MDPI

Article

Piecewise Modeling the Accumulated Daily Growth of COVID-19 Deaths: The Case of the State of São Paulo, Brazil

Erlandson Ferreira Saraiva [1,*] **and Carlos Alberto de Bragança Pereira** [2]

1   Institute of Matematics, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Brazil
2   Institute of Matematics and Statistics, University of São Paulo, São Paulo 05508-090, Brazil; cadebp@gmail.com
*   Correspondence: erlandson.saraiva@ufms.br; Tel.: +55-67-3345-7511

Abstract: The pandemic scenery caused by the new coronavirus, called SARS-CoV-2, increased interest in statistical models capable of projecting the evolution of the number of cases (and associated deaths) due to COVID-19 in countries, states and/or cities. This interest is mainly due to the fact that the projections may help the government agencies in making decisions in relation to procedures of prevention of the disease. Since the growth of the number of cases (and deaths) of COVID-19, in general, has presented a heterogeneous evolution over time, it is important that the modeling procedure is capable of identifying periods with different growth rates and proposing an adequate model for each period. Here, we present a modeling procedure based on the fit of a piecewise growth model for the cumulative number of deaths. We opt to focus on the modeling of the cumulative number of deaths because, other than for the number of cases, these values do not depend on the number of diagnostic tests performed. In the proposed approach, the model is updated in the course of the pandemic, and whenever a "new" period of the pandemic is identified, it creates a new sub-dataset composed of the cumulative number of deaths registered from the change point and a new growth model is chosen for that period. Three growth models were fitted for each period: exponential, logistic and Gompertz models. The best model for the cumulative number of deaths recorded is the one with the smallest mean square error and the smallest Akaike information criterion (AIC) and Bayesian information criterion (BIC) values. This approach is illustrated in a case study, in which we model the number of deaths due to COVID-19 recorded in the State of São Paulo, Brazil. The results have shown that the fit of a piecewise model is very effective for explaining the different periods of the pandemic evolution.

Keywords: COVID-19; growth model; piecewise model; estimation; non-linear minimum square

1. Introduction

Since the discovery of a new coronavirus in the city of Wuhan, China, the virus has spread rapidly worldwide, and created a major global health crisis [1]. In fact, this disease is one of the main health problems in the world and has had an enormous economic and social impact, leading to an increase in poverty and the loss of thousands of lives in many countries. Until 27 April 2021, four million deaths had been registered in the whole world (www.worldometers.info/coronavirus, accessed on 23 July 2021).

This pandemic scenery has increased the interest in statistical models capable of projecting the evolution of the disease in countries, states, or cities. This interest lies mainly in projections that can assist government agencies in making decisions regarding the intensification of social isolation, the acquisition of hospital equipment, an increase in the number of intensive care units in hospitals.

Especially in the year 2020, many articles were published describing modeling procedures for the number of cases and/or deaths due to COVID-19 in many countries. In general, the published works model the accumulated number of cases (or deaths) by using some non-linear growth model. For example, Ref. [2] considered a simple exponential

growth model to analyze the initial phase of the epidemic of COVID-19 in Africa, Ref. [3] modeled the data from Philippines and Taiwan using the generalized logistic model, Ref. [4] applied the Richards growth model to the data collected in China, France, Germany, Iran, Italy, South Korea, and Spain, Ref. [5] calibrated the logistic growth model, the generalized logistic growth model, and the generalized Richards model for the number of cases recorded in China, Ref. [6] considered the Gompertz model to predict the number of deaths in Brazil, Ref. [7] estimated the number of total COVID-19 cases and deaths in the UK, Russia, and Turkey using the Gompertz model, among others.

As discussed by the authors cited above, the growth models present satisfactory performance for modeling the accumulated number of cases (or deaths) of COVID-19. However, these models are generally adequate to model data from a specific period of the pandemic. Since the growth in the cumulative number of cases and deaths by COVID-19 has, in general, presented a heterogeneous evolution over time, this implies that the adjustment of only one of these growth models may not be adequate to explain the entire study period.

In order to overcome this issue, this article introduces a piecewise model in which different growth functions are adjusted for different pandemic periods of time. Our proposed modeling is based on the fit of three growth models, exponential, logistic and Gompertz, for each period of the pandemic. Our preference for these three models is based on studies described in the literature that indicate they are excellent for use in longitudinal data, see for example [8,9] and the references therein. However, the proposed modeling is not restricted to these three growth models and can be easily adapted for other growth models.

To identify the different periods of the pandemic in the course of time, we consider a grid of values with increments of size 1, in which, each point of the grid represents a possible change point of the pandemic. The increment of one unit represents the increase of one day in the pandemic. Then, we fit the three growth models for two sub-datasets obtained according to each possible change point. To choose the best model for each sub-dataset, we consider as a criterion the mean square error (MSE), the Akaike Information criterion [10,11], denoted by AIC, and the Bayesian Information criterion [12], denoted by BIC. The best change point d is that associated with the model with the lowest mean square error.

The fitting procedure of the three models is based on obtaining the point estimates for the parameters of the growth models by using the non-linear least square method, as described by [13,14]. We have used the software R [15] and the command *nls* of the package *nlstools* [16,17]. Based on the fitted model, we find the coordinates of the inflection point and consequently the estimated date for the peak of the pandemic in the current period.

The proposed modeling is illustrated in a case study, in which we model the cumulative number of deaths recorded in the state of São Paulo, Brazil. However, the modeling may be easily adapted for datasets from other countries, states, or cities. We opt for modeling the cumulative number of deaths instead of the cumulative number of recorded cases because these values do not depend on the number of diagnostic tests performed. This allows us to identify the changes in the pandemic's evolution without having to adjust for changes in the number of diagnostic tests. On the other hand, one can argue that the modeling procedure now relies on accurate reporting of the mortality statistics due to COVID-19. Since the health secretary of São Paulo state publishes the mortality statistics daily and the information is checked by a group of news-press companies, we consider that these published data are adequate for our modeling procedure. As an illustration of the versatility of our proposal, we also present the model for the cumulative number of recorded cases in the State of São Paulo, Brazil.

The two main advantages of the proposed modeling approach are: 1. its effectiveness in describing the different periods of the pandemic; and 2. its capacity to explain each period of the pandemic through the estimates for the epidemiological parameters of interest,

such as the growth rate for the number of daily deaths and the date for the occurrence of the peak.

The remainder of the paper is organized as follows: Section 2 presents the statistical model. Section 3 describes the adopted modeling procedure to obtain a piecewise growth model for the cumulative number of recorded deaths in the state of São Paulo, Brazil. In this section, we also present the piecewise model fitted for the cumulative number of cases recorded in the State. Section 4 concludes the paper with the final remarks. Additional details are provided in Appendix A.

2. Statistical Model

Consider D_t to be the number of deaths due to COVID-19 recorded in the state of São Paulo, Brazil, on the t-th day, for $t = 0, \ldots, n$, where $t = 0$ represents the day that the first death was recorded (17 March 2020) and $n = 410$ is the last day considered in the study (27 April 2021). In this period, 96,191 deaths were recorded. This dataset is publicly available on the website www.seade.gov.br/coronavirus/.

Figure 1 shows the recorded D_t values, for $t = 1, \ldots, n$. As one can note, the recorded D_t values have a high variability which makes the modeling procedure somewhat complicated. The sample standard deviation of the D_t values is 237.8288. Due to this, we opt to develop the modeling procedure by considering the cumulative values because these values present a more stable behavior, as can be viewed in Figure 2a. Figure 2b shows the values from Figure 2a in the log-scale.
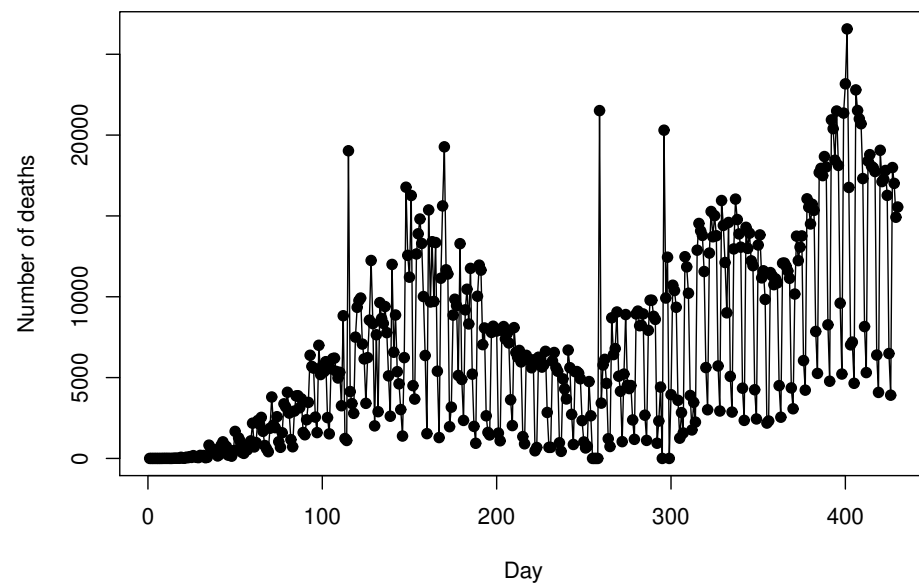

Figure 1. Number of deaths recorded by day.

Thus, let X_t be the number of deaths until the t-th day, in an accumulated way, i.e.,

$$X_0 = D_0 \quad \text{and} \quad X_t = \sum_{i=0}^{t} D_i,$$

for $t = 1, \ldots, n$.

