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.
Original software publication

Source code and secondary data of the stochastic process based COVID-19 simulation model

S.S. Manathunga a, I.A. Abeyagunawardena a,∗, S.D. Dharmaratne b, c

a National Hospital of Sri Lanka, Sri Lanka
b Department of Community Medicine, Faculty of Medicine, University of Peradeniya, Sri Lanka
c Institute for Health Metrics and Evaluation, Department of Health Metrics Sciences, School of Medicine, University of Washington, United States of America

ARTICLE INFO

Keywords:
Novel coronavirus
COVID-19
Epidemic modelling
Basic reproduction number

ABSTRACT

The novel coronavirus disease (COVID-19) culminated in a pandemic with many countries affected in varying stages. We aimed to develop a simulation environment for COVID-19 spread, taking environmental and social factors into account. This program consists of three main components; a stochastic process-based model for simulating epidemics, a basic reproduction number estimation unit and a graphics generator. The model can take a variety of environmental factors as input and simulate expected behaviours of the infection spread, enabling policymakers and the scientific community to test the effects of different mitigation strategies in a sandbox.

Code metadata

Current code version
V2.0
Permanent link to code/repository used for this code version
https://github.com/SoftwareImpacts/SIMPAC-2022-30
Permanent link to reproducible capsule
https://codeocean.com/capsule/1394704/tree/v1
Legal code license
MIT License
Software code languages, tools and services used
R
Compilation requirements, operating environments and dependencies
R packages tidyverse, gridExtra, gganimate, R0, magick, scales, pbapply
If available, link to developer documentation/manual
https://github.com/ssm123ssm/contagion/blob/main/README.md
Support email for questions
Ssm123ssm@gmail.com

Body of the article

The novel coronavirus SARS-CoV-2, first detected in Wuhan, China in November 2019 spread rapidly globally culminating in a pandemic. The disease caused by this virus named coronavirus disease-2019 (COVID-19) caused loss of countless lives. Multiple strategies were implemented to mitigate the spread of infection by nations such as wearing masks and hand hygiene, lockdowns, strict curfews and, vaccination. The pattern of spread of the infection varied greatly across different countries due to the complex interplay of diverse environmental and social factors.

Governments attempted to control the spread of the diseases by actively restricting public mobility and quarantining the infected individuals and close-contacts of them. These mitigation strategies made the assumptions of the conventional epidemiological models such as homogeneous mixing of the population not applicable.

Deterministic epidemiological models divide the population into compartments and describe the dynamics of each with mathematical functions. Even though this approach is simple and more generalizable, active control methods make the reality different from the assumed compartment behaviours. This highlights the necessity of...
stochastic models which can incorporate tunable parameters for the control measures.

The aim of this software is to introduce a simulation environment for COVID-19 spread, taking the complex interactions between environmental and social factors into account. Previously proposed stochastic models incorporate a scarce amount of tunable parameters. The model proposed by Xie takes the average infection rate as the tuning parameter. Triambak et al. proposed a random-walk model which combines the control interventions as a single parameter; the size of each step in the random walk. Our proposed model incorporates randomness of a diverse set of events like public mobility into consideration making it different from deterministic and other stochastic models.

The user can simulate an epidemic by defining the geographical area and parameters for the number of population, the fraction of vaccinated population, the vaccine efficacy and timing, level of the public mobility level and the pattern of change, the sequence of the establishment of quarantine centres and the efficacy of contact tracing, the intrinsic virulence of the virus, the proportion of individuals adhering to proper handwashing and wearing face masks. Once the required parameters are input, the software estimates the evolution of the spread of the virus and calculates the basic reproduction number. Graphical representations of the epidemic with graphs for epidemic curve, dynamics of compartments in the conventional models and graphs depicting the uncertainty of the basic reproduction numbers generated by different models are constructed.

Finally, an animation of the spread of the disease in the simulated community is generated enabling the user to comprehend the dynamics of the spread visually.

The model is also unique due to its modular nature. The time of vaccination, imposing mobility restrictions and quarantining can be explicitly defined by the user. A vast array of ‘what-if’ situations can be modelled through this functionality to aid decision making. The model can be calibrated to produce the epidemic curve of an observed epidemic. This functionality aids in establishing a baseline for parameters and for forecasting the expected behaviour of the epidemic curve.

Results

To demonstrate the model, the spread of the infection in an imaginary population of 40 million was simulated. The calibrate function was utilized to estimate the baseline parameters to achieve a total of 20 million cases over one year simulation period. The geographical area was defined as 26 X 26 space. One quarantine centre was scheduled to be opened at the 20th time step, with a case-detection efficacy of 60%. The initial vaccinated fraction was set to 30% and the efficacy of the vaccine was set as 80%. The mobility level was set constant throughout the simulation period at a level of 25%.

The model estimated the behaviour of the susceptible, infected and removed epidemiological compartments throughout the epidemic and the basic reproduction number was estimated to be 1.14 using the exponential growth rate method.

The above scenario was set as the baseline and, the spread of the infection in different conditions was estimated. The vaccination coverage was increased up to 60%, and quarantining of the identified individuals was performed at an early stage. These changes reduced the total number of infected individuals to 1 million, which is 50% of the previous scenario. The estimated basic reproduction number in this scenario was 1.11.

The epidemic curves observed in California, Texas and Florida were successfully reproduced within the simulation environment. California reported 4 million cases from March 2020 to June 2021. 60% population had received at least one dose of vaccination by the end of June 2021. The model was calibrated with these actual data. An accurate simulation of the pattern of the spread was generated by the model, allowing us to test different ‘what-if’ situations.

The simulated epidemic curve estimated a total of 4 million cases for Texas by November 2021. The actually reported total number was 4.2 million cases, which was comparable with the model estimates. Similarly, the model estimated 3.9 million cases for Florida and the actually reported caseload was 3.6 million.

The generated figures and the estimated reproduction numbers are described in greater details in ‘A stochastic process based modular tool-box for simulating COVID-19 infection spread’ (http://dx.doi.org/10.1016/j.imu.2022.100899) [1].

We are currently working on improving the model by incorporating reinforcement learning-based agents which would be another step ahead of the conventional models towards self-aware agent-based models.

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.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.simpa.2022.100284.

References

[1] http://dx.doi.org/10.1016/j.imu.2022.100899.