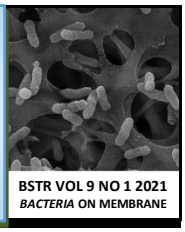
HIBISCUS PUBLISHER

BIOREMEDIATION SCIENCE & TECHNOLOGY RESEARCH

Website: http://journal.hibiscuspublisher.com/index.php/BSTR/index



Outlier and Normality Testing of the Residuals for the Morgan-Mercer-Flodin (MMF) Model Used for Modelling the Total Number of COVID-19 Cases for Brazil

Garba Uba[1]* Nuhu Danladi Zandam[1] , Abdurrashed Mansur[2] and Mohd Yunus Abd Shukor[3]

[1]Department of Science Laboratory Technology, College of Science and Technology, Jigawa State Polytechnic, Dutse, PMB 7040, Nigeria.

[2]Department of Microbiology, Faculty of Science, Gombe State University.

[3]Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.

*Corresponding author:
Garba Uba
Department of Science Laboratory Technology,
College of Science and Technology,
Jigawa State Polytechnic, Dutse,
PMB 7040
Email: garbauba@jigpoly.edu.ng

HISTORY

Received: 20th June 2020
Received in revised form: 4th July 2021
Accepted: 15t July 2021

KEYWORDS

MMF
COVID 19
Outlier
Brazil
Grubbs' Statistic

ABSTRACT

Traditionally, testing for outliers is performed by first creating a null hypothesis, H_0, indicating that the suspected results do not differ significantly from those of other members of the data set, and then rejecting it if the likelihood of getting the experimental results is extremely low (e.g., p=0.05). Similarly, if H_0 can be rejected, the questionable findings may be discarded as outliers as well. If H_0 is retained in the data set, it is important to keep the dubious findings in the data set. In general, in nonlinear regression, the residuals of the curve must be normally distributed before any test for the existence of outliers is performed. This is often accomplished through the use of normalcy tests such as the Kolmogorov-Smirnov, Wilks-Shapiro, D'Agostino-Pearson, and Grubb's tests, the latter of which checks for the presence of an outlier and is the subject of this study. Normality tests for residues used in general nonlinear regression revealed that the usage of the Morgan-Mercer-Flodin (MMF) Model used for Modelling the Total Number of COVID-19 Cases for Brazil was adequate due to lack of an outlier. The critical value of *Z* from statistical table for Grubbs' test for a single outlier using mean and SD was 0.114 (n=50). The Grubbs (Alpha = 0.05) ***g*** value was 3.597. Individual Z value indicates that the residual with a value of -3 (row 3) was far from the rest and is deemed a significant outlier ($p < 0.05$). This outlier was removed, and subsequent Grubb's test show the absence of other outliers. As the Grubbs' test require for the normality of the residuals, several normality tests (Kolmogorov-Smirnov, Wilks-Shapiro, Anderson-Darling and the D'Agostino-Pearson omnibus K2 test) were carried out and the results were found to conform to normality. In addition, a visual inspection of the model's normal probability or Q-Q plot shows a nearly straight and appeared to exhibit no underlying pattern. The resulting histogram overlaid with the ensuing normal distribution curve also reveals that the residuals were truly random and that the model used was adequately fitted.

INTRODUCTION

According to the World Health Organization (WHO), the COVID-19 pandemic is a worldwide coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that was announced by the WHO in March 2020. The new virus was identified in Wuhan, China, in December of 2019; a lockdown in Wuhan and other towns in Hubei province failed to stop the epidemic, which expanded to other parts of mainland China and throughout the world after the outbreak was detected. The outbreak was declared a Public Health Emergency of International Concern by the World Health Organization (WHO) on January 30, 2020, and a pandemic was proclaimed on March 11, 2020. The virus has been causing outbreaks in a number of countries since 2021, with the most hazardous strains being the Delta, Alpha, and Beta variants, which are the most common. As of the 22nd of July, 2021, more than 191 million cases have been documented, with more than 4.12 million

confirmed deaths due to COVID-19, making it one of the deadliest pandemics in history [1,1–5]. People are becoming increasingly conscious of the uneven distribution of SARS-COV-2 mortality among disadvantaged communities as the death toll from COVID-19 continues to rise around the world. In addition to the elderly, individuals who live in densely crowded regions, persons who have poor socioeconomic position, refugees, and members of minority groups should be taken into consideration. Almost every community is in risk of being annihilated. The infection rates in these populations are higher than those in the general population, making them more susceptible to infection and unfavourable disease effects than other groups [6,7].