Then, we assume the following multiplicative growth model:

$$X_t = h(t|\boldsymbol{\theta}) \cdot \nu_t, \tag{1}$$

where $h(t|\boldsymbol{\theta})$ is a nonlinear growth model indexed by the parameter $\boldsymbol{\theta}$ (scalar or vector), ν_t is a random error assumed as being generated from a log-normal distribution with mean μ and variance σ^2, for $t = 1, \ldots, n$.
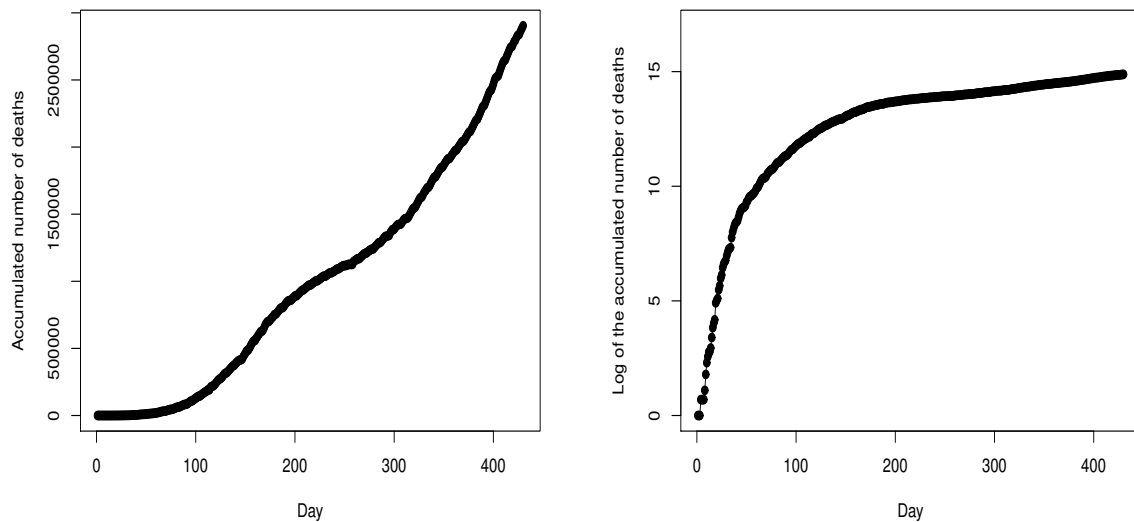
(**a**) Original scale.　　　　　　　　　　　　　　(**b**) Logarithm scale.

Figure 2. Cumulative number of deaths until day t, for $t = 0, 1, \ldots, n$.

To complete the model specification, we need now to set up the nonlinear growth model $h(t|\boldsymbol{\theta})$ in Equation (1), for $t = 1, \ldots, n$. Hereafter, we consider the following three nonlinear growth models: exponential, logistic and Gompertz. Details on these three growth models are presented in Appendix A.

In addition, in order to avoid working with values on the scale of thousands (see Figure 2a), we opt to consider the model Equation (1) on the logarithmic scale. Thus, by taking the logarithm transformation on both sides of Equation (1), we have that:

$$Y_t = log(X_t) = f(t|\boldsymbol{\theta}) + \varepsilon_t \tag{2}$$

where $f(t|\boldsymbol{\theta}) = log(h(t|\boldsymbol{\theta}))$ is the logarithmic transformation of a nonlinear growth model $h(t|\boldsymbol{\theta})$ indexed by the parameter $\boldsymbol{\theta}$ and $\varepsilon_t = log(\nu_t)$ for $t = 1, \ldots, n_d$. Thus, from log-normal properties we have that ε_t is a random error assumed as being generated from a normal distribution with mean μ and variance σ^2, $\varepsilon_t \sim \mathcal{N}(\mu, \sigma^2)$. In addition, we assume that $\mu = 0$ and $cov(\varepsilon_t, \varepsilon_{t'}) = 0$, for $t, t' = 1, \ldots, n$ and $t \neq t'$.

However, since the cumulative number of deaths by COVID-19 has, in general, presented a heterogeneous evolution over time, the adoption of a single function $f(t|\boldsymbol{\theta})$ to explain the entire considered period may not be adequate. Thus, in order to overcome this issue, we consider that the function $f(t|\boldsymbol{\theta})$ is a piecewise function, given by

$$f(t|\boldsymbol{\theta}) = \begin{cases} f_1(t|\boldsymbol{\theta}) & \text{, for } 0 < t \leq d_1; \\ f_2(t|\boldsymbol{\theta}) & \text{, for } d_1 < t \leq d_2; \\ \vdots & , \quad \vdots \quad \vdots \quad ; \\ f_k(t|\boldsymbol{\theta}) & \text{, for } t > d_k. \end{cases}$$

That is, we assume that the entire period of the pandemic is divided into k different periods, being $f_j(t|\boldsymbol{\theta})$ a function considered for recorded data in the period $(d_{j-1}, d_j]$, for $j = 1, \ldots, k$.

Table 1 shows the mathematical functions considered for $f_j(t|\boldsymbol{\theta})$, for $j = 1, \ldots, k$. Other than for the log-exponential model that presents an infinite growth, the log-logistic and log-Gompertz have growth limits. The value $\alpha_1 = log(\alpha_1^*)$ is the upper asymptote for both log-models, where α_1^* is the upper asymptote of the models in the original scale (see Appendixes A.2 and A.3). In the context of COVID-19, the value of $\alpha_1^* = \exp\{\alpha_1\}$ is an estimate for the maximum number of deaths that will occur. The parameters (α_2, α_3) are related to the coordinates $I = (T, f(T|\boldsymbol{\theta}))$ of the inflection point. From a practical

viewpoint, the point I is the peak of the pandemic, i.e., the point in which the curve of the number of deaths (log-transformed) changes its slope (positive to negative), meaning that before the peak the number of deaths grows at an increasing rate (positive slope) and after the peak, the growth is characterized by a decreasing rate (negative slope). That is, after the peak it is expected that the number of deaths is smaller each day. In addition, we have that T is the day on which the peak will occur and $f(T|\theta)$ is the logarithm of the cumulative number of deaths that will be recorded until the day T.

Table 1. Log-transformed growth models.

Model	Parameters	Mathematical Expression	Inflection Point	Growth at Inflection Point
log-exponential	$\theta = (\alpha_1, \alpha_2)$	$Y_t = \alpha_1 + \alpha_2 t$	does not exist	does not exist
log-logistic	$\theta = (\alpha_1, \alpha_2, \alpha_3)$	$Y_t = \alpha_1 - log(1 + \alpha_2 \exp\{-\alpha_3 t\})$	$T = log(\alpha_2)/\alpha_3$	$f(T|\theta) = \alpha_1 - log(2)$
log-Gompertz	$\theta = (\alpha_1, \alpha_2, \alpha_3)$	$Y_t = \alpha_1 - \alpha_2 \exp\{-\alpha_3 t\}$	$T = log(\alpha_2)/\alpha_3$	$f(T|\theta) = \alpha_1 - 1$

In order to get the parameter estimates of the three log-growth models, we adopt the nonlinear least square method. For this, we use the software R [15] and the command *nls* of the package *nlstools* [16,17]. Denote the estimates for the parameters of interest θ by $\hat{\theta}$, where $\hat{\theta} = (\hat{\alpha}_1, \hat{\alpha}_2)$ for the log-exponential model and $\hat{\theta} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3)$ for the log-logistic and log-Gompertz models.

To select the best model for the data of each period of the pandemic, we compare the three log-growth models by using as a criterion the mean square error (MSE) and the model selection criteria AIC and BIC, which are calculated according to the following expressions

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(\hat{Y}_t - Y_t)^2, \quad AIC = -2l(\hat{\theta}|\mathbf{y}) + 2p \quad \text{and} \quad BIC = -2l(\hat{\theta}|\mathbf{y}) + plog(n),$$

where $\hat{Y}_t$ is the estimated value by the model, $l(\hat{\theta}|\mathbf{y}) = log(L(\hat{\theta}|\mathbf{y})$, in which, $L(\hat{\theta}|\mathbf{y})$ and p are the maximum value of the likelihood function and the number of parameters in the model, respectively. The best model is the one that has the smallest MSE, AIC and BIC values. At this point, one could opt to use another criterion of model choice, such as, for instance, the likelihood ratio test (LRT). However, as discussed by [18], there are two complications in the use of LRT for non-nested models: (1) the asymptotic distribution of the LRT under the null hypothesis will not, in general, be chi-squared, and can be difficult to evaluate mathematically; and (2) it is not clear whether model A or model B should be treated as the null model when performing the hypothesis test. In addition, if one assumes that the LRT is chosen from the three considered models, we need to make three tests (exponential $\times$ logistic), (exponential $\times$ Gompertz), and (logistic $\times$ Gompertz). Due to these two issues and the simplicity of obtaining the values of AIC and BIC using the R software, this led us to consider them as model selection criteria instead of the LRT.

Without loss of generality and for the facility of presentation, we opt to describe how to determine the change points d_j in the next section, which describes the updating of the model as the pandemic evolved in the State of São Paulo.

3. Results

In this section, we present the modeling procedure adopted in the course of 411 days (from 17 March 2020 to 30 April 2021) of the pandemic in the state of São Paulo, Brazil. The first analysis was done thirty days after the record of the first death, and then, every thirty days, the model was updated.

