Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Interval nonlinear initial-valued problem using constraint intervals: Theory and an application to the Sars-Cov-2 outbreak

M.S. Ceconello a,⁎, M.T. Mizukoshi b, W. Lodwick c

a Department of Mathematics, Federal University of Mato Grosso, 78075-202 Cuiabá, MT, Brazil
b Institute of Mathematics and Statistic, University Federal of Goiás, 74690-970 Goiânia, GO, Brazil
c Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO, USA

ARTICLE INFO

Article history:
Received 23 January 2021
Received in revised form 7 July 2021
Accepted 12 August 2021
Available online 16 August 2021

Keywords:
SIR model
Constraint interval
Sars-Cov-2 pandemic
Nonlinear regression

ABSTRACT

This article discusses the theory of constraint interval solutions to interval nonlinear initial value problems and applies the notion of constraint interval solutions to analyze the asymptotic behavior of a susceptible-infected-recovered (SIR) epidemiological nonlinear differential equation model, specifically the covid pandemic, in the presence of interval uncertainty to illustrate the efficacy of this approach. Furthermore, constraint interval solutions are used to estimate the intervals for the parameters by fitting solutions to the Brazilian’s Sars-Cov-2 pandemic official data. Simulations and graphical solutions incorporating constraint interval uncertainties are presented to help in the visualization of the pandemic’s behavior.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

Real-valued mathematical analysis of processes and systems that incorporate uncertainty in its theory has a history that dates back to at least Archimedes ([1]) where the perimeter of a circle was approximated via outer and inner polygonal perimeters. Quantitative uncertainties or inaccuracies occur due to data measurement errors, lack of complete information, assumption of physical models, variations of the system, and computational errors. Uncertainty is often represented by distributions ([2]). However, a simpler representation of uncertainty is the interval where only the bounds of the uncertainty are known, but not the distribution within the bounds. Intervals naturally arise in measurement errors and in bounding roundoff error. This study limits itself to quantitative uncertainties represented by intervals.

The context in which these uncertainties will be studied is the nonlinear differential equation initial value problems. That is, of interest to this study are differential equations with interval uncertainties in the initial condition or/and coefficients. For intervals, there are two types of interval arithmetic analyses that can be considered: (1) Standard arithmetic, the War-mus, Sunaga, Moore arithmetic (WSMA) ([3]) where every interval is considered as independent; (2) Constraint interval arithmetic (CIA) ([4]), where it is possible to consider total or partial independence, or total dependence. This study limits itself to the constraint interval approach.

It will be seen that the interval solutions obtained via CIA, whether dependent or independent, can result in solutions that may not have the flow property. This can happen because to obtain an interval solution requires a global optimization problem, which, in turn, means that a solution can be found in the resultant interval that was not in the deterministic case. Even

⁎ Corresponding author.

E-mail addresses: moiseis@ufmt.br (M.S. Ceconello), tuyako@ufg.br (M.T. Mizukoshi), Weldon.Lodwick@ucdenver.edu (W. Lodwick).

https://doi.org/10.1016/j.ins.2021.08.045
0020-0255/© 2021 Elsevier Inc. All rights reserved.
in this case, the stability of the deterministic solution, which, in terms of an interval solution, is a degenerate fixed point (the left endpoint is equal to the right endpoint), can be analyzed considering the interval solution over compact sets in $\mathbb{R}^n$.

There is another constraint interval analysis, a subset of CIA, of interval differential equations using total dependence called single level arithmetic ([5,6]). However, this line of analysis will not be pursued here.

Epidemiological mathematical modeling has been used to assist the decision-making processes and many research articles have been published with the aim of analyzing the dynamics of the epidemic in the population ([7–9]). In Brazil, which is used in the development of the interval SIR model, there are many states with diverse social-economic situations. This means that the information obtained that is used to determine the parameters of the model (infection rates, recovered rates, for example) have varied precision. Even the tests given to determine how many people in the population are infected at any period of time are not deterministic. Thus, the application of uncertainty in the SIR nonlinear differential equation model is not only necessary, it is useful and timely. The parameters for the case of the epidemic in Brazil are obtained via a constrained interval regression procedure on Brazil’s Ministry of Health Sars-Cov-2 data.

This work is organized as the following: in Section 2 we present the basic concepts on the CI theory. In Section 3 we define the concepts of interval solution of initial value problems and asymptotic behavior of such solutions. In Section 4 analyze the SIR’s model interval solution. In Section 5 we estimate the parameters of SIR’s model interval solution using the Sars-Cov-2 data of Brazil.

2. Basic concepts on Constraint Interval theory

This section describes the main concepts on the constraint interval (CI) theory. $\mathbb{R}$ will denote the set of all closed intervals on $\mathbb{R}$ and $\mathcal{P}$ the space of linear functions $p(\gamma) = a + by$ over the interval $[0,1]$ with non-negative slope ($b \geq 0$).

The main feature of the CI theory consists in identifying a given interval $x = [x_1, x_2] \in \mathbb{R}$ with a polynomial $p_x \in \mathcal{P}$ by means of the mapping $I : \mathbb{R} \rightarrow \mathcal{P}$ defined as

$$I(x) = p_x(\gamma) = \gamma_0 + \sum_{i=1}^{n} \gamma_i (x_i - x_0), \quad \gamma_i \in [0,1].$$

It is easy to check that $I(x)$ is a function so that each function $p(\gamma) = a + by$ in $\mathcal{P}$ defines a unique interval $x = I^{-1}(p) = [a, a + b]$ in $\mathbb{R}$. Denotes by $\mathbb{I}^n$ the set of all closed hypercubes on $\mathbb{R}$. In this way, $x = (x_1, x_2, \ldots, x_n) \in \mathbb{I}^n$ means that each interval $[x_i]$ is identified with a polynomial $p_i(\gamma_i) = \gamma_i x_i$, where $w_i = x_i - \underline{x}_i$ and $\gamma_i \in [0,1]$. For $x \in \mathbb{I}^n$ and a vector $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ the notation $a \in [x]$ means that there are scalars $\gamma_i \in [0,1]$ such that $a_i = \gamma_i w_i = \underline{x}_i + \gamma_i (x_i - \underline{x}_i)$.

This mapping approach allows for the concept of extension of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ to be defined, (see [4]). Next, $\mathbb{P}_n$ is the cartesian product of the representation $\mathcal{P}$. That is, $\mathbb{P}_n = \mathbb{P}_{n-1} \times \mathcal{P}$, for $n \geq 2$.

