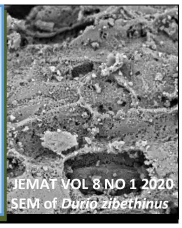
JEMAT, 2020, Vol 8, No 1, 37-41

HIBISCUS PUBLISHER

JOURNAL OF ENVIRONMENTAL MICROBIOLOGY AND TOXICOLOGY

Website: https://journal.hibiscuspublisher.com/index.php/JEMAT



Prediction of Cumulative Death Cases in The United States Due to COVID-19 Using Mathematical Models

Mohd Yunus Abd Shukor[1]*, Ibrahim Alhaji Sabo[2], Salihu Yahuza[3], Bilal Ibrahim Dan-Iya[4], and Surajo Adamu Wada[2]

[1]Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[2]Department of Microbiology, Faculty of Pure and Applied Sciences, Federal University Wukari, P.M.B. 1020 Wukari, Taraba State, Nigeria.
[3]Department of Microbiology and Biotechnology, Federal University Dutse, P.M.B 7156, Dutse, Jigawa state, Nigeria.
[4]Pharmacy Technician Department College of Health Sciences and Technology, Kano State Nigeria.

*Corresponding author:

Prof. Dr. Mohd Yunus Shukor,
Department of Biochemistry,
Faculty of Biotechnology and Biomolecular Sciences,
Universiti Putra Malaysia,
43400 UPM Serdang, Selangor,
Malaysia.
Email: mohdyunus@upm.edu.my

HISTORY

Received: 22nd July 2020
Received in revised form: 28th July 2020
Accepted: 28th July 2020

KEYWORDS

COVID-19
Total infection
pandemic
mathematical model
MMF

ABSTRACT

In this paper, we present different growth models such as Von Bertalanffy, Baranyi-Roberts, Morgan-Mercer-Flodin (MMF), modified Richards, modified Gompertz, modified Logistics and Huang in fitting and analyzing the epidemic trend of COVID-19 in the form of total number of death cases of SARS-COV-2 in The United States as of 20th of July 2020. The MMF model was found to be the best model with the highest adjusted R^2 value with the lowest RMSE value. The accuracy and bias factors values were close to unity (1.0). The parameters obtained from the MMF model include maximum growth of death rate (log) of 0.048 (95% ci from 0.047 to 0.048), curve constant (δ) that affects the inflection point of 2.34 (95% ci from 2.31 to 2.38) and maximal total number of death (y_{max}) of 151,356 (95% ci from 147,911 to 154,525). The MMF predicted that the total number of death cases for The United States on the coming 20th of August and 20th of September 2020 will be 148,183 (95% ci of 149,199 to 147,173) and 153,780 (95% ci of 152,640 to 154,928), respectively. The predictive ability of the model utilized in this study is a powerful tool for epidemiologist to monitor and assess the severity of COVID-19 in The United States in months to come. However, as with any other model, these values need to be taken with caution due to the unpredictability of the COVID-19 situation locally and globally.

INTRODUCTION

The growth curve of viruses and microorganisms on substrates such as nutrients or other organisms, including humans, typically followed a sigmoidal pattern, beginning with the lag section just after t = 0, preceded by the logarithmic section and then entering the stationary period and eventually progressing to the death or decline phase. there are various sigmoidal functions to explain the growth curve of organisms, such as Von Bertalanffy, Baranyi-Roberts, modified Richards, modified Gompertz and modified Logistics [1] including Morgan-Mercer-Flodin (MMF) [2]. The useful parameters for the growth curve include the maximum specific growth rate (μ_m), the lag period and the asymptotic values.

