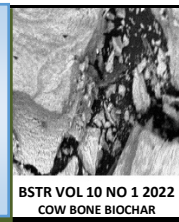


BIOREMEDIATION SCIENCE AND TECHNOLOGY RESEARCH

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



Test of the Randomness of Residuals and Detection of Potential Outliers for the Modified Gompertz Model Used in the Fitting of the Growth of *Shigella flexneri*

Garba Uba^{1*} and Bilal Ibrahim Dan-Iya²

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

²Pharmacy Technician Department, College of Health Sciences and Technology, Kano Nigeria.

*Corresponding author:

Garba Uba,

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

Email: garbauba@jigpoly.edu.ng

HISTORY

Received: 14th April 2022
Received in revised form: 20th June 2022
Accepted: 14th July 2022

KEYWORDS

Wald-Wolfowitz runs test
Modified Gompertz model
Residual data
Shigella flexneri
Grubb's test

ABSTRACT

The formulation of hypotheses and the recommendation of experiments as the subsequent stages of the research process are both brought about as a result of the utilization of complicated computer models that make it possible to represent intricate biological processes. Because these systems rely on random data, this is a necessity for all parametric statistical assessment procedures. When the diagnostic tests reveal that the residuals make up a pattern, there are a few different treatment choices available to choose from. Two of these alternatives include running a nonparametric analysis or switching to a new model. In this study, we use the Wald-Wolfowitz runs test as a statistical diagnosis tool to determine whether or not the randomization conditions have been met. The runs test found that there were 5 total runs, although the randomness assumption predicted 7.46 runs. The null hypothesis is not rejected since the p-value is greater than 0.05; this suggests that there is no convincing evidence of the residuals' non-randomness; rather, the residuals represent noise. In addition, the Grubb's outlier test shows no indication of an outlier, further corroborate the scenario of the adequacy of the modified Gompertz model used in the fitting of the growth of *Shigella flexneri*.

INTRODUCTION

One way to conceptualize a biological system is as a collection of several cellular compartments (such as cell types), each of which is specialized for a certain biological purpose (e.g. white and red blood cells have very different commitments). An object is an elemental unit of some kind that can be examined, but the inside structure of the thing is either unknown or it does not exist. The scale that is applied to the presentation of the system will be decided based on the elemental unit that is selected. The availability of data that represent a variety of biological states and processes, as well as the temporal dependencies of those activities, makes it feasible to conduct research on biological systems at a wide variety of various levels of organization. These levels start at the molecular level and go all the way up to the population level. A model is a representation of a system that may be decoded or comprehended by researchers in general. Models are often used in computer simulations. A model is a description of a system that focuses on the components that make up the system as well as the interactions that occur between those

components [1–7]. The changes that have taken place in the fields of biotechnology and information technology have both sped up the process of recognizing the knowledge that is held inside biological systems, which has resulted in a shortened amount of time needed to complete the procedure. Because of these breakthroughs, the approaches taken in the field of biomedicine to research, development, and application are undergoing significant transformation. With the inclusion of clinical data to biological data, it is feasible to offer detailed descriptions of both healthy and sick states, as well as the evolution of sickness and the body's reaction to treatments. This is made possible by the combination of the two sets of data [8–14].

The availability of data that represent a variety of biological states and processes, as well as the temporal dependencies of those activities, makes it feasible to conduct research on biological systems at a wide variety of various levels of organization. These levels start at the molecular level and go all the way up to the population level. In the field of high-throughput genomics and proteomics research, the application of

mathematical and computational models to help in the comprehension of biological data is becoming increasingly common. The formulation of hypotheses and the recommendation of experiments as the subsequent stages of the research process are both brought about as a result of the utilization of complicated computer models that make it possible to represent intricate biological processes.

Computational models are now employing knowledge discovery strategies in order to make use of the vast amounts of data that are recorded in biomedical databases. This is being done in order to make the most of the information that is available. The availability of data that represent a variety of biological states and processes, as well as the temporal dependencies of those activities, makes it feasible to conduct research on biological systems at a wide variety of various levels of organization. These levels start at the molecular level and go all the way up to the population level. In many instances, the relationship of observed phenomenon to time or concentration can be mathematically modelled using software via least square methods often used in nonlinear regression [8–14].

