Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 2, No. 1, 2024, 17-25
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Partial Derivatives of Gompertz, Logistic, and Weibull Non-Linear Growth Models on Confirmed COVID-19 Cases

Wardhani Utami Dewi[1*], Warsono[2]

[1,2]Postgraduate at Universitas Lampung, Lampung, Indonesia
[*]corresponding author: dewiutamiwardhani@gmail.com

Received April 15, 2024; Received in revised form April 26, 2024; Accepted April 27, 2024

Abstract. *The epidemiological picture of COVID-19 is still unknown, and the number of confirmed cases of COVID-19 varies every day. Researchers have studied COVID-19 a lot, and many of them have used statistical models to estimate the growth of the outbreak. Non-linear statistical models can be used to describe growth behavior, as it varies in time. The aim of this research is to analyze, compare, and find the best model from the Gompertz, Logistic, and Weibull non-linear models. Daily cumulative data on confirmed COVID-19 viruses in Indonesia for 2020-2021 will be used in this research. The results obtained by the Logistic model proved to be very effective in describing the COVID-19 epidemic curve and estimating epidemiological parameters. The Logistic Model provides the best results compared to other growth models applied by Gompertz and Weibull. The R-Square of the logistic model is 0.9990, meaning that the model is able to explain or predict 99.90% of the data and 0.10% is explained by other factors. However, this research cannot explain the turning point of the curve, because there are many factors other than the model. One of them is the nature of the virus carrier from one place to another, then the behavior of the carrier who has not fully implemented the health protocol rules.*
Keywords*: Gompertz model; Logistic model; Weibull model*

INTRODUCTION

Several years ago the whole world was shaken by a deadly virus. The World Health Organization has declared the COVID-19 outbreak a Public Health Emergency of International Concern. The epidemiological picture of this disease is still unknown, and the total number of cases and deaths varies every day (January, 2020). Officially, the first time COVID-19 was recognized and recorded as having entered Indonesia was in March 2020 (Pratikto, 2020). Governments in various countries are starting to take policies in the form of limiting social contact to break the chain of transmission (Çelik, 2021). Likewise, the Indonesian government is limiting large-scale social contact. This is done to reduce the growth rate of confirmed COVID-19 cases. These steps implemented by various countries have proven effective in reducing the spread of disease (Felipe, Ms, Cortés-cortés, & Ms, 2020).

Despite the measures taken, the virus has not been able to be stopped because its infectious power is very fast and high (Science, Phenomena, & Ballı, 2021). The growth of confirmed cases of COVID-19 is increasing day by day. Researchers are seeking understanding of

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 2, No. 1, 2024, 17-25
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

COVID 19, many of them using statistical models to estimate the increase in incidence. This disease began to spread in China, so the first research in this area was carried out in China. Logistical growth modeling of COVID-19 proliferation in China and its international implications (Shen, 2020). Valle (2020) predicted the total number of COVID-19 cases and deaths in Brazil with the Gompertz model. Moreau (2020) estimated the COVID-19 pandemic in Brazil by statistical modeling using the Weibull distribution for daily new cases and deaths. Apart from that, Warsono et al. (2021) studied a general statistical distribution model of PM2.5 concentrations during the COVID-19 pandemic in Jakarta, Indonesia.

Non-linear statistical models have been widely used to describe growth behavior, as it varies in time. A non-linear model is where at least one of the parameters appears non-linear. In this research, non-linear models will be studied to describe the growth of COVID-19 in Indonesia, namely the Gompertz, Logistics and Weibull growth models. Apart from analyzing the non-linear model, this research will also compare and determine the best model from the three models. Growth curve studies in epidemic and disease outbreak studies are analyzed using non-linear models (Ghanim Al-Ani, 2021). For this reason, it is only focused on the indicator of the number of daily confirmed cases of COVID-19. It is hoped that this research can estimate the size of case growth in Indonesia.

