Research Article

Models to Predict the Number of Infected Cases and Deaths from COVID-19 in Chile and Its Most Affected Regions

Francisco Novoa-Muñoz (ID)

Departamento de Estadística, Universidad del Bío-Bío, Concepción, Chile

Correspondence should be addressed to Francisco Novoa-Muñoz; fnovoa@ubiobio.cl

This paper designs and implements a methodology to model the evolution of the COVID-19 pandemic, produced by the SARS-CoV-2 virus, in what was called the first wave in Chile, which lasted from March 2 to 31 October 2020. The models are based on sigmoidal growth curves and can be used to predict the number of daily infections and deaths in future days, making them a useful tool for sanitary authorities to manage an epidemic. The methodology is applied to the entire country and to each of its most affected regions. In addition, the dynamics of these models allow it to be nurtured with the new information that is being produced and forecast a tentative date on which there would be some control over the pandemic. Moreover, these models allow for predicting the total number of infected and deceased people at the time the pandemic is under control. However, the simplicity of these models, which consider only the accumulated data of those infected and deceased, does not contemplate an intervention analysis such as vaccinations, which, as is known, are being effective in controlling the pandemic.

1. Introduction

In his article, the astrophysicist Barrado [1] expresses his opinion regarding the crisis that COVID-19 is leaving and summarizes it with the significant title "Vivimos un punto de inflexión: la generación 2020 y la nueva sociedad," which translated into English is "We live at point of inflection: the 2020 generation and the new society." The author expresses that the health crisis generated by the COVID-19 pandemic, produced by the SARS-CoV-2 virus, is not the first, and unfortunately, it will not be the last that humanity will face. He also mentions that diseases have been powerful levers of historical change, having the ability to change a society, especially when combined with other disturbing elements.

To illustrate these historical changes, recall, for example, that (i) the plagues in Egypt caused notable changes in the way of life of the population since they affected the characteristics of social relationships [2]. (ii) The appearance of the Black Death in 1347 gave rise to an epidemic that covered all of Europe, causing the death of nearly a third of its population. Their socioeconomic structure completely changed [3]. (iii) The encounter between Europeans and Native Americans caused epidemics that devastated the original society. It is said that this was one of the main causes of the destruction of the native culture [4].

Regarding the consequences of these catastrophes, it can be indicated that, for those involved, such as political structures or individuals, the change was dramatic and left innumerable victims, but it also opened up new opportunities. For example, during the birth of modern states, health statistics emerged that kept an accurate record of cases of illness and death in the population, which made it possible to study epidemic phenomena [5].

Since 2020, these studies of epidemic phenomena have become widespread. In particular, just over a year ago, González et al. [6] published an article entitled "COVID-19: Pandemia de modelos matemáticos," which translated into English is "COVID-19: pandemic of mathematical models," the

authors mentioned that "the large number of mathematical models formulated to predict the evolution of the epidemic and the impact of the measures for its control are a fashionable crystal ball throughout the planet, with more or less academic and executive intention."

In turn, in their article, the authors Grillo-Ardila et al. [7] stated that "there are many mathematical models that have been developed to understand the dynamics of COVID-19 infection. However, the difference in sociocultural contexts between countries makes it necessary to specifically adjust these estimates to each scenario."

To the best of our knowledge, publications that predict confirmed cases were found, and accumulated confirmed cases or deaths of COVID-19 are based on (a) SIRS models [8–10], (b) ARIMA model [11], (c) hybrid approaches that include both ARIMA and Wavelet models [12, 13], (d) ARIMA models, cubist regression, random forest, ridge regression, support vector regression and stacking-ensemble learning [14], (e) SARIMA models [15], (f) time series models based on growth curves [16], and (g) Gompertz curves [17].

According to the authors Harvey and Kattuman [16], epidemiological models (such as those mentioned in the previous paragraph) that seek to estimate epidemic trajectories fall into two broad classes:

(1) Compartmental models consist of deterministic models that seek to be faithful to the details of the routes and processes of disease transmission. More specifically, they project how individuals in an initially susceptible (S) population become exposed (E) to the virus, potentially infected (I), and, if infected, either recover (R) from the disease or die from it [16].

In the category of compartmental models are SIR, SIS, SEIS, SIRS, SEIR, MSIR, and MSEIR models, where M represents infants with passive immunity. These models are governed by differential equations in which the compartments S, I, R, E, and M are mutually exclusive, and the sum of all is the total population.

For example, in their SIR model, Kermack and McKendrick obtained the following differential equations that describe the model:

$$S'(t) = -\beta S(t)I(t),$$
$$I'(t) = \beta S(t)I(t) - \gamma I(t), \qquad (1)$$
$$R'(t) = \gamma I(t),$$

where β is the infection rate, and $1/\gamma$ is the average time of infection.

(2) Time series models, such as autoregressive integrated moving average (ARIMA) and seasonal ARIMA, known as SARIMA, use historical data to make predictions.

Given the time series X_t where t is an integer index and X_t are real numbers, an ARIMA(p, d, q) model is a stationary model given by

$$X_t = \phi_0 - \Delta^d X_t + X_t + \sum_{i=1}^{p} \phi_i \Delta^d X_{t-i} - \sum_{i=0}^{q} \theta_i \varepsilon_{t-i}, \qquad (2)$$

where $\Delta^d X_t = X_t - X_{t-d}$, d corresponds to the number of differences that are necessary to make the original series stationary, $\phi_1, \ldots, \phi_p$ are the parameters belonging to the "autoregressive" part of the model, $\theta_1, \ldots, \theta_q$ are the parameters belonging to the "moving average" of the model, $\theta_0 = 1$, ϕ_0 is a constant, and ε_t is the error term (also called innovation or disturbance).

The publications presented in literal (a) correspond to class (1), and the publications identified in literals (b) to (f), with certain modifications and extensions, belong to the class (2).

On the other hand, the publication registered in (f) and the one described in literal (g) are the only ones that use the Gompertz function (sigmoid curve) as a basis to model the spread of COVID-19. In turn, and according to Harvey and Kuttuman [16], the progress of an epidemic typically starts off with the number of cases following an exponential growth path and over time, the growth rate falls, and the total number of cases approaches a final level: the "leveling of the curve."

Since this evolutionary curve that is drawing and characterizing an epidemic such as COVID-19 is captured in all its essence by the sigmoid curves, it was the motivation that gave rise to the present study. In this research, the Gompertz function is used to model the process generated by COVID-19 and is extended to the application of different models of exponential growth curves (sigmoid) that have the ability to capture the evolution of the number of accumulated cases, of infected and dead by COVID-19 in Chile and in its most affected regions. Furthermore, these sigmoid curves have the power to track progress towards an upper limit or saturation level (see [16]). Then, using optimization software, the parameters associated with each curve are estimated, and through comparison techniques, the model that best fits the collected data is achieved.