In the initial period, mathematical modelling research in Wuhan City and Hubei Province total infectious cases was focused on the dynamics of the pandemic[8]. At this early stage, it has taken a significant amount of time and effort to examine surveillance data from China in order to produce parameter estimates such as the basic reproduction number (R0), case fatality rate, and incubation duration [9]. Early attempts at Susceptible-Exposed-Infectious-Recovered (SEIR) style dynamic models were 'borrowed' from what was known about other coronaviruses (SARS-CoV and MERS-CoV) and/or gained through fitting the models to monitoring data gathered during the initial outbreaks [10]. COVID-19 pandemic assessments can be carried out with the help of statistical models, including theoretical, quantitative, and simulation models. Mathematical models are then applied for other affected countries to better understand the mode and spread of infection [1,4,11–18].

Organisms growth including viral infection cases over time usually exhibit a sigmoidal growth profile that exhibits lag time (λ), acceleration to a maximal value (μ_m) and a final phase where the rate decreases and eventually reaches zero or an asymptote (A) is observed [19]. The sigmoidal curve can be fitted by different mathematical functions, such as Logistic [19,20], modified Gompertz [19,21], Richards [19,22], Schnute [19,23], Baranyi-Roberts [24], Von Bertalanffy [19,25–27], Buchanan three-phase [28,29], Huang [30–33] and Morgan-Mercer-Flodin (MMF) [34–43, 43–47]. For the analysis of the COVID-19 pandemic [8], strong predictive ability was employed models, such as updated Gompertz and Bertalanffy and logistics. The total infection case of SARS-CoV-2 in Brazil as of 15th of July 2020 to the 20th of December 2020 was modelled using several primary growth models via nonlinear regression. The MMF models found to be the best [41,42,44,46,48–51].

When linearization is used to smooth out an obviously nonlinear curve, the error structure of the data is disturbed, as is the case in this example. As a result, evaluating the uncertainty of the kinetic parameters, which are often given as a 95 percent confidence interval range, becomes more difficult. In addition, the linearization method leads in the introduction of error into the independent variable as a result of the linearization procedure [52–57]. Additionally, changes in the weights assigned to each data point can occur, which typically results in differences in the fitted parameter values between the linear and nonlinear versions of the kinetics model when compared to the linear version. Nonetheless, in nonlinear regression, the residuals of the curve must be normally distributed, and the residuals must be checked for the presence of outliers [at 95 or 99 percent confidence levels]. In most cases, normality tests such as the Kolmogorov-Smirnov, Wilks-Shapiro, and D'Agostino-Pearson, as well as the Grubb's test, which tests for the presence of an outlier, are used to do this. The Grubb's test is the subject of this study because it tests for the presence of an outlier.

METHOD

Data on the mathematical modelling of the total number of COVID-19 cases for Brazil using the MMF model (**Equation 1**) from our previous works [43] was utilized in this study (Fig. 6).

Residuals

Residuals are very important in assessing the health of a curve from a particular used model. Mathematically, residual for the i^{th} observation in a given data set can be defined as follows (**Eqn. 1**);

$$e_i = y_i - f\left(x_i;\hat{\beta}\right) \quad \text{..... (Eqn. 1)}$$

where y_i denotes the i^{th} response from a given data set while x_i is the vector of explanatory variables to each set at the i^{th} observation corresponding values in the data set.

Grubbs' Statistic

In an average value, a single data point with deformation can lead to gross error in the fitting of a nonlinear curve. Therefore, searching for an outlier is an integral aspect of curve fitting. The Grubbs test is used to evaluate the outlier in the univariate environment and the data is normally distributed [58]. The test can be applied to the maximal or minimal observed data from a Student's t distribution (**Eqn. 2**) and to test for both data simultaneously (**Eqn. 3**).

$$G_{\min} = \frac{\bar{X} - \min(X)}{s}$$

$$G_{\max} = \frac{\max(X) - \bar{X}}{s} \quad \text{(Eqn. 2)}$$

$$p_G = 2n.p_t\left(G\frac{\sqrt{n(n-2)}}{n-1}, n-2, 1\right)$$

$$G_{\text{all}} = \frac{\max(\bar{X} - \min(X), \max(X) - \bar{X}}{s}$$

$$p_G = n.p_t\left(G\frac{\sqrt{n(n-2)}}{n-1}, n-2, 2\right) \quad \text{(Eqn. 3)}$$

Normality test

Residuals from the pseudo-1st order model were subjected to three normality tests- Kolmogorov-Smirnov [59,60], Wilks-Shapiro [61], Anderson-Darling [62] and the D'Agostino-Pearson omnibus K2 test [63]. Using graphical and numerical methods are two ways to search for normality. The simplest and easiest way to assess the normality of data is via graphical methods such as the normal quantile–quantile (Q-Q) plots, histograms or box plots [64]. The normality tests were carried out using the GraphPad Prism® software (Version 6.0, GraphPad Software, Inc., USA).

