

Neural Networks and Sigmoidal Models for the Waves of the Coronavirus

Oliver Amadeo Vilca Huayta

Departamento de Ingeniera de Sistemas, Universidad Nacional del Altiplano, Peru

ovilca@unap.edu.pe

Keywords: Brazil, COVID-19, death number, Italy, logistic sigmoidal model, pandemic, Peru, Switzerland, artificial neural networks

Received: November 1, 2024

The rapid spread of the coronavirus disease 2019 (COVID-19) pandemic has caused enormous problems—mainly a large number of deaths—and, so, it is important to construct epidemiological models for forecasting and prevention. The main objective of this study is to develop a novel and efficient model $F(x)$, treating the cumulative number of deaths due to COVID-19 using two competing approaches: the logistic sigmoidal function and an artificial neural network. We also aim to estimate models for the death rate. This research is retrospective and longitudinal. Data were downloaded from Johns Hopkins University for four countries: Italy, Mexico, Brazil, Switzerland, and Peru. It is shown that the basic logistic function is optimal—that is, it has a running time that cannot be improved—and a new variant of the function is obtained and used. The novel contribution of this work is the general model that can be used on data that are not necessarily epidemiological; and with any function, where, as x approaches positive or negative infinity, the function tends to a constant value. Also, the method of its construction, and the calculation of high-order derivatives, which allow for the development of an elegant and useful model, as well as justification of the term “Wave(s)”. The inflection points (i.e., the place where the concavity changes) are also obtained and validated. Finally, they were applied to real-world datasets. It is concluded that the obtained models have a high fit, with acceptable root mean squared error (RMSE) and Pearson correlation coefficients greater than 0.999. The models are representative, predictive, and optimal and, so, useful to achieve a better understanding of the pandemic and improve future public health response.

Povzetek:

1 Introduction

The coronavirus disease 2019 (COVID-19) is highly contagious disease [1, 2] that has caused a high number of deaths and continues to damage the economy and society, to name just a few of its impacts. On October 4, 2024, the Heads of the International Monetary Fund (IMF), the World Bank Group (WBG), and the World Health Organization (WHO) have agreed on broad principles for cooperation on pandemic preparedness (<https://www.imf.org>; accessed on 7 October 2024). The development and study of epidemiological models can be essential to prevent, predict, or mitigate future epidemics or pandemics [3, 4, 5]. They are also necessary for decision making; for example, allowing public health institutions to take reliable measures in order to preserve people’s lives [6, 7]. In addition, comparison between countries is considered essential for the control of COVID-19 [8].

In this study, four countries with different characteristics were selected for the purposes of modeling, comparison, and research, namely, Italy, Mexico, Brazil,

Switzerland, and Peru, for which COVID-19 has resulted in a total of 188,322, 699,276, 333,188, 14,210, and 219,539 deaths, respectively, according to the accumulated death time-series of Johns Hopkins University (<https://coronavirus.jhu.edu>; accessed on 1 January 2024), without considering the deaths registered after the end of the pandemic. The number of deaths and mortality rates vary between countries [9] and, in this work, we demonstrate that they follow a strict sigmoidal behavior typical of epidemics or pandemics.

The Gompertz model, due to the definition of its formula $f(x) = a \cdot e^{-e^{b-c \cdot x}}$ (i.e., double exponential), requires more elementary operations than the most basic of the sigmoidal functions. Therefore, in this research, we demonstrate that the logistic function is the most efficient and optimal for numerical algorithms (e.g., machine learning and non-linear regression). Therefore, given the structure and functionality of the logistic function, we recommend its use in future work.

Additionally, we adapted and obtained a new variant of the logistic model, which we call logistic 2 or just the logistic model $g(x) = \frac{H}{1+2^{A-Bx}}$, which does

not lose any of the original model's properties or qualities. Furthermore, both functions are equivalent.

One of the major contributions of this work is the general model and the method of its construction. Each function $g(x)_i$ corresponds to wave i . It has the following characteristics or advantages: It does not require any form of integration or union of functions, does not need preliminary work, and it does not require to add a dummy variable to the data (or to the function) as in previous works [10, 11]. Mainly for this reason and the previous ones it is more efficient:

$$F(x) = \frac{H_1}{1 + 2^{A_1 - B_1 x}} + \frac{H_2}{1 + 2^{A_2 - B_2 x}} + \cdots + \frac{H_i}{1 + 2^{A_i - B_i x}} + \cdots + \frac{H_n}{1 + 2^{A_n - B_n x}}$$

where $F(x)$ is the expected cumulative number of deaths, x is the number of days since the first case, n is the number of waves, $1 \leq i \leq n$, H_i is the height of the wave i , A_i and B_i are constants related to the pandemic.

The main objective of this study is to develop a novel and efficient model $F(x)$, treating the cumulative number of deaths due to COVID-19 since the pandemic began as a variable, using the logistic function. Furthermore, we compare the proposed model with an artificial neural network (ANN) which, in this case, constitutes an ideal option.

The first derivative is calculated to explain the death rate and the magnitude of the COVID-19 waves, where even its graph represents the considered characteristics. In this way, justification is provided for the denomination of COVID-19 “waves.” Next, the second derivative is used to calculate the vertical inflection points (i.e., the point where the sigmoidal function changes from concave upward to concave downward). At this point, only one last issue remains to be resolved: the second derivative can report false inflection points. Therefore, the inflection points are validated through the use of the third derivative. The use of high-order derivatives—specifically, the first three derivatives—allows for an elegant and irrefutable mathematical analysis of any model in general. Then, the proposed model is applied to four representative countries, yielding novel results.

In ANN and the proposed model, the differences were minimal. The disadvantage of the ANN is a considerably longer calculation time, when compared to sigmoidal functions.

This paper is structured as follows: Section 2 describes related works, Section 3 presents the materials and methods, and describes the used data. Section 4 demonstrates the results, including those related to the ANN model, the coupled model, high-order derivatives, and the case study. Finally, the conclusions and directions for future work are presented in Section 5.

2 Related works

In relation to key publications: While one study has focused on the Peruvian case, it was limited to the calculation of a function with the Gompertz model and did not apply ANNs, and was carried out with incomplete data as it was conducted during the pandemic period [10]. There is also a study in which $F(x)$ was calculated using the Boltzmann sigmoidal function for a case study (Peru), which was correlated with the social isolation measures in Peru (qualitative variable) [11], Table 1 illustrates the main differences.

Regarding ANN, several works have used artificial neural networks for prediction. In India, research has been carried out on the Multilayer perceptron [12]. In Brazil, a long short-term memory (LSTM) model has been used [13], which obtained a correlation coefficient of 0.96. In the United States and India, various artificial neural networks have been used with data on confirmed infections and deaths from COVID-19, of which the convolutional LSTM stood out, which had greater precision and a lower error [14]. Additionally, in China and other countries, it has been concluded that ANNs are adequate to predict global infections and deaths due to COVID-19 [15].

Regarding the use of the logistic sigmoidal function to analyze pandemics, there have not been many related studies. One study used the extended logistic function in this context [16]; however, the authors did not compare it with an ANN, and did not calculate derivatives or inflection points.

3 Materials and Methods

It should be noted that this study consider other model, another data, and use high-order derivatives of F . And new simple way to couple functions $g(x)$.

3.1 Data Set

The database of the total number of deaths used in the current study is publicly available (open access) from the Coronavirus Disease Data Repository of Johns Hopkins University (<https://github.com/CSSEGISandData/COVID-19>; accessed on 1 January 2024). Complete data for different countries are also available [17].

In order to achieve good results, it is necessary to work with complete data from the beginning to the end of the pandemic. Studying partial data may not be sufficient. In addition, modeling a single wave [18] can be relatively easy. In this work, no function integration method is required [10], and a program is used to process and estimate F .

