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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
ModInterv: An automated online software for modeling epidemics

Arthur A. Brum, Gerson C. Duarte-Filho, Raydonal Ospina, Francisco A.G. Almeida, Antônio M.S. Macêdo, Giovani L. Vasconcelos

Abstract

The COVID-19 pandemic has proven the importance of mathematical tools to understand the evolution of epidemic outbreaks and provide reliable information to the general public and health authorities. In this perspective, we have developed ModInterv, an online software that applies growth models to monitor the evolution of the COVID-19 epidemic in locations chosen by the user among countries worldwide or states and cities in the USA or Brazil. This paper describes the software capabilities and its use both in recent research works and by technical committees assisting government authorities. Possible applications to other epidemics are also briefly discussed.

1. Introduction

The ongoing COVID-19 pandemic is a public health emergency with severe impacts on the lives of people everywhere around the globe. Sadly, the disease has thus far claimed the lives of over 6 million people, a figure that ranks the COVID-19 pandemic among the deadliest in history. Public health authorities play a major role in the fight against COVID-19, making decisions concerning the use of masks, stay-at-home mandates, school closures, among other non-pharmacological interventions. Obtaining reliable information is paramount to guide decision makers in choosing the best course of action to slow down (and eventually stop) the spread of the disease. In this context, mathematical models and data processing softwares are indispensable tools in the fight against epidemics.

Deterministic growth models are phenomenological approaches governed by a single differential equation that often admit analytical solution. Their power lies in the fact that they are capable of capturing the basic features of epidemic dynamics, without the need to describe the underlying epidemiological mechanisms. One growth model that has proven quite effective in describing the complex features of COVID-19 epidemic curves is a generalized logistic model known as the beta logistic model (BLM) [2]. An extension of the BLM with time dependent parameters has also been shown to be able to efficiently model epidemics with multiple waves of infections [3].

In this paper, we present the main features of the software ModInterv which enables the user to monitor the dynamical stages of COVID-19 epidemic curves for both infection cases and deaths by

https://doi.org/10.1016/j.simpa.2022.100409
Received 12 July 2022; Received in revised form 3 August 2022; Accepted 9 August 2022

2665-9638/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
implementing the multiwave BLM growth model mentioned above. More specifically, after the user chooses a type of data (cases or deaths) and a desired location (country or state/city in the USA or Brazil), the software automatically (i.e., without human assistance) performs a model fitting to the empirical data, extracts relevant information about the epidemic dynamics from the theoretical curve, and makes short-term predictions about the disease’s likely progression. The software is written in the Python programming language and made available in an HTML page currently accessed via the address http://fisica.ufpr.br/modinterv. The COVID-19 data used by ModInterv are fetched from public data repositories. Once the data is properly fetched, a plot of the empirical data is displayed and the user then has the option to perform a curve fitting, in which case the result of the best fit is shown superimposed with the empirical data. The software also allows the user to perform additional analysis on the basis of the fitted model.

We note in passing that there are several other softwares developed for application to the COVID-19 pandemic in the most diverse forms. One very common type of software is the dashboard, which is a visual display of a given dataset with the possibility to change between charts, graphs or the data themselves, as is the case of the World Health Organization Coronavirus Dashboard [4] and the Johns Hopkins University Dashboard [5]. Other softwares center their attention to forecasting the evolution of the pandemic, such as the Forecast Hub [6] and the Covidsim from Imperial College London [7], while others focus on the diagnosis of COVID-19 [8,9] and the numerical fitting of multi-wave epidemic models [10]. The app ModInterv described here differs substantially from the aforementioned softwares in that it is a web-based application that enables the user to monitor the evolution of the COVID-19 pandemic in countries around the world as well as for states and cities in Brazil and the US, through the use of mathematical growth models, from which it can determine the current epidemic stage in the chosen location. The app can also make short-term predictions on the evolution of the epidemic curve for a given location based on the numerical fit of the growth model.

2. Software

ModInterv is a software that runs within IPython Notebook [11], an interactive computational environment that combines code execution, rich text, and plots. The code is divided into two similar major cells: one performs analyses of COVID-19 data for countries, while the second does the same for states and cities in the USA and Brazil. In each of these cells there are functions that perform specific tasks, such as fetching the data for the chosen location, plotting the empirical data, performing curve fits, and exporting the results onto figure files with different formats. To ensure its functionality, the code was tested and debugged before it was made available online.

