Proceeding
UKI Press
e-ISSN: ……………………….

Covid-19 Prediction for Raising People Awareness

Leonard Lisapaly
Postgraduate Program, Universitas Kristen Indonesia - Jakarta, Indonesia
leonard.lisapaly@uki.ac.id

Abstract

During the development of the Covid-19 case in Indonesia, an attempt to predict Covid-19 numbers using the normal distribution has been made. The normal (Gaussian) distribution was used because it can mimic the development of Covid-19: a rather flat curve at the beginning of the spread, followed by a rapid increase of the spread, and then followed by the rapid decrease of the spread caused by various human efforts implemented to reduce the spread of the virus, and then finally a rather flat long tail representing the end of the spread. Starting from a model with one normal distribution at the beginning of the spread, we had predicted the Covid-19 case using the stack of multiple normal distributions. This model provided a better prediction than the prediction obtained from a single normal distribution. The method has accurately predicted, for example, a long Covid-19 wave that occurred in Jakarta last year.

Keywords: Covid-19, prediction, Monte-Carlo simulation

Introduction

The year 2020 is marked by the spreading of the Covid-19 virus in many countries globally, starting from the city of Wuhan in China. Indonesia is one of the countries affected by the spreading of this virus. It caused the government of Indonesia to put many resources to cope with the difficulties it caused to the people living in Indonesia. The virus was first observed on 2nd March 2020 when two women were hospitalized, and the first Covid-19 positive case was officially reported. Since then, the number of people infected, positively tested, and hospitalized has increased. In Indonesia, the number of people infected by the virus has been more than 1.6 million cases when this work was presented. With more than 45 thousand people died due to Covid-19, the death-to-total ratio in Indonesia is approximately 2.7%, which is still higher than the global death-to-total ratio of 2.1%. Indonesia is the country with the largest number of Covid-19 patients and the highest death number in Southeast Asia.

Talking about the impact of Covid-19, it has a huge impact on the world. Suppose one try to search for the impact of Covid-19 in Google. In that case, they will find that Covid-19 has affected every aspect of human life in the world: domestic and global economy, education, business, environment, physical and mental health, lifestyle, and many others. Google provided 5.6 billion results when one tries to do the search using the words “impact of covid 19”, explaining the strong impact of Covid-19 on the world.

To illustrate the impact of Covid-19, let us take the figure of global electricity consumption during Covid-19 in 2020 (Figure 1). Recently, International Energy Agency published reports that include, for example, the electricity consumption in 2020 (IEA Report, 2020). We can see that in 2020 the consumption of electricity decreased 5 -20 percent, especially during the lockdown periods at the beginning of the virus spread in various countries (dashed line). There was about 10% decrease in electricity consumption in Indonesia, with the increase in medical centers and residential areas.

 e-ISSN: ……………………….

So, consumption has shifted from industries and offices to houses. Covid-19 has helped reduce industrial carbon emission, but in residential areas, it increases as people cook more as the consequence of working from home and school from home programs. The good thing is that the electricity production remains flat or constant, which means not affected.

