using a non-linear least square approach. The curve's plot was based on the initial parameters and the optimal ones were based on the LGM, then the results of the parameter estimates from the LGM were used as the initial assumptions to create the Richards curve model as given in Figure 4. After determining the parameters of each model, predictions are made for the countries in Figure 4. The prediction results using the LGM and Richards curve models can be seen in Figure 5. By using the cumulative data up to October 12, 2021, and using the initial assumption value of each parameter from the LGM with the addition of variable a, which provides a measure of flexibility in the curvature of the S-shape indicated by the resulting solution curve, the prediction result curve is presented in Figure 5. In Figure 5, it can be seen that Indonesia has passed the peak of the pandemic while other countries are still experiencing a
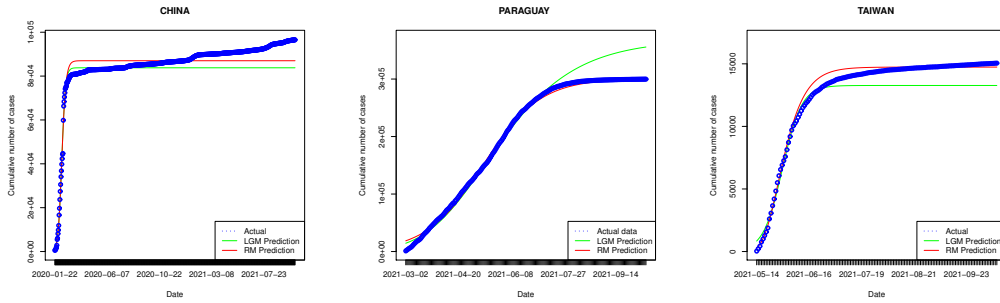
Figure 3: *Comparison of the LGM prediction results and the Richards curve model with actual data for countries that have passed the COVID-19 pandemic.*

Table 1: Initial assumptions on the parameters

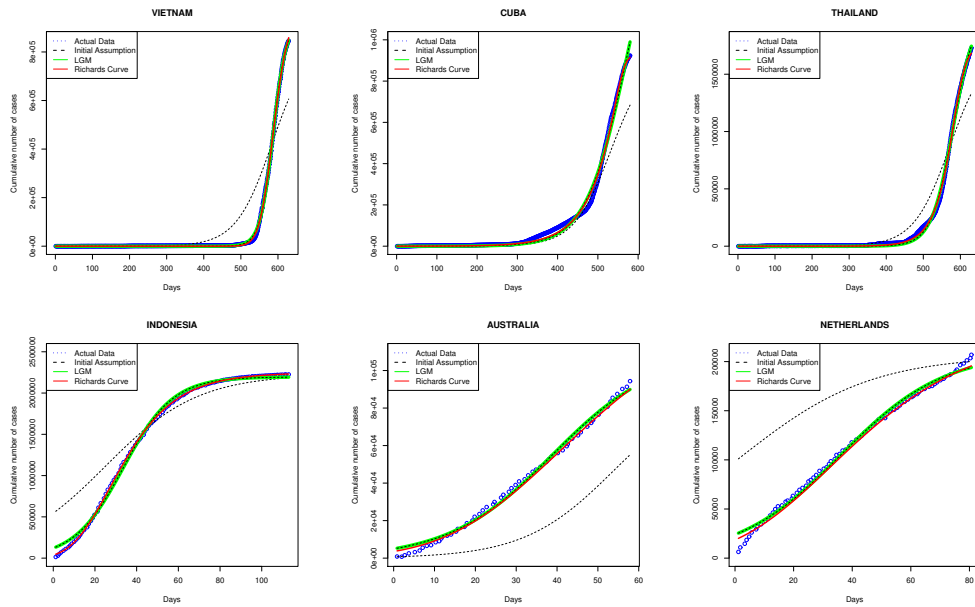| Country	K	r	t_m
Vietnam	846,230	0.0206	584
Cuba	923,966	0.0218	532
Thailand	1,730,364	0.0206	571
Indonesia	2,302,105	0.0458	31
Australia	93,829	0.0909	54
Netherlands	205,966	0.0450	2

Figure 4: *Initial assumptions to create the Richards curve model.*

significant increase in positive cases. Of the 6 countries modeled with the Richards curve model, the parameter estimation results were obtained using the non-linear approach as shown in Table 2.