In spite of this, the residuals of the curve in a nonlinear regression need to have a natural dispersion. This is in contrast to the usual least square's technique, which calls for the residuals of a linear regression to have a normal distribution. This is due to the fact that the principle of least squares forms the foundation of the traditional least square's technique. The residuals, which are significantly more important, have to be random and have the same variance (homoscedastic distribution). It is possible to determine whether or not the randomization procedure was effective by employing the Wald-Wolfowitz runs test. [15]. On the other hand, the residuals of the curve in a nonlinear regression need to have a natural dispersion, whereas in a linear regression, the residues need to have a normal distribution in order for the typical least squares approach to work well. More crucially, the residuals must be random and have the same variance (homoscedastic distribution) (homoscedastic distribution). The residuals must also be random and outliers' absence. The Wald-Wolfowitz runs test is used to establish whether or not the residuals for the modified Gompertz model used in the Fitting of the growth of *Shigella flexneri* is random whilst the Grubb's test is applied to detect the presence of outliers.

METHODOLOGY

One of the utilities of residual information is that it can be utilized to measure the accuracy of any model fitting a curve in nonlinear regression can be achieved by evaluating (D'Agostino, 1986). In the statistical meaning, residual data is calculated by the difference between observed and predicted data, the latter obtained using suitable model and usually carried out using nonlinear regression (Eqn. 1);

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

where y_i is the i^{th} response from a particular data and x_i is the vector of descriptive variables to each set at the i^{th} observation which corresponds to values from the residual data from the modified Gompertz model used in the Fitting of the growth of *Shigella flexneri* [16].

Grubbs' Statistic

The test is a statistical test used to discover outliers in a univariate data set that is believed to have a Gaussian or normal distribution. Grubb's test assumes that the data is regularly distributed. The test is used to discover outliers in a univariate context [17]. The test can be utilized to the maximal or minimal examined data from a Student's t distribution (Eq. 1) and to test for both data instantaneously (Eqn. 2).

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

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

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

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

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

The ROUT method can be employed in the event that there is more than one outlier [18]. The False Discovery Rate is the foundation of the approach (FDR). Q, a probability of (incorrectly) recognising one or more outliers must be explicitly specified. It is the highest desired FDR. Q is fairly comparable to alpha in the absence of outliers. Assumption that all data has a Gaussian distribution is mandatory.

Runs test

When doing a nonlinear regression, it is necessary for the curve's residuals to have a natural distribution. This is in contrast to the requirements of the least squares approach, which ask for the residues to have a regular distribution. Those requirements may be found here. In addition to this, residuals are needed to have the same variance and be random at the same time (homoscedastic distribution). The Wald-Wolfowitz test is utilised for the purpose of detecting whether or not randomization has been accomplished. Because biological systems are intrinsically unexpected, the model can be depended upon to be correct in terms of statistics. [19–21].

This test was applied to the regression residuals in order to find unpredictability in the residuals. The number of sign runs is often stated as a percentage of the greatest number possible. The runs test examines the sequence of residuals, of which they are composed of positive and negative values. A successful run, after running the test, is often represented by the presence of an alternating or adequately balanced number of positive and negative residual values. The runs test computes the likelihood of the residuals data having too many or too few runs of sign (Eq. 3). Too few runs may suggest a clustering of residuals with the same sign or the existence of systematic bias, whereas too many of a run sign may identify the presence of negative serial correlation [15,22].