METHOD

This research uses Indonesian daily cumulative data on confirmed Covid-19 viruses in Indonesia starting from March 2 2020 to May 31 2021. This data is updated every day and shows confirmed cases of this disease for all regions in Indonesia. The source for obtaining the data is the official website www.covid19.go.id, which also displays a graph of the growth of COVID-19 in Indonesia which can be seen in Figure 1.

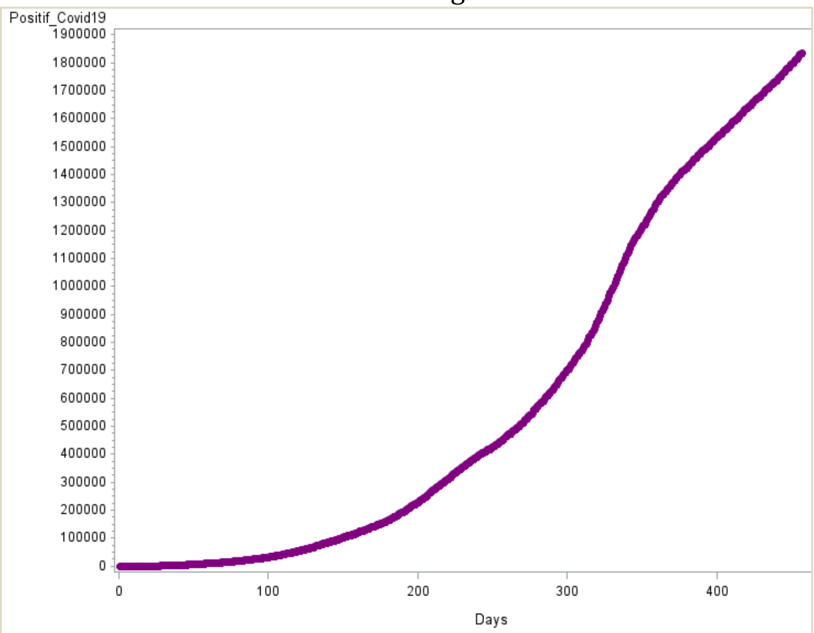

Figure 1. Daily Cumulative Data of Confirmed Covid-19 in Indonesia

The method used to analyze the growth of COVID-19 is the Gompertz, Logistics and Weibull non-linear models with the help of SAS software. There are two variables used, namely the variable t_i where i=1,2,….,457 shows the time or date of confirmed growth of COVID-19 in

Volume 2, No. 1, 2024, 17-25
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Indonesia and the dependent variable, namely the daily cumulative data of confirmed COVID-19 in Indonesia.

The general form of a growth model or non-linear model is (Ghanim Al-Ani, 2021):

$$y_t = g(t_i\,;\,\boldsymbol{\beta}) + \varepsilon_t, \quad i = 1,2, \ldots, n$$

Where,
y_t = dependent variable or response variable (cumulative confirmed COVID-19)
t = independent variable (time)
$\boldsymbol{\beta}$ = vector whose parameter values are unknown, so $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^{\boldsymbol{T}}$
ε_t = random error, so $\varepsilon_t \sim \mathrm{NID}(0, \sigma^2)$

Following are the non-linear functions of the Gompertz, Logistic, and Weibull growth models.

The Gompertz growth model is generally written in the form of the equation below (Blasco, Piles, & Varona, 2003; Piegorsch & Bailer, 2005):

$$g(t_i\,;\,\boldsymbol{\beta}) = \beta_0 \exp\{-e^{-\beta_1-\beta_2 t_i}\} \quad (1)$$

Double exponentials in equations make functions quite complicated. The parameter β_0 shows the number of confirmed cases which is asymptotic and is the limit of the function when t_(i) tends towards infinity. Meanwhile β_1 is the displacement of the t_i axis, and β_2 represents the growth rate (Szyszkowicz, 2021b). In this analysis, the time variable t_i is measured in days. When $\beta_2 > 0$, the function is a strictly increasing sigmoid, with a vertical intercept at $\beta_0 \exp\{-e^{-\beta_1}\}$, , a left horizontal asymptote (where $t \to -\infty$) on the horizontal axis, and a horizontal asymptote right (where $t \to \infty$) at β_0. Meanwhile, if $\beta_2 < 0$, the sigmoid decreases and the asymptote is reversed. In conclusion, β_2 has quite a big influence, because it controls the time function with the response.

