A Novel Metaheuristic Framework Based on the Generalized Boltzmann Distribution for COVID-19 Spread Characterization

JUAN DE ANDA-SUÁREZ[ID][1]**, VALENTÍN CALZADA-LEDESMA**[ID][1]**,**
DAVID ASAEL GUTIÉRREZ-HERNÁNDEZ[ID][2]**, RAÚL SANTIAGO-MONTERO**[ID][2]**,**
LUIS FERNANDO VILLANUEVA-JIMÉNEZ[ID][1]**, AND SERGIO RODRÍGUEZ-MIRANDA**[ID][1]
[1]Tecnológico Nacional de México—Instituto Tecnológico Superior de Purísima del Rincón, Purísima del Rincón, Guanajuato 36413, México
[2]Tecnológico Nacional de México—Instituto Tecnológico de León, León, Guanajuato 37290, México
Corresponding author: Valentín Calzada-Ledesma (valentin.cl@purisima.tecnm.mx)

ABSTRACT Coronavirus disease 2019 (COVID-19) is a highly communicable viral infection caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2), which has spread rapidly throughout the world. From a computer science point of view, research efforts have focused on the use of approaches such as machine learning and curve fitting to predict or simulate disease behavior. However, the mathematical characterization of the spread of COVID-19 is a topic that has not yet been explored by these techniques. In this work, we propose a novel metaheuristic framework called META-COVID19, which merges the Generalized Boltzmann distribution and the family of Jacobi polynomials to automatically characterize the COVID-19 spread without prior knowledge of the data and without involving a human expert. In general terms, the algorithm receives as input a time series of daily reported cases and the output is a polynomial mathematical model. Our framework only needs a single parameter, which is the number of Jacobi polynomials to analyze during the iterative process, and it is capable of proposing polynomials whose adjustment error is close to 1E-3. Finally, we show the applicability of the polynomial models found by META-COVID19, through a theoretical mathematical analysis in order to know attributes of the spread of COVID-19 in different periods of time, allowing to generate better strategies to face it in the future.

INDEX TERMS COVID-19, epidemic spread characterization, generalized boltzmann distribution, Jacobi polynomials, metaheuristic.

I. INTRODUCTION

Historically, humanity has dealt with different pandemics that have taken the lives of millions of people. For example, in the 14th century, the "Black Death" spread across Euroasia and North Africa, causing an estimated 75–200 million deaths, while in the 20th century, 50–100 million people died in Europe and North America due to the "Spanish Flu" [1]. In December 2019, a novel virus called SARS-CoV2 that causes the Coronavirus disease 2019 (COVID-19) emerged in Wuhan, China [2]. In March of 2020, the World Health Organization (WHO) declared the COVID-19 as a pandemic. However, unlike the pandemics mentioned above, globalization caused the virus to spread rapidly on all five continents

in approximately three months. Unfortunately, it has claimed the lives of many people around the world.

From a mathematical and computational point of view, several works have been proposed in the state-of-the-art to understand the behavior of an epidemic. One of the most known tools is the SIR deterministic model, which tries to describe the behavior of an epidemic based on three main factors: the Susceptible population, the Infected population, and the Recovered population [3], [4]. However, the SIR model and its variants, have shown difficulties in modeling the real behavior of the recent pandemic caused by COVID-19, due to its non-deterministic nature and the large number of factors and parameters to consider [5]–[10]; in addition, highly specialized knowledge is needed for its correct implementation. That is why in various investigations, researchers have chosen to use non-deterministic tools such as machine learning, to provide a rapid and approximate overview of

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott[ID].

the behavior of the pandemic, since this type of approach has shown adequate performance in modeling the behavior of other epidemics, such as Ebola, Zika, H1N1 Flu, and Dengue [11]–[15].

After an extensive search on the state-of-the-art, to the best of our knowledge, there are no studies using non-deterministic tools to mathematically characterize the spread of COVID-19. Instead, research efforts have focused on real-time forecasting, computationally simulating the spread of COVID-19, and study regions of possible new outbreaks [16]–[19]. In [20], there was an in-depth discussion of why various methods attempting to forecast COVID-19 have not worked properly, this due to poor data entry, lack of incorporation of epidemiological characteristics, incorrect modeling assumptions, lack of transparency, among others. Likewise, the journal *Nature* published a manifesto to avoid the bad practices of the prediction approaches for COVID-19 [21]. That is why the contribution of our work is not about forecasting. Instead, we focus on performing a characterization.

In general terms, we conceptualize characterization as a way of representing a phenomenon through a model from which attributes can be obtained. To characterize the COVID-19 spread, a data-driven analysis of daily reported cases, either infections or deaths can be conducted. The corresponding data is made up of two independent variables that form a curve (time series) with a non-linear and somewhat stochastic behavior [22], which can be represented mathematically by means of a polynomial curve-fitting. However, choosing the family of polynomials and their respective parameter settings is a very difficult task, since it is desirable that the polynomial be compact, suitable for evaluation, and invariant to small changes in data. It is worth mentioning that, like machine learning, research efforts using curve-fitting to deal with COVID-19 issues have focused on forecasting and not characterization [23]–[28].

Currently, there is a growing trend in the use of statistical mechanics to address computational problems. For example, fluid simulation through Lattice Boltzmann Methods (LBM) [29], [30], which incorporate the Maxwell-Boltzmann distribution, or computational optimization, where an objective function is represented through the Boltzmann distribution [31], [32], and more recently, the use of Boltzmann for COVID-19 forecasting purposes [33], [34]. Although there is a close mathematical relationship between the Maxwell-Boltzmann and Boltzmann distributions, they are different. The Maxwell-Boltzmann distribution gives the probabilities of particle speeds or energies in gases (where the governing equations need to be defined) while the Boltzmann distribution gives the probability that a system will be in a certain state as a function of that state's energy (in Section II we go deeper into this aspect). The Boltzmann distribution is a powerful mathematical tool that has been used successfully in the field of computational optimization, thus creating a research branch of Boltzmann-based metaheuristics.

In this paper, we address the problem of characterizing the spread of COVID-19 from a computational optimization perspective. We propose a novel metaheuristic framework that merges the Generalized Boltzmann distribution [35] and the class of orthogonal Jacobi polynomials [36] (more details will be given in Secction II), specifically designed to automatically perform a polynomial fitting of COVID-19 time series of different countries, without prior knowledge of the data, and without involving a human expert. Generally speaking, the framework input is a COVID-19 time series, either of new confirmed infections or deaths. Subsequently, an iterative process is started, in such a way that the framework performs an automatic search using two simple Selection Operators, which encapsulate the benefits of the Generalized Boltzmann distribution (see Section II-A), on a Characterization Space (Section II-B) where the existence of a Jacobi polynomial that fits with a minimum error to the COVID-19 time series is guaranteed, all this without having prior knowledge about the curve behavior and even for a limited amount of data. Finally, the output is an explicit mathematical model based on the adjusted Jacobi polynomial, which encapsulates the behavior of the non-linear curve; in this way, it is possible to carry out a subsequent mathematical analysis to know specific attributes.

It is important to mention that other polynomial families such as Zernike, Legendre, Gegenbauer, and Chebyshev are special cases of the Jacobi polynomials [36]. Furthermore, under spectral theory, mathematical guarantees are known about the existence of a polynomial function that fits the time series [36]. Because of this mathematical generalization, the study by hand for the selection of the polynomial family that best fit the curve as well as the tuning of its parameters is avoided with our framework. Likewise, the Jacobi polynomials have a large number of mathematical properties [36], that is, from a single polynomial model, it is possible to obtain multiple information that can have an important impact from a statistical inference point of view, which is a cornerstone of epidemiology [37]–[40]. In Section V, three examples of the application of formal mathematical analysis are detailed. However, analyzing all mathematical properties is something that is beyond the scope of the article, and we have framed it as a future job.

The mathematical models found by our metaheuristic framework may be a starting point for other researchers interested in studying the COVID-19 spread. This is of utmost importance, since as anticipated, a new outbreak of COVID-19 (or even new pandemics) may reach the world shortly, so having useful information about the past and present behavior of the pandemic, based on the decisions made at that time, will allow generating better strategies to face it in the future.

The rest of this paper is organized as follows. In Section II the related concepts and the proposed methodology are described. Section III details the algorithmic implementation of the metaheuristic framework. Section IV contains the experimental design, the datasets description, the parameter settings, a discussion, and the results. The mathematical
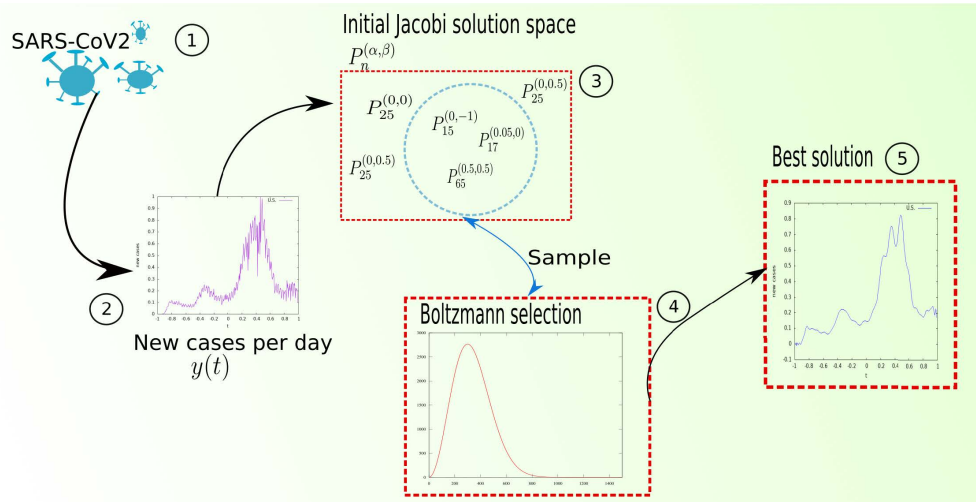
FIGURE 1. META-COVID19 schematic diagram: 1) Daily reported cases of SARS-CoV2; 2) COVID-19 data obtained by country; 3) Characterization Space bounded to the problem domain; 4) Boltzmann-based Selection Operators to sample new solutions; 5) Solutions that best fit the COVID-19 time series by country.

model application of the polynomials proposed by our meta-heuristic framework is shown in Section V. Finally, our conclusions and directions for future work are offered in Section VII.

