medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Prediction of 2019-nCov in Italy based on PSO and inversion analysis

Sun Siyi[1], Zheng Yangping[2]

1 Information and Computational Science Department, University of Science and Technology Beijing, Beijing 100083 CHN
2 College of Physics, Sichuan University, Chengdu, Chengdu 610065 CHN

Abstract

Novel coronavirus (2019-nCov) has swept the world, and all of the world have been harmful. This article makes prediction and suggestions for the Italy. Up to March 11, 2020, 2019-nCov thoroughly broke out in Italy with over 10,000 confirmed cases notwithstanding the gradually block of the country since March 9, 2020. Estimation of possible infection population and prospective suggestion of handling spread based on exist data are of crucial importance. Considering of the biology parameters obtained based on Chinese clinical data in Wuhan, other scholars' work and real spread feature of 2019-nCov in Italy, we built a more applicable model called SEIJR with log-normal distributed time delay to forecast the trend of spreading. Adopting Particle Swarm Optimization (PSO), we estimated the early period average spreading velocity ($\boldsymbol{\alpha_0}$) and conducted inversion analysis of time point ($\boldsymbol{T_0}$) when the virus first hit the Italy. Based on fixed $\boldsymbol{\alpha_0}$ and $\boldsymbol{T_0}$, we then obtained the average spreading velocity $\boldsymbol{\alpha_1}$ after the lock by PSO. For the aim of offering expeditious advice, we generated the prediction trends with different $\boldsymbol{\alpha}$ which we considered would be helpful in addressing the infection. Not only solved the complex, nondifferentiable equation of epidemic model, our research also performs well in inversion analysis based on PSO which conveys informative outcomes for further discussion on precatious action. To conclude, the first day of spread is around February 1, 2020 with the early period average spreading velocity α_0=0.330 which is higher than most cities in China except Wuhan. After locking the country and attaching great attention to public precaution, the α_1 sharply descended to 0.278, indicting the effectiveness of these measures. Furthermore, in order to cope the disease before mid-April, take actions to control the under 0.25 is necessary. Code can be freely downloaded from https://github.com/Summerwork/2019-nCov-Prediction.

1 Introduction

A global epidemic disease known as the novel coronavirus (2019-nCov) had seriously hit the most area around the whole world causing unpredictable loss of manpower and finance during the first quarter of 2020 [21] [1]. The first confirmed 2019-nCov case was reported in Wuhan, China on December 31, 2019 (World Health Organization, 2020a). As of March 19 (21:00 GMT),2020, 2019-nCov has resulted in 81262 confirmed cases and 3250 dead cased in China cumulatively (National Health Commission of the People's Republic of China, 2020). China as the first country faced by the outbreak of the severe disease, it took strict but effective action to contain the spread of 2019-nCov and attained apparent success till now. [21] Many related works have been done in prediction and precaution via constructing proper model and analyzing parameters. [21] [20] [19]

While in the other parts worldwide, the menacing disease just became to spread [7], especially in countries with no preparation and experienced measures for suppressing the possible large-scale infection. In this article, we take Italy which is now experiencing severe situation of 2019-nCov as example to conduct analysis with the aim of offering utilizable suggestion. In the first period (before March 9, 2020), we attempt to inverse the virus spread timeline in Italy and the early period average spreading velocity by adopting PSO to optimize the parameters based on our SEIJR model and existing data (European Centre

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

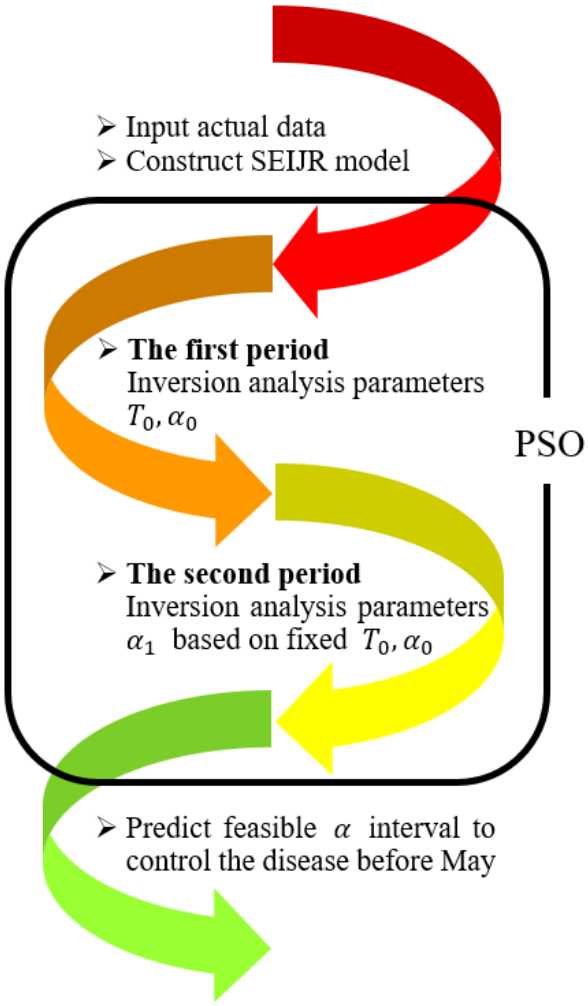


Figure 1. Flow chart

for Disease Prevention and Control). In the second period (only use data from March 9-16, 2020), we optimize the average spreading velocity in order to show the effect of country blockade. In the last period of our research, we demonstrate latent trends of confirmed cases with various average spreading velocity which of vital importance in controlling the disease. The whole flow chart is illustrated in Figure 1.

2 Data and Model

2.1 Data Collection and Processing

We collected the daily reported confirmed diagnosed data from the website of European Centre for Disease Prevention and Control (European CDC) . All these data is public for everyone.

Based on the need of our analysis, we preprocessed the data by adding the daily reported confirmed diagnosed cases to obtain the accumulative amount for following inversion of parameters and prediction. Here are details about the processed data we used in this article (Table 1).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Table 1.
New Confirmed and Accumulative Confirmed Data

Date	New Confirmed	Accumulative Confirmed
2020-02-22	14	14
2020-02-23	62	76
2020-02-24	53	129
2020-02-25	97	226
2020-02-26	93	319
2020-02-27	78	397
2020-02-28	250	647
2020-02-29	238	885
2020-03-01	240	1125
2020-03-02	561	1686
2020-03-03	146	1832
2020-03-04	667	2499
2020-03-05	587	3086
2020-03-06	769	3855
2020-03-07	778	4633
2020-03-08	1247	5880
2020-03-09	1492	7372
2020-03-10	1797	9169
2020-03-11	977	10146
2020-03-12	2313	12459
2020-03-13	2651	15110
2020-03-14	2547	17657
2020-03-15	90 [a]	17747
2020-03-16	6230	23977
2020-03-17	4000	27977
2020-03-18	3526	31503
2020-03-19	4207	35710
2020-03-20	5322	41032
2020-03-21	5986	47018
2020-03-22	6557	52575
2020-03-23	5560	59135
2020-03-24	4789	63924
2020-03-25	5249	69173