According to Laida & Fermin (2022) the Logistics non-linear model is written in the form of the following equation;

$$g(t_i\,;\,\beta) = \frac{\beta_0}{1+e^{-\beta_1-\beta_2 t_i}}\,. \quad (2)$$

Equation (2) can also be written into the equation,

$$g(t_i\,;\,\beta) = \beta_0(1 + e^{\beta_1+\beta_2 t_i})^{-1} e^{\beta_1+\beta_2 t_i}\}.$$

As with the Gompertz curve, (2) has three unknown parameters; however, it differs from Gompertz in many of its operational characteristics. The most important of these is that is sigmoid symmetric: the rate of change is symmetric about the point of inflection. Thus the two curves complement each other, (1) serves to represent an asymmetric growth pattern and (2) serves to represent a symmetric pattern.

Another sigmoid curve based on the exponential operator is the Weibull growth model. The model takes a slightly different approach to relating unknown parameters to predictor

Volume 2, No. 1, 2024, 17-25
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

variables and, as a result, can allow for a variety of curved shapes. The Weibull model is written in the form of the following equation (Dagogo, Nduka, & Ogoke, 2020):

$$g(t_i\,;\,\boldsymbol{\beta}) = \beta_0 + \beta_1 \exp\{-\beta_2 {t_i}^{\beta_3}\}\,. \tag{3}$$

The Weibull model is a flexible and simple function with great potential for COVID-19 growth rate applications. Weibull has been used by several authors to estimate and analyze the spread of COVID-19 (Hembram & Kumar, 2021; Llanes, Rodríguez, Duarte, & María, 2021; Omara & Harby, 2021).

Calculation (3) is limited to $t_i > 0$ unless β_3 is a positive integer. Curves can contain a variety of complex nonlinear patterns. For example, if $\beta_2 > 0$ and $\beta_3 > 0$, then at $= 0$ the curve has a vertical intercept of $g(0;\,\boldsymbol{\beta}) = \beta_0 + \beta_1$ and if t$\rightarrow \infty$ the curve approaches the horizontal asymptote of β_0. However, if $\beta_2 < 0$ then $g(0;\,\boldsymbol{\beta}) = \beta_0 + \beta_1$ and for $\beta_3 > 0$ it diverges to ∞ (if $\beta_1 > 0$) or $-\infty$ (if $\beta_1 < 0$) where $t \rightarrow -\infty$. For $\beta_3 < 0$, $g(0;\,\boldsymbol{\beta})$ diverges to ke ∞, $g(0;\,\boldsymbol{\beta})$ approaches the horizontal asymptote of $\beta_0 + \beta_1$. In this last sense, it approaches simple hyperbole.

The stages carried out in this research:

1. Preparing daily cumulative data on confirmed COVID-19 in Indonesia.
2. Look for partial derivatives of the model equations (1), (2), and (3) for each parameter.
3. Carry out non-linear regression estimates for the three model
4. Evaluate the model based on the *R-Square* value
5. Determining the best model

RESULY AND DISCUSSION

There are two confirmed cases of COVID-19, namely confirmed cases with symptoms (symptomatic) and confirmed cases without symptoms (asymptomatic). The results of the RT-PCR laboratory examination are proof that a person has tested positive for infection. This is summarized in the COVID-19 data, in Indonesia initially there were 2 people confirmed with COVID-19 and every day this increased to millions. Each data represents the t variable which indicates time and the g variable as daily cumulative data on confirmed COVID-19. Statistically, this data can be estimated and analyzed using the Gompertz, Logistics and Weibull growth models. In accordance with the growth curves of the three models, this research uses the non-linear Levenberg-Marquardt algorithm (Moré, 1978). The Gompertz partial derivative in equation (1) which is useful for implementing the algorithm is as follows.

