Heliyon 9 (2023) e12670

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Fractional model analysis of COVID-19 spread based on big data platform

Yanfang Li [a,c], Xianghu Liu [a,b,*]

[a] Department of Arts and Sciences, Suqian College, Jiangsu, Suqian 223800, PR China
[b] Department of Mathematics, Zunyi Normal College, Zunyi, Guizhou 563006, PR China
[c] St. Paul University, Tuguegarao city, Cagayan valley 3500, Philippines

ARTICLE INFO

MSC:
35A24
39D30

Keywords:
COVID-19
FCSEIR model
Big data platform
Prediction
Infected

ABSTRACT

Based on the data of COVID-19, this paper establishes the FCSEIR model for the spread through data analysis and designs the related simulation software. Using the data from Shanghai, the spread of the virus was simulated and predicted, and the process from outbreak to control of this infectious disease was better analyzed.

1. Introduction

For a long time, mankind has never stopped fighting against infectious viruses. From the Black Death to AIDS, from the influenza epidemic to the SARS epidemic, to the current outbreak of the COVID-19, all of them are chapters of the human struggle against infectious viruses. Every outbreak of an infectious virus will have a huge impact on the global economic development and human life and health.

Das and Samanta [1] used a modified compartmental model of susceptible-asymptomatic infection-recovery, taking into account the uncertainty due to limited information on Covid -19, using fractional order calculus theory, and justified the model by comparing real data with simulation results. Baba and Nasidi [2] proposed a fractional order SIR model that relates individuals with mild cases as an interval, called the SMIR model. Omame A et al. [3] considered and analysed an Atangana-Baleanu fractional order model for Covid-19 and TB co-infection. Rezapour et al. [4] provided a SEIR epidemiological model using Caputo to predict the spread of Covid-19, using the approximate solution of the model was obtained using the fractional order Euler method. Tuan et al. [5] gave a mathematical model for the spread of Covid-19 by means of the Caputo fractional order derivatives, obtained an approximate solution using the generalized Adams-Bashforth-Moulton method and finally performed numerical simulations of its spread over the world. Hasib Khan et al. [6] investigated the existence of solutions, Hyers-Ulam stability and computational results for the Covid-19 mathematical model in the fractal-fractional order sense. For more results, one can see [8–19].

* Corresponding author at: Department of Arts and Sciences, Suqian College.
E-mail address: liouxianghu04@126.com (X. Liu).

https://doi.org/10.1016/j.heliyon.2022.e12670
Received 30 May 2022; Received in revised form 21 November 2022; Accepted 19 December 2022

Available online 20 January 2023

Inspired by the above literature, this paper studies the generalized FCSEIR model (Fractional-Covid-19-Susceptible-Exposed-Infectious-Recovered), i.e. a compartmentalised model of susceptibility-asymptomatic infection-sickness-recovery that takes into account factors such as population and climate.

2. Platform design

Since the outbreak of the novel coronavirus in December 2019, project members have acted with urgency to conduct preliminary studies on the novel coronavirus and local temperatures, characteristics of susceptible populations, topographical features, and trajectories of population movements. Collecting and collating the findings with relevant authoritative reports, we set the parameters of the FSEIR prediction model according to age segmentation, climatic conditions, environmental conditions, and the population density and size of human traffic in different locations between different cities, in order to achieve a higher degree of simulation, lower calculation error, etc.

The programming language for this thesis is Python, which was designed by Dutch academic Guido van Rossum in 1990. Its efficient high-level data structures and easy to extend, with the ability to use functions and data types extended by C or C++, have made it an increasingly popular programming language in recent years. The platform development process uses tools such as PyQt5, PyQtChart and Matplotlib in Python for GUI application design and data visualisation programming. PyQt5 was used first to build the GUI framework, PyQtChart to draw line graphs of basic functions, then Matplotlib was used in the core algorithm to build complex algorithms and depth calculations and to draw images, SQL was used to enter and save the collected data, and finally Qt Creator was used to typeset and beautify the platform's interface. Each line of code was designed to take up as little space as possible, to be as fast as possible, to be as accurate as possible, and to be as user-friendly as possible.

