Logistic Wavelets and Their Application to Model the Spread of COVID-19 Pandemic

Grzegorz Rządkowski [1],* and **Giuseppe Figlia [2]**

[1] Department of Finance and Risk Management, Warsaw University of Technology, Narbutta 85, 02-524 Warsaw, Poland

[2] Independent Researcher, Certified Project Manager at P&P SPA Italy, C. Vercelli 168, 10155 Turin, Italy; giuseppe.figlia@pepspa.com

* Correspondence: grzegorz.rzadkowski@pw.edu.pl

Abstract: In the present paper, we model the cumulative number of persons, reported to be infected with COVID-19 virus, by a sum of several logistic functions (the so-called multilogistic function). We introduce logistic wavelets and describe their properties in terms of Eulerian numbers. Moreover, we implement the logistic wavelets into Matlab's Wavelet Toolbox and then we use the continuous wavelet transform (CWT) to estimate the parameters of the approximating multilogistic function. Using the examples of several countries, we show that this method is effective as a method of fitting a curve to existing data. However, it also has a predictive value, and, in particular, allows for an early assessment of the size of the emerging new wave of the epidemic, thus it can be used as an early warning method.

Keywords: logistic wavelet; logistic equation; logistic function; COVID-19 infection; Eulerian number; Riccati's differential equation; CWT scalograms

1. Introduction

The logistic equation is defined as

$$u'(t) = \frac{s}{u_{max}} u(u_{max} - u), \quad u(0) = u_0. \tag{1}$$

where t is time, $u = u(t)$ is the unknown function and the parameters s-steepness and u_{max}-saturation level are constants. The integral curve $u(t)$ fulfilling condition $0 < u(t) < u_{max}$ is known as the logistic function.

After solving (1) we obtain the logistic function in the form

$$u(t) = \frac{u_{max}}{1 + e^{-s(t-t_0)}}, \tag{2}$$

where t_0 is the inflection point, which is related to the initial condition $u(0) = u_0 = \frac{u_{max}}{1 + e^{st_0}}$, therefore $t_0 = \frac{1}{s} \log\left(\frac{u_{max} - u_0}{u_0}\right)$. Putting $t = t_0$ in (2) we see that $u(t_0) = u_{max}/2$.

The logistic function finds applications in many fields, including biology, biomathematics, chemistry, demography, economics, physics, probability, sociology, statistics, and artificial neural networks. The logistic function and the logistic equation, as well as some of their generalizations, have also been widely used in epidemiology to describe various phenomena with a sigmoid trend (see for example papers by Kartono et al. [1] and Pelinovsky et al. [2], and the references therein). Fokas et al. [3] used a generalization of the logistic function for forecasting the number of individuals reported to be infected with COVID-19 in different countries.

Wavelet analysis is now frequently used to extract information from epidemiological and other time series. Grenfell et al. [4] introduced wavelet analysis for characterizing non-stationary epidemiological time series. Cazelles et al. [5] used the Morlet wavelets for applications in epidemiology. Krantz et al. [6] proposed a two-phase procedure (combining discrete graphs and Meyer wavelets) for constructing true COVID-19 epidemic growth. Jose and Bishop [7] used the reverse biorthogonal wavelet pairs for modeling the rotavirus epidemic dynamics. They also mentioned that they obtained similar results when using the Haar wavelet. Wang et al. [8] apply the Daubechies wavelet (db2) and the Coiflets wavelet (coif1) for modeling of pertussis incidence in China from 2004 to 2018. Santos et al. [9] and Zhang et al. [10] use the wavelet analysis to investigate the correlation between the incidence of dengue and weather conditions. Liu [11] applies the Haar wavelet transform and the simplex forecasting method to the dataset of hepatitis A in the United States to give predictions of its incidence.

Lavrova et al. [12] modeled the disease dynamics caused by Mycobacterium tuberculosis in Russia using a sum of two logistic functions (2) (so-called bi-logistic model). Similar method was used by E. Vanucci and L. Vanucci [13] for predicting the end date of COVID-19 disease in Italy.

From a slightly broader point of view, the logistic Equation (1) can be considered in the context of the following non-linear first-order autonomous differential equation

$$\frac{du}{dt} = f(u),\tag{3}$$

where $f(u)$ is a real, continuous function of u, representing a non-linear part of the equation. Several exact forms for $f(u)$ have been studied by Tsoularis [14] and by Tsoularis and Wallace [15]. An advanced theory with important applications of Equation (3) is provided by Kowalski and Steeb [16]. The connections with Lie algebra, Bose–Einstein state, and quantum theory, are put there in evidence.

The method of separated variables gives a first stage of the solution of (3) in terms of $t(u)$. This function is explicit only if the integral of the reciprocal of $f(u)$ is exactly available. The second stage of the solution depends on if $t(u)$ can be inverted in a closed form. Most of the important equations are not exactly solvable in terms of $t(u)$ and we need to solve them with numerical integration and inversion.

The functional inversion can be performed with the Lagrange inversion method for some particular forms of $t(u)$ and with a general procedure described in a paper by Dominici [17] through recursive application of nested derivatives on the kernel $f(u)$. The result is expressed in terms of the Taylor expansions of $u(t)$ at a given point. One of the authors (GF) of the present paper is preparing a new efficient inversion technique based on an auxiliary exact function and through the fast direct derivative of $t(u)$ followed by a linear triangular recursive inversion algorithm. In his method, the nested derivatives are not necessary.

In the present paper, we introduce logistic wavelets and describe their properties in terms of Eulerian numbers. We add the second logistic wavelet into Matlab's Wavelet Toolbox, to be able to use the continuous wavelet transform (CWT). We then perform CWT analysis for the second differences of a smoothed total number of people reported as infected with the COVID-19 virus. We model the total number of people infected with the COVID-19 virus by the sum of several logistic functions (the so-called multilogistic function). The CWT analysis allows to estimate parameters of the successive logistic waves, which together form the multilogistic function. All three parameters (i.e., the inflection point, the steepness and the saturation level) of each logistic function can be read from the CWT scalograms. Then we use the non-linear generalized reduced gradient method, minimizing the RMSE error, to optimize the parameters, mainly the saturation levels. To show the accuracy and effectiveness of our method, we apply it to the cases of COVID-19 infection in several countries, as well as around the world. This method, in addition to the curve-fit property to existing data, has also a predictive value, and, in particular, allows for

an early assessment of the size of the emerging new wave of the epidemic. Thus it can be used as an early warning method.

This paper is organized as follows. In Section 2.1, we discuss basic properties of Riccati's equation, logistic equation, and logistic curve. For this purpose, we use Eulerian numbers. Sections 2.2 and 2.3 are devoted to introduction and study of the logistic wavelets. In Section 3 we model the cumulative number of persons reported to be infected by COVID-19 in several countries and also in the whole world. The results are discussed and concluded in Section 4.

We use the following convention for the Fourier transform:

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-i\xi x}dx, \tag{4}$$

where $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ (the intersection of the space of integrable functions and the space of square integrable functions defined on the set of real numbers $\mathbb{R}$).

2. Materials and Methods

2.1. Logistic Function and Its Derivatives

For the convenience of the reader, we now briefly describe the logistic function and some of its properties, which have been proven mainly in our paper [18].

Equation (1) is a particular case of Riccati's equation with constant coefficients

$$u'(t) = r(u - u_1)(u - u_2). \tag{5}$$

The constants $r \neq 0$, u_1, u_2 can be real or complex numbers.

If $u(t)$ is a solution of (5) then the nth derivative $u^{(n)}(t)$ $(n = 2, 3, 4, \ldots)$ of $u(t)$ can be expressed as a polynomial of the function $u(t)$ itself:

$$u^{(n)}(t) = r^n \sum_{k=0}^{n-1} \left\langle {n \atop k} \right\rangle (u - u_1)^{k+1}(u - u_2)^{n-k} \tag{6}$$

where $n = 2, 3, \ldots$ and $\left\langle {n \atop k} \right\rangle$ denotes the Eulerian number (number of permutations of the set $\{1, 2, \ldots, n\}$ having k, $(k = 0, 1, 2, \ldots, n-1)$ permutation ascents, see Graham et al. [19]). The first few Eulerian numbers are given in the Table 1.