In this work	In Key publication
It is not necessary to estimate any horizontal inflection point as a preliminary procedure to obtain $F(x)$.	It is necessary to calculate the horizontal inflection points.
The logistic sigmoidal function was used. It does not require the dummy variable q .	The Boltzmann sigmoidal function was used. It requires the insertion of a dummy variable q in the model and in the database.
A general model F was obtained, whose elements are extended to functions that are not necessarily sigmoidal, that is, with $\lim_{x \rightarrow -\infty} h(x) = L$ and $\lim_{x \rightarrow \infty} h(x) = R$, where L and R are constants. Finally, it was demonstrated with mathematical induction. It was applied to four case studies, where the Peruvian model is: $F(x) = \frac{85.94}{1+2^{6.07-0.048*x}} + \frac{114.1}{1+2^{15.96-0.041*x}} + \frac{13.22}{1+2^{44.35-0.063*x}} + \frac{3.4}{1+2^{73.16-0.082*x}} + \frac{3.33}{1+2^{46.8-0.045*x}}$	The Peruvian model is: $F(x, q) = \frac{(2-q)(3-q)(4-q)(5-q)}{24} \left\{ \frac{88.05}{1+e^{\frac{(128.95-x)}{30.88}}} \right\} - \frac{(1-q)(3-q)(4-q)(5-q)}{6} \left\{ \frac{113.13}{1+e^{\frac{(390.65-x)}{34.36}}} + 86.49 \right\} + \frac{(1-q)(2-q)(4-q)(5-q)}{4} \left\{ \frac{14.18}{1+e^{\frac{(698.06-x)}{26.51}}} + 199.34 \right\} - \frac{(1-q)(2-q)(3-q)(5-q)}{6} \left\{ \frac{3.61}{1+e^{\frac{(898.54-x)}{17.99}}} + 213.37 \right\} + \frac{(1-q)(2-q)(3-q)(4-q)}{24} \left\{ \frac{2.84}{1+e^{\frac{(1039.82-x)}{25.42}}} + 216.95 \right\}$
$\sigma = 0.999905$ (the largest). The vertical inflection points and higher order derivatives were calculated.	$\sigma = 0.999$. The vertical inflection points and higher order derivatives were not calculated.

Table 1: Main differences in relation to the key publication

3.2 The sigmoidal logistic function

The sigmoidal logistic model is a function that can be applied to a time-series [19], which has different applications such as the representation of epidemic curves. It is represented by the following function:

$$g(x) = \frac{H}{1 + e^{A-Bx}}, \quad (1)$$

where x is the number of days since the first case, $g(x)$ is the expected cumulative number of deaths, H is the height of the wave, A and B are constants related to the pandemic, and e is a mathematical constant (i.e., an irrational and transcendental number approximately equal to 2.718281828459045).

In this work, a new version of the logistic function is used, which we call the logistic2 function or, simply, the logistic function:

$$g(x) = \frac{H}{1 + 2^{A-Bx}}, \quad (2)$$

where A and B are new pandemic-related constants. Instead of the irrational number e (which cannot be expressed as a fraction m/n , where $m, n \in \mathbb{Z}$), we use the natural number 2 as, in computing, we can efficiently calculate 2^k if $k \in \mathbb{N}$. Even a school student could understand it.

The calculation of high-order derivatives was performed by hand and was checked using the Octave Software (version 9.1.0). By doing so, we achieved better factorization and presentation of the functions.

3.3 The program

The program is developed in R (version 4.4.1 - free software environment for statistical computing and graphics), and the integrated development environment (IDE) RStudio Desktop (2024.09.0+375 "Cranberry Hibiscus" release). It is efficient, simple and easy to understand. The algorithm has the following procedures:

1. Load libraries (data structures, nonlinear regression, metrics, graphs, neural networks).
2. Select country and language.
3. Load data.
4. Estimate the model (providing starting values).
5. Show goodness-of-fit reports (σ and RMSE).
6. Export the model $F(x)$, the first derivative $F'(x)$, and the second derivative $F''(x)$.
7. Generate the graph (configuring color, size, resolution, output file, among others).

3.4 Statistical analyses

The Pearson product-moment correlation coefficient or Pearson correlation coefficient (σ for short) is a measure of linear dependence and its direction between two quantitative variables. If there is a positive correlation, when one variable changes, the other variable changes in the same direction; if there is no correlation (zero), there is no relationship between the variables;

and, if there is a negative correlation, the opposite of the positive correlation occurs.

The root mean square error (RMSE) is defined as the measure of the differences between the values that are predicted by a model and those that are actually observed [20]. RMSE is not the only measure for this purpose [21].

All models in this study have two variables—namely, the predicted values and the observed data—therefore, any of the goodness-of-fit measures can be used.

Nonlinear regression was used. For statistical analysis and graph construction, R software was used.

4 Results

4.1 Case studies

Four countries were considered: Italy (Southern and Western Europe), Brazil (South America), the Swiss Confederation (west-central Europe), and Peru (South America).

Table 2 presents the populations of the study countries (year 2020), the number of deaths, the start and end dates of the pandemic (number of days), and the percentage of deaths in relation to the population. The population data were downloaded from Population Pyramids of the World (<https://www.populationpyramid.net>; accessed on 1 September 2024). The last column is the percentage of deaths in relation to the population = $\frac{\text{Deaths}}{\text{Population}} \cdot 100$.

Day one corresponds to the first death. In Italy, the pandemic began in February 2020, while it began in March in Brazil, Switzerland, and Peru. The pandemic lasted until 9 March 2023. Specifically, it lasted 1113 days in Italy, 1088 days in Brazil, and 1099 days in Switzerland and Peru.

From Table 2, it can be seen that the country that had a considerably higher percentage of deaths, according to the total number of inhabitants, was Peru (0.66%), followed by Italy (0.32%), and ending in a tie between Brazil and Switzerland (0.16%).

These countries differ, for example, in terms of their location, population, and number of deaths from COVID-19, among other aspects. However, in relation to the start and end dates of the pandemic, they do not differ greatly.

4.2 Artificial neural networks

In machine learning, a neural network (NN), also called an artificial neural network (ANN), is a model composed of connected units or nodes called neurons. The main feature is their ability to automatically learn and extract characteristics from input data [25].

Long short-term memory (LSTM) is a type of recurrent neural network (RNN) architecture designed

to address the limitations of traditional RNNs, especially when it comes to learning long-term dependencies. The main innovation of LSTM is the use of gating mechanisms to control the flow of information through the network. This allows them to maintain and update their internal state over long periods [26, 27, 28].

For better results, it is convenient to select the most appropriate architecture and configure its parameters.

R provides libraries and frameworks that simplify the process of building and training neural networks, including LSTM networks. The Keras library provides a high-level interface for building and training neural networks. To the best of the authors' knowledge, in most previous works, the Python programming language was used.

To avoid biases in time measurement, we used the computer: HP 11th Generation Intel(R) Core™ i7, and development language: R programming language for statistical computing and graphics.

The data were scaled and split: 90% of the data was used for training and the rest was retained for testing. The LSTM was used with the following characteristics: three layers (1 input neuron, 200 neurons in the hidden layer, and 1 output neuron), the “adam” optimizer, loss = “mse”, epochs = 200, and batch_size = 1.

On one hand, the RMSE of the ANN was quite acceptable, being less than 0.72 (on scaled data, see Table 3). On the other hand, the Pearson correlation coefficients (σ) were also quite good: those for Italy, Brazil, Switzerland, and Peru were 0.9999236, 0.9999440, 0.9999472, and 0.9998413, respectively (usually varying with each execution). In ANNs, training took a considerable running time (i.e., more than 8 minutes), and their performance can be improved by expanding their parameters, at the cost of additional time.

For the ANN, the σ did not differ much from each other; this is because ANNs adapt well to data. However, sigmoidal logistic models require the data to follow a pattern.

In general, for some countries the data fit better; however, all countries were fit fairly well.

Finally, with the parameter settings of an ANN, there is not much that can be done.

4.3 Calculating the number of operations of sigmoidal functions

The execution time of a sigmoidal function depends on the type and number of operations. In general, the order of the execution time of the operations, from smallest to largest, is as follows: (1) addition or subtraction, (2) multiplication or division, and (3) exponentiation with a real number in the exponent; that is, addition/subtraction is calculated faster than the other two, and multiplication/division is generally faster than exponentiation.