2. Materials and Methods

2.1. Data Collection. The data were obtained from the database managed by the Chilean government on its official page https://www.minciencia.gob.cl/covid19. It is important to note how difficult it was to obtain this data at the beginning, and usually in a format that was not editable. In addition, some transformation always had to be done, because apparently the data could be seen, but once it was extracted, it corresponded to another value.

The study of the number of accumulated and daily infections covered from March 2, 2020, when the first confirmed case of COVID-19 appeared in Chile, until October 31, 2020. Figure 1 graphically represents the data used for this investigation. On the other hand, the study of the number of accumulated and daily deaths was considered from March 22, 2020 to November 30, 2020. The gap in dates with respect to the study on those infected is due to the fact that at the beginning, there was no certainty if the people
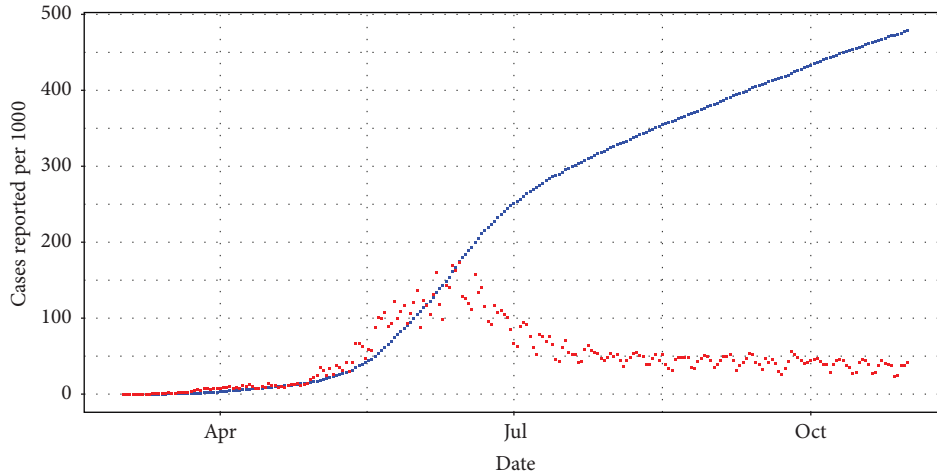
FIGURE 1: Cumulative and daily number of infected by COVID-19 in Chile from March 2 to October 31, 2020. Cumulative infected cases are shown in blue. In red, the daily cases amplified in an appropriate amount are recorded in order to better appreciate them.

who died were due to COVID-19 or another cause. Figure 2 graphically summarizes the information used for the number of deaths.

2.2. Overview. In the beginning, the scheme suggested by the authors Lega and Brown from the year 2016, described in Figure 3, was used. However, with the data that were collected, good fits were not obtained when using step 4 and step 5 of the methodology proposed by these researchers.

Upon further investigation, it was found that the authors Sánchez-Villegas and Daponte [17] used the first two steps of the Lega and Brown [18] methodology; that is,

(1) Fit the accumulated data to the three-parameter Gompertz curve, $G(t)$, given by

$$G(t) = ae^{-e^{b(t-c)}}. \tag{3}$$

(2) Calculate the first derivative, $g(t)$, of the previous curve; that is,

$$\begin{aligned} g(t) &= G'(t), \\ &= -abe^{b(t-c)}e^{-e^{b(t-c)}}, \end{aligned} \tag{4}$$

which is interpreted as the curve of daily cases.

Figure 4 shows the three graphical representations of the study by Sánchez-Villegas and Daponte [17]: (a) the daily accumulated cases, given by the points in black, (b) the Gompertz curve, which predicts the accumulated information, and (c) in red, the change curve.

The two practical interpretations that are obtained from here should be highlighted.

The first is that, through the curve of daily cases in red, the expected values in future days can be calculated which is the way used in this study to make forecasts.

The second is that the coefficient "a" of the Gompertz curve corresponds to its upper asymptote, since if $b < 0$, then

$$a = \lim_{t \longrightarrow \infty} G(t). \tag{5}$$

Parameter "a" could be interpreted as the "horizon" of the pandemic; that is, the number of cases expected when the pandemic is controlled.

As usual, when having new data, they are incorporated into those already available, and the model is recalculated according to the scheme described.

2.3. Proposed Methodology. Considering the two approaches previously analyzed and according to the historical and ongoing information available for this research, the methodology that was implemented is summarized in the following points:

S1: collect national data and for each region of both the infected and the deceased.

S2: a growth curve from three families of curves was fitted to the daily accumulated data, which are detailed in the following subsection. To do this, the parameters were estimated using the "drc" (Dose-Response Curves) package of the *R* (2021) programming language.

S3: obtain the daily change curve, the first derivative of the growth curve. With this curve of change, the number of cases expected in future days was forecast.

S4: calculate daily, the evolution of the horizons (asymptote of the accumulated growth curve), which was becoming increasingly precise. Thus, an estimate of the number of expected cases was obtained when controlling the pandemic.

S5: as soon as the new data are published, steps S2 to S4 are repeated, each time obtaining a more robust model.