Table 1. Eulerian numbers.

n	$\left\langle {n \atop 0} \right\rangle$	$\left\langle {n \atop 1} \right\rangle$	$\left\langle {n \atop 2} \right\rangle$	$\left\langle {n \atop 3} \right\rangle$	$\left\langle {n \atop 4} \right\rangle$	$\left\langle {n \atop 5} \right\rangle$	$\left\langle {n \atop 6} \right\rangle$	$\left\langle {n \atop 7} \right\rangle$
0	1							
1	1	0						
2	1	1	0					
3	1	4	1	0				
4	1	11	11	1	0			
5	1	26	66	26	1	0		
6	1	57	302	302	57	1	0	
7	1	120	1191	2416	1191	120	1	0

Formula (6) was discussed during the Conference ICNAAM 2006 and it appeared, with an inductive proof, in paper [20] (see also [21]). Independently the formula has been considered and proved, with the proof based on generating functions, by Franssens [22]. The polynomial of u, of order $(n + 1)$, appearing on the right-hand side of (1) is known in the literature as the so-called derivative polynomial. In [23], we proved that all $(n + 1)$ zeros of the polynomial are simple and lie in the interval $[u_1, u_2]$. The derivative polynomials have been recently intensively studied.

Formula (6) applied to the particular case of the logistic Equation (1) is as follows:

$$u^{(n)}(t) = \left(-\frac{s}{u_{max}}\right)^n \sum_{k=0}^{n-1} \left\langle {n \atop k} \right\rangle u^{k+1}(u - u_{max})^{n-k}. \tag{7}$$

It is easily seen that if u_0 is a zero of the polynomial on the right-hand side of (7) then u_0/u_{max} is the zero of the polynomial

$$P_{n+1}(u) = (-1)^n \sum_{k=0}^{n-1} \left\langle {n \atop k} \right\rangle u^{k+1}(u - 1)^{n-k}. \tag{8}$$

The first few derivatives (7) (for $u_{max} = 1$, $s = 1$) are as follows

$$u'(t) = u(1 - u) = -u(u - 1) = P_2(u),$$
$$u''(t) = u(u - 1)^2 + u^2(u - 1) = P_3(u),$$
$$u'''(t) = -u(u - 1)^3 - 4u^2(u - 1)^2 - u^3(u - 1) = P_4(u),$$
$$u^{(4)}(t) = u(u - 1)^4 + 11u^2(u - 1)^3 + 11u^3(u - 1)^2 + u^4(u - 1) = P_5(u),$$
$$u^{(5)}(t) = -u(u - 1)^5 - 26u^2(u - 1)^4 - 66u^3(u - 1)^3 - 26u^4(u - 1)^2 - u^5(u - 1) = P_6(u).$$

By direct calculation, one can show that the smallest positive zeros of the polynomials $P_4(u)$, $P_5(u)$, $P_6(u)$ are, respectively, 0.211, 0.0917, 0.0413. From this it follows that if, e.g., t_1 is the smallest root of the equation $u'''(t) = 0$, where $u(t)$ is given by (2), then $u(t_1) = 0.211 u_{max}$. Simultaneously, at $t = t_1$, $u''(t)$ attains its maximum. Similarly, if t_2 and t_3 are the smallest roots of the equations $u^{(4)}(t) = 0$ and $u^{(5)}(t) = 0$, respectively, then $u(t_2) = 0.0917 u_{max}$ and $u(t_3) = 0.0413 u_{max}$.

One can also calculate distances of points t_1, t_2, t_3 from the inflection point t_0. For example

$$u(t_1) = \frac{u_{max}}{1 + e^{-s(t_1 - t_0)}} = 0.211\, u_{max},$$

from which we get

$$t_0 - t_1 = \frac{1.319}{s}. \tag{9}$$

2.2. Wavelets Based on the Second Derivative of the Logistic Function

Let a mother wavelet $\psi_2(x)$ (see Figure 1) be the second derivative of the logistic function $u(x) = \frac{1}{1+e^{-x}}$. Since $u'(x) = -u(u - 1)$, then by (6) or directly we get

$$u''(x) = u(1 - u)(1 - 2u), \tag{10}$$

and by (10) it follows that the wavelet has the following exact form

$$\psi_2(x) = \frac{1}{1 + e^{-x}}\left(1 - \frac{1}{1 + e^{-x}}\right)\left(1 - \frac{2}{1 + e^{-x}}\right) = \frac{e^{-2x} - e^{-x}}{(1 + e^{-x})^3}. \tag{11}$$

By changing the variable $u = \frac{1}{1+\exp(-x)}$, $u'(x) = u(1 - u)$ in the following three integrals we calculate

$$\int_{-\infty}^{\infty} \psi_2(x)dx = \int_0^1 (1 - 2u)du = 0,$$
$$\int_{-\infty}^{\infty} |\psi_2(x)|dx = \int_0^1 |1 - 2u|du = \frac{1}{2},$$
$$\int_{-\infty}^{\infty} (\psi_2(x))^2 dx = \int_0^1 u(1 - u)(1 - 2u)^2 du = \frac{1}{30}, \tag{12}$$

which proves that $\psi_2(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ (the intersection of the space of integrable functions and the space of square integrable functions defined on the set of real numbers $\mathbb{R}$). In fact $\psi_2(x) \in S(\mathbb{R})$ (the space of rapidly decreasing functions on $\mathbb{R}$). We will discuss this in the next section.
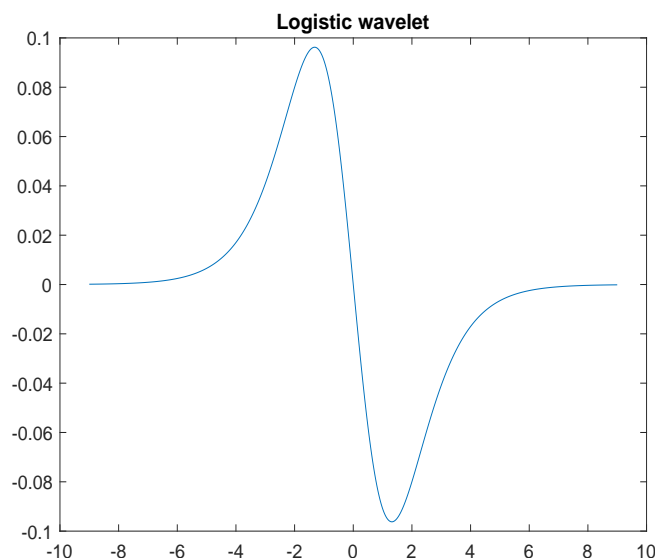

Figure 1. Wavelet $\psi_2(x)$.

The Fourier transform of $\psi_2(x)$ is as follows [24]:

$$\hat{\psi}_2(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_2(x) e^{-i\xi x} dx = \sqrt{\frac{\pi}{2}} \frac{i\xi^2}{\sinh(\pi\xi)}. \tag{13}$$

It is well known (see Daubechies [25]) that a mother wavelet $\psi(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ should satisfy the following admissibility condition

$$2\pi \int_{-\infty}^{\infty} |\xi|^{-1} |\hat{\psi}(\xi)|^2 d\xi < \infty. \tag{14}$$

We will show that for $\psi_2(x)$ the condition (14) is satisfied and even the integral can be expressed in a closed form in terms of the Riemann zeta function. Namely, using (13) and the following formula from Dwight's Tables [26] (item no 860.519):

$$\int_0^{\infty} \frac{x^p}{(\sinh(ax))^2} dx = \frac{\Gamma(p+1)}{2^{p-1}a^{p+1}} \zeta(p), \quad a > 0, \ p > 1, \tag{15}$$

we have

$$2\pi \int_{-\infty}^{\infty} |\xi|^{-1} |\hat{\psi}_2(\xi)|^2 d\xi = \pi^2 \int_{-\infty}^{\infty} \frac{|\xi|^3}{(\sinh(\pi\xi))^2} d\xi = \frac{3\zeta(3)}{\pi^2}. \tag{16}$$

As usual, we generate from ψ_2 a doubly-indexed family of wavelets by dilating and translating,

$$\psi_2^{a,b}(x) = \frac{1}{\sqrt{a}} \psi_2 \left(\frac{x-b}{a} \right),$$

where $a, b \in \mathbb{R}, \ a > 0$.

