Research, Society and Development, v. 9, n. 8, e749986565, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

Previsão a longo prazo do número acumulado de óbitos no Brasil, China, Alemanha, Itália, Espanha, Estados Unidos: uma aplicação aos modelos em forma de S da COVID-19

Long-Term Time Prediction of Cumulative Number of Deaths in Brazil, China, Germany, Italy, Spain, the United States: an application to COVID-19 S-shaped models

Predicción a largo plazo del número acumulado de muertes en Brasil, China, Alemania, Italia, España, Estados Unidos: una aplicación a los modelos con forma de S de COVID-19

Recebido: 09/07/2020 | Revisado: 12/07/2020 | Aceito: 13/07/2020 | Publicado: 30/07/2020

Marcela Portela Santos de Figueiredo
ORCID: https://orcid.org/0000-0002-5042-1247
Universidade Federal Rural de Pernambuco, Brasil
E-mail: portela.marcela.producao@gmail.com
André Luiz Pinto dos Santos
ORCID: https://orcid.org/0000-0002-7703-2102
Universidade Federal Rural de Pernambuco, Brasil
E-mail: andredefensor@hotmail.com
Tiago Alessandro Espínola Ferreira
ORCID: https://orcid.org/0000-0002-2131-9825
Universidade Federal Rural de Pernambuco, Brasil
E-mail: taef.first@gmail.com
Maitê Priscila Lima Jota de Queiroz
ORCID: https://orcid.org/0000-0003-4178-9469
Universidade Federal de Pernambuco, Brasil
E-mail: maitejq@gmail.com

Resumo

Esta pesquisa objetiva ajustar o modelo de regressão não linear de Gompertz e Bertalanffy para as mortes acumuladas pelo COVID-19 em seis países Brasil, Estados Unidos, Alemanha, Itália, China e Espanha. Empregou três medidas de desempenho diferentes no processo de treinamento, coeficiente de determinação ajustado $\left(R^2_{aj.}\right)$, Critério de Informação de Akaike

(AIC) e Quadrado Médio Residual (RMS). O erro da porcentagem absoluta média (MAPE) e o erro relativo (ER) foram usados para selecionar o melhor modelo no conjunto de dados de teste. No conjunto de dados de treinamento, o modelo Bertalanffy foi o que melhor descreveu o crescimento de mortes na China, enquanto o modelo Gompertz foi o melhor para o Brasil, Alemanha, Itália, Espanha e Estados Unidos. Por outro lado, o modelo Bertalanffy foi o melhor para a Espanha no conjunto de dados de teste, de acordo com MAPE e RE. De acordo com o modelo de Gompertz, 214.100 IC (175.929; 267.008) pessoas morrerão no Brasil, atingindo um máximo de 1.577 com um intervalo de previsão [1.367; 1.819] de novas mortes diárias no pico da doença. Os modelos não lineares estudados descreveram satisfatoriamente a curva de crescimento do número de óbitos, fornecendo parâmetros com interpretações práticas. Foram encontradas evidências de que o Brasil pode superar os Estados Unidos em relação ao número total de mortes. A previsão de tempo a curto e longo prazo, bem como o ponto de virada de cada país, são apresentados e comparados com outros modelos preditivos da literatura.
Palavras-chave: Curva S; Pandemia; Coronavirus; Previsão.

Abstract

This research aims to adjust the Gompertz and Bertalanffy nonlinear regression model for the accumulated deaths by COVID-19 in six countries Brazil, United States, Germany, Italy, China, and Spain. It employed three different performance measures in the training process, adjusted determination coefficient $(R^2_{aj.})$, Akaike Information Criterion (AIC), and Residual Mean Square (RMS). The Mean Absolute Percentage Error (MAPE) and the Relative Error (RE) criterion were used to select the best model in the test dataset. On the training dataset, the Bertalanffy model was the one that best described the growth of deaths for China, while the Gompertz model was the best for Brazil, Germany, Italy, Spain, and the United States. In contrast, the Bertalanffy model was the best for Spain in the test dataset, according to MAPE and RE. According to the Gompertz model, 214,100 CI (175,929;267,008) people will die in Brazil, that will reach a maximum of 1,577 with a prediction interval [1,367; 1,819] of daily new deaths at its disease peak. The nonlinear models studied described the number of deaths growth curve satisfactorily, providing parameters with practical interpretations. Evidence was found that Brazil may surpass the United States regarding the total number of deaths. Short and long-term time prediction, as well as the turning point of each country, are presented and compared to other predictive models of the literature.

Keywords: S-curve; Pandemic; Coronavirus; Forecast.

Resumen

Este trabajo tiene como objetivo ajustar el modelo de regresión no lineal de Gompertz y Bertalanffy para las muertes acumuladas por COVID-19 en seis países Brasil, Estados Unidos, Alemania, Italia, China y España. Empleó tres medidas de rendimiento diferentes en el proceso de capacitación, coeficiente de determinación ajustado $\left(R^2_{aj.}\right)$, Criterio de información de Akaike (AIC) y Cuadrado medio residual (RMS). El error de porcentaje absoluto medio (MAPE) y el criterio de error relativo (RE) se utilizaron para seleccionar el mejor modelo en el conjunto de datos de prueba. En el conjunto de datos de entrenamiento, el modelo de Bertalanffy fue el que mejor describió el crecimiento de muertes para China, mientras que el modelo de Gompertz fue el mejor para Brasil, Alemania, Italia, España y los Estados Unidos. Por el contrario, el modelo Bertalanffy fue el mejor para España en el conjunto de datos de prueba, según MAPE y RE. Según el modelo de Gompertz, 214,100 CI (175,929; 267,008) personas morirán en Brasil, que alcanzará un máximo de 1,577 con un intervalo de predicción [1,367; 1,819] de nuevas muertes diarias en su pico de enfermedad. Los modelos no lineales estudiados describieron la curva de crecimiento del número de muertes satisfactoriamente, proporcionando parámetros con interpretaciones prácticas. Se encontró evidencia de que Brasil puede superar a los Estados Unidos con respecto al número total de muertes. La predicción del tiempo a corto y largo plazo, así como el punto de inflexión de cada país, se presentan y comparan con otros modelos predictivos de la literatura.
Palabras clave: S curva; Pandemia; Coronavirus; Pronóstico.

1. Introduction

COVID-19 represents a considerable threat to health on a global scale. (Baumgartner et al., 2020). COVID-19 was first detected in China in late 2019 (Na Zhu et al., 2020), and has spread to Asia, Europe, North America, and, more recently, South America and Africa (Baumgartner et al., 2020).

According to Rodriguez-Morales et al. (2020), the first case of COVID-19 in Latin America occurred in Brazil. The country currently has the highest number of deaths in that region (Prado, 2020; European Center for Disease Prevention and Control, 2020a). On April 23rd, 20 European countries already showed evidence that they had passed the peak of transmission (European Center for Disease Prevention and Control, 2020b). However, in early

June, in Brazil, the number of death still grows (Ministry of Health, 2020).