Mathematical models can be used to conduct COVID-19 pandemic analyzes including theoretical, quantitative, and simulation for the total number of death cases and deaths. models such as the modified Gompertz, Von Bertalanffy and Logistics have been used with strong predictive capacity to model the covid-19 pandemic [3]. The objective of this work is to evaluate several available models such as Logistic [1,4], Gompertz [1,5], Richards [1,6], Morgan-Mercer-Flodin (MMF) [2], Baranyi-Roberts [7], Von Bertalanffy [8,9], Buchanan three-phase [10] and more recently Huang model [11], in fitting and evaluating the COVID-19 outbreak pattern as of 20th of July 2020 in the form of the cumulative SARS-COV-2 death case in The United States.

MATERIALS AND METHODS

Data for the cumulative or total number of death cases from The United States as of 20th of July 2020 was acquired from Worldometer [12]. data were first converted to logarithmic values and the time after first death were utilized for time zero.

Statistical analysis

Statistical significant difference between the models was calculated through various methods including the adjusted coefficient of determination (R^2), accuracy factor (AF), bias factor (BF), Root-mean-square error (RMSE) and corrected AICc (Akaike information criterion) as before [13].

The RMSE was calculated according to Eqn. (1), where pd_i are the values predicted by the model and ob_i are the experimental data, n is the number of experimental data, and p is the number of parameters of the assessed model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Pd_i - Ob_i)^2}{n-p}} \quad \text{(eqn. 1)}$$

The adjusted R^2 is used to calculate the quality of nonlinear models according to the formula where rms is residual mean square and s_y^2 is the total variance of the y-variable ad calculated as follows:

$$Adjusted\left(R^2\right) = 1 - \frac{RMS}{s_Y^2} \quad \text{(eqn. 2)}$$

$$Adjusted\left(R^2\right) = 1 - \frac{\left(1-R^2\right)(n-1)}{(n-p-1)} \quad \text{(eqn. 3)}$$

The Akaike information criterion (AIC) [14] was calculated as follows;

$$AICc = 2p + n\ln\left(\frac{RSS}{n}\right) + 2(p+1) + \frac{2(p+1)(p+2)}{n-p-2} \quad \text{(eqn. 4)}$$

where n is the number of data points and p is the number of parameters of the model. the model with the smallest AICc value is highly likely correct [15]. The accuracy factor (AF) and bias factor (BF) as suggested by Ross [16] were calculated as follows;

$$\text{Bias factor} = 10^{\left(\sum_{i=1}^{n} \frac{\log(Pd_i/Ob_i)}{n}\right)} \quad \text{(eqn. 5)}$$

$$\text{Accuracy factor} = 10^{\left(\sum_{i=1}^{n} \frac{\log\left|(Pd_i/Ob_i)\right|}{n}\right)} \quad \text{(eqn. 6)}$$

Fitting of the data

Fitting of the bacterial growth curve using various growth models (**Table 1**) was carried out using GraphPad prism (v 8.0 trial version).

Table 1. models used in this study.

Model	p	Equation
modified Logistic	3	$y = \frac{A}{\left\{1+\exp\left[\frac{4\mu_m}{A}(\lambda-t)+2\right]\right\}}$
modified Gompertz	3	$y = A\exp\left\{-\exp\left[\frac{\mu_m e}{A}(\lambda-t)+1\right]\right\}$
modified Richards	4	$y = A\left\{1+v\exp(1+v)\exp\left[\frac{\mu_m}{A}(1+v)\left(1+\frac{1}{v}\right)(\lambda-t)\right.\right.$
Morgan-Mercer-Flodin (MMF)	4	$y = y_{max} - \frac{(y_{max}-\beta)}{1+(\mu_m t)^\delta}$
Baranyi-Roberts	4	$y = A + \mu_m x + \frac{1}{\mu_m}\ln\left(e^{-\mu_m x} + e^{-h_0} -\right.$ $-\ln\left(1+\frac{e^{\mu_m x + \frac{1}{\mu_m}\ln\left(e^{-\mu_m x}+e^{-h_0}-e^{-\mu_m x - h_0}\right)}-1}{e^{(y_{max}-A)}}\right)$
Von Bertalanffy	3	$y = K\left[1-\left[1-\left(\frac{A}{K}\right)^3\right]\exp^{-\left(\mu_m x/3K^{\frac{1}{3}}\right)}\right]^3$
Huang	4	$y = A + y_{max} - \ln\left(e^A + \left(e^{y_{max}} - e^A\right)e^{-\mu_m B(x)}\right)$ $B(x) = x + \frac{1}{\alpha}\ln\frac{1+e^{-\alpha(x-\lambda)}}{1+e^{\alpha\lambda}}$
Buchanan three-phase linear model	3	Y = A, IF X < LAG Y=A + K(X−λ), IF λ ≤ X ≥ X_{MAX} Y = Y_{MAX}, IF X ≥ X_{MAX}