3. Model, parameters, data

3.1. General model concepts

Susceptible (S) refers to a healthy person who lacks immunity to certain viruses, such as a certain probability of infection after contact with an infected person.

Exposed (E) refers to a person who has been in contact with an infected person but is not infectious for the time being. Such infectious virus has a certain incubation period.

Infectious (I) refers to an infected person with strong contagiousness, and when it comes into contact with S, it has a high probability of spreading the virus to S, making it E or I.

Recovered (R) refers to people who are immune after the infected person recovers. If such a virus is a life-long immune infectious virus, the recovered person cannot be changed to S, E or I again. If the immunity period is limited, the recovered person can change to S again.

The establishment of FCSEIR model: Characteristics of the Covid-19 transmission model: There are 4 groups of people in this model: susceptible, sick, recovered, and exposed. The virus is in an incubation period and the sick get lifelong immunity after being cured.

3.2. Model setting

When the susceptible person is in effective contact with the sick person, the susceptible person becomes the exposed person, the exposed person becomes the sick person after the average incubation period of the virus, the sick person can be cured and becomes the recovered person, and Those who recover are immune to the virus for life, and some ideas come from [7].

Suppose the isolation ratio is q, the infection probability is β, and the contact rate is c. The conversion rate of susceptible S to isolated susceptible S_q, isolated latent E_q, latent E is $cq(l-\beta)$, $c\beta q$ and $c\beta(l-q)$. At the same time, considering the impact of the non-isolated infected person I and latent person E on the susceptible population. The susceptible person S_q released from isolation is re-transformed to S. Due to the generally low climate in the country in the first quarter of 2022, the colder weather makes the inherent defense capacity of the respiratory tract is reduced, the virus stays in the respiratory tract for a longer time, and the probability of infection increases. The population and climate factors were $\mu = 0.01$, and the contagion model was determined to be FCSEIR, as showed in Fig. 1.

The fractional order governing equation of the number of susceptible persons is: $D_s^\alpha = -[c\beta + cq(1-\beta)]S(I+\theta E) + pS_q + \mu$, where D_s^α represents the fractional derivative of Caputo type S with respect to t, θ is the ratio of the infectivity of the latent to the infected, and it is assumed that the infectivity of the latent patients and the symptomatic patients is the same, that is, $\theta = 1$, p is the rate of isolation release, set $p = 1/14$ (quarantine duration is 14 days).

$$\begin{cases} D_S^\alpha = -[c\beta + cq(1-\beta)]S(I+\theta E) + \lambda S_q + \mu \\ D_E^\alpha = c\beta(1-q)S(I+\theta E) - \sigma E \\ D_I^\alpha = \sigma E - (\delta_1 + \alpha + \gamma_1)I \\ D_{S_q}^\alpha = cq(1-\beta)S(I+\theta E) - \lambda S_q \\ D_{E_q}^\alpha = c\beta q S(I+\theta E) - \delta_q E_q \\ D_H^\alpha = \delta_1 I + \delta_q E_q - (\alpha + \gamma_H)H \\ D_R^\alpha = \gamma_1 I + \gamma_H H \end{cases} \tag{3.1}$$

Fig. 1. FCSEIR model

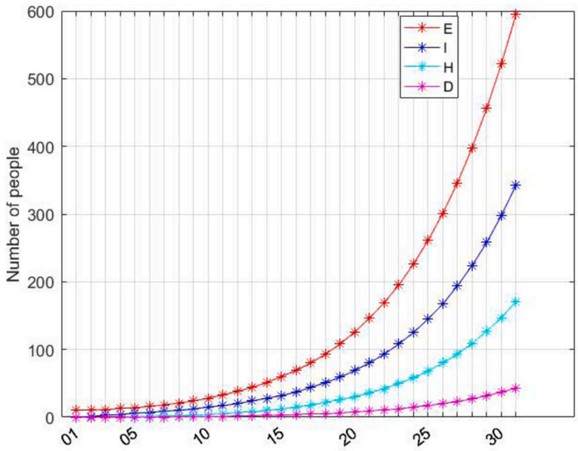

Fig. 2. FCSEIR model data simulation (random)