The software workflow and its dependencies with the main Python modules used are shown in Fig. 1. IPywidgets (also known as Jupyter Widgets) [12] is a framework that provides interactive controls for Jupyter Notebooks through the use of GUI elements, the so-called widgets, which are tools displayed as part of the output of the code that allow the user to provide inputs to the software. The main widgets used by the ModInterv software are as follows: dropdown menus, to allow the user to select the type of epidemic data (cases or deaths) to be analyzed and the desired location (countries or states and cities in the USA and Brazil); checkboxes, to select which outputs of the fit the user wants to see; and sliders, to select the final date used in the fits and the time range for a short term prediction of the epidemic curve. Once the user chooses a place and type of data, the respective empirical data are fetched from a specific data repository and imported into the software as a Pandas DataFrame [13]. Empirical data for countries as well as for states and counties in the USA are fetched from the database by the Johns Hopkins University [14], whereas data for Brazilian states and cities are fetched from the GitHub database maintained by Wesley Cota [15]. After the data has been imported, the corresponding time series is then transformed into a NumPy array [16] and plotted using the Matplotlib [17], a standard graphics module in Python. If the user chooses to perform the curve fit (by clicking on the appropriate button), the non-linear least square minimization problem is dealt with via the Levenberg–Marquardt algorithm implemented in Python by the LMFIT package [18]. From the best-fit curve, the software then determines the epidemic’s previous and current stages in the chosen location by computing certain characteristic dynamical points (zeros of the second to the fourth derivatives) [19] using the SciPy package [20]. While the software’s front-end is built using the IPython and IPywidgets frameworks, its back-end is provided by a binder [21] server, which builds a Docker image from an IPython notebook into an interactive executable that can be converted to an HTML file, thus rendering the software accessible for online use.

3. Impact overview

The ModInterv was developed by the MODINTERV-COVID19 Research Network, a research group formed by the present authors at the beginning of the COVID-19 pandemics. The software is licensed under a MIT license. The main purpose of the software is to offer an online tool for monitoring the epidemic evolution in countries around the world as well as in states and cities in Brazil and the USA, by means of reliable mathematical analyses (rather than mere visual inspection of epidemic curves). The software implements in a automated fashion phenomenological growth models to describe the empirical data and extract relevant information about the dynamics of the epidemic in the chosen location, thus contributing to a more precise assessment of the current epidemic situation in that location and its future evolution.

Since its inception, the software has been employed by members of our research network to study the COVID-19 dynamics in countries worldwide [2,3] as well as in states and cities in Brazil [22,23]. The software has also been used as a complementary tool of analysis by the
Scientific Committee to Combat Coronavirus (C4NE in its Portuguese acronym) of the Consórcio Nordeste, a Brazilian interstate consortium formed by the Northeastern state governments [24] with the purpose of promoting sustainable development in the region and enabling the implementation of joint public policies. Over the past two years, the C4NE has published several technical bulletins about the COVID-19 situation in the Brazilian Northeastern [25] in which the ModInterv software was used (together with other methods) to analyze the epidemic evolution in the region. Extensive use of the ModInterv was made by the COVID-19 data analysis team of the Prefecture of Teresina, capital of the Brazilian state Piauí, to monitor the epidemic there. Perhaps more importantly, they have taken the model’s predictions into account in their decision making process (e.g., as to when to relax control measures), particularly during the first wave of the epidemic [26].

4. Software limitations and future improvements

Notwithstanding its proven successes, as mentioned in the preceding section, the ModInterv software can be further improved to remove some of its minor limitations. The main points that can be improved concern the wave detection algorithm and the numerical fitting procedure, as briefly discussed below.

Regarding wave detection, the main difficulty lies in finding which of the (many) local maxima and minima detected by the Python’s native signal processing module are actually local maxima and minima of the epidemic curve, rather than random fluctuations. In the current version of ModInterv we have implemented a filter based on the relative amplitudes of nearby maxima/minima to reject those that are likely due to fluctuations [19]. We are currently investigating the possibility of using a more selective algorithm recently introduced in the literature for detecting relevant maxima and minima [27]. Another difficulty stems from the often large number of parameters that need to be determined from the chosen empirical data (there are $3N - 2$ free parameters for an epidemic curve with $N$ waves). To avoid overfitting, we have implemented a multi-step procedure in which we begin by fitting the first wave and then successively include data for each subsequent wave, whereby each next partial fit uses the fitted parameters from the previous step as initial guesses [19], until the entire dataset is included. This multi-step routine may, on occasion, render the numerical fit rather time consuming, especially for curves with four or more waves. It remains an open question to find more efficient optimization algorithms to speed up the best-fit routine without incurring in excessive over-fitting.