II. MATERIALS AND METHODS

The general methodology scheme that we followed to build our proposed metaheuristic is shown in Figure 1. As a first step, we obtained the SARS-CoV2 data available by country. Then, we established the problem domain, focused on the characterization of COVID-19 time series $y(t)$ (description is provided in Section IV-A). Subsequently, we proposed a Characterization Space (Section II-B) bounded to the problem domain, in such a way that a sample of promising solutions based on orthogonal Jacobi polynomials (dotted blue circle) can be selected automatically by our metaheuristic in order to fit the time series. As a four-step, we mathematically proposed the Selection Operators μ and ν (Section II-A) based on Boltzmann, which are capable of being iteratively refined to sample better and offer better solutions. Finally, we obtained the adjusted Jacobi polynomials that best fit the COVID-19 time series by country. Next, we will give a detailed description of the previously exposed methodology, as well as the theoretical and experimental implications that we found during the development of this research.

The metaheuristic optimization algorithms have been widely used to solve modern problems in various research areas such as applied engineering, biology, computer vision, data science, physics, and other areas of knowledge [41]–[45]. In general terms, these algorithms have been designed to find approximate solutions to optimization problems in which deterministic techniques are excluded due to the characteristics of the problem.

In general, metaheuristics can be classified according to the type of search they perform, either exploitation or exploration. In the context of the search by exploitation,

a diversification of the proposed solutions is generated in the domain of the local neighborhood, while in the search by exploration the previous information of the solutions found is used to generate a new set of optimal solutions [46], [47]. Particularly in the latter approach, the use of the Boltzmann distribution has been proposed [32], [48], [49] since this distribution guarantees that the probability of finding a new set of solutions that performs worse than the previous set of solutions is approximately zero, which is desirable in a search by exploration scheme.

Broadly speaking, Boltzmann-based metaheuristics build explicit probabilistic models that are iteratively refined to produce increasingly better solutions for a particular problem. These approaches employ an approximation to the Boltzmann distribution through a Gaussian distribution with μ and ν parameters, in order to establish a random number generation mechanism that is used to sample new promising solutions, this is because it is difficult to sample random numbers directly with the Boltzmann distribution. The approximation is generally derived from the minimization of a divergence measure, calculated between the Gaussian and Boltzmann distributions [32], [48], [49]. However, these types of metaheuristics use the original Boltzmann distribution, which was developed to treat thermodynamic problems by means of statistical mechanics following a set of restrictions of the system under study [35], but it is possible to find problems that, due to their characteristics, are outside the proposed analysis restrictions in said distribution, for example, the stochastic behavior of a COVID-19 time series [22].

In this paper, we propose to use the Generalized Boltzmann distribution to power the search for the best Jacobi polynomial, in order to automatically characterize the COVID-19 time series effectively. However, like the previously mentioned Boltzmann-based approaches, to achieve this, we must mathematically obtain the parameters μ and ν, derived from the minimization of a divergence measure between the Gaus-

sian and Generalized Boltzmann distributions. As far as we know, this is something that has not yet been explored by any Boltzmann-based approach.

In the next section, we explain in detail how we obtained mathematically the minimum difference between the Gaussian and the Generalized Boltzmann distribution using a divergence measure, in order to demonstrate the mathematical origin of the Selection Operators (μ and ν) that we simply apply in our metaheuristic framework (see Section III).

A. SELECTION OPERATORS

The Generalized Boltzmann distribution for a system is defined in the following equation [35]:

$$P(x) = P_x = \frac{1}{Z} g(f_x) \exp(\beta f_x). \tag{1}$$

Given the particular problem of characterizing the COVID-19 spread, we will describe the terms of (1) in the context of the problem to be addressed. That said, we have that f_x (a shorthand notation for $f(x)$) is the error of adjustment between a Jacobi polynomial and the COVID-19 time series (this is described in detail in Section III), $\beta = \frac{1}{T}$, where T could be considered as the characteristic temperature of the COVID-19 time series of each country analyzed, Z is the system partition function, and finally $g(f_x)$ is a function dependent on f_x.

In general terms, by using (1), it is possible to find a set of Jacobi polynomials that fit the COVID-19 time series. To achieve this, there must be a random number generation mechanism with which the Jacobi polynomials can be selected from a Characterization Space (Section II-B). As mentioned in the previous section, it is well-known that (1) does not have a theoretical estimate that guarantees the generation of random numbers. Therefore, we propose a way to estimate the minimum difference between the Gaussian and the Generalized Boltzmann distributions by using the parameters μ and $\sigma^2 = \nu$, since the Normal distribution has mathematical guarantees in the estimation of random numbers. To carry out the aforementioned, we use Jeffrey's Divergence equation (2), which is a divergence measure between two probability distributions:

$$D_J(Q_x||P_x) = \int_x P_x \left(\frac{Q_x}{P_x} - 1 \right) \log \left(\frac{Q_x}{P_x} \right) dx. \tag{2}$$

To estimate a minimum difference between the distributions, we propose Theorem 2.1, which is the partial derivative with respect to the parameters μ and ν of (2).

Theorem 2.1: If Q_x and P_x are the Gaussian and the Generalized Boltzmann distributions respectively, and $\theta_i = [\mu, \nu]$, then the derivative of the measure $D_J(Q_x||P_x)$ with respect to θ_i is given by (3):

$$\frac{\partial D_J(Q_x||P_x)}{\partial \theta_i} = -\int_\chi \frac{P_x}{Q_x} \frac{\partial Q_x}{\partial \theta_i} dx$$
$$+ \int_\chi \left[1 + \log \left(\frac{Q_x}{P_x} \right) \right] \frac{\partial Q_x}{\partial \theta_i} dx. \tag{3}$$

Proof of Theorem 2.1: Deriving (2) with respect to θ we have (4):

$$\frac{\partial D_J(Q_x||P_x)}{\partial \theta_i} = \int_x \left(\frac{\partial Q_x}{\partial \theta_i} - \frac{\partial P_x}{\partial \theta_i} \right) \log \left(\frac{Q_x}{P_x} \right) dx$$
$$+ \int_x (Q_x - P_x) \frac{\partial Q_x}{\partial \theta_i} \frac{1}{Q_x} dx$$
$$- \int_x (Q_x - P_x) \frac{\partial P_x}{\partial \theta_i} \frac{1}{P_x} dx. \tag{4}$$

Due to the fact that the Boltzmann function does not depend on the θ_i parameter, then $\frac{\partial P_x}{\partial \theta_i} = 0$; therefore, Theorem 2.1 is proved.

From Theorem 2.1 we can estimate the μ parameter that has the minimum difference with respect to the behavior of the Generalized Boltzmann distribution, so we establish Theorem 2.2.

Theorem 2.2: If Q_x and P_x are the Gaussian and the Generalized Boltzmann distributions respectively, and μ is the Gaussian mean, then the minimum μ with respect to P_x is given by (5):

$$\mu = \frac{\int_\chi P_x x \, dx + \int_\chi \log[g(f_x)] Q_x x \, dx + \beta \int_\chi f_x Q_x x \, dx}{1 + \int_\chi \log[g(f_x)] Q_x \, dx + \beta \int_\chi f_x Q_x \, dx}. \tag{5}$$

Proof of Theorem 2.2: Given the result of Theorem 2.1 for the parameter μ and taking the fact that $\frac{\partial Q_x}{\partial \mu}$, we obtain (6):

$$\frac{\partial D_J(Q_x||P_x)}{\partial \mu} = -\frac{1}{\nu} \int_\chi P_x(x - \mu) dx$$
$$+ \frac{1}{\nu} \int_\chi Q_x(x - \mu) dx$$
$$+ \frac{1}{\nu} \int_\chi Q_x \log(Q_x)(x - \mu) dx$$
$$- \frac{1}{\nu} \int_\chi Q_x \log(P_x)(x - \mu) dx. \tag{6}$$

Substituting the results $\log(Q_x) = \log(\frac{1}{\sqrt{2\pi\nu}}) - \frac{(x-\mu)^2}{2\nu}$ and $\log(P_x) = -\log(Z) + \log[g(f_x)] + \beta f_x$ in (6), we get (7):

$$\frac{\partial D_J(Q_x||P_x)}{\partial \mu} = -\frac{1}{\nu} \int_\chi P_x(x - \mu) dx + \frac{1}{\nu} \int_\chi Q_x(x - \mu) dx$$
$$+ \frac{1}{\nu} \log(\frac{1}{\sqrt{2\pi\nu}}) \int_\chi Q_x(x - \mu) dx$$
$$- \frac{1}{2\nu^2} \int_\chi Q_x(x - \mu)^3 dx$$
$$+ \frac{\log(Z)}{\nu} \int_\chi Q_x(x - \mu) dx$$
$$- \frac{1}{\nu} \int_\chi \log[g(f_x)] Q_x(x - \mu) dx$$
$$- \frac{\beta}{\nu} \int_\chi f_x Q_x(x - \mu) dx. \tag{7}$$

The result of (7) can be simplified by the central moments of the Gaussian distribution given by $\int_\chi Q_x(x - \mu) dx = 0$,

$\int_\chi Q_x(x-\mu)^3 dx = 0$, and equating $\frac{\partial D_J(Q_x\|P_x)}{\partial \mu}$ to zero, we obtain (8):

$$\mu\left[\int_\chi \log[g(f_x)]Q_x dx + \beta\int_\chi f_x Q_x dx\right] - \int_\chi P_x x\,dx$$
$$- \int_\chi \log[g(f_x)]Q_x dx - \beta\int_\chi f_x Q_x x\,dx = 0. \quad (8)$$

Finally, solving for μ, we obtain the result of Theorem 2.2, thus being demonstrated.

The second important parameter is the estimation of variance ν since it provides diversification of the search for polynomials that best fit the COVID-19 time series, so the minimum variance is estimated from Theorem 2.1, resulting in Theorem 2.3.