Our choice for updating the model to every thirty days is based on the fact that, on average, the first symptoms of COVID-19 manifest between 4 to 14 days; and that the patients with grave symptoms remain an average time of 21 days in intensive clinical care. Then, it is very plausible that the changes in the evolution in the number of deaths occur between 25 to 35 days. Due to this, we opt to update the model every 30 days. However,

this does not restrict the approach and a user can update the model by using another period of time, such as 20, 10 or 5 days.

3.1. Fitting of a Single Growth Model

For the first analysis, consider the cumulative number of deaths recorded in the first thirty days since the recording of the first death, i.e., the period from $t = 0$ (17 March 2020) to $t = 29$ (15 April 2020). Let $\mathbb{D}_1 = \{y_0, \ldots, y_{29}\}$ be the cumulative number of deaths log-transformed.

Once the dataset $\mathbb{D}_1$ is defined, we fit the three log-growth models described in Table 1. Table 2 shows the MSE, AIC and BIC values for the three fitted log-growth models. The smallest values are highlighted in bold. As one can note, the log-Gompertz model is the best model.

Table 2. MSE, AIC and BIC values for the fitted models for dataset $\mathbb{D}_1$.

Model	Criterion		
	MSE	AIC	BIC
log-exponential	0.2776	52.6845	56.8881
log-logistic	0.0968	23.0901	28.6949
log-Gompertz	**0.0306**	**−11.5014**	**−5.8966**

Table 3 shows the point estimates and the confidence intervals (95%) for the parameters of the log-Gompertz model. The fitted model is given by

$$\hat{Y}_t^{(1)} = 7.0025 - 6.97 \exp\{-0.0861t\},$$

for $t \geq 0$, where $\hat{Y}_t^{(1)}$ denotes the fitted model in the first analysis.

Table 3. Parameter estimates for model $Y_t^{(1)}$.

Values	Parameter		
	α_1	α_2	α_3
Estimates	7.0025	6.9700	0.0861
C. interval	(6.6901, 7.3981)	(6.6779, 7.2712)	(0.0734, 0.0994)

Figure 3a shows the values of the dataset $\mathbb{D}_1$ and the fitted model (black line) for a period of 60 days (from day 0 to day 59), with 30 days of fitting and 30 days of projections. Figure 3b shows the graphic in the original scale. According to the point-estimates presented in Table 3, the projection for the maximum number of deaths is 1,099 (rounded $\exp\{7.0024\}$ value) and the ordinate of the inflection point is $T = \frac{log(6.97)}{0.0861} = 22.5507$, i.e., the peak of the pandemic was projected to occur on the 23rd day (08 April 2020) with 484 deaths. Until the 23rd day, 496 deaths were registered. That is, the projected value for the 23rd day presented an absolute percentage error of 2.42% in relation to the recorded value. Although the projections do not indicate a very critical future situation, these results should be interpreted with great caution, since the recorded values on these thirty days represent only the beginning of the pandemic.
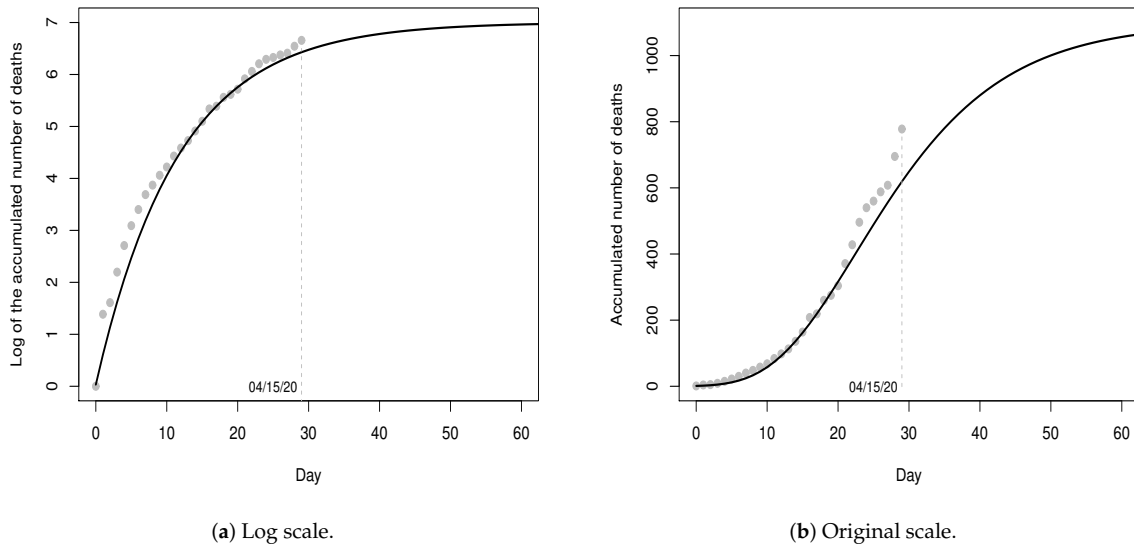
(**a**) Log scale.

(**b**) Original scale.

Figure 3. Data from 17 March 2020 to 15 April 2020 and fitted model $\hat{Y}_t^{(1)}$, for $t \geq 0$.

Unfortunately, the recorded values in the next thirty days (from 16 April 2020 to 15 May 2020) did not follow the projections of the initial fitted model since the recorded values were all above the projections, as shown in Figure 4. Due to this, we insert these thirty cumulative values (log-transformed) into the dataset $\mathbb{D}_1$ and obtain the dataset $\mathbb{D}_1^u = \{\mathbb{D}_1\} \cup \{y_{30}, \ldots, y_{59}\}$, which we refer to as updated $\mathbb{D}_1$. Then, we update the model considering the dataset $\mathbb{D}_1^u$.

Once the dataset $\mathbb{D}_1^u$ is defined, we repeat the fitting procedure of the three log-growth models. Table 4 shows the MSE, AIC and BIC values for the three fitted log-growth models. The smallest values are highlighted in bold. Again, the log-Gompertz model is the best model.
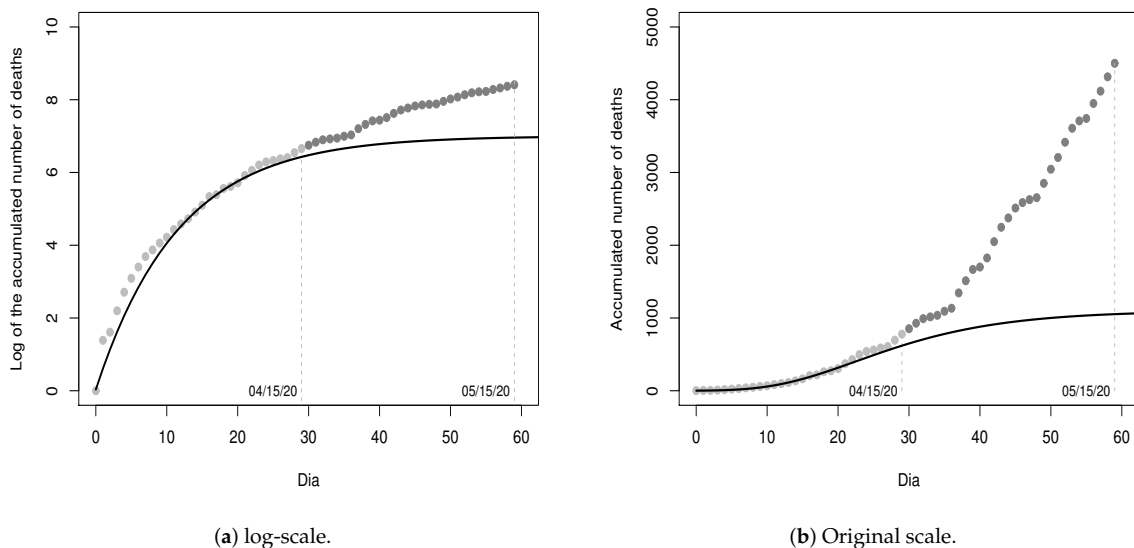

(**a**) log-scale.

(**b**) Original scale.

Figure 4. Data from 17 March 2020 to 15 May 2020 and fitted model $\hat{Y}_t^{(1)}$, for $t \geq 0$.

Table 5 shows the point estimates and the confidence intervals (95%) for the parameters of the log-Gompertz model. The fitted model is given by

$$\hat{Y}_t^{(2)} = 8.5646 - 7.4705 \exp\{-0.05t\},$$

for $t > 0$, where $\hat{Y}_t^{(2)}$ denotes the fitted model in the second analysis.

Table 4. MSE, AIC and BIC values for the fitted models for dataset $\mathbb{D}_1^u$.

Model	Criterion		
	MSE	AIC	BIC
log-exponential	0.4724	131.2807	137.5637
log-logistic	0.1700	71.9576	80.3350
log-Gompertz	**0.0473**	**−4.8170**	**3.5604**

Table 5. Parameter estimates for model $Y_t^{(2)}$.