The estimation results of the Richards curve parameters given in Table 2 will be substituted into Equation (2.4) to predict cases in the following several periods. The K-value in Table 2 shows the
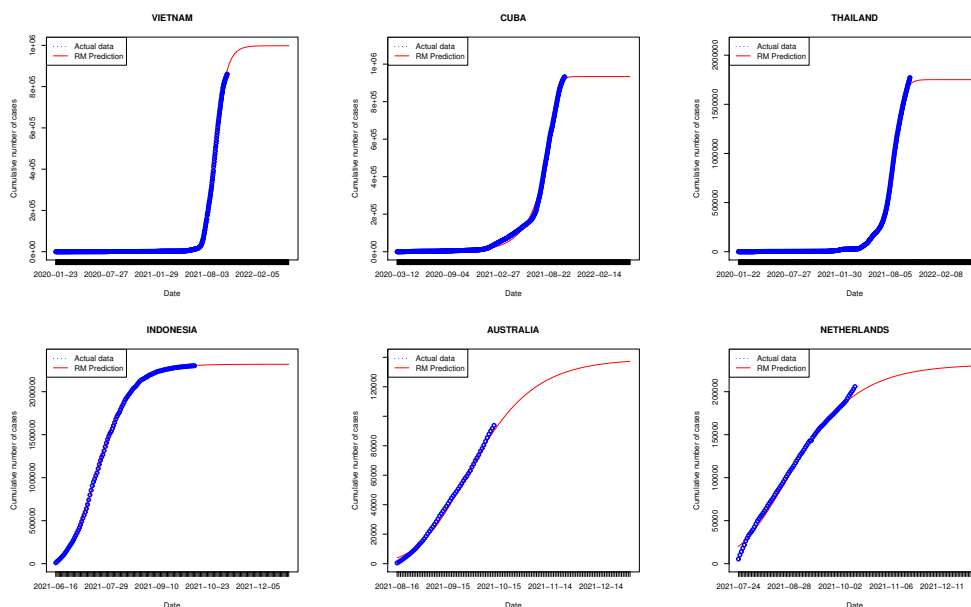
Figure 5: *Prediction results for countries that are still experiencing the COVID-19 pandemic.*

Table 2: The results of the estimated parameters from Richards curve models

Country	K	r	a	t_m
Vietnam	998,000	0.0428	0.5340	570.90
Cuba	933,600	0.1024	6.2080	560.20
Thailand	1,751,000	0.0660	2.5110	595.60
Indonesia	2,319,000	0.0600	0.2540	10.77
Australia	138,900	0.0456	0.2843	13.35
Netherlands	231,900	0.0366	0.2039	10.77

Table 3: Prediction results

Country	Pandemic peak	End of pandemic	Maximum number of cases
Vietnam	August 2021	March 2022	998,000
Cuba	September 2021	December 2021	933,600
Thailand	August 2021	January 2022	1,751,000
Indonesia	July 2021	December 2021	4,246,708
Australia	September 2021	February 2022	178,515
Netherlands	August 2021	February 2022	2,094,417

maximum number of COVID-19 cases in each country. Table 3 presents the prediction results, including when the pandemic peak will occur and when the pandemic will end along with the maximum number of cases at the peak of the pandemic in the countries.

4. Discussion

The COVID-19 pandemic, which has spread to all countries in the world, was grouped into three categories, namely countries that have passed the pandemic period, countries that are still experiencing a single-wave pandemics, and countries that are experiencing multi-wave pandemics. In this paper,

Table 4: AIC value for each model

Country	Logistic growth model	Richards curve model
Vietnam	12270.62	12049.23
Cuba	13094.48	12743.30
Thailand	14217.28	13802.51
Indonesia	2833.56	2521.85
Australia	1071.83	1060.25
Netherlands	1616.92	1601.27

three countries were used to represent each group. Countries in Group 1, China, Paraguay, and Taiwan, have gone through a pandemic, so for this study, the countries in Group 1 were used as a benchmark to determine which model, between the LGM and the Richards curve model, is the most effective for the actual COVID-19 data. Figure 3 shows that the Richards curve model gives predictions that are closer to the actual data than the LGM. This conclusion is reinforced by the AIC value of each model shown in Table 4.

Data on positive COVID-19 cases for countries that are still going through a pandemic wave, namely Vietnam, Cuba, and Thailand, is used to find the initial values that can be seen in Table 1. Meanwhile, for countries experiencing a multi-wave pandemic, it is necessary to use the PELT and BS method to determine the changepoint in order to make it easier to discern the data that will be used to make predictions on the last wave of the pandemic. The initial details of the last wave for countries experiencing multi-waves based on the changepoint in Figure 3 are Indonesia with 472 data starting on June 16, 2021, Australia with 569 data starting on August 16, 2021, and the Netherlands with 514 data starting on July 24, 2021.

The initial assumptions in determining the parameters for multi-wave countries are also shown in Table 1. Based on Figure 4, it can be seen that the LGM curve and Richards curve data plots have the same curve shape using the initial assumption values in Table 1. The data used to produce Figure 5 started with the detection of the first confirmed cases in each country and continued to the highest number of cases during the pandemic period.