DEFINITION 1 (CI extension). Consider a function $f : U \rightarrow \mathbb{R}, U \subset \mathbb{R}^n$ and intervals $[x_i] \in \mathbb{P}$ such that $x = ([x_1], [x_2], \ldots, [x_n]) \subset \mathbb{P} \cap \mathbb{I}^n$. The CI extension of $f$ is the function $\hat{f} : \mathbb{I}^n \rightarrow \mathbb{R}$ such that $\hat{f}([x])$ is computed by the following procedure:

a) Mapping to representation space: First, the vector $x \in \mathbb{P}$ is embedded into $\mathbb{P}_n$ by means of the map $I_n : \mathbb{I}^n \rightarrow \mathbb{P}_n$ defined as

$$I_n(x) = P_n(\gamma) = (p_1(\gamma_1), p_2(\gamma_2), \ldots, p_n(\gamma_n))$$

$$= (x_1 + \gamma_1 w_1, x_2 + \gamma_1 w_2, \ldots, x_n + \gamma_n w_n),$$

where $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ and $w_i = x_i - \underline{x}_i$.

b) Evaluating the function: Next, define $g : \mathbb{I}^n \rightarrow \mathcal{F}$ by computing

$$g(\gamma) = (f \circ I_n)([x]) = f(x_1 + \gamma_1 w_1, x_2 + \gamma_1 w_2, \ldots, x_n + \gamma_n w_n).$$

$\mathcal{F}$ is a space of multivariate functions.

c) Mapping back: Finally, if it is desired, define $\hat{f}([x]) \in \mathbb{R}$ by computing

$$\hat{f}([x]) = \left[ \min_{\gamma} g(\gamma), \max_{\gamma} g(\gamma) \right],$$

where $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ and $w_i = x_i - \underline{x}_i, \gamma_i \in [0,1]$.

Given a function $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, the CI extension $\hat{f}$ is defined taking the CI extension on each one of the coordinates $f_i$ of $f$. That is, $\hat{f}$ at $[x] \in \mathbb{P}^n$ is the vector of interval $\hat{f}([x]) = (\hat{f}_1([x]), \hat{f}_2([x]), \ldots, \hat{f}_m([x]))$. It is noted and not difficult to check that $\hat{f}([x])$ is not always equal to $f([x])$. For any $[x] \in \mathbb{P}^n$ it turns out that $\hat{f}([x])$ is a hypercube on $\mathbb{R}$ whereas for $f([x])$ it is not always a hypercube. Moreover, it is always true that $\hat{f}([x]) \subset \hat{f}([x])$.
The approach considered in Definition 1 allows \( \mathbb{R} \) to be endowed with the basic arithmetic operations, say, addition, subtraction, multiplication and division, by considering the following functions, respectively: 
\[
f_{(+)}(x,y) = x + y, f_{(-)}(x,y) = x - y, f_{(.)}(x,y) = xy \text{ and } f_{(-)}(x,y) = x/y.
\]

Furthermore, \( \mathbb{R}^n \) can be also turned into a metric space by considering the Pompieu-Hausdorff distance for compact sets given by:
\[
d_{\text{PH}}([a],[b]) = \max \left\{ \sup_{x \in [a]} \|x - b\|, \sup_{y \in [b]} \|a - y\| \right\},
\]
for all \( [a],[b] \in \mathbb{R}^n \) ([10]).

Now, consider \( A \) and \( B \) subsets of \( \mathbb{R}^n \), the projection \( \pi_i : \mathbb{R}^n \rightarrow \mathbb{R} \) is such that \( \pi_i(x_1,x_2, \cdots, x_n) = x_i \). Thus \( \pi_i(A) \) and \( \pi_i(B) \) are unions of points of the metric space \( \mathbb{R}^n \) and, since \( |x_i - y_i| \leq \|x - y\| \) for all \( x \in A \) and \( y \in B \), then
\[
\sup_{x \in A, y \in B} |x_i - y_i| \leq \sup_{x \in A} \|x - y\|
\text{ and } \sup_{x \in A} |x_i - y_i| \leq \sup_{x \in B} \|x - y\|.
\]

Therefore, it follows that:
\[
d_H(\pi_i(A), \pi_i(B)) \leq d_H(A, B),
\]
where \( d_H \) denotes the Pompieu-Hausdorff distance.

Next result is useful in what follows.

**Theorem 1.** Consider continuous functions \( f : [0, \infty) \times U \rightarrow \mathbb{R}^m \) and \( g : U \rightarrow \mathbb{R}^m \) and \( |x| \in \mathbb{R}^n \), \( x \subset U \). If \( f(t,x) \) converges uniformly to \( g(x) \) on \( \mathbb{R}^m \) when \( t \to \infty \) then \( f(t, |x|) \) converges to \( g(|x|) \) on \( \mathbb{R}^m \) when \( t \to \infty \).

**Proof.** Since \( |x| \in \mathbb{R}^n \), the continuity of \( f \) and \( g \) ensures that \( A_t = f(t, |x|) \) and \( B = g(|x|) \) are in \( \mathbb{R}^m \). We must show that, for a given \( \varepsilon \) there is a \( T > 0 \) such that
\[
d_{\text{PH}}(A_t, B) = \max \left\{ \sup_{b \in |x|} \|f(t,b) - g(b)\|, \sup_{a \in |x|} \|f(t,a) - g(a)\| \right\} < \varepsilon,
\]
for all \( t > T \). Now,
\[
\sup_{b \in |x|} \|f(t,b) - g(b)\| \leq \sup_{b \in |x|} \|f(t,b) - g(b)\|
\]
\[
\sup_{a \in |x|} \|f(t,a) - g(a)\| \leq \sup_{a \in |x|} \|f(t,a) - g(a)\|.
\]

By the uniform convergence hypothesis, given \( \varepsilon > 0 \) there exist \( T > 0 \) such that \( \|f(t,x) - g(x)\| < \varepsilon \) for all \( t > T \) and \( x \in U \).

Therefore, for these conditions, it turns out that
\[
d_{\text{PH}}\left(f(t, |x|), g(|x|)\right) \leq \max \left\{ \sup_{b \in |x|} \|f(t,b) - g(b)\|, \sup_{a \in |x|} \|f(t,a) - g(a)\| \right\} \leq \varepsilon
\]
for all \( t > T \), proving the statement.  \( \square \)

### 3. Interval solution of initial value problems

The main goal in this section is to consider constraint interval uncertainties on initial value problems as well as to analyze the impact of these uncertainties in the asymptotic behavior of the solution. To this end, we are going to consider two cases: i) - uncertainties only on initial conditions; ii) - uncertainties on both initial conditions and parameters.

#### 3.1. Constraint interval uncertainties on initial conditions

Consider an initial value problem
\[
\frac{dx}{dt} = f(x), \quad x(0) = x_0,
\]
for all \( t > 0 \), \( U \subset \mathbb{R}^n \) where \( x : U \rightarrow U \) denotes the (deterministic) flow of the initial value problem defined by Eq. (4).
Now, consider an initial condition $x_0$ subject to generalized uncertainties. The CI extension of the flow $x_t$ is $\tilde{x}_t : \mathbb{R}^n \to \mathbb{R}^n$ such that for each $[x_0] \in \mathbb{R}^n$ we associate the constraint interval vector $\tilde{x}_t([x_0]) \in \mathbb{R}^n$ for all $t \in \mathbb{R}$. In this way, $\tilde{x}_t([x_0])$ is defined as the constraint interval solution, for the interval initial condition $[x_0]$, of Eq. (4).