Theorem 2.3: If Q_x and P_x are the Gaussian and the Generalized Boltzmann distributions respectively, and ν is the Gaussian variance, then the minimum ν with respect to P_x is given by (9):

$$\nu = \frac{\begin{array}{c}\int_\chi P_x(x-\mu)^2 dx + \int_\chi \log[g(f_x)]Q_x(x-\mu)^2 dx \\ + \beta\int_\chi f_x(x-\mu)^2 Q_x dx\end{array}}{\int_\chi \log[g(f_x)]Q_x dx + \beta\int_\chi f_x Q_x dx}. \quad (9)$$

Proof of Theorem 2.3: We take the result of Theorem 2.1 with respect to the parameter ν, and considering that $\frac{\partial Q_x}{\partial \nu} = \left[\frac{(x-\mu)^2}{2\nu^2} - \frac{1}{2\nu}\right]$, we get (10):

$$\frac{\partial D_J(Q_x\|P_x)}{\partial \nu} = -\frac{1}{2\nu^2}\int_\chi P_x(x-\mu)^2 dx + \frac{1}{2\nu}\int_\chi P_x dx$$
$$- \frac{1}{2\nu}\int_\chi Q_x dx + \frac{\log\left(\frac{1}{\sqrt{2\pi\nu}}\right)}{2\nu^2}$$
$$\cdot \int_\chi Q_x(x-\mu)^2 dx - \frac{\log\left(\frac{1}{\sqrt{2\pi\nu}}\right)}{2\nu}\int_\chi Q_x dx$$
$$- \frac{1}{4\nu^3}\int_\chi (x-\mu)^4 Q_x dx$$
$$+ \frac{1}{4\nu^2}\int_\chi (x-\mu)^2 Q_x dx$$
$$- \frac{\log(Z)}{2\nu^2}\int_\chi (x-\mu)^2 Q_x + \frac{\log(Z)}{2\nu}\int_\chi Q_x dx$$
$$- \frac{1}{2\nu^2}\int_\chi \log[g(f_x)](x-\mu)^2 Q_x dx$$
$$+ \frac{1}{2\nu}\int_\chi \log[g(f_x)]Q_x dx$$
$$- \frac{\beta}{2\nu^2}\int_\chi f_x(x-\mu)^2 Q_x dx$$
$$+ \frac{\beta}{2\nu}dx\int_\chi f_x Q_x dx. \quad (10)$$

Using the results of the central moments for the Gaussian distribution $\int_\chi (x-\mu)^2 Q_x dx = \nu$, $\int_\chi (x-\mu)^4 Q_x dx = 3\nu^2$, and $\frac{\partial D_J(Q_x\|P_x)}{\partial \nu} = 0$, (10) can be reduced to (11), so that

solving for ν (9) of Theorem 2.3 is proved:

$$\nu\left[\int_\chi \log[g(f_x)]Q_x + \beta\int_\chi f_x Q_x dx\right] - \int_\chi P_x(x-\mu)^2 dx$$
$$- \int_\chi \log[g(f_x)]Q_x(x-\mu)^2 dx - \beta\int_\chi f_x(x-\mu)^2 Q_x dx = 0. \quad (11)$$

The results of Theorem 2.2 and 2.3 are the theoretical elements of the Selection Operators proposed in our metaheuristic framework and will be connected to the Characterization Space to find the optimal fitting polynomial.

B. CHARACTERIZATION SPACE

From the set of Jacobi polynomials, which are described as a class of orthogonal polynomials from which other polynomials widely studied in the state-of-the-art are included, such as the Zernike, Legendre, Gegenbauer, and Chebyshev polynomials, it is possible to find, with certain mathematical guarantees, a Jacobi polynomial that fits the COVID-19 time series.

The Jacobi polynomials in their general form are defined as follows:

$$P_n^{(\alpha,\beta)}(z) = \binom{n+\alpha}{n}F\left(-n, n+\alpha+\beta+1; \alpha+1; \frac{1-z}{2}\right) \quad (12)$$

where $F(\cdot)$ is the hypergeometric function [36]. It should be mentioned that due to the nature of the COVID-19 time series, in this proposal we only use the real part of the Jacobi polynomials (as shown in (13)) to characterize the curves, whose result has its domain at $[-1, 1]$ for α and $\beta > -1$.

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}$$
$$\cdot \frac{d^n}{dx^n}\left[(x-1)^{\alpha+n}(1+x)^{\beta+n}\right]. \quad (13)$$

Some properties of the Jacobi polynomials that are used in the present work are listed below [36]:

1) The weighting function:

$$w_n(x) = (1-x)^\alpha(1+x)^\beta. \quad (14)$$

2) The boundary values:

$$P_n^{\alpha,\beta}(-1) = (-1)^n\binom{n+\beta}{n} \quad (15)$$

$$P_n^{\alpha,\beta}(1) = \binom{n+\alpha}{n}. \quad (16)$$

3) The derivative:

$$\frac{d}{dx}[P_n^{\alpha,\beta}(x)] = \frac{1}{2}(n+\alpha+\beta+1)P_{n-1}^{\alpha+1,\beta+1}(x). \quad (17)$$

In this work, we propose a Characterization Space, which is defined by the product of modulating coefficients with respect to the Jacobi polynomials of degree n, as shown in (18):

$$\mathscr{F}(t; \alpha, \beta) = \Sigma_{i=1}^\infty a_i P_i^{(\alpha,\beta)}(t). \quad (18)$$

We have changed the variable x by t for the Jacobi polynomials due to the time variable we are studying. Likewise, the a_i coefficients in (18) are calculated in a deterministic way by evaluating the input data (i.e., the COVID-19 time series) in the Jacobi polynomial. Subsequently, the result is orthogonalized using the Gram-Schmidt method, thus generating a new orthogonal space. In this space, the COVID-19 time series is approximated using the least squares technique, resulting in the adjustment coefficients in the orthogonal space. However, these coefficients are not associated with the orthogonal Jacobi base; therefore, they must be mapped to this base to obtain the polynomial mathematical model [50]. It is important to note that the α, β, and n parameters in (18) belong to an infinite space of combinatorial possibilities.

In the next section, we present our metaheuristic approach in which the Selection Operators (proposed in Section II-A) and the Characterization Space interacts in order to search a Jacobi polynomial that best fits the COVID-19 time series.

III. METAHEURISTIC FRAMEWORK

In this section, we will describe in detail the elements of our proposed metaheuristic approach called META-COVID19, specifying the discrete mathematical model of the theoretical results presented in Section II, the connection between the Selection Operators and the Characterization Space, as well as the algorithm description.

We start from the Characterization Space, as mentioned in the previous section, its parameters belong to an infinite space of combinatorial possibilities, which in computer practice leads to infinite execution times. To deal with that, we define an approximation of this space as shown below:

$$\hat{\mathscr{F}}(t; \alpha, \beta, n) = \Sigma_{i=1}^{n} a_i P_i^{(\alpha, \beta)}(t). \tag{19}$$

Thus, the space is constrained by the fundamental parameters α, β, and n (denoting the polynomial degree); in this way, different Jacobi polynomials can be selected from the Characterization Space by setting these parameters. To know how good the fit of each Jacobi polynomial $\hat{\mathscr{F}}(t; \alpha, \beta, n)$ is with respect to the COVID-19 time series $y(t)$, we use (20) as evaluation function.

$$f_x = \frac{1}{2}\left(y(t) - \hat{\mathscr{F}}(t; \alpha, \beta, n)\right)^2. \tag{20}$$

Regarding the Selection Operators μ and ν ((5) and (9) respectively), they must be discretized because the results of Theorems 2.2 and 2.3 involves continuous spaces, which cannot be computationally treated without an approximation. There exist different techniques for the integral approximations; however, we used the Monte-Carlo method as in [45], [49].

The approximation of the μ parameter using the Monte-Carlo method, weighed by Q_x, is presented below:

$$\mu \approx \frac{\frac{b-a}{m}\sum_{i=1}^{m}\left[\frac{1}{Z}g(f_{x_i})\exp[\beta f_{x_i}]x_i + \log[g(f_{x_i})]x_i + \beta f_{x_i}x_i\right]}{1 + \frac{b-a}{m}\sum_{i=1}^{m}\left[\log[g(f_{x_i})] + \beta f_{x_i}\right]} \tag{21}$$

where f_{x_i} is the performance of the ith Jacobi polynomial whose fundamental parameters $x_i = (\alpha, \beta, n)$ are in the domain $[a, b]$, m is the number of Jacobi polynomials, and $\beta = \frac{1}{f_{x_{best}}}$ [45]. Applying the same criteria of the previous equation, it is possible to approximate the ν parameter as shown below:

$$\nu \approx \frac{\sum_{i=1}^{m}\left[\begin{array}{l}\frac{1}{Z}g(f_{x_i})\exp[\beta f_{x_i}](x_i - \mu)^2 \\ + \log[g(f_{x_i})](x_i - \mu)^2 + \beta f_{x_i}(x_i - \mu)^2\end{array}\right]}{\sum_{i=1}^{m}\left[\log[g(f_{x_i})] + \beta f_{x_i}\right]}. \tag{22}$$

To clarify and avoid confusion, note that the Jacobi polynomial notation involves the use of a β parameter; this parameter is different from the β parameter of the generalized Boltzmann distribution from which the discretized Selection Operators ((21) and (22)) are derived.

Generally speaking, in our metaheuristic framework, the fundamental Jacobi parameters $x_i = (\alpha, \beta, n)$ are set with random numbers sampled from a distribution with μ and ν parameters. Then, during the iterative process, μ and ν are updated using (21) and (22) respectively. Note that in the calculation m Jacobi polynomials are needed, these polynomials should be the top-ranked at each iteration; in this way, the convergence of the algorithm is guaranteed, thus allowing the selection of new Jacobi polynomials whose adjustment will be increasingly closer to the behavior of the COVID-19 time series.

A. ALGORITHM

At the start of the algorithm, an initial Jacobi subset J_j^i is selected from the Characterization Space, this subset is made up by $j = 0, 1, \ldots, S$ Jacobi polynomials $\hat{\mathscr{F}}(\cdot)$ with different α, β and n parameters, which are initially set using uniform random numbers $\mathcal{U}(\mu^0, \nu^0)$, where $\mu^0 = 0$ and $\nu^0 = 1$. Then, each Jacobi polynomial from J_j^i is evaluated using (20) and sorted in decreasing order (to reduce computational cost) according to their performance; note that the first element J_0^i of the sorted Jacobi subset, is the best polynomial $\hat{\mathscr{F}}_{best}$ at iteration i, whose performance is $f_{x_{best}}$. Subsequently, m Top-ranked Jacobi polynomials from J_j^i are chosen, thus creating the subset T^i, and with these, the Selection Operators are updated via (21) and (22) respectively. Finally, a New Jacobi subset J_j^{i+1} is selected from the Characterization Space by setting the parameters α, β and n with Gaussian random numbers $\mathcal{N}(\mu^{i+1}, \nu^{i+1})$. The META-COVID19 algorithm remains in this converging cycle until a stopping condition is met or a predefined number of iterations is reached. Figure 2 shows the META-COVID19 flowchart, and the complete pseudocode is presented in Algorithm 1.

IV. EXPERIMENTAL SETUP

This section describes the experimental setup followed in this work to test the effectiveness of our metaheuristic framework, including the datasets and parameter settings.
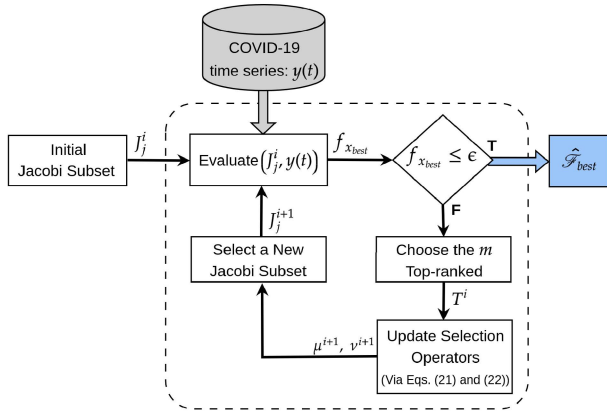
FIGURE 2. META-COVID19 flowchart to automatically characterize COVID-19 time series.