Note:
a= maximum no of death cases lower asymptote.
y_{max}= maximum no of death cases upper asymptote.
μ_m= maximum specific growth rate of death.
v= affects near which asymptote maximum no of death cases occurs.
λ=lag time
e = exponent (2.718281828)
t = time after first death case is reported
α,β,δ and k = curve fitting parameters
h_0 = a dimensionless parameter quantifying the initial Physiological state of the reduction process. the lag time (h^{-1}) or (d^{-1}) can be calculated as $h_0=\mu_m$
when data at time zero is 0 (day after 1st death case log 1=0 for COVID-19) the MMF is reduced to a 3-parameter model

RESULTS AND DISCUSSION

All curves tested except for the Buchanan-3-phase model demonstrate visually appropriate fit (**Figs 1** to **6**). The best result was the MMF model which had the lowest value for RMSE, AICc and the highest value for adjusted R^2. The AF and BF values were both excellent for the model with their values being the nearest to 1.0. The poorest performance was the Buchanan-3-phase model (**TABLE 2**). The coefficients for the MMF model are shown in **TABLE 3**.

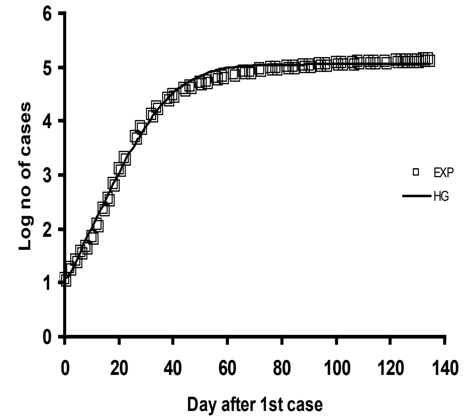


Fig. 1. Total no of SARS-COV-2 death cases in The United States as of 20th of July 2020 as modelled using the Huang model.

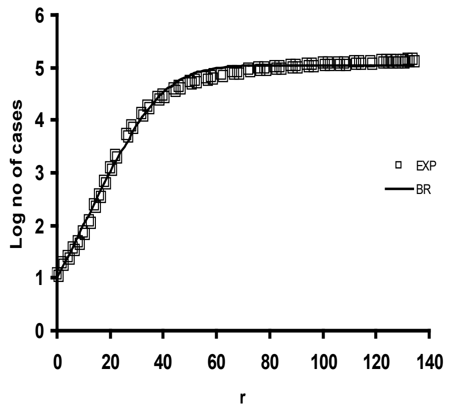


Fig. 2. Total no of SARS-COV-2 death cases in The United States as of 20th of July 2020 as modelled using the Baranyi-Roberts model.

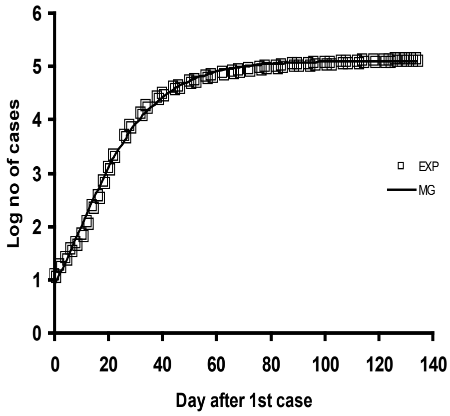


Fig. 3. Total no of SARS-COV-2 death cases in The United States as of 20th of July 2020 as modelled using the modified Gompertz model.

