Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/150403

Estimating Parameters of a SEIRD Model Applied to SARS-CoV-2 Infections in Germany based on the Particle Swarm Optimization Method

*Evandro A. Nakajima *, Antonio A. Ignacio, Denise Lange and Erika Izumi*

Universidade Tecnológica Federal do Paraná (UTFPR), Campus Santa Helena, PR, Brazil

Received: 9 Sep. 2020, Revised: 2 Feb. 2021, Accepted: 18 May 2021
Published online: 1 Jul. 2021

Abstract: The present paper describes a particle swarm optimization (PSO) method used to estimate parameters of a Susceptible-Exposed-Infected-Recovered-Dead (SEIRD) model applied to predict SARS-CoV-2 transmission in Germany, based on data from February 15th to April 25th, 2020, considering that the lockdown in the country started on March 23rd. The model estimated patients' mortality (4.92%) and recovery rates (95.08%), virus incubation (8.54 days), infection periods (18.65 days), as well as the basic virus reproduction number before (R_0 = 11.60) and after (R_0 = 0.39) lockdown. The predicted values were accurate until the 70th day. The performances achieved by the model were 0.98 for infected, 0.97 for the recovered and 0.97 for the dead, asserting the model's great performance ($>$ 0.75). The model also suggests that on February 15th, 2020, there were 67 infected individuals in the incubation period. We believe that this model can help other studies to better understand and accurately predict epidemic curves, mainly in countries where the new coronavirus has recently started to spread. It also may guide public health policies that aim to control the disease.

Keywords: COVID-19, Epidemiology, Modelling, Pandemic, PSO

1 Introduction

SARS-CoV-2 is the novel coronavirus, which is responsible for a worldwide pandemic that currently accounts for almost 6 million infected people and approximately 360 thousand deaths in more than 200 countries and territories [1]. Infection occurs through respiratory droplets and contact routes during the incubation period, which enables its rapid geographic spread and makes it difficult to estimate and control the disease [2,3]. Although the disease (COVID-19) is asymptomatic in approximately 25% of individuals, most of the infected population presents mild-to-moderate unspecific symptoms that could be misdiagnosed as other diseases and hinder COVID-19 diagnosis [4,5,6].

Severely- and critically-ill patients represent the minority of positive cases. However, they require hospitalization for several weeks; approximately 10% of infected individuals require Intensive Care Unit (ICU) treatment in some countries [7]. Given the limited number of hospital beds, mainly of ICU beds, many countries have adopted early control measures to reduce viral transmission and to avoid healthcare system overload due to prolonged hospitalizations [8,9].

Preventing the spread of infectious diseases is a difficult task, especially in the present globalized world. Factors such as socio-economic conditions, access to healthcare services, social behavior, personal hygiene habits, environmental conditions, among others, can influence epidemic outbreaks [10]. It is essential to isolate the infected individuals and conduct mass diagnostic tests to determine the real number of infected people and death rate to better control the pandemic [11].

Predicting SARS-CoV-2 transmission is a current challenge for public health policy planning. Thus, the application of predictive mathematical models to infectious disease epidemiological dynamics can help better understand pathogen transmission and propagation in different countries and regions [12,13].

The present paper aims to use the Particle Swarm Optimization method (PSO) to estimate parameters of a SEIRD (Susceptible-Exposed-Infected- Recovered-Dead) model [14] to predict the number of infected, recovered and dead people throughtout SARS-CoV-2 transmission.

* Corresponding author e-mail: enakajima@utfpr.edu.br

To achieve the research objective, we used epidemic data from Germany because the country adopted lockdown and mass testing procedures, which provided numerical data close to the real context.

2 Methods

Real data about the number of SARSCoV-2-infected, recovered and dead individuals, as well as about the total population of Germany in 2020, were provided by the World of Meters [15] and corroborated by data released by Johns Hopkins University [16].

The SEIRD model was used to check and model SARS-CoV-2 transmission among the German population from February 15$^{\text{th}}$ to April 25$^{\text{th}}$ (70 days) based on five infection stages: susceptible (S), exposed (E), infected (I), recovered (R) and dead (D), as shown in Fig. 1.