Among them, σ is the transformation rate of the latent person to the infected person, taking $\sigma = 1/14$ (the incubation period is 14 days), α is the mortality rate, δ_1 is the isolation rate of the infected person, and γ_1 is the recovery rate of the infected person. δ_q is the conversion rate of isolated latents to isolated infected, and γ_H is the recovery rate of isolated infected. The above model is called the generalized FSEIR model. According to the FCSEIR differential equation model, substituting $S = 9990, E = 10, \alpha = 0.2, \sigma = 1/14, \delta_1 = 0.7, \gamma_1 = 0.8, \delta_q = 0.143, \gamma_H = 0.1$ into the algorithm to calculate the virus transmission infection. The model can be solved numerically with given parameter values. It can be seen from the Fig. 2 that if mandatory isolation is not established and residents are not restricted from entering and leaving, the number of infected people will increase exponentially, and the spread of the virus will be difficult to control. Therefore, relevant government departments should make emergency plans for the spread of the virus to effectively control the spread of the virus.

4. Numerical simulation

In order to test the effectiveness of the algorithm, this section will use the following examples to verify. We used the FCSEIR model to conduct a study and analysis on the data of patients with Covid-19 in Shanghai. The data collection period was taken from March 1, 2022 to April 30, 2022, a total of 61 days. As a potent economic city in China, Shanghai has a large population base, complex population flow, and severe epidemics, which may lead to a relative shortage of medical resources and manpower. Compared with the strains spread in other provinces in China, the strain of this large-scale infection is Omicron, which mutates faster, spreads faster, makes a wider range, and is more harmful to people's health. When the epidemic began to spread, it was difficult for relevant government departments and citizens to deal effectively and immediately. Therefore, we divided the outbreak in Shanghai into four stages.

The first stage is when the first infected person spreads the virus without knowing he or she is carrying the virus. At this time, the government's control is relatively lax, and the protection measures for citizens are relatively low. The duration is 14 days. The second stage is when the first positive case is confirmed, the relevant government departments will control it, citizens will take preliminary protective measures, and the susceptible will gradually become the sick or exposed, and the duration will be 14 days. The third stage is a large-scale virus outbreak. The government has stepped up control and citizens have greatly increased prevention and control measures. The duration is 14 days. The fourth stage is the stage of compulsory isolation for the whole people. The number of sick people continues to grow. The government uses compulsory isolation to prevent the further spread of the virus.

Based on the global COVID-19 data and the development of the local epidemic situation in Shanghai, the parameters in FCSEIR are initially set to $\sigma = 0.071, \alpha = 0.013, \gamma_H = 0.274, \gamma_1 = 0.013, S = 95000000$. Since there had been hidden transmission in the population for several weeks before the outbreak of the new crown epidemic, we set $E = 500, I = 100, R = 0$. In the first stage, due to the relatively weak control and protective measures in this situation, but the number of patients is relatively low, based on the infection rate of

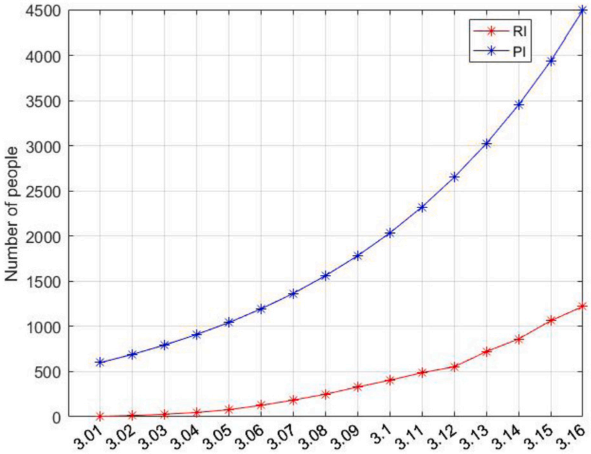

Fig. 3. FCSEIR model data simulation (3.01-3.16)