The test statistic is

H_0 = the sequence was produced randomly
 H_a = the sequence was not produced randomly

$$Z = \frac{R - \bar{R}}{sR} \quad (\text{Eqn. 3})$$

Where Z is the test statistic, \bar{R} indicates the anticipated number of runs, sR is the standard deviation of the runs and R is the observed number of runs and (Eq. 4). The calculation of the respective values of \bar{R} and sR (n_1 is positive while n_2 is negative signs) is as follows.

$$\bar{R} = \frac{2n_1 \cdot n_2 + 1}{n_1 + n_2} \quad (\text{Eqn. 4})$$

$$s^2 R = \frac{2n_1 \cdot n_2 (2n_1 \cdot n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)} \quad (\text{Eqn. 5})$$

As an example

Test statistic: $Z = 3.0$

Significance level: $\alpha = 0.05$

Critical value (upper tail): $Z_{1-\alpha/2} = 1.96$

Critical region: Reject H_0 if $|Z| > 1.96$

If the test statistical value (Z) is greater than the critical value, then the rejection of the null hypothesis at the 0.05 significance level shows that the sequence was not generated randomly.

RESULTS AND DISCUSSION

It might be difficult to locate an appropriate model for biological and even chemical processes.

The process of modelling is challenging in and of itself, and mistakes are not an extremely uncommon occurrence. The modelling technique is in and of itself a process that adheres to a loosely formalized set of guidelines. The process is based on the completion of four large phases. The first step is to get a solid grasp of the issue at hand, which involves precisely defining the queries that are posed to the model. The second phase is to develop a strategy for addressing the problem, which entails outlining a sequence of activities that need to be carried out in order to locate an accurate model of the system that is the subject of the investigations. In this step, you will acquire knowledge and data from specialists in the field as well as from published works, model structure, model hypothesis, conceptual model, appropriate mathematical formalism selection, solving the formal model, obtaining the results, checking to see if the results of the model match the data that is available, and other similar tasks.

The third phase is to put the plan into action, which involves doing the processes from the previous two steps, determining whether or not the solution is accurate, and finally refining the model. This last step is a significant test to examine the hypothesis that was developed prior to the setting of the model. Ultimately all models will need to be subjected to mathematical curve fitting and this is where nonlinear regression comes into place. Data known as residuals play a crucial part in the statistical analysis of nonlinear regression, which is done using nonlinear regression. The discrepancy between the data that was actually collected and the data that was expected to be collected is denoted by the residuals.

The term "residuals" refers to the disparities that exist between the values that are predicted by a mathematical model and the values that are actually seen in the data. These disparities may be seen when comparing the model's predictions to the data. The residuals need to be analyzed statistically in order to determine whether or not they are sufficiently random, do not include any outliers, adhere to the normal distribution, and do not exhibit autocorrelation. This may be done by determining whether or not they do not have any outliers. It is common practice to display the data on residues in the form of positive and negative values. This is an essential component for establishing that the data are balanced, and it can be visually observed before any tests are carried out.

When carrying out nonlinear regression, it is standard practice to disregard the findings of any residual tests that are carried out. As a general rule, the quality of a model is deemed to be worse when there is a bigger difference between the values that were predicted and those that were actually observed. This is because there is a lesser degree of correlation between the two sets of data, which explains why this is the case [23]. The residuals for the modified Gompertz model used in the Fitting of the growth of *Shigella flexneri* are shown in **Table 1**.

Table 1. Residuals for the modified Gompertz model used in the Fitting of the growth of *Shigella flexneri*.

Time (h)	Residual
0	0.036517
1	-0.02864
2	-0.00736
3	0.065861
4	-0.02456
5	-0.05027
6	-0.02806
7	-0.01429
8	0.006728
9	0.017126
10	0.013392
11	0.017306
12	0.015911

Grubbs' test indicated that there was no hint of an outlier when applied to previously released data. This indicates that the model was able to properly describe the data. When attempting to fit a nonlinear curve, it is easy to introduce considerable inaccuracies if either the mean or a single data point from a triple is warped. Both of these circumstances get the same result. The Grubbs test is capable of detecting a single abnormality across any specified time period. It is critical to seek for and delete any outlying data points when fitting curves [24–30]. Because this specific data point was considered to be an outlier, it was eliminated from the collection, and the analysis was repeated until no further outliers were found. Because the test consistently identifies the great majority of points as outliers, sample sizes of six or less are not recommended. Furthermore, repeating the test numerous times might vary the likelihood that it will detect something.