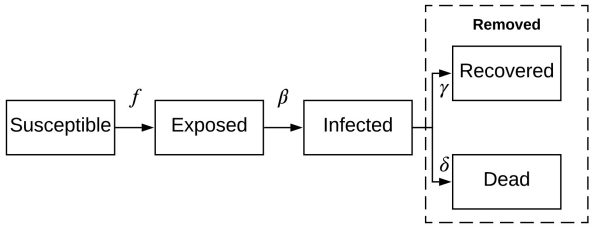

Fig. 1: SEIRD compartimental model flowchart

Overall, the SEIRD model enables a system comprising five differential equations (1-5), which rule the disease transmission dynamics. The model application considers that Germany has implemented lockdown system on March 23$^{\text{rd}}$, 2020 [17], which is indicated by functions f (change in α parameter after the incubation period) and g (reduction by 70% in the number of susceptible individuals [18]) in the differential equation system below:

$$\frac{dS}{dt}(t) = -f(t)\cdot S(t)\cdot I(t)\cdot N^{-1} + g(t), \quad (1)$$

$$\frac{dE}{dt}(t) = f(t)\cdot S(t)\cdot I(t)\cdot N^{-1} - \beta\cdot E(t), \quad (2)$$

$$\frac{dI}{dt}(t) = \beta\cdot E(t) - \gamma\cdot I(t), \quad (3)$$

$$\frac{dR}{dt}(t) = \gamma\cdot(1-\delta)\cdot I(t), \quad (4)$$

$$\frac{dD}{dt}(t) = \gamma\cdot\delta\cdot I(t). \quad (5)$$

Where

$$f(t) = \begin{cases} \alpha_1, & 0 \le t \le 39 + \dfrac{1}{\beta} \\ \alpha_2, & \text{otherwise} \end{cases}, \quad (6)$$

$$g(t) = \begin{cases} -\dfrac{N}{1.4}, & 38 \le t \le 39 \\ 0, & \text{otherwise} \end{cases}, \quad (7)$$

- $N = 8.3\cdot 10^7$ is the Germany total population;
- α_1 and α_2 are the transmission rates;
- β is the inverse of the incubation period [days^{-1}];
- γ is the inverse of the infection period [days^{-1}];
- $100\cdot\delta$ is the mortality rate [%].

The PSO method, implemented on Maple 19 [19], can be divided into two big stages: initialization and iteration. First, we established the domain to estimate the six parameters : $\alpha_1 \in [0,1]$, $\alpha_2 \in [0,1]$, $\beta \in [0,1]$, $\gamma \in [0,1]$, $\delta \in [0,1]$ and $K \in [0,1000]$, where K is the number of infected at $t = 1$. Then, each particle k at iteration 1 was defined by a six-component vector, each of which was a random number from its respective domain represented by:

$$X_1^k = (\alpha_{1,1}^k, \alpha_{2,1}^k, \beta_1^k, \gamma_1^k, \delta_1^k, K_1^k). \quad (8)$$

The system of equations (1-7) was solved based on the Fehlberg fourth-fifth order Runge-Kutta method of Maple 19. We defined the fitness-function F as the sum of the absolute error among $I(t)$, $R(t)$, $D(t)$ and the number of infected, recovered, dead individuals, respectively, for t ranging from 1 to 70, which was calculated based on parameters of each particle X_i^k.

After the test, we defined the constants of the PSO method presented by [20] as: number of iterations = 200, number of particles = 200, $c_1 = 0.5$, $c_2 = 0.5$, $\omega_0 = 0.9$, $\omega_{200} = 0.4$, $v_0^k = 0$. The random constants c_1 and c_2 represent the weight of individual and collective learning respectively, the constants w_0 and w_200 represent the initial and final control of the velocity of the particles and v_0^k represents the initial velocity of the particle k.