2.3. Wavelets Based on Higher Derivatives of the Logistic Function

Similarly, as in the previous section we define a wavelet $\psi_n(x)$ to be the nth ($n = 3, 4, \ldots$) derivative of the logistic function $u(x) = \frac{1}{1+\exp(-x)}$. Thus, (6) gives

$$u^{(n)}(x) = (-1)^n \sum_{k=0}^{n-1} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle u^{k+1}(u-1)^{n-k},$$

and then $\psi_n(x)$ can be explicitly expressed as

$$\psi_n(x) = (-1)^n \sum_{k=0}^{n-1} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \left(\frac{1}{1+e^{-x}} \right)^{k+1} \left(\frac{1}{1+e^{-x}} - 1 \right)^{n-k} = \frac{(-1)^n \sum_{k=0}^{n-1} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle (-e^{-x})^{n-k}}{(1+e^{-x})^{n+1}}. \tag{17}$$

By definition, the function $\psi_n(x)$ is an even function for odd n and an odd function when n is even. The numerator of the expression (17) is a polynomial of degree n of the variable e^{-x}, while the denominator of degree $n+1$. Therefore, for any polynomial $p(x)$ we have $\lim_{x \to -\infty} p(x)\psi_n(x) = 0$. Since $\psi_n(x)$ has the symmetry property then also $\lim_{x \to \infty} p(x)\psi_n(x) = 0$. The last conclusion results also from multiplying the numerator and the denominator of (17) by $e^{(n+1)x}$. From this and from the fact that $\psi_{k+1}(x) = \psi_k'(x)$ for any integer $k \geq 2$ it follows that $\psi_n(x) \in S(\mathbb{R})$, ($n = 2, 3, \ldots$).

By (13) we have

$$\hat{\psi}_n(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_n(x) e^{-i\xi x} dx = \sqrt{\frac{\pi}{2}} \frac{(i\xi)^{n-1}\xi}{\sinh(\pi\xi)}. \tag{18}$$

Now, using (18) and once again formula (15) we can calculate the integral of the admissibility condition (14) as follows:

$$2\pi \int_{-\infty}^{\infty} |\xi|^{-1} |\hat{\psi}_n(\xi)|^2 d\xi = \pi^2 \int_{-\infty}^{\infty} \frac{|\xi|^{2n-1}}{(\sinh(\pi\xi))^2} d\xi = \pi^2 \frac{2\Gamma(2n)}{2^{2n-2}\pi^{2n}} \zeta(2n-1)$$

$$= \frac{(2n-1)!}{2^{2n-3}\pi^{2n-2}} \zeta(2n-1). \tag{19}$$

As usually we generate a doubly-indexed family of wavelets from ψ_n by dilating and translating,

$$\psi_n^{a,b}(x) = \frac{1}{\sqrt{a}} \psi_n \left(\frac{x-b}{a} \right),$$

where $a, b \in \mathbb{R}$, $a > 0$, $n = 2, 3, \ldots$.

3. Results

We model the cumulative number of persons reported to be infected by COVID-19 virus, using a sum of logistic functions (multilogistic function) of the form

$$f(x) = \sum_{j=1}^{k} \frac{u_{j,\,max}}{1 + \exp(-\frac{x-b_j}{a_j})}, \tag{20}$$

where k is the number of all considered logistic waves.

Denote by y_n^* the total cumulative number of individuals reported to be infected up to nth day in a country or a region and by y_n the 7-day moving arithmetic average for the sequence y_n^*, i.e.,

$$y_n = \frac{1}{7} \sum_{i=-6}^{0} y_{n+i}^*.$$

We will look, in the sequence (y_n), for points (days n) corresponding to zeros of the second or the third derivative of the logistic function. This is equivalent to detect points, where the sequence of second differences,

$$\Delta^2 y_n = y_{n+1} - 2y_n + y_{n-1},$$

takes a value close to zero or a maximum, respectively.

In order to calculate the continuous wavelet transform (CWT) coefficients for $(\Delta^2 y_n)$, we implemented $\psi_2(x)$ into Matlab's Wavelet Toolbox giving the following definition of the Logistic wavelet:

function [psi,t] = logist(LB,UB,N,∼)
% LOGISTIC Logistic wavelet.
% [PSI,T] = LOGIST(LB,UB,N) returns values of
% the Logistic wavelet on an N point regular
% grid in the interval [LB,UB].
% Output arguments are the wavelet function PSI
% computed on the grid T.
% This wavelet has [-7 7] as effective support.
% See also WAVEINFO.
% Compute values of the Logistic wavelet.
t = linspace(LB,UB,N); % wavelet support.
psi = (exp(- 2* t)-exp(-t))./(1+exp(-t)).^3;
end

Observations show that successive, relatively large, logistic waves arise at fairly long time intervals. Therefore, usually, the second differences corresponding to all previous waves are small as the next wave unfolds. However, the largest wave (with relatively high values of both the saturation level u_{max} and the parameter a), may be an exception. Therefore, at least for some countries, we perform the next wavelet analysis and calculate the CWT coefficients after subtracting the largest wave. Hence, we calculate the second differences and perform their wavelet analysis for the following new time series

$$z_n = y_n - \frac{u_{i,\,max}}{1 + \exp(-\frac{x-b_i}{a_i})},$$

where for $j = i$, the parameters $a_i, b_i, u_{i,\,max}$ correspond to the largest wave in (20). Moreover, if necessary, we also carry out the wavelet analysis for shorter periods of time.

We find the above mentioned points and parameters either directly by observing the sequence of second differences $(\Delta^2 y_n)$ or read them from the scalograms of the CWT transform of Matlab's Wavelet Toolbox by using the Logistic wavelet. From the considerations in Section 2.1 and from (9) it follows that parameter b should be determined as that point where the sequence $(\Delta^2 y_n)$ changes sign. Parameter a should be chosen in such a way that the distance between the zero and the maximum of $(\Delta^2 y_n)$ was approximately $1.319a$ (see (9)). Thus, we obtain two parameters defining a logistic function approximating the time series (y_n). It remains to determine the third parameter of the wave, i.e., its saturation level u. Assuming that (y_n) follows locally a logistic function $y_n \approx y(n) = \dfrac{u_{max}}{1 + \exp(-\frac{n-b}{a})}$

and since by definition it holds

$$y''(x) = \frac{u_{max}}{a^{3/2}} \psi_2^{a,b}(x),$$

then by (12) we get successively

$$\sum_n \Delta^2 y_n \psi_2^{a,b}(n) \approx \sum_n \Delta^2 y(n) \psi_2^{a,b}(n) \approx \int_{-\infty}^{\infty} y''(x)\psi_2^{a,b}(x)dx = \int_{-\infty}^{\infty} \frac{u_{max}}{a^{3/2}} \psi_2^{a,b}(x)\psi_2^{a,b}(x)dx$$
$$= \frac{u_{max}}{a^{3/2}} \int_{-\infty}^{\infty} (\psi_2^{a,b}(x))^2 dx = \frac{u_{max}}{30a^{3/2}}. \tag{21}$$

Using (21) we can estimate the saturation level u_{max} as follows

$$u_{max} \approx 30a^{3/2} \sum_n \Delta^2 y_n \psi_2^{a,b}(n), \tag{22}$$

Parameters a and b in (20) can also be estimated by maximizing locally the integral (i.e., the CWT coefficient) on the left-hand side of (21). Hence, we find in the sequence $\Delta^2 y_n$ the best pattern corresponding to the wavelet $\psi_2^{a,b}$. The saturation level of a logistic wave can also be estimated as twice the value of the sequence (y_n) at the point where $(\Delta^2 y_n)$ changes signs (inflection point) or its maximal value multiplied by $1/0.211$ (zero of the third derivative). This approach is especially important for the last wave of the multilogistic function (20), when it is still in its initial or middle stage and has not yet terminated. After this we use the non-linear generalized reduced gradient method, minimizing the RMSE error, to optimize the parameters, mainly saturation levels. All data were collected from the Our World in Data website [27].