As in Mellan et al. (2020) from Imperial College, in their report of May 8, at the beginning of the pandemic in Brazil, an individual infected an average of three to four other people. After the implementation of closing schools and reducing the mobility of the population, the effective reproduction number (Ro) was reduced. However, its value remains above one (Ro=1.47 with a 95% confidence interval (CI) of [1.34,1.59] from São Paulo state, for instance), which indicates that the epidemic was not yet under control in this period and would continue to grow. According to De Lemos Menezes, Garner, and Valenti (2020), Brazil will be the new epicenter of the COVID-19 pandemic in the world, surpassing the USA in the number of death.

According to Vasconcelos et al. (2020), the accumulated death curve is more reliable than that of confirmed cases, because of underreporting, since many carriers of the virus are asymptomatic or develop mild symptoms, and therefore will not be detected unless they are tested. How commented by them, the fraction of confirmed cases concerning the total number of infections depends heavily on the testing policy employed, what it will depend on each mitigation disease policies used by each country. There is no uniformization of social and health policies, which makes it difficult to compare the evolution of the confirmed cases between different countries (Vasconcelos et al., 2020). Thus, it would be more appropriate to analyze the number of deaths.

Studies involving the calculation of the adjustments of curves and predictions from regressions are essential in most recent researches involved with COVID-19. Ghosal et al. (2020) predicted the number of deaths due to Covid-19 using linear regression in data from India for a six-week interval. Yang et al. (2020b) used linear and nonlinear models (exponential model) to model the cumulative death curve of different Chinese provinces, obtaining suitable adjustments. Melik-Huseynov et al. (2020) produced linear models of daily deaths in the USA, China, and Italy. Santiago et al. (2020a) ran several nonlinear models to adjust the accumulated death curve speed, i.e., the daily deaths due to COVID-19, in Brazil, Italy, and the world. The best models, according to the AIC, BIC, and RMSE criteria, were Rational, Exponential modified, and Weibull, respectively, for the world, Brazil and Italy.

Utkucan and Tezcan (2020) had used grey model (GM), nonlinear grey Bernoulli model (NGBM) and fractional nonlinear grey Bernoulli model (FANGBM) to adjust data from the coronavirus outbreak, furthermore, they forecast thirty-steps ahead the number of confirmed cases of COVID-19 in Italy, the United Kingdom (UK) and the United States of America (USA). Also, Utkucan and Tezcan (2020) state that according to the criterion, FANGBM was the best prediction model. They noted that the cumulative number of cases of COVID-19 in Italy, UK, and USA were forecasted as approximately 233000, 184000, and 1090000, respectively, on May 22, 2020, by using FANGBM (1,1).

Ribeiro et al. (2020) analyzed several forecasting models, such as autoregressive integrated moving average (ARIMA), cubist regression (CUBIST), random forest (RF), ridge regression (RIDGE), support vector regression (SVR), and stacking-ensemble learning. In this work, they executed time series forecasting with one, three, and six-days ahead for COVID-19 cumulative confirmed cases in Brazilian states. This study concluded that the ranking of models, from the best to the worst, were: SVR, stacking ensemble learning, ARIMA, CUBIST, RIDGE, and RF models.

However, the sigmoid logistic functions, Gompertz, Richards, and Hill models are more suitable for epidemics forecasting (Yang et al., 2020a). According to them, linear, exponential, cubic, and quadratic functions cannot adequately capture the typical S-shaped curve. Kucharavy and De Guio (2011) noted that the application of an S-curve (Gompertz, Richards, Bertalanffy, for example) with the purpose of forecasting, induces the correct measurement of the growth process in most cases. The S-shape model can be utilized to identify the law of natural growth quantitatively and inform the ceiling (upper limits of growth) and the steepness of the growth (slope of the curve). Indeed, one can determine a more accurate ceiling and steepness with a bigger data set (Kucharavy and De Guio, 2011).

Vasconcelos et al. (2020) used Richards' nonlinear model to analyze the death curve from China, Germany, Iran, France, South Korea, and Spain. For Brazil, an exponential growth model was carried. Das Neves (2020) adjusted the Gompertz model for the accumulated confirmed COVID-19 case data for nine countries: Brazil, Austria, China, Germany, Italy, New Zealand, Spain, United Kingdom, and the United States, not analyzing the accumulated deaths curve. Shen (2020) applied a logistic growth model to new daily cases due to COVID-19 from China, South Korea and Iran, France, America, Germany, Italy, Japan, Singapore, and Spain. Moreover, this work demonstrates that the logistic model suits the data in China exceedingly well.

Based on the models tested and its respective forecast power, this work aims to adjust Gompertz and Bertalanffy S-shaped nonlinear models to the data of confirmed accumulated deaths of COVID-19 for six countries: Brazil, The United States, Germany, Italy, China, and Spain. With this study, it is possible to choose the best model for each country among the observed models. Also, with this study, it is possible to make short and long-term predictions of the number of deaths with a 95% confidence interval to these countries and estimate the start date of the peak death rate.

The contributions of this paper can be summarized as follows:

• The first contribution is related to the presentation of the parameters of S-shaped models (Gompertz and Bertallanfy models) from Brazil, the United States, Spain, Italy, Germany, and China. To measure the models' quality and choose the best model for describing the pandemic in each country;

• The second contribution, is the measure of the quality of the short-term prediction of the cumulative confirmed deaths of COVID-19, providing by S-shaped models (Gompertz and Bertalaffy models);

• Also, this paper evaluates models forecasting in a multi-day-ahead forecasting strategy with 95% confidence interval and estimation of turning point of curve, whose accuracy of the models may assists theses countries in decision-making to implement social distance, quarantine or isolation strategies. The Gompertz model suggests that Brazil will surpass the United States regarding the total number of deaths from COVID-19;

This paper is organized as follows: Section 2 gives a brief description of the employed data and the forecasting models. Subsequently, it illustrates the estimation of nonlinear model parameters and the quality measures for adjustment models. Section 3 presents the results obtained, where it is divided into three subsections: training results (Section 3.1), test results (Section 3.2), and long-term time prediction results (Section 3.3). The Training results comprehend the evaluation of the best model of the cumulative number of deaths in each country. Moreover, it brings the asymptotic number of deaths with 95% CI. The Test Results (Section 3.2) measure the quality of short-term forecasting in the test dataset and compare the models Gompertz and Betalanffy in each country. Section 3.3 Long-term Time Prediction Results elaborates a long-term prediction of the cumulative number of deaths' S-shaped models and provides predictions with 95%-Confidence interval. Finally, Section 4 concludes this work.

2. Methodology

A case study is a detailed analysis of a case that presents some particularity, making it unique (Pereira et a., 2018). Thus, this research is a case study with a quantitative nature that evaluates the Cumulative number of deaths due to COVID-19 in six different countries, with unique specificities each. Data collection was performed by documentary research. The data source will be explored in the next section (2.1 Data).