The algorithm was performed 200 times - with mean and standard deviation of 115.52 ± 14.79 iterations until the standard deviation of each parameter was half of its mean – among all particles. This criterion was adopted to assure that particles were close to each other to avoid generating significant difference in the results.

Mann-Whitney U-test was applied to check whether there were differences in the distribution between observed and predicted data. Linear regressions were performed between the observed and predicted data, and the d_r index, presented by Willmott [21], was used to validate the model and to analyze its performance, based

on:

$$d_r = \begin{cases} 1 - \frac{1}{2}\frac{\sum_{i=1}^{n}|P_i - O_i|}{\sum_{i=1}^{n}|O_i - \overline{O}|}, & \text{if } \sum_{i=1}^{n}|P_i - O_i| \leq 2\sum_{i=1}^{n}|O_i - \overline{O}| \\ 2 \cdot \frac{\sum_{i=1}^{n}|O_i - \overline{O}|}{\sum_{i=1}^{n}|P_i - O_i|} - 1, & \text{otherwise} \end{cases} \quad (9)$$

Where

- P_i is the i-th predicted data;
- O_i is the i-th observed data;
- $\overline{O}$ is the mean of the observed data;
- $n = 70$ for this study.

Then, the performance coefficient [22] is given by:

$$PI = r \cdot d_r \quad (10)$$

Where r is the Pearson correlation coefficient.

3 Results and Discussion

Parameters determined by the PSO method in equations (1-7) and their respective standard deviations were: $\alpha_1 = 0.6223 \pm 0.0155$, $\alpha_2 = 0.0198 \pm 0.0120$, $\beta = 0.1171 \pm 0.0035$, $\gamma = 0.0536 \pm 0.0010$, $\delta = 0.0492 \pm 0.0023$ and $K = 67.5633 \pm 12.4460$.

Based on these values, it was possible to use the model to estimate significant SARS-CoV-2-related numbers with their respective standard deviations: mortality rate $(100 \cdot \delta)$ reached 4.92 ± 0.27, recovery rate $(100 \cdot (1-\delta))$ was 95.08 ± 0.27, incubation period (β^{-1}) comprised 8.54 ± 0.03 days, the infectious period (γ^{-1}) encompassed 18.65 ± 0.04 days, basic reproduction number before lockdown $(R_0^{(1)} = \alpha_1 \cdot \gamma^{-1})$ was 11.60 ± 0.26 and basic reproduction number after lockdown $(R_0^{(2)} = \alpha_2 \cdot \gamma^{-1})$ was 0.37 ± 0.03. The model also suggests that on February 15th, 2020 there were 67 ± 12 infected individuals in the incubation period. For those parameters, the SEIRD model was adopted and graphics of functions I(t), R(t) and D(t) are shown in Fig. 2.

Similar and linear (Fig. 4) distribution of real and predicted data about infected, recovered and dead individuals ($U = 2309$, $p = 0.5582$, $r = 0.9964$; $U = 1998$, $p = 0.0599$, $r = 0.9960$; $U = 2003$, $p = 0.0622$, $r = 0.9956$, respectively) throughout the 70 days was used to estimate the evaluated parameters. Dispersion measures of estimated data are available in Table 1. Results of Willmott index and performance coefficient ($d_r = 0.9892$, $PI = 0.9856$; $d_r = 0.9815$, $PI = 0.9776$; $d_r = 0.9806$, $PI = 0.9762$ for the infected, recovered and dead, respectively) have asserted the model's great performance ($PI > 0.75$, [22]).

