2021 International Conference on Information Technology and Biomedical Engineering (ICITBE) | 978-1-6654-0099-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICITBE54178.2021.00063

Analysis and prediction of COVID-19 based on the SIR-B model

Yijun Shen
Emory College of Arts and Sciences,
Emory University,
Atlanta, USA;
e-mail: yshe255@emory.edu

Xin Guo
School of Computer and Information Technology,
Shanxi University,
Taiyuan, China;
e-mail: mymql123@163.com

***Abstract*—With the spread of the global COVID-19 epidemic, an increasing number of asymptomatic infectors are being detected, and are having an increasingly significant impact on the epidemic spread. To address this problem, a modified SIR-B model based on the time-varying is proposed, which takes into account the presence of asymptomatic infectors on the basis of the traditional SIR model, and predicts the impact of asymptomatic infectors on the subsequent development of epidemic by a Particle Swarm Optimization (PSO) algorithm which changes the adaptation function. Simulation experiments show that the SIR-B model has about one-third more infectors than the SIR model, which is closer to the actual situation, and that the SIR-B model is more adaptive and more accurate in predicting the epidemic than the traditional SIR model.**

Keywords：data analysis; COVID-19; SIR model; SIR-B model; particle swarm optimization algorithm

I. INTRODUCTION

Corona Virus Disease 2019, also known as "COVID-19", refers to the pneumonia caused by the 2019 coronavirus infection[1]. The World Health Organization (WHO) names this virus as "SARS-CoV-2". After 2019, cases emerged in various countries. Up to April 2021, there have been 216 countries and regions affected by COVID-19, the number of confirmed cases worldwide has reached 130 million, and the cumulative number of deaths has exceeded 2.9 million, which has caused panic around the world. After the Hubei epidemic was declared to be ended on April 8, 2020, the epidemic broke out in Xinjiang on July 16 of the same year; and on April 7, 2021, another round of epidemic broke out in Ruili, Yunnan. The National Health Commission strengthened the screening of asymptomatic infectors when releasing the latest epidemic data of two places, which was the first statistical record of asymptomatic infectors in China[2]. Then the Chinese government immediately deployed strategies such as cities lockdown, quarantining, and limiting public entertainment venues, to prevent the virus spread. Different from the 2003 Severe Acute Respiratory Syndrome (SARS) infectious disease, COVID-19 has a special feature: asymptomatic infector, that is, the patient has no relevant clinical symptoms, but the pathogenic test of COVID-19 specimens, such as the respiratory tract, is positive. This feature makes that the asymptomatic infected people do not realize that they have been infected, at the same time, they have the ability to infect others, which will cause more people to be infected.

With the continuous development of COVID-19, there have been many studies and predictions. Some modified SIR models are used to predict the peak and end time of the epidemic spreading in different situations, and there are also predictions of cumulative cases in different provinces[3-6]. However, when facing sudden outbreaks, governments are normally unable to accurately screen and count the infected people at the beginning. Therefore, the initial data of epidemic used in the experiment is often inaccurate. Part of the process can only be analyzed by using hypothesis, which leads to a big difference between the predicted result and the actual situation. There are also SEIR models used under different population densities to predict the transmission trend and the peak arrival time, and further analyze the inflection point and disappearance time of epidemic[7-11]. However, because the National Health Commission failed to consider the incubating patients at the beginning of epidemic, there is a lack of data for such patients. Researchers can only reversely calculate the incubation period of the patients and add many unknown variables in the model, which increases the prediction uncertainty. In response to such shortcomings, some researchers have combined the Artificial Intelligence(AI) technology with the SEIR model to estimate the change of infection rate, analyze and predict the tendencies of transmission and development, and further predict the peak of cumulative cases to help the government to take corresponding measures[12-16]. After applying the AI technology, it is sure that the error rate reduces effectively and the fitting accuracy increases. However, all the above studies have a common problem, that is, they have not considered the existence of asymptomatic infectors in the epidemic and their impact on the epidemic.

With the rapid spread of COVID-19, more and more asymptomatic infectors have appeared. Although these infected people have no clinical symptoms, they are contagious and play a major role in the epidemic spread. Due to the lack of understanding of COVID-19 in the early stage of epidemic and the lack of statistical data on the asymptomatic infectors, all of the above-mentioned studies do not carry out the analysis of asymptomatic infectors. However, in the outbreaks in Xinjiang and Yunnan, the National Health Commission calculated the number of asymptomatic patients from the beginning and conducted

978-1-6654-0099-2/21/$31.00 ©2021 IEEE
DOI 10.1109/ICITBE54178.2021.00063

medical isolation on the asymptomatic infectors. Therefore, this article uses the asymptomatic patient data released in Xinjiang and Yunnan, and conducts experiments to analyze and study its impact on the epidemic. The study uses a time-varying correction-based SIR dynamic model combined with the intelligent optimization algorithm to predict the cumulative cases and the end time of epidemic, and analyzes the impact of asymptomatic infectors on the epidemic.