2.1 Data

The information used in this work to adjust the models came from the World Health Organization (WHO) database (https://covid19.who.int/). Firstly, the raw observation data are split into training and test datasets. The training dataset was adopted for the adjustments of the curves. The training datasets employed in this study were composed of daily observation on March 3, 2020, to May 23, 2020, for Brazil. For China, January 11, 2020, to May 23, 2020. For Germany, March 9, 2020, to May 23, 2020. For Italy, February 22, 2020, to May 23, 2020. For Spain, March 4, 2020, to May 23, 2020. And for the United States, March 3, 2020, to May 23, 2020. The test datasets are composed of the next eight daily observations for all countries, from May 24, 2020, to May 31, 2020. These Test sets were used to calculate the Percent Relative Error (PRE) and Mean Absolute Percentage Error (MAPE) of the models concerning the observed values.

2.2 Models

The Gompertz and Bertalanffy models were adjusted to estimate the increase in the number of deaths as a function of time (days). The equations of the nonlinear models are,

Gompertz $$W(t) = Ae^{-Be^{(-kt)}} + \varepsilon \quad (1)$$

Bertalanffy $$W(t) = A\{1 - Be^{-kt}\}^3 + \varepsilon \quad (2)$$

In the models showed on equations (1 and 2): $W(t)$ is the confirmed COVID 19 number of deaths in time t. k indicates the curve growth speed to asymptotic death toll. A is the COVID 19 number of deaths when t tends to infinite; B is a shape parameter; e is the exponential function, and ε is the associated random error following a normal distribution with zero mean and constant variance.

Research, Society and Development, v. 9, n. 8, e749986565, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

2.2.1 Estimation of nonlinear models parameters

The obtaining of the nonlinear models' parameters was performed through the least squares method using the iterative Gauss-Newton process, using the NLS procedure of the R software stats package (R Development Core Team, 2018).

2.3 Adjustment quality assessor

The precision of the adjustments was considered according to the following criteria to determine the most appropriate models to describe the number of deaths growth.

2.3.1 The adjusted coefficient of determination $\left(R^2_{aj.}\right)$, expressed by:

$$R^2_{aj.} = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - p}\right],$$

in which, $R^2 = 1 - \frac{RSS}{TSS}$ is the determination coefficient, RSS the residual sum of squares, and TSS the total sum of squares.

2.3.2 The Residual Mean Squares (RMS)

$$RMS = \sum_{i=1}^{n} |y_i - \hat{y}_i|/n,$$

where y_i is the observed number of death; $\hat{y}_i$ is the estimated number of deaths (prediction of y_i); n is the number of observations.

2.3.3 The Akaike information criteria is given by:

$$AIC = -2logL + 2(p + 1).$$

where L represents the maximized likelihood log, and p stands for the number of parameters in the model. According to Pan, (2001) AIC is a powerful and widely used model selection criterion.

2.3.4 Absolute Growth Rate

The absolute growth rate (AGR) is obtained from the first derivative of the adjusted model, regarding the time $(\partial W(t) / \partial t)$. AGR represents, in fact, the daily new deaths.

2.3.5 Mean Absolute Percentage Error

The Mean Absolute Percentage Error (MAPE) is a measure of regression quality. It is defined as the mean of the ratio $|g(x) - y| / |y|$, wherein x is the input of the regression model (time, for instance), g is the regression model, and y represents the observed value. It is generally used because it has a very intuitive interpretation of relative error (De Myttenaere et al., 2016).

3. Results

In this section, the Training Results (Section 3.1), Test Results (Section 3.2), and Long-term Time Prediction Results (Section 3.3) will be described in the following three passages.

3.1 Training Results

Table 1 shows the confidence intervals (CI), the estimates of the parameters, and the criteria used for the two evaluated models. The Bertalanffy model for Brazil did not present estimates, as it did not converge. Santos et al. (2020) report in their study that the Bertalanffy model presented problems regarding the convergence with the data in the State of Ceará in Brazil. It is still possible to see that the Bertalanffy model was the one that best fitted the data from China. Through this table, it is possible to notice that the Gompertz and Bertalanffy models presented useful amplitudes in the range for the estimated parameters.

The asymptotic number of deaths A and the relative growth rate k are considered the most critical parameters of the Gompertz and Bertalanffy models. k represents the decline rate in the relative growth rate. The higher is value, the lower is the growth in the death rate

(Santos et al., 2020). Gompertz model presented the highest values for this parameter, while Bertalanffy presented the lowest ones.

For most countries, the asymptotic number of deaths represented by A was higher in the Bertalanffy model, like the for the United States that represents 207,400 IC (192,877;225,378) deaths which suggest that the Bertalanffy model overestimated the death rate. According to the Gompertz model, 214,100 CI (175,929;267,008) people will die in Brazil, in line with the study by Imperial College (2020), which estimated the total number of death for Brazil using the suppression model (1.6 deaths per 100,000), finding a total of 206,087 deaths. On the other hand, the Gompertz model predicts that 120,400 CI (116,614;124,638) deaths will occur in the United States. Therefore, the Gompertz model brings indications that Brazil will overtake the United States in the number of death with a 95%-confidence interval in an extended period.

According to Vasconcelos et al (2020) 4,620 IC (4,615; 4,625) people will die in China due COVID-19, this result agrees with the finding of this work (Table 1), 4,609 CI(4,415;4,837) from Gompertz model and 4,732 CI(4,522;4,985) from Bertalanffy. Vasconcelos et al (2020) used Richard’s model. Moreover, Çelik, Ankarali and Pasin (2020) calculated several models: Weibull, Negatif üstel, Janoschek, Lundqvist-Korf and Sloboda, thus the cumulative number of deaths was predicted in a range of 3,547 and 5,031 deaths.

Table 1. Model parameters estimates, 95%-confidence interval (CI), and selection criteria.

		Parameters			Criteria		
Countries	Models	A	k		$R^2_{aj.}$	RMS	AIC
Brazil	Gompertz	214100	0.02		**0.999**	**13024**	**829**
		CI(175929;267008)	CI(0.01;0.02)				
China	Gompertz	4609	0.952	132267			1965
		CI(4415;4837)					
	Bertalanffy	4732	0.03		**0.957**	**116721**	**1948**
		CI(4522;4985)	CI(0.02;0.04)				
Germany	Gompertz	8811	0.06		**0.999**	**2316**	**809**
		CI(8752;8870)	CI(0.06;0.07)				
	Bertalanffy	10570	0.04		0.996	39847	1025
		CI(10256;10917)	CI(0.03;0.04)				
Italy	Gompertz	33930	0.05		**0.999**	**60805**	**1279**
		CI(33709;34165)	CI(0.05;0.06)				
	Bertalanffy	37940	0.03		0.998	327423	1434
		CI(37334;38576)	CI(0.03;0.04)				
Spain	Gompertz	28460	0.07		**0.999**	**128683**	**1187**
		CI(28196;28739)	CI(0.07;0.08)				
	Bertalanffy	30900	0.05		0.997	325563	1262
		CI(30438;31395)	CI(0.04;0.05)				
United States	Gompertz	120400	0.04		**0.999**	**1303950**	**1392**
		CI(116614;124638)	CI(0.04;0.05)				
	Bertalanffy	207400	0.021		0.994	2930419	1479
		CI(192877;225378)	CI(0.020;0.022)				

Source: Authors.