Country	Population	Number of deaths	Start and end date of the pandemic (number of days)	Percentage of deaths
Italy	59,500,579	188,322	From 21 February 2020 to 9 March 2023 (1113 days)	0.32 %
Brazil	213,196,304	333,188	From 17 March 2020 to 9 March 2023 (1088 days)	0.16 %
Switzerland	8,638,613	14,210	From 06 March 2020 to 9 March 2023 (1099 days)	0.16 %
Peru	33,304,756	219,539	From 06 March 2020 to 9 March 2023 (1099 days)	0.66 %

Table 2: Population of the study countries (year 2020), the number of deaths, start and end dates of the pandemic (number of days), and the percentage of deaths in relation to the population

Country	Number of days	Correlation coefficient R	Determination coefficient R^2	Root mean square error
Italy	1113	0.9999236	0.9998472	0.713694600
Brazil	1088	0.9999440	0.9998881	0.008530093
Switzerland	1099	0.9999472	0.9998945	0.003714717
Peru	1099	0.9998413	0.9996826	0.009235257

Table 3: Measures of the ANN

Then, for evaluation, the functions were grouped into types: those based on the Gompertz function (representing the double exponential functions), based on the hyperbolic tangent function (representing the extended functions of the logistic function), and based on the logistic function. The number of operations in each group was then calculated, as detailed in Table 4.

The Gompertz function was found to be the least suitable, as it has two exponential operations, two multiplication operations, and one subtraction operation. Meanwhile, the hyperbolic tangent function has an additional subtraction operation in relation to the logistic function. The Boltzmann function has two divisions (generally, in computing, the division operation costs more than multiplication). Finally, the logistic function was selected as it has the fewest operations. It could barely be equaled by some extended sigmoidal functions and, therefore, it was considered to be optimal in terms of running time.

4.4 The model

First of all, it was necessary to select an optimal and efficient sigmoidal function. Subsequently, a general model was developed.

The fundamental element of the model is the logistic sigmoidal function, equation (2). Specifically, the model $F(x)$ is a series of ordered elements, that is, a summation of logistic functions, starting with the function $g(x)_1$ for the first wave, $g(x)_2$ for the second wave, and so on up to $g(x)_n$ for the last wave. The basic structure is:

$$F(x) = g(x)_1 + g(x)_2 + \dots + g(x)_i + \dots + g(x)_n \quad (3)$$

It has the following characteristics or advantages: First, it does not require any form of integration or

union of functions, and second, it does not need to add a dummy variable to the data or to the function as in previous publications [10, 11].

Theorem 1. The coupled function with n sigmoidal logistic functions is:

$$F(x) = \frac{H_1}{1 + 2^{A_1 - B_1 x}} + \frac{H_2}{1 + 2^{A_2 - B_2 x}} + \dots + \frac{H_i}{1 + 2^{A_i - B_i x}} + \dots + \frac{H_n}{1 + 2^{A_n - B_n x}}, \quad (4)$$

where $F(x)$ is the expected cumulative number of deaths, x is the number of days since the first case, n is the number of waves, $1 \leq i \leq n$, H_i is the height of the wave i , A_i and B_i are constants related to the pandemic.

Furthermore, $x_0 < x_1 < x_2 < \dots < x_n$ are consecutive real numbers that determine the intervals of the logistic functions. The function $g(x)_i$ is defined on the interval $x \in [x_{i-1}, x_i]$; $\forall x > x_i$, $g(x)_i$ tends to H_i , and $\forall x < x_{i-1}$, $g(x)_i$ tends to 0 (limits of the sigmoidal logistic function).

Proof. Let $x_0 < x_1 < x_2 < \dots < x_n$ be consecutive real numbers determining the intervals of logistic functions. The function $g(x)_1$ is defined on the initial interval $x \in [x_0, x_1]$, and for $x > x_1$, $g(x)_1$ tends to the constant H_1 (right asymptote of a sigmoidal function, $\lim_{x \rightarrow \infty} g(x)_1 = H_1$). In general, the function $g(x)_i$ is defined on the interval $x \in [x_{i-1}, x_i]$ (the i -th "wave"); $\forall x > x_i$, $g(x)_i$ tends to H_i , and $\forall x < x_{i-1}$, $g(x)_i$ tends to 0.

Proof of $F(x)$ by mathematical induction: Base case (trivial), for two functions; if $x < x_0$, then $F(x) \approx 0$; if $x \in [x_0, x_1]$, then $F(x) \approx g(x)_1$; if $x \in [x_1, x_2]$, then $F(x) \approx H_1 + g(x)_2 \approx g(x)_1 + g(x)_2$, and if $x > x_2$, then $F(x) \approx H_1 + H_2$. Induction hypothesis: $F(x) \approx \sum_{k=1}^{i-1} \left(\frac{H_k}{1 + 2^{A_k - B_k x}} \right)$, $\forall x \in$

Sigmoidal function	Number of power operations	Number of multiplication or division operations	Number of addition or subtraction operations
Gompertz function: $G(x) = a \cdot e^{-e^{b-c \cdot x}}$	2	2	1
Hyperbolic tangent function: $\tanh(x) = U - \frac{H}{1+e^{2(x-A)}}$	1	2	3
Boltzmann function: $F(x) = \frac{H}{1+e^{(C-x)/D}}$	1	2	2
Logistic function: $F(x) = \frac{H}{1+e^{A-Bx}}$	1	2	2

Table 4: Number of operations in sigmoidal functions

$[x_0, x_{i-1}]$. Then, $F(x) \approx 0$, $\forall x < x_0$; $F(x) \approx \sum_{k=1}^i \left(\frac{H_k}{1+2^{A_k-B_k x}} \right)$, $\forall x \in [x_0, x_i]$ (by the induction hypothesis); and $F(x) \approx \sum_{k=1}^i H_k$, $\forall x > x_i$, because, $\lim_{x \rightarrow -\infty} g(x)_k = 0$, $\forall x > x_i$, and $\forall_k > i$.

Finally, $F(x)$ is defined by the equation (5). \square

The model is understandable and efficient in running time.

The construction of F extends to other functions, and it is not necessary that each $g(x)$ be defined exclusively on an interval. The only condition is that the limits of $F(x)$ as x tends to $\pm\infty$ are constant.

Theorem 2. The coupled function with n functions is:

$$F(x) = h(x)_1 + h(x)_2 + \cdots + h(x)_i + \cdots + h(x)_n, \quad (5)$$

where, each $h(x)_i$ is a function with left and right horizontal asymptotes, that is,

$\lim_{x \rightarrow -\infty} h(x) = L$, and $\lim_{x \rightarrow \infty} h(x) = R$, where L and R are constants.

The proof is similar to Theorem 1. If there are two or more functions with total or partial intersection of their intervals, that is, they are defined in the same interval, they are added.

Figure 1 shows the coupling of four functions: Gompertz, Logistic, algebraic, and wave function. Only the first two are properly sigmoidal, the third is an algebraic function with sigmoidal behavior, and the last is the wave function. It is possible to include different functions that meet the conditions explained, for example, the normal distribution function. On the other hand, the quadratic function $y = x^2$, cannot be coupled.