The Grubbs' test statistic zeroes in on the sample value with the largest absolute deviation from the sample mean, as measured by the sample's standard deviation, to select the winner. If the test statistic g yields a result that is bigger than the critical value, the result is referred to as an outlier. This is due to the fact that the critical value is the lowest value that can be tolerated. The results of Grubbs's test suggested the absence of an outlier (**Table 2**). This suggests that the model was adequate.

Table 2. Detection of outlier for the residuals for the modified Gompertz model used in the Fitting of the growth of *Shigella flexneri*.

Row	Value	Z	Significant Outlier?
1	0.0365	1.1207	
2	-0.0286	0.9652	
3	-0.0074	0.2842	
4	0.0659	2.0601	Furthest from the rest, but not a significant outlier ($P > 0.05$).
5	-0.0246	0.8347	
6	-0.0503	1.6578	
7	-0.0281	0.9466	
8	-0.0143	0.5060	
9	0.0067	0.1670	
10	0.0171	0.4999	
11	0.0134	0.3803	
12	0.0173	0.5056	
13	0.0159	0.4610	

A possible outlier is an extreme data point that the investigator believes is implausible because it does not meet a number of specific characteristics. An outlier is a figure that stands out as being notably different from the rest of the data in a sample. For example, an outlier is defined as a maximum that is statistically considerably larger than the maximum distribution anticipated by the population model employed in engineering. This criterion is used to assess if the maximum is an outlier or not. The Chauvenet's criterion, the 3-sigma criterion, and the Z-score may all be used to identify potential measurement outliers. The Z-score is commonly used in chemometrics in conjunction with the 3-sigma criteria.

A boxplot is a simple way for identifying potential outliers in measurement data [26]. A statistical test is recommended for evaluating whether or not a data set contains an outlier, despite the fact that the methods in question are simple, quick, and can pass visual inspections. The Dixon's Q-test and the Grubbs' ESD-test are two specific tests that may be used to identify whether or not a person is an outlier. The specific value of the predicted number of outliers, denoted by k , needs to be provided before the Grubbs test may be accepted as valid. This is the most significant restriction of the exam. If k is not accurately reflected in the test, it is quite possible that the findings of the test will be changed. Rosner's generalised Extreme Studentized Deviate, also known as the ESD-test, or the ROUT methodology may be utilised in circumstances in which there are several outliers, or the precise number of outliers cannot be determined. Both of these methodologies are usually referred to as the ESD-test [18] are recommended [31]. Of the two, the ROUT method, which combines robust regression and outlier removal is increasingly being employed in removal of multiple outliers [24,32–35]

The runs test found that there were 5 total runs, although the randomness assumption predicted 7.46 runs (**Table 3**). This implies that the residuals collection only includes runs that are marginally relevant. The Z-value is the number of standard errors by which the actual number of runs differs from the projected number of runs, and the p-value denotes the degree to which this Z-value is significant. The interpretation is the same as for any other statistic that includes p-values. If the p-value is less than 0.05, which is the threshold at which the null hypothesis may be rejected, it is possible to conclude that the residuals are not truly random. The null hypothesis is not rejected since the p-value is marginally greater than 0.05; this suggests that there is no convincing evidence of the residuals' non-randomness; rather, the residuals represent noise. When there are a large number of occurrences of a certain run sign, this may indicate a negative serial correlation; when there are a small number of runs, this

may indicate a clustering of residuals with the same sign or the presence of a systematic bias [22]. When a specific model is used, the runs test can detect a systematic deviation from the curve, such as an overestimation or underestimation of the sections. This is performed by comparing the actual and projected values.

Table 3. Runs test for randomness.

Runs test	Residual data set
R=	5
n0=	6
n1=	7
n=	13
E(R)=	7.46
Var(R)=	2.94
StDev(R)=	1.71
Z=	-1.44
p-value=	0.0756

This may be performed by comparing the model's predictions to the actual values. The runs test is used to detect if there are an excessive number of sign runs or whether there are insufficient runs overall. The runs test was conducted to the regression residuals to see whether or not there was evidence of nonrandomness. It is feasible to build a model with an ordered variance of the curve that is more or less than the estimate. This is only one of the numerous ways this is possible. As part of the run test, which is used to establish whether or not a chemical poses a risk to human health, a comparison is done between a drug's often negative sequence of residues and its generally positive sequence.