$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_0} = \exp\{-e^{-\beta_1-\beta_2 t_i}\}$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_1} = \beta_0 \exp\{-\beta_1 - \beta_2 t_i - e^{-\beta_1-\beta_2 t_i}\}$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_2} = t_i\,\beta_0 \exp\{-\beta_1 - \beta_2 t_i - e^{-\beta_1-\beta_2 t_i}\}$$

For the Logistic model equation (2) is also derived partially, so it is obtained,

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 2, No. 1, 2024, 17-25
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_0} = \left(1 + e^{\beta_1+\beta_2 t_i}\right)^{-1}$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_1} = \beta_0 e^{-\beta_1-\beta_2 t_i}\left(1 + e^{\beta_1+\beta_2 t_i}\right)^{-2}\}$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_2} = t_i\,\beta_0 e^{-\beta_1-\beta_2 t_i}\left(1 + e^{\beta_1+\beta_2 t_i}\right)^{-2}\}$$

Like the Gompertz da Logistik model, (3) also fits to the data via the non-Linear Levenberg-Marquardt algorithm (Moré, 1978). The partial derivative in question is.

$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_0} = 1$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_1} = \exp\{\beta_2 {t_i}^{\beta_3}\}$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_2} = -{t_i}^{\beta_3}\beta_1 \exp\{\beta_2 {t_i}^{\beta_3}\}$$
$$\frac{\partial g(t_i\,;\,\boldsymbol{\beta})}{\beta_3} = -{t_i}^{\beta_3}\log(t_i)\beta_1\beta_2 \exp\{\beta_2 {t_i}^{\beta_3}\}$$

After all models have been partially derived, parameter estimation will be carried out using SAS software. The parameter estimation results are presented in Table 1.

Table 1. Parameter estimates for the Gompertz, Logistic, and Weibull models

Model	Estimated Coefficient				SSR	SST	R^2
	β_0	β_1	β_2	β_3			
Gompertz	3025105	-2.3380	0.00675	-	3.268	3.275	0.9978
Logistik	2056142	-5.4801	0.0164	-	3.272	3.275	0.9990
Weibull	13165.9	18864439	337.0	-0.8158	1.683	1.693	0.9940

The Levenberg–Marquardt iterative method was chosen because it is intermediate between the linearization of the Gauss–Newton method and the steepest descent method and appears to combine the best features of both while avoiding their existing limitations. Using the initial values of the parameters above, the estimates of the learned model are:

Gompertz Model Estimation
$$\hat{y}_t = 3025105 \exp\left[-e^{23380-0.00675t}\right]$$
Logistic Model Estimation
$$\hat{y}_t = \frac{2056142}{1 + e^{5.4801+0.0164}}$$
Weibull Model Estimation
$$\hat{y}_t = 13165.9 + 18864439 \exp[-337t^{-0.8158}]$$

So that the confirmed prediction results for COVID-19 are obtained which are presented in the following Table 2, 3, and Table 4.

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 0, No. 0, 20xx, 00-00
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Table. 2 Prediction Results with the Gompertz Model

Days	Positif_Covid19	pred	resid
1	2	102.66	-100.66
2	2	110.02	-108.02
3	2	117.85	-115.85
4	2	126.18	-124.18
5	2	135.04	-133.04
6	4	144.46	-140.46
⋮	⋮	⋮	⋮
454	1815591	1865510.31	-49919.31
455	1822156	1871587.72	-49431.72
456	1828721	1877643.85	-48922.85
457	1834836	1883678.63	-48842.63