Algorithm 1 META-COVID19 Pseudocode

Require: S: Jacobi subset size, $y(t)$: COVID-19 time series
Ensure: $\hat{\mathscr{F}}_{best}$: Best Jacobi
1:   $i \leftarrow 0$: Current iteration
2:   $J_j^i \leftarrow$ InitialJacobiSubset(S, μ^0, ν^0)
3:   Evaluate$(J_j^i, y(t))$
4:   $\hat{\mathscr{F}}_{best} \leftarrow J_0^i$: Get Best Jacobi
5:   **while** $(f_{x_{best}} \leq \epsilon)$ **do**
6:     $T^i \leftarrow$ Top-ranked (J_j^i, m)
7:     $\mu^{i+1}, \nu^{i+1} \leftarrow$ UpdateSelectionOperators(T^i)
8:     $J_j^{i+1} \leftarrow$ NewJacobiSubset$(S, \mu^{i+1}, \nu^{i+1})$
9:     Evaluate$(J_j^{i+1}, y(t))$
10:    $\hat{\mathscr{F}}_{best} \leftarrow J_0^{i+1}$: Get Best Jacobi
11:    $i \leftarrow i + 1$
12: **end while**
13: **return** $\hat{\mathscr{F}}_{best}$

A. DATASETS

We used the compilation of datasets presented by Ortiz-Ospina *et al.* (2020) maintained by Our World in Data, and posted at https://github.com/owid/covid-19-data/tree/master/public/data. It is updated daily and includes data on daily confirmed cases, deaths, hospitalizations, testing, and vaccinations as well as other variables of potential interest [51].

For this work, we will focus on the time series of daily confirmed cases reported in nine countries: the United States of America (USA), Spain, United Kingdom (UK), India, Chile, Peru, Brazil, Russia, and Mexico. However, reported data of deaths, hospitalizations, tests, and vaccinations, can also be characterized by our framework. The only restriction is that the data must be time series.

It is worth mentioning that the experimentation was carried out using the data available at the date of creation of this article (May 2021), but it can be applied with the data collected at any time. In Figure 3 are shown some examples of time series

of new daily cases confirmed as positive to SARS-CoV2, which for research purposes will offer a better perspective on the behavior of the disease, and will show the power of our proposal to characterize time series; allowing adequate decision-making by the health, economic, and social sectors.

B. PARAMETER SETTINGS

The settings of the META-COVID19 parameters are shown below, to carry out the corresponding experimentation. The Jacobi polynomial subset is made up of $S = 50$ elements. To compute the Selection Operators μ and ν, we recommend that the subset T^i be at least 40% of the size of the subset J^i, since this allows conducting the search for the best polynomial in an adequate way. Using 100% of subset J^i is possible; however, this leads to higher computational cost. On the other hand, if less than 40% were used, the convergence of the algorithm would be slower. Specifically, for these experiments $m = 20$ top-ranked polynomials are contained in T^i. In order to delimit the size of the Characterization Space to be explored, the following domains are established: $n = [1, 60]$, $\alpha = [-1, 1]$ and $\beta = [-1, 10]$. It should be noted that the α and β domains were established from the mathematical explanation in Section II-B. Likewise, after a stage of meticulous experimentation, it was observed that polynomials of degree greater than $n = 60$ imply an expensive computational cost, and the adjustment obtained with respect to others of a lesser degree is not significant. In every experiment, the convergence criterion is set to $\epsilon \leq 1 \times 10^{-3}$, and the number of iterations is set to 50. Therefore, the limit of function calls is 2500.

To ensure the consistency of the experimental results, each experiment is executed 31 times independently, and the best Jacobi Polynomial found each time is recorded. Finally, the median and standard deviation of these 31 experiments are reported in Table 2. The experimentation was carried out on a conventional computer with an Ubuntu operating system, an Intel i5 processor, and 8GB of RAM. The META-COVID19 was developed in the Python programming language under a non-parallel programming paradigm. The following section shows the results obtained under these parameter settings.

C. RESULTS

Table 1 shows the best parameters of the Jacobi polynomial whose performance was the median of the 31 experiments, as well as their corresponding fit value f_x for each country. As can be seen, the degree of all Jacobi polynomials did not surpass 60. Likewise, the polynomial found for India was the greater with 59, and the smallest was for Chile with a degree of 21. Finally, for these experiments, the best polynomial found was for India with a fit value of 1.20E-2.
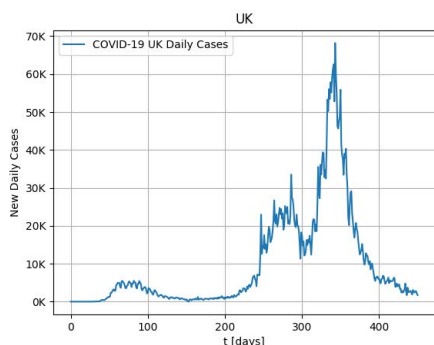
To graphically exemplify the performance of the proposed metaheuristic, we have selected two representative examples reported in Table 1. Figure 4 shows the average fit of 31 independent experiments (orange line) found by META-COVID19 for the USA (Figure 4a) and India (Figure 4b),
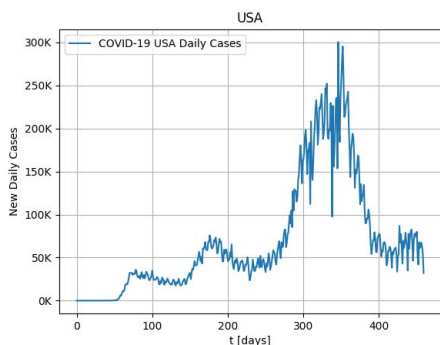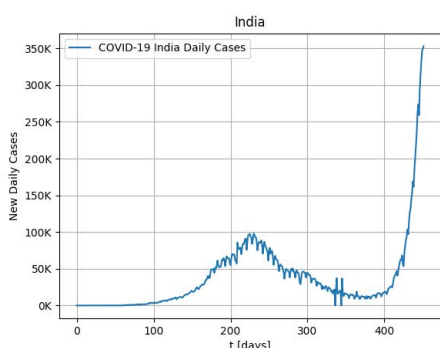
(a) USA Time Series
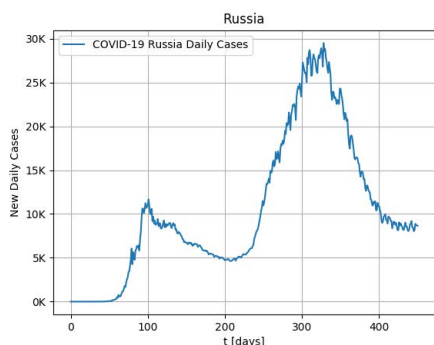
(b) Russia Time Series

(c) UK Time Series

(d) India Time Series

FIGURE 3. Time series of new daily cases confirmed as positive for SARS-CoV2, over a period of more than 400 days, reported in different countries.

in contrast to the time series of daily confirmed cases (line shaded in blue) for each country, respectively; likewise, a subplot of the area of greatest variability of the experiments

TABLE 1. The best polynomial parameters and their fit value f_x for each country.

| Country	Best Polynomial Parameters			
	α	β	n	f_x
USA	0.0612	-0.0002	51	4.83E-2
Spain	-0.1214	0.0505	33	7.46E-2
UK	-0.0597	-0.0123	57	3.26E-2
India	0.0115	0.0019	59	1.20E-2
Chile	-0.0197	0.0676	21	6.54E-2
Peru	0.0264	0.0009	45	1.12E-1
Brazil	0.3842	-0.0501	23	1.37E-1
Russia	-0.0180	-0.0452	58	1.71E-2
Mexico	0.1200	0.0174	29	8.91E-2

carried out is shown (shaded yellow region). As can be seen, the solutions found by the metaheuristics show a precise fit for the two scenarios presented, even for India, which has a considerable change of direction between the first and second waves of infection. In general, the results presented in Figure 4 show that our metaheuristic approach is capable of achieving an adjustment in each of the time series used in this study. However, it is worth mentioning that due to the different data capture processes in the health departments of the countries, the information presented directly influences the result of the approximation. Based on the presented results, it can be said that the Selection Operators proposed Theorems 2.1 y 2.2, were able to automatically find suitable Jacobi polynomials for the characterization of COVID-19 spread.

### D. STATISTICAL CONTRAST AGAINST ARTIFICIAL NEURAL NETWORKS

To evidence the performance of META-COVID19, we present a statistical contrast between our proposal and a methodology for time series fitting: Artificial Neural Networks (ANN). The ANN was applied on the data described in Section IV-A, with an architecture of 4 hidden layers, 100 neurons in each layer, a ReLU (Rectified Linear Unit) activation function, and to train the Neural Network, Adam's optimization algorithm was used (sequentially) with 1000 epochs. The implementation was carried out in Keras, which is an open-source software library that provides a Python interface for ANNs.

For a comparison with META-COVID19, 31 experiments were performed independently; the median and standard deviation of these 31 experiments are reported in Table 2. The experiments ran on a computer with the following characteristics: Ubuntu operating system, Intel i7 processor, and 16GB of RAM.