RESULTS AND DISCUSSION

Statistics often used in nonlinear regressions rely on the use of residual data, which is the difference between the expected and the actual values. Statistical analyses should be done to evaluate the adequacy of residues in randomness, do not include outliers, obey normality, and do not demonstrate autocorrelation. Usually, the greater the discrepancy between the expected and the observable values, the less well off the model. [65]. The Grubbs' test deals with one aspect at a time. Outliers are eliminated and test replicated before test passes without revealing any outliers. As a general rule, sample sizes of 6 or less results in biased data sets. Many variations of the same model alter the probability of

identification. The Grubb's test was applied to the residual results (**Table 1**). Grubbs test statistic defines the highest absolute variance from the survey mean in the sample standard deviation units. The critical value of *Z* from statistical table for Grubbs' test for a single outlier using mean and SD was 0.114 (n=50). The Grubbs (Alpha = 0.05) ***g*** value was 3.597. Individual Z value indicates that the residual with a value of -3 (row 3) was far from the rest and is deemed a significant outlier ($p < 0.05$) (**Table 2**). This outlier was removed, and subsequent Grubb's test show the absence of other outliers (Data not shown).

Table 1. Residual data from the pseudo-1st order model.

	Residuals
	0.0000
	-0.0600
	-0.1500
	-0.2000
	-0.4200
	-0.0100
	-0.1400
	0.1500
	0.0000
	0.1300
	0.2400
	0.2000
	0.1500
	0.1000
	0.1100
	0.0500
	0.06
	-0.01
	0
	-0.06
	-0.11
	-0.13
	-0.13
	-0.12
	-0.13
	-0.12
	-0.1
	-0.12
	-0.09
	-0.09
	-0.1
	-0.05
	-0.04
	-0.01
	-0.01
	0.01
	0.01
	0.02
	0.02
	0.03
	0.04
	0.04
	0.06
	0.06
	0.07
	0.07
	0.07
	0.07
	0.07
	0.07
Average	-0.010
Std deviation	0.12

Table 2. Descriptive statistics and calculated Z value for residual data. Only the first seven data are shown.

Row	Value	Z	Significant Outlier?
	0.0000	0.087720	
2	-0.0600	0.438598	
3	-0.1500	1.228074	A significant outlier ($P < 0.05$).
4	-0.2000	1.666672	
5	-0.4200	3.596503	
6	-0.0100	0.000000	
7	-0.1400	1.140354	

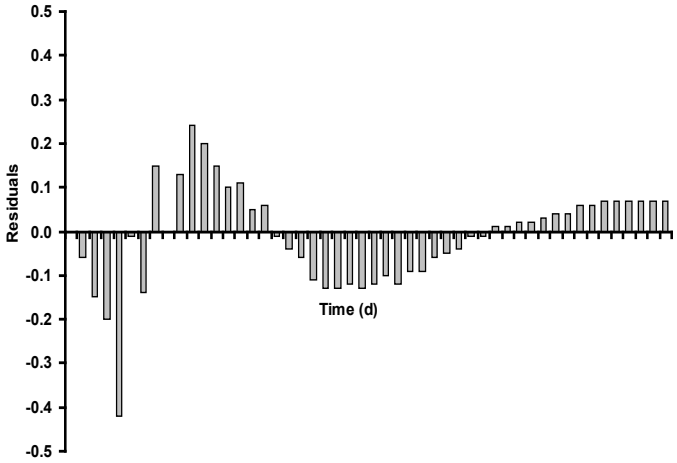


Fig. **1**. Residual plot for the pseudo-1st order model model.

Table 3. Numerical normality test for the residual from the pseudo-1st order model after removal of an outlier.

Test for normal distribution	
Anderson-Darling test	
A2*	0.5764
P value	0.1273
Passed normality test (alpha=0.05)?	Yes
P value summary	ns
D'Agostino & Pearson test	
K2	9.896
P value	0.0071
Passed normality test (alpha=0.05)?	No
P value summary	**
Shapiro-Wilk test	
W	0.9510
P value	0.0376
Passed normality test (alpha=0.05)?	No
P value summary	*
Kolmogorov-Smirnov test	
KS distance	0.1200
P value	0.0693
Passed normality test (alpha=0.05)?	Yes
P value summary	ns
Number of values	50

It is common practise to calculate the fitness of a mathematical model precisely through the use of residual measurements. As defined by a certain mathematical model, residuals are the difference between the sum expected and the total actually observed. The underlying premise is that the wider the disparity between the expected and observed values, the poorer the model is considered to be. The residual plot (observed-predicted) was tested, and the results of the investigation demonstrated that the data for all experiments exhibit an outlier (Fig. 1). Evidently, a probable outlier is a data point that is out of the ordinary and that the researcher determines to be impossible based on a range of specific criteria.

More specifically, an outlier in a study may be a unique characteristic that is far too uncommon in comparison to the rest of the population. To give an example, most outliers are only considered outliers if they are statistically excessively high for the distribution to the limit in the sample model, which is not always the case [66].