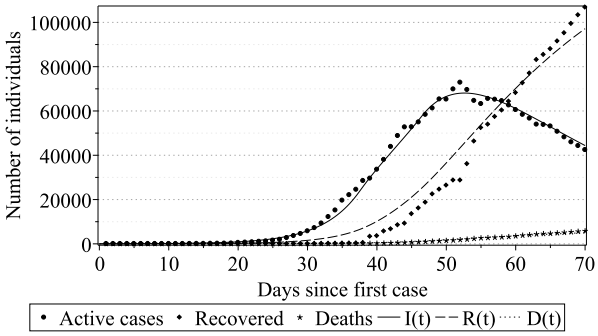

Fig. 2: Predicted and Real data about the number of infected, recovered and dead individuals for 70 days

Estimated parameters were used to predict the following 30 days (Fig. 3); the distribution of predicted and real data throughout 100 days was also similar and linear (Fig. 5) among infected, recovered and dead individuals ($U = 4785$, $p = 0.6002$, $r = 0.9965$; $U = 4859$, $p = 0.7314$, $r = 0.9960$; $U = 4902$, $p = 0.8116$, $r = 0.9955$) respectively. The model has also shown great performance in this case ($d_r = 0.9816$, $PI = 0.9780$; $d_r = 0.9636$, $PI = 0.9598$; $d_r = 0.9588$, $PI = 0.9545$ for the infected, recovered and dead individuals, respectively).

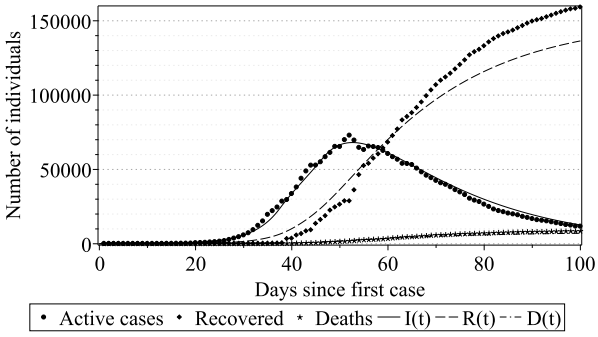

Fig. 3: Predicted and Real data about the number of infected, recovered and dead individuals for 100 days

Our model has accurately estimated data about SARS-CoV-2 infected, dead and recovered individuals in Germany. The distribution of the estimated data was similar to that of the real data and presented high validation rate. The herein estimated parameters were compatible to data provided by the World Health Organization, as well as by other studies on the novel coronavirus (Table 1).

On April 19th [25], the German Government started a mass testing to enable finding the exact number of

Table 1: Information about recent studies using modeling for SARS-CoV-2 in the literature

Local	Incubation Period	Infectious period	Lethality rate	Method
Germany [this study]	**8.54 days**	**18.54 days**	**4.92%**	**SEIRD + PSO**
China/WHO(1) [23]	**1-14 days**	**14-42 days**	**3.80%**	**Observed data**
U.S. [14]	2.56 days	17.82 days	0.56%	SEIRD
Italy [24]	3.77 days	0.5 – 58.82 days	3.80%	SEIR+PSO
Spain [24]	1.06 days	6.41 – 22.73 days	4.60%	SEIR+PSO
South Korea [24]	0.50 days	20 days	7.80%	SEIR+PSO

(1) Data from China considered the current refence by WHO.

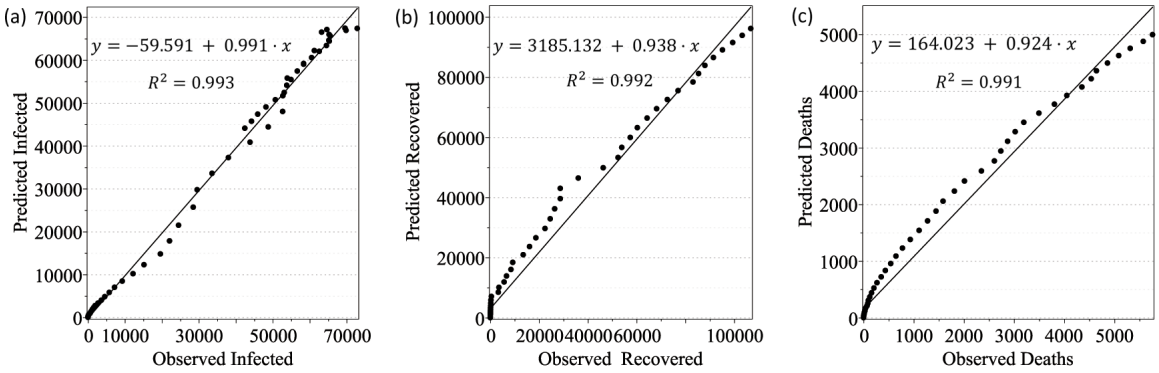

Fig. 4: Regression line between observed and predicted data for Infected (a), Recovered (b) and Death (c) for 70 days