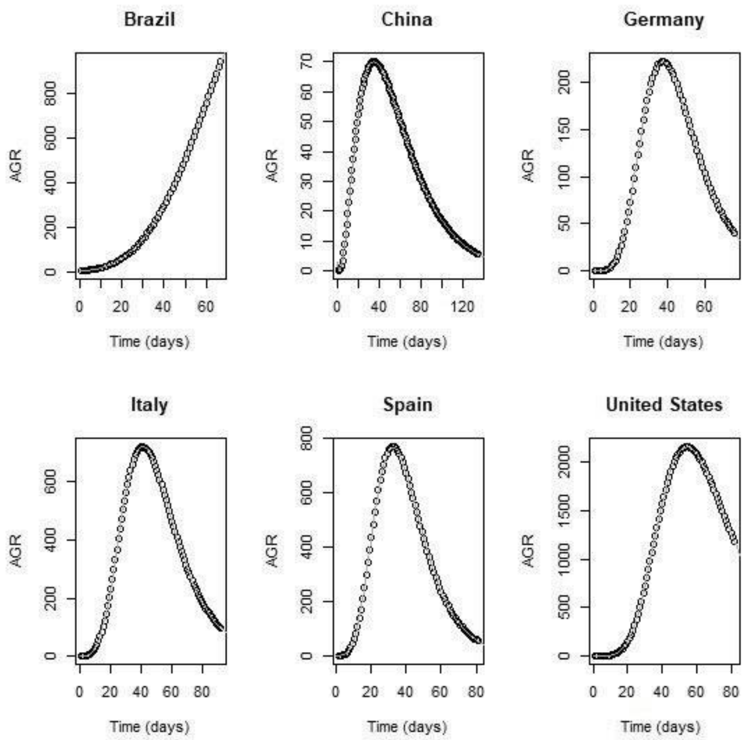
Table 1 informs the prediction of the total number of deaths for Germany, 8,811 CI(8,752;8,870), from Gompertz model and 10,570 CI(10,256;10,917). Vasconcelos et al (2020) founds 8,400 IC (8,300; 8,500) deaths from Richard's model in the first wave of the outbreak. Gompertz model, in Table 1, gives an information that Italy will reach 33,930 CI(33,709;34,165) deaths, futhermore Bertalanffy provides 37,940 CI(37,334;38,576) deaths in the first wave. Vasconcelos et al (2020) indicate 32,000 IC (31,700; 32,300) deaths for this country. Table 1 presents the total number of deaths for Spain, 28460 CI(28,196;28,739), from Gompertz model and 30,900 CI(30,438;31,395) from Bertalanffy. Vasconcelos et al (2020), show 27,100 IC (26,900; 27,300) deaths for Spain by Richard's model.

The tested models adequately represented the number of deaths in the training dataset, presenting high $R^2_{aj.}$, indicating useful adjustments (Table 1). Gompertz model' RMS and AIC were lower than Bertalanffy model' (results in bold, Table 1). Then, according to RMS and AIC criteria, Gompertz was the best model for Brazil, Italy, Spain, Germany, although Bertalanffy was the best for China in the training dataset. Consequently, the Gompertz model fits the data better than the Bertalanffy model in most countries of this study in the training dataset.

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

Figure 1 shows the total death absolute growth rate (AGR) in the training dataset for Brazil (March 18, 2020, to May 23, 2020). China from January 1, 2020, to May 23, 2020. Germany from March 9, 2020, to May 23, 2020. Italy from February 22, 2020, to May 23, 2020. Spain from March 4, 2020, to May 23, 2020. And the United States from March 3, 2020, to May 23, 2020.

Figure 1. The absolute growth rate of the total number of deaths by countries due to COVID-19 according to the selection criteria in the training dataset.



Source: Authors.

The Gompertz model AGR for Brazil indicates it is still in full growth, i.e., the number of cases has not yet reached the tipping point. The evolution of the Bertalanffy model AGR for China, the Gompertz model for Germany, Italy, Spain, and the United States, indicate a decrease from the forty-day, thirty-seventh day, forty-first day, thirty-second day and fifty-fourth day, respectively (Figure 1).

3.2 Test Results

In this subsection, the model prediction was calculated from the twenty-fourth to the thirty-first day of May, two thousand and twenty, in the test dataset, as it can be seen in Table

2. The Gompertz model (G. Est) for Brazil in the first two days of prediction shows lower values than observed. However, this model presents augmented values in the subsequent days (from the twenty-sixth day to the thirty-first day of May of two thousand and twenty) with a Relative Error (RE) ranging from (0.41 to 1.38%) and a MAPE of 1.09% in that 8-day interval.

Table 2 also shows the Bertalanffy model estimative of the cumulative number of deaths (B. Est), and the estimative of Gompertz model in Germany, China and Brazil one, two, three, to eight-days ahead, which presented values close to that observed with an RE ranging from 1.01 to 1.74%; 0.05 to 0.70%, and a MAPE of 0.37; 0.45%, respectively into the interval of eight days.

Table 2. Prediction of the number of deaths by COVID-19 in Brazil, China, and Germany from May 24, 2020, to May 31, 2020, in the test dataset.

	Brazil		China			Germany		
Date	Obs.	G. Est.	Obs.	G. Est.	B. Est.	Obs.	G. Est.	B. Est.
05/24	21048	20776	4645	4510	4564	8247	8243	8623
ER		1.29%		2.91%	**1.74%**		**0.05%**	4.56%
05/25	22013	21754	4645	4514	4569	8257	8279	8695
ER		1.18%		2.82%	**1.63%**		**0.23%**	5.30%
05/26	22666	22758	4645	4517	4575	8302	8313	8764
ER		0.41%		2.75%	**1.51%**		**0.13%**	5.56%
05/27	23473	23787	4645	4521	4580	8349	8345	8831
ER		1.38%		2.67%	**1.40%**		**0.05%**	5.77%
05/28	24512	24840	4645	4524	4585	8411	8375	8895
ER		1.34%		2.60%	**1.29%**		**0.43%**	5.75%
05/29	25598	25918	4645	4528	4589	8450	8403	8957
ER		1.25%		2.52%	**1.21%**		**0.56%**	6%
05/30	26754	27020	4645	4531	4594	8489	8430	9017
ER		0.99%		2.45%	**1.10%**		**0.70%**	6.22%
05/31	27878	28146	4645	4534	4598	8500	8455	9075
ER		0.96%		2.39%	**1.01%**		**0.53%**	6.76%
MAPE		**1.09%**		2.64%	**1.36%**		**0.34%**	5.74%

Source: Authors.

To summarize, for China, It´s important to notice that Bertalanffy model had better quality on prediction in test dataset than the Gompertz model regarding the MAPE and ER performance (in bold, Table 2), whereas for Germany, Gompertz model was the best for the test dataset.