Values	Parameter		
	α_1	α_2	α_3
Estimates	8.5646	7.4705	0.0500
C. interval	(8.3497, 8.8117)	(7.2472, 7.6960)	(0.0452, 0.0551)

Figure 5 shows the values of $\mathbb{D}_1^u$ and the curves of the models fitted on the 30th and 60th day, respectively, in the log-scale and in the original scale for a period of 90 days, with sixty days of fitting and thirty days of projection.
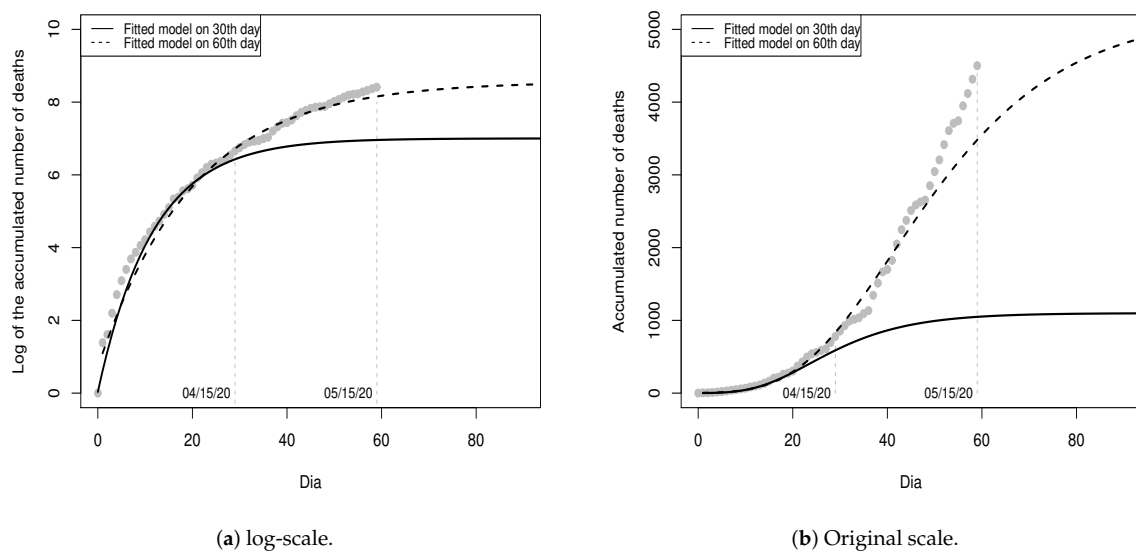

(**a**) log-scale.

(**b**) Original scale.

Figure 5. Data from 17 March 2020 to 15 May 2020 and fitted models on 30th $\left(\hat{Y}_t^{(1)} \right)$ and 60th day $\left(\hat{Y}_t^{(2)} \right)$.

The MSE of the updated model $\hat{Y}_t^{(2)}$ is 0.0473, for $t \geq 0$. That is, an MSE value greater than that of the first fitted model. Besides, as one can note, the last ten recorded values are far from the projected values by the updated model and residuals are all positive, making them serially correlated and, hence, violating the initial assumption that residuals are uncorrelated. This result made us conjecture that the fit of a single growth model may not be adequate since a new period with a growth rate and/or growth type different from the first 30 days could be beginning.

In order to empirically verify our conjecture, we registered the cumulative values in the next thirty days after the date 15 May 2020 (from 16 May 2020 to 14 June 2020) and plot these values as shown in Figure 6. As one can note, our conjecture is very plausible. Due to this, hereafter, we adopt the fit of a piecewise growth model.
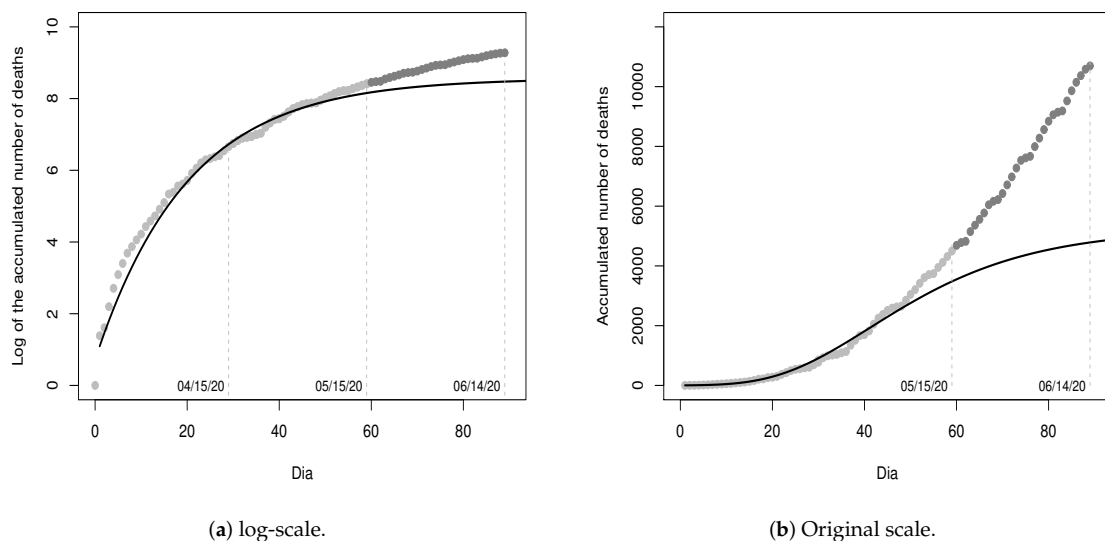
(**a**) log-scale.

(**b**) Original scale.

Figure 6. Data from 17 March 2020 to 14 June 2020 and fitted models on 30th $\left(\hat{Y}_t^{(1)}\right)$ and 60th day $\left(\hat{Y}_t^{(2)}\right)$.

3.2. Fitting of a Piecewise Growth Model

Consider $\mathbb{D} = \{y_0, \ldots, y_{89}\}$ as the cumulative number of deaths, log-transformed, recorded in the first ninety days since the first case. Let $\mathbb{D}_1$ and $\mathbb{D}_2$ be two sub-datasets of $\mathbb{D}$ representing two distinct periods of the pandemic. In order to define the change point d and to obtain the sub-datasets $\mathbb{D}_1$ and $\mathbb{D}_2$, we adopt the following approach. Let $\mathcal{G} = \{18, \ldots, 58\}$ be a grid from 18 to 58 with increments of size 1. Each increment of one unit represents the increase of one day in the pandemic. Then, we define the following 41 scenarios: $\mathbb{D}_1 = \{y_0, \ldots, y_d\}$ and $\mathbb{D}_2 = \{y_{d+1}, \ldots, y_{89}\}$, for $d \in \mathcal{G}$.

For each of the scenarios, we fit the three growth models to sub-datasets $\mathbb{D}_1$ and $\mathbb{D}_2$. We select the best model for each of the sub-datasets by using the MSE, AIC and BIC as a criterion. In addition, the best d value is the one associated with the fitted model that has the lowest MSE value.

According to the criteria considered, the best model is composed of a log-Gompertz model for $\mathbb{D}_1$ and $\mathbb{D}_2$. Figure 7 shows the MSE values of the fitted piecewise model for each d value considered, $d \in \mathcal{G}$. The point $d = 20$ has lead to a piecewise model with the smallest MSE value. Due to this, we set up $d = 20$ as the separation point for $\mathbb{D}_1$ and $\mathbb{D}_2$.

Thus, let $\mathbb{D}_1 = \{y_0, \ldots, y_{20}\}$ and $\mathbb{D}_2 = \{y_{21}, \ldots, y_{89}\}$. Table 6 shows the MSE, AIC and BIC values for the three fitted growth models for $\mathbb{D}_1$ and $\mathbb{D}_2$. The smallest values are highlighted in bold, indicating the log-Gompertz model for $\mathbb{D}_1$ and $\mathbb{D}_2$.

Table 7 shows the point estimates and the confidence intervals (95%) for parameters of the log-Gompertz model for $\mathbb{D}_1$ and $\mathbb{D}_2$, respectively. The fitted model is given by the following piecewise model

$$\hat{Y}_t^{(3)} = \begin{cases} 6.0242 - 6.3704 \exp\{-0.1227t\} & \text{, for } 0 \leq t < 21; \\ 10.0677 - 4.3578 \exp\{-0.0242t\} & \text{, for } t \geq 21; \end{cases},$$

where $\hat{Y}_t^{(3)}$ denotes the fitted model in the third analysis.

Table 6. MSE, AIC and BIC values for the models fitted for datasets $\mathbb{D}_1$ and $\mathbb{D}_2$.

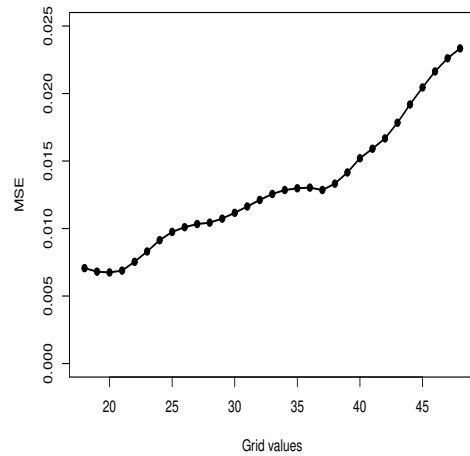| Model	Dataset $\mathbb{D}_1$			Dataset $\mathbb{D}_2$		
	MSE	AIC	BIC	MSE	AIC	BIC
log-exponential	0.2012	30.6927	33.6799	0.0456	−11.5210	−4.7755
log-logistic	0.0716	12.0195	16.0024	0.0046	−169.9950	−161.0010
log-Gompertz	**0.0242**	**−9.6880**	**−5.7051**	**0.0017**	**−237.1040**	**−228.1100**

Figure 7. MSE values for fitted models according to the grid $\mathcal{G}$.

Table 7. Estimates for the piecewise model parameters $Y_t^{(3)}$.