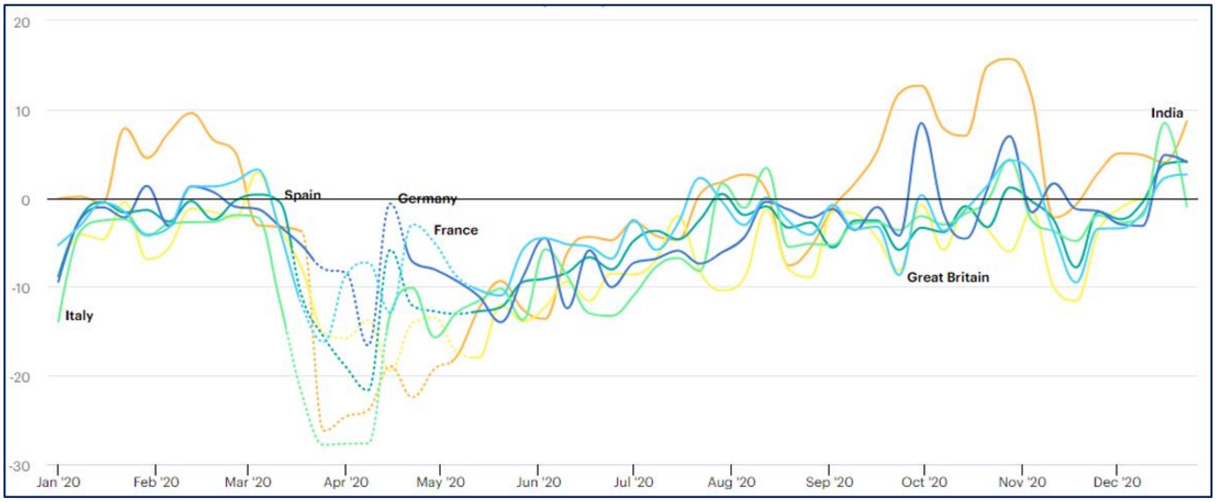


Figure 1. Electricity consumption in several countries during the Covid-19 period in 2020 (*Source: IEA*).

Covid-19 has the most significant effect on human health, especially during the rapid increase of the Covid-19 spread. Therefore, proper planning of the medical resource distribution is needed: number of hospitals, doctors, nurses, and medicines. To help to have good planning, we predict Covid-19 development. There are several benefits of making Covid-19 prediction: (1) To simulate what will happen in the future based on the analysis of existing data; (2) To plan actions to deal with the predicted problems; (3) To properly distribute medical resources like medical resources; and (4) To help in making decisions. This study aims to provide the government or authority with some predictions or indications, so the government or authority can react to Covid-19 development properly.

Research Methods

To model the development of Covid-19, we used a single normal (Gaussian) distribution before being replaced by a stack or sum of several normal distributions. The normal (Gaussian) was used because it can describe the development of Covid-19 in several stages: First, a rather flat curve at the beginning of the spread representing early Covid-19 spread. It is followed by a rapid increase of the spread among people who are not well prepared for the Covid-19 infection. When people start to be aware of the danger of Covid-19, they react to avoid or prevent the virus from spread using various methods such as social distancing, wearing the mask, washing hands, avoiding the crowd, limiting their travelling, and implementing lockdown. Currently, vaccination has been added as one of the methods to fight the virus. As a result, a rapid decrease in the spread will occur. Thus, it can be seen as a steep decreasing curve of normal distribution. The final stage is seen as a slow decrease of virus spread showing that people are still aware of the virus existence, but they start to be involved in their normal activity, causing the low amount of new Covid-19 cases. The normal distribution is shown in Figure 2.

e-ISSN: ……………………….

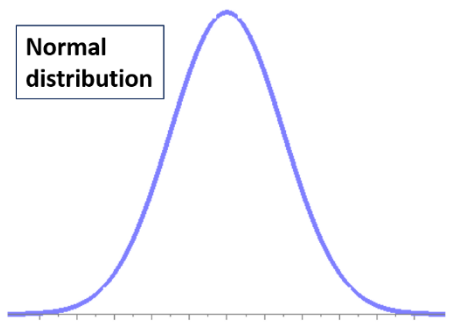


Figure 2. A normal distribution curve

A normal distribution can be written mathematically as (see for example Walpole *et al*., 2012):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

where σ is standard deviation, and μ is the mean or expectation.

To model the Covid-19 spread, the author has introduced a constant *A* that represents the strength of Covid-19. Therefore, replacing $f(x)$ by $c(t)$, the equation used for Covid-19 prediction as a function of time can be written as:

$$c(t) = \frac{A}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-T}{\sigma}\right)^2}$$

where

$c(t)$ = Covid-19 daily case as function of time
t = time
A = Covid-19 strength or Covid-19 amplitude
σ = a standard deviation representing the “duration” of Covid-19
T = Covid-19 time lag, representing the time of Covid-19 peak

It is not easy to find the shape of a normal distribution in the observed data from many countries in the world. The left side of the normal distribution may be observed at the beginning of the spread. However, later it changes because of the contribution of various areas where Covid-19 started to infect people. However, a nice example of the shape of the normal distribution can be observed in Figure 4, showing the daily Covid-19 curve in Britain after the vaccination began in December 2020. Figure 4a shows the daily Covid-19 numbers, and Figure 4b shows the daily death numbers. Combined with some lockdowns and any other forms of prevention, Covid-19 has passed its peak in January 2021 and started to diminish in Britain.