[a]The data downloaded from European CDC of March 15, 2020 is indeed 90 persons, while we not sure it consistent with the real situation.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

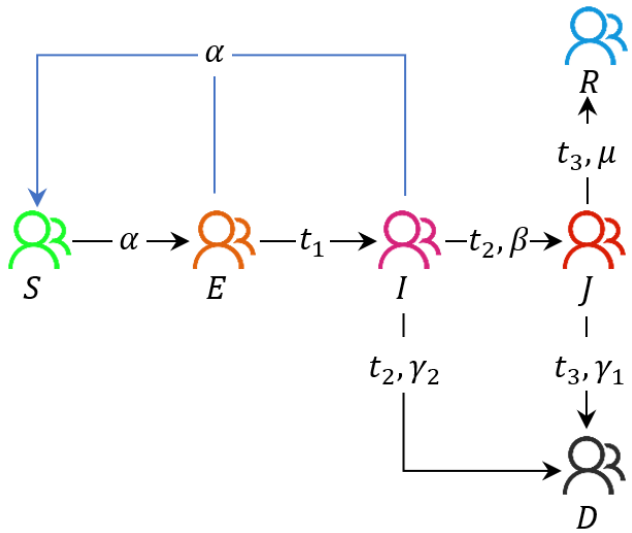


Figure 2. SEIJR model

2.2 SEIJR Model

Disease transmission is a complex process with multiple variables and uncertainties making it unable to be accurately solved and predicted [16], [11] [6]. However, models are feasible in forecasting for infectious diseases when different characteristics parameters [18] [17] like transmission mode, immunization mode, mortality and average spreading velocity are offered. Classical models for infectious diseases include SIR model [2] and SEIR model, etc. Considering the actual situation in Italy and the transmission characteristics of 2019-nCov obtained in China, this paper built the SEIJR model with log-normal distributed time-delay terms [15] [3]based on the SEIR model [14] [12]. Figure 2 illustrates the SEIJR model. The model describes the problem by assuming population consists of six types (accumulative value): susceptible population S, exposed population E, infectious population I, confirmed J, recovered population R and dead population D. α is average spreading velocity, β is diagnose rate, γ_1, $\gamma 2$ are die rate and μ is cure rate. $t_1(t)$ is the time of incubation period [5] E need to become I, $t_2(t)$ is the time of waiting period I need to become J and $t_3(t)$ is the duration of hospitalization [5] J need to become R. $S(t)$, $E(t)$, $I(t)$, $J(t)$, $R(t)$ and $D(t)$ are dependent variables of time t, respectively. α, β, γ, μ are constants only related to the actual situation. Time-delay t_1,t_2,t_3 are only depend on time t, respectively, which can be written as $t_1(t)$, $t_2(t)$ and $t_3(t)$.

2.3 Interpretation

In the model, only E and I have ability to infect S which is a process of contact infection with transient time. E indeed infected but shows no symptoms of 2019-nCov and then transfers to I after an incubation period t_1. I has symptoms like fever, cough and shortness of breath. Because the pre-virus symptoms are not obvious [5] and the uneven medical facilities in Italy, I will be confirmed as J after a period t_2. Due to the seriousness and infectivity of the virus, it can be considered that when it becomes J, J will be immediately isolation and lose its ability of infection. Treatment will be started immediately after confirmed. J will recovery and become R after a duration of hospitalization t_3. Because of the 2019-nCov has certain lethality, it needs some more assumptions:

2.3.1 Assumption 1

The number of deaths during the period of E is too tiny to be considered, only need to consider the mortality during the period of I and J;

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

2.3.2 Assumption 2

The data during the period of I is unavailable. We consider that I to D and I to J have the same delay time t_2 which means the only difference between them is proportion;

2.3.3 Assumption 3

The official organization and medical institutions do not give any information on how long for patients in the treatment stage will die, but there exists clinical information that for a confirmed patient how long the patient is needed to be cured [5]. As the same way J to R, J to D have the same delay time t_3 only with different proportion.

People in I can be diagnosed and receive treatment with β diagnose rate. Otherwise, they will die with the die rate $\gamma_1 = 1-$. The recover rate for J is μ and the dead rate is $\gamma_2 = 1 - \mu$. The total mortality proportion is $\gamma = \gamma_1 + \gamma_2$.

2.3.4 Assumption 4

After recovering, people in R will not go out because they are in a frail state and will be considered as isolation.

2.4 Time-delay Function

Combining the basic principles of epidemiology and etiology, it can get that the time t_2 from infection to diagnosis, which approximately follows the log-normal distribution [8]. The assumption can be applied to t_1 and t_2 without loss of generality. For t_1, due to the lack of Italian clinical information statistics, it's more accurate and reliable to use the clinical data of Zhong NS et.al: the median incubation period is 4 days [5], the quartile is 2 and 7 days. With the log-normal distribution's formula:

$$\begin{aligned} P(T_S \leq 4) &= P(\frac{lnT_S - \mu}{\sigma} \leq \frac{ln4 - \mu}{\sigma}) = 0.5, \\ P(T_S \leq 7) &= P(\frac{lnT_S - \mu}{\sigma} \leq \frac{ln7 - \mu}{\sigma}) = 0.75. \end{aligned} \tag{1}$$

Get $\mu = ln4$, $\sigma = \frac{ln7 - ln4}{Z_{0.75}} = 0.8291$ where $Z_{0.75}$ is normal distribution quartile. Then, the log-normal density function corresponding to t_1 is:

$$s_1(t) = \begin{cases} \frac{0.4812}{t} exp(-\frac{(lnt - ln4)^2}{1.3747}) & , t > 0, \\ 0 & , t = 0. \end{cases} \tag{2}$$

For t_2, the data in Italy is still unavailable, but the relevant distribution function is given in the Chinese research report. In the paper of Yang ZW et.al, the relevant statistics of "the time interval between the most recent stay in Hubei Province and the confirmed diagnosis" are given [4]. The end point of this period corresponds to the start point of J in the SEIJR model, but there is no corresponding start point. It can only be known that it's in the middle of the time between t_1 and t_2, it is convincing to regard it as the point when e enters I, which means the time given in that article corresponds to the t_2 in this article. As the same way, the log-normal density function corresponding to t_2 is:

$$s_2(t) = \begin{cases} \frac{1.0224}{t} exp(-\frac{(lnt - ln10)^2}{0.3045}) & , t > 0, \\ 0 & , t = 0. \end{cases} \tag{3}$$

For t_3, the Italian official organizations and medical institutions still lack clinical information and official statistics, so we also assumes that t_3 follows the log-normal distribution. However, Zhong NS et.al gave some relevant data: the duration of hospitalization which is t_3, the median is 12 days, and the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

quartiles are 10 and 14 days [5]. Adopting the same method used in t_1 estimation, we can get the distribution density function corresponding to t_3 is:

$$s_3(t) = \begin{cases} \frac{1.7469}{t} exp(-\frac{(lnt - ln12)^2}{0.1043}) & , t > 0, \\ 0 & , t = 0. \end{cases} \tag{4}$$

Now give the definition of log-distributed normal time-delay function in SEIJR model:

Definition 1 (Time-delay Function in SEIJR Model) *The time-delay function describes how many people change this stage from the previous stage at time t.*

The general formula is written as:

$$\Gamma(x(t), s_i(t)) = \int_0^t x(t - t') \dots s_i(t')dt', i = 1, 2, 3 \tag{5}$$

where $x(t)$ is the term which has time-delay, and $s_i(t)$ is the distribution density mentioned before. Here we finally get the precise differential equation of SEIJR model as:

$$\begin{aligned} \frac{dS}{dt} &= -\alpha(E(t) + I(t))\frac{S(t)}{S(t) + E(t) + I(t)}, \\ \frac{dE}{dt} &= \alpha(E(t) + I(t))\frac{S(t)}{S(t) + E(t) + I(t)} - \Gamma(E(t), s_1(t)), \\ \frac{dI}{dt} &= \Gamma(E(t), s_1(t)) - \Gamma(I(t), s_2(t)), \\ \frac{dJ}{dt} &= \beta\Gamma(I(t), s_2(t)) - \Gamma(J(t), s_3(t)), \\ \frac{dR}{dt} &= \mu\Gamma(J(t), s_3(t)), \\ \frac{dJ}{dt} &= \gamma_1\Gamma(I(t), s_2(t)) - \gamma_2\Gamma(J(t), s_3(t)). \end{aligned} \tag{6}$$

3 Methods

3.1 Runge-Kutta Methods

The previous differential Eq. (6) corresponding to the SEIJR model is a form with integrals and independent variables on the integral limit which from the time-delay function term,so there is no analytical solution. For this case, numerical methods are useful. We combine iteration and degree four Runge-Kutta to generate the numerical solution with the initial condition. Here is the main principles of Runge–Kutta fourth-order method [8], let the differential equation have the form as follow:

$$\begin{aligned} \frac{dy}{dt} &= f(t, y), \\ y(t_0) &= y_0. \end{aligned}$$

Then its iterative formula is:

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

where

$$
\begin{aligned}
k_1 &= f(t_n, y_n),\\
k_2 &= f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1),\\
k_3 &= f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_2),\\
k_4 &= f(t_n + h, y_n + hk_3).
\end{aligned}
$$

3.2 Particle Swarm Optimization

Usually the epidemic model's descriptive function was derivative-based without considering time delay or just assume fixed linear time delay, such as SIR model. In this article, the SEIJR model solved by combining iteration and degree four Runge-Kutta mentioned before. Accounting for this, the function of least square method (LSE) employed in addressing this problem for different periods are denoted as follow:

$$
\begin{aligned}
f_1(\alpha_0, T_0) &= \| J_{pre}(\alpha_0, T_0) - J_{act} \|_2,\\
f_2(\alpha_1) &= \| J_{pre}(\alpha_1) - J_{act} \|_2
\end{aligned}
\tag{7}
$$

where f_1 is LSE of the first period. f_2 is LSE of the second period. $J^*_{pre1}(\alpha_0, T_0)$, J_{act1} are predicted and actual value from February 22 to March 9, 2020. $J^*_{pre2}(\alpha_1)$, J_{act2} are predicted and actual value from March 9 to March 25, 2020. Both of them are nonlinear, complex, discontinuous and nondifferentiable [13]. Traditional method based on derivation is infeasible for minimization. While approaches such as annealing algorithm [10] to search for parameters is more calculative expensive than the particle swarm optimization (PSO).

PSO [14] is a population-based search algorithm sparked by the forage behavior of birds within a flock. Individuals gain the ability of searching for better solution areas by learning the fitness information from the environment. In PSO algorithm, the velocity of individual is dynamically changed considering its previous flying experience.

The algorithm consists of three main parts: individual best, global best and individual optimization based on the best particle of whole population [9]. In this article, f_1 and f_2 are the fitness function during the first and second period. The main process of conducting this algorithm and basic parameters are respectively shown in Figure 3 and Table 2.

Table 2.
Parameters for PSO

Parameter	Value
c_1	2
c_2	2
ω	0.6

4 Results

4.1 The first period

During this period of our work, we adopt PSO using f_1 as its fitness function to generate the optimal α_0,T_0. As shown in the Figure 4, the best result is T_0=21 (after rounding) and α_0=0.33. After obtained the greatest α_0,T_0, we draw the prediction curves for further comparison in the following section. Illustrated in Figure 5a is the prediction value J^*_{pre1} with the optimal parameter and the actually value

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

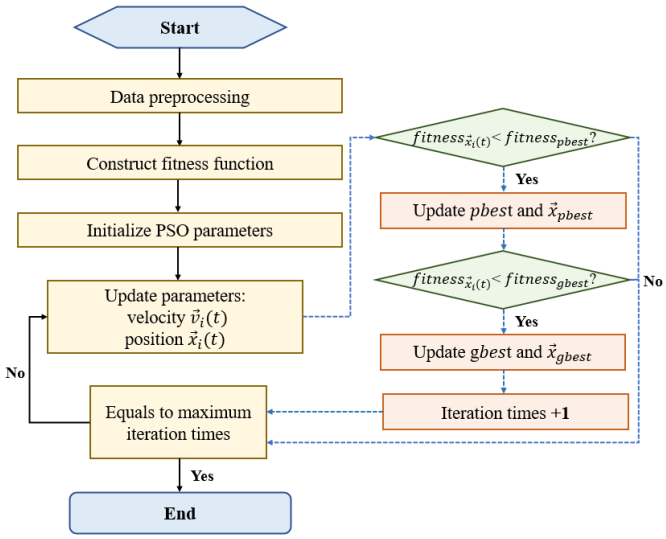


Figure 3. PSO flow chart

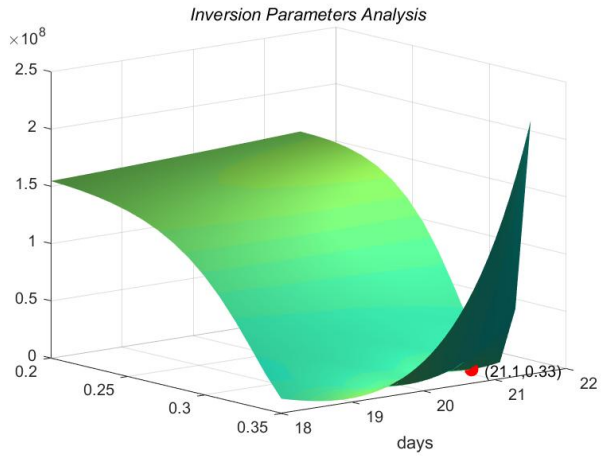
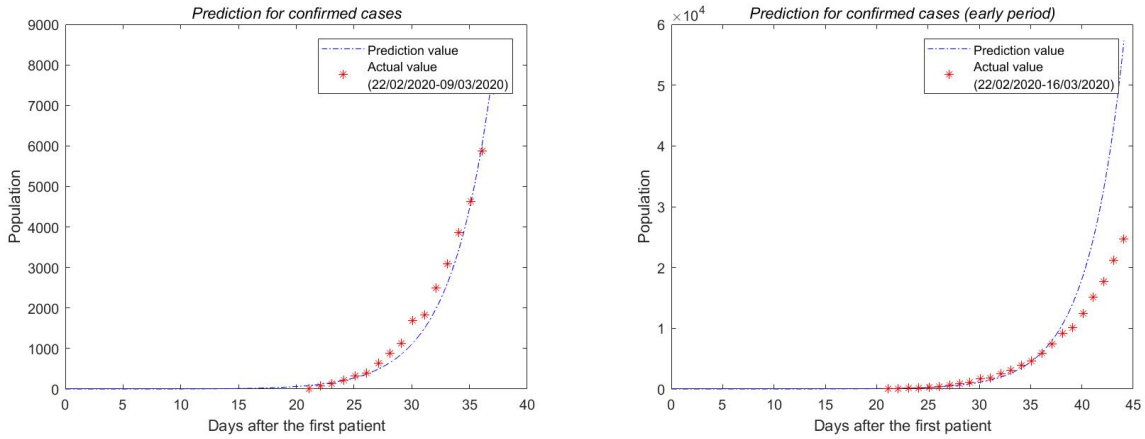