Parameter	Dataset	
	$\mathbb{D}_1$	$\mathbb{D}_2$
α_1	6.0242	10.0677
	(5.6457, 6.5511)	(9.9692, 10.1756)
α_2	6.3704	4.3578
	(6.0233, 6.7403)	(4.2767, 4.4464)
α_3	0.1227	0.0242
	(0.0984, 0.1484)	(0.0230, 0.0254)

Figure 8 shows the values of the sub-datasets $\mathbb{D}_1$ and $\mathbb{D}_2$ and the fitted model $\hat{Y}_t^{(3)}$ (black line) for a period of 120 days (from day 0 to day 119), with 90 days of fitting and 30 days of projection, for $t \geq 0$. The MSE of the fitted piecewise model is 0.0067. This value is smaller than the MSE of the model fitted in Section 3.1, meaning that the piecewise model fits the data better.

At this point, it is important to note that future projections are given by the log-Gompertz model fitted for sub-dataset $\mathbb{D}_2$. According to point estimates for the parameters of this model, the projection for the maximum number of cases is 23,569 (rounded $\exp\{10.0677\}$ value) and the ordinate of the inflection point is $T = \frac{log(4.3578)}{0.0242} = 60.8251$, i.e., the peak of the pandemic was projected to occur on the 81st day $(20 + 61)$ day (17 May 2020) with 8710 deaths. There were 8842 deaths recorded by the 81st day, an absolute percentage error of 1.49% in relation to the real value.
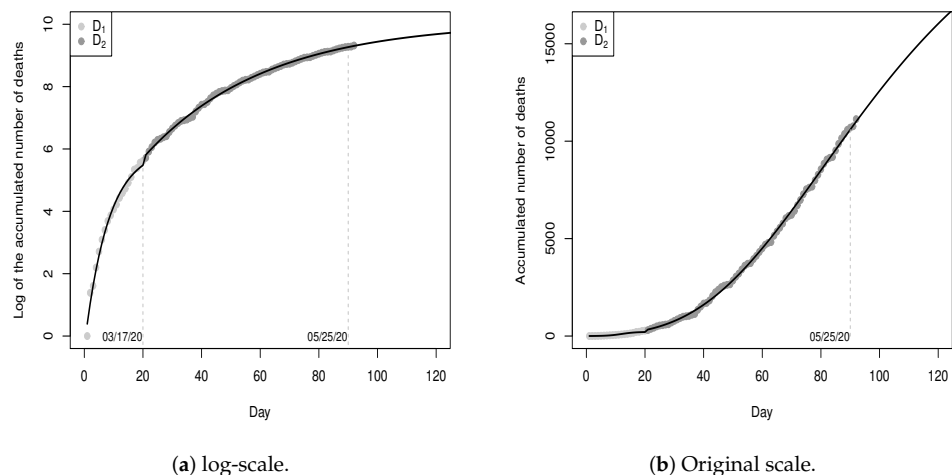

(**a**) log-scale.

(**b**) Original scale.

Figure 8. Data from 26 February 2020 to 25 May 20 and fitted piecewise model $\hat{Y}_t^{(3)}$, $t \geq 0$.

3.3. Updates of the Piecewise Model

Following the procedure described in Section 3.2, we update the piecewise model every thirty days. However, in this section, we focus on describing the modeling procedure adopted to obtain the piecewise model. That is, we focus on the periods at which the change in the pandemic's growth behavior happened.

In the updating ran on the 120th day (fourth analysis), the log-Gompertz model for $\mathbb{D}_2$ was maintained. However, on the update run on the 150th day (fifth analysis), it was possible to identify a third period of the pandemic. For the identification of this third period, we consider the following procedure. Let $\mathbb{D}_2 = \{y_{21}, \ldots, y_d\}$ and $\mathbb{D}_3 = \{y_{d+1}, \ldots, y_{149}\}$, for $d \in \mathcal{G} = \{60, \ldots, 110\}$, where $\mathcal{G}$ is a grid from 60 to 110 with increments of size 1, representing possible change points of the pandemic. Repeating the procedure of fitting the three growth models for each sub-dataset, we obtain that the log-Gompertz is the best model for both sub-datasets.

Figure 9 shows the MSE values from models fitted for $d = \{90, \ldots, 100\}$. As one can note, $d = 92$ is the best change point, i.e., the point with the smallest MSE value. Thus, we set $\mathbb{D}_2 = \{y_{21}, \ldots, y_{92}\}$ and $\mathbb{D}_3 = \{y_{93}, \ldots, y_{149}\}$.

The fitted model is given by the following piecewise model

$$\hat{Y}_t^{(5)} = \begin{cases} 6.0242 - 6.3704 \exp\{-0.1227\} & 0 \leq t < 21; \\ 10.0707 - 4.3603 \exp\{-0.0241t\} & \text{, for } 21 \leq t < 92; \\ 10.8892 - 1.5467 \exp\{-0.0132t\} & \text{, for } t \geq 92; \end{cases}$$

where $\hat{Y}_t^{(5)}$ denotes the fitted model in the 5-th analysis.

Figure 10 shows the values of the sub-datasets $\mathbb{D}_1$, $\mathbb{D}_2$ and $\mathbb{D}_3$ and the fitted model $\hat{Y}_t^{(5)}$ (black line) for a period of 180 days (from day 0 to day 179), with 150 days of fitting and 30 days of projection. The MSE of the fitted piecewise model is 0.0004811.

Following this procedure, after 14 analysis steps, we identify seven different periods of the pandemic with change points on the days $d = \{20, 92, 240, 295, 363, 381\}$. The piecewise growth model has the following configuration: log-Gompertz for $\mathbb{D}_1 = \{y_0, \ldots, y_{20}\}$, $\mathbb{D}_2 = \{y_{21}, \ldots, y_{92}\}$, $\mathbb{D}_3 = \{y_{93}, \ldots, y_{240}\}$, log-exponential for $\mathbb{D}_4 = \{y_{241}, \ldots, y_{295}\}$, $\mathbb{D}_5 = \{y_{296}, \ldots, y_{363}\}$, $\mathbb{D}_6 = \{y_{364}, \ldots, y_{381}\}$ and log-Gompertz for $\mathbb{D}_7 = \{y_{382,\ldots}\}$. Its mathematical expression is given by

$$\hat{Y}_t^{(14)} = \begin{cases} 6.0242 - 6.3704 \exp\{-0.1227t\} & \text{, for } 0 \leq t < 21; \\ 10.1376 - 4.4034 \exp\{-0.0229t\} & \text{, for } 21 \leq t < 93; \\ 10.7378 - 3.1148 \exp\{-0.0158t\} & \text{, for } 93 \leq t < 241; \\ 10.5955 + 0.0030t & \text{, for } 241 \leq t < 296; \\ 10.7666 + 0.0043t & \text{, for } 296 \leq t < 364; \\ 11.0592 + 0.0091t & \text{, for } 364 \leq t < 382; \\ 11.7423 - 0.5207 \exp\{-0.0226t\} & \text{, for } t \geq 382; \end{cases} \tag{3}$$

where $\hat{Y}_t^{(14)}$ denotes the fitted model in the 14-th analysis step.

Figure 11 shows the recorded values and the curve of the fitted piecewise model $\hat{Y}_t^{(14)}$, in the log and original scale, for a period of 600 days. The MSE of the fitted model is 0.0042, $t \geq 0$. At this point of the analysis, the future projection is given by the log-Gompertz model fitted for sub-dataset $\mathbb{D}_7$. According to the point estimated for the parameters of this model, the projection for the maximum number of deaths is 122,296 ($\exp\{11.7142\}$). In other words, if the current scenario is maintained, it is expected that there will be additional 26,105 deaths.
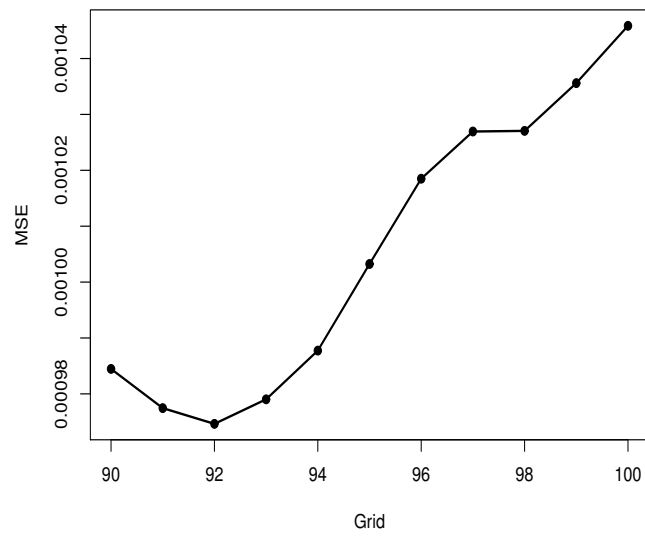
Figure 9. MSE values for fitted models according to the change point *d*.
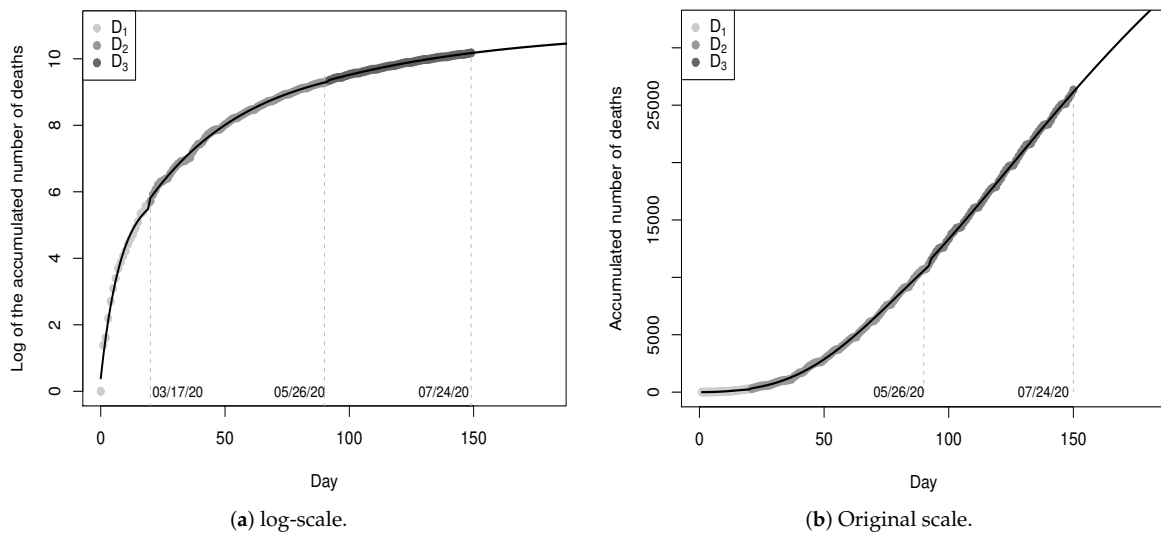

(**a**) log-scale.