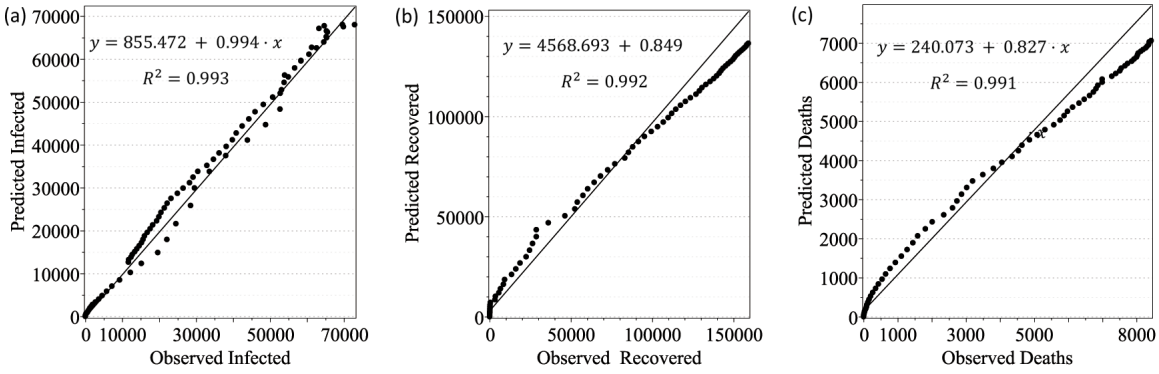

Fig. 5: Regression line between observed and predicted data for Infected (a), Recovered (b) and Death (c) for 100 days

asymptomatic individuals. This approach increased the number of recovered patients from that date on; these patients did not join the group of infected individuals and made the predicted number of recovered individuals relatively smaller than the actual number. There was also difference in the number of deaths due to the speed at which the tests were carried out.

Recent studies have used the data available in the global literature on the novel coronavirus in different modelling approaches to help better understand the viral dissemination and its consequences in several countries. However, no SEIRD-PSO method had yet been applied in Germany [26,27,28,29,30] to predict the evolution of the novel coronavirus epidemic curve.

In this paper, the SEIRD model based on the use of the PSO method as parameter estimator was effective in predicting the number of infected and dead individuals, as well as in indicating important features of the epidemic behavior of the novel coronavirus on a specific population. The relative simplicity of the SEIRD model -

when it comes to the need of preliminary data such as number of infected, recovered and dead individuals, and to determining its parameters - turns it into a good modelling option to follow the evolution of the disease outbreak in a shorter period-of-time due to its significant accuracy. This simplicity in association with the low computational cost of the PSO method and its effectiveness (evidenced in the current study) makes the herein presented model highly recommended.

Table 2: Nomenclature

Symbol	
α	transmission rate
β^{-1}	incubation period (days)
γ^{-1}	inverse of the infection period (days)
δ	mortality rate
K	number of infected at February 15^{th}, 2020
N	total population of Germany
S	number of susceptibles
E	number of exposed
I	number of infected
R	number of recovered
D	number of deaths
f	change in α parameter after the incubation period
g	reduction by 70% in the number of susceptible individuals
d_r	Willmott index
P_i	i-th predicted data
O_i	i-th observed data
$\overline{O}$	mean of observed data
n	data size

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] World Health Organization. Coronavirus disease (COVID-2019) situation reports: report 130. 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. Accessed 30 May 2020.

[2] D. Tang, P. Comish and R. Kang. The hallmarks of COVID-19 disease. *PLoS Pathogens*, **16(5)**, 1-24 (2020).