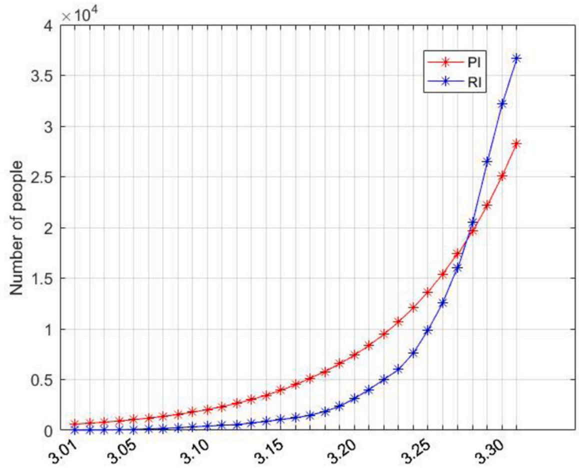

Fig. 4. FCSEIR model data simulation (3.01-3.31)

0.15 provided by the WHO, considering the dense population in Shanghai, there will be a greater risk of transmission, so the infection rate is set to $\delta_q = 0.2$, and the isolation rate is set to $\delta_1 = 0.15$ for a period of 16 days. Under the current parameters, we can obtain the simulation Fig. 3, RI and PI are the actual number of patients and the number of patients predicted by the model, respectively, and the following is the same definitions.

Depending on the prediction of the FCSEIR model algorithm, 16 days after entering the first stage, $E = 4492, I = 1221$. The number of latent people is increasing exponentially, and the number of sick people is also gradually increasing. It can be seen from Fig. 3 that the actual number of infected people detected is lower than the number predicted by the model. The reason is that Shanghai's initial testing capacity is limited and isolation measures are not sufficient, resulting in large errors in the data.

In the second stage, the relevant government departments have carried out a preliminary investigation and isolation according to the movement trajectory of the confirmed person and the contact person, and strengthened the control. Citizens received relevant reports and adopted certain protective measures, but the number of sick people and the number of latent people increased at this time. There is a greater risk of infection, so we do not change the infection rate in the model, increase the isolation strength and set it to $\delta_1 = 0.2$ for 15 days, under the current parameters, we can obtain the simulation Fig. 4.

According to the prediction of the SEIR model algorithm, 15 days after the end of the first stage, $E = 29259, I = 8769$. This is the rising stage of the Shanghai epidemic. The number of sick people is growing rapidly, and the growth rate of the latent population has achieved its peak. It can be seen from Fig. 4 that the government dispatched medical staff from all over the country for testing and treatment while the isolation intensity was increasing. Therefore, the actual number of infections showed a sharp increase., which is still show non-linear growth.

In the third stage, the relevant government departments conduct mandatory home isolation for citizens. We have increased the isolation intensity and set it to $\delta_1 = 0.3$. Although citizens are under the age of mandatory isolation, the base of sick and latent populations is very large, and there is no difference between people. There may be no contact at all, causing the risk of infection to remain high, so continue to maintain the existing infection rate for 30 days. Under the current parameters, we can

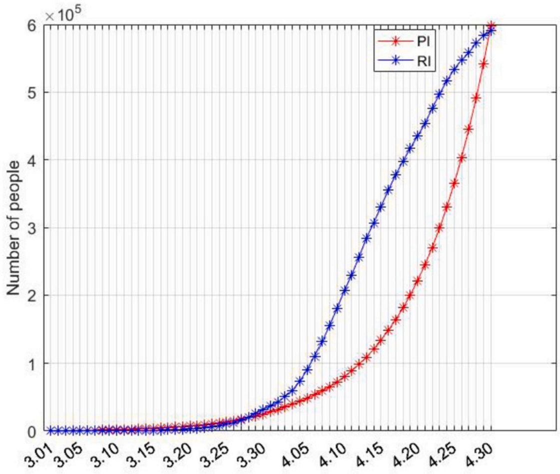

Fig. 5. FCSEIR model data simulation (3.01-4.30)