This scheme was used to obtain an estimate of the number of infected and another for the number of deaths at the national level and for each of the regions of Chile.
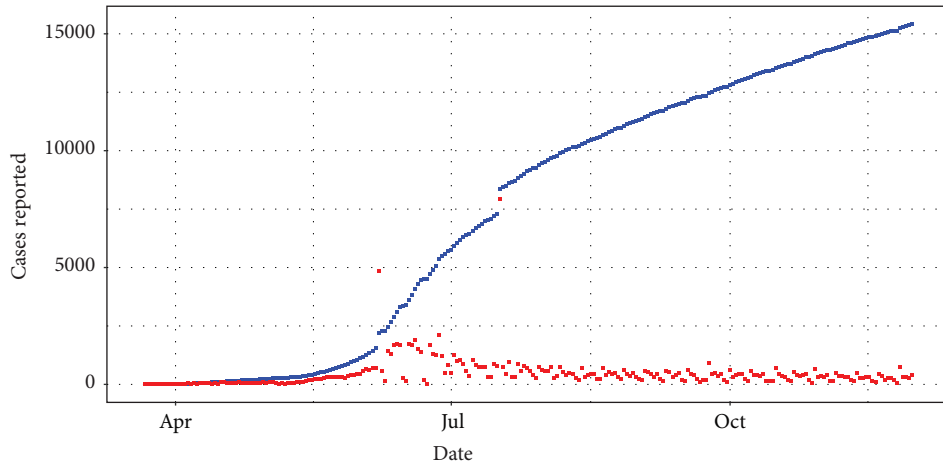
FIGURE 2: Cumulative and daily number of deaths from COVID-19 in Chile from March 22 to November 30, 2020. Cumulative death cases are shown in blue. In red, the daily number of deaths is recorded, amplified in an appropriate amount to be able to appreciate them better.
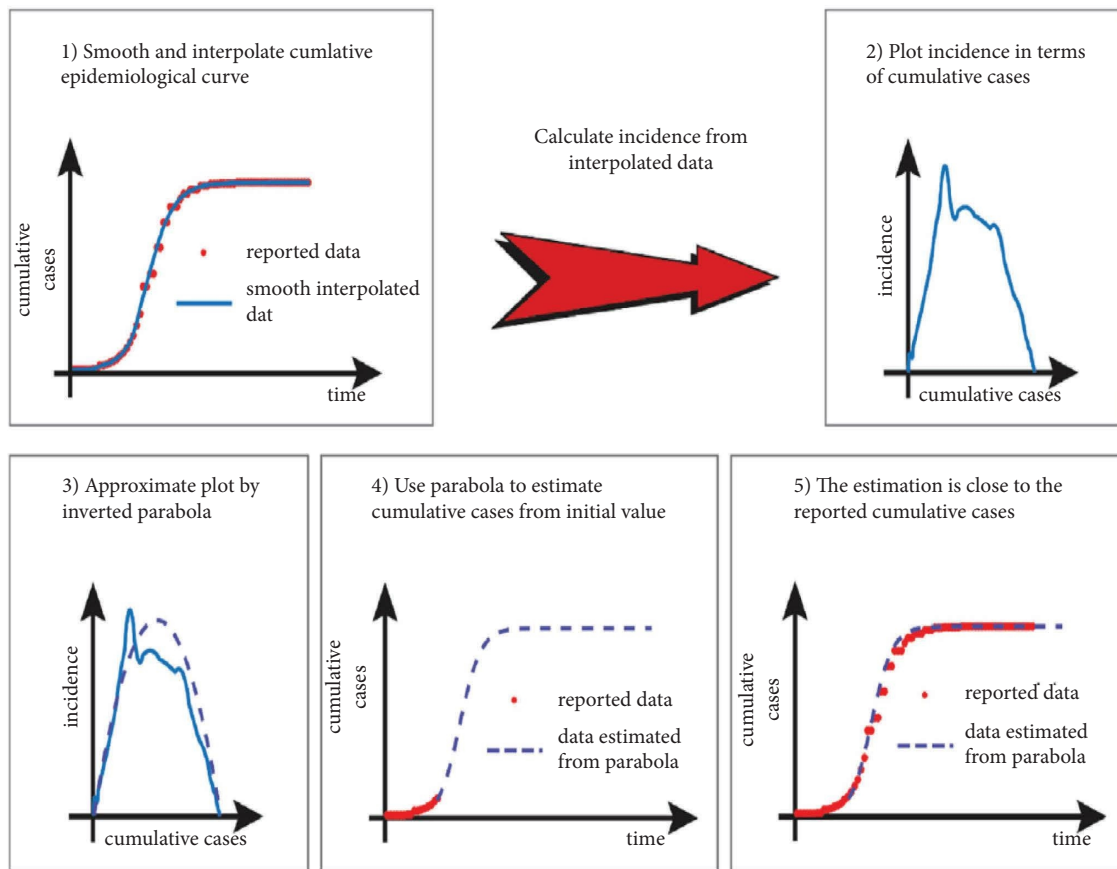

FIGURE 3: Methodology used by Lega and Brown [18].

Since the three-parameter Gompertz curve used by the authors Sánchez-Villegas and Daponte [17] did not fit the data in this study well, other curves were used and are discussed below.

2.4. Sigmoidal Growth Model. The growth curves analyzed are sigmoid curves or "S"-shaped curves, which represent a typical biological growth curve. They symbolize the growth

of organisms in a new and favorable environment. These curves represent a variable that first increases slowly, then speeds up, and finally slows down, eventually growing very little or declining [19]. The three stages of the sigmoid curve are called the exponential phase, the linear phase, and the senescence phase. The sigmoid curves studied in this research are (A) Log-Logistic curve, (B) Gompertz curve, and (C) Weibull curve. These curves are detailed as follows:
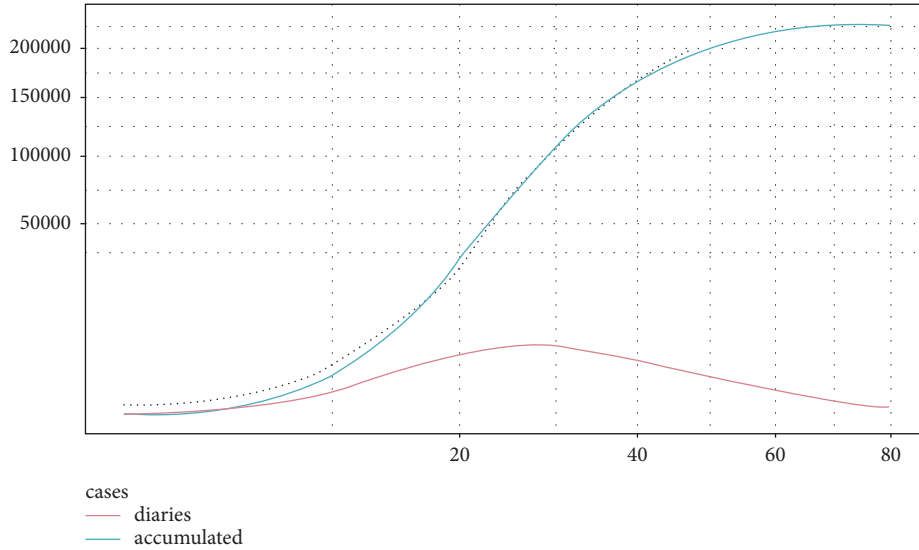
FIGURE 4: Estimated cases in Spain according to the model on April 19 (day 47 of the series of registered data) by Sánchez-Villegas and Daponte [17].

(A) Log-logistic curve that is commonly used to model outbreaks, as they can capture the initial slow growth of the pandemic, followed by a period of rapid growth and a period of slowdown, such as the one shown with the blue line in Figure 5 and denoted by $LL(x)$. The curve with 4 parameters was used since it was the one that best fits the data, which is defined by the authors of [19]

$$LL(x) = b + (c - b)\left(1 + e^{a(\ln(x) - \ln(d))}\right)^{-1}. \quad (6)$$

The red line in Figure 5 is the corresponding change curve, expressed as

$$LL'(x) = \frac{-a(c - b)e^{a(\ln(x) - \ln(d))}}{x\left(1 + e^{a(\ln(x) - \ln(d))}\right)^2}, \quad (7)$$

this change curve is multiplied by an appropriate amount so that it can be displayed on the graph.

Note that, for interpretation purposes, the upper asymptote of $LL(x)$ corresponds to the parameter "c", that is, if $a < 0$, then

$$\lim_{x \to \infty} LL(x) = c. \quad (8)$$

(B) Another curve that was also studied was the Gompertz curve, in blue in Figure 6, denoted by $G(x)$ and defined by the authors of [19]

$$G(x) = b + (c - b)e^{-e^{a(x-d)}}. \quad (9)$$

It is similar to the logistic curve, with the difference that it grows faster at the beginning, which makes it more appropriate to describe biological and epidemiological growth. Furthermore, with 4

parameters, the parameter "c" is its upper asymptote, which is given by

$$\lim_{x \to \infty} G(x) = c, \quad (10)$$

provided that $a < 0$.