II. MODEL AND METHOD

A. Traditional SIR Model and Problem Description

In the dynamics model of infectious disease[17], the SIR model divides the total population into three categories in a closed system that does not consider the population mobility factor, where S stands for the susceptible people, I stands for the infected people, and R stands for the cured people. And the specific explanations can be defined as follows:

Definition 1. The susceptible people, whose number is recorded as $S(t)$, which means the number of people who are not infected at time t, but may be infected by COVID-19 in the future; the infective people, whose number is recorded as $I(t)$, which means that they have been infected at time t and have infectivity; the remove people, whose number is recorded as $R(t)$, which represents the number of people who are removed from the infected population at time t, including the number of people who die after infection and the number of people who recover. Suppose the total population is N, then there is $N = S(t) + I(t) + R(t)$. The relationship diagram between the three is as follows:

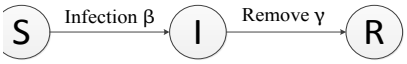

Figure 1. Basic SIR model diagram

According to the SIR relationship model diagram, the model establishment needs to meet the following three conditions:

- Regardless of population dynamics such as birth, death, and mobility, the population in a closed system is always a constant $N(t) = K$.
- A patient who carries the virus must have a certain degree of infectivity when it makes contact with a susceptible person, and the infection rate is β.
- At time t, the rate of transfer from the collection of infected persons to the collection of removers per unit time is γ, which is called the removal rate. The coefficient is proportional to the number of patients.

Based on the above model relationship diagram and conditions, the basic SIR model is as following formula:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta S(t) I(t) \\ \frac{dI(t)}{dt} &= \beta S(t) I(t) - \gamma I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t) \end{aligned} \tag{1}$$

The traditional SIR model has the following shortcomings: the classification of infected population is not detailed, and the situation of asymptomatic infectors is not considered. In this model, there are two key variables: the infection rate β and the removal rate γ, which are considered as constant. The infection rate β refers to the rate at which a person contacts and infects others in a unit of time. The removal rate γ refers to the rate at which a patient moves from the infected person collection to the remover collection. Since the SIR model doesn't use a feedback mechanism, the time-varying characteristics of parameters will be ignored, resulting in low prediction accuracy.

Aiming at the problems existing in SIR model, a SIR-B model based on time series change is proposed, in which the infection rate β and the removal rate γ are both functions of time t. At the same time, the factor of asymptomatic patients is also added to the model. Then the intelligent optimization algorithm is used to fit the patient data to obtain the infection rate $\beta(t)$ and the removal rate $\gamma(t)$. At last, the improved model can be applied to predict the epidemic development.

B. SIR-B Model and Analysis

Aiming to the shortcomings of the SIR model and combining the actual situation, there are the following improvements to modify the traditional SIR model:

- Divide the infected person into two parts, one part is the symptomatic infectors $I_1(t)$, and the other part is the asymptomatic infectors $I_2(t)$, so the total number of cases is recorded as $I(t) = I_1(t) + I_2(t)$.
- Suppose the rate at which a susceptible infector is infected and becomes a patient is $\beta_0(t)$, and the probability is k for patients with symptoms, and for asymptomatic patients, the probability is $1-k$.

After the above improvements, the revised SIR-B relationship model is shown in Fig. 2.

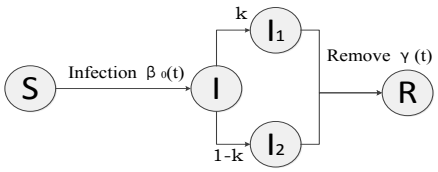

Figure 2. SIR-B relationship model diagram

Based on the above improvements, the infection mechanism of the SIR-B model can be obtained as follows:

$$\begin{aligned} S(i) + I_1(j) + I_2(j) &\xrightarrow{\beta_0(t)} I(i) + I_1(j) + I_2(j) \\ I(i) &\xrightarrow{\gamma(t)} R(i) \end{aligned} \tag{2}$$

It can be seen from the above conditions that when the susceptible individuals are in sufficient contact with the infected individuals, the process of the susceptible people aggregation moving in and out of the infected people aggregation can be expressed by the differential equations as follows:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\beta_0(t)S(t)(I_1(t)+I_2(t)) \\
\frac{dI_1(t)}{dt} &= k\beta_0(t)S(t)I(t)-\gamma(t)I_1(t) \\
\frac{dI_2(t)}{dt} &= (1-k)\beta_0(t)S(t)I(t)-\gamma(t)I_2(t) \\
\frac{dR(t)}{dt} &= \gamma(t)(I_1(t)+I_2(t))
\end{aligned} \tag{3}$$