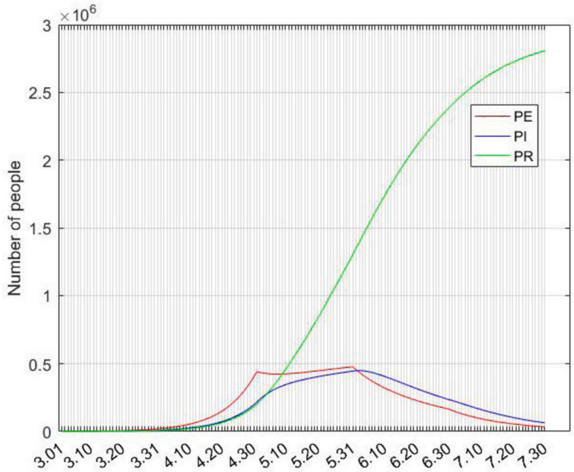

Fig. 6. FCSEIR model data simulation (3.01-7.30)

obtain the simulation in Fig. 5. According to the prediction of the FCSEIR model algorithm, 30 days after the end of the second stage, $E = 401234$, $I = 196774$. Depending on the real-time data of the National Health Commission, Shanghai had a total of 53,375 confirmed positive infections on April 30, a total of 537,730 asymptomatic infections, and a total of 591,105 confirmed cases as of April 30. There are 598,008 people and lurkers in total, and the calculation error is about 1.16%. The error value is less than 5%, which is within the acceptable range. It can also be seen from the real data in Fig. 5 that in April, the number of real infections increased by more than 20,000 cases per day, and the growth trend was violent, which fully shows that the hidden spread of the virus in the population is very serious.

The fourth stage, the ideal prediction stage, we predict that the Shanghai epidemic will increase to 0 after 30 days of the outbreak, that is, the inflection point of the function image will be reached. Under the concerted efforts of relevant government departments and citizens to fight the epidemic, the latent population will be controlled by means of nucleic acid testing for all people. Separate isolation, we divide this stage into two periods, the infection rate of the first period is set to $\delta_q = 0.08$, the isolation strength is set to $\delta_1 = 0.6$, and the period is 30 days, the infection rate of the second period is set to $\delta_q = 0.05$, and the isolation strength is set to for $\delta_1 = 0.6$, for a period of 30 days, we can obtain the simulation Fig. 6 under the current parameters, where PE, PI, PR are predicted Exposed, predicted Infectious, predicted Recovered respectively.

In the prediction chart of the FCSEIR model algorithm, we can note that the number of infected people did not increase sharply on April 30 and the number of recoveries continues to increase, as is the current reality. After the balance period in May, the turning point appeared in mid-June, and the epidemic was completed in early August. After the establishment of perfect isolation measures, in the next 60 days, the latent and sick people will decrease rapidly, and the goal of clearing in Shanghai will soon be achieved.

5. Conclusions

In this paper, a simple and efficient prediction software is designed using mathematical algorithms and big data software. This study acts as a propeller for the application of prediction of big data in the direction of prevention and control field, and also provides more data support for infectious disease prevention and control managers based on the continuous development and exploration of the platform.

The fractional-order model used in this paper is relatively simple and does not take into account several other factors, such as flexible epidemic prevention policies and environment, which will be the direction of our future research.

Funding statement

The work is supported by funds of Guizhou Province (Qian Jiao He KY Zi [2020]093), Qian Ke He Ping Tai Ren Cai [2018]5784-08, Innovation and Entrepreneurship Training Program for College Students 202110664021, Suqian College School 2022XRC067, SuQian Sci&Tech Program (Grant No. K202225).