We will use the theory to build models for the total cumulative number of individuals reported to be infected by COVID-19 successively in Germany, Italy, Poland, the United Kingdom, the United States, and in the world.

3.1. Germany

Using the example of Germany, we will show in detail the steps leading to the approximating function $f(x)$ (23). We analyze the data over a period of 496 days from 7 March 2020 ($n = 1$) to 15 July 2021 ($n = 496$). On the scalogram (Figure 2a) showing CWT coefficients, we can distinguish three large logistic waves. We read for them initially (before optimization) the values of parameters a and b, as well as the value of the CWT coefficient (i.e., the integral (21)). Then, we calculate their estimated saturation levels using Formula (22). We have:

1. First wave, $a = 10$, $b = 30$, CWT coeff $= 248$. Thus, we estimate the saturation level as follows:

$$u_{max} = 248 \times 30 \times 10 \times \sqrt{10} = 235,273;$$

2. Second wave, $a = 28$, $b = 280$, CWT coeff $= 519$, $u_{max} =$2,306,883;
3. Third wave, $a = 16$, $b = 413$, CWT coeff $= 557$, $u_{max} =$1,069,440.

Next, we optimize the values of these parameters by minimizing the root mean square error (RMSE) value, which gives:

$$\text{RMSE} = \sqrt{\frac{1}{496} \sum_{n=1}^{496} (y_n - f(n))^2} = 15,336$$

with the following approximating function

$$f(x) = \frac{190,654}{1 + \exp(-\frac{x-35}{11})} + \frac{2,265,636}{1 + \exp(-\frac{x-283}{25})} + \frac{1,308,386}{1 + \exp(-\frac{x-407}{17})}. \tag{23}$$

The approximating function $f(x)$ (23) is shown in Figure 2b. In the case of Germany, no large, newly emerging logistic wave is visible at that time.
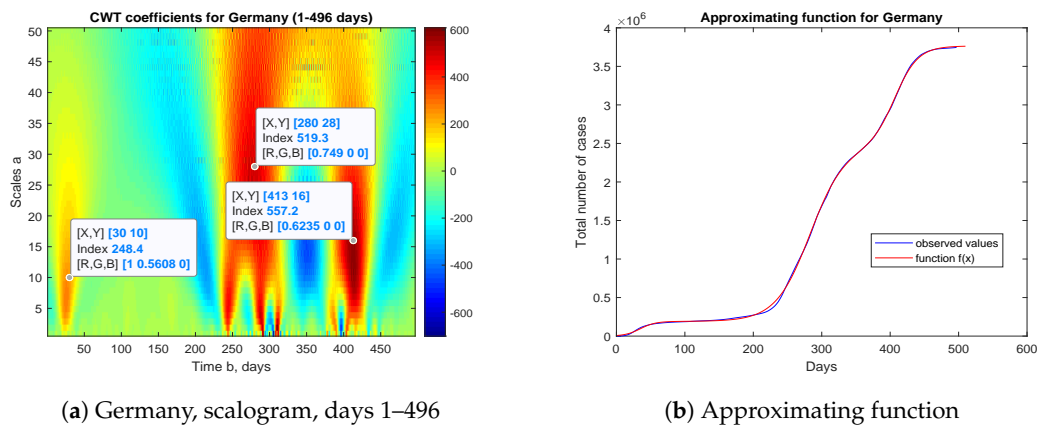
(**a**) Germany, scalogram, days 1–496

(**b**) Approximating function

Figure 2. Scalogram and the approximating multilogistic function (23) for Germany.

3.2. Italy

In the case of Italy, we analyzed a period of 502 days from 1 March 2020 ($n = 1$) to 15 July 2021 ($n = 502$) and obtained the following approximating function

$$f(x) = \frac{253,302}{1 + \exp(-\frac{x-41}{13})} + \frac{1,517,214}{1 + \exp(-\frac{x-259}{12})} + \frac{693,808}{1 + \exp(-\frac{x-311}{15})} + \frac{1,815,925}{1 + \exp(-\frac{x-389}{22})}, \quad (24)$$

with the error RMSE = 12,926. Figure 3 shows scalograms with CWT coefficients and the approximating function (24).
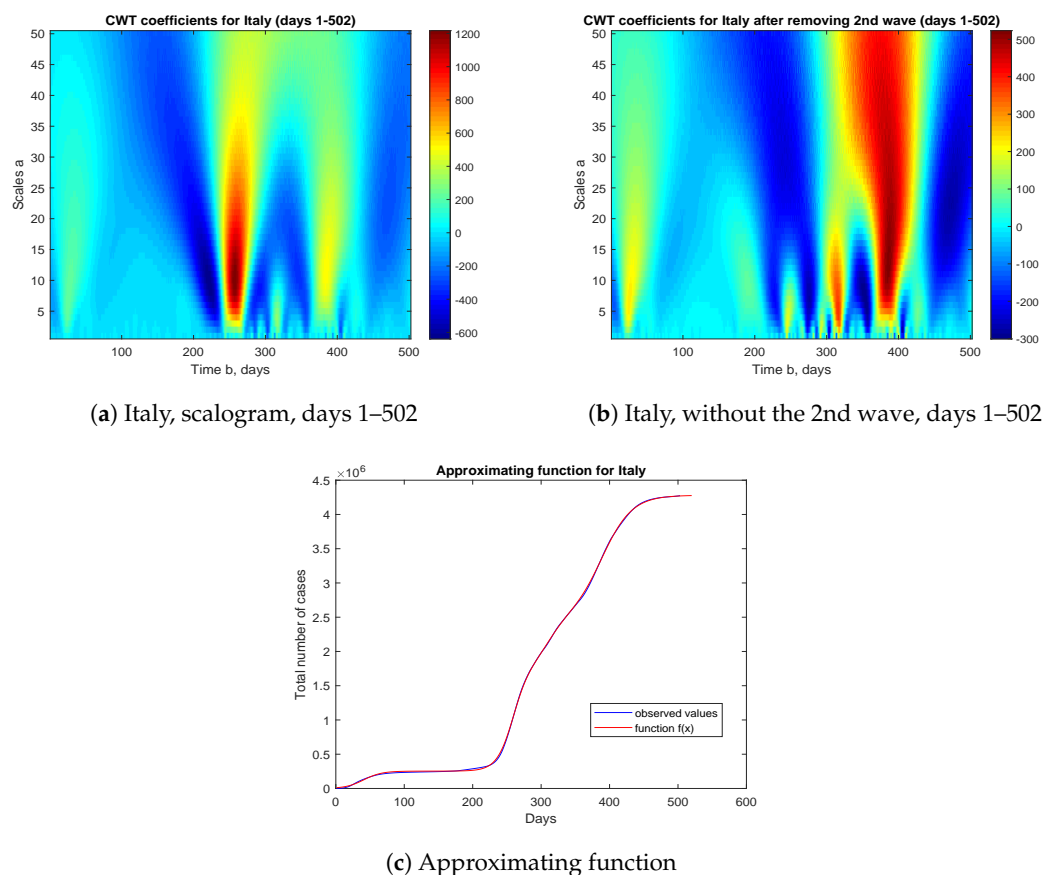

(**a**) Italy, scalogram, days 1–502

(**b**) Italy, without the 2nd wave, days 1–502

(**c**) Approximating function

Figure 3. Scalograms and the approximating multilogistic function (24) for Italy.