The line in red in Figure 6 is the corresponding change curve, represented by

$$G'(x) = -a(c - b)e^{a(x-d)}e^{-e^{a(x-d)}}, \quad (11)$$

magnified by an appropriate amount so that it could be seen on the graph.

(C) The third curve studied was the Weibull curve, in blue in Figure 7, denoted by $W(x)$ and defined by the authors of [19]:

$$W(x) = b + (c - b)e^{-e^{a(\ln(x) - \ln(d))}}. \quad (12)$$

It is commonly used to model survival data used in biomedical applications.

The red line in Figure 7 is the corresponding change curve, also amplified by an appropriate amount, which is given by

$$W'(x) = \frac{-a(c - b)e^{a(\ln(x) - \ln(d))}}{xe^{e^{a(\ln(x) - \ln(d))}}}. \quad (13)$$

The respective asymptote is

$$\lim_{x \to \infty} W(x) = c, \quad (14)$$

provided that $a < 0$.

Once the methodology described above was implemented, certain results began to be obtained, which improved as more information was incorporated. The
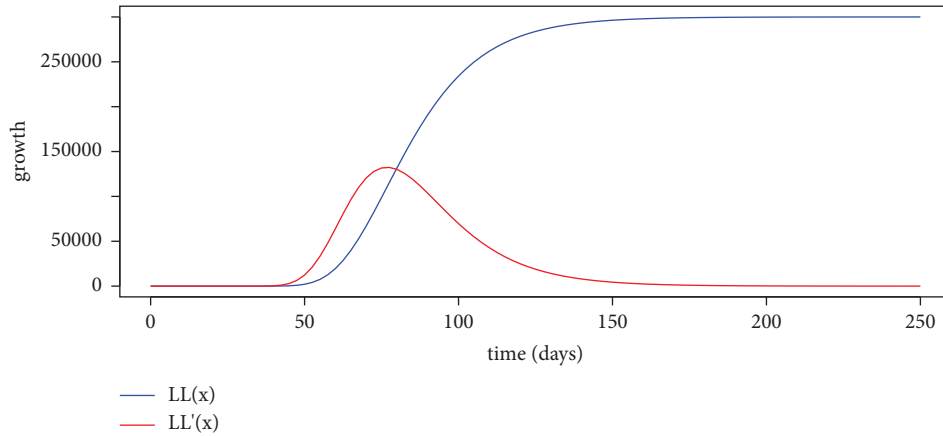
FIGURE 5: Log-logistic curve with final level of 300,000. The red line shows the corresponding change curve multiplied by 10.
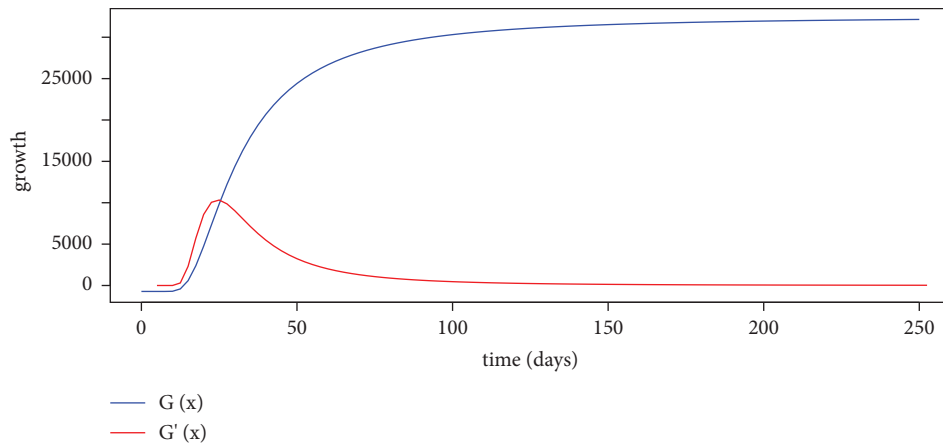

FIGURE 6: Gompertz curve with final level of 300,000. The red line shows the corresponding change curve multiplied by 10.
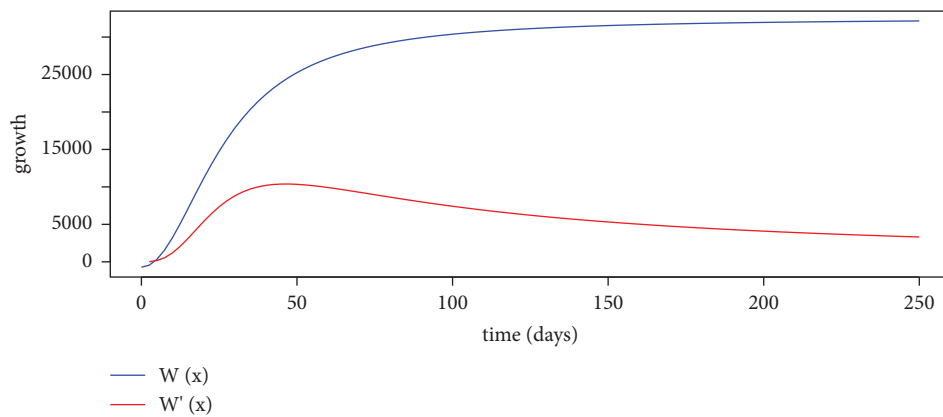

FIGURE 7: Weibull curve with final level of 300,000. The red line shows the corresponding change curve multiplied by 10.

following section presents the curves fitted to the data used and a description of the results obtained.

3. Results

To estimate the parameters of the three growth curves detailed in the subsection sigmoidal growth model, the "drc"

package of the R language was used. Moreover, to quantify the fit error, the mean absolute percentage error (MAPE) was used, which is defined by (see [20])

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{x_i - \widehat{x}_i}{x_i} \right|, \quad (15)$$

where x_i is the observed value, and $\widehat{x}_i$ is the estimated value, which is a measure of the accuracy of the prediction.

Table 1 shows the values of the estimated parameters for each of the growth curves studied when these curves are adjusted to the number of people infected by COVID-19 accumulated in Chile. In addition, their respective MAPEs are presented.

With the estimated values of the parameters (given in Table 1), the graphs of the three growth curves that model the accumulated number of people infected by COVID-19 in Chile were made, and these curves are presented in Figure 8. The points in black correspond to the number of people infected by COVID-19 accumulated in Chile, which were already presented in blue in Figure 1. The fits obtained when using the estimated parameters are given explicitly by the following equations:

(i) Gompertz curve in red is as follows:

$$G_{AI}(x) = -2,464.54 + \frac{873,750.05}{e^{e^{-0.02(x-107.26)}}}. \tag{16}$$

(ii) Weibull curve in blue is as follows:

$$W_{AI}(x) = -0.68 + \frac{1,279,981.20}{e^{e^{-0.00132(\ln(x)-\ln(0.0001))}}}. \tag{17}$$

(iii) Log-logistic curve in yellow is as follows:

$$LL_{AI}(x) = -5,001.27 + \frac{995,260.70}{1 + e^{-3.79(\ln(x)-\ln(125.21))}}, \tag{18}$$

where the subscript AI refers to the accumulated infected and is used to differentiate these curves from those that describe the cases of accumulated deaths (AD) that will be seen as follows.