Figure 4. Vistualization of PSO searched results

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .



(a) (b)

Figure 5.
J^*_{pre1} and J_{act1},J_{act} with α_0=0.330

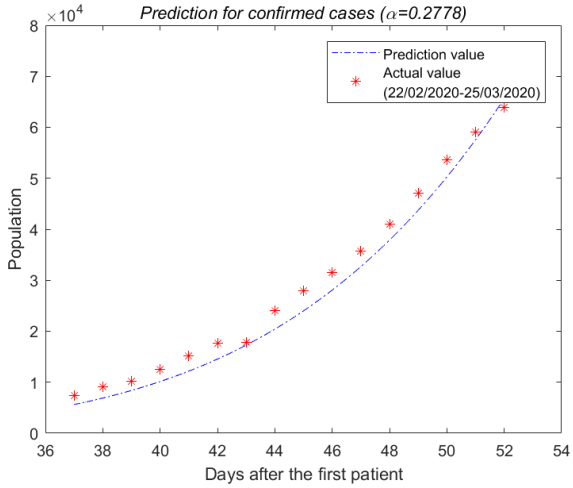


Figure 6. J^*_{pre1} and J_{act2} with α_0=0.278

J_{act1}; (b) is the same curve with whole actual value J_{act} if the average spreading velocity remain unchanged.

4.2 The second period

Considering of the precautious action like blockade taken by Italy government and the information Figure 5b conveyed, we assumed the average spreading velocity changed after these actions. Based on fixed α_0,T_0, we applied PSO to the optimization problem and got the result of α_1=0.278 which is cut down from α_0. The Figure 6 shows the predicted population J^*_{pre2} with α_1 from March 9 to 25, 2020 and J_{act2}.

4.3 Predicted J trends with different α

Figure 7 gives prediction curves with different α with the aim for further discussion about controlling the disease before mid-April.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

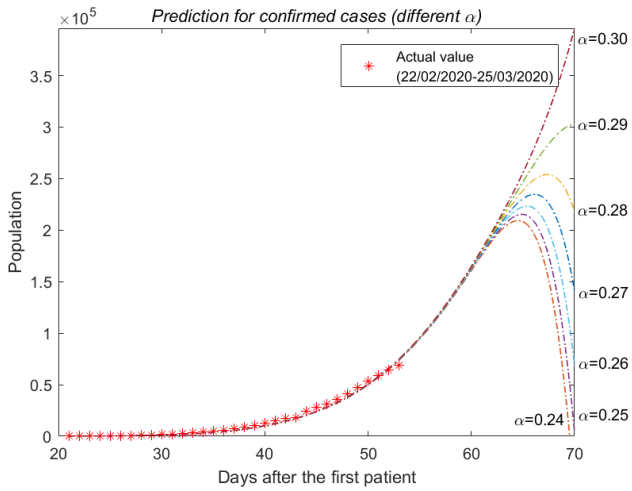


Figure 7. J_{pre} with different α_0

5 Discussion

5.1 Reliability of SEIJR model

It is clearly seen that the fitting curve of J in the SEIJR model fits well with the confirmed diagnosis data of the Italian official statistics. $T_0 = 21$ represents the initial value of the model which indicates the first E appeared 21 days before February 22 that is February 1. According to the information reported by the Italian government, the first case in Italy, when J appeared, was January 31. Two patients from Wuhan, China arrived in Italy by air at January 23 and visited other cities in Italy. They finally arrived in Rome, feeling physical discomfort at January 30 and was confirmed diagnosis then isolated at January 31. It is reasonable to speculate that they had carried the 2019-nCov in Wuhan before arriving in Italy in January 23, and continued to spread in Italy for eight days after January 23 until January 31. The first case of E reported in official data appeared at January 23, and the model's conclusion appeared at February 1. In fact, because the two senior travelers are 67 and 66 years old with poor mobility, unfamiliar surroundings and the simple factors of interpersonal relationships, they have a lower α which also leads to real initial time point is earlier than the initial time point of the theoretical model. To sum up, the model has higher accuracy and stronger credibility. [11]

5.2 Effect of blockade

Prime Minister Giuseppe Conte extended the quarantine lockdown to cover all the region of Lombardy and 14 other northern provinces on March 8, and all region on March 10. At the same time, the Italian government further banned rallies and sports activities nationwide, announced a national blockade, and unnecessarily stopped going out. Compared the average spread velocity $\alpha_1 = 0.278$ with $\alpha_0 = 0.330$ at the early stage(before March 9, 2020), a decrease existed after the series of actions which verifies the effectiveness of this measure.

5.3 Advice to Italy government

According to the model's prediction, when α is in a suitable range, the number of confirmed patients will gradually decrease. It is obvious that the smaller the value of α is, the faster the number of confirmed diagnoses decreases. Since Italian government hopes to end the 2019-nCov by mid-April, it can control α around 0.25 according to the prediction of the model. Based on this conclusion, Italy government is supposed to strengthen the isolation, reduce gathering activities and make people understand the importance of precautionary measures to constrain α under 0.25.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

6 Conclusion

Applying PSO to our SEIJR model with log-normal distributed time delay, we obtained the convincing start time (around February 1, 2020) of 2019-nCov and the average spreading velocity (α_0=0.330) at the early stage. We compared the average spreading velocity during the early period and following period, a conspicuous decrease attributed to the effective measures was found. Based on the prediction interval of possible infected population of different α, we strongly recommend Italy to keep α under 0.25 if they want the situation take a turn for the better even ended before mid-April.

Acknowledgment

The authors would like to thank Dongmei Ai and Yan Xu at University of Science and Technology Beijing for helpful suggestions and discussion.

References

1. A. Al-Mandhari, D. Samhouri, A. Abubakar, and R. Brennan. Coronavirus disease 2019 outbreak: preparedness and readiness of countries in the eastern mediterranean region. *Eastern Mediterranean health journal= La revue de sante de la Mediterranee orientale= al-Majallah al-sihhiyah li-sharq al-mutawassit*, 26(2):136, 2020.