Currently, the official data are all using day as the unit, so it is to consider using the discrete equations instead of the situation of continuous changes. Discretize the above differential equations to obtain the difference equations as follows:

$$\begin{aligned}
S(t+1) &= S(t)-\beta_0(t)S(t)(I_1(t)+I_2(t)) \\
I_1(t+1) &= I_1(t)+k\beta_0(t)S(t)I(t)-\gamma(t)I_1(t) \\
I_2(t+1) &= I_2(t)+(1-k)\beta_0(t)S(t)I(t)-\gamma(t)I_2(t) \\
R(t+1) &= R(t)+\gamma(t)(I_1(t)+I_2(t))
\end{aligned} \tag{4}$$

The basic regeneration number σ represents the average number of new infectors directly generated by an infected person in one unit of time. $\sigma<1$ means that the epidemic has gone to the retreat phase; $\sigma>1$ means that the epidemic is in the growth phase.

$$\sigma(t)=\frac{\beta(t)}{\gamma(t)} \tag{5}$$

According to the above analysis, to obtain the SIR-B model, it is necessary to calculate $\beta(t)$, $\beta_0(t)$, $\gamma(t)$ in the model, which are affected by the epidemic prevention and control work, and change with time. In order to obtain the three time-varying parameters, the PSO algorithm is used to find their optimal values. At last the SIR-B model is determined.

III. PARTICLE SWARM OPTIMIZATION ALGORITHM

The basic PSO algorithm was proposed by American engineers Eberhart and Kennedy[18]. It is a group-based intelligent optimization searching method. In high-dimensional space functions, the PSO algorithm has the advantages of fast convergence speed, high quality, and good robustness in finding the optimal solution. At the same time, the PSO has the characteristics of simple procedures, few parameters, and easy implementation. Therefore, the PSO algorithm will be used for the fitting calculation to obtain the three important parameters $\beta(t)$, $\beta_0(t)$, $\gamma(t)$ in the SIR-B model.

Algorithm idea: First initialize a group of random particles, expressed as the random solution, and then find the optimal solution through iterative calculation. In each iteration, the updating information of particle's own velocity and position is achieved by tracking two "extrema", and at the same time, the exchange of information between particles can also be obtained by tracking the extrema. The first solution is the individual extreme value *pBest* (self-experience), which is the optimal solution found by the particle itself. The second solution is the group extreme value *gBest* (social-experience), which is the optimal solution found by the entire group at present. The mathematical description of the standard PSO algorithm is as follows [19]:

Definition 2. Assuming that the search space is M-dimensional and the particle number is N, the population is denoted as $X=\{X_1,X_2,\ldots,X_n\}$. Each particle has a position, denoted as $X_i=\{X_{i1},X_{i2},\ldots,X_{in}\}$. Each particle has a velocity, denoted as $V_i=\{V_{i1},V_{i2},\ldots,V_{in}\}$. The best position of $i-th$ particle among $P_i=\{P_{i1},P_{i2},\ldots,P_{in}\}$ is *pBest*, and among all $P_i(i=1,2,\ldots,n)$ the best individual is *gBest*. Each particle updates its own speed according to (6):

$$V_{id}^{k+1}=\omega V_{id}^{k}+c_1r_1(P_{id}^{k}-X_{id}^{k})+c_2r_2(G_{id}^{k}-X_{id}^{k}) \tag{6}$$

The position update formula of each particle is as (7):

$$X_{id}^{k+1}=X_{id}^{k}+V_{id} \tag{7}$$

Where V_{id}^{k+} represents the updated position, P_{id} represents the individual extreme value *pBest*, G_{id}^{k} represents the group extreme value *gBest*, and ω represents the inertia factor. r_1 and r_2 take the random value between 0 and 1. c_1 is the learning factor and c_2 is the social factor. When $c_1>c_2$, the particle is easy to fall into the local search; when $c_1<c_2$, it is helpful to search for the best particle in the integrated population. Therefore, c_1 should be taken as 1.8, and c_2 is 2. The position change range of the $d-th(1<d<M)$ dimension is limited to $[X_{\min,d},X_{\max,d}]$, and the speed change range is limited in $[V_{\min,d},V_{\max,d}]$ (that is, if X_{id} , V_{id} exceed the boundary value in the iteration, the velocity or position of this dimension is limited to the maximum velocity or boundary position of this dimension).

The experiment uses (4) as the fitness function of PSO algorithm to calculate the particle fitness. The PSO algorithm divides the entire particle group into several small particle groups that partially overlap each other. In each iteration, each particle first calculates its own fitness, and then selects the particle with the best fitness from each small particle group. Finally, the best particle in the entire particle group is selected from all the small particle groups, and this particle is *gBest*. Formula (8) can be obtained from (3), and the result contains two columns of data, which are the values of β and γ respectively.

$$gBest=\{\beta \quad \gamma\} \tag{8}$$

Theorem 1. Based on the SIR-B model, the missed detection rate ε of infected patients decreases.