A simple strategy to identify potential outliers in testing is to include a boxplot, although more complex methodologies, such as the Chauvenet criteria in engineering and the 3-sigma criterion, coupled with the Z-score in chemometrics, are frequently employed. Despite the fact that these approaches are simple and quick, there is a considerably more effective way of employing the statistical test for outlier discovery than the methods described above. With the exception of one outlier, relevant assessments differ from the Dixon Q-test or the Grubbs ESD-test.

A variety of conditions have benefitted from the use of the Grubbs test to detect the presence of outliers [67–77]. The most significant restriction of the Grubbs test is that the thinking quantity of the outliers, denoted by the letter k, must be given explicitly. A failure to properly clarify the variable "k" can result in distorted results from the trials. A test called the Rosner Generalized Severe Studentized Deviate (ESD-test) is used when there are several outliers or when the exact number of outliers cannot be determined [78]. For example, if there is more than one outlier in a sample, the findings of the Grubbs test will be distorted, and when this occurs, the Ferguson sample skew test is more resistant to the misleading impact than the Grubbs test [79].

The number of bins and samples assessed determined the shape of the distribution. The W_2 statistic in the Wilks-Shapiro test is calculated using the anticipated values of the order statistics between identically distributed random variables as well as their independent covariance, as well as the regular normal distribution. The agreement is refused if the test statistics-W2 have a significant impact on the outcome. According to Royston, formalised euphemism is The Kolmogorov-Smirnov statistic, when applied to data, computes the cumulative residual frequency, which is a non-parametric numerical test [61]. It evaluates the link between the model and the observed values. It can also be used to compare two sets of data to see how they differ. The p value is derived using the difference between two combined distributions as well as the sample population size.

On a more general level, the Central Limit Theorem (CLT) claims that as n approaches infinite (in actuality, n>30), the probability frequency distribution tends to fit the Gaussian distribution on any continuous variable (even for discrete variables such as Binomial or Poison distributions) [80,81]. The skewness and kurtosis of the distributions were analysed as a technique of quantifying the difference between the sample distributions and the usual distribution in order to determine the significance of the results in the D'Agostino-Pearson normality test method. Following that, the p-value of the sum of these inconsistencies or discrepancies is calculated. D'Agostino developed a variety of normality tests, the most extensively used of which is the omnibus K2 test [63]. More and more nonlinear regression curve fitting exercise works are reporting an extensive testing for the normality of the residuals [82–90].

Graphical diagnostic of residuals normality

After removal of the outlier, the model's normal probability or Q-Q plot was nearly straight and appeared to exhibit no underlying pattern (**Fig. 4**). The resulting histogram (**Fig. 5**) overlaid with the ensuing normal distribution curve reveals that the residuals were truly random and that the model used was adequately fitted.

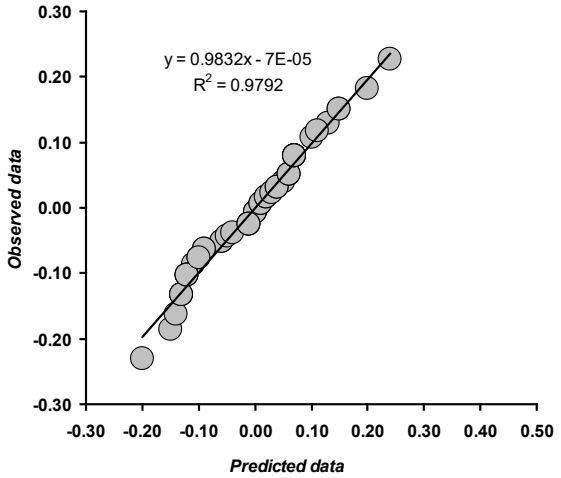


Fig 4. Normal Q-Q plot for the observed sample against theoretical quantiles after outlier removal.

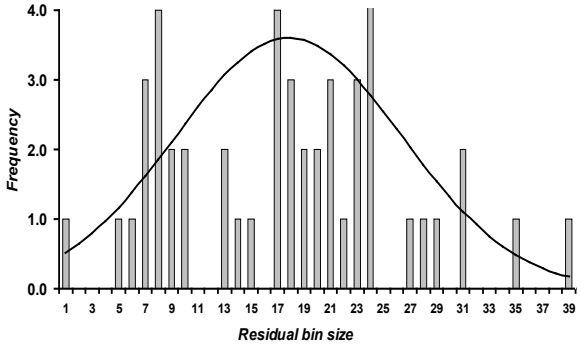


Fig. **5**. Histogram of residual for the pseudo-1st order model overlaid with a normal distribution (mean 0.00163 and standard deviation 0.0984).