2. B. Cantó, C. Coll, and E. Sánchez. Estimation of parameters in a structured sir model. *Advances in Difference Equations*, 2017(1):33, 2017.

3. Y. Chen, J. Cheng, Y. Jiang, and K. Liu. A time delay dynamical model for outbreak of 2019-ncov and the parameter identification. *arXiv preprint arXiv:2002.00418*, 2020.

4. K. J. e. a. Ding Z, Liu Y. A probability model for estimating the expected number of the newly infected and predicting the trend of the diagnosed. *Operations Research Transactions*, 24(1):1–12, 2020.

5. W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D. S. Hui, et al. Clinical characteristics of coronavirus disease 2019 in china. *New England Journal of Medicine*, 2020.

6. X.-N. Han, S. J. De Vlas, L.-Q. Fang, D. Feng, W.-C. Cao, and J. D. F. Habbema. Mathematical modelling of sars and other infectious diseases in china: a review. *Tropical Medicine & International Health*, 14:92–100, 2009.

7. B. Haynes, N. E. Messonnier, and M. S. Cetron. First travel-related case of 2019 novel coronavirus detected in united states: press release, tuesday, january 21, 2020. 2020.

8. J. Kalbfleisch and J. Lawless. Estimating the incubation period for aids patients. *Nature*, 333(6173):504–505, 1988.

9. J. Kennedy and R. Eberhart. Particle swarm optimization. 4:1942–1948, 1995.

10. A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya. Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE transactions on neural networks*, 22(3):337–346, 2010.

11. J. Koopman. Modeling infection transmission. *Annu. Rev. Public Health*, 25:303–326, 2004.

12. P. E. Lekone and B. F. Finkenstädt. Statistical inference in a stochastic epidemic seir model with control intervention: Ebola as a case study. *Biometrics*, 62(4):1170–1177, 2006.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.08.20095869; this version posted December 22, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

13. S. A. Levin, B. Grenfell, A. Hastings, and A. S. Perelson. Mathematical and computational challenges in population biology and ecosystems science. *Science*, 275(5298):334–343, 1997.

14. M. Y. Li and J. S. Muldowney. Global stability for the seir model in epidemiology. *Mathematical biosciences*, 125(2):155–164, 1995.

15. G. Medley, L. Billard, D. R. Cox, and R. M. Anderson. The distribution of the incubation period for the acquired immunodeficiency syndrome (aids). *Proceedings of the Royal society of London. Series B. Biological sciences*, 233(1272):367–377, 1988.

16. T. W. Ng, G. Turinici, and A. Danchin. A double epidemic model for the sars propagation. *BMC Infectious Diseases*, 3(1):19, 2003.

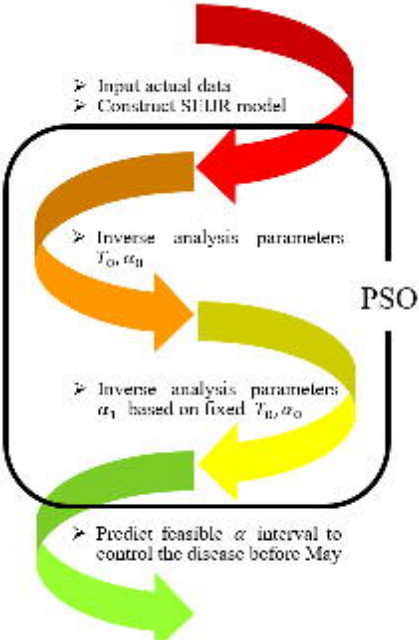
17. J. T. Wu, K. Leung, and G. M. Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395(10225):689–697, 2020.

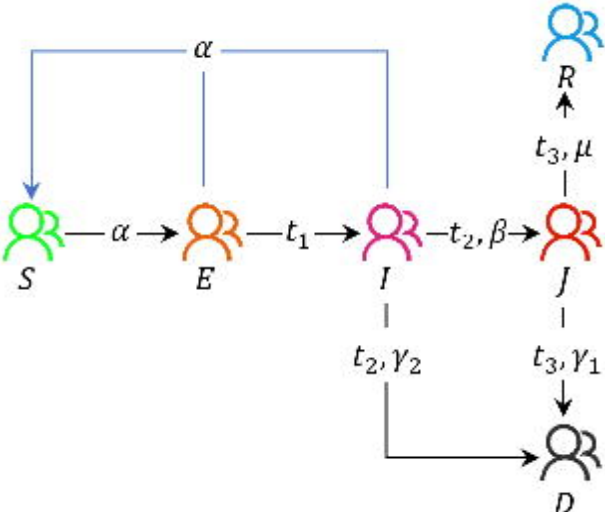
18. P. S. Yip, K. Lam, Y. Xu, P. Chau, J. Xu, W. Chang, Y. Peng, Z. Liu, X. Xie, and H. Lau. Reconstruction of the infection curve for sars epidemic in beijing, china using a back-projection method. *Communications in Statistics—Simulation and Computation®*, 37(2):425–433, 2008.

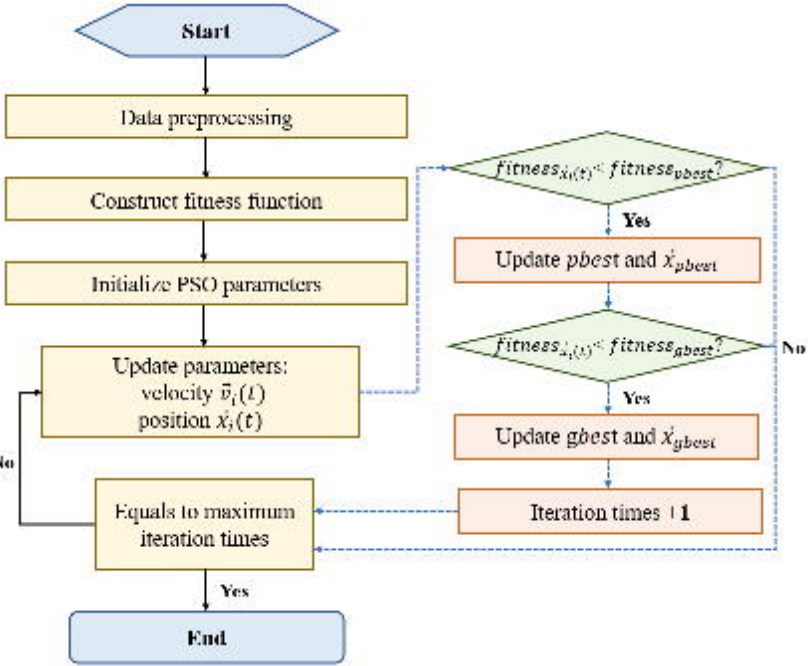
19. S. Zhang, M. Diao, W. Yu, L. Pei, Z. Lin, and D. Chen. Estimation of the reproductive number of novel coronavirus (covid-19) and the probable outbreak size on the diamond princess cruise ship: A data-driven analysis. *International Journal of Infectious Diseases*, 2020.