Proof: Let the initial set of infected persons be I_0. The number of newly-added symptomatic infectors per day is ΔI_j, and the number of newly-added asymptomatic infectors per day is ΔI_k, and $\Delta(I_j+I_k)$ is the total number of newly infected patients per day. According to (1) in the SIR model, let the number of symptomatic patients detected in the total population be I, the formal expression is as follows:

$$I_1=I_0+\sum_{d=0}^{D}\Delta I_j \tag{9}$$

In this formula, t represents the epidemic duration, and D represents the end date of epidemic.

According to (3) in the SIR-B model, it can be seen that the total number of patients with symptomatic and asymptomatic infectors detected in the total population is I_2, the formal expression is as follows:

$$I_2 = I_0 + \sum_{d=0}^{D} \Delta\left(I_j + I_k\right) \tag{10}$$

Assuming that in the total population, the total number of infected patients is I. In the SIR model, the missed detection rate θ_1 is:

$$\theta_1 = \frac{I - I_1}{I} \tag{11}$$

In the SIR-B model, the missed detection rate θ_2 is:

$$\theta_2 = \frac{I - I_2}{I} \tag{12}$$

Obviously $\theta_1 - \theta_2 > 0$, so the missed detection rate of infected patients based on the SIR-B model is lower.

Theorem 2. The value of the basic regeneration number R_0 in the SIR-B model is higher than R_0 of the SIR model.

Proof: According to (5), the value of R_0 depends on the ratio of the infection rate β to the recovery rate γ. And

$$\beta \propto \mathrm{I} \tag{13}$$

In the SIR model:

$$I = I_1 \tag{14}$$

In the SIR-B model:

$$I = I_1 + I_2 \tag{15}$$

In above formulas, I_1 represents the symptomatic patients, while I_2 represents the asymptomatic patients. The value of γ depends on the medical level and has nothing to do with the model used. So it is a fixed value.

The pseudo code of PSO algorithm is expressed as follows:

Algorithm 1 algorithm1

Input: Daily cases
Output: fitted value
Function PSO($fit, pBest, gBest, N$)
  for $i = 0 \rightarrow N$ **do**
    Initialize velocity Vi and position Xi for particle i
    Evaluate particle i and set $pBest = Xi$
  end for
  $gBest = min[pBest]$
  while not stop **do**
    for $i = 1 \rightarrow N$ **do**
      Update the velocity and position of particle i
      Evaluate particle i
      if $fit(Xi) < fit(pBest)$ **then**
        $pBest = Xi$
      end if
      if $fit(pBest) < fit(gBest)$ **then**
        $gBest = pBest$
      end if
    end for
  end while
  return $gBest$
End function

Therefore, when γ is fixed and β increases, δ of the SIR-B model is higher than δ of the SIR model. It can be seen that the addition of asymptomatic infectors increases the number of new patients who are infected by infected people within a unit time, leading to the epidemic expansion. So predicting more accurately the number of infected people, strengthen screening, and finding and isolating the asymptomatic infectors in an effective manner can productively reduce the value of δ and control the epidemic spread.

IV. Experiments and Results

The experiment uses the Python platform to simulate and predict the two sets of statistical data respectively. The simulation experiment is divided into two parts: Section A to Section D is data collection and fitting calculation of parameters in the model; Section E is the epidemic prediction based on the value of existing parameters.

A. Data Sources

The experiment uses the statistical data of Xinjiang and Yunnan, issued by the National Health Commission and including the cumulative number of cases and the number of new asymptomatic patients per day, to analyze the epidemic situation. Since the strict prevention and control measures have been implemented in the outbreak areas, these areas have entered a closed status monitored by the government strongly since the beginning of epidemic. Therefore, the population movement factor will not be considered in the following analysis.

B. PSO Actual Fitting Data

The experiment divides the epidemic statistics into two groups. The first group contains only the number of symptomatic patients, and the second group contains the total number of infectors including both symptomatic and asymptomatic patients.

First, the PSO algorithm combining with the SIR-B model is used to fit the first set of data, as shown in Fig. 3 and Fig. 4. In the figures, the abscissa represents the date of statistical data, and the ordinate represents the number of confirmed cases. The square in the figures represents the actual data in the epidemic, the red dashed line represents the result of fitting the confirmed infection cases data, and the blue dashed line represents the trend change of S set in the model. It can be seen from the figure that the red fitting curve is closer to the change trend of actual data. Fig. 3 is the fitting of the first set of data in Xinjiang epidemic. It can be seen that the number of new cases each day begin to decrease from the 15th day, which indicates that under the strict prevention and control measures of the country, the epidemic has entered a flat situation and is changing from the initial exponential stage into the sub-exponential growth phase. Fig. 4 is the fitting of the first set of data in Yunnan. It can also be seen that after the sixth day, the epidemic has entered a mitigation period.