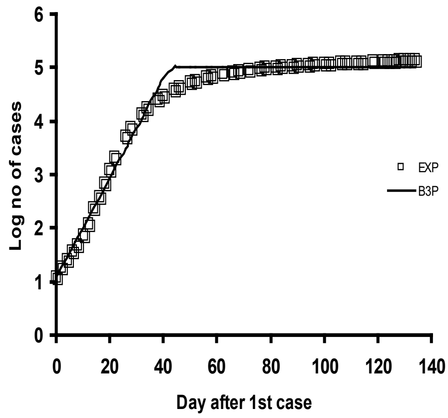


Fig. 4. Total no of SARS-COV-2 death cases in The United States as of 20th of July 2020 as modelled using the Buchanan-3-phase model.

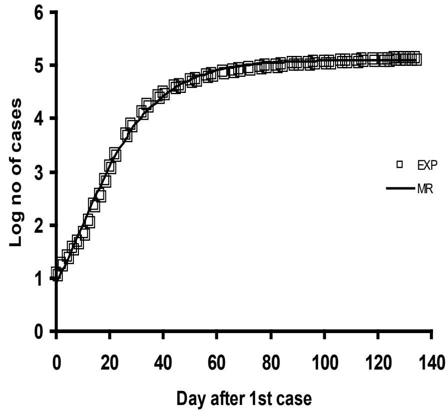


Fig. 5. Total no of SARS-COV-2 death cases in The United States as of 20th of July 2020 as modelled using the modified Richard model.

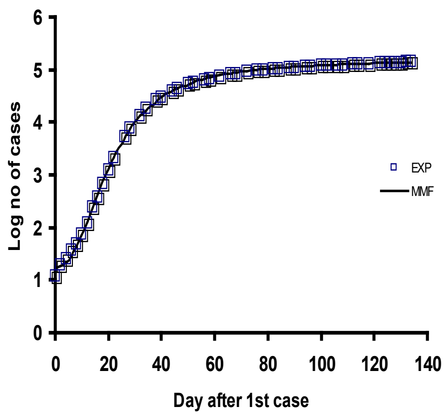


Fig. 6. Total no of SARS-COV-2 death cases in The United States as of 20th of July 2020 as modelled using the MMF model.

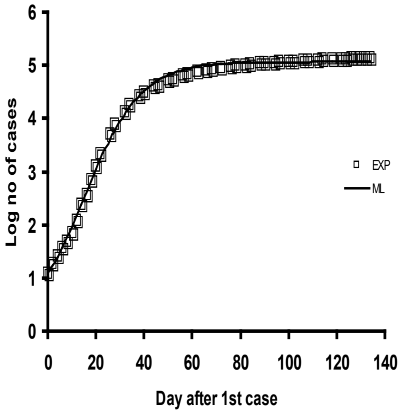


Fig. 7. Total no of SARS-COV-2 death cases in The United States as of 20th of july 2020 as modelled using the modified Logistics model.

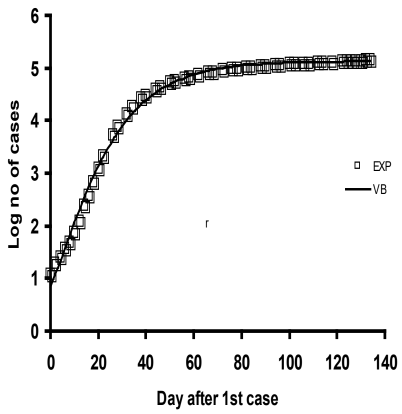


Fig. 8. Total no of SARS-COV-2 death cases in The United States as of 20th of july 2020 as modelled using the Von Bertalanffy model.