3.3. Poland

We analyzed data over a period of 488 days from 15 March 2020 ($n = 1$) to 15 July 2021 ($n = 488$). The multilogistic approximating function has the following form

$$f(x) = \frac{27{,}659}{1 + \exp(-\frac{x-29}{7})} + \frac{30{,}235}{1 + \exp(-\frac{x-160}{15})} + \frac{1{,}118{,}196}{1 + \exp(-\frac{x-247}{13})} + \frac{237{,}375}{1 + \exp(-\frac{x-301}{7})}$$
$$+ \frac{1{,}398{,}841}{1 + \exp(-\frac{x-378}{14})}, \tag{25}$$

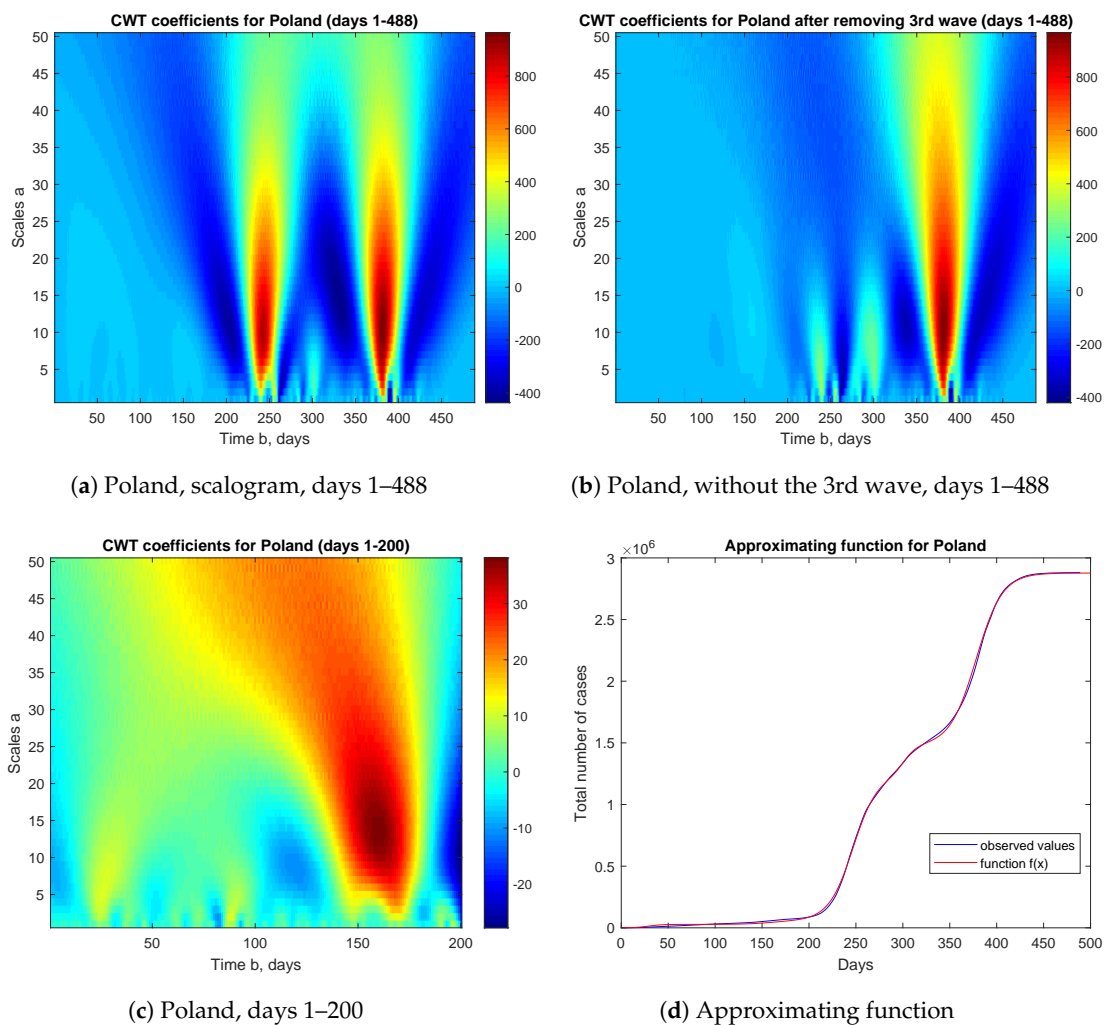with the error RMSE = 13,433. Figure 4 shows scalograms with CWT coefficients and the approximating function (25).

(**a**) Poland, scalogram, days 1–488

(**b**) Poland, without the 3rd wave, days 1–488

(**c**) Poland, days 1–200

(**d**) Approximating function

Figure 4. Scalograms and the approximating multilogistic function (25) for Poland.

3.4. The United Kingdom

For the UK we analyze the period of 491 days from 12 March 2020 ($n = 1$) to 15 July 2021 ($n = 491$). Figure 5 shows scalograms with CWT coefficients.
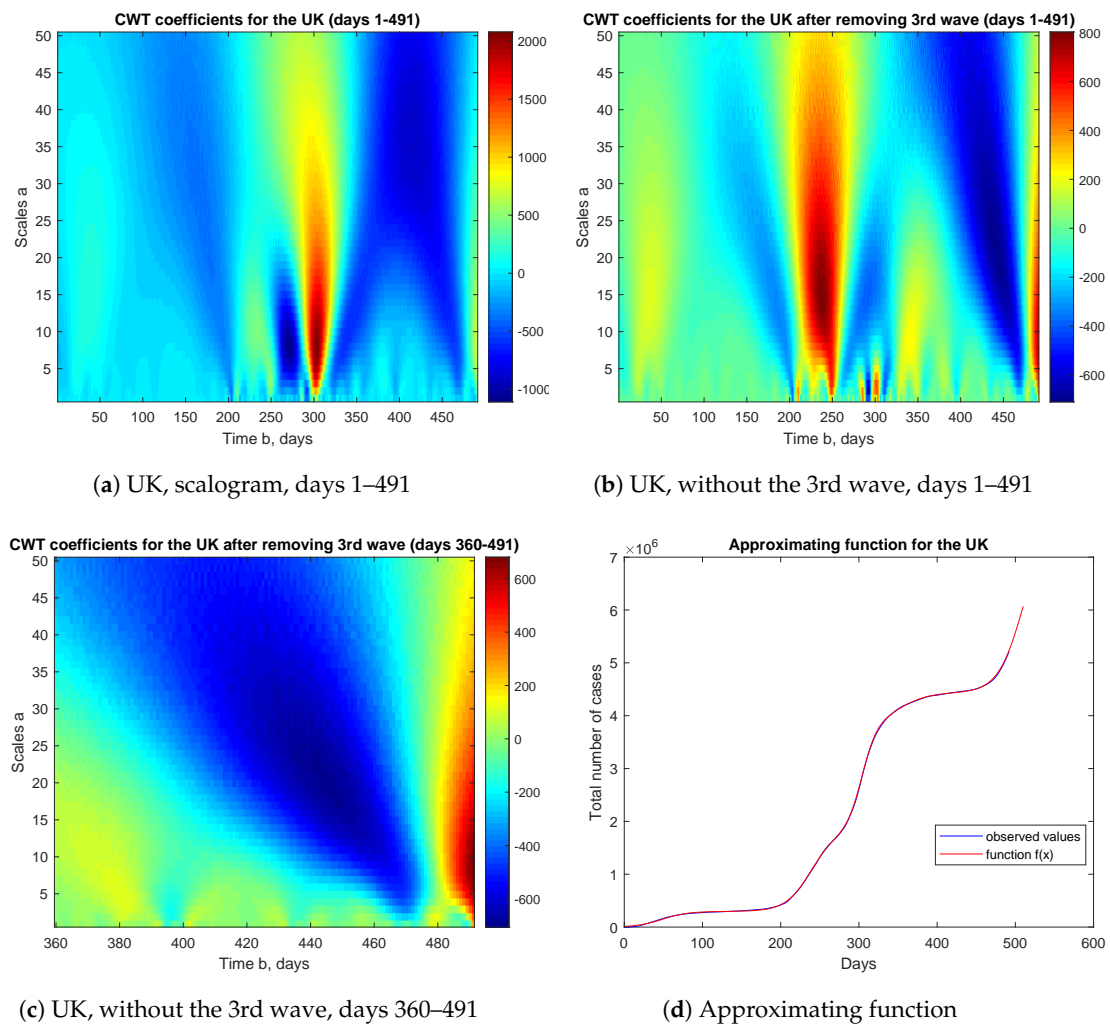
(**a**) UK, scalogram, days 1–491

(**b**) UK, without the 3rd wave, days 1–491

(**c**) UK, without the 3rd wave, days 360–491

(**d**) Approximating function

Figure 5. Scalograms and the approximating multilogistic function (26) for the UK.