Then the second set of data in the two places is fitted respectively to obtain Fig. 5 and Fig. 6. Considering that the asymptomatic infectors are inherently contagious and due to

individual differences, the individuals infected by the asymptomatic infectors may not be all asymptomatic, so the experiment will analyze the asymptomatic infectors and the symptomatic patients at the same time. Comparing Fig. 3 and Fig. 5, Fig. 4 and Fig. 6, it can be seen that the fitting curve becomes steeper and the peak value also increases after adding the asymptomatic infectors. It shows that if the asymptomatic infectors are not considered in the early stage of epidemic, the patients infected by them will be ignored, which will delay the arrival of the peak of confirmed cases in the epidemic and prolong the duration of epidemic.

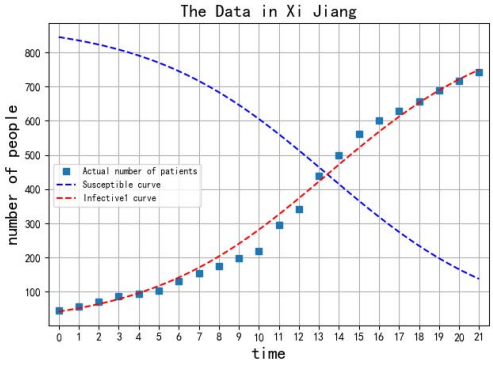

Figure 3. Fitting of the first set of data in Xinjiang

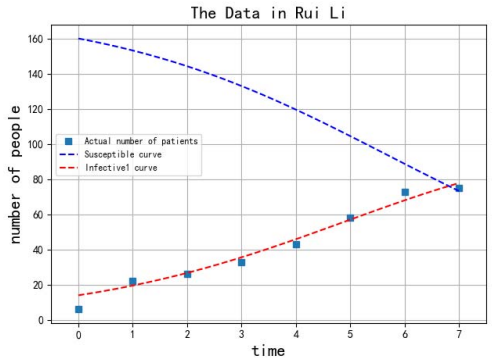

Figure 4. Fitting of the first set of data in Yunnan

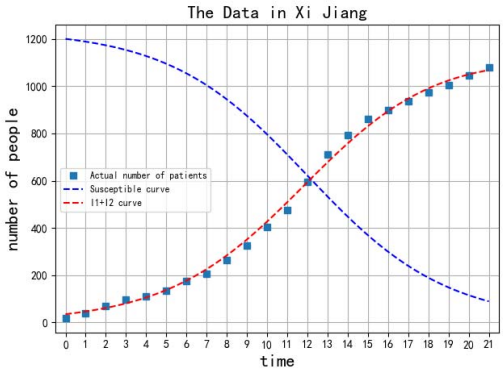

Figure 5. Fitting of the second set of data in Xinjiang

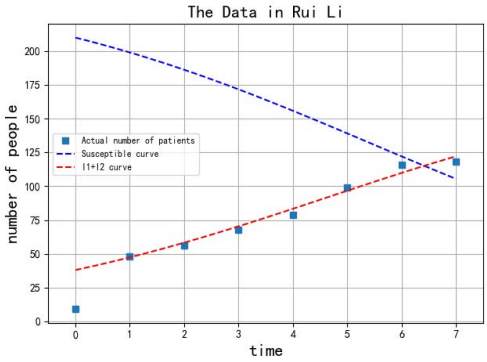

Figure 6. Fitting of the second set of data in Yunnan

C. *Time-Varying Parameters in SIR-B Model*

In the fitting of the two sets of statistical data in Fig. 3 to Fig. 6, the PSO algorithm performs multiple optimization calculations on their infection rate. The first group is the infection rate $\beta(t)$ only considering the symptomatic infectors. And the second group is the infection rate $\beta_0(t)$ both considering the asymptomatic and asymptomatic infectors. First, get the values that $\beta(t)$ changes with time according to the first set of data, as shown in the blue square in Fig. 7. And use the polynomial fitting method to fit the values to get the curve that $\beta(t)$ changes with time in Fig. 7, where the abscissa represents the calculation numbers along the time increment, and the ordinate represents the calculated value. Fig. 8 shows the change curve of the infection rate $\beta_0(t)$ in the second set of data. By observing Fig. 7 and Fig. 8, it can be seen that the fitting curves overall show a rapid downward trend, indicating that the prevention measures taken by the government have effectively controlled the transmission rate of infected persons. The initial value of curve in Fig. 8 is higher than that in Fig. 7, which indicates that if the asymptomatic infectors were not considered at the beginning of epidemic, the infection rate would be underestimated. The later value in Fig. 8 is lower than that in Fig. 7, which indicates that the isolation of asymptomatic infectors will help to reduce the infection rate of epidemic. If the asymptomatic infectors are not considered, the infectivity of these people will be ignored, which will cause more susceptible people to be infected and expand the spread of epidemic. Therefore, the screening of asymptomatic infectors is very necessary. Once discovered, they should be isolated and treated immediately so as to effectively control the further expansion of epidemic.