According to Wu et al. (2020), with more data, scenarios become more accurate, especially when using data after the peak. Santiago et al. (2020b) studied the evolution of the

COVID-19 epidemic with data up to March 31st, 2020, two thousand and twenty, in Brazil, Italy, and the world, and concluded that the results did not allow safe projections on when the number of deaths and confirmed cases would decrease. In comparison with Santiago et al. (2020b) study, Wu et al. (2020) employed two more months in the analysis. However, the data do not show the passage of the pandemic peak in Brazil.

Table 3 shows the prediction on the total number of deaths for Italy, Spain, and the United States, from May 24th to May 31st, a short prediction horizon of eight days. It is noticed that the model of Italy by Gompertz has smaller RE (1.18% to 1.32%), as well as lower MAPE (1.29%) concerning Bertalanffy. The same is right in the United States, but for Spain, the reverse is true. The Bertalanffy model has lower RE (0.07% to 1.01%) and a lower MAPE (0.37%) in the short eight-day prediction horizon. Although the Gompertz Spain model has been better adjusted throughout the training data (March 4, 2020, to May 23, 2020) according to AIC, RMS, and $R^2_{aj.}$ criteria, in the test data, the Bertalanffy adjusted better to the observations according to the RE and MAPE.

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

Table 3. Prediction on the number of deaths by COVID-19 in Italy, Spain, and the United States from May 24 to May 31, in the test dataset.

	Italy			Spain			United States		
Date	Obs.	G. Est.	B.Est.	Obs.	G. Est.	B. Est.	Obs.	G. Est.	B. Est.
05/24	32735	32308	33440	28678	27734	28697	94011	93868	98484
ER		1.30%	2.15%		3.29%	0.07%		0.15%	4.76%
05/25	32785	32397	33602	28752	27785	28802	95863	94985	100271
ER		1.18%	2.49%		3.36%	0.17%		0.92%	4.60%
05/26	32877	32481	33758	28834	27832	28902	96909	96060	102040
ER		1.20%	2.68%		3.48%	0.24%		0.88%	5.29%
05/27	32955	32561	33909	29035	27876	28998	97529	97097	103791
ER		1.20%	2.89%		3.99%	0.32%		0.44%	6.42%
05/28	33072	32636	34054	29036	27917	29089	98119	98094	105525
ER		1.32%	2.97%		3.85%	0.18%		0.03%	7.55%
05/29	33142	32707	34195	29037	27955	29175	98889	99053	107240
ER		1.31%	3.18%		3.73%	0.48%		0.17%	8.44%
05/30	33229	32775	34331	29039	27991	29258	100304	99976	108936
ER		1.37%	3.32%		3.61%	0.75%		3.22%	5.45%
05/31	33340	32839	34461	29043	28024	29337	101567	100863	110613
ER		1.59%	3.36%		3.51%	1.01%		0.69%	8.91%
MAPE		**1.29%**	2.88%		3.60%	**0.37%**		**0.45%**	6.82%

Source: Authors.

To sum up, the Bertalanffy model was the best for Spain and China in the test dataset, according to MAPE and RE criteria. In contrast, the Gompertz model was the best for Germany, Italy, and the United States. The Bertalanffy model for Brazil did not converge. However, the Gompertz model for Brazil presented reasonable estimations in training and test dataset based on criteria (Table 3).

An overview of the seletion criteria of non-linar models in literature related to COVID-19 disease is given in Table 4.

Table 4. Sumarization of the selection criteria results of non-linear models in litterature.

Author	Variable	Country	Model	MAPE	MSE	RMSE	R^2
Çelik, Ankarali and Pasin (2020)	Cumulative number of deaths	China	Weibull	–	88568.282	–	0.949
			Negatif üstel	–	146041.563	–	0.949
			Bertalanffy	–	80308.717	–	0.953
			Janoschek	–	88568.282	–	0.949
			Lundqvist-Korf	–	75008.636	–	0.956
			Sloboda	–	76161.284	–	0.956
Utkucan and Tezcan (2020)	Cumulative number of cases	Italy	GM	9.2422%	9846	–	0.947
			NGBM	1.9690%	2418	–	0.997
			FANGBM	0.9174%	1223	–	0.999
		USA	GM	99.582%	118786	–	0.958
			NGBM	6.0082%	6002	–	0.999
			FANGBM	4.8950%	5767	–	0.999
Santiago et al (2020b)	Cumulative number of deaths	Italy	Gaussian	–	–	60.7108	0.999
			Logistical power	–	–	53.6634	0.999
		Brazil	Hoerl	–	–	2.35996	0.999
			Hyperbolic decline model	–	–	2.58327	0.999
Results of this paper	Cumulative number of deaths	Brazil	Gompertz	1.09%	–	–	0,999
		China	Gompertz	2.64%	–	–	0.952
			Bertalanffy	1.36%	–	–	0.957
		Germany	Gompertz	0.34%	–	–	0,999
			Bertalanffy	5.74%	–	–	0,996
		Italy	Gompertz	1.29%	–	–	0,999
			Bertalanffy	2.88%	–	–	0,998
		Spain	Gompertz	3.60%	–	–	0,999
			Bertalanffy	0.37%	–	–	0,997
		United States	Gompertz	0.45%	–	–	0,999
			Bertalanffy	6.82%	–	–	0,994

Source: Authors.

Table 4 demonstrates former literature about prediction quality indicators of the models adjusted on COVID-19 data. It´s important to notice that the MAPE criteria should be next to zero such as Utkucan and Tezcan (2020), and the R^2 close to 1, like Çelik, Ankarali and Pasin (2020), Utkucan and Tezcan (2020) and Santiago et al (2020b). The results of this work (Table 1, 2 and 3) are as valuable as previous analysis.

3.3. Long-term Time Prediction Results

Table 5 shows the extrapolation to a broader horizon of days, to August 4th, 2020. The Institute for Health Metrics and Evaluation (IHME) from the University of Washington made a forecast for the total number of deaths in Brazil on August 4th, 2020, obtaining a value of 125,833 CI(68,311;221,078) deaths from COVID-19 (IHME, 2020b), a value close to that obtained by the Gompertz model for Brazil 122,387 CI(109,066;135,708) shown in Table 5. Given that the forecast horizon is long, the uncertainty associated with the value is substantial. For the United States, the IHME (2020c) predicts 147,040 CI(113,182;226,971) as a total number of deaths on August 4th, 2020.

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

Table 5. Forecast on the cumulative number of deaths by COVID-19 time series in Brazil, China, Germany, Italy, Spain, and the United States multi-step ahead with 95%-Confidence Interval.