In this case, unlike in previous countries, a new, developing logistic wave is visible on the right in the Figure 5b,c. On July 16 (day 492), this wave has not yet reached its inflection point, as shown by the second differences, which are all positive (Table 2).

Table 2. A part of the working table for the UK.

Day, n	y_n	$\Delta^2 y_n$
$\vdots$	$\vdots$	$\vdots$
482	4,896,238	913
483	4,923,476	636
484	4,951,350	1288
485	4,980,512	1054
486	5,010,728	1069
487	5,042,013	1100
488	5,074,398	1125
489	5,107,908	1446
490	5,142,864	2363
491	5,180,183	2295
492	5,219,797	

With the CWT analysis alone, it would be difficult to determine the parameters of this wave. However, in order to do it we can use the results of Section 2.1. Assuming that the second differences reach their maximum on July 14 (day 490), i.e., it is the zero ($t_1 = 490$) of the third derivative of the last wave, we can estimate its saturation level. Subtracting from $y_{490} = 5,142,864$ the saturation levels of all previous four waves, we get about 700,000 cases. Thus, the saturation level of the last wave can be estimated as $u_{max} = 700,000/0.211 \approx 3,300,000$. Moreover, assuming that the parameter a is similar to that which was for the other waves, say $a = 16$, by (9) we have $b = t_0 = t_1 + 1.319 \times 16 \approx 511$. After the optimization we get the following approximating function

$$f(x) = \frac{296,755}{1 + \exp(-\frac{x-49}{16})} + \frac{1,492,663}{1 + \exp(-\frac{x-238}{16})} + \frac{212,6051}{1 + \exp(-\frac{x-304}{10})} + \frac{519,365}{1 + \exp(-\frac{x-355}{18})} + \frac{3,267,991}{1 + \exp(-\frac{x-510}{16})}, \tag{26}$$

with the error RMSE = 12,174. Figure 5 shows also the approximating function (26).

In Table 3 (Added in proof), we compare the 10-day forecast with actual data. The last two columns show the forecast error, both absolute and relative. It can be seen that the function $f(n)$ predicts the trend quite well for a short period. However, it seems more important that the discussed method can be used for early warning of the appearance of new waves. This is in line with Hu et al. [28] comments on the use of wavelet analysis to study the development of infectious diseases.

Table 3. Forecasting for the UK.

Day, n	Date	y_n	$f(n)$	$|y_n - f(n)|$	$|y_n - f(n)|/y_n$
493	2021-07-17	5,262,604	5,273,915	11,311	0.00214
494	2021-07-18	5,307,763	5,313,502	5739	0.00108
495	2021-07-19	5,353,760	5,354,247	487	0.00009
496	2021-07-20	5,401,212	5,396,120	5092	0.00094
497	2021-07-21	5,448,935	5,439,082	9853	0.00180
498	2021-07-22	5,495,427	5,483,091	12,336	0.00224
499	2021-07-23	5,539,655	5,528,095	11,560	0.00208
500	2021-07-24	5,580,660	5,574,038	6622	0.00118
501	2021-07-25	5,618,963	5,620,858	1895	0.00033
502	2021-07-26	5,655,018	5,668,486	13,468	0.00238

3.5. The United States

In the case of the US, we analyzed the data over a period of 490 days from 13 March 2020 ($n = 1$) to 15 July 2021 ($n = 490$). The multilogistic approximating function has the following form

$$f(x) = \frac{1,489,667}{1 + \exp(-\frac{x-42}{11})} + \frac{4,777,642}{1 + \exp(-\frac{x-135}{19})} + \frac{22,527,836}{1 + \exp(-\frac{x-284}{26})} + \frac{996,663}{1 + \exp(-\frac{x-311}{4})} + \frac{3,648,578}{1 + \exp(-\frac{x-400}{16})} + \frac{1,228,591}{1 + \exp(-\frac{x-499}{17})}, \tag{27}$$

with the error RMSE = 59,106.

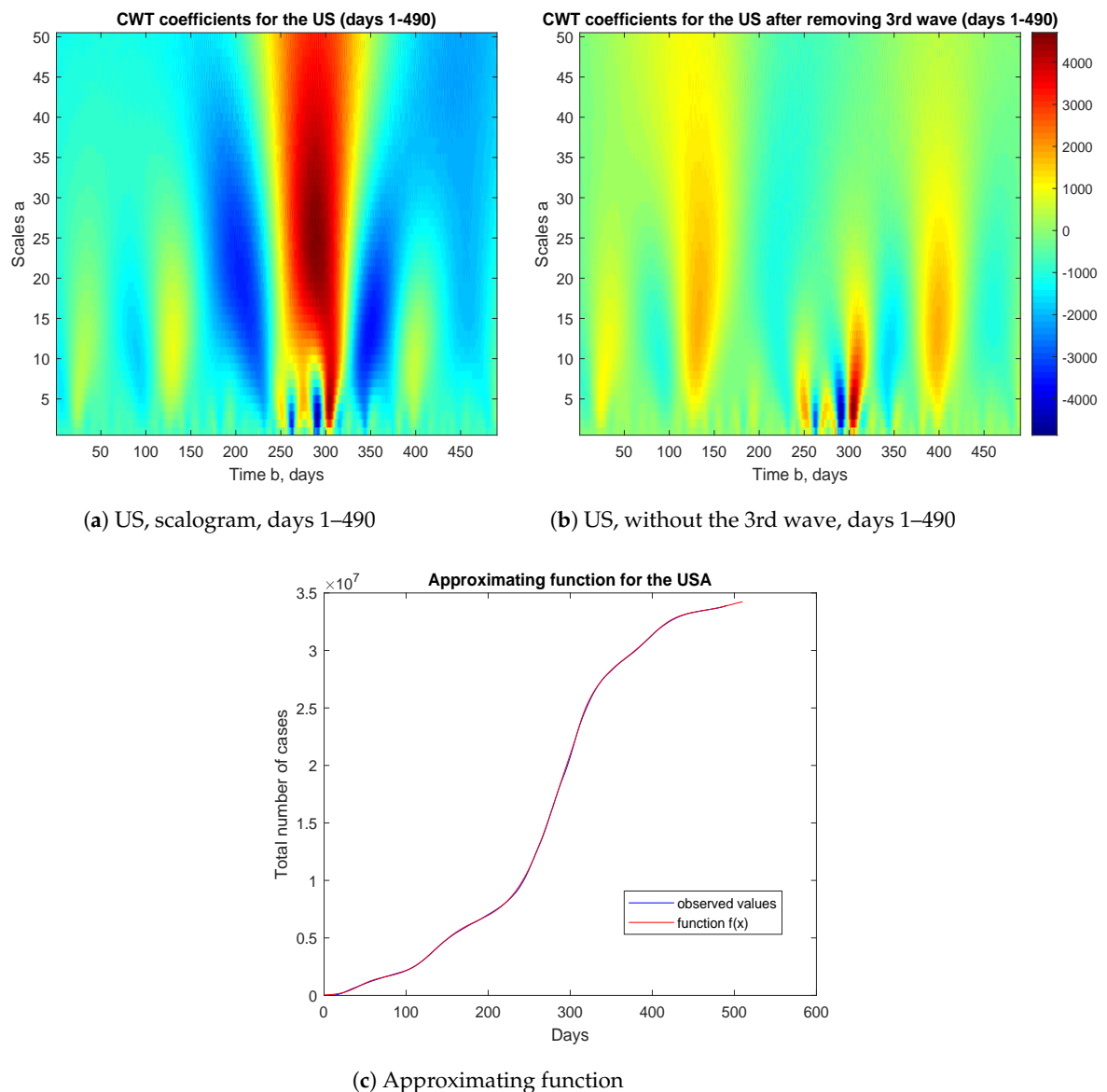The scalograms with CWT coefficients and the approximating function (27) are shown in Figure 6.