[3] R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yank and J. Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, **368(6490)**, 489-493 (2020).

[4] J. Y. Noh, J. G. Yoon, H. Seong, W. S. Choi, J. W. Sohn, H. J. Cheong, W. J. Kim and J. Y. Song. Asymptomatic infection and atypical manifestations of COVID-19: comparison of viral shedding duration. *J. Infect.*, **81(5)**, 816-846 (2020).

[5] X. Tang, R. Du, R. Wang, T. Cao, L. Guan, C. Yang, Q. Zhu, M. Hu, X. Li, Y. Li, L. Liang, Z. Tong, B. Sun, P. Peng and H. Shi. Comparison of Hospitalized Patients With ARDS Caused by COVID-19 and H1N1. *Chest*, **158(1)**, 195-205 (2020).

[6] C. Lorenz, T. S. Azevedo and F. Chiaravalloti-Neto. COVID-19 and dengue fever: A dangerous combination for the health system in Brazil. *Travel Med. Infect. Dis.*, **35**, 101659 (2020).

[7] J. Phua, L. Weng, L. Ling, M. Egi, C. M. Lim, J. V. Divatia, B. R. Shrestha, Y. M. Arabi, J. Ng, C. D. Gomersall, M. Nishimura, Y. Koh and B. Du. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations. *Lancet Respir. Med.*, **8(5)**, 506-517 (2020).

[8] X. Ma and D. Vervoort. Critical care capacity during the COVID-19 pandemic: Global availability of intensive care beds. *J. Crit. Care.*, **58**, 96-97 (2020).

[9] B. J. Cowling and A. Aiello. Public health measures to slow community spread of COVID19. *J. Infect. Dis.*, **221(11)**, 1749-1751 (2020).

[10] H. Salje, J. Lessler, K. K. Paul, A. S. Azman, M. W. Rahman, M. Rahman, D. Cummings, E. S. Gurley and S. Cauchemez. How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study. *Proc. Natl. Acad. Sci. USA.*, **113(47)**, 13420-13425 (2016).

[11] J. Peto. Covid-19 mass testing facilities could end the epidemic rapidly. *BMJ.*, **368**, 1 (2020).

[12] A. Huppert and G. Katriel. Mathematical modelling and prediction in infectious disease epidemiology. *Clin. Microbiol. Infect.*, **19(11)**, 999-1005 (2013).

[13] F. Brauer. Mathematical epidemiology: Past, present, and future. *Infect. Dis. Model.*, **2(2)**, 113-127 (2017).

[14] A. Maugeri, M. Barchitta, S. Battiato and A. Agodi. Estimation of Unreported Novel Coronavirus (SARS-CoV-2) Infections from Reported Deaths: a Susceptible Exposed Infectious Recovered Dead Model. *J. Clin. Med.*, **9(5)**, 1350 (2020).

[15] The Worldometer: COVID-19 CORONAVIRUS PANDEMIC (2020). https://www.worldometers.info/coronavirus/country/germany. Accessed 30 May 2020.

[16] John Hopkins University and Medicine: CORONAVIRUS RESOURCE CENTER (2020). https://coronavirus.jhu.edu/map.html. Accessed 30 May 2020.

[17] D. Bundesregierung. Coronavirus in Deutschland (2020). https://www.bundesregierung.de/breg-de/themen/coronavirus/besprechung-derbundeskanzlerin-mit-den-regierungschefinnen-und-regierungschefs-derlaender-1733248. Accessed 30 May 2020.

[18] Our World in Data: COVID-19: Government Response Stringency Index (2020). https://ourworldindata.org/grapher/covid-stringency-index. Accessed 30 May 2020.

[19] Waterloo Maple Inc: Maple (2019).

[20] M. Schwaab, E. C. Biscaia, J. L. Monteiro and J. C. Pinto. Nonlinear parameter estimation through particle swarm optimization. *Chem. Eng. Sci.*, **63(6)**, 1542-1552 (2008).