A single normal distribution in this study happened only at the beginning of the Covid-19 spread in Indonesia. Following series of accurate prediction, Lisapaly (2020a) and Lisapaly (2020b) initially predicted that the peak of Covid-19 in Indonesia would be on 24th April 2020 based on the use of single normal distribution. The method also predicted increased active cases, predicting the need for medical resources between the first and third weeks of May 2020, which became a reality. Using the method, we predicted a total positive Covid-19 number at the end of

 e-ISSN: ………………………..

April was eight days before to be 10,003 – 10,280 while the actual number was 10,118, showing a good agreement.

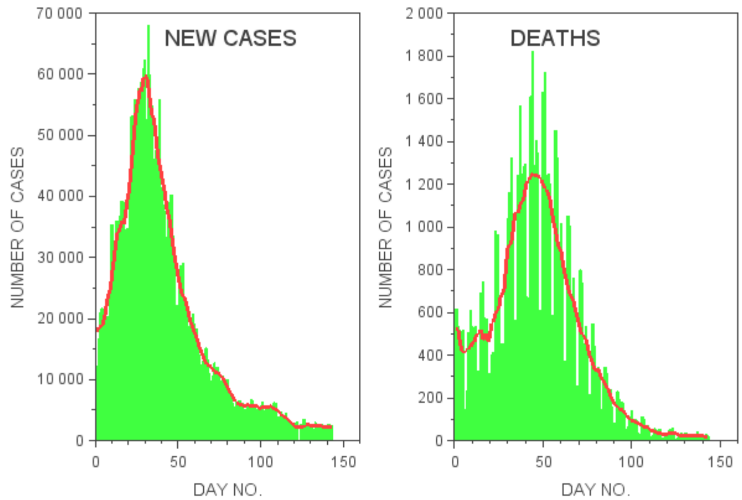


Figure 3. Covid-19 daily case in Britain after the vaccination started at the beginning of December 2020. Figure (a) shows the daily Covid-19 numbers, and Figure (b) shows the daily death numbers.

Soon, a single normal distribution proved to be less accurate after two months of Covid-19 spread in Indonesia. An analysis using digital signal processing that utilized the use of Fourier transformation has shown that the approach for Covid-19 prediction should not be a simple normal distribution (Lisapaly, 2020c). Therefore, the author modified the equation to involve more normal distributions. The assumption is that Covid-19 may occur in different areas in Indonesia with different strength, different duration, and different peak time. The concept is explained in Figure 4, and Java island in Indonesia illustrates the example. Let us assume that the first Covid-19 spread occurred in Jakarta, and it is normally distributed with certain statistical parameters described above. After several days, one infected person travels to Bandung and cause the development of a new normal distribution with different strength and probably different duration. Following that, infected persons travel to Semarang and Surabaya and build new virus centers there. Individually, all these Covid-19 spreads may look like a normal distribution. However, if statistics for Java island needs to be calculated, we can obtain the total curve by simply collecting all distribution and summing them up. The resulting curve does not look like a simple distribution, and multiple normal distributions can be justified.

 e-ISSN: ……………………….