(**a**) US, scalogram, days 1–490

(**b**) US, without the 3rd wave, days 1–490

(**c**) Approximating function

Figure 6. Scalograms and approximating multilogistic function (27) for the US.

3.6. The World

In the case of the whole world, we analyzed the data over a period of 530 days from 2 February 2020 ($n = 1$) to 15 July 2021 ($n = 530$). The multilogistic approximating function has the following form

$$f(x) = \frac{2,951,000}{1 + \exp(-\frac{x-83}{12})} + \frac{30,199,966}{1 + \exp(-\frac{x-188}{31})} + \frac{42,251,210}{1 + \exp(-\frac{x-293}{23})} + \frac{37,503,735}{1 + \exp(-\frac{x-348}{19})}$$
$$+ \frac{67,293,529}{1 + \exp(-\frac{x-451}{21})} + \frac{27,772,754}{1 + \exp(-\frac{x-541}{15})}, \tag{28}$$

with the error RMSE = 248,525.

The scalograms with CWT coefficients and the approximating function (28) are shown in Figure 7.

(**a**) The world, scalogram, days 1–530

(**b**) The world, without 3rd and 5th waves
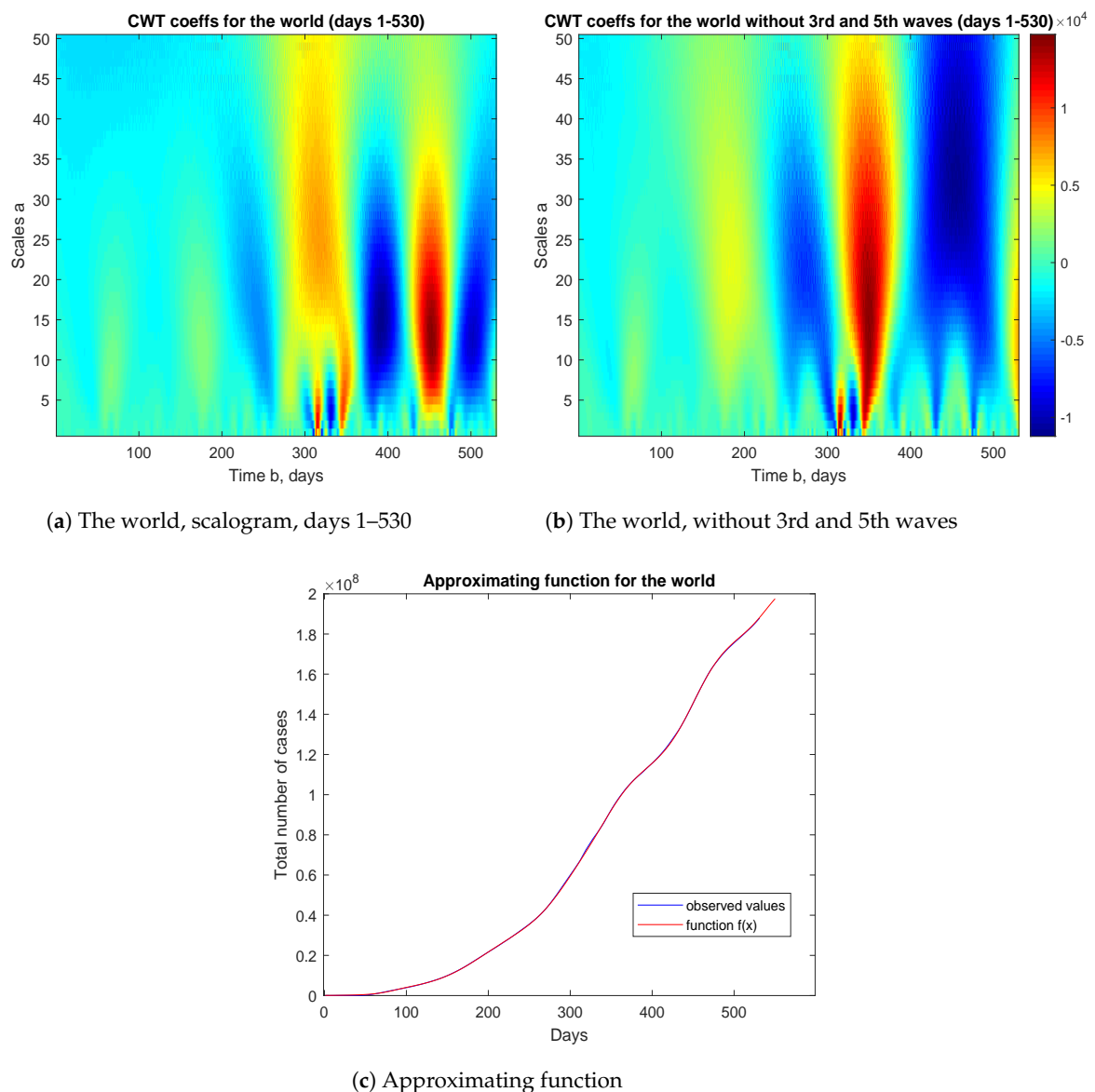

(**c**) Approximating function

Figure 7. Scalograms and the approximating multilogistic function for the world.

4. Discussion

The article deals with the application of the logistic function to the description of phenomena that follow many overlapping sigmoidal trends. We used some properties of the zeros of successive derivatives of the logistic function, in particular their relation to the saturation level.

By using the Eulerian numbers, we defined logistic wavelets of any order and examined their properties, checking for them the general admissibility condition for wavelets. We have added a second order logistic wavelet to the Matlab's Wavelet Toolbox. Then we performed a wavelet analysis of the time series, whose terms are the second differences of the smoothed total number of individuals infected with the COVID-19 virus in several countries. As a result, we obtained the continuous wavelet transform (CWT) scalograms, from which we could read the distribution of successive logistic waves, and their parameters. Note that the three-dimensional CWT scalograms allow the simultaneous identification of all three parameters of consecutive logistic waves. This has been described in detail using the example of Germany Section 3.1. Other approximation methods do not provide such possibilities. We then optimized the parameters (mainly the saturation levels) minimizing the RMSE error. We have shown in the examples that the multilogistic

function, obtained in this way, well approximates the total number of infections. The theory and procedure can be applied to model the total number of infections in any country or a region.

We limited ourselves to identifying the largest logistic waves visible on the scalograms. There are also visible smaller waves, which we have omitted, because they do not contribute much to the explanation of the phenomenon. These could be considered in addition and then the RMSE error would probably be smaller.

The CWT analysis described above makes it possible to estimate the parameters of the logistic waves that have already terminated. In the case of an ongoing, new wave, what we have seen on the example of the UK Section 3.4, we should determine its level of development. In practice, it can be done by determining whether the logistic wave has already reached the inflection point (zero of the second derivative, 50% of the saturation level), which corresponds to a change in the sign in the series of second differences ($\Delta^2 y_n$). If the inflection point has not yet been reached, we can try to determine whether the wave has reached the previously located zero of the third derivative. This corresponds to a maximum in a series of second differences (this point is about 21% of the saturation level).