The function $y = F(x)$ is described by equation (6):

$$y = \underbrace{2 \cdot e^{-e^{2-x}}}_{\text{Gompertz f.}} + \underbrace{\frac{2}{1+e^{10-x}}}_{\text{Sigmoidal f.}} + \underbrace{\frac{(x-20)}{1+\sqrt{1+(x-20)^2}}}_{\text{Algebraic function}} - \underbrace{4e^{-\frac{(x-30)^2}{4}} \sin\left(\frac{x-30}{4}\right)}_{\text{Wave function}} + 1 \quad (6)$$

The wave function satisfies the requirement of Theorem 2. Proof: By definition $-1 \leq \sin(x) \leq 1$, it follows that, $-\frac{1}{e^{x^2}} \leq \frac{\sin(x)}{e^{x^2}} \leq \frac{1}{e^{x^2}}$. Since the limits of both functions are equal, $\lim_{x \rightarrow \infty} -\frac{1}{e^{x^2}} = \lim_{x \rightarrow \infty} \frac{1}{e^{x^2}} = 0$, it follows that $\lim_{x \rightarrow \infty} e^{-x^2} \sin(x) = 0$ (squeeze theorem), and in the same way we can obtain $\lim_{x \rightarrow -\infty} e^{-x^2} \sin(x) = 0$. \square

Finally, the code to graph $F(x)$ in R is:

```

y = function(x){
  2*e^(-e^(2-x)) + 2/(1+e^(10-x)) +
  (x-20)/sqrt(1+(x-20)^2) -
  4*e^(-(x-30)^2/4)*sin((x-30)/4) + 1
}
plot(y, 0, 37, col="green")

```

4.5 The Logistic function of the cumulative number of deaths

The logistic function $F(x)$ is a set of functions, which are ordered sequentially and do not overlap. Each function corresponds to a wave of COVID-19.

Figure 2 represents the cumulative number of deaths from COVID-19. The dataset (observed data) are represented in gray, and the logistic function in blue. Remark, while the sigmoidal shape of the later waves cannot be distinguished due to the scale (thousands on the ordinate axis), it is visible if these (one or two) waves are plotted exclusively.

In all cases, the first two or three waves were relatively larger, and they lasted longer (their length on the x-axis is the duration in days); in addition, they

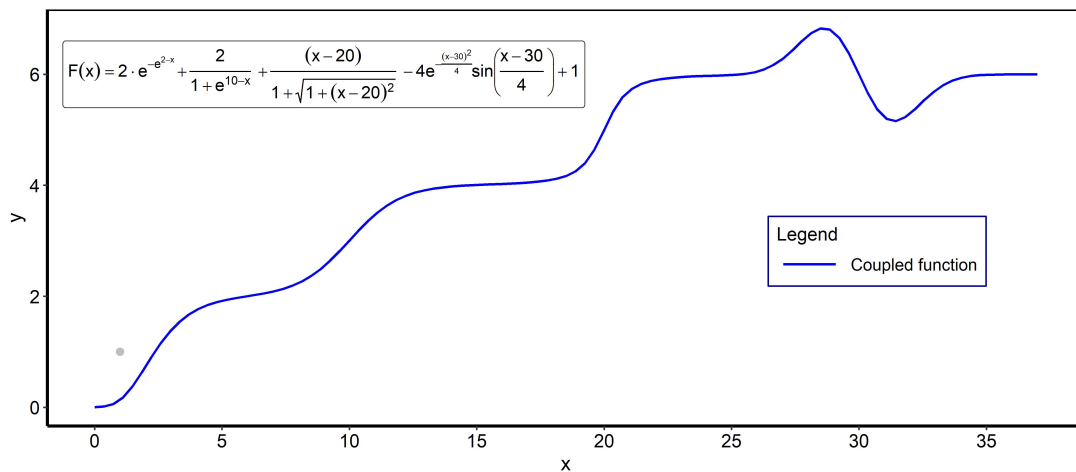


Figure 1: The coupled function $F(x)$ is made up of four functions: Gompertz, Logistic, algebraic, and wave. Only the first two are properly sigmoidal, the third is an algebraic function with sigmoidal behavior and the last is the wave function.

Country	Number of days	Pearson Correlation coefficient σ	Coefficient of Determination	Root square error	mean
Italy	1113	0.9994238	0.9988479	1.989923	
Brazil	1088	0.9999283	0.9998566	3.013632	
Switzerland	1099	0.9999178	0.9998356	0.061266	
Peru	1099	0.9999050	0.99981	1.022324	

Table 5: Function adjustment measures.

had a greater number of victims (represented by the height on the ordinate axis). Meanwhile, subsequent waves were considerably shorter in duration and had fewer victims. In each country, the last wave ends almost flat (the slope tends to zero); under these circumstances, the pandemic ended.

The Pearson correlation coefficients were greater than $R = 0.999$ (see the third column of Table 5), indicating very strong positive linear associations; that is, a high degree of association between the observed data and the function. In the scatter plot, the observed data are very close to the curve. The best adjustment was obtained for Brazil while that for Italy was the worst, but the differences were minimal. The coefficients of determination (fourth column) indicate that the functions obtained are explained or determined by the differences in the days.

The results are very good. Finally, the following hypothesis was accepted: The existence of a correlation that is greater than zero or a statistically significant correlation (with p-value = 2.2×10^{-16} , less than the significance level of 0.01).

Equations (7), (8), (9), and (10) represent the functions $F(x)$:

$$\underbrace{F(x)}_{Italy} = \frac{31.64}{1 + 2^{6.6-0.141*x}} + \frac{98.11}{1 + 2^{10.94-0.033*x}} + \frac{35.82}{1 + 2^{32.31-0.045*x}} + \frac{12.5}{1 + 2^{45.46-0.051*x}} + \frac{10.71}{1 + 2^{57.88-0.056*x}} \quad (7)$$

$$\underbrace{F(x)}_{Brazil} = \frac{152.06}{1 + 2^{5.62-0.045*x}} + \frac{464.04}{1 + 2^{11.9-0.029*x}} + \frac{48.72}{1 + 2^{51.47-0.074*x}} + \frac{21.45}{1 + 2^{56.44-0.066*x}} + \frac{14.2}{1 + 2^{50.35-0.049*x}} \quad (8)$$

$$\underbrace{F(x)}_{Switzerland} = \frac{1.77}{1 + 2^{6.49-0.192*x}} + \frac{7.51}{1 + 2^{22.68-0.082*x}} + \frac{1.5}{1 + 2^{14.71-0.041*x}} + \frac{3.1}{1 + 2^{24.01-0.036*x}} + \frac{0.36}{1 + 2^{33.8-0.035*x}} \quad (9)$$