A change or combination of movements between negative and positive residual values usually separates a notable occurrence from other probable outcomes. A notable conclusion is typically distinguished by a metamorphosis or succession of shifts [15]. A common practice in this field is to employ the greatest number of indicators that can be counted. The run's test is used to assess whether a big number of sign passes are more likely, or a low number of sign passes are more likely. Negative serial correlation may be implied by run signs, but it is also conceivable that residues are associated with the same sign or that systemic biases are impacting the results. [22].

When testing time-series regression models to see if autocorrelation exists, it is common practise to use the run technique. Run-time testing, according to the results of Monte Carlo simulation studies, causes unequal error rates in both tails of the distribution. This finding implies that run-time autocorrelation research may be unstable, and it forecasts that the Durbin-Watson methodology will become the most often utilised strategy to analysing autocorrelation in the near future [36]. It has been established that the technique used in this analysis, which was taken from previous research on the unpredictability of residuals, is dependable. Modeling the growth curve of algae, for example, using the Baranyi-Roberts model, which establishes statistical sufficiency [37], *Moraxella* sp. B on monobromoacetic acid (MBA) [20] and the Buchanan-three-phase model used in the fitting the growth of *Paracoccus* sp. SKG on acetonitrile [38]. For lead (II) absorption by alginate gel bead, the runs tests on the residuals for the Sips and Freundlich models were found to be sufficient [39].

It has been established that the technique used in this analysis, which was taken from previous research on the unpredictability of residuals, is dependable. Modeling the growth curve of algae, for example, using the Baranyi-Roberts model, which establishes statistical sufficiency [40]. In the body of

academic research, different applications of the runs test of residual may be found for the purpose of evaluating the validity of the nonlinear regression [41–45].

CONCLUSION

In the framework of this investigation, the Wald-Wolfowitz runs test was performed. The observations led to the conclusion that the residual series had an acceptable number of runs; this was the result of the observations. The p-value indicates how significant this z-value is, whereas the z-value discloses the number of typical errors that were detected when the number of runs actually found exceeded the number of runs predicted to be found. The runs test found that there were 5 total runs, although the randomness assumption predicted 7.46 runs. The null hypothesis is not rejected since the p-value is greater than 0.05; this suggests that there is no convincing evidence of the residuals' non-randomness; rather, the residuals represent noise. In addition, the Grubb's outlier test shows no indication of an outlier, further corroborate the scenario of the adequacy of the model used in fitting the data.