(**b**) Original scale.

Figure 10. Data from 26 February 2020 to 25 May 20 and fitted piecewise model $\hat{Y}_t^{(5)}$, $t \geq 0$.
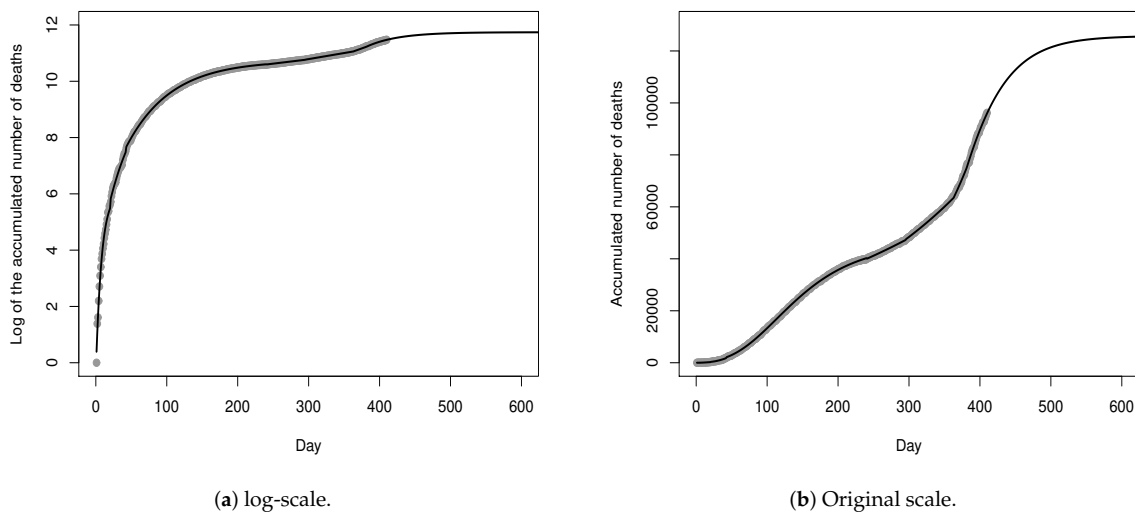

(**a**) log-scale.

(**b**) Original scale.

Figure 11. Data from 26 February 2020 to 30 April 2021 and fitted piecewise model $\hat{Y}_t^{(14)}$, $t \geq 0$.

Table [8] shows the projections of the model for the cumulative number of deaths in the last ten days of the sub-dataset $\mathbb{D}_7$ and the recorded values. The fourth line of this Table shows the absolute percentage error in relation to the real value. The biggest percentage error was 0.7877%.

Table 8. Projections and percentage error for the period from 11/09/20 to 11/16/20.

Date	04/13/21	04/14/21	04/15/21	04/16/21	04/17/21	04/18/21	04/19/21	04/20/21	04/21/21	04/22/21
Projection	90,279	90,949	91,609	92,259	92,899	93,529	94,149	94,759	95,359	95,950
Real value	90,627	90,810	91,673	92,548	92,693	92,798	93,842	94,656	95,532	96,191
error	348	139	64	289	206	731	307	103	173	241
% error	0.3840	0.1531	0.0698	0.3123	0.2222	0.7877	0.3271	0.1088	0.1811	0.2505

In addition to modeling the different pandemic periods, the fitted piecewise model also allows linking the model change points to facts that affect the speed of the epidemic spreading. For example, on the 334th day (13 February 2021) the first case of COVID-19 by the P_1 variant was registered in the state of São Paulo. After thirty-one days, on the 364th day (15 March 2021), the proposed approach has indicated a change point; changing from an exponential model with a growth rate of 0.0043 to an exponential model with a growth rate of 0.0091. That is, a change for a more aggressive pandemic phase occurred. Since the disease has an average time from 4 to 14 days to manifest the first symptoms and the patients with grave symptoms remain an average time of 21 days in intensive clinical care; then, it is very plausible to consider that the P_1 variant is one of the possible factors related to the change point.

On the other hand, due to an excessive increase in the number of deaths due to the COVID-19 in the first months of 2021, the government of the state of São Paulo has published on 6 March 2021 (45-th day) a decree implementing an emergency phase to contain the transmission of the disease. In this decree, a curfew from 8 pm to 5 am and the closing of commerce was adopted. Twenty-eight days after the decree, the proposed approach has identified a change from the exponential model to a Gompertz model (day 382—4 February 2021). That is, a change to a "little better" phase occurred because, in this new phase, at least, it is possible to make a projection for the occurrence of a peak. Although the modeling procedure does not allow us to associate directly the publication of the decree with the "new" pandemic phase, it is very plausible to consider that the adoption of the decree has contributed to this change.

To finalize this section, we inform the reader that the fitted piecewise model in Equation ([3]) is not absolutely continuous. However, in the opposite to the fit of a single growth model, the absolute continuity of a piecewise model is not a fundamental property because the aim is to identify change points at which model changes happen. Thus, the final model may have a jump in the change points. However, as is expected for a growth model, the piecewise fitted model has the following properties:

- $f(t)$ is non-decreasing;
- $f(t)$ is continuos to the right side of the change points;
- $\lim\limits_{t \to -\infty} f(t) = 0$ and $\lim\limits_{t \to +\infty} f(t) = \alpha_1^*$, where α_1^* is the upper asymptote of the last function.

3.4. Piecewise Growth Model for the Cumulative Number of Cases

As another illustration of the good performance of the proposed method, we fit a piecewise model for the cumulative number of cases recorded in the State of São Paulo, Brazil, in the period from 26 February 2020 to 30 April 2021. In this period, 2,903,709 cases of COVID-19 were recorded.

The cumulative number of cases also presented a heterogeneous behavior over time. In the course of analysis, we identify eight different periods for the number of cases with change points on the days $d = \{34, 93, 144, 256, 310, 346, 376\}$. However, at this point, it is important to emphasize that the number of cases is directly connected with the number of

diagnostic tests performed. Thus, when a change in the pandemic's behavior is detected, this may be due to increased testing.

The fitted piecewise model has the following mathematical expression

$$
\hat{Z}_t^{(14)} = \begin{cases}
7.5950 - log(1 + 4214.0455 \exp\{0.2850\}) & \text{, for } 0 \le t < 35; \\
13.1173 - 5.2413 \exp\{-0.0192t\} & \text{, for } 35 \le t < 94; \\
14.3067 - 2.8624 \exp\{-0.0143t\} & \text{, for } 94 \le t < 145; \\
14.0132 - 1.1061 \exp\{-0.0228t\} & \text{, for } 145 \le t < 257; \\
13.9311 + 0.0048t & \text{, for } 257 \ge t < 311; \\
14.1852 + 0.0067t & \text{, for } 311 \ge t < 347; \\
14.4121 + 0.0047t & \text{, for } 347 \ge t < 377; \\
15.7549 + 1.1918 \exp\{-0.0061t\} & \text{, for } t \ge 377;
\end{cases}
\tag{4}
$$

where $\hat{Z}_t^{(14)}$ denotes the fitted model for the number of cases in the 14-th analysis.