The software impact can be further augmented by extending its applications to other epidemics, such as monkeypox, measles, dengue, zika, etc. Indeed, the software behind the ModInterv app can be readily applied to any infectious disease outbreak, so long as the data are available in public repositories.

5. Conclusions

ModInterv is an online software written in Python which enables the user to monitor COVID-19 epidemic curves around the world. The software fits empirical epidemiological data with a multi-wave generalized logistic model in a fully automated manner (i.e., without human assistance), thus making a sophisticated mathematical analyses accessible to any user. In this sense, the ModInterv is a unique informatics tool for monitoring the epidemic evolution and as such it should be of interest to researchers and health authorities as well as to the general public. The ModInterv has been employed by our COVID-19 Research Network in several published research papers (http://fisica.ufpr.br/redecovid19). It has also been used by technical committees (assisting government authorities) as an auxiliary tool to analyze the evolution of the epidemic in Northeastern Brazil.

CRediT authorship contribution statement

Arthur A. Brum: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review and editing, Supervision. Gerson C. Duarte-Filho: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review and editing. Raydonal Osypina: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review and editing, Supervision. Francisco A.G. Almeida: Conceptualization, Methodology, Validation, Investigation, Writing – review and editing. Giovani L. Vasconcellos: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, writing – original draft, Writing – review and editing, Supervision.

Declaration of competing interest

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

Acknowledgments

This work was supported in part by the following Brazilian agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), under Grants No. 312985/2020-7 (GLV), No. 312612/2019-2 (AMSS), No. 305305/2019-0 (RO) and No. 167348/2018-3 (AAB), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). GLV also acknowledges partial funding from UFPR through the COVID-19/PROIND-2020 Research Program. All authors have read and agreed to the published version of the manuscript.

References

[1] G. Chowell, L. Safrenspiel, S. Bansal, C. Viboud, Mathematical models to characterize early epidemic growth: A review, Phys. Life Rev. 18 (2016) 66–97.
[2] G.L. Vasconcelos, A.M.S. Macêdo, G.C. Duarte-Filho, A.A. Brum, R. Osypina, F.A.G. Almeida, Power law behaviour in the saturation regime of fatality curves of the COVID-19 pandemic, Sci. Rep. 11 (1) (2021) 4619, http://dx.doi.org/10.1038/s41598-021-84165-1.
[3] G.L. Vasconcelos, A.A. Brum, F.A.G. Almeida, A.M.S. Macêdo, G.C. Duarte-Filho, R. Osypina, Standard and anomalous waves of COVID-19: A multiple-wave growth model for epidemics, Braz. J. Phys. S1 (6) (2021) 1687–1683, http://dx.doi.org/10.1007/s13538-021-00996-3.
[4] World Health Organization, WHO COVID-19 dashboard, 2020, Available online: Available online: (Last accessed on 02 August 2022).
[5] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, Lancet Infect. Dis. 20 (5) (2020) 533–534.
[6] E.Y. Cramer, Y. Huang, Y. Wang, E.L. Ray, M. Cornell, J. Bracher, A. Brennen, A.J. Castro Rivadeneira, A. Gerding, K. House, D. Jayawardena, A.H. Kanji, A. Khandelwal, K. Le, J. Niemi, A. Stark, A. Shah, N. Wattanachit, M.W. Zorn, N.G. Reich, U.C.-F.H. Consortium, The United States COVID-19 forecast hub dataset, 2021, http://dx.doi.org/10.1101/2021.11.04.21265867, MedRxiv, URL https://www.medrxiv.org/content/10.1101/2021.11.04.21265867v1.
[7] MRC Centre for Global Infectious Disease Analysis, Imperial College London, COVID-19 scenario analysis tool, 2020, www.covidsim.org, (Last accessed on 2 August 2022).
[8] A. Tavakolian, F. Hajati, A. Rezaee, A.O. Fasakhodi, S. Uddin, Source code for optimized parallel inception: A fast COVID-19 screening software, Softw. Impacts 13 (2022) 100337, http://dx.doi.org/10.1016/j.simpa.2022.100337, URL https://www.sciencedirect.com/science/article/pii/S2665963822000616.
[9] S. Hasani, H. Naini, COV-ADSX: An automated detection system using X-ray images, deep learning, and XBRoot for COVID-19, Softw. Impacts 11 (2022) 100210, http://dx.doi.org/10.1016/j.simpa.2021.100210, URL https://www.sciencedirect.com/science/article/pii/S2665963822000932.
[10] A. Cunha, F. da Conceição Batista, P.R. de Lima Gianfelice, R.S. Oyarzabal, J.M.V. Graybowski, E.E. Macau, epidWave: A code for fitting multi-wave
epidemic models, Softw. Impacts (2022) 100391, http://dx.doi.org/10.1016/j.simpa.2022.100391, URL https://www.sciencedirect.com/science/article/pii/S2665963822000884.