[21] C. J. Willmott, S. M. Robesonb and K. Matsuura. A refined index of model performance. *Int. J. Climatol.*, **32(13)**, 2088-2094 (2012).
[22] C. A. Alvares, J. L. Stape, P. C. Sentelhas and J. L. de Moraes Gonçalves. Modeling monthly mean air temperature for Brazil. *Theor. Appl. Climatol.*, **113(3-4)**, 407-427 (2013).
[23] WHO. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). Tech. Rep. (2020). https://www.who.int/docs/defaultsource/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf. Accessed 30 May 2020.
[24] A. Godio, F. Pace and A. Vergnano. SEIR Modeling of the Italian Epidemic of SARS-CoV2 Using Computational Swarm Intelligence. *Int. J. Environ. Res. Public. Health.*, **17(10)**, 3535 (2020).
[25] Mark Armstrong: Germany starts mass-testing for coronavirus antibodies in bid to learn more about COVID-19 (2020). https://www.euronews.com/2020/04/19/germany-starts-mass-testing-for-coronavirus-antibodies-in-bid-to-learn-more-about-covid-19. Accessed 30 May 2020.
[26] E. Karadag. Increase in COVID-19 cases and case-fatality and case-recovery rates in Europe: A cross-temporal meta-analysis. *J. Med. Virol.*, **92(9)**, 1511-1517 (2020).
[27] V. Verma, R. K. Vishwakarma, A. Verma, D. C. Nath and H. T. A. Khan. Time-to-Death approach in revealing Chronicity and Severity of COVID-19 across the World. *PLoS ONE*, **15(5)**, e0233074 (2020).
[28] R. K. Singh, M. Rani, A. S. Bhagavathula, R. Sah, A. J. Rodriguez-Morales, H. Kalita, C. Nanda, S. Sharma, Y. D. Sharma, A. A. Rabaan, J. Rahmani and P. Kumar. Prediction of the COVID-19 Pandemic for the Top 15 Affected Countries: Advanced Autoregressive Integrated Moving Average (ARIMA) Model. *JMIR Public Health Surveill.*, **6(2)**, e19115 (2020).
[29] S. Sanchez-Caballero, M. A. Selles, M. A. Peydro and E. Perez-Bernabeu. An Efficient COVID-19 Prediction Model Validated with the Cases of China, Italy and Spain: Total or Partial Lockdowns? *J. Clin. Med.*, **9(5)**, e1547 (2020).
[30] P. Khrapov and A. Loginova. Comparative analysis of the mathematical models of the dynamics of the coronavirus COVID-19 epidemic development in the different countries. *Int. J. Open Inf. Technol.*, **8(5)**, 17-22 (2020).

Evandro A. Nakajima is Associate Professor of Universidade Tecnológica Federal do Paraná since 2014 and received the Ms degree in Mathematics at Institute of Mathematical and Computer Sciences, Univeristy of São Paulo (2013). His research areas of interest are mathematics applied to physical-chemical phenomena, mathematical modeling of infectious diseases and bioprocess modeling.

Antonio A. Ignacio is a student in the Biological Sciences Degree course at the Universidade Tecnológica Federal do Paraná. His areas of interest are infectious diseases, bioprocess modeling, ecological interactions and microbiology.

Denise Lange completed her doctorate in Ecology at Universidade Federal de Uberlândia, Brazil, in 2012. She is professor in Universidade Tecnológica Federal do Paraná since 2015. She is referee in several international and national journals. Her main research interests are: ecological interactions, interaction networks and population and community ecology.

Erika Izumi obtained her BSc degree in Biological Sciences in 2004 and her PhD degree in Microbiology in 2010 from the Universidade Estadual de Londrina, Brazil. She undertook postdoctoral research on Biotechnology studying molecular targets for drug action in 2011. She worked as R&D Scientist from 2012 to 2015 with immunological and molecular diagnosis of infectious diseases. Currently she is an Associate Professor of the Universidade Tecnológica Federal do Paraná, Brazil, and her research interests are infectious diseases focusing on the development of new drugs.