Although the application $\tilde{x}_t : \mathbb{R}^n \to \mathbb{R}^n$ is not necessarily a flow over $\mathbb{R}^n$ (see Fig. 2) its asymptotic behavior can be analyzed by looking at the behavior of $x_t$ over compact sets of $U \subseteq \mathbb{R}^n$. For a dynamical system $x_t$ defined over a metric space $X$ we say that $S \subseteq X \subseteq \mathbb{R}^n$ is an invariant set if for $x_t(S) = S$ for all $t \in \mathbb{R}$. Two of the most ordinary examples of invariant sets are the equilibrium points and periodic orbits.

An invariant set $S$ for a dynamical system $x_t$ is said to be stable if for every open neighborhood $W$ of $S$ there is a neighborhood $W' \subseteq W$ such that for all $x_0 \in W', x_t(x_0) \in W$ for all $t \in \mathbb{R}$. Furthermore, $S$ is said to be asymptotically stable if $S$ is stable and there is a neighborhood $W_S$ such that for all $x_0 \in W'$, dist$(x_0, S) \to 0$ when $t \to \infty$. Moreover, according to [11] in the case of $S$ being uniformly stable, then for any neighborhood $V$ of $S$ and a compact $B \subseteq W'$ there is $T > 0$ such that $x_t(B) \subseteq V$ for all $t > T$. Finally, an invariant set $S$ is unstable when it is not stable.

The next statements establish a relationship between the flow $x_t$ and its CI extension $\tilde{x}_t$.

**Theorem 2.** Consider the flow $x_t : \mathbb{R} \to \mathbb{R}$ given by Eq.(4) its CI extension $\tilde{x}_t : \mathbb{R}^n \to \mathbb{R}^n$. The following statements are true:

a) A point $x \in \mathbb{R}^n$ is an equilibrium for $x_t$ if and only if $[x] \in \mathbb{R}^n$, in which $[x] = [x, x]$, is an equilibrium for $\tilde{x}_t$.

b) An equilibrium $x \in \mathbb{R}^n$ is stable for $x_t$ if and only if $[x] \in \mathbb{R}^n$, in which $[x] = [x, x]$, is stable for $\tilde{x}_t$.

c) An equilibrium $x \in \mathbb{R}^n$ is uniformly stable for $x_t$ if and only if $[x] \in \mathbb{R}^n$, [x] = [x, x], is uniformly stable for $\tilde{x}_t$.

**Proof.** Let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $[x] = (\{x_1\}, \{x_2\}, \ldots, \{x_n\}) \in \mathbb{R}^n$ such that $[x] = [x, x]$. The first part of this statement follows directly from the identity $x_t([x]) = \tilde{x}_t([x])$. To prove the second and third parts of the statement it is enough to check that for any neighborhood $W_x = \{y \in \mathbb{R}^n : ||x - y|| < \varepsilon\}$ of $x$ the set $W = \{y \in \mathbb{R}^n : [y] \subseteq W_x\}$ is also a neighborhood of the $x$ when $x$ is seen as a point in $\mathbb{R}^n$. The converse is also true. The neighborhood $W = \{y \in \mathbb{R}^n : 0 \leq \text{dist}(x, [y]) < \varepsilon\}$ of $[x] \in \mathbb{R}^n$ defines a $\varepsilon$-neighborhood of $x \in \mathbb{R}^n$. Thus, the statements follows directly from the definition of stable and uniformly stable. □

**Theorem 3.** Let $S$ be an invariant set under $x_t$ and consider

$$S = \{[x] \in \mathbb{R}^n : \pi_S([x]) = [x], s \subseteq S, s \text{ connected and compact}\}. \tag{5}$$

For a given neighborhood $V$ of $S$ it is true that:

a) If $S$ is stable then there are neighborhoods $V$ of $S$ and $V'$ of $S$ such that for all $[x_0] \in \mathbb{R}^n$ satisfying $[x_0] \subset V'$, $\tilde{x}_t([x_0]) \subset V$ for all $t > 0$.

b) If $S$ is uniformly stable then there is a neighborhood $V'$ of $S$ such that for any neighborhood $W$ of $S$ and $[x_0] \in \mathbb{R}^n$, $[x_0] \subset V'$, there is $T > 0$ such that $\tilde{x}_t([x_0]) \subset W$ for all $t > T$.

**Proof.** To prove a), given $\varepsilon > 0$ let us consider the $\varepsilon$-neighborhood $V'$ of $S$ defined as $V = \{x \in \mathbb{R}^n : ||x - y|| < \varepsilon, y \in S\}$. The set $V$ defines a neighborhood $V'$ of $S$ and, since $S$ is stable, there is a neighborhood $V''$ of $S$ such that $x_t(B) \subseteq V$ for all $B \subseteq V''$. In particular, for all $[x_0] \in \mathbb{R}^n$ such that $[x_0] \subset V'$ it is true that $x_t([x_0]) \subset V$ for all $t > 0$. Defining

$$V = \{[x] \in \mathbb{R}^n : [x] = \pi_V([x]), \mu \subset V, \mu \text{ connected and compact}\}$$

it is clear that $V$ is a neighborhood of $S$ and since $\tilde{x}_t([x_0])$ is a compact connected subset of $V$ for all $t > 0$ then $\tilde{x}_t([x_0]) \in V$ for all $t > 0$ and the statement is proved.

To prove b), suppose $S$ uniformly stable. Thus, $S$ is stable and for a given neighborhood $V$ of $S$ there is a neighborhood $V'$ of $S$ such that for any neighborhood $W$ of $S$ and $B \subseteq V'$ there is $T > 0$ such $x_t(B) \subset W$ for all $t > T$. Now, given the neighborhood $W$ defined as

$$W = \{[x] \in \mathbb{R}^n : [x] = \pi_V([x]), \mu \subset V, \mu \text{ connected and compact}\}$$

then, for a given $[x_0] \in \mathbb{R}^n$ such that $[x_0] \subset V'$ there is $T > 0$ such $\tilde{x}_t([x_0]) \subset W$ for all $t > T$. Consequently, $\tilde{x}_t([x_0]) \in W$ for all $t > T$ and the statement is proved. □

**Example 1.** Let us consider the two-dimensional system defined by the equations ([12])

$$\frac{dx_1}{dt} = -x_2 + \mu x_1 (k - x_1^2 - x_2^2), \quad \frac{dx_2}{dt} = x_1 + \mu x_2 (k - x_1^2 - x_2^2) \tag{6}$$