Table 2 presents the median fit errors for 31 independent experiments, as well as their corresponding standard deviation; the best results in the table are highlighted in a bold typeface, whenever two or more algorithms produced the same median value, the result with the smallest standard deviation is highlighted. We can see that META-COVID19 and ANN coincide in obtaining results in the order of 1E-2. Regarding the reported standard deviation for META-COVID19, we can
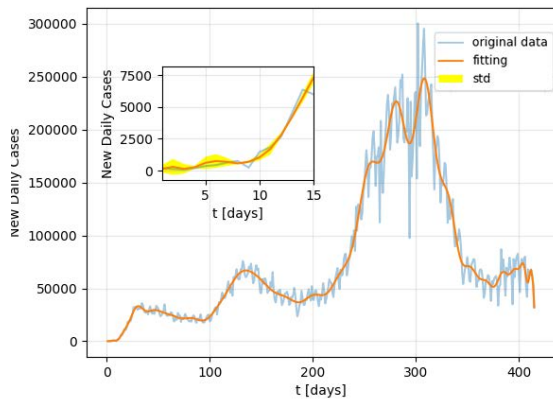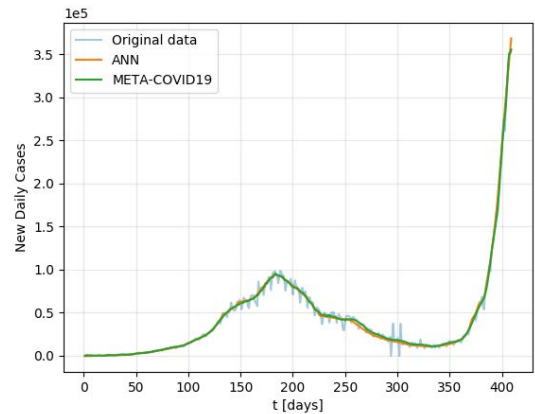
(a) USA Time Series
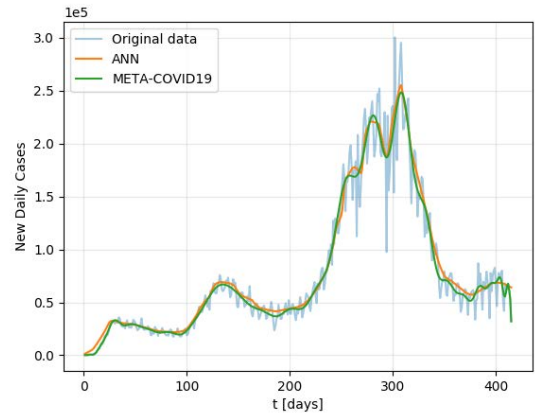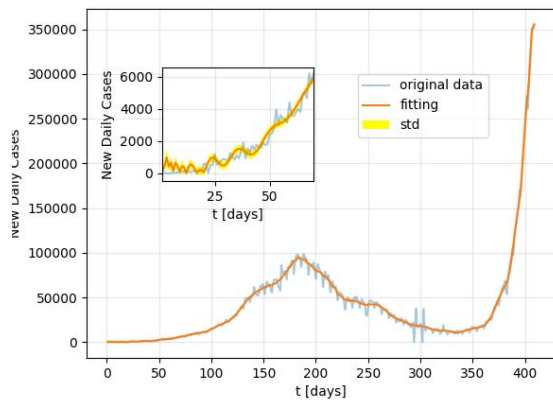
(b) India Time Series

FIGURE 4. Average fit of 31 independent experiments (orange line) found by META-COVID19 for the USA and India, in contrast to the time series of daily confirmed cases (lineshaded in blue) for each country.

(a) USA Time Series

(b) India Time Series

FIGURE 5. Comparison of time series adjustment of daily confirmed cases in the USA and India. The blue line represents the original data, and the orange and green lines show the fit of ANN and META-COVID19 correspondingly.

TABLE 2. Median ± Standard Deviation of the fit errors obtained by META-COVID19 and ANN for 31 independent experiments.

Country	META-COVID19	ANN
USA	**4.83E-2 ± 9.83E-6**	4.92E-2 ± 3.8E-3
Spain	**7.46E-2 ± 0**	7.6E-2 ± 5.6E-3
UK	3.26E-2 ± 3.39E-5	**3.1E-2 ± 8.4E-3**
India	**1.21E-2 ± 5.78E-5**	1.43E-2 ± 3.3E-3
Chile	**6.54E-2 ± 0**	6.55E-2 ± 2.2E-3
Peru	**1.12E-1 ± 3.76E-5**	1.12E-1 ± 2.9E-3
Brazil	1.37E-1 ± 0	**1.31E-1 ± 3.27E-3**
Russia	**1.72E-2 ± 3.70E-5**	2.1E-2 ± 4.4E-3
Mexico	8.91E-2 ± 0	**8.36E-2 ± 3.6E-3**

observe values close to or equal to 0; this means that our metaheuristic approach is able to consistently find a Jacobi polynomial whose performance is similar in all 31 experiments.

In Figure 5, we provide the graphical results obtained by META-COVID19 and ANN in a single experiment. Figure 5a shows the fit for the USA time series, while Figure 5b shows the case for India. We observe that both methodologies achieve a similar smoothed approximation of the time series

of daily reported cases, suggesting that META-COVID19 could be as good as an ANN for time series fitting purposes.

To determine if there are differences in the performance of META-COVID19 and ANN in terms of the fit error obtained, we carried out a Wilcoxon signed-rank test [52]; this is a non-parametric test to compare the mean rank of two related samples and determine if there are significant differences between them [52].

We performed a two-tailed Wilcoxon signed-rank test on the data presented in Table 2, using $n = 8$ non-zero differences, at a significance level of $\alpha = 0.05$. The null hypothesis H_0 is that META-COVID19 and ANN perform equally, and the alternative hypothesis H_1 is that they perform differently. The test was carried out in *RStudio*, through the *wilcox.test* function available in the *stats* package, obtaining a Wilcoxon test statistic: $W = 17$, a critical value for W: $C_W = 3$, and finally, a *p*-value = 0.945313. Since $W > C_W$ at $n = 8$, W is in the 95% acceptance region. This can be corroborated with a *p*-value $> \alpha$, therefore H_0 is accepted. This means that the

probability of a type I error (i.e., rejecting a correct H_0) is high: 0.9453 (94.53%).

From the results obtained by the Wilcoxon signed-rank test, we can conclude that for this experimental design, META-COVID19 and ANN have similar performance for time series fitting. However, META-COVID19 is capable of finding polynomial models that allow us to design propagation indicators for COVID-19 and other types of diseases: in Section V, we present some examples of application; while in Section IV-E the advantages of our proposal in contrast to other methodologies are discussed.

E. DISCUSSION

It is essential to emphasize that the objective of this work is to mathematically characterize the COVID-19 time series through an explicit polynomial model, thus providing descriptive information on the spread of the pandemic.

In Section IV-D, we compared META-COVID19 with ANN, concluding that for the proposed experimental design, they have similar performance. However, there are differences between them in terms of the information obtained. META-COVID19 generates Jacobi polynomial models that have known mathematical properties (see Section II-B), which allows offering analytical descriptions. In contrast, attributes such as the derivative can be obtained with an ANN (such as Sparse Functional Multilayer Perceptron, and Long Short-Term Memory) [53], and a specific network could even be trained to calculate differential equations [54]; however, to obtain more mathematical or statistical information about the time series, specific neural networks would have to be trained, which implies programming time and computational cost, not to mention multiple hyperparameters need to be configured for the approach to work properly. Therefore, we consider that an ANN is an excellent tool to perform time series fitting, but it might not be the appropriate technique to mathematically characterize an epidemic spread, this is where our proposal has its main advantage.

In addition to ANNs, we identify other two main approaches to perform time series fitting.

- Numerical: Regression and interpolation methods [55], [56].
- Symbolic regression: Genetic Programming (GP) approaches [57], [58].

Regarding regression and interpolation methods, they can offer an explicit analyzable function. However, there is the problem that the parameters for each method must be known [55], [56], which would imply carrying out an exhaustive study by hand, and that is precisely what we are avoiding with our proposal, by implementing a metaheuristic framework capable of finding a polynomial on a Jacobi space that fits the time series automatically. Therefore, we consider that benchmarking against those methods is inappropriate.

On the other hand, Symbolic Regression with GP approaches involves metaheuristic algorithms, can be applied to time series, and offers an explicit function. It is well known that these approaches work well for noise-free and low-information synthetic time series, but their performance decreases considerably when dealing with real-world time series. To address this issue, in [57] the time series is divided into windows to adjust each of them independently, and in [58], a scalable chromosome encoding scheme that is capable of representing multiple solutions simultaneously is used. However, these approaches lead to excessive computational cost (up to 2 million fitness evaluations [58]) and the construction of several explicit mathematical models whose mathematical guarantees might not be known, so obtaining useful information from them can be complicated. Furthermore, there is an inherent problem of methods based on GP called ''Bloat'' [57], exhausting computer resources, not counting the large number of hyperparameters that must be configured. The computational cost of our proposal is considerably low, in this experimentation the maximum limit of fitness evaluations was 2500, reaching an error close to 1E-3. Therefore, we consider that GP might not be an adequate technique to characterize an epidemic spread.

This discussion intends to show the strengths that our proposal has in contrast to other state-of-the-art methodologies. META-COVID19 only needs a single parameter, which is the size of the subset of the Jacobi polynomials, unlike the symbolic regression with GP approaches and ANNs, where multiple hyperparameters must be tuned, which implies either empirical experimentation or a priori knowledge of the problem. Furthermore, it is important to mention that with the polynomial models shown in Section IV-C, multiple information can be inferred without the need to perform the time series fitting again. In the next section, we show three application examples, which show how information that characterizes the epidemic spread can be extracted mathematically.

V. MATHEMATICAL MODEL APPLICATION

To show the applicability of the polynomial mathematical models proposed by our metaheuristic framework, in this section different examples are carried out, where useful information can be known.

A. EXAMPLE OF APPLICATION

1) EXAMPLE 1

We have chosen the mathematical model for the USA, as it shows three waves of contagion and allows us to extract relevant information. In Appendix, we show the complete Jacobi polynomial (summarized in (23)) derived from the fundamental parameters reported in Table 1 which were obtained by META-COVID19.

$$\hat{\mathscr{F}}(t; 0.0612, -0.0002, 51). \tag{23}$$

As a direct application, we can calculate the derivative of this mathematical model with respect to time, using the derivative property of Jacobi polynomials shown in (17). The derivative of $\hat{\mathscr{F}}$ provides descriptive information about the change of propagation of new cases with respect to the elapsed time. In Figure 6, we show the derivative of (23) respect to t.
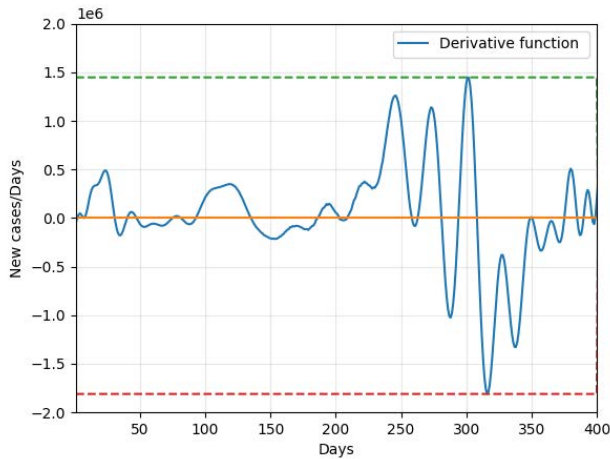
FIGURE 6. Derivative of $\hat{\mathscr{F}}$ with respect to t in the domain of [1,400].
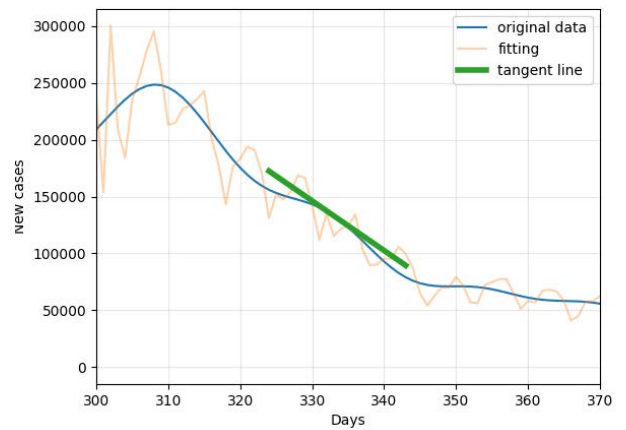
TABLE 3. Results of the derivative of $\hat{\mathscr{F}}$ with respect to t for the USA, showing the different slope changes.

| Figure	t [Days]	$\frac{d\hat{\mathscr{F}}(t;\alpha,\beta,n)}{dt}$
7a	333	-928457.659
7b	248	1182487.163
7c	107	43925.92

In the solid orange line, we observe the zeros of the derivative, which represent the resulting maxima and minima in our mathematical model. If we focus on the two regions (dotted lines) above and below the origin, these regions provide the information of the amplitude of change of new cases with respect to time—In the particular case of the USA, under the mathematical model a slightly greater amplitude in the negative region was found compared to the positive one—which in general terms allows us to observe that the social and political behavior has led the new cases downward against the propagation. Likewise, it is interesting to note that when the blue line coincides with the solid orange line, for a certain time, it can be said that this epidemic cycle has ended.