CONCLUSION

In conclusion, the normality checks performed on the residues used in this study revealed that the use of the pseudo-1st order model in the fitting of Modelling the Total Number of COVID-19 Cases for Brazil was satisfactory due to the absence of an outlier, indicating that the model was appropriate for the data. Many studies on the application of the model used in the mathematical diagnostic of residues have been published, but it is widely known that they have not gone any farther in their exploration. In the case of a Gaussian or regular distribution, this may result in a data violation. Most of the parametric predictive estimate approaches used in nonlinear regression rely on this assumption as a necessary but not sufficient condition. On the basis of residuals that adhere to the normal distribution, methods such as the root mean square error, Pearson correlation coefficient (either standard or modified), the F-test, and the t-test are used. Type I and Type II mistakes might be avoided if these assumptions were followed. Additionally, in the event that diagnostic tests reveal that the data from the total number of COVID-19 Cases have violated any of the assumptions, the issue may be rectified in the field by implementing numerous nonparametric treatments or changing the form of the therapies in question.

REFERENCES

1. Areepong Y, Sunthornwat R. Predictive models for cumulative confirmed COVID-19 cases by day in Southeast Asia. CMES - Comput Model Eng Sci. 2020;125(3):927–42.
2. Andreopoulos P, Kalogeropoulos K, Tragaki A, Stathopoulos N. Could historical mortality data predict mortality due to unexpected events? ISPRS Int J Geo-Inf. 2021;10(5).
3. Chalkiadakis I, Yan H, Peters GW, Shevchenko PV. Infection rate models for COVID-19: Model risk and public health news sentiment exposure adjustments. PLoS ONE. 2021;16(6 June).
4. Gupta R, Pandey G, Pal SK. Comparative analysis of epidemiological models for COVID-19 pandemic predictions. Biostat Epidemiol. 2021;
5. Salehi M, Arashi M, Bekker A, Ferreira J, Chen D-G, Esmaeili F, et al. A Synergetic R-Shiny Portal for Modeling and Tracking of COVID-19 Data. Front Public Health. 2021;8.
6. Shadmi E, Chen Y, Dourado I, Faran-Perach I, Furler J, Hangoma P, et al. Health equity and COVID-19: global perspectives. Int J Equity Health [Internet]. 2020 Jun 26 [cited 2020 Jul 23];19(1):104. Available from: https://doi.org/10.1186/s12939-020-01218-z
7. Dorn A van, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. The Lancet [Internet]. 2020 Apr [cited 2020 Jul 23];395(10232):1243–4. Available from: https://linkinghub.elsevier.com/retrieve/pii/S014067362030893X
8. Jia L, Li K, Jiang Y, Guo X, zhao T. Prediction and analysis of Coronavirus Disease 2019. ArXiv200305447 Q-Bio [Internet]. 2020 Mar 16 [cited 2020 Jul 20]; Available from: http://arxiv.org/abs/2003.05447
9. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med [Internet]. 2020 Apr [cited 2020 Jul 23];26(4):450–2. Available from: http://www.nature.com/articles/s41591-020-0820-9
10. Tuite AR, Fisman DN, Greer AL. Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. Can Med Assoc J [Internet]. 2020 May 11 [cited 2020 Jul 23];192(19):E497–505. Available from: http://www.cmaj.ca/lookup/doi/10.1503/cmaj.200476
11. Català M, Alonso S, Alvarez-Lacalle E, López D, Cardona P-J, Prats C. Empirical model for short-time prediction of COVID-19 spreading. PLoS Comput Biol. 2020;16(12).
12. Mahanty C, Kumar R, Mishra BK, Hemanth DJ, Gupta D, Khanna A. Prediction of COVID-19 active cases using exponential and non-linear growth models. Expert Syst. 2020;
13. Shukor MY, Alam MS. Mathematical Modelling of the Growth of SARS-CoV-2 (COVID-19) and SARS-CoV (SARS) Viruses in Vero E6 Cells. Journalf Environ Microbiol Toxicol. 2020;8(1):1–4.
14. Spanakis M, Zoumpoulakis M, Patelarou AE, Patelarou E, Tzanakis N. Covid-19 epidemic: Comparison of three european countries with different outcome using gompertz function method. Pneumon. 2020;33(2):1–6.
15. Valencia M, Becerra JE, Reyes JC, Castro KG. Assessment of early mitigation measures against COVID-19 in Puerto Rico: March 15-May 15, 2020. PLoS ONE. 2020;15(10 October).
16. Ahmadi A, Fadaei Y, Shirani M, Rahmani F. Modeling and Forecasting Trend of COVID-19 Epidemic in Iran. medRxiv. 2020 Mar 27;2020.03.17.20037671.
17. Liang K. Mathematical model of infection kinetics and its analysis for COVID-19, SARS and MERS. Infect Genet Evol. 2020 Aug 1;82:104306.
18. Sunthornwat R, Areepong Y. Predictive models for the number of cumulative cases for spreading coronavirus disease 2019 in the world. Eng Appl Sci Res. 2021;48(4):432–45.