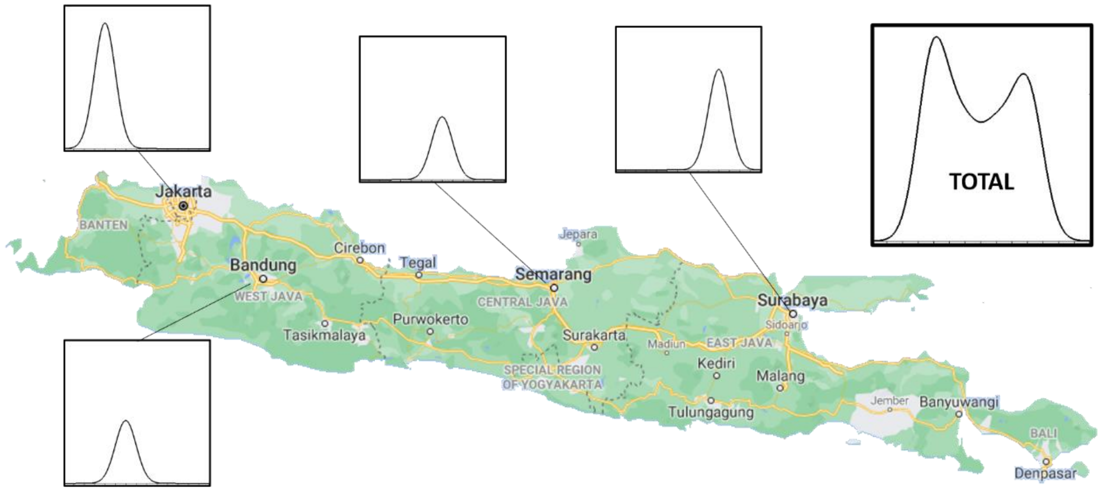


Figure 4. A simple illustration shows how different Covid-19 distribution centers with different Covid-19 strength, duration, and time lag will contribute to a Covid-19 curve of arbitrary shape.

The modified equation for Covid-19 now becomes

$$c(t) = \sum_{i=1}^{N} \frac{A_i}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-T_i}{\sigma_i}\right)^2}$$

where N is the number of normal distributions representing the number of the virus distribution center. It is important to notice that several adjacent spread centers may construct a pseudo normal distribution. In other words, they are difficult to be distinguished.

To do the prediction, all parameters A, σ, and T_i need to be determined. Because the equation is not a simple linear equation, a linear regression cannot solve the parameters. Instead, a Monte-Carlo simulation was used to model the Covid-19 synthetic response as a function of the parameters and time. The Monte-Carlo simulation is a mathematical technique that is used to estimate the possible outcomes of an uncertain event. The number of the normal distributions increase from time to time based on the progress of the Covid-19.

Monte-Carlo simulation has been used for analyzing uncertainties in various problems. For example, it has been used in structural engineering for stability analysis, as described in Lin and Su (2021), Jahani *et al*. (2013), and Papadrakakis *et al*. (1995).

The parameters are randomly determined in the Monte-Carlo simulation, and the Covid-19 response is calculated based on those parameters. Given the calculated response, the error between the observed data and the calculated data can be determined as a least-squares error. After a certain number of iterations, the parameters that provided the smallest error was accepted as an initial model. Monte-Carlo simulation is good to avoid the local minima in a serious non-linear case like the one found in this study. The longer the spread, the more the normal distributions used in this study.

Further parameters improvement was made using the Marquardt-Levenberg method described by Smith and Shanno (1971) and Leite and Leao (1985). This method was used here because of the non-linear behavior of the problem. First, the initial model generated from the previous Monte-

 e-ISSN: ……………………….

Carlo simulation was used to calculate the synthetic response and the error was generated. The error was then used to correct the initial model using the Marquardt-Levenberg method. Next, a new parameters set was calculated by summing the initial model and the correction and setting the results to be a new initial model. It is done iteratively until an error cutoff was reached, which was set to be 0.001.