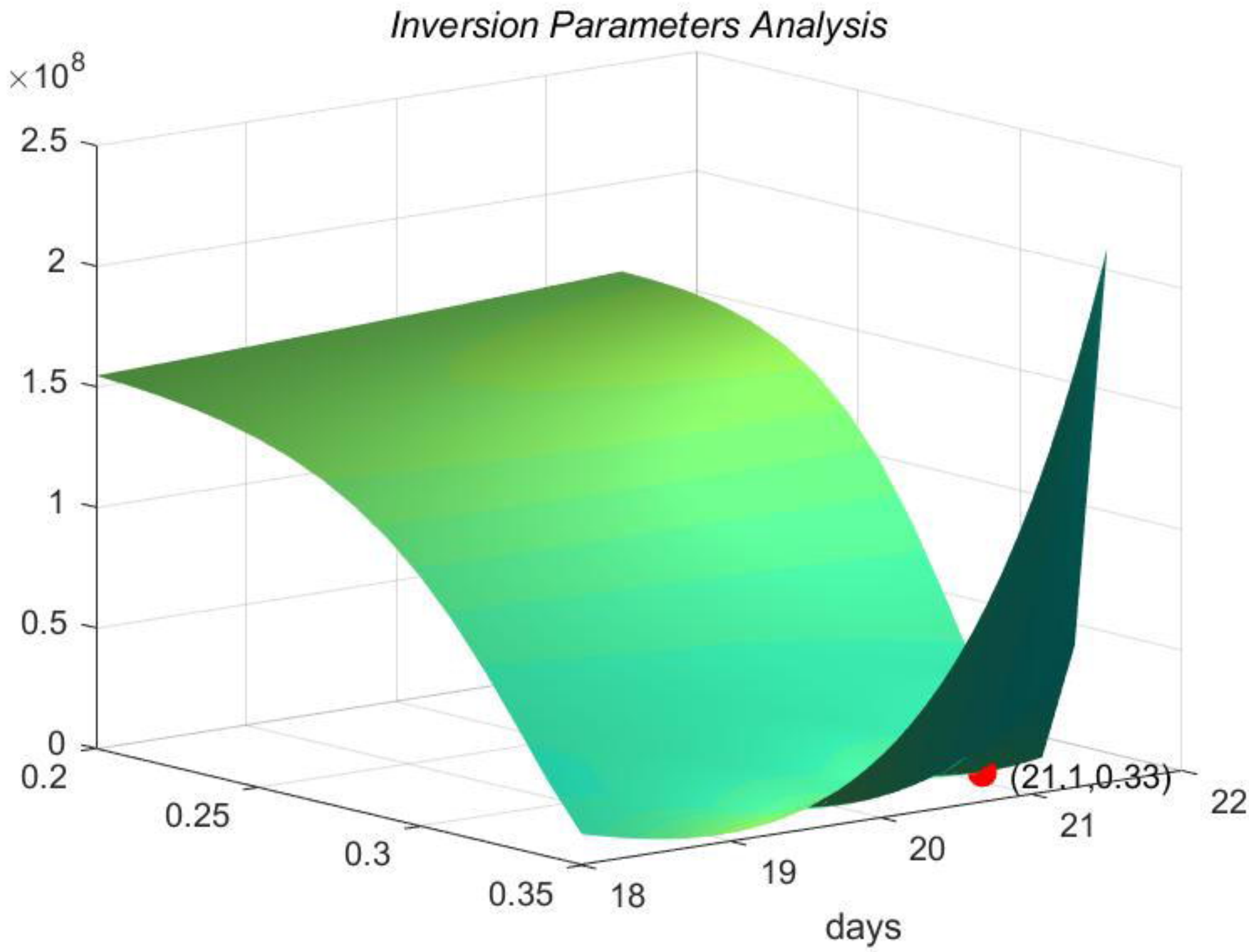
20. S. Zhao, Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, Y. Lou, D. Gao, L. Yang, D. He, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-ncov) in china, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases*, 92:214–217, 2020.

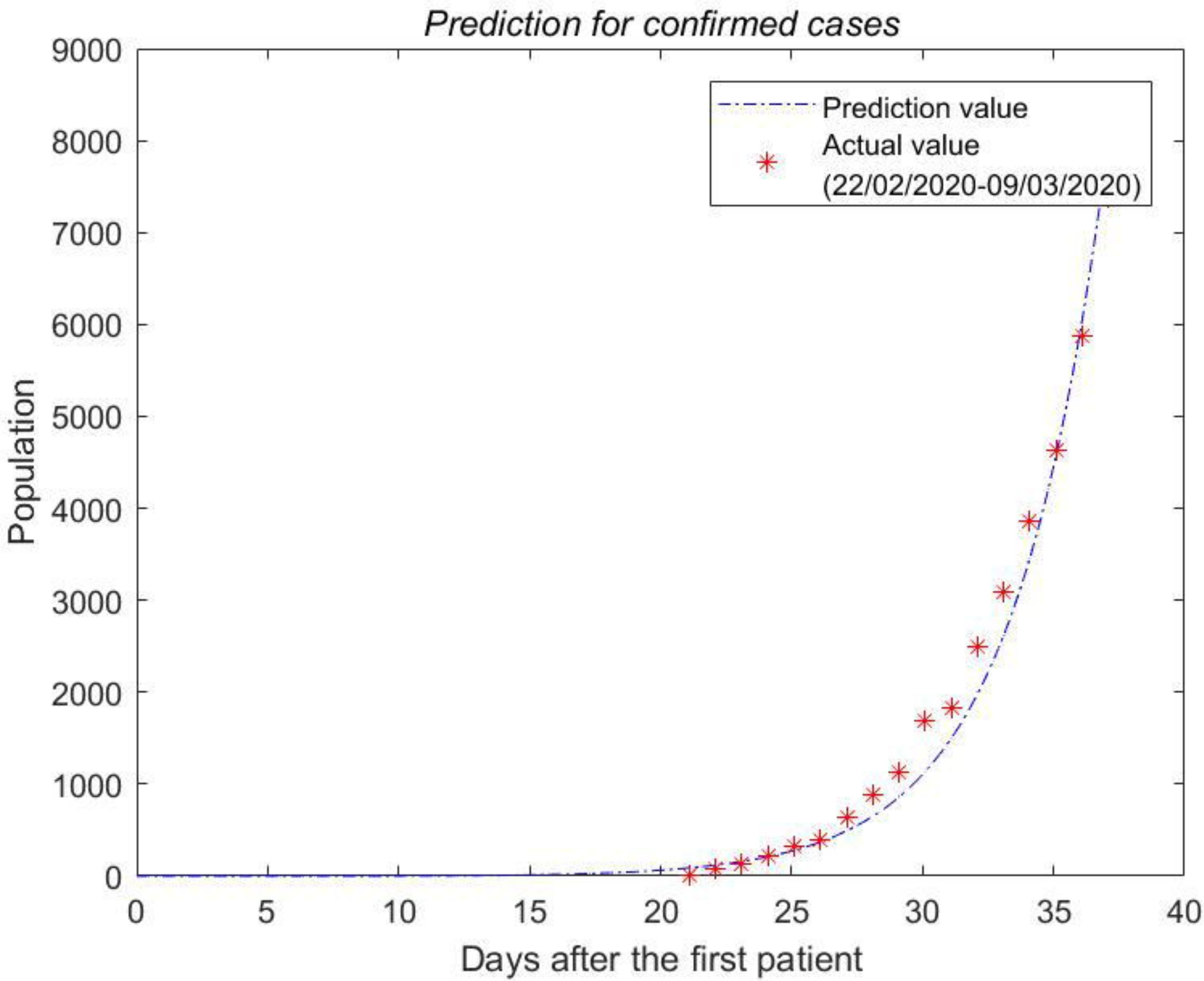
21. J. Zheng. Sars-cov-2: an emerging coronavirus that causes a global threat. *Int J Biol Sci*, 16(10):1678–1685, 2020.