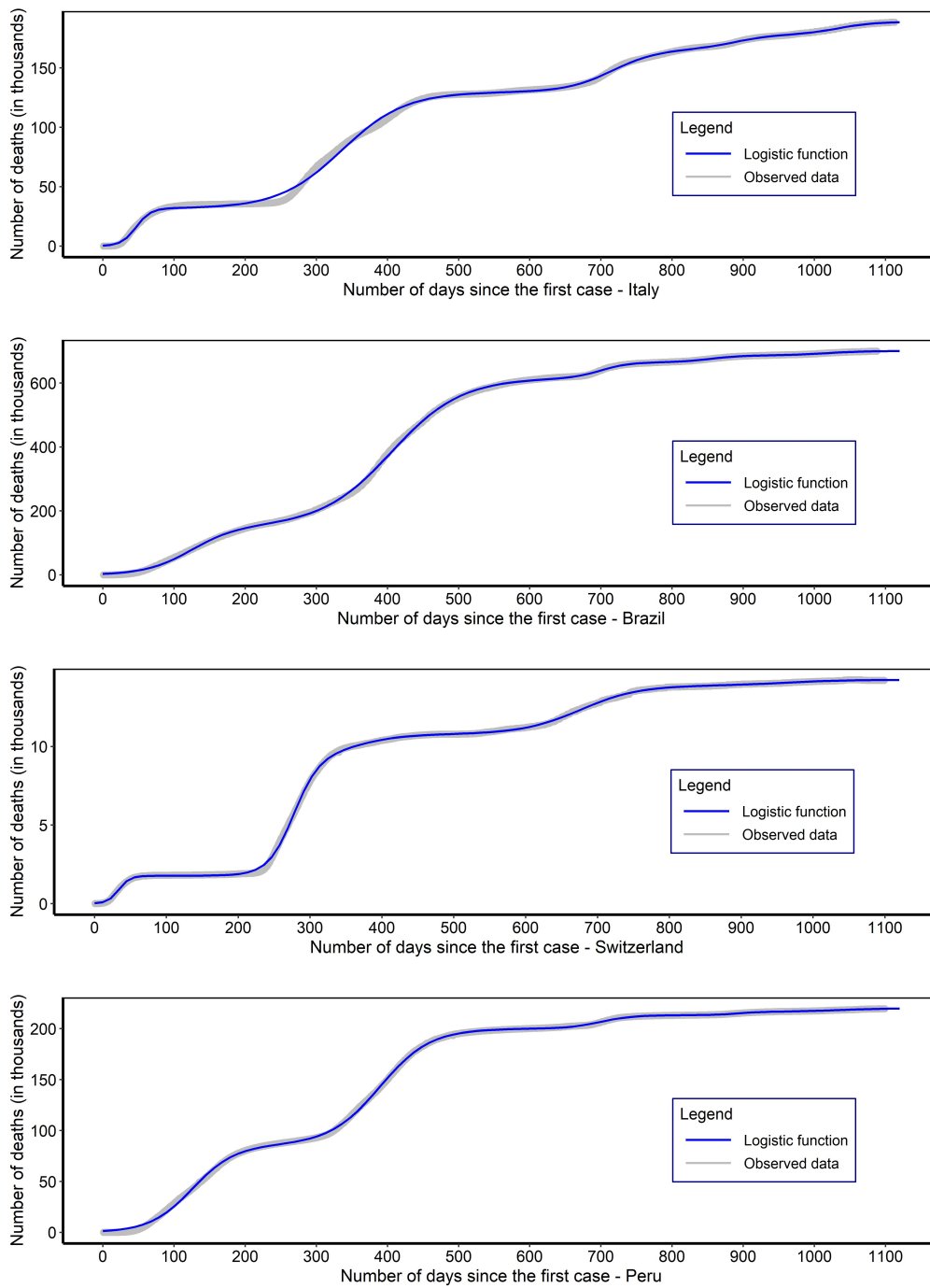


Figure 2: Cumulative number of deaths from COVID-19 (in thousands) by day.

4.6 First derivative

The previous subsection was limited to describing the waves through visual means. To better interpret the mortality rate—represented by the slope or gradient of the curves—the first and second derivative of the functions F were calculated. The third derivative was also calculated, in order to confirm or check the inflection points. In this way, a correct and elegant interpretation of the results was possible [22].

$$\underbrace{F(x)}_{Peru} = \frac{85.94}{1 + 2^{6.07 - 0.048 * x}} + \frac{114.1}{1 + 2^{15.96 - 0.041 * x}} + \frac{13.22}{1 + 2^{44.35 - 0.063 * x}} + \frac{3.4}{1 + 2^{73.16 - 0.082 * x}} + \frac{3.33}{1 + 2^{46.8 - 0.045 * x}} \quad (10)$$

The first derivative of the Logistic function F or equation (2) is given in the equation (11):

$$F'(x) = \frac{B H \ln(2) 2^{A-Bx}}{(1 + 2^{A-Bx})^2}, \quad (11)$$

where x is the number of days since the first case, $F'(x)$ is the death rate, H is the height of the wave, A and B are the pandemic constants, and \ln is the natural logarithm of base e ; that is, $\ln(x) = \log_e(x)$.

The first derivative is useful as it represents the slope of the curve or the speed of spread. It is also used to calculate the second derivative.

4.7 Second derivative

To obtain the inflection point, it is necessary to calculate the second derivative.

The equation (12) is the second derivative of the logistic function $F(x)$; that is, the derivative of the equation (11):

$$F''(x) = -\frac{B^2 H \ln(2)^2 2^{A-Bx} (1 - 2^{A-Bx})}{(1 + 2^{A-Bx})^3}, \quad (12)$$

where x is the number of days since the first case and $F''(x)$ is the second derivative (also called the acceleration) of the cumulative number of deaths.

4.8 Third derivative

The third derivative $F'''(x)$ is:

$$\frac{B^3 H \ln(2)^3 2^{A-Bx} (1 - 4 \cdot 2^{A-Bx} + 2^{2(A-Bx)})}{(1 + 2^{A-Bx})^4} \quad (13)$$

4.9 Inflection point

A point c on a curve $y = F(x)$ is called an inflection point if F is continuous there and the curve changes from concave upward to concave downward (or vice versa) at c [23, 24].

To obtain the inflection point(s), it is first necessary to find the critical point of $F'(x)$; that is, to solve the following equation:

$$F''(x) = 0. \quad (14)$$

The solution to the equation (14) is given by $x_0 = \frac{A}{B}$ and $y_0 = \frac{H}{2}$.

Finally, $F'''(x_0)$ is evaluated as follows:

$$\begin{aligned} F'''(x_0) &= -\frac{B^3 H \ln(2)^3}{8}, \\ F'''(x_0) &= -0.0416 B^3 H. \end{aligned} \quad (15)$$

The result is different from zero (with $B \neq 0$ and $H \neq 0$), which indicates that $(\frac{A}{B}, \frac{H}{2})$ is an inflection point. It should be highlighted that, according to the method of Higher Derivatives, if $F''(x_0) = 0$ holds, then x_0 must be substituted into the third derivative. Then, if $F'''(x_0) \neq 0$ holds, then it is an inflection point.

4.10 Applying high order derivatives

We continued by deriving the functions $F(x)$. The results are equations (16), (17), (18), and (19), represent the first derivatives for Italy, Mexico, Brazil, Switzerland, and Peru, respectively, using equation (11).

$$\begin{aligned} \underbrace{F'(x)}_{Italy} &= \frac{3.1 \cdot 2^{6.6-0.141*x}}{(1 + 2^{6.6-0.141*x})^2} + \frac{2.22 \cdot 2^{10.94-0.033*x}}{(1 + 2^{10.94-0.033*x})^2} \\ &+ \frac{1.12 \cdot 2^{32.31-0.045*x}}{(1 + 2^{32.31-0.045*x})^2} + \frac{0.44 \cdot 2^{45.46-0.051*x}}{(1 + 2^{45.46-0.051*x})^2} + \frac{0.42 \cdot 2^{57.88-0.056*x}}{(1 + 2^{57.88-0.056*x})^2} \end{aligned} \quad (16)$$

$$\begin{aligned} \underbrace{F'(x)}_{Brazil} &= \frac{4.76 \cdot 2^{5.62-0.045*x}}{(1 + 2^{5.62-0.045*x})^2} + \frac{9.44 \cdot 2^{11.9-0.029*x}}{(1 + 2^{11.9-0.029*x})^2} \\ &+ \frac{2.48 \cdot 2^{51.47-0.074*x}}{(1 + 2^{51.47-0.074*x})^2} + \frac{0.98 \cdot 2^{56.44-0.066*x}}{(1 + 2^{56.44-0.066*x})^2} + \frac{0.49 \cdot 2^{50.35-0.049*x}}{(1 + 2^{50.35-0.049*x})^2} \end{aligned} \quad (17)$$

$$\begin{aligned} \underbrace{F'(x)}_{Switzerland} &= \frac{0.24 \cdot 2^{6.49-0.192*x}}{(1 + 2^{6.49-0.192*x})^2} + \frac{0.43 \cdot 2^{22.68-0.082*x}}{(1 + 2^{22.68-0.082*x})^2} \\ &+ \frac{0.04 \cdot 2^{14.71-0.041*x}}{(1 + 2^{14.71-0.041*x})^2} + \frac{0.08 \cdot 2^{24.01-0.036*x}}{(1 + 2^{24.01-0.036*x})^2} + \frac{0.01 \cdot 2^{33.8-0.035*x}}{(1 + 2^{33.8-0.035*x})^2} \end{aligned} \quad (18)$$

$$\begin{aligned} \underbrace{F'(x)}_{Peru} &= \frac{2.87 \cdot 2^{6.07-0.048*x}}{(1 + 2^{6.07-0.048*x})^2} + \frac{3.24 \cdot 2^{15.96-0.041*x}}{(1 + 2^{15.96-0.041*x})^2} \\ &+ \frac{0.58 \cdot 2^{44.35-0.063*x}}{(1 + 2^{44.35-0.063*x})^2} + \frac{0.19 \cdot 2^{73.16-0.082*x}}{(1 + 2^{73.16-0.082*x})^2} + \frac{0.1 \cdot 2^{46.8-0.045*x}}{(1 + 2^{46.8-0.045*x})^2} \end{aligned} \quad (19)$$

To calculate the second derivative of the function $F(x)$, the equation (12) was used.