Figure 12 shows the recorded values and the curve of the fitted piecewise model $\hat{Z}_t^{(14)}$, in the original scale, for a period of 1000 days. The MSE of the fitted model is 0.0042. At this point of the analysis, the future projections are given by the log-Gompertz model fitted for sub-dataset $\mathbb{D}_8$. According to point estimates for parameters of this model, the projection for the maximum number of cases is 6,954,504 ($\exp\{15.7549\}$).
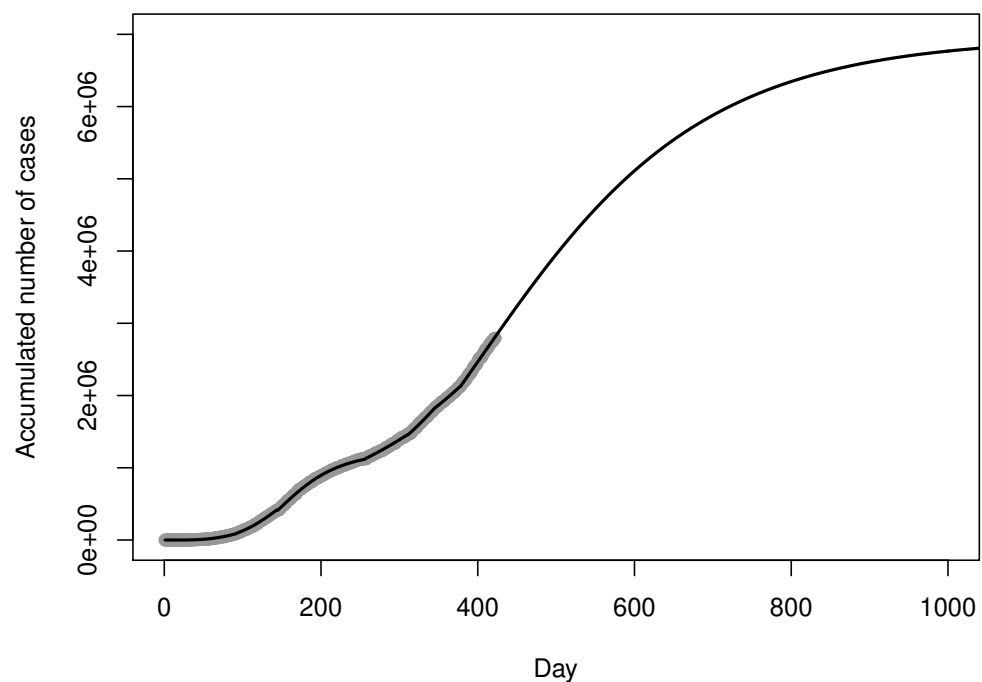

Figure 12. Fitted piecewise model for the cumulative number of cases.

4. Final Remarks

In this paper, we describe a case study on the evolution of the COVID-19 pandemic in the state of São Paulo, Brazil. The main aim of the study was to fit a piecewise growth model in order to be able to explain the different periods of the pandemic's evolution. In addition, there is also the interest in being able to give a short-term forecast for the cumulative number of cases and predict the peak point of the pandemic for each period of the pandemic.

The modeling procedure was developed in the course of time by updating the model every thirty days. In addition, for each update, we also verify the existence of two different periods of the pandemic. In the affirmative case, we separate the dataset of the period into two sub-datasets and fit a growth model for each of the sub-datasets. Overall, we identify seven periods of the pandemic for the cumulative number of deaths. To each period of the

pandemic, we fit three non-linear growth models: exponential, logistic and Gompertz. In order to select the best model for each sub-dataset, we consider the MSE, the AIC and BIC as criteria.

The fitted piecewise model for the number of deaths is a mix of the Gompertz model and the exponential model. From a practical viewpoint, this can be viewed as the main advantage of the proposed method, because, it is capable of explaining each period by different values of pandemic parameters, such as the peak and the growth rate for each period. In addition, the results also show that the fit of a piecewise model is more effective than the fit of a single growth model.

From a practical viewpoint, the identification of the change points shows for government agencies that some containment procedures of the transmission of the disease need to be implemented; or, if some containment procedures implemented are working. For example, if the data from a period are explained by a Gompertz model, then there is a projection for the peak of the pandemic (inflection point) and an estimate for the maximum number of cases (upper asymptote). Under this scenery, the government's agents may elaborate better strategies for containment of the transmission of the virus. However, if a change point is identified, in which, the recorded data after this change point are explained by an exponential model, then there is a new situation without a projection for a peak. That is, the situation changed to a more aggressive phase of the pandemic. This scenery shows that that more restrictive containment strategies need to be implemented as soon as possible.

On the other hand, if we identify a change point from a period explained by an exponential model for a period explained by a Gompertz model; this indicates that a change happened towards "a little better" situation because now, at least, it is possible to project the peak of the pandemic. This information may be used by government agencies in order to elaborate and justify the adoption of containment strategies of the transmission of the disease, and, in this way, to obtain a flattening of the Gompertz curve and, consequently, anticipate the occurrence of the peak and reduce the number of deaths (decrease the upper asymptote value).

Both examples described above show that it is important to consider the fit of a piecewise growth model, such as the one in our proposal, in contrast to fitting a single growth model. Although the paper has been developed considering only three growth models (exponential, logistic, and Gompertz), the procedure presented can be easily adapted for other kinds of growth models. The computational codes used for fitting the models are in the R language and can be requested by contacting the authors via e-mail.

Author Contributions: E.F.S. and C.A.d.B.P. have developed the whole theoretical part of the research. E.F.S. has implemented the computational codes in the R software. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: https://www.seade.gov.br/coronavirus (accessed on 13 July 2021).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Growth Model

In order to model the cumulative number of deaths recorded in the State of São Paulo Brazil, let Y_t as defined in Equation (2) of the paper.

Appendix A.1. Exponential Growth Model

One of the best known nonlinear models to describe the growth of an epidemic disease is the exponential model. Assuming this model, the cumulative number of deaths is given by

$$X_t = f(t|\theta) = \alpha_1^* \exp\{\alpha_2 t\}, \tag{A1}$$

for $\theta = (\alpha_1^*, \alpha_2)$, where α_1^* is the number of deaths in the initial time, $t = 0$, and α_2 is the growth rate, for $t \geq 0$.

Taking the logarithmic transformation on both sides of Equation (A1), one obtains the following linearized form of the model, called the log-exponential model,

$$Y_t = log(X_t) = \alpha_1 + \alpha_2 t,$$

where $\alpha_1 = log(\alpha_1^*)$, for $t \geq 0$.

Figure A1 shows the exponential and log-exponential functions for an initial value $\alpha_1^* = 2$ and growth rate $\alpha_2 = \{0.10, 0.20, 0.40\}$. As one can note in Figure A1, increasing the value of α_2 increases the rate of growth curve, meaning a fast growth of the number of disease death. In addition, the curve of this model grows infinitely regardless of the population size. This can be viewed as a practical problem since the number of cumulative deaths is bound, for instance, by the population size.
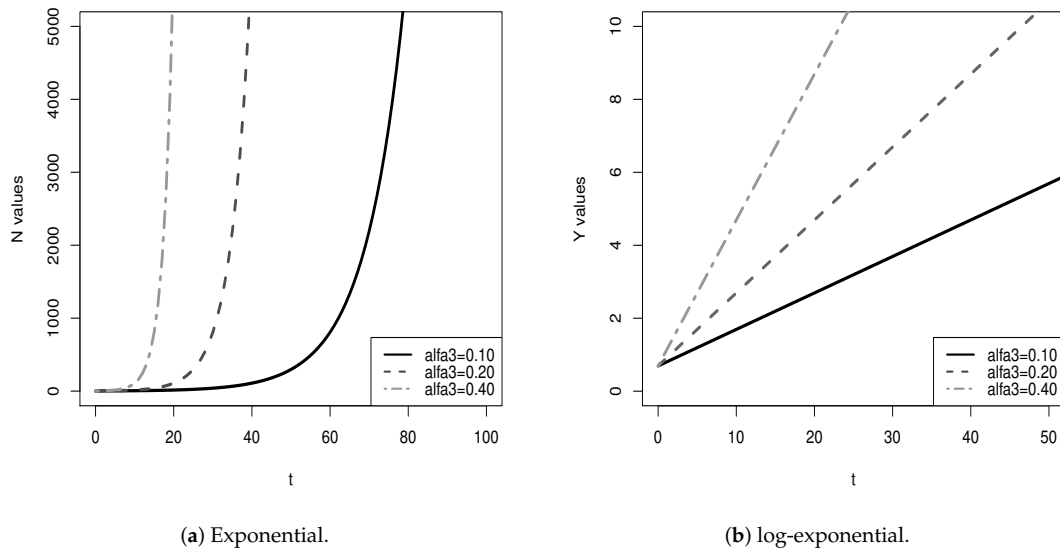

(**a**) Exponential. (**b**) log-exponential.

Figure A1. Exponential and log-exponential models.

Appendix A.2. Logistic Growth Model

Consider now the logistic growth model [19]. In contrast to the exponential model, this model has an S-shape curve and, consequently, a growth limit. The cumulative number of deaths is given by

$$X_t = f(t|\theta) = \frac{\alpha_1^*}{1 + \alpha_2 \exp\{-\alpha_3 t\}}, \tag{A2}$$

where $\theta = (\alpha_1^*, \alpha_2, \alpha_3)$ are the model parameters, for $t \geq 0$. The parameter α_1^* is the upper asymptote. In the context of the COVID-19, the value of α_1^* is an estimate for the maximum number of deaths. The parameter α_2 is related to the coordinates (t_m, N_{t_m}) of the inflection point, where $t_m = \frac{ln(\alpha_2)}{\alpha_3}$ and $N_{t_m} = \frac{\alpha_1^*}{2}$. The parameter α_3 is the intrinsic growth rate at the inflection point. The time-point to obtain $100(1-p)\%$ of the upper asymptote α_1^* is $t_p = \frac{log(p\alpha_2) - log(100-p)}{\alpha_3}$, for $p \in (0,1)$.

Taking the logarithmic transformation on both sides of Equation (A2), we obtain the log-logistic model,

$$Y_t = log(X_t) = \alpha_1 - log(1 + \alpha_2 \exp\{-\alpha_2 t\}),$$

where $\alpha_1 = log(\alpha_1^*)$, for $t \geq 0$.