Table 2. Statistical tests for the various models utilized in modelling the total no of SARS-COV-2 death cases in The United States as of 20th of July 2020.

Model	*p*	RMSE	R^2	adR^2	AF	BF	AICc
Huang	4	0.085	0.996	0.995	1.012	1.00	-231.47
Baranyi-Roberts	4	0.086	0.996	0.995	1.013	1.00	-230.18
Modified Gompertz	3	0.072	0.997	0.997	1.023	1.00	-251.91
Buchanan-3-phase	3	0.152	0.990	0.989	1.023	1.00	-176.30
Modified Richards	4	0.072	0.997	0.997	1.006	1.00	-247.42
MMF	4	0.037	0.999	0.999	1.002	1.00	-314.34
Modified Logistics	3	0.065	0.998	0.997	1.010	1.00	-262.26
Von Bertalanffy	3	0.089	0.995	0.995	1.007	1.00	-230.31

Note: p is no of parameter

Table 3. Coefficients as modelled using the MMF model.

Parameters	Value	95% Confidence Interval
μ_m	0.048	0.047 to 0.048
δ	2.34	2.31 to 2.38
Y_{max}	151,356	147,911 to 154,525
B	1.28	1.25 to 1.30

Table 4. Predictions of COVID-19 pandemic for united states based on the MMF model.

Prediction	Mean	95% Confidence Interval
Maximum number of total cases by the end of covid-19	151,356	147,911 To 154,525
Maximum number of total cases by 20th of august 2020	148,183	147,173 To 149,199
Maximum number of total cases by 20th of September 2020	153,780	152,640 To 154,928

The parameters obtained from the MMF model include maximum growth of death rate (log) of 0.048 (95% ci from 0.047 to 0.048), curve constant (δ) that affects the inflection point of 2.34 (95% ci from 2.31 to 2.38) and maximal total number of death (y_{max}) of 151,356 (95% ci from 147,911 to 154,525). The MMF model predicted that COVID-19 deaths will end about 272 days (95% ci of 208 to 1327) days from 20th of July 2020 based on the lower bound of the 95% ci from the calculated maximum number of total cases (y_{max}) while the mean and upper 95% CI bound values failed to be predicted by the software for their number of days. The MMF predicted that the total number of death cases for The United States on the coming 20th of August and 20th of September 2020 will be 148,183 (95% ci of 149,199 to 147,173) and 153,780 (95% CI of 152,640 to 154,928), respectively.this prediction has to be taken with caution since the model failed to predict the number of days for the mean and upper 95% CI values and the number of days for COVID-19 to end may be much larger.

The Morgan-Mercer-Flodin or MMF model was originally designed to explain a wide variety of nutrient-response relationships in higher organisms [2] including ruminantsm microorganisms [17–21], yield of oil palm [22], alcohol [23] and even in financial growth [24]. Whether the predicted data is correct or not will depend on a case by case basis and include effectiveness of lockdown, mutation of the virus that increases the infectivity rate of the virus to name a few. Certainly, the models will be checked every few months to remodel the data so a better prediction can be obtained.

CONCLUSION

In conclusion, the MMF model was the best model in modelling number of deaths due to COVID-19 in the US based on statistical tests such as corrected AICc (akaike information criterion), bias factor (BF), adjusted coefficient of determination (R^2) and Root-mean-square error (RMSE). Parameters obtained from the fitting exercise were maximum growth rate (μ_m), the curve constants (δ) and maximal total number of death cases (y_{max}). The parameters obtained from the MMF model include maximum growth of death rate (log) of 0.048 (95% ci from 0.047 to 0.048), curve constant (δ) that affects the inflection point of 2.34 (95% CI from 2.31 to 2.38) and maximal total number of death (y_{max}) of 151,356 (95% CI from 147,911 to 154,525). The MMF predicted that the total number of death cases for The United States on the coming 20th of august and 20th of September 2020 will be 148,183 (95% CI of 149,199 to 147,173) and 153,780 (95% ci of 152,640 to 154,928), respectively. The model allows for prediction of total number of death cases and this prediction will vary according to various number of factors. Despite this, the predictive ability of the model utilized in this study is a powerful tool for epidemiologist to monitor and assess the severity of COVID-19 in the United states in months to come.