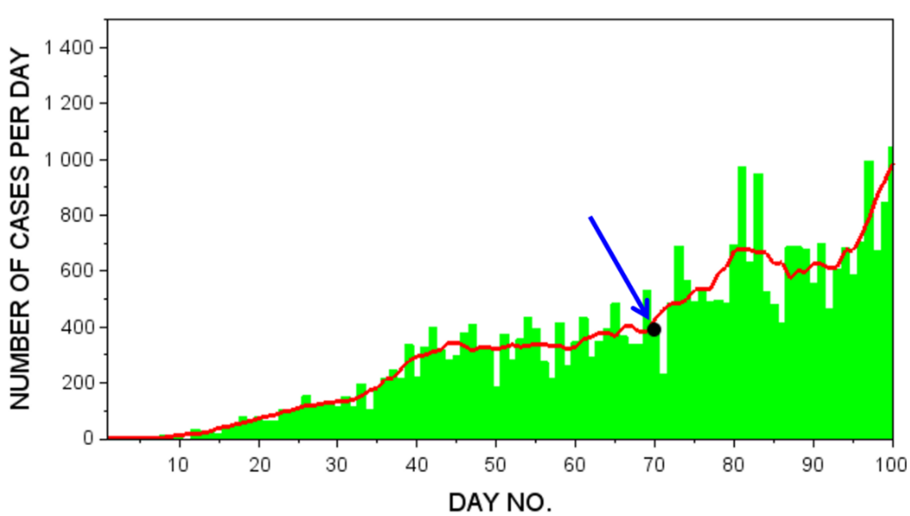
Results and Discussion

The work done here was started at the end of March 2020. As discussed in the earlier part of this article, a single normal distribution model was used to predict the Covid-19 data when Covid-19 spread just occurred and was applied until the end of April 2020 before a less accurate prediction was observed. Some of the results of this early work are repeated here (Lisapaly, 2020a and Lisapaly, 2020b):

- A Covid-19 may be observed on 24th April 2020
- Increase of active cases between the first and third week of May 2020
- Accurate daily prediction
- At 22 April 2020 predicted 10,003-10,280 case at the end of April 2020 (Actual 10,118)

Following the less accurate prediction of the Covid-19 case, the author introduced multiple normal distributions to predict the Covid-19 development after an investigation using digital signal processing by Fourier transform analysis (Lisapaly, 2020c). This analysis indicated that a sum of multiple normal distributions should approach the prediction.

The new model consisting of multiple normal distributions was tested to predict the Covid-19 development using the Indonesia Covid-19 data. After two months of observation, Lisapaly (2020d) predicted that the end of Covid-19 would be delayed. Figure 5 shows the Covid-19 development in Indonesia during the first three months of observation. The black dot shows the time when the prediction was made and published. It can be seen that after the prediction, the number of daily cases increases, and it reached the peak close to the end of May 2020. Later, another increase occurred that made Covid-19 termination further delayed.



 e-ISSN: ………………………..

Figure 5. Covid-19 development in Indonesia during the first three months of observation showing that the end of Covid-19 will be delayed. The black dot marks the day number when Lisapaly (2020d) predicted the delay.

The method described in this work has also been used to monitor the Covid-19 development in the Jakarta area, the capital city of Indonesia (Lisapaly, 2020e). After three months of observation, the tendency of the Covid-19 case in Jakarta started to decrease, as can be seen in Figure 6 from day 45 to 60. This decrease has created optimism that Covid-19 in Jakarta will reach its end. However, Lisapaly (2020f) predicted that there would be a new increase in Jakarta although the Covid-19 curve from day 45 to day 90, when the prediction was made, showed a flat or constant rate. However, looking at the curve from day 90 and beyond, one can see an indication of the Covid-19 increase in Jakarta.