Figure 7 shows different scenarios of the change of direction by means of the slope of the tangent line to the curve of new cases per day. In Table 3, we present the results of these slopes on representative days. Note that, Figure 7a has a negative slope, which is a desirable value in mitigating the spread of SARS-CoV2, while in Figure 7b, the direction of the slope is positive, alerting of an increase of new cases per day; the magnitude of the slope is a parameter of information about the rate of increases of daily reported cases per day. Finally, Figure 7c shows a case where the slope tends to be horizontal, resulting from a plateau daily contagion spread; such plateau is synonymous with partial control of the spread of the virus.

Examples 2 and 3 show the advantages of using our proposal since different mathematical analyzes can be performed using Jacobi polynomials properties [36]. Particularly, we focus on elements of the statistical inference, which is essential in the study of pandemics spread [39], [40].

(a) A negative slope of change of confirmed cases with respect to t.
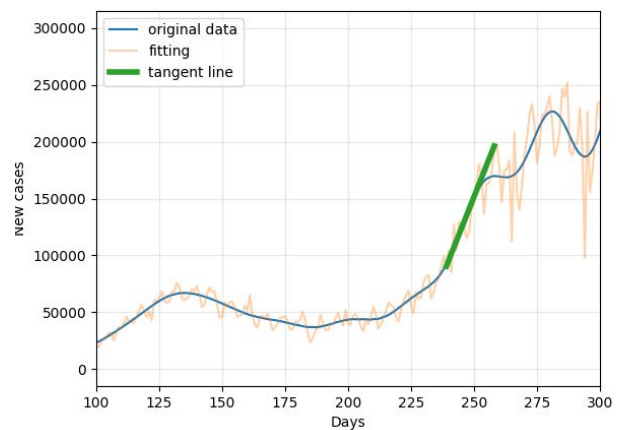
(b) A positive slope of change of confirmed cases with respect to t.
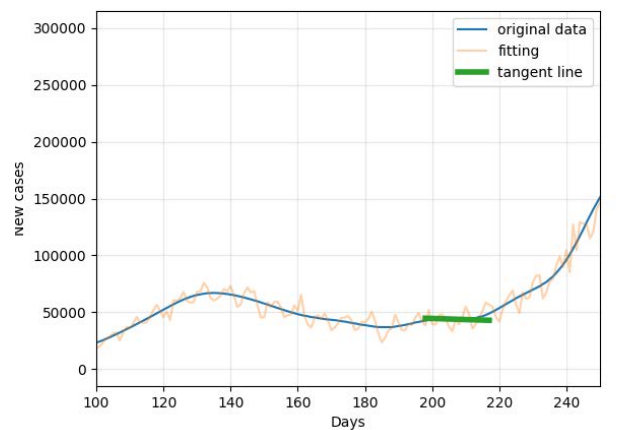
(c) An unchanged slope of confirmed cases with respect to t.

FIGURE 7. Example of the behavior of the derivative of $\hat{\mathscr{F}}$ with respect to t for the USA, at different changes of direction.

2) EXAMPLE 2

We start from a model $\hat{\mathscr{F}}(t, \alpha, \beta, n)$ proposed by META-COVID19. If this model is normalized by (24) to (25) in the

domain $t \in [c, d]$, then (26) gives the Probability Distribution Function of the time series:

$$A \int_c^d \hat{\mathscr{F}}(t, \alpha, \beta, n) dt = 1 \tag{24}$$

$$A = \frac{1}{\sum_i a_i \int_c^d P_n^{\alpha,\beta}(t)\, dt} = \frac{1}{\sum_i a_i \int_c^d P_n^{\alpha,\beta}(t)\, dt}$$

$$= \frac{\alpha + \beta + n}{2 \sum_i a_i P_{n+1}^{\alpha-1,\beta-1}(t)\big|_c^d} \tag{25}$$

$$\text{PDF}(t) = \frac{\alpha + \beta + n}{2 \sum_i a_i P_{n+1}^{\alpha-1,\beta-1}(t)\big|_c^d} \hat{\mathscr{F}}(t, \alpha, \beta, n) \tag{26}$$

where:

$$\int_a^b P_n^{\alpha,\beta}(t) dt = \frac{2 P_{n+1}^{\alpha-1,\beta-1}(t)\big|_a^b}{\alpha + \beta + n}, \quad \text{for } \alpha, \beta, n \neq 0.$$

A direct application of (26) is to calculate the cumulative probability $\hat{P}$ given by (27) in the domain $f, g \in [c, d]$, where clearly $\hat{P}(f \leq t \leq g)$ provides probabilistic information of contagions in a given period, thus generating quantitative information from the model without involving extra processing:

$$\hat{P}(f \leq t \leq g) = \int_f^g \text{PDF}(t) dt$$

$$= \frac{\sum_i a_i P_n^{\alpha,\beta}(t)\big|_f^g}{\sum_i a_i P_{n+1}^{\alpha-1,\beta-1}(t)\big|_c^d}. \tag{27}$$

3) EXAMPLE 3

One of the most important statistical elements is the expected value, which we can calculate for a given time window $t \in [c, d]$ of the virus spread. From (26) the expected value is calculated in terms of the model found by META-COVID19 and in (28) we show the mathematical result:

$$E[t] = \int_c^d t \cdot \text{PDF}(t) dt$$

$$= \frac{\alpha + \beta + n}{2 \sum_i P_{n+1}^{\alpha-1,\beta-1}\big|_a^b} \sum_i a_i \int_c^d t \cdot P_n^{\alpha,\beta}(t) dt \tag{28}$$

where:

$$\int_c^d t \cdot P_n^{\alpha,\beta}(t) dt = \frac{2 \left[t \cdot P_{n+1}^{\alpha-1,\beta-1}(t)\big|_c^d - \frac{2 P_{n+2}^{\alpha-2,\beta-2}(t)\big|_a^b}{\alpha+\beta+n-1} \right]}{\alpha + \beta + n}. \tag{29}$$

Therefore, (30) estimates the expected value of daily contagion spread:

$$E[t] = \frac{\sum_i a_i t_i P_{n+1}^{\alpha-1,\beta-1}(t)\big|_c^d - 2 \frac{P_{n+2}^{\alpha-2,\beta-2}(t)\big|_c^d}{\alpha+\beta+n-1}}{\sum_i a_i P_{n+1}^{\alpha-1,\beta-1}(t)\big|_c^d}. \tag{30}$$

Equations (26), (27) and (30) are examples of quantitative estimators obtained from the model fitted with our methodology and provide descriptive information about the virus

spread. Finally, from this analysis, the assumptions can be deepened to generate more inferential statistics and support assertive decision-making with mathematical foundations, which could be combined and strengthened with other computational models in favor of society.

VI. LIMITATIONS OF THE STUDY

In this research, the mathematical analysis was limited only to the polynomial models found by META-COVID19, which were adjusted to the time series available at the date of creation of the article (but it can be applied with the data collected at any time). Likewise, our proposed metaheuristic is bound to the Jacobi polynomial family because it includes other polynomials widely studied in the state-of-the-art, such as Zernike, Legendre, Gegenbauer, and Chebyshev.

For the experimentation, the time series of new daily cases confirmed as positive for SARS-CoV2 were used since they show greater variability of the temporal data compared to the time series of deaths, which allows us to show the efficacy of our proposal. It is worth mentioning that due to the different data capture processes in the health departments of the countries studied, the information presented directly influences the result of the approximation.

Finally, the objective of our proposal is not the explicit treatment of information to predict the short-term spread of COVID-19. Instead, we focus on a mathematical characterization that allows us to obtain information to generate better strategies for decision-making and thus prevent the future spread of this and other diseases.

VII. CONCLUSION

A novel metaheuristic framework called META-COVID19 that merges the Generalized Boltzmann distribution and the orthogonal Jacobi polynomials were proposed to automatically perform a data-driven mathematical characterization of the spread of COVID-19, without prior knowledge of the data and without involving a human expert. We use our framework to characterize the time series of new daily cases confirmed as positive for SARS-CoV2, over a period of more than 400 days, reported in nine countries. The results presented in Section IV-C show that our metaheuristic approach is capable of adjusting to the time series with slight variability, reaching an error close to 1E-3 with a standard deviation less than 6E-5.

META-COVID19 only needs one fundamental parameter, which is the number of Jacobi polynomials to be analyzed at each iteration, providing an algorithm with simplicity in its technical treatment. Furthermore, the Selection Operators proposed in Section II-A, allows finding Jacobi polynomials at a low computational cost since only 2500 function calls are needed to perform the search in the proposed Characterization Space. In this sense, our metaheuristic framework can be used in low-resource computer architectures.

According to the Wilcoxon signed-rank test performed in Section IV-D, our proposal is as good as other state-of-the-art techniques such as Artificial Neural Networks (ANNs)

for time series fitting purposes. However, it is important to note that ANNs might not provide an explicit mathematical model from which more information can be obtained. On the other hand, methodologies such as symbolic regression with Genetic Programming approaches can build explicit mathematical models, but these might not have known mathematical guarantees, and their construction implies a computational cost much higher than that of META-COVID19. Finally, we have shown in Section V the applicability of the polynomial mathematical models found by META-COVID19, where the assumptions can be deepened to generate more inferential statistics and support assertive decision-making with mathematical foundations.

As future work, we will carry out a study on the time series of more countries, performing a segmentation by cities in different time windows. Likewise, we will study in more detail the properties of the polynomial models found. A new outbreak or even new pandemics may reach the world shortly. Therefore, having useful information on the propagation behavior provided by different approaches such as META-COVID19, will allow generating better strategies for decision-making.