19. Zwietering MH, Jongenburger I, Rombouts FM, Van't Riet K. Modeling of the bacterial growth curve. Appl Environ Microbiol. 1990;56(6):1875–81.
20. Ricker FJ. Growth rates and models. In: Hoar WS, Brett JR, Randall DJ, editors. Fish Physiology. New York: Academic Press; 1979. p. 677–743. (Bioenergetics and Growth; vol. Volume 8).
21. Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos Trans R Soc. 1825;115:513–85.
22. Richards, F.J. A flexible growth function for empirical use. J Exp Bot. 1959;10:290–300.
23. Schnute J. A versatile growth model with statistically stable parameters. Can J Fish Aquat Sci. 1981;38:1128–40.
24. Baranyi J, Roberts TA. A dynamic approach to predicting bacterial growth in food. Int J Food Microbiol. 1994;23(3–4):277–94.
25. von Bertalanffy L. heoretische Biologie, Zweiter Band: Stoffwechsel,Wachstum. A FranckeAG Verlag, Bern, Switzerland; 1951. 418 p.
26. Anderson MJ, Millar RB, Blom WM, Diebel CE. Nonlinear multivariate models of successional change in community structure using the von Bertalanffy curve. Oecologia. 2005;146(2):279–86.
27. Schofield MR, Barker RJ, Taylor P. Modeling individual specific fish length from capture-recapture data using the von bertalanffy growth curve. Biometrics. 2013;69(4):1012–21.
28. Buchanan RL. Predictive food microbiology. Trends Food Sci Technol. 1993;4(1):6–11.
29. Peleg M, Corradini MG, Normand MD. The logistic (Verhulst) model for sigmoid microbial growth curves revisited. Food Res Int. 2007;40(7):808–18.
30. Huang L, Hwang C-A, Phillips J. Evaluating the effect of temperature on microbial growth rate-the Ratkowsky and a Bělehrádek-Type models. J Food Sci. 2011;76(8):M547–57.
31. Fang T, Gurtler JB, Huang L. Growth Kinetics and Model Comparison of Cronobacter sakazakii in Reconstituted Powdered Infant Formula. J Food Sci. 2012;77(9):E247–55.
32. Halmi MIE, Shukor MS, Johari WLW, Shukor MY. Evaluation of several mathematical models for fitting the growth of the algae *Dunaliella tertiolecta*. Asian J Plant Biol. 2014;2(1):1–6.
33. Huang L. Direct construction of predictive models for describing growth of Salmonella Enteritidis in liquid eggs - A one-step approach. Food Control. 2015;57:76–81.
34. Morgan PH, Mercer LP, Flodin NW. General model for nutritional responses of higher organisms. Proc Natl Acad Sci. 1975 Nov 1;72(11):4327–31.
35. Khamis A, Ismail Z, Haron K, Mohammed AT. Nonlinear Growth Models for Modeling Oil Palm Yield Growth. J Math Stat. 2005 Sep 30;1(3):225–33.
36. Topal M, Bolukbasi ŞC. Comparison of nonlinear growth curve models in broiler chickens. J Appl Anim Res. 2008 Dec 1;34(2):149–52.
37. Tariq M, Iqbal F, Eyduran E, Bajwa M, Huma Z, Waheed A. Comparison of non-linear functions to describe the growth in Mengali sheep breed of Balochistan. Pak J Zool. 2013 Jun 1;45:661–5.
38. Wijeratne AW, Karunaratne JA. Morgan-Mercer-Flodin model for long term trend analysis of currency exchange rates of some selected countries. Int J Bus Excell. 2013 Dec 2;7(1):76–87.
39. Augustine A, Imelda J, Paulraj R, David NS. Growth kinetic profiles of *Aspergillus niger* S14 a mangrove isolate and *Aspergillus oryzae* NCIM 1212 in solid state fermentation. Indian J Fish. 2015;62(3):100–6.
40. Kemper CM. Growth and development of the brush-tailed rabbit-rat (*Conilurus penicillatus*), a threatened tree-rat from northern Australia. Aust Mammal. 2020 Jun 5;
41. Aisami A, Shukor MYA. Predictive Mathematical Modelling of the Total Number of COVID-19 Cases for the Kingdom of Saudi Arabia. J Environ Microbiol Toxicol. 2020 Jul 31;8(1):11–5.
42. Shukor MYA, Sabo IA, Yahuza S, Dan-Iya BI, Wada SA. Prediction of Cumulative Death Cases in The United States Due to COVID-19 Using Mathematical Models. J Environ Microbiol Toxicol. 2020 Jul 31;8(1):37–41.
43. Uba G, Yakasai HM, Abubakar A, Shukor MYA. Predictive Mathematical Modelling of the Total Number of COVID-19 Cases for Brazil. J Environ Microbiol Toxicol. 2020 Jul 31;8(1):16–20.
44. Uba G, Yakasai HM, Abubakar A, Shukor MYY. Prediction of Cumulative Death Cases in Brazil Due to Covid-19 Using Mathematical Models. Bull Environ Sci Sustain Manag. 2020 Jul 31;4(1):13–9.
45. Yahuza S, Sabo IA, Dan-Iya BI, Shukor MYY. Prediction of Cumulative Death Cases in Nigeria Due to COVID-19 Using Mathematical Models. Bull Environ Sci Sustain Manag. 2020 Jul 31;4(1):20–4.
46. Yakasai HM, Shukor MYA. Predictive Mathematical Modelling of the Total Number of COVID-19 Cases for The United States. Bioremediation Sci Technol Res. 2020 Jul 31;8(1):11–6.