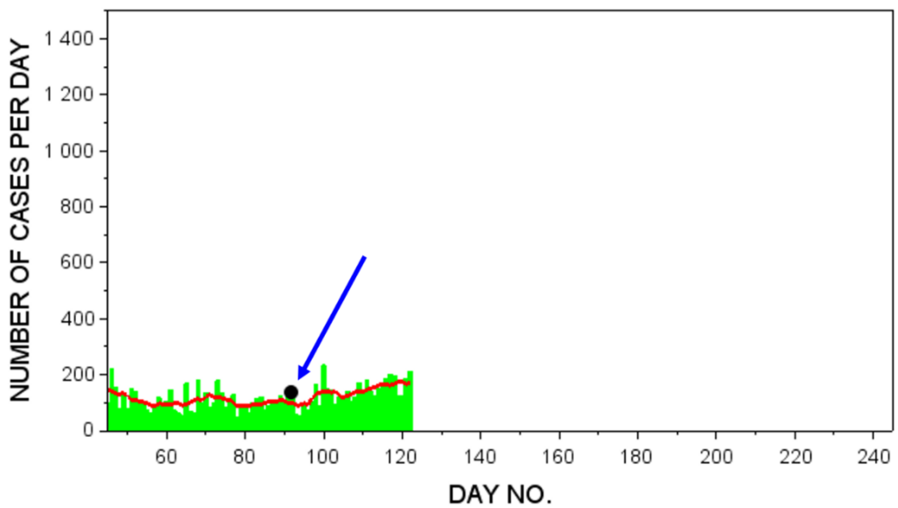


Figure 6. Covid-19 development during the first three months of observation showing that the Covid-19 was decreasing in Jakarta. The black dot marks the day number when Lisapaly (2020f) predicted that Covid-19 would increase again in Jakarta.

One month later, after analyzing new data, Lisapaly (2020g) predicted that a long Covid-19 wave would arrive in Jakarta. Figure 7 displays the Covid-19 curve in Jakarta during the latest eight months. The black dot indicates when the prediction and publication were made. The curve clearly shows that there was a significant increase in the Covid-19 case in Jakarta after the prediction. Therefore, given an accurate prediction, the government should prepare resources management to avoid unexpected problems.

 e-ISSN: ……………………….

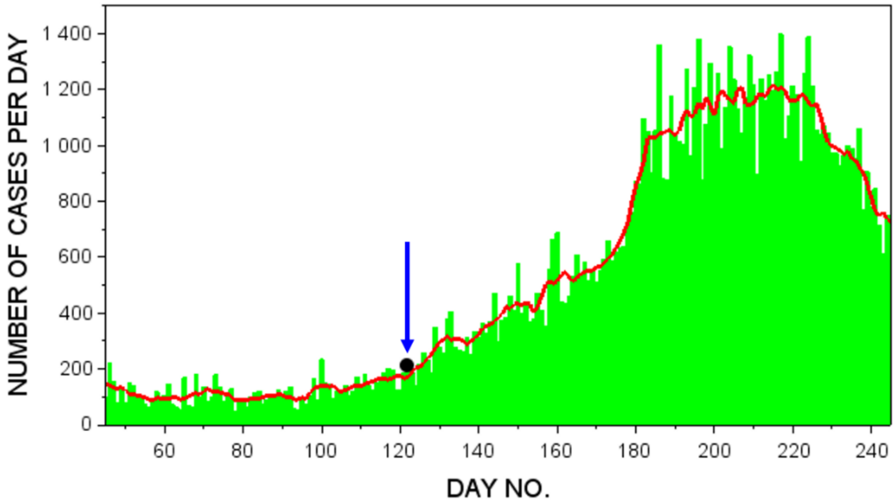


Figure 7. Covid-19 development during the first eight months of observation showing that a long Covid-19 wave was arriving in Jakarta. The black dot marks the day number when Lisapaly (2020f) predicted that Covid-19 would increase again in Jakarta and take a long duration.

Figure 8 provides the latest prediction on Covid-19 development in Indonesia when this study is presented. It predicts that the Covid-19 case in Indonesia will not decrease for a while. It will remain flat or constant and may increase a little bit to 6,000 – 6,500 per day. The Indonesia government will need to work harder, and strong cooperation from the Indonesian people is needed to reduce the daily Covid-19 case in Indonesia.