ACKNOWLEDGEMENT

On behalf of other authors, I would like to express my special gratitude to the Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, University Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia, for giving me the golden opportunity to publish this article.

REFERENCES

1. Zwietering MH, Jongenburger I, Rombouts FM, Van't Riet K. Modeling of the bacterial growth curve. Appl Environ Microbiol. 1990;56(6):1875–81.
2. Morgan PH, Mercer LP, Flodin NW. (YlimXn)/(K. 1975;72(11):4327–31.
3. Jia L, Li K, Jiang Y, Guo X, zhao T. Prediction and analysis of Coronavirus Disease 2019. 2020;(December).
4. Ricker, F.J. 11 Growth Rates and Models. Fish Physiol. 1979;Volume 8:677–743.
5. On X. Esq. F. R. S. &c. By. 1825;
6. Richards FJ. A flexible growth function for empirical use. J Exp Bot. 1959;10(2):290–301.
7. Baranyi J, Roberts TA. Mathematics of predictive food microbiology. Int J Food Microbiol. 1995;26(2):199–218.
8. 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.
9. Babák L, Šupinová P, Burdychová R. Growth models of thermus aquaticus and thermus scotoductus. Acta Univ Agric Silvic Mendel Brun. 2012;60(5):19–26.
10. Buchanan RL. Predictive food microbiology. Trends Food Sci Technol. 1993;4(1):6–11.
11. Huang L. Optimization of a new mathematical model for bacterial growth. Food Control. 2013;32(1):283–8.
12. Worldometer. COVID-19 Corona virus pandemic. 2020.
13. Halmi MIE, Shukor MS, Johari WLW, Shukor MY. Modeling the growth curves of *Acinetobacter* sp. strain DRY12 grown on diesel. J Environ Bioremediation Toxicol. 2014;2(1):33–7.
14. Akaike H. Factor analysis and AIC. Psychometrika. 1987;52(3):317–32.
15. Motulsky HJ, Ransnas L a. Fitting curves nonlinear regression : review a practical. FASEB J. 1987;1(5):365–74.
16. Ross T, McMeekin TA. Predictive microbiology. Int J Food Microbiol. 1994;23(3–4):241–64.
17. Santos SA, Da Silva E Souza G, De Oliveira MR, Sereno JR. Using nonlinear models to describe height growth curves in pantaneiro horses. Pesqui Agropecu Bras. 1999;34(7):1133–8.
18. Topal M, Bolukbasi C. Comparison of nonlinear growth curve models in broiler chickens. J Appl Anim Res. 2008;34(2):149–52.
19. Tariq MM, Iqbal F, Eyduran E, Bajwa MA, Huma ZE, Waheed A. Comparison of non-linear functions to describe the growth in Mengali sheep breed of Balochistan. Pak J Zool. 2013;45(3):661–5.
20. Asha-Augustine, Imelda-Joseph, Raj PR, David NS. Growth kinetic profiles of Aspergillus niger S(1)4 a mangrove isolate and Aspergillus oryzae NCIM 1212 in solid state fermentation. Indian J Fish. 2015;62(3):100–6.
21. Kemper CM. Growth and development of the brush-tailed rabbit-rat (Conilurus penicillatus), a threatened tree-rat from northern Australia. J Aust Mammal Soc. 2020;
22. Khamiz. Nonlinear Growth Models for Modeling Oil Palm Yield Growth. J Math Stat. 2005;1(3):225–33.
23. Germec M, Turhan I. Ethanol production from acid-pretreated and detoxified tea processing waste and its modeling. Fuel. 2018;231(January):101–9.
24. 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. 2014;7(1):76–87.