Countries	Date	Models Gompertz	Models Bertalanffy
Brazil	8/4	122,387 CI(109,066;135,708)	– –
China	8/4	4,603 CI(4.426;4.780)	4,717 CI(4,524;4,909)
Germany	8/4	8,806 CI(8,748;8,864)	10,453 CI(10,117;10,788)
Italy	8/4	33,908 CI(33,685;34,131)	37,634 CI(37,007;38,260)
Spain	8/4	28,459 CI(28,195;28,723)	30,842 CI(30.341;31.342)
United States	8/4	119,520 CI(115,752;123,288)	179,630 CI(168,924;190,336)

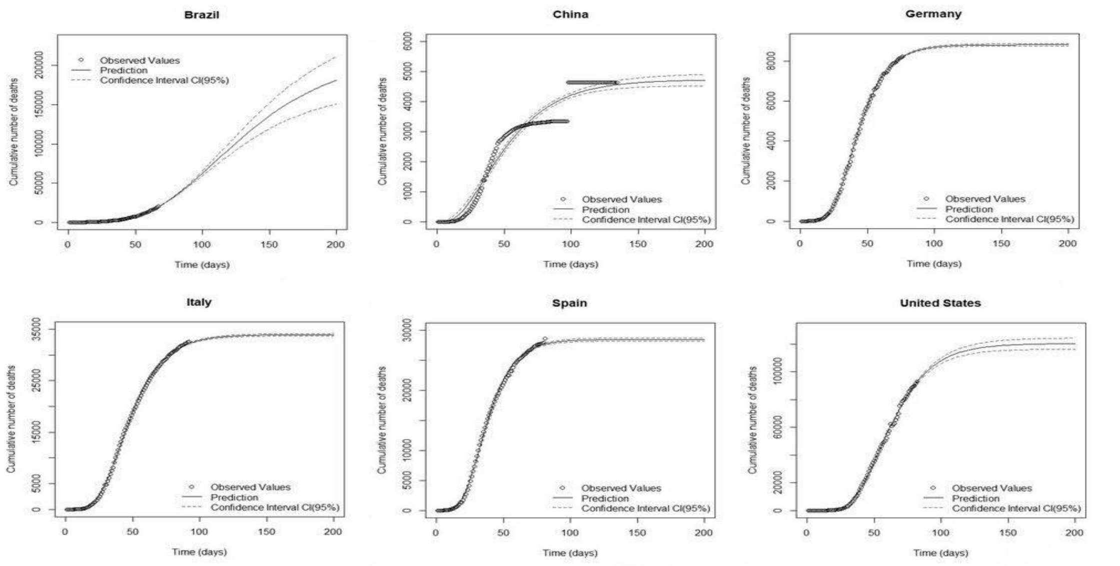
Source: Authors.

Table 5 shows the Gompertz and Bertalanffy model forecasts, which predict the total number of deaths 119,520 CI(115,752;123,288) and 179,630 CI(168,924;190,336), respectively, on the same date. Regarding Spain, the death predicted by the IHME (2020d) is 27,727 CI(25,720;32,130). In this study, the number of deaths is 28,459 CI(28,195;28,723) and 30,842 CI(30.341;31.342) for the Gompertz and Bertalanffy model, respectively. The

IHME (2020d) evaluates 31,458 CI(29,605;34,969) deaths in Italy, while this work predicts 33,908 CI(33,685;34,131) deaths for the Gompertz model and 37,634 CI(37,007;38,260) for Bertalanffy model. For IHME (2020d), in Germany, there would be 8,543 CI(7,006;12,150) deaths on August 4th, while according to Gompertz 8,806 CI(8,748;8,864) and Betalanffy 10,453 CI(10,117;10,788). The results are close to the studied countries, considering the significant fluctuation associated with the data.

Figure 2 shows the extrapolation for the total number of deaths due to COVID-19, according to the Gompertz model for Brazil, Italy, The United States, Spain, and Germany, whereas Bertalanffy model for China.

Figure 2. Extrapolation of the total number of death from COVID-19 by countries.



Source: Authors.

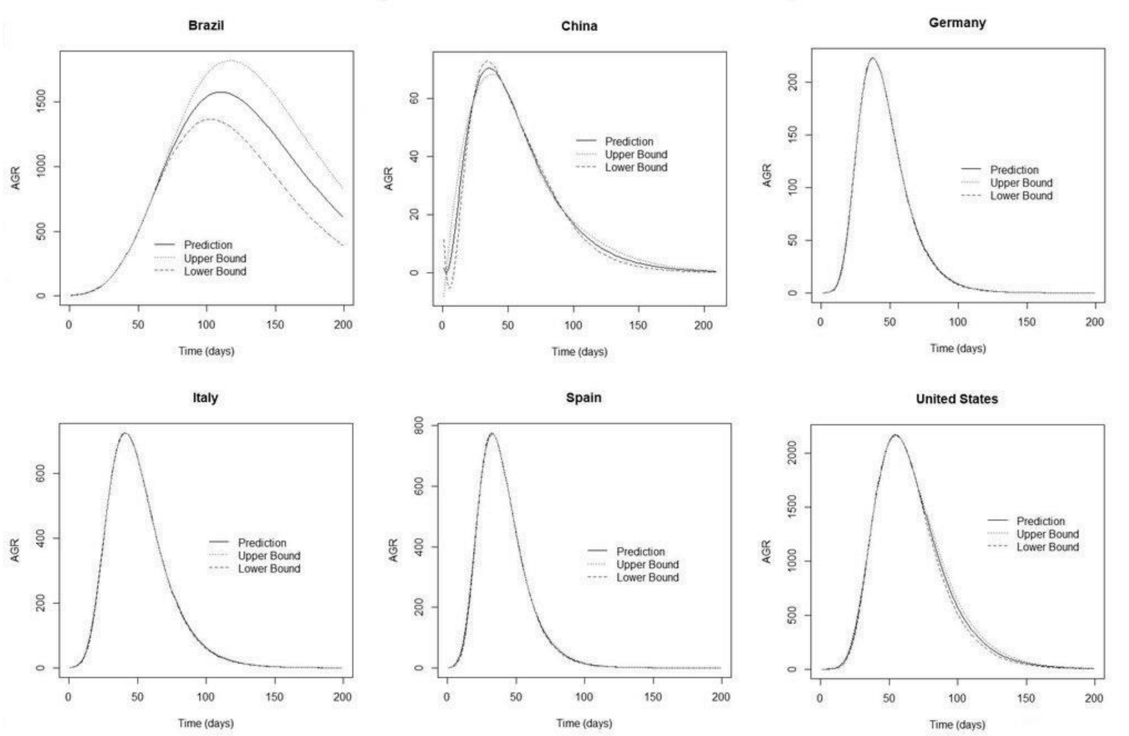
The curve from Brazil does not show signs of stabilization, while in other analyzed countries, the same happened 100 days before the beginning of the pandemic in these countries (Figure 2). Brazil Confidence Interval (CI) was wider than the one from other countries considered in this work, furthermore, the uncertainty on the prediction of the cumulative number of deaths in Brazil is greater, although there may still be enough precision to make decisions about the authorities intervention usefulness.

Figure 3 shows the AGR, that is, the daily new deaths ($deathsx\,day^{-1}$). The AGR was calculated by the First-difference of the S-shaped model's cumulative number of deaths. Then, Figure 3 shows the AGR provided by First-difference of the Gompertz model for Brazil, the United States, Germany, Italy, Spain, in addition to the AGR provided by the First-

Research, Society and Development, v. 9, n. 8, e749986565, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

difference of Bertalanffy model for China. Figure 3 shows the China AGR maximum of 71 [69; 73] daily new deaths. Moreover, Germany attains 224 new deaths per day, although Italy AGR hits 725 [724; 726] daily new deaths. Furthermore, Spain comes to 774 [772;776] daily new deaths at the Turing Point of the cumulative number of deaths.