APPENDIX
ORTHOGONAL JACOBI POLYNOMIAL OF USA

The complete Jacobi polynomial found by META-COVID19 that characterize the USA time series for $t = 422$ days, with an approximation error of $g(f_x) = 4.83\text{E-}2$, is shown in (31).

$$\hat{\mathscr{F}}(t, 0.0612, -0.0002, 51)$$
$$= 0.2398919806 P_0^{(\alpha,\beta)} + 0.19580159 P_1^{(\alpha,\beta)}$$
$$-0.1840274078 P_2^{(\alpha,\beta)} - 0.2836806116 P_3^{(\alpha,\beta)}$$
$$-0.1918465289 P_4^{(\alpha,\beta)}$$
$$+0.1273633146 P_5^{(\alpha,\beta)} + 0.1383303565 P_6^{(\alpha,\beta)}$$
$$-0.0347301828 P_7^{(\alpha,\beta)}$$
$$-0.3676392751 P_8^{(\alpha,\beta)} - 0.2709421146 P_9^{(\alpha,\beta)}$$
$$-0.0799113774 P_{10}^{(\alpha,\beta)}$$
$$-0.0777634871 P_{11}^{(\alpha,\beta)} - 0.1771552104 P_{12}^{(\alpha,\beta)}$$
$$-0.0997154247 P_{13}^{(\alpha,\beta)}$$
$$-0.2160231213 P_{14}^{(\alpha,\beta)} - 0.1752544954 P_{15}^{(\alpha,\beta)}$$
$$-0.2412128278 P_{16}^{(\alpha,\beta)}$$
$$-0.1634194009 P_{17}^{(\alpha,\beta)} - 0.230539758 P_{18}^{(\alpha,\beta)}$$
$$-0.0961695459 P_{19}^{(\alpha,\beta)}$$
$$-0.1919655918 P_{20}^{(\alpha,\beta)} - 0.2306250073 P_{21}^{(\alpha,\beta)}$$
$$-0.2769133376 P_{22}^{(\alpha,\beta)}$$
$$-0.1617292943 P_{23}^{(\alpha,\beta)} - 0.1840142091 P_{24}^{(\alpha,\beta)}$$
$$-0.1315583219 P_{25}^{(\alpha,\beta)}$$
$$-0.209756872 P_{26}^{(\alpha,\beta)} - 0.1777921369 P_{27}^{(\alpha,\beta)}$$
$$-0.2271212321 P_{28}^{(\alpha,\beta)}$$

$$-0.1438255534 P_{29}^{(\alpha,\beta)} - 0.1727163355 P_{30}^{(\alpha,\beta)}$$
$$-0.1186929923 P_{31}^{(\alpha,\beta)}$$
$$-0.1448777975 P_{32}^{(\alpha,\beta)} - 0.145347906 P_{33}^{(\alpha,\beta)}$$
$$-0.1683693389 P_{34}^{(\alpha,\beta)}$$
$$-0.1029743957 P_{35}^{(\alpha,\beta)} - 0.071850483 P_{36}^{(\alpha,\beta)}$$
$$-0.0449190244 P_{37}^{(\alpha,\beta)}$$
$$-0.1284396092 P_{38}^{(\alpha,\beta)} - 0.1532636351 P_{39}^{(\alpha,\beta)}$$
$$-0.1297923669 P_{40}^{(\alpha,\beta)}$$
$$-0.0405234397 P_{41}^{(\alpha,\beta)} + 0.0052702233 P_{42}^{(\alpha,\beta)}$$
$$-0.0359805926 P_{43}^{(\alpha,\beta)}$$
$$-0.1199310014 P_{44}^{(\alpha,\beta)} - 0.1315914891 P_{45}^{(\alpha,\beta)}$$
$$-0.051470919 P_{46}^{(\alpha,\beta)}$$
$$+0.0279153584 P_{47}^{(\alpha,\beta)} + 0.0094849109 P_{48}^{(\alpha,\beta)}$$
$$-0.050232562 P_{49}^{(\alpha,\beta)}$$
$$-0.0807345798 P_{50}^{(\alpha,\beta)} - 0.0499763644 P_{51}^{\alpha,\beta}. \quad (31)$$

REFERENCES

[1] S. A. Alchon, *A Pest in the Land: New World Epidemics in a Global Perspective*. Albuquerque, NM, USA: UNM Press, 2003, pp. 109–143.

[2] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, "COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses," *J. Adv. Res.*, vol. 24, pp. 91–98, Mar. 2020.

[3] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proc. Roy. Soc. London A*, vol. 115, pp. 700–721, Aug. 1927.

[4] A. Huppert and G. Katriel, "Mathematical modelling and prediction in infectious disease epidemiology," *Clin. Microbiol. Infection*, vol. 19, no. 11, pp. 999–1005, Nov. 2013.

[5] W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li, "Why is it difficult to accurately predict the COVID-19 epidemic?" *Infectious Disease Model.*, vol. 5, pp. 271–281, Mar. 2020.

[6] T. Carletti, D. Fanelli, and F. Piazza, "COVID-19: The unreasonable effectiveness of simple models," *Chaos, Solitons Fractals, X*, vol. 5, Mar. 2020, Art. no. 100034.

[7] E. B. Postnikov, "Estimation of COVID-19 dynamics 'on a back-of-envelope': Does the simplest SIR model provide quantitative parameters and predictions?" *Chaos, Solitons Fractals*, vol. 135, Jun. 2020, Art. no. 109841.

[8] I. Cooper, A. Mondal, and C. G. Antonopoulos, "A SIR model assumption for the spread of COVID-19 in different communities," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110057.

[9] B. Malavika, S. Marimuthu, M. Joy, A. Nadaraj, E. S. Asirvatham, and L. Jeyaseelan, "Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models," *Clin. Epidemiol. Global Health*, vol. 9, pp. 26–33, Jan. 2021.

[10] R. F. Reis, B. de Melo Quintela, J. de Oliveira Campos, J. M. Gomes, B. M. Rocha, M. Lobosco, and R. W. dos Santos, "Characterization of the COVID-19 pandemic and the impact of uncertainties, mitigation strategies, and underreporting of cases in South Korea, Italy, and Brazil," *Chaos, Solitons Fractals*, vol. 136, Jul. 2020, Art. no. 109888.

[11] F. Koike and N. Morimoto, "Supervised forecasting of the range expansion of novel non-indigenous organisms: Alien pest organisms and the 2009 H1N1 flu pandemic," *Global Ecol. Biogeogr.*, vol. 27, no. 8, pp. 991–1000, Apr. 2018.

[12] S. Anno, T. Hara, H. Kai, M.-A. Lee, Y. Chang, K. Oyoshi, Y. Mizukami, and T. Tadono, "Spatiotemporal dengue fever hotspots associated with climatic factors in Taiwan including outbreak predictions based on machine-learning," *Geospatial Health*, vol. 14, no. 2, pp. 1–12, Nov. 2019.

[13] S. S. Chenar and Z. Deng, "Development of genetic programming-based model for predicting oyster norovirus outbreak risks," *Water Res.*, vol. 128, pp. 20–37, Jan. 2018.

[14] R. Liang, Y. Lu, X. Qu, Q. Su, C. Li, S. Xia, Y. Liu, Q. Zhang, X. Cao, Q. Chen, and B. Niu, "Prediction for global African swine fever outbreaks based on a combination of random forest algorithms and meteorological data," *Transboundary Emerg. Diseases*, vol. 67, no. 2, pp. 935–946, Mar. 2020, doi: 10.1111/tbed.13424.

[15] S. Mezzatesta, C. Torino, P. D. Meo, G. Fiumara, and A. Vilasi, "A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis," *Comput. Methods Programs Biomed.*, vol. 177, pp. 9–15, Aug. 2019.

[16] A. I. Saba and A. H. Elsheikh, "Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks," *Process Saf. Environ. Protection*, vol. 141, pp. 1–8, Sep. 2020.

[17] P. Melin, J. C. Monica, D. Sanchez, and O. Castillo, "Multiple ensemble neural network models with fuzzy response aggregation for predicting COVID-19 time series: The case of Mexico," *Healthcare*, vol. 8, no. 2, p. 181, Jun. 2020.

[18] P. Wang, X. Zheng, J. Li, and B. Zhu, "Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110058.

[19] S. Tuli, S. Tuli, R. Tuli, and S. S. Gill, "Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing," *Internet Things*, vol. 11, Sep. 2020, Art. no. 100222.

[20] J. P. A. Ioannidis, S. Cripps, and M. A. Tanner, "Forecasting for COVID-19 has failed," *Int. J. Forecasting*, Aug. 2020, doi: 10.1016/j.ijforecast.2020.08.004.

[21] A. Saltelli *et al.*, "Five ways to ensure that models serve society: A manifesto," *Nature*, vol. 582, no. 7813, pp. 482–484, Jun. 2020.

[22] K. N. Nabi, "Forecasting COVID-19 pandemic: A data-driven analysis," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110046.

[23] A. Hernandez-Matamoros, H. Fujita, T. Hayashi, and H. Perez-Meana, "Forecasting of COVID19 per regions using ARIMA models and polynomial functions," *Appl. Soft Comput.*, vol. 96, Nov. 2020, Art. no. 106610.

[24] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons Fractals*, vol. 135, Jun. 2020, Art. no. 109864.

[25] A. Tomar and N. Gupta, "Prediction for the spread of COVID-19 in India and effectiveness of preventive measures," *Sci. Total Environ.*, vol. 728, Aug. 2020, Art. no. 138762.

[26] C. Lee, Y. Li, and J. Kim, "The susceptible-unidentified infected-confirmed (SUC) epidemic model for estimating unidentified infected population for COVID-19," *Chaos Solit. Fractals*, vol. 139, Oct. 2020, Art. no. 110090.

[27] D. Okuonghae and A. Omame, "Analysis of a mathematical model for COVID-19 population dynamics in Lagos, Nigeria," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110032.

[28] C. H. Nkwayep, S. Bowong, J. J. Tewa, and J. Kurths, "Short-term forecasts of the COVID-19 pandemic: A study case of Cameroon," *Chaos, Solitons Fractals*, vol. 140, Nov. 2020, Art. no. 110106.

[29] M. Sheikholeslami, "Numerical investigation for CuO-H₂O nanofluid flow in a porous channel with magnetic field using mesoscopic method," *J. Mol. Liquids*, vol. 249, pp. 739–746, Jan. 2018.

[30] M. Sheikholeslami and M. Sadoughi, "Mesoscopic method for MHD nanofluid flow inside a porous cavity considering various shapes of nanoparticles," *Int. J. Heat Mass Transf.*, vol. 113, pp. 106–114, Oct. 2017.