The great utility of these curves and the objective of this research are that with these functions, it is possible to estimate, for future days, the number of people infected accumulated by COVID-19 in Chile; that is, forecasts can be made of what could happen in the days following October 31, 2020, which was the last day contemplated in the study of the accumulated infected. This day corresponds to the 244th pandemic in Chile, with its beginning on March 2, 2020.

As an illustration, using the Weibull curve, $W_{AI}(x)$, and knowing the data until October 31, 2020, the number of accumulated infected for the following day (November 1, 2020) can be forecast, which corresponds to the day 245 of the pandemic in Chile, which gives an estimate of $W_{AI}(245) = 481,659.1$; that is, approximately 481,659 accumulated infected by COVID-19 in Chile are projected. The real value as of November 1, 2020 was 480,085 accumulated infected, yielding a prediction error of approximately 0.3%, considered acceptable. Forecasts can also be made for the following days, but the accuracy of the estimate decreases.

From Table 1, it can be seen that the Weibull curve yields a lower MAPE; therefore, the following analysis and what is illustrated in Figures 9 and 10 are carried out considering the Weibull curve. Table 1 shows that the estimator of the parameter c (upper limit) for the Weibull curve is

Table 1: Values of the parameters estimated for the three growth models studied when fitting the number of people infected accumulated by COVID-19 in Chile. The last row contains the MAPEs for the various fits.

Parameters	Gompertz	Weibull	Log-logistic
a	−0.02	-1.32×10^{-3}	−3.79
b	−2,464.54	− 0.68	−5,001.27
c	871,285.51	1,279,980.52	990,259.43
d	107.26	1.00×10^{-4}	125.21
MAPE	27.99	**21.53**	35.21

1,279,980.52, which means that, with the information available at the time of completing the study of accumulated infected, they were projected around of 1,279,981 people infected by COVID-19 accumulated in Chile by having the pandemic under control. Value is far from reality, because, at the time of completing this work, the pandemic was not yet under control, and there were already more than 2,900,000 confirmed cases of COVID-19 in Chile.

On the other hand, to get an idea of the evolution of daily cases, the first derivative of the sigmoid curves is used. Since, as already mentioned, the Weibull curve shows a lower MAPE, then the work continues with said curve, which is shown in red in Figure 9, amplified to be able to visualize it on the graph. An expanded view of the evolution of daily cases is presented in Figure 10, where, in blue, the data on the number of daily infected people are shown, and in red, the fit using the first derivative of the Weibull curve, $W_{AI}'(x)$. The utilities of this curve of daily cases are as follows:

(i) The first is that with this curve the number of new cases in future days can be forecast.

(ii) The second is that with this curve, it is possible to have a forecast of the date on which there would be some control over the pandemic which corresponds when $W_{AI}'(x)$ is less than one, that is, when the number of new cases is less than one person, which would be an indication that the pandemic would be coming to an end since there would no longer be someone who could spread.

Using the parameters of Table 1, equation (1), and the information available at the end of the study of accumulated infected, it is obtained that

$$W_{AI}'(x) = \frac{1,689.5752 e^{-0.00132(\ln(x)-\ln(0.0001))}}{x \, e^{e^{-0.00132(\ln(x)-\ln(0.0001))}}}. \tag{19}$$

From this equation, it follows that $W_{AI}'(x) < 1$ for $x \geq 622$; that is, according to the Weibull model, it was projected that on November 13, 2021 (corresponding to day 622), the pandemic in Chile would be under control. But this was not fulfilled, because at the time of completing this investigation, the cases of contagion in Chile were still continuing. The reason why the model was not assertive was due to the fact that what was called the second wave of contagion began in Chile and the sigmoid curves only manage to capture one wave.
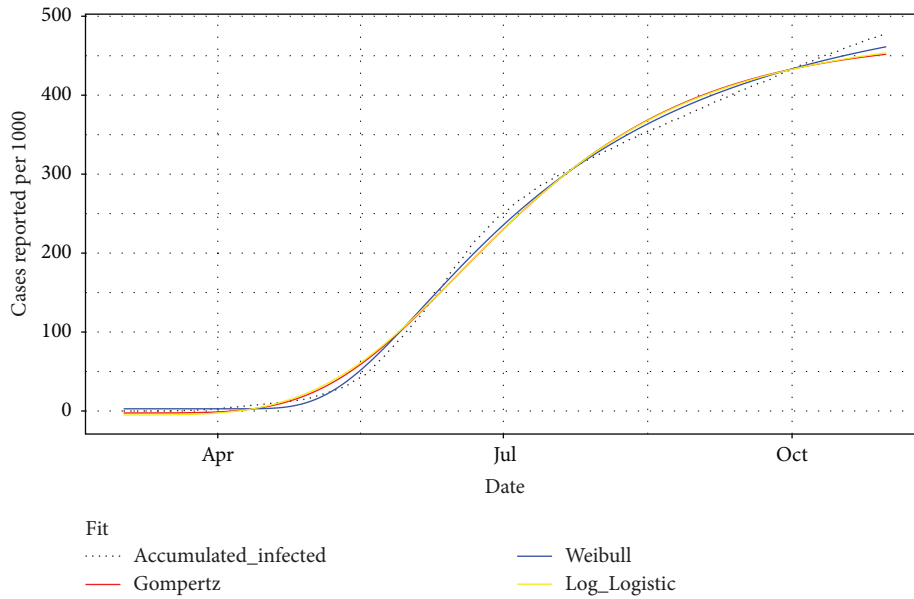
FIGURE 8: Fit of the growth curve models to the accumulated number of infected by COVID-19 in Chile (using data from March 2 to October 31, 2020).
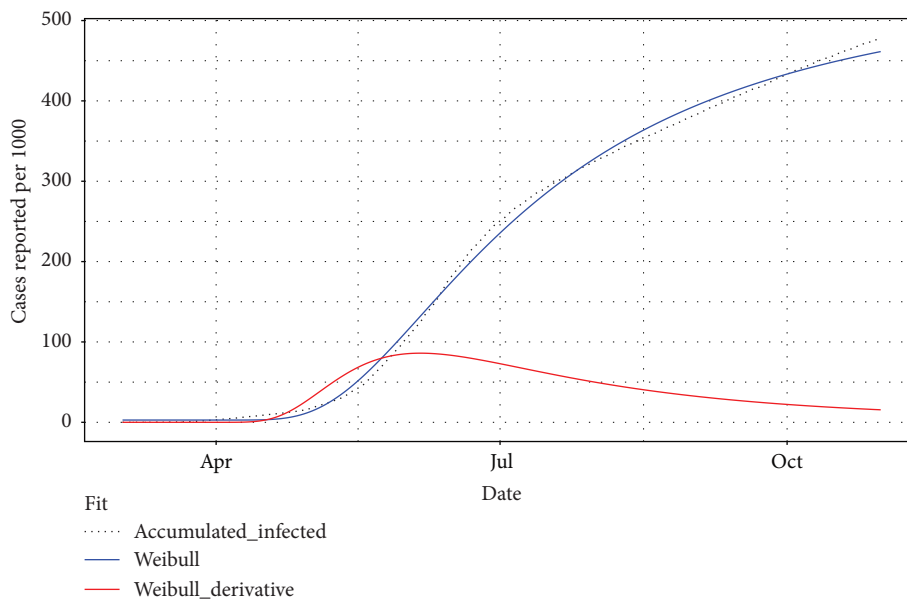

FIGURE 9: Fit of the Weibull model to the number of cases accumulated by COVID-19 in Chile (in blue) and fit of the evolution of daily cases (in red), using data from March 2 to October 31, 2020.