REFERENCES

- Brul S, Mensonides FIC, Hellingwerf KJ, Teixeira de Mattos MJ. Microbial systems biology: New frontiers open to predictive microbiology. *Int J Food Microbiol.* 2008;128(1):16–21.
- McMeekin T, Bowman J, McQuestin O, Mellefont L, Ross T, Tamplin M. The future of predictive microbiology: Strategic research, innovative applications and great expectations. *Int J Food Microbiol.* 2008;128(1):2–9.
- Membré JM, Lambert RJW. Application of predictive modelling techniques in industry: From food design up to risk assessment. *Int J Food Microbiol.* 2008;128(1):10–5.
- Velugoti PR, Bohra LK, Juneja VK, Huang L, Wesseling AL, Subbiah J, et al. Dynamic model for predicting growth of *Salmonella* spp. in ground sterile pork. *Food Microbiol.* 2011;28(4):796–803.
- Yilmaz MT. Identifiability of Baranyi model and comparison with empirical models in predicting effect of essential oils on growth of *Salmonella* Typhimurium in rainbow trout stored under aerobic, modified atmosphere and vacuum packed conditions. *Afr J Biotechnol.* 2011;10(38):7468–79.
- Xiao H, Li B, Zheng X, Ou J. Predicting the growth of bacteria in chilled beef based on Artificial Neural Networks. In 2012. p. 449–52.
- Zhou K, George SM, Li PL, Baranyi J. Effect of periodic fluctuation in the osmotic environment on the adaptation of *Salmonella*. *Food Microbiol.* 2012;30(1):298–302.
- Caleb OJ, Mahajan PV, Al-Said FAJ, Opara UL. Modified atmosphere packaging technology of fresh and fresh-cut produce and the microbial consequences-A Review. *Food Bioprocess Technol.* 2013;6(2):303–29.
- Giltner CL, Saeki S, Bobenchik AM, Humphries RM. Rapid detection of campylobacter antigen by enzyme immunoassay leads to increased positivity rates. *J Clin Microbiol.* 2013;51(2):618–20.
- Koziel M a, Corcoran D b, O'Callaghan I b, Sleator RD a, Lucey B a b. Validation of the EntericBio Panel II® multiplex polymerase chain reaction system for detection of *Campylobacter* spp., *Salmonella* spp., *Shigella* spp., and verotoxigenic *E. coli* for use in a clinical diagnostic setting. *Diagn Microbiol Infect Dis.* 2013;75(1):46–9.
- Lobacz A, Kowalik J, Tarczynska A. Modeling the growth of *Listeria monocytogenes* in mold-ripened cheeses. *J Dairy Sci.* 2013;96(6):3449–60.
- Ioannidis AG, Kogkaki EA, Natskoulis PI, Nychas GJE, Panagou EZ. Modelling the influence of temperature, water activity and sodium metabisulphite on the growth and OTA production of *Aspergillus carbonarius* isolated from Greek wine grapes. *Food Microbiol.* 2015;49:12–22.
- Mukherjee A, Chettri B, Langpoklakpam JS, Basak P, Prasad A, Mukherjee AK, et al. Bioinformatic Approaches Including Predictive Metagenomic Profiling Reveal Characteristics of Bacterial Response to Petroleum Hydrocarbon Contamination in Diverse Environments. *Sci Rep.* 2017 Apr 24;7(1):1–22.
- Altuntas S, Cinar A, Altuntas V. Modelling of *Listeria monocytogenes* growth and survival in presence of royal jelly, a promising anti-biofilm agent. *J Food Nutr Res.* 2020 Jan;59(1):7–15.
- 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.
- Uba G, Yakubu A, Baba AM. The Effect of the plant *Adiantum philippense* Extracts on Biofilms Formation and Adhesion to *Shigella flexneri*: A Predictive Modelling Approach. *J Biochem Microbiol Biotechnol.* 2021 Dec 31;9(2):34–9.
- Grubbs F. Procedures for detecting outlying observations in samples. *Technometrics.* 1969;11(1):1–21.
- 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.
- Lazaridis NK, Asouhidou DD. Kinetics of sorptive removal of chromium(VI) from aqueous solutions by calcined Mg–Al–CO₃ hydrotalcite. *Water Res.* 2003 Jul 1;37(12):2875–82.
- Sabullah MK, Shukor MS, Masdor NA, Shamaan NA, Shukor MY. Test of randomness of residuals for the Buchanan-three-phase model used in the fitting the growth of *Moraxella* sp. B on monobromoacetic acid (MBA). *Bull Environ Sci Manag.* 2015;3(1):13–5.
- Uba G, Mansur A, Manogaran M, Shukor MY. Runs Test for the Residuals of The Morgan-Mercer-Flodin (MMF) Model Used for Modelling the Total Number of Covid-19 Cases for Brazil. *Bull Environ Sci Sustain Manag.* 2021 Jul 31;5(1):7–11.
- Draper NR, Smith H. *Applied Regression Analysis.* Wiley, New York; 1981.
- 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.
- Tietjen GL, Moore RH. Some Grubbs-Type Statistics for the Detection of Several Outliers. *Technometrics.* 1972;14(3):583–97.
- Bendre SM, Kale BK. Masking effect on tests for outliers in normal samples. *Biometrika.* 1987 Dec 1;74(4):891–6.
- Barnett V, Lewis T. *Outliers in Statistical Data.* 3rd ed. Chichester ; New York: Wiley; 1994. 604 p.
- Jain RB. A recursive version of Grubbs' test for detecting multiple outliers in environmental and chemical data. *Clin Biochem.* 2010;43(12):1030–3.
- Fong Y, Wakefield J, De R, Frahm N. A Robust Bayesian Random Effects Model for Nonlinear Calibration Problems. *Biometrics.* 2012;68(4):1103–12.
- 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.
- Miller JN. Using the Grubbs and Cochran tests to identify outliers. *Anal Methods.* 2015;7(19):7948–50.
- Rosner B. *Fundamentals of biostatistics.* 7th ed. Boston: Brooks/Cole; 2011.
- Horn R, Stewart AJ, Jackson KV, Dryburgh EL, Medina-Torres CE, Bertin FR. Clinical implications of using adrenocorticotrophic hormone diagnostic cutoffs or reference intervals to diagnose pituitary pars intermedia dysfunction in mature horses. *J Vet Intern Med.* 2021;35(1):560–70.
- Shukor MY. Outlier and Normality Testing of the Residuals from the Carreau-Yasuda Model in Fitting the Rheological Behavior of the Non-Newtonian fluid TF2N. *Bioremediation Sci Technol Res.* 2021 Jul 31;9(1):20–6.
- Uba G, Zandam ND, Mansur A, Shukor MY. 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. *Bioremediation Sci Technol Res.* 2021 Jul 31;9(1):13–9.
- Li J. A Practical $\sqrt{N} \log N$ Outlier Removal Method for Correspondence-Based Point Cloud Registration. *IEEE Trans Pattern Anal Mach Intell.* 2022 Aug;44(8):3926–39.