[31] M. Gallagher and M. Frean, "Population-based continuous optimization, probabilistic modelling and mean shift," *Evol. Comput.*, vol. 13, no. 1, pp. 29–42, Mar. 2005.

[32] Y. Cai, X. Sun, and P. Jia, "Probabilistic modeling for continuous EDA with Boltzmann selection and Kullback–Leibeler divergence," in *Proc. GECCO*, Seattle, WA, USA, 2006, pp. 389–396.

[33] Y. Gao, Z. Zhang, W. Yao, Q. Ying, C. Long, and X. Fu, "Forecasting the cumulative number of COVID-19 deaths in China: A Boltzmann function-based modeling study," *Infection Control Hospital Epidemiol.*, vol. 41, no. 7, pp. 841–843, Apr. 2020.

[34] C. Tsallis and U. Tirnakli, "Predicting COVID-19 peaks around the world," *Frontiers Phys.*, vol. 8, p. 217, May 2020.

[35] W. Greiner, D. Rischke, L. Neise, and H. Stöcker, *Thermodynamics and Statistical Mechanics* (Classical Theoretical Physics). New York, NY, USA: Springer, 2012, pp. 208–239.

[36] G. Szegő, *Orthogonal Polynomials*. Providence, RI, USA: American Mathematical Society, 1939, pp. 23–159.

[37] N. P. Jewell, *Statistics for Epidemiology*. Boca Raton, FL, USA: CRC Press, 2003, pp. 9–123.

[38] A. Stang, M. Deckert, C. Poole, and K. J. Rothman, "Statistical inference in abstracts of major medical and epidemiology journals 1975–2014: A systematic review," *Eur. J. Epidemiol.*, vol. 32, no. 1, pp. 21–29, Nov. 2016.

[39] A. Baker, I. Biazzo, A. Braunstein, G. Catania, L. Dall'Asta, A. Ingrosso, F. Krzakala, F. Mazza, M. Mézard, A. P. Muntoni, M. Refinetti, S. S. Mannelli, and L. Zdeborová, "Epidemic mitigation by statistical inference from contact tracing data," in *Proc. PNAS*, Princeton, NJ, USA, 2021, pp. 1–6.

[40] M. O'Driscoll, C. Harry, C. A. Donnelly, A. Cori, and I. Dorigatti, "A comparative analysis of statistical methods to estimate the reproduction number in emerging epidemics, with implications for the current coronavirus disease 2019 (COVID-19) pandemic," *Clin. Infectious Diseases*, vol. 73, pp. e215–e223, Oct. 2020.

[41] X.-S. Yang, A. H. Gandomi, S. Talatahari, and A. H. Alavi, "Optimization and metaheuristic algorithms in engineering," in *Metaheuristics in Water, Geotechnical and Transport Engineering*. Walthman, MA, USA: Elsevier, 2013, pp. 1–23.

[42] M. A. Remli, M. S. Mohamad, S. Deris, S. Napis, R. Sinnott, and M. F. Sjaugi, "Metaheuristic optimization for parameter estimation in kinetic models of biological systems—Recent development and future direction," *Current Bioinf.*, vol. 12, no. 4, pp. 286–295, Jul. 2017.

[43] S. Fong, S. Deb, and X.-S. Yang, "How meta-heuristic algorithms contribute to deep learning in the hype of big data analytics," in *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*, P. K. Sa, M. N. Sahoo, M. Murugappan, Y. Wu, and B. Majhi, Eds. Singapore: Springer, 2018, pp. 3–25.

[44] V. Calzada-Ledesma, H. J. Puga-Soberanes, M. Ornelas-Rodríguez, A. Rojas-Domínguez, J. M. Carpio-Valadez, A. Espinal, J. A. Soria-Alcaraz, and M. A. Sotelo-Figueroa, "Evolutionary design of problem-adapted image descriptors for texture classification," *IEEE Access*, vol. 6, pp. 40450–40462, 2018.

[45] J. de Anda-Suárez, S. Jeyakumar, M. Carpio, H. J. Puga, A. Rojas-Domínguez, L. Cruz-Reyes, and J. F. Mosiño, "Parameter optimization for the smoothed-particle hydrodynamics method by means of evolutionary metaheuristics," *Comput. Phys. Commun.*, vol. 243, pp. 30–40, Oct. 2019.

[46] X.-S. Yang, "Metaheuristic optimization: Algorithm analysis and open problems," in *Experimental Algorithms*, P. M. Pardalos and S. Rebennack, Eds. Berlin, Germany: Springer, 2011, pp. 21–32.

[47] M. Abdel-Basset, L. Abdel-Fatah, and A. K. Sangaiah, "Metaheuristic algorithms: A comprehensive review," in *Computational Intelligence for Multimedia Big Data on the Cloud With Engineering Applications: Intelligent Data-Centric Systems*, A. K. Sangaiah, M. Sheng, and Z. Zhang, Eds. Cambridge, MA, USA: Academic, 2018, pp. 185–231.

[48] J. De Anda-Suárez, J. M. Carpio-Valadez, H. J. Puga-Soberanes, V. Calzada-Ledesma, A. Rojas-Domínguez, S. Jeyakumar, and A. Espinal, "Symmetric-approximation energy-based estimation of distribution (SEED): A continuous optimization algorithm," *IEEE Access*, vol. 7, pp. 154859–154871, 2019.

[49] S. I. Valdez, A. Hernández, and S. Botello, "A Boltzmann based estimation of distribution algorithm," *Inf. Sci.*, vol. 236, pp. 126–137, Jul. 2013.

[50] D. Malacara-Hernandez, M. Carpio-Valadez, and J. J. Sanchez-Mondragon, "Wavefront fitting with discrete orthogonal polynomials in a unit radius circle," *Opt. Eng.*, vol. 29, p. 672, Jun. 1990.

[51] E. O.-O. M. Roser, H. Ritchie, and J. Hasell. Coronavirus pandemic (COVID-19). Our World in Data. Accessed: May 2021. [Online]. Available: https://ourworldindata.org/coronavirus

[52] R. F. Woolson, "Wilcoxon signed-rank test," in *Wiley Encyclopedia of Clinical Trials*. Hoboken, NJ, USA: Wiley, 2008, pp. 1–3.

[53] Q. Wang, A. Farahat, C. Gupta, and S. Zheng, "Deep time series models for scarce data," *Neurocomputing*, vol. 456, pp. 504–518, Oct. 2021.

[54] J. Sirignano and K. Spiliopoulos, "DGM: A deep learning algorithm for solving partial differential equations," *J. Comput. Phys.*, vol. 375, pp. 1339–1364, Dec. 2018.

[55] M. Fernández-Delgado, M. S. Sirsat, E. Cernadas, S. Alawadi, S. Barro, and M. Febrero-Bande, "An extensive experimental survey of regression methods," *Neural Netw.*, vol. 111, pp. 11–34, Mar. 2019.

[56] M. Lepot, J.-B. Aubin, and F. Clemens, "Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment," *Water*, vol. 9, no. 10, p. 796, Oct. 2017.

[57] N. Wagner, Z. Michalewicz, M. Khouja, and R. R. McGregor, "Time series forecasting for dynamic environments: The DyFor genetic program model," *IEEE Trans. Evol. Comput.*, vol. 11, no. 4, pp. 433–452, Aug. 2007.

[58] J. Zhong, L. Feng, W. Cai, and Y.-S. Ong, "Multifactorial genetic programming for symbolic regression problems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 11, pp. 4492–4505, Nov. 2020.

JUAN DE ANDA-SUÁREZ received the B.Sc. degree in physics from the University of Guanajuato, in 2012, the M.Sc. degree in computer science from the TecNM—Instituto Tecnológico de León, in 2014, and the Ph.D. degree in computer science from the TecNM—Instituto Tecnológico de Tijuana, in 2019. He is currently a full-time Professor with the Electromechanical Engineering Division, TecNM—Instituto Tecnológico Superior de Purísima del Rincón. His research interests include estimation of distribution algorithms, information theory, and computational physics. He is a member of the National System of Researchers of the National Council of Science and Technology of Mexico (SNI-CONACYT).
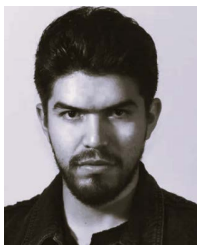
VALENTÍN CALZADA-LEDESMA received the M.Sc. degree in computer science from the TecNM—Instituto Tecnológico de León, in 2009, and the Ph.D. degree in computer science from the TecNM—Instituto Tecnológico de Tijuana, in 2019. He is currently a full-time Professor with the Computer Engineering Division, TecNM—Instituto Tecnológico Superior de Purísima del Rincón. His research interests include evolutionary algorithms, computer vision, and machine learning. He is a member of the National System of Researchers of the National Council of Science and Technology of Mexico (SNI-CONACYT).

DAVID ASAEL GUTIÉRREZ-HERNÁNDEZ received the Ph.D. degree in physics from the Autonomous University of Sinaloa. He has a post-doctorate in the area of natural sciences. He is currently the Head of the Division of Postgraduate Studies and Research, TecNM—Instituto Tecnológico de León. His research interests include optoelectronics applied to biomedic engineering and applied bioinformatics.

RAÚL SANTIAGO-MONTERO received the B.Sc. degree in electronics and communications engineering from the National Polytechnic Institute, Mexico, in 2000, the M.Sc. degree in computer science from CIC-IPN, Mexico, in 2003, and the Ph.D. degree from IIMAS-UNAM, Mexico, in 2008. Since 2009, he has been with the TecNM—Instituto Tecnológico de León, Mexico, where he is currently a full-time Professor. His research interests include images analysis, pattern recognition, shape analysis, and digital geometry. He has won the National Prize for the best master thesis in computer science, in 2004.

LUIS FERNANDO VILLANUEVA-JIMÉNEZ received the M.I.I. degree in industrial engineering from the TecNM—Instituto Tecnológico de Celaya, in 2013. He is currently pursuing the Ph.D. degree with the Centro de Innovación Aplicada a Tecnología Competitivas A.C. (CIATEC). He is the Head of the Research and Postgraduate Department, TecNM—Instituto Tecnológico Superior de Purísima del Rincón. His research interests include decision-making based on artificial intelligence and multi-objective optimization.

SERGIO RODRÍGUEZ-MIRANDA received the B.E.M. degree from the University of Guadalajara, in 2011, and the M.E.E. degree from the University of Guanajuato, in 2014. He is currently pursuing the Ph.D. degree with the Centro de Innovación Aplicada a Tecnología Competitivas A.C. (CIATEC). He is a full-time Professor with the Automotive Systems Engineering Division, TecNM—Instituto Tecnológico Superior de Purísima del Rincón, México. His research interests include computer vision and robotics.

● ● ●