874
For \( \mathbf{x}_0 = (x_{01}, x_{02}) \in \mathbb{R}^2 \) the solution is \( \mathbf{x}_t(\mathbf{x}_0) = (x^1_t(\mathbf{x}_0), x^2_t(\mathbf{x}_0)) \) in which \( x^i_t(\mathbf{x}_0) = r(t) \cos \theta(t), x^2_t(\mathbf{x}_0) = r(t) \sin \theta(t) \) and

\[
r(t, \mathbf{r}_0) = \frac{r_0 \sqrt{k}}{\left[ r_0^2 + (k - r_0^2) e^{-2\mathbf{r}_0^2 t} \right]^{\frac{1}{2}}}, \quad \theta(t, \mathbf{r}_0) = t + \theta_0,
\]

\[
r_0 = \sqrt{x^2_{01} + x^2_{02}}, \quad \theta_0 = \arcsin \left( \frac{x_{02}}{r_0} \right).
\]

Now, let us consider an initial condition \( [\mathbf{x}_0] \in \mathbb{R}^2 \) defined by \( p_1(\gamma_1) = x_{01} + \gamma_1 w_i \) and \( p_2(\gamma_2) = x_{02} + \gamma_2 w_2, w_i = x_{0i} - x_{0i}, i = 1, 2 \). Thus, the constraint interval solution is given by:

\[
\hat{x}_1^i([\mathbf{x}_0]) = \left[ \min x^i_1(\gamma), \max x^i_1(\gamma) \right], \quad \hat{x}_2^i([\mathbf{x}_0]) = \left[ \min x^i_2(\gamma), \max x^i_2(\gamma) \right],
\]

\[\gamma = (\gamma_1, \gamma_2), \gamma_i \in [0, 1], i = 1, 2.\] It is not difficult to check that \( x_t([\mathbf{x}_0]) \) spirals towards the circumference \( S \) of radius \( k \) for all \( x_0 \neq (0, 0) \) so that \( S \) is an asymptotically stable invariant set. Therefore, Theorem 3 ensures that the set

\[ S = \{ [\mathbf{x}] \in \mathbb{R}^2 : [\mathbf{x}_1] = \pi_1(S), [\mathbf{x}_2] = \pi_2(S), s \subset S, S \text{ connected and compact} \}
\]

attracts the constraint interval solution \( x_t([\mathbf{x}_0]) \) for all \([\mathbf{x}_0] \in \mathbb{R}^2 \) such that \((0, 0) \neq [\mathbf{x}_0]\).

The behavior of the constraint interval solution \( x_t([\mathbf{x}_0]) \) for \( \mu = 0.2, k = 1 \) and interval initial condition \([\mathbf{x}_0] = ([0.1, 0.3], [-0.1, 0])\) is shown in Fig. 1. In Fig. 2 shows the behavior of constraint interval solution \( x_t([\mathbf{x}_0]) \) of Eq. (6) on the plane. Given the initial condition \([\mathbf{x}_0] = ([0.1, 0.3], [-0.1, 0])\) we compute the interval solution \( x_t([\mathbf{x}_0]) \) at \( \tau = 2\pi \). As we clearly see, it turns out that \( x_{2\pi}([\mathbf{x}_0]) \) (the blue box on the right) is not equal to \( x_{2\pi}([\mathbf{x}_{2\pi}/2([\mathbf{x}_0])]) \) (the yellow box) so that \( x_t \) is not a flow over \( \mathbb{R}^2 \). The red boxes are the set \( x_t([\mathbf{x}_0]) \) at \( t = \pi/2 \) (left) and \( t = 2\pi \) (right).

### 3.2. Constraint interval uncertainties on initial conditions and parameters

The results previously presented can also be used to analyze the behavior of initial value problems with uncertainties on parameters. In fact, given the equation

\[
\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, \mathbf{p}), \quad \mathbf{x}(0) = \mathbf{x}_0,
\]

\[\text{(7)}\]

**Fig. 1.** Behavior of the constraint interval solution \( x_t([\mathbf{x}_0]) \) to Eq. (6) for \( \mu = 0.2, k = 1 \) and \([\mathbf{x}_0] = ([0.1, 0.3], [-0.1, 0])\).

**Fig. 2.** Constraint interval solution \( x_t([\mathbf{x}_0]) \) to Eq. (6) showing the geometrical properties \( x_t([\mathbf{x}_0]) \) and \( x_t([\mathbf{x}_0]) \) for \( s = \pi/2 \) and \( t = 3\pi/2 \).
where \( f : U \times V \to \mathbb{R}^n, U \subset \mathbb{R}^n, V \subset \mathbb{R}^k \) is a sufficiently smooth function such that (7) has a solution \( x_t(x_0), \forall t \geq 0. \) Given (7), we can define the equation

\[
\begin{align*}
\frac{dx}{dt} &= f(x, p), \quad x(0) = x_0, \\
\frac{dp}{dt} &= 0, \quad p(0) = p;
\end{align*}
\]

and the parameters of Eq. (7) are turned into initial conditions of Eq. (8). Furthermore, if \( x_t(x_0) \) is the solution of Eq. (7) then \( y_t(x_0, p) = (x_t(x_0), p) \) is the solution of Eq. (8).

Thus, for the generalized uncertainties in the initial conditions and parameters, \( x_0 \) and \( p \) of Eq. (7) we consider the constraint interval solution \( y_t([x_0], [p]) \) of Eq. (8).

### 4. Constraint Interval solution to the SIR model

Our main goal in this section is to consider interval uncertainties in the SIR model as well as to analyze the impact of these uncertainties in the asymptotic behavior of the solution. To this end, we are going to consider interval uncertainties for both initial conditions and parameters.

#### 4.1. SIR model

A classic mathematical model to describe the spread of an infectious disease on a population as function of time is the SIR model, proposed by Kermack - McKendrick in 1927 [13]. Despite of its apparent simplicity, the SIR model encompasses the main features of many contagious diseases like the Sars-Cov-2 infection.

This model divides individuals of a population in three categories: susceptible (S), infectious (I) and recovered (R). The proportion of each category in relation to the total population, evolve in time according to the following equation

\[
\begin{align*}
\frac{dS}{dt} &= -\alpha SI, \quad S(0) = S_0, \\
\frac{dI}{dt} &= \alpha SI - \beta I, \quad I(0) = I_0; \\
\frac{dR}{dt} &= \beta I, \quad R(0) = R_0;
\end{align*}
\]

where \( \alpha \) and \( \beta \) are both positive parameters [14].