Table. 3 Prediction Results with Logistic Models

Days	Positif_Covid19	pred	resid
1	2	8677.06	-8675.06
2	2	8820.09	-8818.09
3	2	8965.47	-8963.47
4	2	9113.23	-9111.23
5	2	9263.42	-9261.42
6	4	9416.07	-9412.07
⋮	⋮	⋮	⋮
454	1815591	1805406.25	10184.75
455	1822156	1808998.68	13157.32
456	1828721	1812546.58	16174.42
457	1834836	1816050.31	18785.69

Table. 4 Prediction Results with the Weibull Model

Days	Positif_Covid19	pred	resid
1	2	13165.91	-13163.91
2	2	13165.91	-13163.91
3	2	13165.91	-13163.91
4	2	13165.91	-13163.91
5	2	13165.91	-13163.91
6	4	13165.91	-13161.91
⋮	⋮	⋮	⋮
454	1815591	1920948.50	-105357.50
455	1822156	1928803.58	-106647.58

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 0, No. 0, 20xx, 00-00
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

Days	Positif_Covid19	pred	resid
456	1828721	1936659.52	-107938.52
457	1834836	1944516.27	-109680.27

Each model displays a curve of predicted results for confirmed COVID-19 with actual cases.

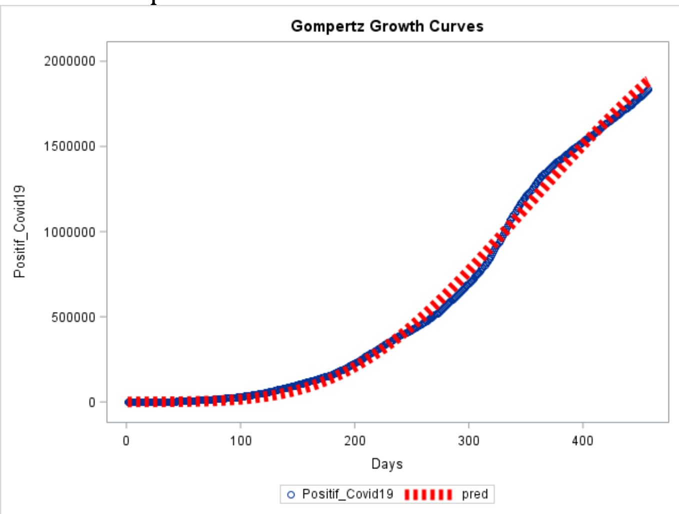

Figure 2. Gompertz Growth Model Curve

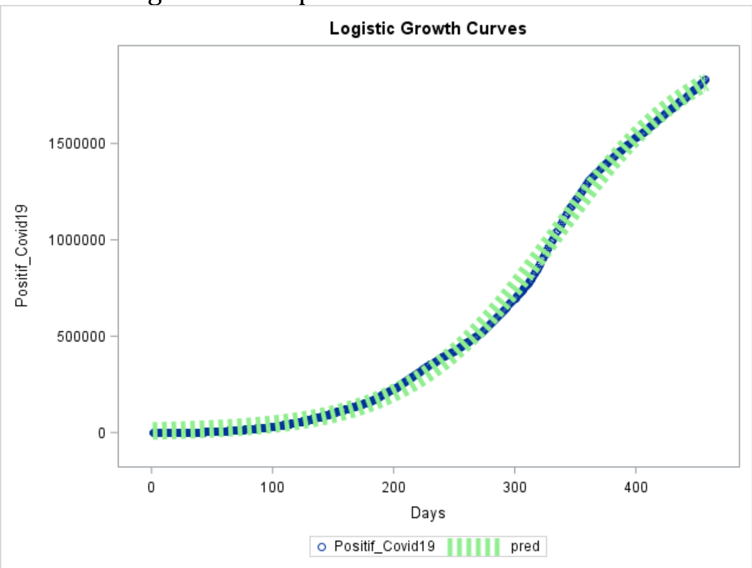

Figure 3. Logistic Growth Model Curve

Sciencestatistics
Journal of Statistics, Probability, and Its Application
Available at https://scholar.ummetro.ac.id/index.php/sciencestatistics/index