After estimating the parameters with a non-linear least square approach, the actual data was included in the equations for each model, and their results can be seen in Figure 5. Prediction using the Richards curves model from the 6 countries showed better results than the predictions using the LGM. In the Richards formula, parameter a was added, where a is the magnitude of the deviation from the standard logistic curve. From the estimation results of the two types of models, which determines which model is suitable for each country, we needed to calculate the value of the AIC (Akaike's Information Criterion) below in Table 4.

Table 4 shows that the AIC value for the Richards curve was smaller compared to that for the LGM. To get a better prediction result, a model with a smaller AIC value was used (Bozdogan, 2000), so that the countries in Table 4 will only be predicted using the Richards curve model. The peak of the pandemic and when it will end is shown in Table 3.

Using the initial assumptions in Table 1, actual data, plots, and Richards curves can be obtained using data up to October 12, 2021, and the results can be seen in Figure 4. If the curve is not sloping, the spread of COVID-19 will continue increase. From the 6 data plots, it can be seen that all countries have passed the peak of the pandemic. For countries with multi-waves, the actual data in the last wave, which is marked by the changepoint from the PELT and BS methods, is used. From the results of the changepoint detection, the PELT method produces more changepoints than the BS method. The same changepoint generated by both methods is used in this study with the expectation of getting more accurate prediction results. The parameter estimation results, which are obtained by using the

Richards curve model, are used to make predictions as shown in Figure 5.

Table 3 shows that the peak of the pandemic in the 6 countries occurred between July−September 2021. The prediction results for the country with the highest number of confirmed positive COVID-19 cases was Indonesia, where 4.2 million people were infected and the country with the lowest number of confirmed positive COVID-19 cases was Australia. In addition, predictions of when the COVID-19 pandemic will end can be seen in Table 3. The country with the longest pandemic period is Vietnam, which was predicted to end in March 2022 and the countries that will soon experience an end to the pandemic are Cuba and Indonesia. The prediction results obtained will depend on the actual data used (Leon *et al.*, 2020). In addition, government intervention also influences when a pandemic in a country will end.

5. Conclusion

The development of the Richards curve model in this study can overcome the problem of predicting if an area is experiencing a multi-wave pandemic. The initial assumption value that uses a non-linear least squares approach in the LGM model can produce an estimate for the Richards curve model parameters, and thus, obtain better prediction results because it is close to the actual data and has a smaller AIC value. The steps for making predictions, which use the Richards curve model, begin with plotting data on the confirmed positive COVID-19 cases, and then looking at the data patterns using the interpolation method to find out if a country is still experiencing a single-wave or a multi-wave pandemic. If a multi-wave pandemic is detected, the PELT and BS methods are used to determine the changepoint so that the initial data can be obtained to make predictions. For more accurate results, the same changepoint results from both methods are used.

The prediction results for the country with the highest number of positive confirmed cases of COVID-19 is Indonesia, where 4.2 million people were infected and the country with the lowest number of positive confirmed COVID-19 cases is Australia. In addition, predictions of when the COVID-19 pandemic will end were also obtained through the Richards curve model. The country with the longest pandemic period is Vietnam, which was predicted to end in March 2022 and the countries where the pandemic period will end soon are Indonesia and Cuba. The prediction results obtained will depend on the actual data used in determining the initial assumptions. In addition, government intervention from each country also affects the rise and fall of positive confirmed COVID-19 cases, which has an impact on determining when an epidemic will end in a country.

Acknowledgement

References

Bozdogan H (2000). Akaike's information criterion and recent developments in information complexity, *Journal of Mathematical Psychology*, **44**, 62–91.

Darmawan G, Rosadi D, and Ruchjana BN (2022). Hybrid model of singular spectrum analysis and ARIMA for seasonal time series data, *CAUCHY–Jurnal Matematika Murni dan Aplikasi*, **7**, 302–315.

Eckley IA, Fearnhead P, and Killick R (2011). *Bayesian Time Series Models*, Cambridge: Cambridge University Press, New York, *Analysis of changepoint models. In D. Barber, A. Cemgil, and S. Chiappa (Eds.)*, 205–224,

Edwards B, Biddle N, Gray M, and Sollis K (2021). COVID-19 vaccine hesitancy and resistance: Correlates in a nationally representative longitudinal survey of the Australian population, *PLoS ONE*, **16**, e0248892.

Feehan J and Apostolopoulos V (2021). Is COVID-19 the worst pandemic?, *Maturitas*, **149**, 56–58.

Guven O, Eftekhar A, Kindt W, and Constandinou TG (2014). Realtime ECG baseline removal: An isoelectric point estimation approach, *IEEE 2014 Biomedical Circuits and Systems Conference, BioCAS 2014 - Proceedings*, Washington, D.C., 29–32.

Haddawy P, Lawpoolsri S, Sa-Ngamuang C, *et al.* (2021). Effects of COVID-19 government travel restrictions on mobility in a rural border area of Northern Thailand: A mobile phone tracking study, *PLoS ONE*, **16**, e0245842.