Since that \( S(t) + I(t) + R(t) = 1 \), defining \( Z(t) = I(t) + R(t) \) the SIR model can be rewritten as

\[
\begin{align*}
\frac{dI}{dt} &= \alpha(1-Z)I - \beta I, \quad I(0) = I_0, \\
\frac{dZ}{dt} &= \alpha(1-Z)I, \quad Z(0) = Z_0,
\end{align*}
\]

where the solution \( x_t(I_0, Z_0) \) of Eq. (10) is defined on the set

\[ U = \{ (I, Z) \in \mathbb{R}^2 : I, Z \geq 0, Z \leq 1, Z \geq 1 \}. \]

The reproductive number \( \rho = \frac{\alpha}{\beta} \), which is the ratio, defines the asymptotic behavior dynamics of Eq. (10) and, according to [15], it is a key parameter determining whether an infectious disease will persist. In fact, looking at the equations of Eq. (10), and considering \( I_0 > 0 \), it can be seen that \( Z(t) \) is an increasing function of time. Yet, for \( I(t) \) it turns out that

\[
\frac{dI}{dt} > 0 \iff Z \leq 1 - \frac{1}{\rho}.
\]

Thus, since \( Z \in [0, 1], I(t) \) is decreasing if \( \rho \leq 1 \). On the other hand, when \( \rho > 1 \) then \( I(t) \) is increasing while \( Z(t) < 1 - 1/\rho \) and decreasing otherwise. Therefore, \( I(t) \) reaches a maximum at the time \( Z(t) = 1 - 1/\rho \).

It is not difficult to check that every point \( E = (0, z) \subset S = 0 \times [1 - \rho^{-1}, 1] \subset U \) is an equilibrium point to the model described by Eq. (10). The Jacobian matrix of Eq. (10) at \( E = (0, z) \) is given by

\[
J(E) = \begin{bmatrix} \alpha(1-z) - \beta & 0 \\ \alpha(1-z) & 0 \end{bmatrix}
\]

so that the eigenvalues are \( \lambda_1 = \alpha(1-z) - \beta \) and \( \lambda_2 = 0 \). We can easily check that \( z < 1 - 1/\rho \) implies \( \lambda_1 > 0 \) so that \( E = (0, z) \) is unstable. Therefore, we can conclude that the set \( S \) is invariant and attracts all solutions of Eq. (10). In fact, since the dynamics of Eq. (10) are bounded on \( U \) then for every compact set \( B \subset U - S \) the \( \omega \)-limit, \( \omega(B) \), is a nonempty, compact and invariant subset of \( U \) ([16]). Since \( S \subset U \) is the a invariant set on \( U \) then we must have \( \omega(B) \subset U \) and, therefore, \( S \) attracts compact subsets \( B \) of \( U \).

Using the chain rule for derivatives it turns out that

\[
\frac{dl}{dz} = 1 - \frac{\beta}{\alpha(1-z)}
\]

whos the solution at \( (l_0, Z_0) \) is given by
\begin{equation}
I(I_0, Z_0, Z) = Z - Z_0 + I_0 + \frac{1}{\rho} \ln \left( \frac{1 - Z}{1 - Z_0} \right).
\end{equation}

Therefore, the equilibrium for Eq. (10) is the solution of \( I(I_0, Z_0, Z) = 0 \), which can be expressed as
\begin{equation}
E(I_0, Z_0, \alpha, \beta) = (0, Z_e), \quad Z_e = 1 + \frac{W(-\rho e^{-\rho(1-c)})}{\rho},
\end{equation}

with \( W \) being the Lambert \( W \) function \((17)\) and \( c = Z_0 - I_0 + \rho^{-1} \ln (1 - Z_0) \).

To account for uncertainties on the parameters we follow the recipe presented in the previous section and consider the expanded system
\begin{equation}
\begin{cases}
\frac{dx}{dt} = \alpha(1 - Z)I - \beta I, \\
\frac{dI}{dt} = \alpha(1 - Z)I,
\end{cases}
\end{equation}

so that \( y_t(I_0, Z_0, \alpha, \beta) = (x_t(I_0, Z_0), \alpha, \beta) \) is the solution to Eq. (13).

Using the results discussed earlier it turns out that \( y_t(I_0, Z_0, \alpha, \beta) \) converges uniformly to \( E(I_0, Z_0, \alpha, \beta) = (0, Z_e, \alpha, \beta) \) and the \( S = \{0\} \times [1 - 1/\rho, 1] \times [\alpha] \times [\beta] \) is the attracting invariant set for \( y_t \).

### 4.2. Behavior of the constraint interval solution

Based on the asymptotic behavior of \( x_t \) and \( y_t \), together with Theorem 1 and Theorem 3 the following is behavior of the CI extensions \( \hat{x}_t \) and \( \hat{y}_t \). To this end, let \( U \subset \mathbb{R}^2 \) be the set
\[ U = \{(I, Z) \in \mathbb{R}^2 : I, Z > 0, Z \leq 1, Z \geq I\} \]
and as before \( \pi_i \) is the orthogonal projection onto the \( i \)th axis.

**Theorem 4.** Let \( x_t \) be the flow of Eq. (10) and \( \hat{x}_t \) be its CI extension. Suppose \( \alpha, \beta \) positive scalars and \( [x_0] = ([I_0], [Z_0]) \in \mathbb{R}^2 \), \( [x_0] \in U \). Considering the set \( S = \{0\} \times [1 - 1/\rho] \) the following are true:

a) The set \( S = \{[x] \in \mathbb{R}^2 : x_1 = \pi_1(s), s \in S, \text{connected and compact} \} \) attracts the constraint interval solution \( \hat{x}_t([x_0]) \).

b) The constraint interval solution \( \hat{x}_t([x_0]) \) converges to \( \hat{E}([x_0]) \), with \( E \) defined in Eq. (10).

**Proof.** Part a) follows directly from the Theorem 3. To prove b), it is enough to prove that \( x_t \) converges uniformly to \( E \) on compact subsets \( K \) of \( U \). But, once \( Z(t) \) is an increasing function of \( t \), Dini’s theorem ensures that \( Z(t) \) converges uniformly to the equilibrium \( Z_e \) \((18)\). Since \( I(t) \) is given by Eq. (11) then \( x_t(I_0, Z_0) \) converges uniformly to the equilibrium \( E(I_0, Z_0) \) for all \( (I_0, Z_0) \in K \). Therefore, Theorem (1) ensures that the CI extension \( \hat{x}_t([x_0]) \) converges to the CI extension \( \hat{E}([x_0]) \). \( \square \)

**Theorem 5.** Let \( y_t \) be the flow of Eq. (13) and \( \hat{y}_t \) be its CI extension. Consider the intervals \( [x] = [x_1, x_2], [\beta] = [\beta_1, \beta_2] \) and \( [x_0] = ([I_0], [Z_0]) \in \mathbb{R}^2 \), \( [x_0] \subset U \). Defining the set \( S \subset \mathbb{R}^4 \) as
\[ S = \{(x_1, x_2, x_3, x_4) : x_1 = 0, x_2 = 1 - x_4/x_3, x_3 \in [\alpha], x_4 \in [\beta]\}, \]
then the following are true:

a) The set \( S = \{[x] \in \mathbb{R}^4 : x_1 = \pi_1(s), s \in S, \text{connected and compact} \} \) attracts the constraint interval solution \( \hat{y}_t([x_0], [x], [\beta]) \).

b) The constraint interval solution \( \hat{y}_t([x_0], [x], [\beta]) \) converges to \( \hat{E}([x_0], [x], [\beta]) \) where \( E((I_0, Z_0), a, b) = (0, Z_e, a, b) \) and \( Z_e \) as in Eq. (12).