In addition, the removal rate γ(refers to the rate at which patients are transferred from the infected population to the removed population) in the system can be obtained according to the fitting data. Since both asymptomatic and symptomatic infected persons will be isolated and treated, the removal rate is the removal rate of asymptomatic and symptomatic infected patients together, as shown in Fig. 9. It can be seen that the removal rate curve rises rapidly in the

early stage and is nearly stable in the later stage. It shows that after being tested and confirmed, the infected patients can receive isolation treatment immediately. Due to the individual differences, each person's recovery time is different. And because the infected patients have just started treatment, the early removal rate is low. However, with the lengthening of treatment time and the improvement of medical level, the removal rate of infected patients has increased overall. As the epidemic stabilizes in the later period, the removal rate reaches a relatively stable value. Comparing Fig. 9 with Fig. 7 and Fig. 8, it can be seen that when a relatively stable state is reached in the later stage, the removal rate is greater than the infection rate, which indicated that the epidemic has tended to subside.

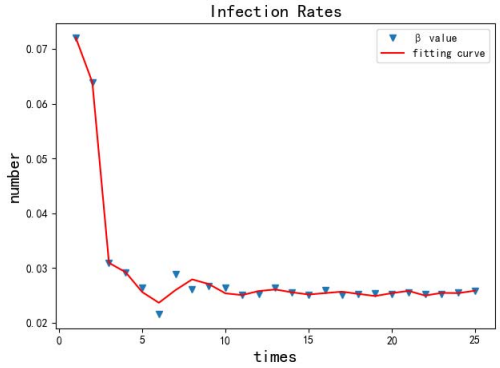

Figure 7. Fitting of infection rate of symptomatic infector

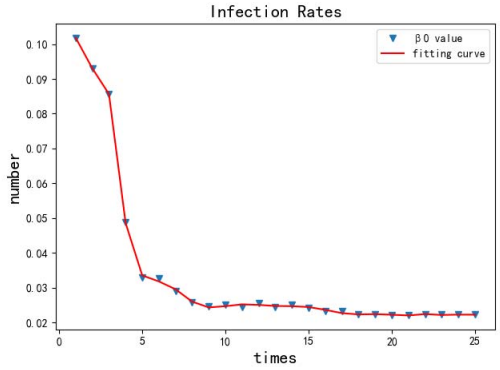

Figure 8. Fitting of infection rate of symptomatic and asymptomatic infector

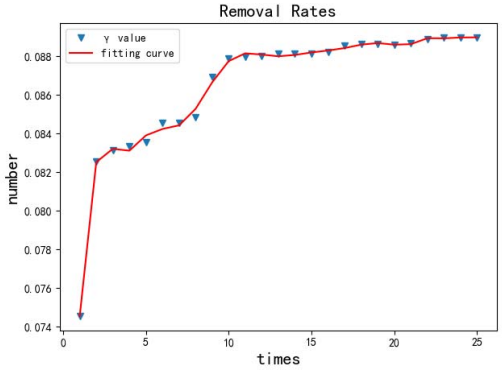

Figure 9. Removal rate of infected patients

D. *Basic Reproduction Number* δ

The basic reproduction number δ reflects the degree of the epidemic spread. $\delta<1$ indicates that the epidemic is expanding; $\delta<1$ indicates that the epidemic trend is weakening; $\delta=1$ indicates that the infection rate and the recovery rate of epidemic are in a balanced state, that is, neither increasing nor decreasing. Fig. 7 and Fig. 8 calculate the values of $\beta(t)$ and $\beta_0(t)$, and Fig. 9 calculates the value of $\gamma(t)$. The basic regeneration number δ can be calculated by (6). The results are shown in Fig. 10 and Fig. 11.