Hsieh YH (2009). Richards Model: A simple procedure for Real-time prediction of outbreak severity, *Modeling and Dynamics of Infectious Diseases*, (pp. 216–236), Co-Published with higher education press, Singapore.

Hsieh YH and Cheng YS (2006). Real-time forecast of multiphase outbreak, *Emerging Infectious Diseases*, **12**, 122–127.

James NA and Matteson DS (2014). ecp: An R package for nonparametric multiple changepoint analysis of multivariate data, *Journal of Statistical Software*, **62**, 1–25.

Killick R and Eckley IA (2014). changepoint: An R package for changepoint analysis, *Journal of Statistical Software*, **58**, 1–19.

Killick R, Fearnhead P, and Eckley IA (2012). Optimal detection of changepoints with a linear computational cost, *Journal of the American Statistical Association*, **107**, 1590–1598.

Kim JE, Lee JH, Kang YG, Lee SH, Shin H, Rönnebeck N, Rönnebeck R, and Nam EW (2021). Depression in public officials during the COVID-19 pandemic in Paraguay: a web-based study, *BMC Public Health*, **21**, 1835, Available from: https://doi.org/10.1186/s12889-021-11860-z

León UAP, Ángel GCP, and Avila-Vales E (2020). A data driven analysis and forecast of an SEIARD epidemic model for COVID-19 in Mexico, *Big Data and Information Analytics*, **5**, 14–28.

Lo SY, Li SCS, and Wu TY (2021). Exploring psychological factors for COVID-19 vaccination intention in Taiwan, *Vaccines*, **9**, 764.

Ma L, Grant AJ, and Sofronov G (2020). Multiple changepoint detection and validation in autoregressive time series data, *Statistical Papers*, **61**, 1507–1528.

Maleki M, Mahmoudi MR, Wraith D, and Pho KH (2020a). Time series modelling to forecast the confirmed and recovered cases of COVID-19, *Travel Medicine and Infectious Disease*, **37**, 101742.

Meier K, Glatz T, Guijt MC, *et al.* (2020). Public perspectives on protective measures during the COVID-19 pandemic in the Netherlands, Germany and Italy: A survey study, *PLoS ONE*, **15**, e0236917.

Nguyen LH, Hoang MT, Nguyen LD, *et al.* (2021). Acceptance and willingness to pay for COVID-19 vaccines among pregnant women in Vietnam, *Tropical Medicine and International Health*, **26**, 1303–1313.

Nishiura H and Chowell G (2009). The effective reproduction number as a prelude to statistical

estimation of time-dependent epidemic trends, *Mathematical and Statistical Estimation Approaches in Epidemiology*, (pp. 103–121), Springer, Netherlands.

Ritchie H, Ortiz-Ospina E, Beltekian D, *et al.* (2021). Coronavirus pandemic (COVID-19), *Our World in Data*, Available from: https://ourworldindata.org/coronavirus.

Roberts DL, Rossman JS, and Jarić I (2021). Dating first cases of COVID-19, *PLOS Pathogens*, **17**, e1009620, Available from: https://dx.plos.org/10.1371/journal.ppat.1009620

Roosa K, Lee Y, Luo R, *et al.* (2020). Short-term forecasts of the COVID-19 epidemic in Guangdong and Zhejiang, *Journal of Clinical Medicine*, **9**, 596.

Wang X, Liu S, and Huang Y (2016). A study on the rapid parameter estimation and the grey prediction in Richards Model, *Journal of Systems Science and Information*, **4**, 223–234.

Worldometer (2021). Coronavirus update (Live): cases and deaths from COVID-19 virus pandemic, *Worldometers*, (pp. 1), Worldometer, Available from: https://www.worldometers.info/coronavirus/

Wylie LL (2021). Cuba's response to COVID-19: lessons for the future, *Journal of Tourism Futures*, **7**, 356–363.

Zuhairoh F and Rosadi D (2020). Real-time forecasting of the COVID-19 epidemic using the Richards model in South Sulawesi, Indonesia, *Indonesian Journal of Science and Technology*, **5**, 456–462.

Zuhairoh F and Rosadi D (2022). Data-driven analysis and prediction of COVID-19 infection in Southeast Asia using a Phenomenological Model, *Pakistan Journal of Statistics and Operation Research*, **18**, 59–69.

Zuhairoh F, Rosadi D, and Effendie AR (2021). Determination of basic reproduction numbers using transition intensities multi-state SIRD Model for COVID-19 in Indonesia, *Journal of Physics: Conference Series*, **1821**, 012050.

Zuhairoh F, Rosadi D, and Effendie AR (2022EL). Multi-state discrete-time Markov Chain SVIRS Model on the spread of COVID-19, *Engineering Letters*, **30**, 598–608.