The inflection points play an important role, as they can be considered as so-called “happy days;” that is, the day from which the infection rate begins to decrease (mathematically, the slope that was growing

Wave	Brazil	Italy	Switzerland	Peru
1	124.38	46.68	33.87	126.08
2	405.60	335.82	277.2	389.42
3	699.89	716.04	358.53	699.39
4	854.16	888.31	676.27	892.89
5	1018.16	1034.38	964.09	1037.25

Table 6: Inflection points by country.

begins to decrease). On the graph, this is the point where it changes from concave up to concave down (or vice versa, in other circumstances).

Finally, the inflection points are detailed in Table 6.

4.11 Interpretation

The first derivative of the function $F(x)$ allows us to describe the behavior of the rate or speed of the number of deaths, and is also understood as the slope or gradient of the function $F(x)$.

Figure 3 shows the graphs of the first derivatives of $F(x)$ for each country as blue curves, where the inflection points are indicated by red vertical lines. For comparison between waves or countries, both axes have the same scale. It can be observed that each wave begins with an increasing rate, reaches an inflection point (change), and then begins to decrease. This provides an elegant and analytically based justification for the term "wave".

Definition 4.1 (x). The value of the abscissa x are the days elapsed since the first case or cases. Then, the range (i.e., the number of days from an initial value x_i to another final value x_j , with $x_i \leq x_j$) is the duration or the number of days ($x_j - x_i + 1$). A full wave, x is the width (or the duration in days). This is valid with respect to the function $F(x)$, the first derivative $F'(x)$, all of the derivatives in general, and their respective graphs.

Definition 4.2 ($F'(x)$). The value of the ordinate $F'(x)$ is the slope of the original curve $F(x)$ at point x . In the case study, the slope is the change of the death rate at the point x (i.e., on day x , if $x \in \mathbb{N}$). It is also the velocity and, so, if this value is higher than that on another day, it can be immediately understood as indicating higher lethality.

Definition 4.3 ($F''(x)$). The value of the ordinate $F''(x)$ is the change of rate of $F'(x)$ at point x . In the case study, is the acceleration at the point x (i.e., on day x , if $x \in \mathbb{N}$).

The second wave in Brazil was the largest (in the first figure, it can be observed that it led to more than 400,000 deaths), and its greater height implies that it had a greater lethality.

The first waves in Italy and Switzerland reached a lower death rate and shorter duration, which may

be for different reasons, including strict social isolation measures and population differences. In Peru and Brazil, it lasted more than 200 days and had fatality rates below and above one, respectively. Using the proposed model, it was possible to characterize the waves of COVID-19, and it is expected that it can be extended to other countries (which also present similar patterns, according to our pilot study).

Figure 4 represents the graphs of the second derivative of $F(x)$ (for Italy and Brazil only) as blue curves, with the inflection points indicated by dashed vertical lines in red. It can be understood as the acceleration of the number of deaths accumulated due to COVID-19. As the second derivative represents the concavity, when the graph crosses the abscissa axis, it indicates that point as an inflection point; here, the second derivative must change sign, which is interpreted as a change in concavity (positive indicates upward concavity and negative indicates downward concavity). Finally, the second derivatives $F''(x)$ are given as equations (20), and (21).

$$\underbrace{F''(x)}_{Italy} = \frac{0.3 \cdot 2^{6.6-0.141*x}(1 - 2^{6.6-0.141*x})}{(1 + 2^{6.6-0.141*x})^3} + \quad (20)$$

$$\begin{aligned}
 & \frac{0.05 \cdot 2^{10.94-0.033*x}(1 - 2^{10.94-0.033*x})}{(1 + 2^{10.94-0.033*x})^3} + \\
 & \frac{0.04 \cdot 2^{32.31-0.045*x}(1 - 2^{32.31-0.045*x})}{(1 + 2^{32.31-0.045*x})^3} + \\
 & \frac{0.02 \cdot 2^{45.46-0.051*x}(1 - 2^{45.46-0.051*x})}{(1 + 2^{45.46-0.051*x})^3} + \\
 & \frac{0.02 \cdot 2^{57.88-0.056*x}(1 - 2^{57.88-0.056*x})}{(1 + 2^{57.88-0.056*x})^3}
 \end{aligned}$$

$$\underbrace{F''(x)}_{Brazil} = \frac{0.15 \cdot 2^{5.62-0.045*x}(1 - 2^{5.62-0.045*x})}{(1 + 2^{5.62-0.045*x})^3} + \quad (21)$$

$$\begin{aligned}
 & \frac{0.19 \cdot 2^{11.9-0.029*x}(1 - 2^{11.9-0.029*x})}{(1 + 2^{11.9-0.029*x})^3} + \\
 & \frac{0.13 \cdot 2^{51.47-0.074*x}(1 - 2^{51.47-0.074*x})}{(1 + 2^{51.47-0.074*x})^3} + \\
 & \frac{0.05 \cdot 2^{56.44-0.066*x}(1 - 2^{56.44-0.066*x})}{(1 + 2^{56.44-0.066*x})^3} + \\
 & \frac{0.02 \cdot 2^{50.35-0.049*x}(1 - 2^{50.35-0.049*x})}{(1 + 2^{50.35-0.049*x})^3}
 \end{aligned}$$

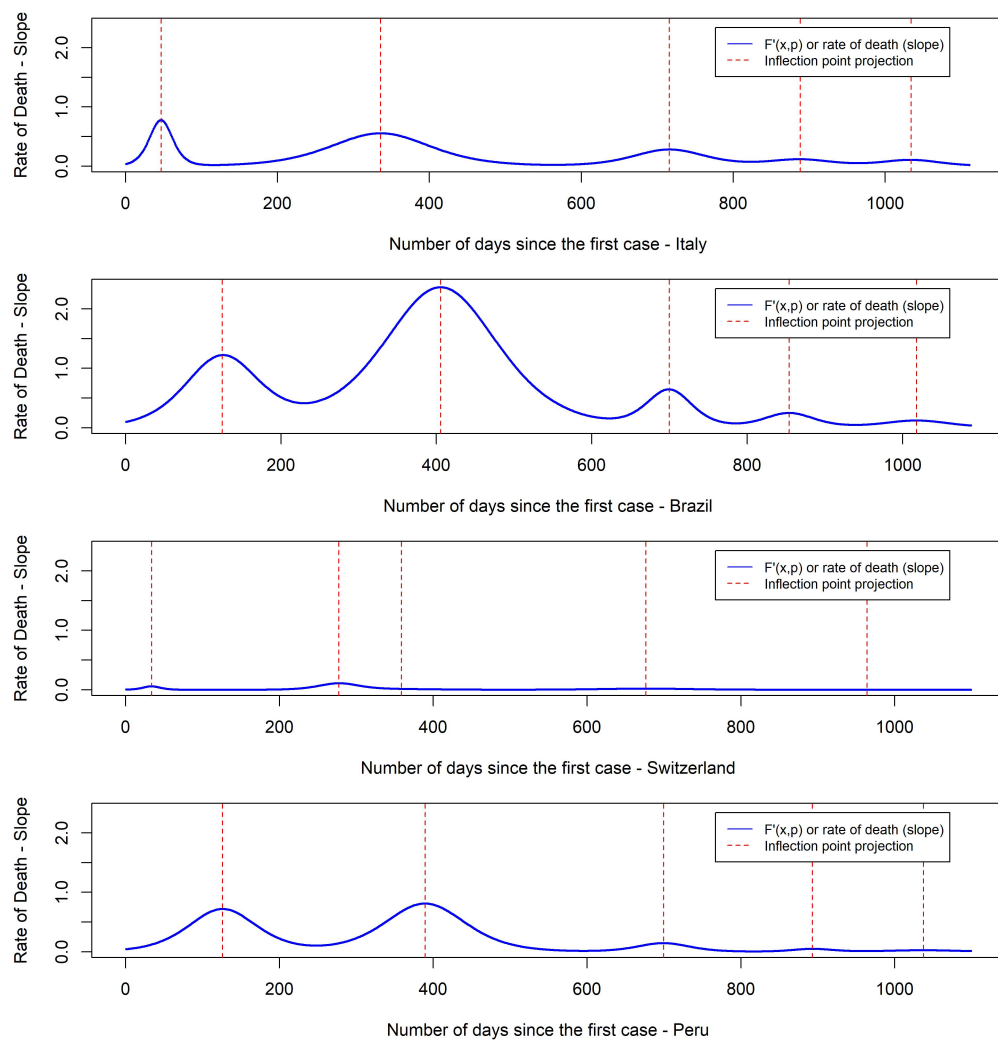


Figure 3: The first derivative of the COVID-19 logistic function (blue curve) and the vertical projection of the inflection points (red dashed line).