[11] F. Pérez, B.E. Granger, IPython: A system for interactive scientific computing, Comput. Sci. Eng. 9 (3) (2007) 21–29, http://dx.doi.org/10.1109/MCSE.2007.53, URL https://ipython.org.

[12] J.W. Community, Ipywidgets, a GitHub repository, 2015, Retrieved from https://github.com/jupyter-widgets/ipywidgets.

[13] J. Reback, W. McKinney, J. Van Den Bosche, T. Augspurger, P. Cloud, A. Klein, S. Hawkins, M. Roeschke, J. Trainor, C. She, et al., Pandas-dev/pandas: Pandas 1.0. 5, 2020, Zenodo.

[14] JHU, Coronavirus COVID-19 global cases by the center for systems science and engineering (CSSE) at Johns Hopkins university (JHU), 2020, https://coronavirus.jhu.edu/map.html. (Accessed 29 March 2020).

[15] W. Cota, Monitoring the number of COVID-19 cases and deaths in Brazil at municipal and federative units level, 2020, SciELO Preprint, https://covid19bc.wcota.me/.

[16] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature 585 (7825) (2020) 357–362, http://dx.doi.org/10.1038/s41586-020-2649-2.

[17] J.D. Hunter, Matplotlib: A 2D graphics environment, Comput. Sci. Eng. 9 (3) (2007) 90–95, http://dx.doi.org/10.1109/MCSE.2007.55.

[18] M. Newville, T. Stensitzki, D.B. Allen, A. Ingargiola, LMFIT: Non-linear least-square minimization and curve-fitting for python, 2014, http://dx.doi.org/10.5281/zenodo.11813, Zenodo.

[19] A.A. Brum, G.L. Vasconcelos, G.C. Duarte-Filho, R. Ospina, F.A.G. Almeida, A.M.S. Macêdo, ModInterv COVID-19: An online platform to monitor the evolution of epidemic curves, 2022, http://dx.doi.org/10.1101/2022.01.31.22270192, MedRxiv, URL https://www.medrxiv.org/content/early/2022/06/14/2022.01.31.22270192.

[20] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, I. Polat, Y. Feng, E.W. Moore, I. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in python, Nature Methods 17 (2020) 261–272, http://dx.doi.org/10.1038/s41592-019-0686-2.

[21] Project Jupyter, M. Bussonnier, J. Forde, J. Freeman, B. Granger, T. Head, C. Holdgraf, K. Kelley, G. Valnarte, A. Onserohr, M. Pacer, Y. Panda, F. Perez, B. Ragan-Kelley, C. Willing, Binder 2.0 - Reproducible, interactive, sharable environments for science at scale, in: F. Akici, D. Lippa, D. Niederhut, M. Pacer (Eds.), Proceedings of the 17th Python in Science Conference, 2018, pp. 113–120, http://dx.doi.org/10.25080/Majora-4af1f417-011.

[22] G.L. Vasconcelos, G.C. Duarte-Filho, A.A. Brum, R. Ospina, F.A.G. Almeida, A. Macêdo, Analysis of COVID-19 epidemic curves via generalized growth models: Case study for the cities of recife and Teresina (in Portuguese), 2020, SciELO Preprints.

[23] G.L. Vasconcelos, G.C. Duarte-Filho, A.A. Brum, R. Ospina, F.A.G. Almeida, A. Macêdo, Situation of COVID-19 in Brazil in august 2020: An analysis via growth models as implemented in the ModInterv system for monitoring the pandemic, J. Control Autom. Electr. Syst. (2022) 1–19.

[24] Consórcio nordeste official web page, 2022, http://www.consorcionordeste-ne.com.br/. (Accessed 28 June 2022).

[25] The technical bulletins are published in the section comitê científico of the following web page, 2022, http://www.consorcionordeste-ne.com.br/compras-conjuntas-combate-a-pandemia/. (Accessed 28 June 2022).

[26] F.S. Soares-Filho, G.S. Câncio Jr., Private communication, 2021.