**Proof.** The first item follows directly from the Theorem 3. To prove b), it is enough to see that \( y_t \) converges uniformly to \( E \) for all \( (I_0, Z_0, a, b) \in K \times [x] \times [\beta], K \) compact subsets over \( U \). Therefore, Theorem (1) ensures that the CI extension \( \hat{y}_t([x_0], [x], [\beta]) \) converges to the CI extension \( \hat{E}([x_0], [x], [\beta]) \). \( \square \)
5. Regression procedure

Several policies to mitigate the spread of the disease have been proposed by Brazilian authorities. Moreover, such policies have changed over time as well as with geographical areas and such changes can be an obstacle to fit the model to the data. In this section we are going to use the constraint interval solution of the model here presented in order to predict the time evolution of the Sars-Cov-2 disease in Brazil.

5.1. The data

We are mainly interested in the cumulative of cases, or total cases, of Sars-Cov-2 in Brazil. Among many other useful information, the cumulative number of individuals diagnosed with Sars-Cov-2 is daily updated by official authorities and reported in the official coronavirus website https://covid.saude.gov.br ([19]).

Let \( a_k \) denote the cumulative of cases at time \( k \), \( k \in \{1, 2, \ldots, N\} \), \( k \) ranging from 26 February 2020 to 10 July 2021. From the time series data of total cases we can infer the daily new cases just by computing the difference of two consecutive total cases data, that is \( n_k = a_k - a_{k-1} \). We can also infer the number of currently infected by computing the difference \( i_k = a_k - a_{k-\tau} \), where \( \tau = \beta^{-1} \) is the infectious period of Sars-Cov-2.

It is clear the presence of a second wave of infection as shown in Fig. 3. In order to better fit the model to the data, we split the data set in two parts: the first one ranging from 26 February 2020 to 03 November 2020 and the second one ranging from 04 November 2020 to 10 July 2021.

To the regression procedure we are considering the relative values (cumulative cases and currently infected) to the total population, currently estimated at 211 million individuals. Furthermore, following the classical approach used in machine learning ([20]), we split the first part of the data set (from 26 February 2020 to 03 November 2020) into training and testing sets (85% - 15% of the data, respectively) in order to validate the model.

5.2. Infectious period and Under-reporting

The infectious period is key parameter to any epidemiological model and several studies have been made in order to estimate it to Sars-Cov-2. A study conducted in [21] suggests that infectiousness may start 1 to 3 days prior to the onset of symptoms and declines within seven days. Furthermore, transmission after 7 to 10 days of illness is unlikely [22].

It is also important to consider the under-report aspect of the data published by Official Brazilian government sources. According to [23], under-reported infections range from 11 to 30 times the number of confirmed cases and in [24] suggests that reporting rate is at 9.2% (IC95: 8.8% - 9.6%) of confirmed cases. We use these estimates to explore some scenarios, comparing the results with the best-fit scenario.

5.3. Objective function

Let \( a_k \) and \( i_k \) be the proportions of reported total cases and infectious at time \( k \), respectively. As we have described previously, the behavior of the dynamics relies on the ratio \( \rho = \alpha/\beta \) so that we suppose uncertainty just on \( \alpha \), considering \( \beta \) a deterministic unknown. Therefore, we assume that the initial conditions \( I_0 \) and \( Z_0 \), with \( Z_0 = I_0 \), as well as \( \alpha \) are unknown intervals.

Let \( [\tilde{I}_t, \bar{I}_t] = \tilde{I}_t([I_0], [Z_0], [\alpha], [\beta]) \) and \( [\tilde{Z}_t, \bar{Z}_t] = \tilde{Z}_t([I_0], [Z_0], [\alpha], [\beta]) \) be the CI extension of solutions \( I(t) \) and \( Z(t) \) from Eq. (13) and let us define \( \tilde{I}_t \) and \( \tilde{Z}_t \) as

\[
\tilde{I}_t = \frac{\bar{I}_t + \tilde{I}_t}{2}, \quad \tilde{Z}_t = \frac{\bar{Z}_t + \tilde{Z}_t}{2}.
\]

Fig. 3. Time evolution of Sars-Cov-2 pandemic in Brazil. The 7-days moving average on the daily new cases shows the onset of a second wave of infections at 03 November 2020.
To estimate the unknown intervals \( [Z_0] = [\bar{Z}_0, \tilde{Z}_0] \) and \( [\alpha] = [\bar{\alpha}, \tilde{\alpha}] \), we minimize, over the training set, the Mean Squared Error (MSE) (\([20]\)) given by the following objective function

\[
F(Z_0, \bar{Z}_0, \tilde{Z}_0) = \frac{1}{N} \sum_{k=1}^{N} \left[ a \left( \frac{i_k}{s} - \bar{I}_k \right)^2 + (1 - a) \left( \bar{a}_k - \tilde{Z}_k \right)^2 \right]
\]

where \( s > 0 \) is a constant accounting for under-report of cases and \( a > 0 \) is trade-off parameter.

For the optimization process we use the Octave’s fmsearch algorithm ([25]), setting both TolFun and TolX options as 1e-7.

5.4. The instantaneous reproduction number \( R_t \)

The pandemic behavior, as previously mentioned, is defined by the reproductive number \( \rho = \alpha/\beta \) and the previous procedure allows us to determine the reproductive number for the whole data set. However, due to changes in both policies and behavior of the population the reproduction number can change over time and nevertheless we would like to measure it.

The main idea is to track the value of reproduction number over time consists in finding the CI solution that best fits a subset of the data set. That is, to compute the instantaneous reproduction number \( R_{k+1} \) at a time \( k + 1 \), the objective function

\[
F(\bar{s}_{k+1}, \tilde{s}_{k+1}) = \frac{1}{2} \sum_{j=0}^{2} \left[ a \left( \bar{i}_{k,j} - \bar{I}_j \right)^2 + (1 - a) \left( \bar{a}_{k,j} - \tilde{Z}_j \right)^2 \right]
\]

is minimized where \( \bar{I}_j \) and \( \tilde{Z}_j \) are as in Eq. (14) and the CI solution is computed considering \( [i_0] = [\bar{i}_{k,j}, \tilde{i}_{k,j}] \) and \( [\alpha] = [\bar{\alpha}_k, \tilde{\alpha}_k] \). Finally, \( R_{k+1} \) is given by \( R_{k+1} = \frac{\bar{a}_{k,j}}{\tilde{\alpha}_k} \). That is, in the above procedure, for each \( k \) in the range of the data set, given the initial condition \( (i_0, Z_0) = (\bar{i}_0, \tilde{Z}_0) \) the interval \( [\alpha_{k+1}] = [\bar{s}_{k+1}, \tilde{s}_{k+1}] \) is found such that the CI solution best fits the data \((\bar{i}_{k,j}, \tilde{i}_{k,j}, \bar{a}_{k,j}, \tilde{\alpha}_k)\) and \( (\bar{i}_0, \tilde{Z}_0) \).