Next, the study carried out for the cases of deceased is presented. The methodology is the same that was applied to estimate the number of accumulated infected. First, the "drc" package of the R language was used to estimate the parameters of the three growth models already described. Table 2 contains the estimated parameters for the different growth curves studied when these curves are adjusted to the cumulative number of deaths from COVID-19 in Chile. The respective MAPEs are also presented.

Figure 11 presents the graphs of the respective growth curves that are obtained by using the estimated values of the parameters given in Table 2. Visually, it seems that the curves

do not fit the data in black as well. This is apparently because the vertical scale was not altered as in the case of the accumulated infected. Consequently, it is being seen more closely. In order to have an objective measurement, Table 2 shows that the three curves effectively reduced their percentage error, and as before, the lowest MAPE was obtained by using the Weibull curve, so the study continues using this curve, whose equation is given by

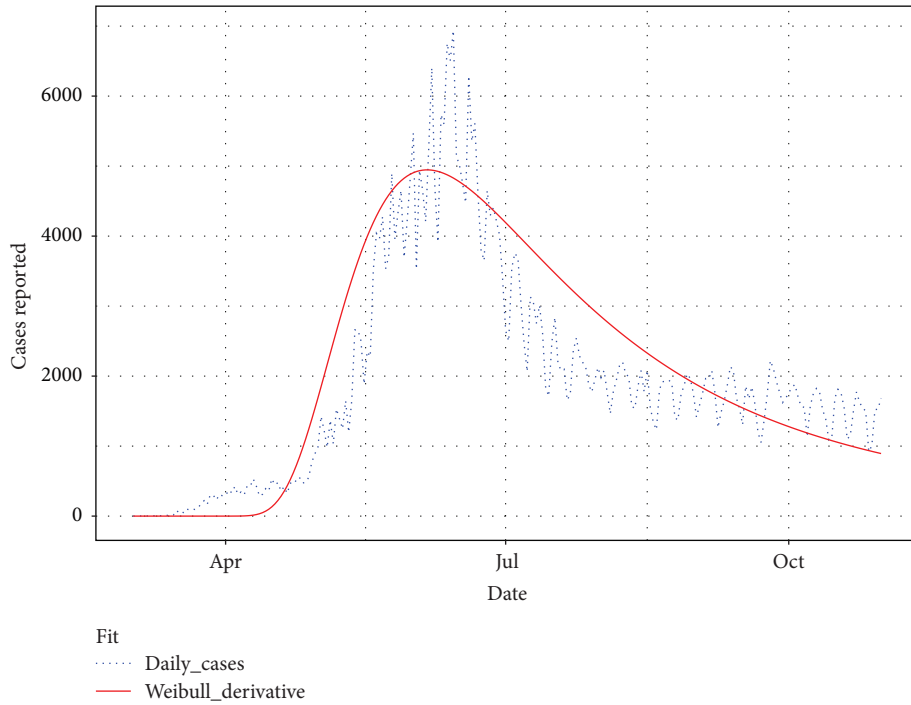$$W_{AD}(x) = -0.0659 + \frac{26,432.3459}{e^{e^{-0.085(\ln(x)-\ln(0.172))}}}. \tag{20}$$

FIGURE 10: Fit of the first derivative of the Weibull curve (in red) to the number of daily infections (in blue) by COVID-19 in Chile (using data from March 2 to October 31, 2020).

TABLE 2: Values of the estimated parameters for the three growth models studied when fitting the number of accumulated deaths by COVID-19 in Chile. The last row contains the MAPEs for the various fits.

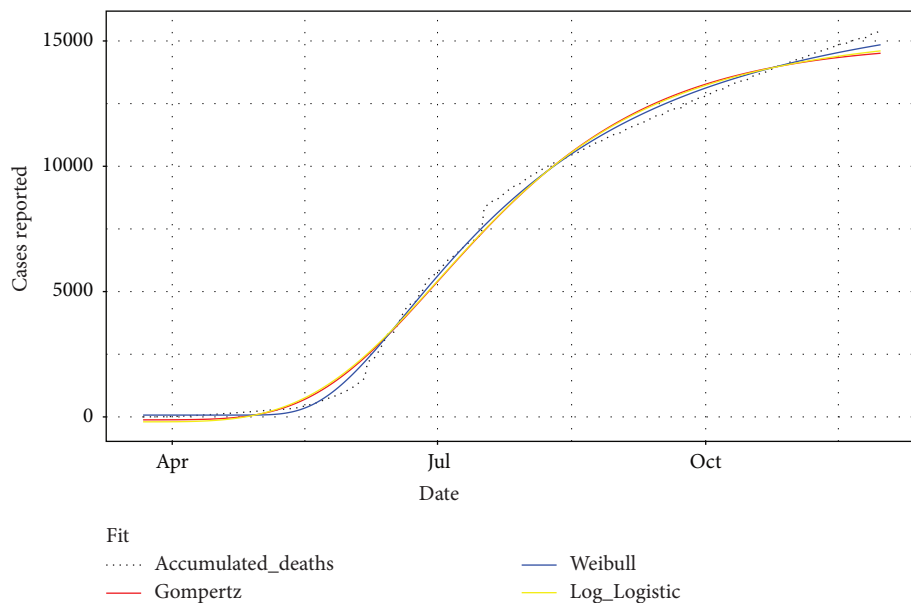| Parameters	Gompertz	Weibull	Log-logistic
a	-0.02	-8.50×10^{-2}	-3.64
b	-118.24	-6.59×10^{-2}	-196.26
c	$21,937.64$	$26,432.28$	$24,574.49$
d	102.17	0.17	120.02
MAPE	21.95	**10.15**	35.10