It may be considered in a further work, whether the CWT analysis, based on the use of only the positive part of the second-order logistic wavelet, would be helpful in determining the parameters of ongoing waves.

In our further work, we intend to use, in a similar way, the logistic wavelets of higher order (see Section 2.3). Moreover, using appropriate special numbers we are going to define analogous wavelets for the Gompertz function (see some initial calculations [29,30]) or for some generalizations of the logistic function (for preliminary theorems see [31]).

Author Contributions: Conceptualization, G.R. and G.F.; methodology, G.R.; software, G.R.; validation, G.R. and G.F.; formal analysis, G.R.; investigation, G.R. and G.F.; resources, G.R. and G.F.; data curation, G.R.; writing—original draft preparation, G.R. and G.F.; writing—review and editing, G.R.; visualization, G.R.; supervision, G.R. and G.F.; project administration, G.R.; funding acquisition, G.R. 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: The source data we used in the article are publicly available on the website Our World in Data, https://ourworldindata.org/coronavirus-source-data (accessed on 17 July 2021). The Our World in Data website relies on data from the Johns Hopkins University. The data processed by us can be available upon request.

Conflicts of Interest: The authors declare that they do not have any conflict of interest in their submitted manuscript.

References

1. Kartono, A.; Wahyudi, S.T.; Setiawan, A.A.; Sofian, I. Predicting of the coronavirus disease 2019 (COVID-19) epidemic using estimation of parameters in the logistic growth model. *Infect. Dis. Rep.* **2021**, *13*, 465–485. [CrossRef]
2. Pelinovsky, E.; Kurkin, A.; Kurkina, O.; Kokoulina, M.; Epifanova, A. Logistic equation and COVID-19. *Chaos Solitons Fractals* **2020**, *140*, 110241. [CrossRef]
3. Fokas, A.S.; Dikaios, N.; Kastis, G.A. Mathematical models and deep learning for predicting the number of individuals reported to be infected with SARS-CoV-2. *J. R. Soc. Interface* **2020**, *17*, 20200494. [CrossRef] [PubMed]
4. Grenfell, B.T.; Bjørnstad, O.N.; Kappey, J. Travelling waves and spatial hierarchies in measles epidemics. *Nature* **2001**, *414*, 716–723. [CrossRef] [PubMed]
5. Cazelles, B.; Cazelles, K.; Chavez, M. Wavelet analysis in ecology and epidemiology: Impact of statistical tests. *J. R. Soc. Interface* **2014**, *11*, 20130585. [CrossRef]

6. Krantz, S.G.; Polyakov, P.; Rao, A.S.R.S. True epidemic growth construction through harmonic analysis. *J. Theor. Biol.* **2020**, *494*, 110243. [CrossRef]
7. José, M.V.; Bishop, R.F. Scaling properties and symmetrical patterns in the epidemiology of rotavirus infection. *Philos. Trans. R. Soc. Lond.* **2003**, *358*, 1625–1641. [CrossRef] [PubMed]
8. Wang, Y.; Xu, C.; Wang, Z.; Zhang, S.; Zhu, Y.; Yuan, J. Time series modeling of pertussis incidence in China from 2004 to 2018 with a novel wavelet based SARIMA-NAR hybrid model. *PLoS ONE* **2018**, *13*, e0208404. [CrossRef]
9. Santos, C.A.G.; Guerra-Gomes, I.C.; Gois, B.M.; Peixoto, R.F.; Keesen, T.S.L.; da Silva, R.M. Correlation of dengue incidence and rainfall occurrence using wavelet transform for João Pessoa city. *Sci. Total Environ.* **2019**, *647*, 794–805. [CrossRef]
10. Zhang, Q.; Chen, Y.; Fu, Y.; Liu, T.; Zhang, Q.; Guo, P.; Ma, W. Epidemiology of dengue and the effect of seasonal climate variation on its dynamics: A spatio-temporal descriptive analysis in the Chao-Shan area on China's southeastern coast. *BMJ Open* **2019**, *9*, e024197. [CrossRef]
11. Liu, W.M. Applications of wavelet transform in forecasting incidence of infectious diseases. In Proceedings of the 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Baltimore, MD, USA, 3–6 November 1994; Volume 1, pp. A30–A31. [CrossRef]
12. Lavrova, A.I.; Postnikov, E.B.; Manicheva, O.A.; Vishnevsky, B.I. Bi-logistic model for disease dynamics caused by Mycobacterium tuberculosis in Russia. *R. Soc. Open Sci.* **2017**, *4*, 171033. [CrossRef]
13. Vanucci, E.; Vanucci, L. Forecast Covid-19 End Date in Italy by Logistics Waves. Available online: https://www.researchgate.net/publication/341104205 (accessed on 10 July 2021).
14. Tsoularis, A. Analysis of logistic growth models. *Res. Lett. Inf. Math. Sci.* **2001**, *2*, 23–46. [CrossRef]
15. Tsoularis, A.; Wallace, J. Analysis of logistic growth models. *Math. Biosci.* **2002**, *179*, 21–55. [CrossRef]
16. Kowalski, K.; Steeb, W.H. *Nonlinear Dynamical Systems and Carleman Linearization*; World Scientific Publishing Co. Pte. Ltd.: Singapore; London, UK, 1991.
17. Dominici, D. Nested derivatives: A simple method for computing series expansions of inverse functions. *Int. J. Math. Math. Sci.* **2003**, *58*, 3699–3715. [CrossRef]
18. Rządkowski, G.; Głażewska, I.; Sawińska, K. Logistic function as a tool of planning. *Found. Manag.* **2014**, *6*, 57–70. [CrossRef]
19. Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics: A Foundation for Computer Science*; Addison Wesley: Reading, MA, USA, 1994.
20. Rządkowski, G. Eulerian numbers and Riccati's differential equation. In Proceedings of the ICNAAM 2006, Hersonnisos, Greece, 15–19 September 2006; Simos, T.E., Ed.; Wiley–VCH Verlag: Berlin, Germany, 2006; pp. 291–294. ISBN 3-527-40743-X.
21. Rządkowski, G. Derivatives and Eulerian numbers. *Am. Math. Mon.* **2008**, *115*, 458–460. [CrossRef]
22. Franssens, G.R. Functions with derivatives given by polynomials in the function itself or a related function. *Anal. Math.* **2007**, *33*, 17–36. [CrossRef]
23. Rządkowski, G. On a family of polynomials. *Math. Gaz.* **2006**, *518*, 283–286. [CrossRef]
24. Bilge, A.H.; Ozdemir, Y. The Fourier transform of the first derivative of the generalized logistic growth curve. *arXiv* **2015**, arXiv:1502.07182.
25. Daubechies, I. *Ten Lectures on Wavelets*, 2nd ed.; CBMS-NSF Regional Conference Series in Applied Mathematics 61; SIAM: Philadelphia, PA, USA, 1992.
26. Dwight, H.B. *Tables of Integrals and Other Mathematical Data*, 4th ed.; The Macmillan Company: New York, NY, USA, 1961.
27. Our World in Data. Available online: https://ourworldindata.org/coronavirus-source-data (accessed on 17 July 2021).
28. Hu, S.; Cheng, X.; Chen, D. Comparative study on early warning methods of infectious diseases. *E3S Web Conf.* **2021**, *251*, 03084. [CrossRef]
29. Rządkowski, G.; Rządkowski, W.; Wójcicki, P. On some connections between the Gompertz function and special numbers. *J. Nonlinear Math. Phys.* **2015**, *3*, 374–380. [CrossRef]
30. Rządkowski, G.; Głażewska, I.; Sawińska, K. The Gompertz function and its applications in management. *Found. Manag.* **2015**, *7*, 185–190. [CrossRef]
31. Rządkowski, G.; Urlińska, M. Some applications of the generalized Eulerian numbers. *J. Comb. Theory Ser. A* **2019**, *163*, 85–97. [CrossRef]