6. Results

The results of this section are obtained following a machine learning approach on the first part as follows: for each \( \beta \) and \( s \) fixed, the objective function (15) is minimized over the training set (\([20]\)). Next, we compute the Mean Absolute Percentage Error (MAPE) using

\[
\text{MAPE} = \frac{1}{2R} \sum_{k=1}^{K} \left( \frac{\bar{a}_k}{s} - \tilde{Z}_k \right) + \left( \frac{\bar{i}_k}{s} - \bar{I}_k \right)
\]

over the testing set in order to find the reporting rate \( s \) that best predicts the data on the testing set for that fixed \( \beta \) (\([20]\)).

Simulations on a \( 400 \times 400 \) grid of pairs \((\beta, s)\) are performed in which \( \beta \in [0.1, 0.5] \) and \( s \in [0.01, 0.75] \). Some relevant results are compiled in Table 1.

We use the estimations for \( \beta \) and the reporting rate \( s \) of the first part of the data set to find the parameter \( \alpha \) that best fits the model to the data of the second part of the data set.

Table 1 shows the CI solution that best fits the reported data indicates an infectious period of \( \tau = 8.24 \) days and a reproductive number interval \([\rho] = [1.20, 1.28] \). For this CI solution, the reported rate is \( s = 8.3\% \). In the second best-fitted scenario, it turns out that \( \tau = 7.26 \) days and \([\rho] = [1.17, 1.24] \) with a reported rate of \( s = 9.2\% \).

As well know, parameters \( \alpha \) and \( \beta \) are paramount to describe the dynamics of spread and they are usually not known at the beginning of a new infectious disease, as it is the case of the Sars-Cov-2 infection. Estimations as presented in Table 1 could be informative to predict and control the spread of the disease as well as be useful in comparison to other estimation methods.

### Table 1

| \( Z_0 \)   | \( Z_0 \)   | \( \bar{Z}_0 \) | \( \tilde{Z}_0 \) | \( \bar{\alpha} \)   | \( \tilde{\alpha} \) | \( \beta \)   | \( s \)   | MAPE      |
|-------------|-------------|----------------|----------------|---------------------|---------------------|-------------|--------|---------|
| 5.93E-04    | 7.54E-04    | 0.13300        | 0.14295        | 0.10918             | 0.07680             | 0.05012     |        |         |
| 4.20E-04    | 6.69E-04    | 0.14595        | 0.15567        | 0.12143             | 0.08292             | 0.04848     |        |         |
| 4.02E-04    | 5.84E-04    | 0.16143        | 0.17090        | 0.13776             | 0.09158             | 0.04925     |        |         |
| 3.07E-04    | 4.55E-04    | 0.18530        | 0.19409        | 0.16122             | 0.10352             | 0.05273     |        |         |
| 3.19E-04    | 3.49E-04    | 0.20945        | 0.22548        | 0.19388             | 0.14506             | 0.08136     |        |         |
| 1.71E-04    | 2.43E-04    | 0.23868        | 0.27291        | 0.23980             | 0.15547             | 0.10490     |        |         |
| 1.37E-04    | 1.48E-04    | 0.33653        | 0.35178        | 0.32143             | 0.23339             | 0.14520     |        |         |
| 5.19E-05    | 5.94E-05    | 0.49632        | 0.51143        | 0.47857             | 0.29733             | 0.19299     |        |         |
To illustrate the results, in Fig. 4 we present the CI solution that best predicts the data on the test set as indicated in Table 1. The reproductive numbers estimated from Table 1 are obtained by fitting the CI solution to the first part of the data set. Thus, these estimates refer to an overall dynamics and, since social distancing policies as well as the behavior of the population change over time we should expect fluctuations on the reproductive number as time evolves. Since policies of mass testing can also change over time, the reporting rate $s$ can also be affected. Therefore, although the parameters in Table 1 give reasonable predictive values on the test set the CI solutions based on them is not able to predict the second wave of infections.

To the second wave of infections we use previous estimates for $\beta$ and then the objective function is optimized over the second part of the data set (from 04 November 2020 to 10 July 2021) to find the $[x]$ and $s$ that best fits the model to the data from 04 November 2020 to 10 July 2021. Fig. 5 shows the behavior of the CI solution that best fit the data in this scenario. The graphical representation of the constrain interval solution that best fits the second wave’s data is in Fig. 5. This CI solution is obtained considering $\beta = 0.1214$ and $[x] = [0.1419, 0.1449]$ and so, the reproductive number is $[\rho] = [1.1684, 1.1931]$. The initial condition in this case is the final condition of the CI solution in Fig. 4.
Fig. 6 shows the 7-day moving average of the daily reproductive number $R_t$ estimated following the recipe described in SubSection 5.4. The mean of daily reproductive number $R_t$ by epidemiological week of 2020–2021 is also shown. Although the time-varying $R_t$ presents a lot of fluctuations over time we can observe an increasing trend in the days. Considering the 7-day moving average, the reproductive number was $1.40$ ($[R_t] = [1.27, 1.53]$) on 13 November 2020, the highest value since 30 April 2020. The minimum value for the reproductive number is $0.61$ ($[R_t] = [0.38, 0.85]$) on 07 September followed by $0.63$ and $0.63$ on 08 of September and 02 of November, respectively. The maximum value for the reproductive number is $2.62$ ($[R_t] = [1.46, 3.77]$) on 07 March followed by $2.60$ on 05 and 06 March. On 13 November 2020 the reproductive number was $1.48$ ($[R_t] = [1.15, 1.80]$), the highest value since 30 April 2020. We remark also a peak on 09 January 2021 in which the reproductive number was $1.42$ ($[R_t] = [1.27, 1.58]$).

Now, considering mean by epidemiological week, the reproductive number was $1.48$ ($[R_t] = [1.14, 1.80]$) at week 46 (08–14 November 2020), the highest value since week 18 (26 April – 02 May), and it was $1.40$ ($[R_t] = [0.97, 1.82]$) on the first week of 2021. The minimum value for the reproductive number is $0.74$ ($[R_t] = [0.34, 1.20]$) at week 37 followed by $0.76$ and $0.82$ at weeks 52 and 45, respectively. The maximum value for the reproductive number is $2.99$ ($[R_t] = [2.16, 3.81]$) at week 10 followed by $2.66$ and $2.65$ at weeks 11 and 12, respectively. Furthermore, the mean weekly $R_t$ is not greater than 1 at weeks 28–29, 32–37, 39–42, 44–45, 52–53, 56, 58, 66, 69 and 79. Of these weeks, all the interval $[R_t]$ is less than 1 only at weeks 36, 44, 52, 66 and 79.