CRediT authorship contribution statement

Yanfang Li: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data. **Xianghu Liu:** Contributed reagents, materials, analysis tools or data; Wrote the paper.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data included in article/supp. material/referenced in article.

Acknowledgements

We would like to express our gratitude to all those who helped me during the writing of this thesis.

References

[1] M. Das, G. Samanta, Stability analysis of a fractional ordered COVID-19 model, Comput. Math. Biophys. 9 (2021) 22–45.
[2] I. Baba, B. Nasidi, Fractional order model for the role of mild cases in the transmission of COVID-19, Chaos Solitons Fractals 142 (2021) 10.
[3] A. Omame, M. Abbas, C. Onyenegecha, A fractional-order model for COVID-19 and tuberculosis co-infection using Atangana-Baleanu derivative, Chaos Solitons Fractals (2021) 13.
[4] S. Rezapour, H. Mohammadi, M. Samei, SEIR epidemic model for COVID-19 transmission by Caputo derivative of fractional order, Adv. Differ. Equ. 490 (2020) 19.
[5] N. Tuan, H. Mohammadi, S. Rezapour, A mathematical model for COVID-19 transmission by using the Caputo fractional derivative, Chaos Solitons Fractals 140 (2020) 11.
[6] K. Hasib, A. Farooq, et al., On fractal-fractional Covid-19 mathematical model, Chaos Solitons Fractals 157 (2022) 111937.
[7] C. Jingan, l. Jinlong, G. Songbai, Dynamical model of emerging infectious diseases-applied to COVID-19 transmission, Acta Math. Appl. Sin. 43 (2) (2020) 147–155.
[8] National Health Commission, As of 24:00 on April 4, the latest situation of the new coronavirus, pneumonia [EB/OL], [2022-04-30].
[9] A. Kouidere, L. Youssoufifi, et al., Optimal control of mathematical modeling of the spread of the COVID-19 pandemic with high lighting the negative impact of quarantine on diabetics people with cost-effectiveness, Chaos Solitons Fractals 145 (2021) 110777.
[10] O. Kolebaje, O. Vincent, U. Vincent, P. McClintock, Nonlinear growth and mathematical modelling of COVID-19 in some African countries with the Atangana-Baleanu fractional derivative, Commun. Nonlinear Sci. Numer. Simul. 105 (2022) 27.
[11] R. Ullah, M. Waseem, N. Rosli, J. Kaflfle, Analysis of COVID-19 fractional model pertaining to the Atangana-Baleanu-Caputo fractional derivatives, J. Funct. Spaces (2021) 16.
[12] M. Das, G. Samanta, Stability analysis of a fractional ordered COVID-19 model, Comput. Math. Biophys. 9 (2021) 22–45.
[13] K. Furati, I. Sarumi, A. Khaliq, Fractional model for the spread of COVID-19 subject to government intervention and public perception, Appl. Math. Model. 95 (2021) 89–105.
[14] I. Baba, B. Nasidi, Fractional order model for the role of mild cases in the transmission of COVID-19, Chaos Solitons Fractals 142 (2021) 10.
[15] A. Akgl, N. Ahmed, A. Raza, et al., A fractal fractional model for cervical cancer due to human papillomavirus infection, Chaos Solitons Fractals 29 (5) (2021) 2140015-19.
[16] A. Akgl, S. Ahmad, et al., A novel method for analysing the fractal fractional integrator circuit, Alex. Eng. J. 60 (4) (2021) 3721–3729.
[17] R. Ibrahim, D. Baleanu, Analytic solution of the Langevin differential equations dominated by a multibrot fractal set, Fractal Fract. 5 (2) (2021) 50.
[18] A. Tassaddiq, S. Qureshi, et al., A new three-step root-fifinding numerical method and its fractal global behavior, Fractal Fract. 5 (4) (2021) 204.
[19] A. Akgl, D. Baleanu, Analysis and applications of the proportional Caputo derivative, Adv. Differ. Equ. 1 (2021) 1–12.