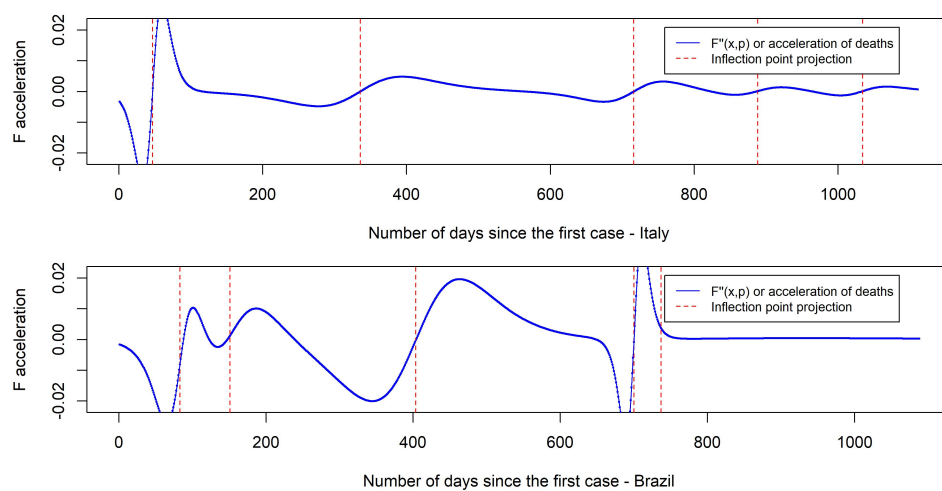


Figure 4: Graph of the second derivative of the logistic function of cumulative COVID-19 deaths in blue (Italy and Brazil) and the projections of the inflection points as red vertical lines.

5 Conclusion

In almost all countries, and particularly in the study countries (i.e., Italy, Brazil, Switzerland, and Peru), the COVID-19 waves presented a sigmoidal behavior. Therefore, the proposed model obtained high Pearson correlation coefficients.

Through analysis of the type and quantity of operations, it was shown that the logistic function is optimal; that is, it has the shortest execution time compared to similar functions and it is not possible to reduce its operations. Additionally, a new version of the logistic function $g(x)$ was obtained and used, without the addition of operations.

The first three derivatives of the sigmoidal logistic function were calculated, allowing for a precise and elegant analysis. Moreover, the inflection points were obtained (i.e., the point at which the concavity of the function F changes). The derivatives allowed the waves of COVID-19 to be described, analyzed, and compared, as well as providing justification for the well-known term “waves.”

Highly representative and expressive sigmoidal logistic functions were obtained, which stand out for their efficiency in running time, far surpassing that of the tested artificial neural network. A good fit was obtained with the function, as the waves presented sigmoidal behavior which is typical of epidemics or pandemics.

The ANN can adapt to any curve without the requirement that it presents a specific pattern. Therefore, such a model would be preferable if atypical behaviors were present in the observed data, but at the expense of additional time. Also, are suitable for large data sets and high dimensionality in the input variables.

As a final point, neither model is better than the other, and both have their advantages and disadvantages. We consider both to be good for modeling epidemics, particularly when the data follow a sigmoidal pattern.

The main contribution of this work is the general model that can be used on data that are not necessarily epidemiological; and with any function $h(x)$ with constant left and right asymptotes (that is, $\lim_{x \rightarrow -\infty} h(x) = L$, and $\lim_{x \rightarrow \infty} h(x) = R$, where L and R are constants, proven by mathematical induction). Also, the method of its construction, and the calculation of high-order derivatives, which allow for the development of an elegant and useful model, as well as justification of the term “Wave(s)”.

Future work may extend the proposed model to other countries. To build a model $F(x)$, combine different functions (not necessarily sigmoidal). The study of the effects of the waves on different variables, such as the resources used during social isolation measures, is also pending.

Funding statement

This research was funded by Universidad Nacional del Altiplano (UNA), IPN SIP 20231387 project (RR. N. 1589-2023-R-UNA)

Conflict of interest

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Acknowledgement

To the administrative staff and the Vice-Rectorate of Research of the Universidad Nacional del Altiplano for their combined efforts and support.

References

- [1] A. N. Sajed and K. Amgain, “Corona Virus Disease (COVID-19) Outbreak and the Strategy for Prevention,” *Europasian Journal of Medical Sciences*, vol. 2, no. 1, pp. 1–3, july 2020, number: 1. [Online]. Available: <http://nepmed.nhrc.gov.np/index.php/EJMS/article/view/836>
- [2] A. Din, Y. Li, T. Khan, and G. Zaman, “Mathematical analysis of spread and control of the novel corona virus (COVID-19) in China,” *Chaos, Solitons & Fractals*, vol. 141, p. 110286, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960077920306822>
- [3] O. O. Ogunleye, D. Basu, D. Mueller, J. Sneddon, R. A. Seaton, A. F. Yinka-Ogunleye, J. Wamboga, N. Miljković, J. C. Mwita, G. M. Rwegerera, A. Massele, O. Patrick, L. L. Niba, M. Nsaikila, W. M. Rashed, M. A. Hussein, R. Hegazy, A. A. Amu, B. B. Boahen-Boaten, Z. Matsebula, P. Gwebu, B. Chirigo, N. Mkhabela, T. Dlamini, S. Sithole, S. Malaza, S. Dlamini, D. Afriyie, G. A. Asare, S. K. Amponsah, I. Sefah, M. Oluka, A. N. Guantai, S. A. Opanga, T. V. Sarele, R. K. Mafisa, I. Chikowe, F. Khuluza, D. Kibuule, F. Kalemeera, M. Mubita, J. Fadare, L. Sibomana, G. M. Ramokgopa, C. Whyte, T. Maimela, J. Hugo, J. C. Meyer, N. Schellack, E. M. Rampamba, A. Visser, A. Alfadl, E. M. Malik, O. O. Malande, A. C. Kalungia, C. Mwila, T. Zaranyika, B. V. Chaibva, I. D. Olaru, N. Masuka, J. Wale, L. Hwenda, R. Kamoga, R. Hill, C. Barbui, T. Bochenek, A. Kurdi, S. Campbell, A. P. Martin, T. N. T. Phuong, B. N. Thanh, and B. Godman, “Response to the Novel Corona Virus (COVID-19)