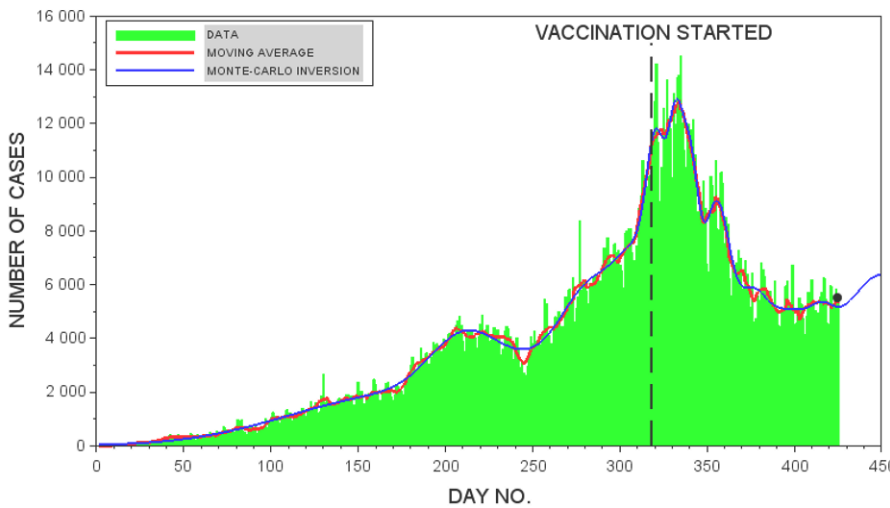


Figure 8. The latest prediction of Covid-19 development in Indonesia.

Conclusion

This study has shown that to predict the Covid-19 cases, a sum of the stack of multiple normal distributions provides a better result than the single normal distribution. Thus, the method has been able to predict potential new virus distribution. However, it needs at least data from the last

e-ISSN: ……………………….

three days to detect new virus distribution. From observation, it is easier to predict the development of the virus spread with a long duration.

Indonesia has progressed well enough to cope with the Covid-19 spread. However, the death ratio needs to be lowered, and reducing the number of daily cases should be speeded up for various reasons.

Reference

Jahani, E., M.A. Shayanfar, and M.A. Barkhordari (2012), "A New Adaptive Importance Sampling Monte Carlo Method for Structural Reliability." *KSCE Journal of Civil Engineering 17(1):210-215.*

Leite, L.W.B. and J.W.D. Leao (1985). "Ridge regression applied to the inversion of two-dimensional aeromagnetic anomalies." *Geophysics 50:1294-1306.*

Lin, W. and Cheng Su (2021). "An Efficient Monte-Carlo Simulation for the Dynamic Reliability Analysis of Jacket Platforms Subjected to Random Wave Loads." *Journal of Marine Science and Engineering.*

Lisapaly, L. (2020a). "The Prediction of Covid-19 Cases Development in Indonesia for Hospital and Human Resources Management Using the Curve Fitting of the CDF of Normal Distribution." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 110*.

Lisapaly, L. (2020b). "Heading for the end of Covid-19 in Indonesia." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 271*.

Lisapaly, L. (2020c). "Investigation of The Potential Existence of Covid-19 Data Clustering In Indonesia for a Better Prediction of Covid-19 Cases Development Using Digital Signal Processing." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 136.*

Lisapaly, L. (2020d). "Covid-19 Potential Delay in Indonesia." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 121.*

Lisapaly, L. (2020e). "Covid-19 Prediction for Jakarta Area." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 133.*

Lisapaly, L. (2020f). "New peak preventing early termination of Covid-19 cases in Jakarta." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 194.*

Lisapaly, L. (2020g). "A Moderate Long Covid-19 Wave is Happening in Jakarta." *SINTA Database of Researches, Publications, and Innovations on Coronavirus Disease (Covid-19), Research Abstract No. 314.*

Papadrakakis, M., V. Papadopoulos, and N.D. Lakaros (1995). "Structural reliability analysis of elastic-plastic structures using neural networks and Monte-Carlo simulation." *Computer method in Applied Mechanics and Engineering 136:145-163.*

Smith, F.B. and D.F. Shanno (1971). "An improved Marquardt procedure for non-linear regressions." *Technometric 13:63-74.*

Walpole, E.R., R.H. Myers, S.L. Myers, and K.E. Ye (2012). "Probability and Statistics for Engineers and Scientists." *Pearson*.

 e-ISSN: ………………………….