47. Umar AM, Shukor MYA. Predictive Mathematical Modelling of the Total Number of COVID-19 Cases for Indonesia. J Environ Microbiol Toxicol. 2020 Aug 1;8(1):27–31.
48. Nyoni T. Prediction of daily new Covid -19 cases in Indonesia using artificial neural networks. Int J Adv Res Innov Ideas Educ. 2020 Dec 15;6(6):2174–87.
49. Supriatna AK, Husniah H. Modeling Covid-19 cummulative data in Indonesia using Morgan-Mercer-Flodin growth equation. In: Proceedings of the International Conference on Industrial Engineering and Operations Management. Harare, Zimbabwe; 2020.
50. Uba G, Yakasai HM, Abubakar A, Shukor MYY. Prediction of Cumulative Death Cases in Brazil Due to Covid-19 Using Mathematical Models. Bull Environ Sci Sustain Manag. 2020 Jul 31;4(1):13–9.
51. Yahuza S, Sabo IA, Dan-Iya BI, Shukor MYY. Prediction of Cumulative Death Cases in Nigeria Due to COVID-19 Using Mathematical Models. Bull Environ Sci Sustain Manag. 2020 Jul 31;4(1):20–4.
52. Motulsky HJ, Brown RE. Detecting outliers when fitting data with nonlinear regression - A new method based on robust nonlinear regression and the false discovery rate. BMC Bioinformatics. 2006;7.
53. Bolster CH, Hornberger GM. On the use of linearized langmuir equations. Soil Sci Soc Am J. 2007;71(6):1796–806.
54. El-Khaiary MI. Least-squares regression of adsorption equilibrium data: Comparing the options. J Hazard Mater. 2008;158(1):73–87.
55. Fong Y, Wakefield J, De R, Frahm N. A Robust Bayesian Random Effects Model for Nonlinear Calibration Problems. Biometrics. 2012;68(4):1103–12.
56. Hu W, Xie J, Chau HW, Si BC. Evaluation of parameter uncertainties in nonlinear regression using Microsoft Excel Spreadsheet. Environ Syst Res. 2015 Mar 24;4(1):4.
57. Rout PR, Bhunia P, Dash RR. Evaluation of kinetic and statistical models for predicting breakthrough curves of phosphate removal using dolochar-packed columns. J Water Process Eng. 2017 Jun 1;17:168–80.
58. Grubbs F. Procedures for detecting outlying observations in samples. Technometrics. 1969;11(1):1–21.
59. Kolmogorov A. Sulla determinazione empirica di una legge di distribuzione. G Dell' Ist Ital Degli Attuari. 1933;4:83–91.
60. Smirnov N. Table for estimating the goodness of fit of empirical distributions. Ann Math Stat. 1948;19:279–81.
61. Royston P. Wilks-Shapiro algorithm. Appl Stat. 1995;44(4):R94.
62. Stephens MA. Tests of fit for the logistic distribution based on the empirical distribution function. Biometrika. 1979 Dec 1;66(3):591–5.
63. D'Agostino RB. Tests for Normal Distribution. In: D'Agostino RB, Stephens MA, editors. Goodness-Of-Fit Techniques. Marcel Dekker; 1986.
64. Motulsky HJ, Ransnas LA. Fitting curves to data using nonlinear regression: a practical and nonmathematical review. FASEB J Off Publ Fed Am Soc Exp Biol. 1987;1(5):365–74.
65. López S, Prieto M, Dijkstra J, Dhanoa MS, France J. Statistical evaluation of mathematical models for microbial growth. Int J Food Microbiol. 2004;96(3):289–300.
66. Barnett V, Lewis T. Outliers in Statistical Data. 3rd ed. Chichester ; New York: Wiley; 1994. 604 p.
67. Tietjen GL, Moore RH. Some Grubbs-Type Statistics for the Detection of Several Outliers. Technometrics. 1972;14(3):583–97.
68. Pestiaux P, Pic C. The limitations of the cochran and grubbs outlier tests in round robin testing. SAE Tech Pap. 2004;
69. Jain RB. A recursive version of Grubbs' test for detecting multiple outliers in environmental and chemical data. Clin Biochem. 2010;43(12):1030–3.
70. Cohn TA, England JF, Berenbrock CE, Mason RR, Stedinger JR, Lamontagne JR. A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series. Water Resour Res. 2013;49(8):5047–58.
71. Urvoy M, Autrusseau F. Application of grubbs' test for outliers to the detection of watermarks. In: IH and MMSec 2014 - Proceedings of the 2014 ACM Information Hiding and Multimedia Security Workshop. 2014. p. 49–60.