Volume 0, No. 0, 20xx, 00-00
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

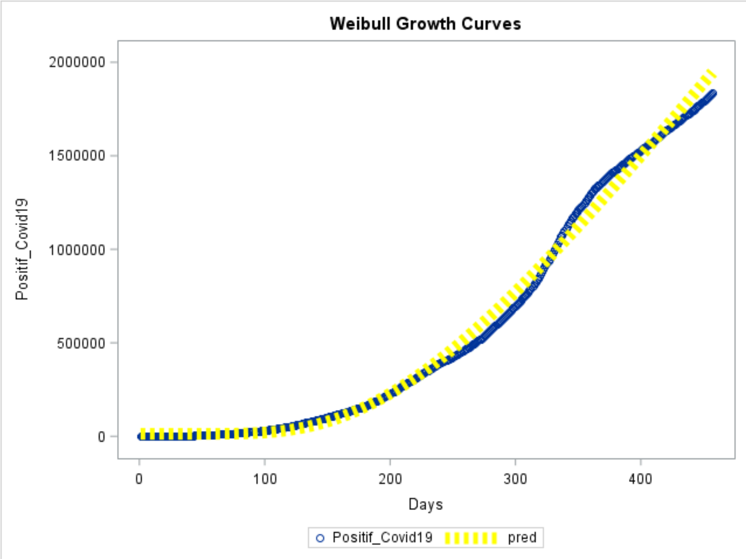

Figure 4. Weibull Growth Model Curve

The three images above adjust the growth model to the daily growth of COVID-19 cases in Indonesia. Non-linear functions are shown in the figure in red for the Gompertz model, green for the Logistic model, and yellow for the Weibull model. Meanwhile, the original data is illustrated in blue circles. The curve explains that the data spreads out and even overlaps the estimation line from the model. This shows that the three non-linear models are able to predict the daily cumulative confirmed COVID-19. Determining the best model can be done by evaluating model suitability. One of them uses R-Square (R^2) with the aim of measuring the fit between the model and the data. The results obtained are the Logistics model that best describes the growth of COVID-19 in Indonesia, namely with an R^2 value of 0.9990. In accordance with table 1. the Gompertz model also has a fairly large R^2 and is almost close to the R^2 value in the Logistic model, namely 0.9978. Meanwhile, Weibull only provides an R^2 value of 0.9940.

CONCLUSION

The Logistic Model proved to be very effective in describing the COVID-19 epidemic curve and estimating important epidemiological parameters, such as the peak time of the curve for daily cases, thereby enabling practical and efficient monitoring of the epidemic evolution. In this research, the Logistic model provides the best results compared to other growth models applied by Gompertz and Weibull. The R-Square of the logistic model is 0.9990, meaning that the model is able to explain or predict 99.90% of the data and 0.10% is explained by other factors. However, this research cannot explain the turning point of the curve, because there are many factors outside of confirmed COVID-19. One of them is the nature of the virus carriers from one place to another, then the behavior of carriers who have not fully implemented government regulations.

REFERENCE

Blasco, A., Piles, M., & Varona, L. (2003). A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. *Genetics Selection Evolution*, *35*(1), 21–41.

Çelik, Ş. (2021). *Modeling of COVID-19 cases and deaths in Uruguay with different , nonlinear growth models*. *7*(11), 15–18.

Dagogo, J. E., Nduka, C., & Ogoke, U. P. (2020). Comparative Analysis of Richards, Gompertz

Volume 0, No. 0, 20xx, 00-00
ISSN 2963-9875 (Online)
ISSN 2964-2884 (Print)

and Weibull Models. *IOSR Journal of Mathematics*, *16*(1), 15–20. https://doi.org/10.9790/5728-1601041525

Felipe, J., Ms, J. F. M., Cortés-cortés, M., & Ms, M. C. (2020). *COVID-19 Forecasts for Cuba Using Logistic Regression and Gompertz Curves*.