Finally, we compare our approach to estimate the time-varying $R_t$ with the method presented in [26]. Denoting by $d_t$ the number of new cases at time $t$, the authors in [26] define the time-varying reproductive number by $R_t = d_t / \left( \sum_{s=1}^{t} w_s d_{t-s} \right)$ in which $w_s$ is a probability measurement related to infectivity. Considering $\beta = 0.1214$ then the infectious at time $t$ are they who get infected up to 8 $(1/\beta)$ days earlier. The red lines in Fig. 6 are the time-varying reproductive number estimated according to this approach (CFFC method) assuming $w_s = 1/8$ for $s = 1, 2, \ldots, 8$. The mean value for the $R_t$ estimated by the CFFC method is $1.14$ (CI 95%: 1.09–1.20) and the data defining the black and red lines on the left in Fig. 6 haas $0.96$ (p-value$=1.71e^{-263}$) correlation coefficient.

7. Conclusion

This study presented the theoretical foundations of the constraint interval representation in nonlinear differential equations and least squares problems under interval uncertainties in initial conditions, parameters, and data. Using the constraint representation of interval uncertainty, this study showed how to define and obtain resulting solutions to interval nonlinear differential equations and interval least squares problems. The efficacy and efficiency of the constraint interval approach was demonstrated in the analysis of a current nonlinear differential equations applied problem, the SIR analysis of the current COVID pandemic in Brazil, where the interval parameters and initial conditions were calculated using interval least squares approaches together with machine learning.

CRediT authorship contribution statement

M.S. Cecconello: Conceptualization, Methodology, Formal analysis, Writing - original draft. M.T. Mizukoshi: Formal analysis, Writing - original draft, Writing - review & editing. W. Lodwick: Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T.L. Heath, On the measurement of the circle, Cambridge, 1897.
[2] T. Allahviranloo, W. Pedrycz, Soft Numerical Computing in Uncertain Dynamic Systems, Academic Press, 2020.
[3] R.E. Moore, R.B. Kearfott, M.J. Cloud, Introduction to interval analysis, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009.
[4] W.A. Lodwick, Constrained interval arithmetic, CCM report 138, February 1999 (1999).
[5] Y. Chalco-Cano, A. Rufián-Lizana, H. Román-Flores, M.D. Jiménez-Gamero, Calculus for interval-valued functions using generalized Hukuhara derivative and applications, Fuzzy Sets Syst. 219 (2013) 49–67.
[6] Z. Liao, P. Lan, Z. Liao, Y. Zhang, S. Liu, TW-SIR: time-window based SIR for COVID-19 forecasts, Sci. Rep. 10 (1).
[7] W. Jing Zhu, S. Feng Shen, An improved sir model describing the epidemic dynamics of the covid-19 in china, Res. Phys. 25.
[8] M. Hirsch, S. Smale, Differential equations, dynamical systems, and linear algebra, Pure and applied mathematics, Acad. Press, 1974.
[9] O.N. Bjørnstad, K. Shea, M. Krzywinski, N. Altman, Modeling infectious epidemics, Nat. Methods 17 (5) (2020) 455–456.
[10] L. Stefanini, B. Bede, Generalized Hukuhara differentiability of interval-valued functions and interval differential equations, Nonlinear Anal. 71 (3–4) (2009) 1311–1328.
[11] W. Jing Zhu, S. Feng Shen, An improved sir model describing the epidemic dynamics of the covid-19 in china, Res. Phys. 25.
[12] M. Hirsch, S. Smale, Differential equations, dynamical systems, and linear algebra, Pure and applied mathematics, Acad. Press, 1974.
[13] W.O. Kermack, A.G. McKendrick, Contributions to the mathematical theory of epidemics—i, Bull. Math. Biol. 53 (1–2) (1991) 33–55.
[14] L. Edelstein-Keshet, Mathematical Models in Biology, Society for Industrial and Applied Mathematics, USA, 2005.
[15] J.T. Griffin, T. Garske, A.C. Ghani, P.S. Clarke, Joint estimation of the basic reproduction number and generation time parameters for infectious disease outbreaks, Biostatistics 12 (2) (2010) 303–312.

[16] J. Hale, H. Koçak, Dynamics and Bifurcation, Text in Applied Mathematics, Vol. 3, Springer - Verlag, New York, 1991.

[17] L. Euler, De serie Lambertina Plurimisque eius insignibus proprietatibus, Vol. 2, 1763.

[18] R.G. Bartle, D.R. Sherbert, Introduction to real analysis, Vol. 2, Wiley, New York, 2000.

[19] Coronavirus panel, Brasília, 2020.https://covid.saude.gov.br.

[20] T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learning: data mining, inference and prediction, 2nd Edition., Springer, 2009.

[21] J.Y. Chun, G. Baek, Y. Kim, Transmission onset distribution of COVID-19, Int. J. Infect. Dis. 99 (2020) 403–407.

[22] H.-Y. Cheng, S.-W. Jian, D.-P. Liu, T.-C. Ng, W.-T. Huang, H.-H.L. and Contact tracing assessment of COVID-19 transmission dynamics in taiwan and risk at different exposure periods before and after symptom onset, JAMA, Int. Med. 180 (9) (2020) 1156.

[23] L. Veiga e Silva, M.D.P. de Andrade Abi Harb, A.M. Teixeira Barbosa dos Santos, C.A. de Mattos Teixeira, V.H. Macedo Gomes, E.H. Silva Cardoso, M.S da Silva, N.L. Vijaykumar, S. Venâncio Carvalho, A. Ponce de Leon Ferreira de Carvalho, C.R. Lisboa Frances, Covid-19 mortality underreporting in brazil: Analysis of data from government internet portais, J Med Internet Res 22 (8).

[24] M. Prado, B. Anrunes, L. Bastos, I. Peres, A. Silva, L. Dantas, F. Baião, P. Maçaira, S. Hamacher, F. Bozza, Análise da subnotificação de COVID-19 no Brasil., Rev. Bras. Ter. Intensiva 32 (2).

[25] J.W. Eaton, D. Bateman, S. Hauberg, R. Wehbring, GNU Octave version 4.4.0 manual: a high-level interactive language for numerical computations (2018). https://www.gnu.org/software/octave/doc/v4.4.1/.

[26] A. Cori, N.M. Ferguson, C. Fraser, S. Cauchemez, A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics, Am. J. Epidemiol. 178 (9) (2013) 1505–1512.