Figure A2 shows the logistic and log-logistic models for an upper asymptote $\alpha_1^* = 10,000$, $\alpha_2 = 5$ and $\alpha_3 = \{0.10, 0.15, 0.30\}$. Similar to the exponential model, by increasing the value of the parameter α_3, the curve is more inclined. In these Figures, the symbol • represents the inflection point and the symbol ⋄ represents the time-point needed to reach 80% of the α_1 value.

Appendix A.3. Gompertz Growth Model

As the third growth model, consider the Gompertz model [20,21]. This model is defined as

$$X_t = f(t|\theta) = \alpha_1^* \exp\{-\alpha_2 \exp\{-\alpha_3 t\}\}, \tag{A3}$$

where $\theta = (\alpha_1^*, \alpha_2, \alpha_3)$ are the model parameters. The interpretation of the parameters is similar to that described for the logistic model.

The main difference in relation to the logistic model is that the curve of the Gompertz model is not symmetric in relation to the inflection point. While the ordinate value of the inflection point of the logistic model is $\frac{\alpha_1^*}{2}$, for the Gompertz model, this value is approximately 37% of the α_1^* value. The time-point to obtain $100(1 - p)\%$ of the upper asymptote α_1^* is $t_p = \frac{log(\alpha_2) - log(-log(p))}{\alpha_3}$, for $p \in (0, 1)$.

Taking the logarithmic transformation on both sides of Equation (A3), we obtain the log-Gompertz model,

$$Y_t = log(X_t) = \alpha_1 - \alpha_2 \exp\{-\alpha_3 t\}, \quad \text{for } t \geq 0.$$

where $\alpha_1 = log(\alpha_1^*)$, for $t \geq 0$.
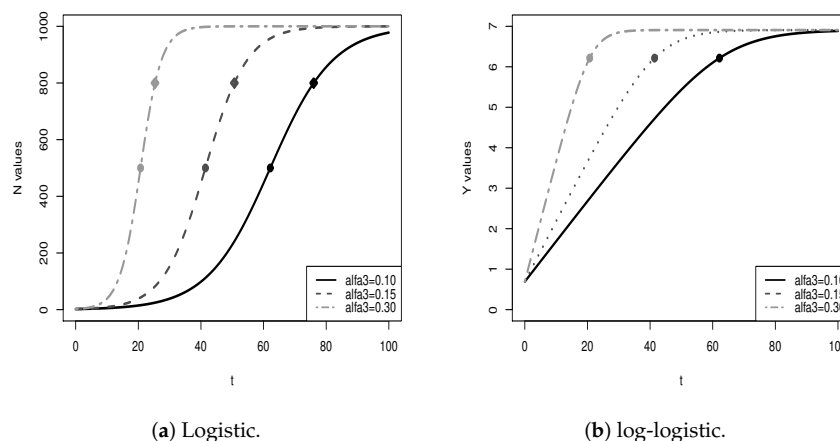

(**a**) Logistic.　　　　　　　　(**b**) log-logistic.

Figure A2. Logistic and log-logistic models.

Figure A3 shows the Gompertz and log-Gompertz models for an upper asymptote $\alpha_1^* = 10,000$, $\alpha_2 = 5$ and $\alpha_3 = \{0.10, 0.20, 0.40\}$. Similar to the exponential and logistic models, by increasing the value of the parameter α_3, the curve is more inclined. In these Figures, the symbol • represents the inflection point and the symbol ⋄ represents the time-point needed to reach 80% of the α_1^* value.

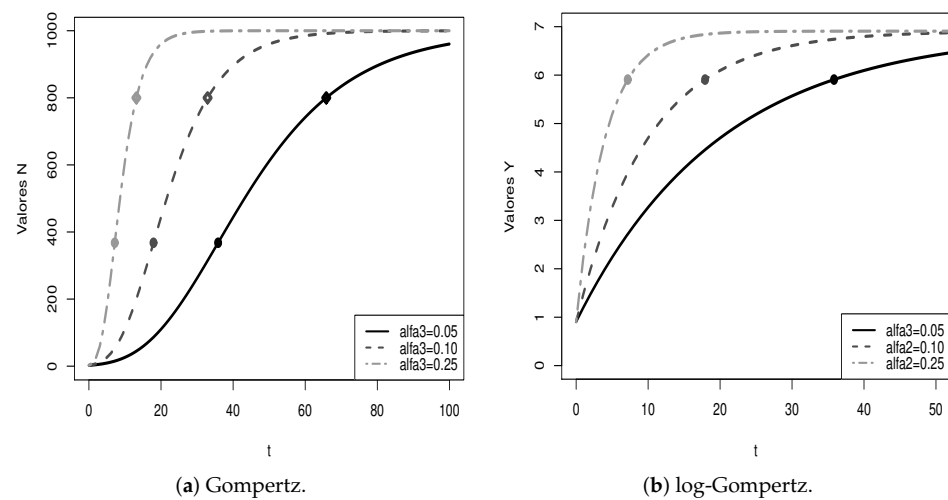(**a**) Gompertz. (**b**) log-Gompertz.

Figure A3. Gompertz and log-Gompertz models.

References

1. World Health Organization (WHO). Director-General's Opening Remarks at the Media Briefing on COVID-19—30 July of 2020. 2020. Available online: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---30-july-2020 (accessed on 1 February 2021).
2. Musa, S.S.; Zhao, S.; Wang, M.H. Estimation of exponential growth rate and basic reproduction number of the coronavirus disease 2019 (COVID-19) in Africa. *Infect. Dis. Poverty* **2020**, *9*, 96. [CrossRef] [PubMed]
3. Aviv-Sharon, E.A. Generalized logistic growth modeling of the COVID-19 pandemic in Asia. *Infect. Dis. Model.* **2020**, *5*, 502–509. [CrossRef] [PubMed]
4. Vasconcelos, G.L.; Macêdo, A.M.S.; Ospina, R.; Almeida, F.A.G.; Duarte-Filho, G.C.; Brum, A.A.; Souza, I.C.L. Modelling fatality curves of COVID-19 and the effectiveness of intervention strategies. *PeerJ* **2020**, *8*, e9421. [CrossRef] [PubMed]
5. Wu, K.; Darcet, D.; Wang, Q. Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in China and in the rest of the world. *Nonlinear Dyn.* **2020**, *101*, 1561–1581. [CrossRef] [PubMed]
6. Valle, J.A.M. Predicting the number of total COVID-19 cases and deaths in Brazil by the Gompertz model. *Nonlinear Dyn.* **2020**, *102*, 2951–2957. [CrossRef] [PubMed]
7. Mazurek, J.; Rico, C.P.; Garcia, C.F. Forecasting the Number of COVID-19 Cases and Deaths in the World, UK, Russia and Turkey by the Gompertz Curve. 2020. Available online: https://www.researchgate.net/publication/341132093_Forecasting_the_number_of_COVID-19_cases_and_deaths_in_the_World_UK_Russia_and_Turkey_by_the_Gompertz_curve?channel=doi&linkId=5eb042d6299bf18b9594bc43&showFulltext=true (accessed on 23 July 2021).
8. Budimulyati, L.S.; Norr, R.R.; Saefuddin, A.; Talib, C. Comparison on accuracy of Logistic, Gompertz and Von Bertalanffy models in predicting growth of new born calf until first mating of holstein friesian heifers. *J. Indones. Trop. Anim. Agric.* **2012**, *3*, 151–160. [CrossRef]
9. Saraiva, E.F.; Sauer, L.; Pereira, B.B.; Pereira, C.A.B. A piecewise growth model for modeling the accumulated number of COVID-19 cases in the city of Campo Grande. *Rev. Barsileira Biom.* **2021**, *39*, 240–265. [CrossRef]
10. Akaike, H.A. New look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
11. Bozdogan, H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrica* **1987**, *52*, 345–370. [CrossRef]
12. Schwarz, G.E. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
13. Vieira, S.; Hoffmann, R. Comparison of the Logistic and the Gompertz growth functions considering additive and multiplicative error terms. *Appl. Stat.* **1977**, *26*, 143–148. [CrossRef]
14. Hsieh, Y.H. Temporal patterns and geographic heterogeneity of zika virus (zikv) outbreaks in french polynesia and central america. *PeerJ* **2017**, *5*, e3015. [CrossRef] [PubMed]
15. Cran, R. *R: A Language and Environment for Statistical Computing*; R foundation for Statistical Computing, Vienna, Austria, 2020; ISBN 3-900051-07-0. Available online: http://R-project.org (accessed on 23 July 2021).
16. Florent, B.; Ritz, C.; Charles, M.; Brustsche, J.P.; Flandrois, P.; Delignette-Muller, M.L. A toolbox for nonlinear regression in R: The package nlstools. *J. Stat. Softw.* **2015**, *66*, 1–21.
17. Pinheiro, J.; Bates, D.; Debroy, S.; Sakar, D.; R Core Team. Linear and Nonlinear Mixed, R Package Version 3.1-147. 2020. Available online: https://cran.r-project.org/web/packages/nlme/ (accessed on 23 July 2021)
18. Lewis, F.; Butler, A.; Gilbert, L. A unified approach to model selection using the likelihood ratio test. *Methods Ecol. Evol.* **2010**, *2*, 155–162. [CrossRef]

19. Blumberg, A.A. Logistic Growth Rate Functions. *J. Theor. Biol.* **1968**, *21*, 42–44. [CrossRef]
20. Gompertz, B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philos. Trans. R. Soc. Lond. Biol. Sci.* **1825**, *182*, 513–585.
21. Winsor, C.P. The Gompertzcurve as a growth curve. *Proc. Natl. Acad. Sci. USA* **1932**, *18*. [CrossRef] [PubMed]