By observing Fig. 10 and Fig. 11, it can be seen that the basic regeneration rate drops rapidly in the early stage and reaches a relatively stable state in the later stage. It shows that the epidemic situation is well controlled in the early stage, and the epidemic prevention measures are strengthened in the later stage so that the epidemic situation has been in a state of gradual subsidence. Fig. 11 is the basic reproductive number comprehensively considering the symptomatic infectors and the asymptomatic infectors. Although at the beginning δ is greater than 1, and the epidemic has expanded, δ soon drops to less than 1, indicating that the anti-epidemic personnel strictly screen and isolate the infectors so that the value of δ are effectively controlled. Since Fig. 10 does not consider the role of asymptomatic infectors in the epidemic, the number of patients infected by the asymptomatic inflections is ignored, which results in a low initial value δ. In the later stage of epidemic, after considering the impact of asymptomatic infectors, δ in Fig. 11 is basically stable at about 0.3. While Fig. 10 does not consider the asymptomatic infectors, δ is basically stable at about 0.25. This indicates that during the epidemic spread, the asymptomatic infectors cannot be detected in time because they have no clinical symptoms, so that the susceptible people will have a high chance of being exposed to the asymptomatic infectors, which makes the basic regeneration number larger.

E. *Epidemic Prediction*

Section B and Section C have determined all unknown parameters. This section combines the SIR-B model with the PSO algorithm to simulate the development trend of epidemic in the future. The simulation experiment will separately predict the two sets of data in Xinjiang and Yunnan, and use the experimental results to verify the impact of asymptomatic infectors on the epidemic. The prediction results obtained through experiments are shown in Fig. 12 to Fig. 15. The abscissa is the date, the ordinate on the left represents the cumulative number of confirmed infectors, and the ordinate on the right represents the number of new diagnoses per day. The square in the figures represents the actual data, the dashed line represents the predicted data curve of infected patients, and the solid line represents the daily increase number of infected patients. It can be seen from Fig. 12 that when only the symptomatic infectors are considered, the cumulative number of confirmed infectors in Xinjiang will reach its peak on the 37th day, with a peak value of about 860 patients, and the

number of new infectors each day will reach its peak on the 15th day. Fig. 13 shows the prediction of the two types of patients.

Comparing Fig. 12 and Fig. 13, the presence of asymptomatic patients will increase the number of confirmed patients. The prediction curve of the second group is steeper than that of the first group, and the epidemic is developing faster. Fig. 14 and Fig. 15 show the two sets of forecasts for the Yunnan epidemic. Because the asymptomatic patients are taken into account in Fig. 15, the duration of epidemic is longer than that in Fig. 14, and the number of confirmed cases is larger.