- Pandemic Across Africa: Successes, Challenges, and Implications for the Future,” *Frontiers in Pharmacology*, vol. 11, 2020.
- [4] S. Bhagat, N. Yadav, J. Shah, H. Dave, S. Swaraj, S. Tripathi, and S. Singh, “Novel corona virus (COVID-19) pandemic: current status and possible strategies for detection and treatment of the disease,” *Expert Review of Anti-infective Therapy*, vol. 20, no. 10, pp. 1275–1298, oct 2022. [Online]. Available: <https://doi.org/10.1080/14787210.2021.1835469>
 - [5] S. A. Tabish, “Covid-19 Pandemic: Emerging Perspectives and Future Trends,” *Journal of Public Health Research*, vol. 9, no. 1, p. jphr.2020.1786, june 2020, publisher: SAGE Publications. [Online]. Available: <https://doi.org/10.4081/jphr.2020.1786>
 - [6] S. Kumar, “Monitoring Novel Corona Virus (COVID-19) Infections in India by Cluster Analysis,” *Annals of Data Science*, vol. 7, no. 3, pp. 417–425, sep 2020. [Online]. Available: <https://doi.org/10.1007/s40745-020-00289-7>
 - [7] M. A. Alsalem, A. H. Alamoodi, O. S. Albahri, K. A. Dawood, R. T. Mohammed, A. Alnoor, A. A. Zaidan, A. S. Albahri, B. B. Zaidan, F. M. Jumaah, and J. R. Al-Obaidi, “Multi-criteria decision-making for coronavirus disease 2019 applications: a theoretical analysis review,” *Artificial Intelligence Review*, vol. 55, no. 6, pp. 4979–5062, aug 2022. [Online]. Available: <https://doi.org/10.1007/s10462-021-10124-x>
 - [8] N. Pearce, D. A. Lawlor, and E. B. Brickley, “Comparisons between countries are essential for the control of COVID-19,” *International Journal of Epidemiology*, vol. 49, no. 4, pp. 1059–1062, jun 2020, eprint: <https://academic.oup.com/ije/article-pdf/49/4/1059/34275473/dyaa108.pdf>. [Online]. Available: <https://doi.org/10.1093/ije/dyaa108>
 - [9] M. Jawad Hashim, A. R. Alsuwaidi, and G. Khan, “Population Risk Factors for COVID-19 Mortality in 93 Countries,” *Journal of Epidemiology and Global Health*, vol. 10, no. 3, pp. 204–208, september 2020. [Online]. Available: <https://doi.org/10.2991/jegh.k.200721.001>
 - [10] O. A. Vilca-Huayta and U. Y. Tito, “Efficient function integration and a case study with gompertz functions for covid-19 waves,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 8, 2022. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2022.0130863>
 - [11] O. A. Vilca Huayta, A. C. Jimenez Chura, C. B. Sosa Maydana, and A. J. Martínez García, “Analysis of the epidemic curve of the waves of covid-19 using integration of functions and neural networks in peru,” *Informatics*, vol. 11, no. 2, 2024. [Online]. Available: <https://www.mdpi.com/2227-9709/11/2/40>
 - [12] R. Sujath, J. M. Chatterjee, and A. E. Hassanien, “A machine learning forecasting model for COVID-19 pandemic in india,” *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 7, pp. 959–972, 2020. [Online]. Available: <https://doi.org/10.1007/s00477-020-01827-8>
 - [13] F. Fernandes, S. F. Stefenon, L. O. Seman, A. Nied, F. C. S. Ferreira, M. C. M. Subtil, A. C. R. Klaar, and V. R. Q. Leithardt, “Long short-term memory stacking model to predict the number of cases and deaths caused by COVID-19,” *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 6, pp. 6221–6234, 2022, publisher: IOS Press.
 - [14] S. Shastri, K. Singh, S. Kumar, P. Kour, and V. Mansotra, “Time series forecasting of covid-19 using deep learning models: India-usa comparative case study,” *Chaos, Solitons and Fractals*, vol. 140, p. 110227, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960077920306238>
 - [15] S. Namasudra, S. Dhamodharavadhani, and R. Rathipriya, “Nonlinear Neural Network Based Forecasting Model for Predicting COVID-19 Cases,” *Neural Processing Letters*, vol. 55, no. 1, pp. 171–191, 2023.
 - [16] S. Abolmaali and S. Shirzaei, “A comparative study of SIR Model, Linear Regression, Logistic Function and ARIMA Model for forecasting COVID-19 cases,” *AIMS Public Health*, vol. 8, no. 4, pp. 598–613, aug 2021. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8568588/>
 - [17] J. H. University, “Github - cssegisanddata/covid-19: Novel coronavirus (covid-19) cases, provided by jhu csse.” [Online]. Available: <https://github.com/CSSEGISandData/COVID-19>
 - [18] M. Haouari and M. Mhiri, “A particle swarm optimization approach for predicting the number of COVID-19 deaths,” *Scientific Reports*, vol. 11, no. 1, p. 16587, Aug. 2021. [Online]. Available: <https://www.nature.com/articles/s41598-021-96057-5>

- [19] M. Visková and M. Vik, “Transition Temperature of Color Change in Thermochromic Systems and Its Description Using Sigmoidal Models,” *Materials*, vol. 16, no. 23, p. 7478, jan 2023, number: 23 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: <https://www.mdpi.com/1996-1944/16/23/7478>
- [20] D. K. Sharma, M. Chatterjee, G. Kaur, and S. Vavilala, “Deep learning applications for disease diagnosis,” *Deep Learning for Medical Applications with Unique Data*, pp. 31–51, 1 2022.
- [21] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation,” *PeerJ Computer Science*, vol. 7, pp. 1–24, 7 2021. [Online]. Available: <https://peerj.com/articles/cs-623>
- [22] K. Heusser, R. Heusser, J. Jordan, V. Urech, A. Diedrich, and J. Tank, “Baroreflex Curve Fitting Using a WYSIWYG Boltzmann Sigmoidal Equation,” *Frontiers in Neuroscience*, vol. 15, 2021. [Online]. Available: <https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2021.697582>
- [23] G. Gu, P. Zhang, S. Chen, Y. Zhang, and H. Yang, “Inflection point: a perspective on photonic nanojets,” *Photonics Research*, vol. 9, no. 7, pp. 1157–1171, july 2021, publisher: Optica Publishing Group. [Online]. Available: <https://opg.optica.org/prj/abstract.cfm?uri=prj-9-7-1157>
- [24] Y. A. Hong and J. W. Ha, “Enhanced refractive index sensitivity of localized surface plasmon resonance inflection points in single hollow gold nanospheres with inner cavity,” *Scientific Reports*, vol. 12, no. 1, p. 6983, apr 2022, number: 1 Publisher: Nature Publishing Group. [Online]. Available: <https://www.nature.com/articles/s41598-022-11197-6>
- [25] M. Dai and Z. Huang, “Research on fault diagnosis of drilling pump fluid end based on time-frequency analysis and convolutional neural network,” *Processes* 2024, Vol. 12, Page 1929, vol. 12, p. 1929, 9 2024. [Online]. Available: <https://www.mdpi.com/2227-9717/12/9/1929/htmhttps://www.mdpi.com/2227-9717/12/9/1929>
- [26] I. D. Mienye, T. G. Swart, and G. Obaido, “Recurrent neural networks: A comprehensive review of architectures, variants, and applications,” *Information* 2024, Vol. 15, Page 517, vol. 15, p. 517, 8 2024. [Online]. Available: <https://www.mdpi.com/2078-2489/15/9/517/htmhttps://www.mdpi.com/2078-2489/15/9/517>
- [27] R. C. Torcasio, S. Federico, E. Avolio, I. Ebtehaj, and H. Bonakdari, “Cnn vs. lstm: A comparative study of hourly precipitation intensity prediction as a key factor in flood forecasting frameworks,” *Atmosphere* 2024, Vol. 15, Page 1082, vol. 15, p. 1082, 9 2024. [Online]. Available: <https://www.mdpi.com/2073-4433/15/9/1082/htmhttps://www.mdpi.com/2073-4433/15/9/1082>
- [28] C. Turchetti, L. Falaschetti, Y. J. Kim, J. H. Song, K. H. Cho, J. H. Shin, J. S. Kim, J. S. Yoon, and S. J. Hong, “Improved plasma etch endpoint detection using attention-based long short-term memory machine learning,” *Electronics* 2024, Vol. 13, Page 3577, vol. 13, p. 3577, 9 2024. [Online]. Available: <https://www.mdpi.com/2079-9292/13/17/3577/htmhttps://www.mdpi.com/2079-9292/13/17/3577>