FIGURE 11: Fit of the growth curve models to the number of deaths accumulated by COVID-19 in Chile (using data from March 22 to November 30, 2020).
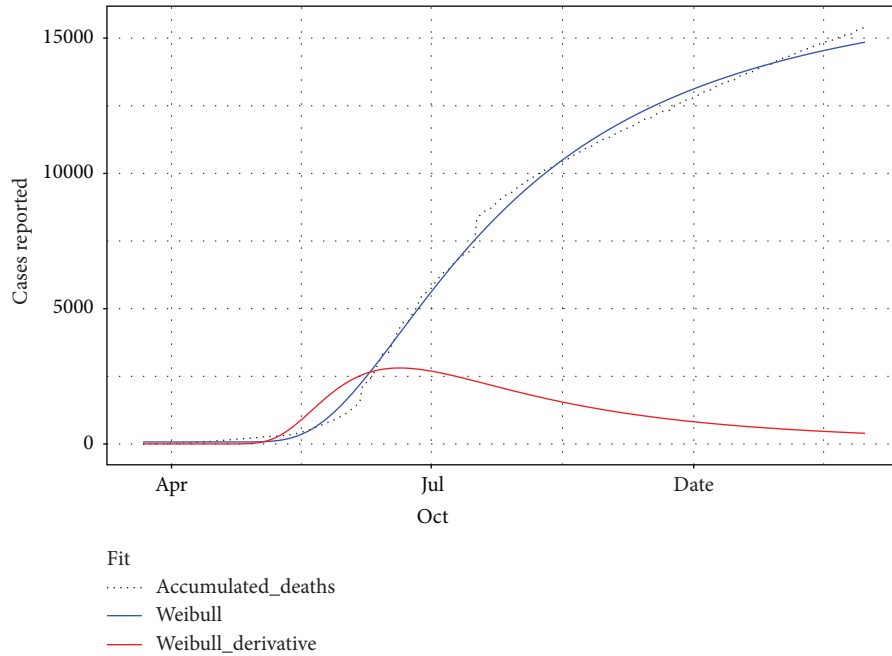
FIGURE 12: Fit of the Weibull model to the number of deaths accumulated by COVID-19 in Chile (in blue) and fit of the evolution of the number of daily deaths (in red), using data from March 22 to November 30 of 2020.
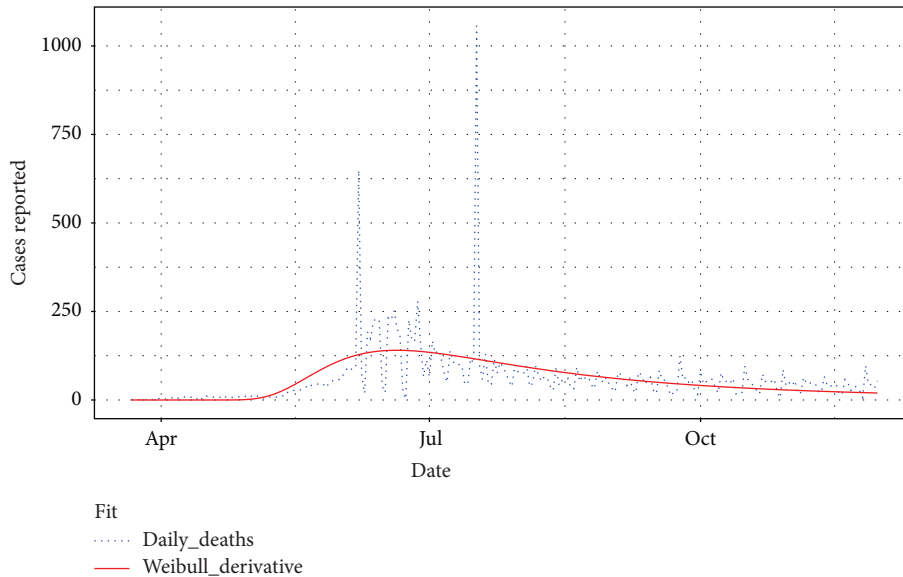

FIGURE 13: Fit of the first derivative of the Weibull curve (in red) to the number of daily deaths (in blue) from COVID-19 in Chile (using data from March 22 to November 30, 2020).

With this function, forecasts can be made of what could happen in the days following November 30, 2020, which was the last day considered in the study of accumulated deaths. This day corresponds to 254 considering March 22, 2020, the first day on which deaths were recorded. Taking this into account, the number of accumulated deaths can be forecast for the following day (December 1, 2020), which corresponds to day 255, which gives an estimate of $W_{AD}(255) =$ 15,440.26; that is, approximately 15,440 deaths accumulated by COVID-19 in Chile are projected. The actual value

as of December 1, 2020, was 15,430 cumulative deaths, very close to the predicted value, yielding a prediction error of approximately 0.06%.

On the other hand, from Table 2, it is obtained that the estimator of the parameter c (upper asymptote of the curve) for the Weibull curve is 26,432.28, which means that, with the information available at the time of completing the study of deceased accumulated, around 26,432 people were projected to die from COVID-19 in Chile when the pandemic was under control. Value is far from reality, because, at the
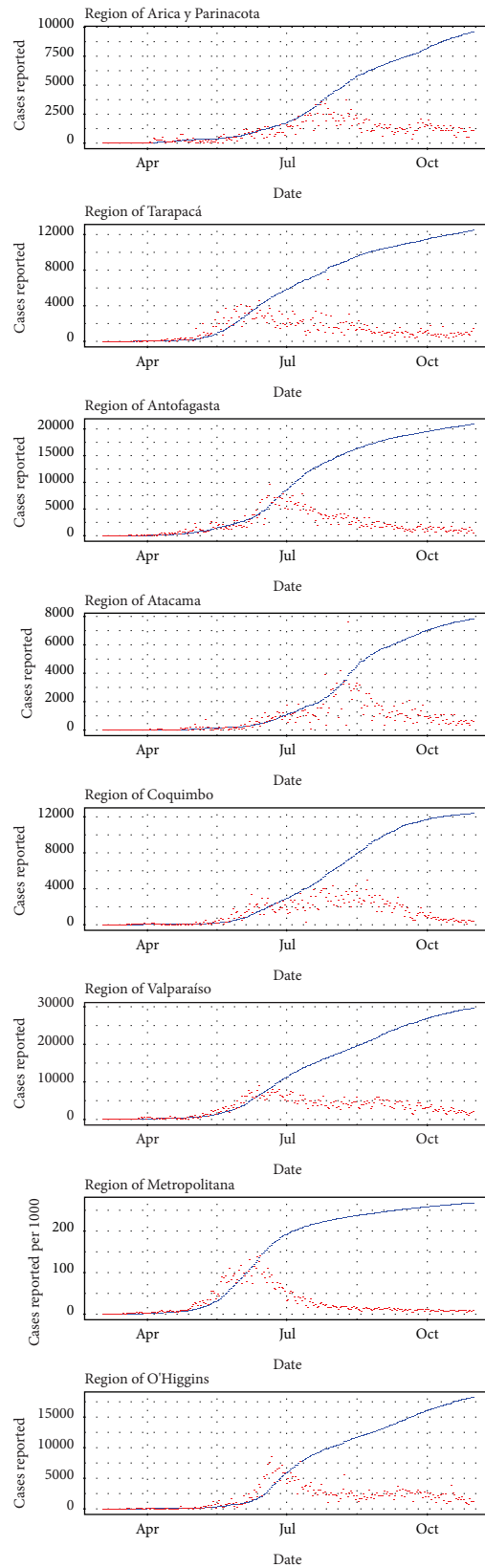
FIGURE 14: Cumulative (in blue) and daily (in red) number of infected by COVID-19 in the most affected regions of Chile. The data record covers from March 22 to November 30, 2020. The daily cases are magnified by 25 to better appreciate them.

time of completing this work, the pandemic was not yet under control, and there were already more than 42,000 cases of COVID-19 deaths in Chile.