Figure 3. AGR for the cumulative number of deaths from COVID-19 in countries.



Source: Authors.

The AGR implies that Brazil will reach a maximum of 1,577 with a prediction interval [1,367; 1,819] daily new deaths at its disease peak. Besides, there is a broader curve at its base compared to other countries. Although the United States obtains a rate of 2167 [2,160; 2,174] daily deaths at its peak, Brazil has a broader curve base than the United States. On April 11th, the United States reached the quota of 2000 daily deaths, according to BBC News (2020). That rate continued for days in April in the United States, according to Shumaker (2020). As claimed by the Wordometers website (2020), the United States peaked at 2683 deaths on April 21st. Thus, the AGR of the Gompertz Model for the United States is underestimating the rate. Figure 3 shows Bertalanffy's AGR for China.

The Bertalanffy model AGR in the United States was 1997 [1996;2199] daily new deaths. However, close to this value, the Bertalanffy model's curve is wider at the base than in

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.6565

the Gompertz AGR for the same country. Thus, the Bertalanffy model points to a considerable increase in the number of deaths predicted by the Bertalanffy model.

The turning point of the cumulative number of death curves is the death rate critical point, i.e., when the derivative of the death rate is equal to zero, the daily new deaths peak. The turning point for each model was calculated by numerical analysis in Table 6.

Table 6. Turning points and prediction intervals for the cumulative number of deaths according to S-shaped models by countries.

Country	Model	Turning Point
Brazil	Gompertz	7/7/2020 [6/30/2020 - 07/14/2020]
China	Gompertz	2/21/2020 [2/20/2020-2/23/2020]
	Bertalanffy	2/16/2020 [2/15/2020-2/19/2020]
Germany	Gompertz	4/17/2020 [4/17/2020- 4/17/2020]
	Bertalanffy	4/16/2020 [4/15/2020- 4/16/2020]
Italy	Gompertz	4/04/2020 [4/04/2020- 4/04/2020]
	Bertalanffy	4/02/2020 [4/02/2020- 4/02/2020]
Spain	Gompertz	4/7/2020 [4/7/2020- 4/8/2020]
	Bertalanffy	4/4/2020 [4/4/2020- 4/4/2020]
United States	Gompertz	4/28/2020 [4/28/2020- 4/28/2020]
	Bertalanffy	5/7/2020 [5/6/2020; 5/10/2020]

Source: Authors.

The findings in this study are similar to those presented by the Fokas study, Dikaios and Kastis (2020), where they found the turning point on April 15th for Germany, April 4th for Italy, April 6th for Spain and April 17th to the United States using the Logistics model. Vasconcelos et al. (2020) report in their study that the number of accumulated deaths in Brazil attributed to COVID-19 until May 8th, 2020, already showed concern for not reaching the turning point. Dr. Murray, IHME director, notes that "the IHME is predicting that the number of deaths in Brazil will continue to rise, there will be a shortage of critical hospital resources, and the death peak may not occur until mid-July" (IHME, 2020a). According to the AGR of the Brazil Gompertz model, the turning point will occur in outset-July.

4. Conclusion

This article analyzes the accumulated death curve of countries that have surpassed the pandemic (Spain, Italy, China, Germany, the United States), and Brazil, which has not reached its peak yet. It is possible to establish that the best model to be adjusted in the training data according to the criteria of AIC, RMS, and $R^2_{aj.}$: Gompertz for all countries, except for China, which had the Bertalanffy model more adequately adjusted to the data. Also, it is possible to indicate the best model to the test dataset according to the criteria of MAPE and RE: Gompertz for the United States, Italy, whereas Bertallanfy for China and Spain. The Bertalanffy model for Brazil did not converge. However, the Gompertz model for Brazil presented reasonable estimations in training and test dataset based on employed criteria.

Based on the Gompertz and Bertalanffy models, it is possible to predict estimates regarding the number of death and the death rate with 95% confidence interval, as well as the turning point of the curve and prediction interval, with considerable associated uncertainty, but compatible with results found by other research groups such as IMHE and Imperial College. According to the Gompertz model, 214,100 CI (175,929;267,008) of people will die in Brazil during the first wave of the disease. Brazil will reach a maximum of 1,577 with a prediction interval [1,367; 1,819] of daily new deaths at its disease peak.

According to the results presented, it is concluded that the epidemic is still far from being controlled in Brazil, and there is significant uncertainty in the forecasts from the data. Even so, evidence was found that Brazil may surpass the United States in the number of the total deaths, according to the Gompertz S-shaped model. As projected, the daily new deaths of Brazil will start to decrease in outset-July, although further studies are needed to update the parameters as the data are collected for more accurate results. For this reason, the safest situation is to maintain isolation as long as possible. At the same time, the population has not been immunized, promoting non-pharmacological measures such as closing schools, increasing the number of tests carried out, and restricting public meetings to limit the disease dissemination.

For future works, it is intended (i) Update the database and perform the analyzes to check the error of the long-term time prediction, (ii) Employ a two-compartmental model to forecast the cumulative number of deaths before and after the turning point.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. The authors thank the reviewers for their fruitful comments for the manuscript.

References

Baumgartner, M. T., Lansac-Toha, F. M., Coelho, M. T. P., Dobrovolski, R., & Diniz-Filho, J. A. F. (2020). Social distancing and movement constraint as the most likely factors for COVID-19 outbreak control in Brazil. medRxiv. doi: 10.1101/2020.05.02.200880

BBC News (2020, April 11). *Coronavirus: US death toll passes 2,000 in a single day*. BBC News. Retrieved from https://www.bbc.com/news/world-us-canada-52249963

Celik, S., Ankarali, H., & Pasin, O. (2020). Modelling of COVID-19 outbreak indicators in china between january and april. medRxiv.

Das Neves, K. D. Predictive analysis of covid-19 confirmed cases in brazil and eight countries based on the Gompertz nonlinear model. doi: 10.1590/SciELOPreprints.451

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean absolute percentage error for regression models. Neurocomputing, 192, 38-48.

European Centre for Disease Prevention and Control (2020a). Recuperado de https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

European Centre for Disease Prevention and Control (2020b). Recuperado de https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-rapid-risk-assessment-coronavirus-disease-2019-ninth-update-23-april-2020.pdf

De Lemos Menezes, P., Garner, D. M., & Valenti, V. E. (2020). Brazil is projected to be the next global COVID-19 pandemic epicenter. medRxiv. doi: 10.1101/2020.04.28.20083675

Fokas, A. S., Dikaios, N., & Kastis, G. A. (2020). COVID-19: Predictive Mathematical Models for the Number of Deaths in South Korea, Italy, Spain, France, UK, Germany, and USA. MedRxiv. doi: 10.1101/2020.05.08.20095489

Ghosal, S., Sengupta, S., Majumder, M., & Sinha, B. (2020). Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14, 2020). Diabetes & metabolic syndrome, 14(4), 311–315.