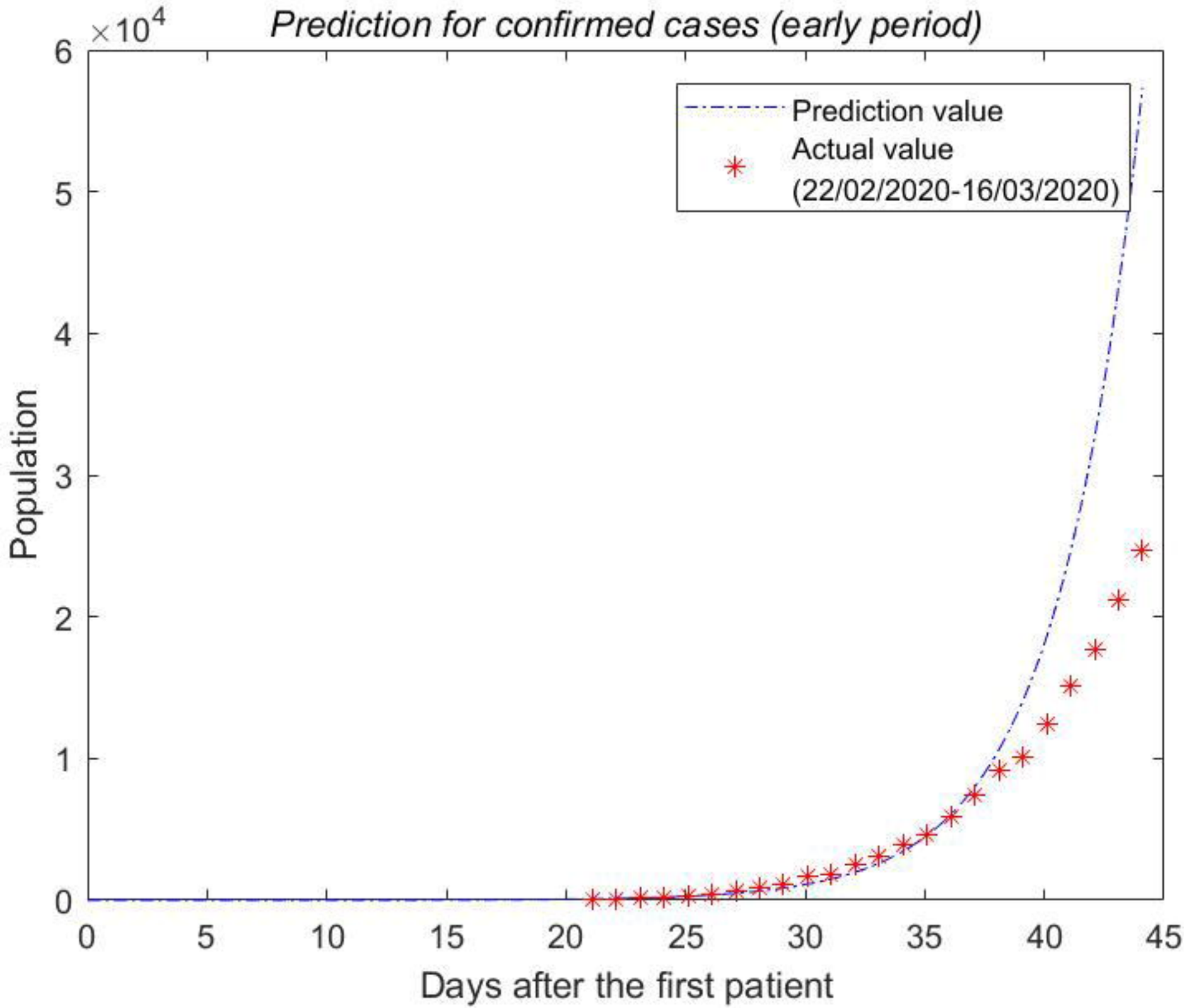

Input actual data
Construct SEIIR model
Inverse analysis parameters T_0, α_0
PSO
Inverse analysis parameters α_1 based on fixed T_0, α_0
Predict feasible α interval to control the disease before May