36. Huitema BE, McKean JW, Zhao J. The runs test for autocorrelated errors: unacceptable properties. *J Educ Behav Stat.* 1996;21(4):390–404.
37. Halimi 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.
38. Gunasekaran B, Shukor MS, Masdor NA, Shamaan NA, Shukor MY. Test of randomness of residuals for the Buchanan-three-phase model used in the fitting the growth of *Paracoccus* sp. SKG on acetonitrile. *J Environ Bioremediation Toxicol.* 2015;3(1):12–4.
39. Cataldo S, Gianguzza A, Merli M, Muratore N, Piazzese D, Turco Liveri ML. Experimental and robust modeling approach for lead(II) uptake by alginate gel beads: Influence of the ionic strength and medium composition. *J Colloid Interface Sci.* 2014 Nov 15;434:77–88.
40. Umar G. Test of the Randomness of Residuals for the Pseudo-1st Order Kinetic Modelling of Adsorption of the Brominated Flame Retardant 4-bromodiphenyl Ether onto Biochar-immobilized *Sphingomonas* sp. *Bioremediation Sci Technol Res.* 2021 Jul 31;9(1):27–31.
41. Cooper SM, Baker JS, Eaton ZE, Matthews N. A simple multistage field test for the prediction of anaerobic capacity in female games players. *Br J Sports Med.* 2004;38(6):784–9.
42. Worthington AC, Higgs H. Efficiency in the Australian stock market, 1875-2006: A note on extreme long-run random walk behaviour. *Appl Econ Lett.* 2009;16(3):301–6.
43. Abu GA, Abachi PT, Oloja-Ojabo ED. Long-run relationship between agricultural crop prices and supply response in Benue State, Nigeria: 1990-2010. *Eur J Soc Sci.* 2011;24(4):565–75.
44. Burns RD, Hannon JC, Brusseau TA, Eisenman PA, Shultz BB, Saint-Maurice PF, et al. Development of an aerobic capacity prediction model from one-mile run/walk performance in adolescents aged 13-16 years. *J Sports Sci.* 2016;34(1):18–26.
45. Gardiner SK, Mansberger SL. Effect of restricting perimeter testing algorithms to reliable sensitivities on test-retest variability. *Invest Ophthalmol Vis Sci.* 2016;57(13):5631–6.