Analogous to the study on daily infections, to get an idea of the evolution of daily deaths, the first derivative of the Weibull curve is used (because it has a lower MAPE), which is presented in red in Figure 12 and amplified by an appropriate amount to display it on the graph. Another view is presented in Figure 13, where, in blue, the number of daily deaths is shown, and in red, the fit using the first derivative of the Weibull curve, $W_{AD}'(x)$, whose applications are as follows:

(i) That with this curve, it is possible to forecast the number of people who died on a certain day in the future

(ii) That it is possible to have a forecast of the date on which there would be some control over the deaths from the pandemic, which would happen when $W_{AD}'(x) < 1$, that is, when the number of deaths is less than one.

Using the parameters of Table 2, equation (1), and the information available at the end of the study of the accumulated deaths, it is obtained that

$$W_{AD}'(x) = \frac{2246.7494 e^{-0.085(\ln(x) - \ln(0.172))}}{x\, e^{e^{-0.085(\ln(x) - \ln(0.172))}}}. \qquad (21)$$

From this equation, it is obtained that $W_{AD}'(x) < 1$ for $x \geq 678$; that is, according to the Weibull model, it was projected that on January 28, 2022 (corresponding to day 678), the cases of deaths in Chile would end. But this was not fulfilled, because at the time of completing this investigation, deaths from COVID-19 continued to occur in Chile.

Regarding the study of the most affected regions, it can be mentioned that the situation is similar to that which occurred in the entire country. Figure 14 only presents, for some regions of Chile, the graphs of the accumulated and daily reported cases from March 2 to October 31, 2020. Therefore, the same methodology described in subsection proposed methodology was applied, obtaining results similar to those detailed in the national situation. In some regions, a 4-parameter Gompertz curve was a better fit.

As for the regional analysis of the number of accumulated and daily deaths, it is necessary to highlight that the situation is similar to the country case, but, for space reasons, the results are not shown.

4. Conclusions

In this research, a methodology was designed and implemented to model epidemics with sigmoidal growth curves. Three different models were used and compared, instead of a single model as cited in the literature. One strength of the models used is that they only use the data on the number of infected and accumulated deaths, which are then used to predict the number of infected and deceased daily in future days, without incorporating external variables. This, on the

one hand, could be considered a weakness by not adding more information to the model, such as periods of confinement or vaccination, for example. On the other hand, the absence of more information means that the quality of the predictions depends a lot on the quality of the data, a fundamental aspect, since, at the beginning of the pandemic, it was not certain whether the patients were due to COVID-19. The growth models studied were adjusted to the evolution of the pandemic recorded in what was called the first wave, which is undoubtedly a valuable aid for decision-making by government agencies responsible for health policies. However, these models do not contemplate an intervention analysis such as vaccinations, which, as is known, are controlling the pandemic.

Data Availability

The data used to support the findings of this study are available at the link https://www.minciencia.gob.cl/covid19/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. Barrado, "We Live at point of Inflection: The 2020 Generation and the New Society," 2020, https://telos.fundaciontelefonica.com/punto-de-inflexion-la-generacion-2020-y-la-nueva-sociedad/.

[2] P. Norrie, *A History of Disease in Ancient Times: More Lethal than War*, Springer, London, UK, 2016.

[3] B. P. Zietz and H. Dunkelberg, "The history of the plague and the research on the causative agent yersinia pestis," *International Journal of Hygiene and Environmental Health*, vol. 207, no. 2, pp. 165–178, 2004.

[4] F. Guerra, "Origin of epidemics during the conquest of America," *Quinto Centenario*, vol. 14, pp. 43–51, 1988.

[5] S. López-Moreno, F. Garrido-Latorre, and M. Hernández-Avila, "Desarrollo histórico de la epidemiología: su formación como disciplina científica," *Salud Pública de México*, vol. 42, no. 2, pp. 133–143, 2000.

[6] B. González, L. Tomaino, L. Serra, and P. Barber, S. Rodríguez-Mireles, COVID-19: Pandemic of Mathematical Models," 2020, https://theconversation.com/covid-19-pandemia-de-modelos-matematicos-136212.

[7] E. K. Grillo-Ardila, L. E. Bravo, R. Guerrero, and J. Santaella-Tenorio, "Mathematical model and COVID-19," *Colombia Médica*, vol. 51, no. 2, pp. e4277–e4279, 2020.

[8] A. Jajarmi, A. Yusuf, D. Baleanu, and M. Inc, "A new fractional HRSV model and its optimal control: a non-singular operator approach," *Physica A: Statistical Mechanics and Its Applications*, Article ID 123860, 2020.

[9] A. Ajbar, R. T. Alqahtani, and M. Boumaza, "Dynamics of an SIR-based COVID-19 model with linear incidence rate,

nonlinear removal rate, and public awareness," *Frontiers in Physics*, vol. 9, 2021.

[10] R. u. Din and E. A. Algehyne, "Mathematical analysis of COVID-19 by using SIR model with convex incidence rate," *Results in Physics*, vol. 23, Article ID 103970, 2021.

[11] T. Dehesh, H. A. Mardani-Fard, and P. Dehesh, "Forecasting of COVID-19 confirmed cases in different countries with ARIMA models," *medRxiv*, vol. 14, no. 9, pp. 971–976, 2020.

[12] S. Singh, K. Singh, J. Kumar, and S. Jitendra, "Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19. Chaos," *Solitons & Fractals*, vol. 135, Article ID 109866, 2020.

[13] T. Chakraborty and I. Ghosh, "Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis," *Chaos, Solitons & Fractals*, vol. 135, Article ID 109850, 2020.

[14] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, and L. d. S. Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: perspectives for Brazil," *Chaos, Solitons & Fractals*, vol. 135, Article ID 109853, 2020.

[15] S. Yadav and P. K. Yadav, "The peak of COVID-19 in india," *medRxiv Preprint*, 2020.

[16] A. C. Harvey and P. Kattuman, "Time series models based on growth curves with applications to forecasting coronavirus," *Harvard Data Science Review Special Issue 1-COVID-19: Unprecedented Challenges and Chances*, 2020.

[17] P. Sánchez-Villegas and A. Daponte, "Predictive models of the COVID-19 epidemic in Spain with Gompertz curves," *Gaceta Sanitaria*, 2020.

[18] J. Lega and H. E. Brown, "Data-driven outbreak forecasting with a simple nonlinear growth model," *Epidemics*, vol. 17, pp. 19–26, 2016.

[19] G. A. F. Seber and C. J. Wild, *Nonlinear Regression*, Wiley & Sons, New York, 1989.

[20] J. D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey, USA, 1994.

[21] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021.