Institute for Health Metrics and Evaluation (IHME) (2020a). Recuperado de http://www.healthdata.org/sites/default/files/files/Projects/COVID/PressRelease_Brasil-Portuguese-0525.pdf

Institute for Health Metrics and Evaluation (IHME) (2020b). Recuperado de http://www.healthdata.org/sites/default/files/files/Projects/COVID/Estimation_update_052520.pdf

Institute for Health Metrics and Evaluation (IHME) (2020c). Recuperado de http://www.healthdata.org/news-release/new-ihme-covid-19-projections-first-forecasts-select-nations-latin-america-asia-and

Institute for Health Metrics and Evaluation (IHME) (2020d). Recuperado de http://www.healthdata.org/news-release/new-ihme-forecast-projects-nearly-135000-covid-19-deaths-us

Imperial College (2020). Recuperado de https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-Global-unmitigated-mitigated-suppression-scenarios.xlsx

Kucharavy, D., & De Guio, R. (2011). Application of S-shaped curves. Procedia Engineering, 9, 559-572.

Melik-Huseynov, D. V., Karyakin, N. N., Blagonravova, A. S., Klimko, V. I., Bavrina, A. P., Drugova, O. V., & Kovalishena, O. V. (2020). Regression Models Predicting the Number of Deaths from the New Coronavirus Infection. Современные технологии в медицине, 12(2 (eng)).

Mellan, T. A., Hoeltgebaum, H. H., Mishra, S., Whittaker, C., Schnekenberg, R. P., Gandy, A., & Faria, N. R. (2020). Report 21: Estimating COVID-19 cases and reproduction number in Brazil. medRxiv. doi 10.1101/2020.05.09.20096701

Ministério da Saúde [MS]. (2020). *Painel Coronavírus*. Recuperado de https://covid.saude.gov.br/

Na Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G. F., & Tan, W. A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine. 382, 727-733. doi: 10.1056/NEJMoa2001017

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. Biometrics, 57(1), 120-125.

Pereira, A. S., Shitsuka, D. M., Parreira, F. J., & Shitsuka, R. (2018). *Metodologia da pesquisa científica. [e-book]. Santa Maria. Ed.* UAB/NTE/UFSM. Recuperado de https://repositorio.ufsm.br/bitstream/handle/1/15824/Lic_Computacao_Metodologia-Pesquisa-Cientifica.pdf?sequence=1.

Prado, B. (2020). COVID-19 in Brazil: “So what?”. TheLancet, 395, 10235, 1461.

Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., and dos Santos Coelho, L. (2020).

Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. Chaos, Solitons & Fractals, 135, 109853.

Rodriguez-Morales, A. J., Gallego, V., Escalera-Antezana, J. P., Méndez, C. A., Zambrano, L. I., Franco-Paredes, C., Suárez, J. A., Rodriguez-Enciso, H. D., Balbin-Ramon, G. J., Savio-Larriera, E., Risquez, A., & Cimerman, S. (2020). COVID-19 in Latin America: The implications of the first confirmed case in Brazil. Travel medicine and infectious disease, 35, 101613. https://doi.org/10.1016/j.tmaid.2020.101613

Santiago, E. J. P., da Silva Freire, A. K., de Almeida Ferreira, D. S., de Amorim, J. F., Cunha, A. L. X., de Freitas, J. R., Santiago, E. J. P., da Silva Freire, A. K., de Almeida Ferreira, D. S., de Amorim, J. F., Cunha, A. L. X., de Freitas, J. R., & Cunha Filho, M. (2020a). Velocity of deaths and confirmed cases of COVID-19 in Brazil, Italy and worldwide. Research, Society and Development, 9(7), 263974085. doi: http://dx.doi.org/10.33448/rsd-v9i7.4085

Santiago, E. J. P., da Silva Freire, A. K., Cunha Filho, M., Moreira, G. R., de Almeida Ferreira, D. S., & Cunha, A. L. X. (2020). (2020b). Non-linear models applicable to mortality and cases of COVID-19 in Brazil, Italy and the world. Research, Society and Development, 9(6), 117963561. doi: http://dx.doi.org/10.33448/rsd-v9i6.3561

Santos, A.L.P., de Figueiredo, M. P. S., Ferreira, T. A. E., Gomes-Silva, F., Moreira, G. R., Silva, J. E., & de Freitas, J. R. (2020). Analysis and forecasting of the evolution of COVID-19 death numbers in the state of Pernambuco and Ceará using regression models. Research, Society and Development, 9(7), 1-24, e602974551. doi: http://dx.doi.org/10.33448/rsd-v9i7.4551

Shen, C. Y. (2020). A logistic growth model for COVID-19 proliferation: experiences from China and international implications in infectious diseases. International Journal of Infectious Diseases. 96, 582-589. doi: https://doi.org/10.1016/j.ijid.2020.04.085

Shumaker, L. (2020, April 24). *COVID-19 killing 2,000 Americans a day in April; US death toll exceeds Korean War. National Post*. Recuperado de https://nationalpost.com/news/world/covid-19-killing-2000-americans-a-day-in-april-toll-exceeds-korean-wars/

Utkucan, Ş., & Tezcan, Ş. (2020). Forecasting the cumulative number of confirmed cases of COVID-19 in Italy, UK and USA using fractional nonlinear grey Bernoulli model. Chaos, Solitons & Fractals, 139, 109948. https://doi.org/10.1016/j.chaos.2020.109948

Vasconcelos, G. L., Macêdo, A. M., Ospina, R., Almeida, F. A., Duarte-Filho, G. C., & Souza, I. C. (2020). Modelling fatality curves of COVID-19 and the effectiveness of intervention strategies. medRxiv.

Wordometers (2020). Retrieved from https://www.worldometers.info/coronavirus/country/us/

Wu, K., Darcet, D., Wang, Q., & Sornette, D. (2020). Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. arXiv preprint arXiv:2003.05681. https://arxiv.org/abs/2003.05681, 2020.

Yang, W., Zhang, D., Peng, L., Zhuge, C., & Hong, L. (2020a). Rational evaluation of various epidemic models based on the COVID-19 data of China. arXiv preprint arXiv:2003.05666.

Yang, S., Cao, P., Du, P., Wu, Z., Zhuang, Z., Yang, L., Yu, X., Zhou, Q., Feng, X., Wang, X., Li, W., Liu, E., Chen, J., Chen, Y., & He, D. (2020b). Early estimation of the case fatality rate of COVID-19 in mainland China: a data-driven analysis. Annals of translational medicine, 8(4), 128. https://doi.org/10.21037/atm.2020.02.66

Percentage of contribution of each author in the manuscript

Marcela Portela Santos de Figueiredo – 30%

André Luiz Pinto dos Santos – 30%

Tiago Alessandro Espínola Ferreira – 20%

Maitê Priscila Lima Jota de Queiroz – 20%