Ghanim Al-Ani, B. (2021). Statistical modeling of the novel COVID-19 epidemic in Iraq. *Epidemiologic Methods*, *10*(s1), 1–16. https://doi.org/10.1515/em-2020-0025

Hembram, K. P. S. S., & Kumar, J. (2021). Epidemiological study of novel coronavirus (COVID-19): macroscopic and microscopic analysis. *International Journal Of Community Medicine And Public Health*, *8*(3), 1364. https://doi.org/10.18203/2394-6040.ijcmph20210828

January, B. (2020). *Modeling of COVID-19 Outbreak Indicators in China*. 223–231. https://doi.org/10.1017/dmp.2020.323

Laida, D. G., & Fermin, E. (2022). Hospital preparedness during epidemics using simulation : the case of COVID-19. *Central European Journal of Operations Research*, *30*(1), 213–249. https://doi.org/10.1007/s10100-021-00779-w

Llanes, C. O., Rodríguez, R. O., Duarte, R. F., & María, P. (2021). Comparison of Linear ROR Vs Nonlinear Weibull Model for COVID- 19 in Iraq Himalayan Journal of Applied Medical Sciences and Research Comparison of Linear ROR Vs Nonlinear Weibull Model for COVID- 19 in Iraq. *Himalayan Journal of Applied Medical Sciences and Research*, *2*(5), 88–96. https://doi.org/10.47310/hjamsr.2021.v02i05.018

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis* (pp. 105–116). Springer.

Moreau, V. H. (2020). *Forecast predictions for the COVID-19 pandemic in Brazil by statistical modeling using the Weibull distribution for daily new cases and deaths*. 1109–1115.

Omara, T. M., & Harby, K. A. (2021). Using Mathematical and Statistical Model to Forecast the Path of Infection by COVID-19 in the Kingdom of Saudi Arabia. *IIUM Medical Journal Malaysia*, *20*(2), 195–203. https://doi.org/10.31436/IMJM.V20I2.1683

Piegorsch, W. W., & Bailer, A. J. (2005). *Analyzing environmental data*. Southern Gate, Chichester: John Wiley & Sons.

Pratikto, F. R. (2020). *Prediksi Akhir Pandemi COVID-19 di Indonesia dengan Simulasi Berbasis Model Pertumbuhan Parametrik*. *9*(2).

Science, N., Phenomena, C., & Ballı, S. (2021). *Data analysis of COVID-19 pandemic and short-term cumulative case forecasting using machine learning time series methods*. *142*. https://doi.org/10.1016/j.chaos.2020.110512

Shen, C. Y. (2020). Logistic growth modelling of COVID-19 proliferation in China and its international implications. *International Journal of Infectious Diseases*, *96*, 582–589. https://doi.org/10.1016/j.ijid.2020.04.085

Szyszkowicz, M. (2021a). *Advances in Environmental and Engineering Research Modelling the Cumulative Number of COVID-19 Cases*. https://doi.org/10.21926/aeer.2102014

Szyszkowicz, M. (2021b). Modelling the Cumulative Number of COVID-19 Cases. *Advances in Environmental and Engineering Research*, *2*(2), 1.

Valle, J. A. M. (2020). Predicting the number of total COVID-19 cases and deaths in Brazil by the Gompertz model. *Nonlinear Dynamics*, *102*(4), 2951–2957. https://doi.org/10.1007/s11071-020-06056-w

Warsono, W., Antonio, Y., Yuwono, S. B., Kurniasari, D., Suroso, E., Yushananta, P., … Hadi, S. (2021). Modeling generalized statistical distributions of pm2.5 concentrations during the COVID-19 pandemic in Jakarta, Indonesia. *Decision Science Letters*, *10*(3), 393–400. https://doi.org/10.5267/j.dsl.2021.1.005