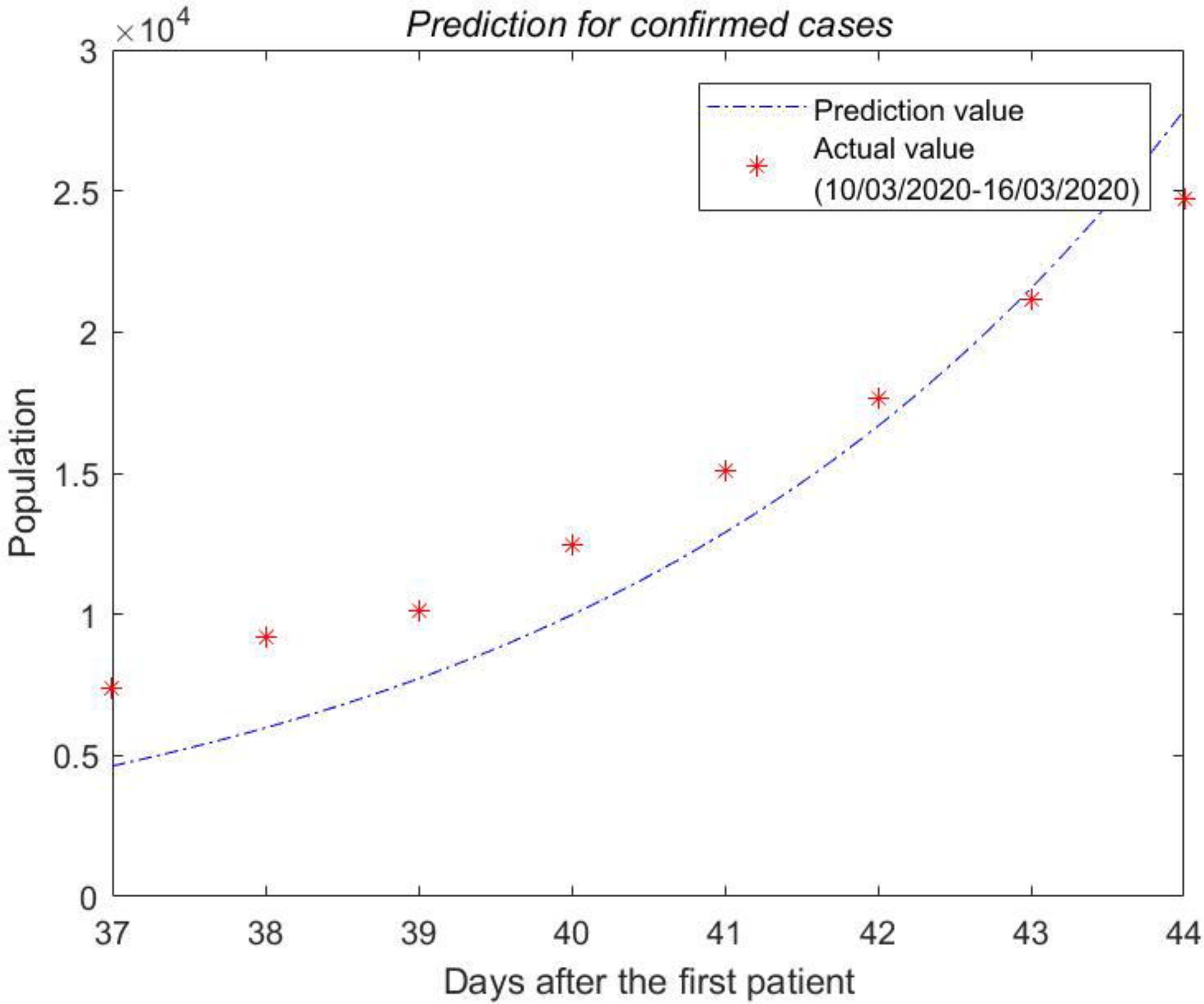

α
R
t_3, μ
S
α
E
t_1
I
t_2, β
J
t_2, γ_2
t_3, γ_1
D

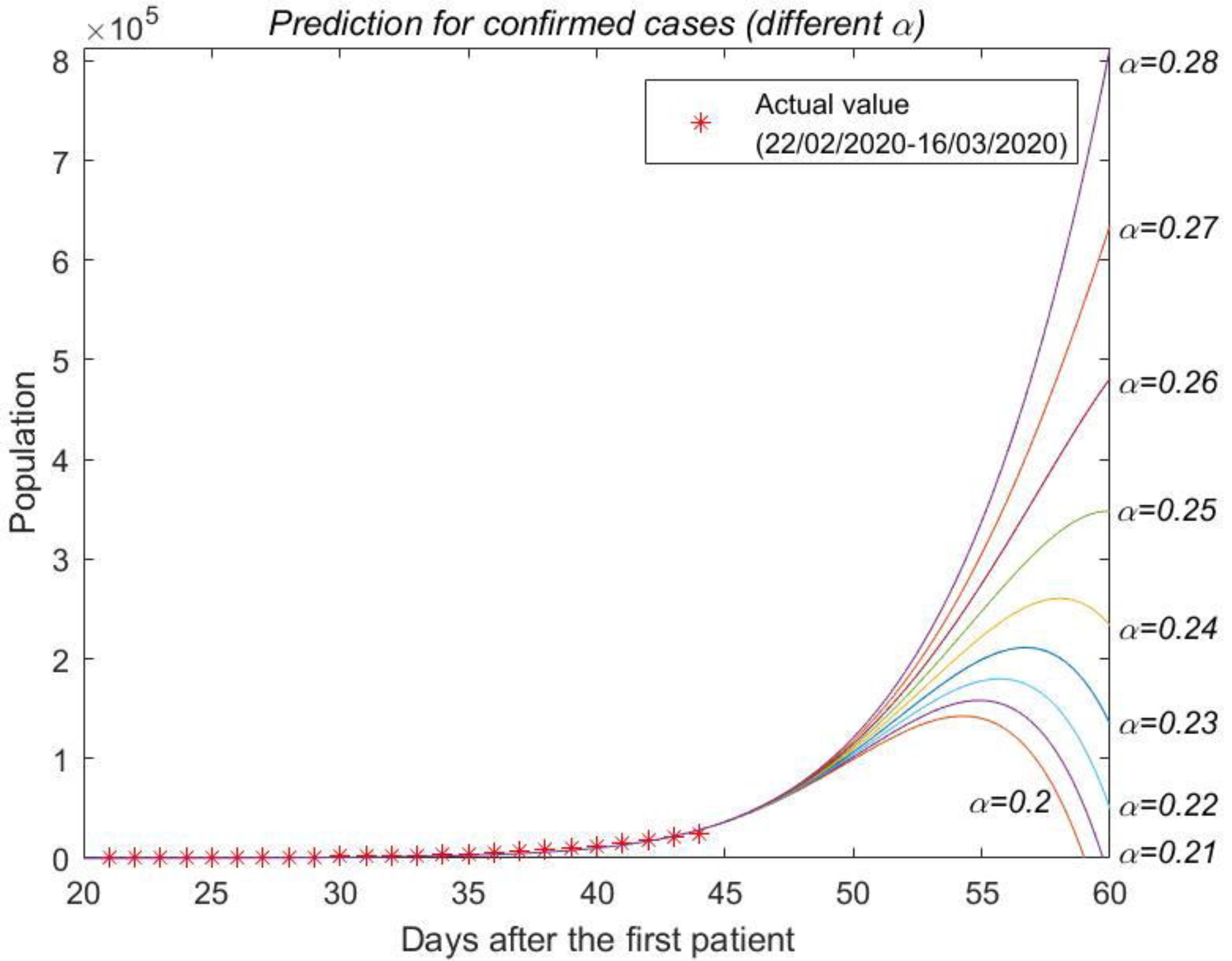

Start
Data preprocessing
Construct fitness function
Initialize PSO parameters
Update parameters:
velocity $\vec{v}_i(t)$
position $\vec{x}_i(t)$
$fitness_{\vec{x}_i(t)} < fitness_{pbest}$?
Yes
Update $pbest$ and $\vec{x}_{pbest}$
$fitness_{\vec{x}_i(t)} < fitness_{gbest}$?
No
Yes
Update $gbest$ and $\vec{x}_{gbest}$
Iteration times + 1
Equals to maximum iteration times
No
Yes
End


Inversion Parameters Analysis
×10^8
2.5
2
1.5
1
0.5
0
0.2
0.25
0.3
0.35
18
19
20
21
22
days
(21.1,0.33)


Prediction for confirmed cases
Prediction value
Actual value
(22/02/2020-09/03/2020)
Population
Days after the first patient
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
0
5
10
15
20
25
30
35
40


Prediction for confirmed cases (early period)
×10^4
6
5
4
3
2
1
0
Population
0
5
10
15
20
25
30
35
40
45
Days after the first patient
Prediction value
Actual value
(22/02/2020-16/03/2020)


Prediction for confirmed cases
×10^4
3
2.5
2
1.5
1
0.5
0
Population
37
38
39
40
41
42
43
44
Days after the first patient
Prediction value
Actual value
(10/03/2020-16/03/2020)


Prediction for confirmed cases (different α)
$\times 10^5$
Actual value
(22/02/2020-16/03/2020)
α=0.28
α=0.27
α=0.26
α=0.25
α=0.24
α=0.23
α=0.22
α=0.21
α=0.2
Population
Days after the first patient
0
1
2
3
4
5
6
7
8
20
25
30
35
40
45
50
55
60