72. Adikaram KKLB, Hussein MA, Effenberger M, Becker T. Data transformation technique to improve the outlier detection power of grubbs' test for data expected to follow linear relation. J Appl Math. 2015;2015.
73. Miller JN. Using the Grubbs and Cochran tests to identify outliers. Anal Methods. 2015;7(19):7948–50.
74. Lamontagne JR, Stedinger JR, Yu X, Whealton CA, Xu Z. Robust flood frequency analysis: Performance of EMA with multiple Grubbs-Beck outlier tests. Water Resour Res. 2016;52(4):3068–84.
75. Shiryaeva LK. On distribution of Grubbs' statistics in case of normal sample with outlier. Russ Math. 2017;61(4):72–88.
76. Aslam M. Introducing Grubbs's test for detecting outliers under neutrosophic statistics – An application to medical data. J King Saud Univ - Sci. 2020;32(6):2696–700.
77. Ding K, Zhang J, Ding H, Liu Y, Chen F, Li Y. Fault detection of photovoltaic array based on Grubbs criterion and local outlier factor. IET Renew Power Gener. 2020;14(4).
78. Rosner B. Fundamentals of biostatistics. 7th ed. Boston: Brooks/Cole; 2011.
79. Bendre SM, Kale BK. Masking effect on tests for outliers in normal samples. Biometrika. 1987 Dec 1;74(4):891–6.
80. Fischer H. A History of the Central Limit Theorem: From Classical to Modern Probability Theory [Internet]. New York: Springer-Verlag; 2011 [cited 2021 Jan 4]. (Sources and Studies in the History of Mathematics and Physical Sciences). Available from: https://www.springer.com/gp/book/9780387878560
81. Pandoo P. Which normality test is more appropriate on residuals with sample size 1000? [Internet]. ResearchGate. 2014 [cited 2021 Jan 4]. Available from: https://www.researchgate.net/post/Which-normality-test-is-more-appropriate-on-residuals-with-sample-size-1000. Accessed date 4th January 2021.
82. Legendre P. Comparison of permutation methods for the partial correlation and partial Mantel tests. J Stat Comput Simul. 2000;67(1):37–73.
83. Neumann MB, Gujer W. Underestimation of uncertainty in statistical regression of environmental models: Influence of model structure uncertainty. Environ Sci Technol. 2008;42(11):4037–43.
84. Greghi GF, Barcelos B, Saran Netto A, Vilela FG, Rodrigues PHM, Marino CT. Contribution of citrus pulp and soybean hulls for silage quality of brewery waste [Contribuição da inclusão de polpa cítrica e casca de soja para a qualidade da silagem de resíduo úmido de cervejaria]. Arq Bras Med Vet E Zootec. 2014;66(1):277–83.
85. Kupka TW, Nowak J, Szczesio A, Kopacz K, Fronczek-Wojciechowska M, Sokołowski J. Effect of addition of antimicrobial triclosan on selected properties of water-activated glass ionomer cement. J Stomatol. 2016;69(5):492–500.
86. MacKinnon MC, Pearl DL, Carson CA, Parmley EJ, McEwen SA. A comparison of modelling options to assess annual variation in susceptibility of generic *Escherichia coli* isolates to ceftiofur, ampicillin and nalidixic acid from retail chicken meat in Canada. Prev Vet Med. 2018;160:123–35.
87. Naji O, Bowtell L, Al-Juboori RA, Aravinthan V, Ghaffour N. Effect of air gap membrane distillation parameters on the removal of fluoride from synthetic water. Desalination Water Treat. 2018;124:11–20.
88. Castro VS, Rosario DKA, Mutz YS, Paletta ACC, Figueiredo EES, Conte-Junior CA. Modelling inactivation of wild-type and clinical *Escherichia coli* O26 strains using UV-C and thermal treatment and subsequent persistence in simulated gastric fluid. J Appl Microbiol. 2019;127(5):1564–75.
89. Deveci C, Tuzuner T, Cinar C, Odabas M, Buruk C. Short-term antibacterial activity and compressive strength of biodentine containing chlorhexidine/cetirimide mixtures. Niger J Clin Pract. 2019;22(2):227–31.
90. Potysz A, Pędziwiatr A, Hedwig S, Lenz M. Bioleaching and toxicity of metallurgical wastes. J Environ Chem Eng. 2020;8(6).