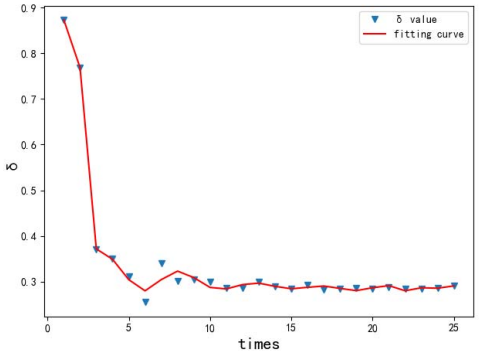

Figure 10. δ only considering symptomatic infector

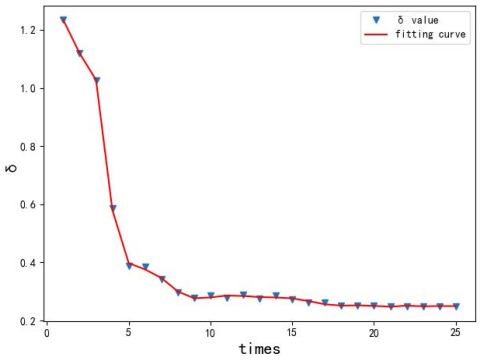

Figure 11. δ considering both asymptomatic and symptomatic infector

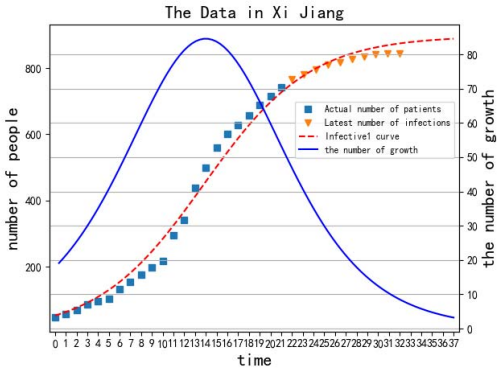

Figure 12. The first group of Xinjiang forecast

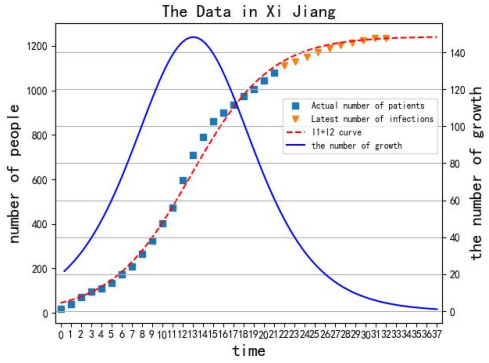

Figure 13. The second group of Xinjiang forecast

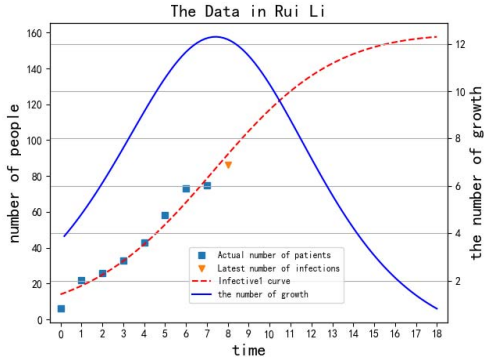

Figure 14. The first group of Yunnan forecast

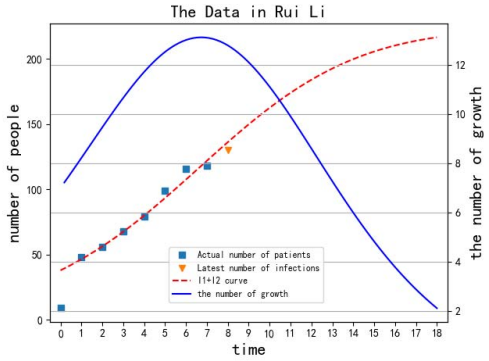

Figure 15. The second group of Yunnan forecast

In order to compare the SIR model with the SIR-B model, their prediction results are analyzed to obtain the following Table 1 and the intuitive diagram Fig. 16. Table 1 lists the prediction result data gotten from the simulation experiment, and based on this data, the histogram shown in Fig. 16 can be obtained.

TABLE I. DATA COMPARISON TABLE

August 20	SIR model	SIR-B mode
Xinjiang	850	1230
Yunnan	147	250

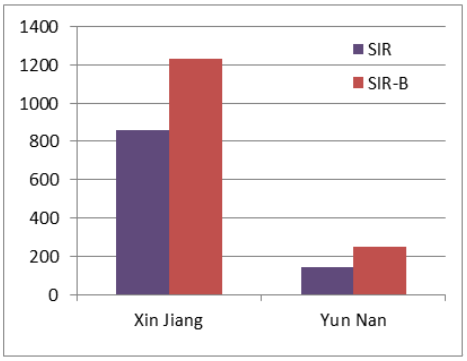

Figure 16. Statistical data result diagram

V. CONCLUSION

The main research question in this article is the impact of asymptomatic infector in the model on the follow-up development of epidemic. Based on the existing official data and the improved SIR-B model of infectious disease dynamics, with the PSO algorithm performing data fitting, the parameters obtained from the fitting are used to predict the subsequent development of epidemic. The result shows that the prediction curve is closer to the actual situation.

The above experiments have verified that the screening and isolation of asymptomatic infectors are very necessary in the early stage of epidemic. A series of epidemic prevention measures adopted by the government can also effectively curb the epidemic spread. Therefore, when a new epidemic occurs in a certain place, the epidemic prevention personnel should carry out screening and testing for the susceptible people in the place and those who enter and exit the epidemic area. Although the asymptomatic infectors have no clinical symptoms, which enhance the screening difficulty, such measures are essential. Under such strict prevention and control, finding the asymptomatic patients early for isolation and treatment can not only reduce their exposure time but also reduce the number of infectors so that the epidemic will end as soon as possible and people's production and life will return to normal sooner.

REFERENCES

[1] "Coronavirus disease (covid-19) outbreak," Jan 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[2] (2020).Xinjiang Municipal Health Commission Infection Data. [Online] Available: http://wjw.xinjiang.gov.cn/hfpc/fkxxfyfkxx/fkxxfy_ list.shtml

[3] Zhong Lin-hao, Mu Lin, Li Jing, et al. Early Prediction of the 2019 Novel Coronavirus Outbreak in the Mainland China Based on Simple Mathematical Model[J].IEEE access, 2020, PP(8):51761-51769.

[4] Yi-Cheng Chen, Ping-En Lu, Cheng-Shang Chang, et al. A Time-dependent SIR model for COVID-19 with Undetectable Infected Persons [J]. IEEE Transactions on Network Science and Engineering, 2020, 4(7): 3279-3294.

[5] Benjamin F, Maier, Dirk Brockmann. Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in Mainland China [J]. Medical Letter on the CDC & FDA, 2020, 368(6492):742-746.

[6] Leung Kathy, Wu Joseph T, Liu Di, et al. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modeling impact assessment [J]. The Lancet, 2020, 0(Pre-publish):1382-1393.

[7] Rocklöv J, Sjödin H, Wilder-Smith A, et al. COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures [J]. Journal of travel medicine, 2020, 27(3):1-7.

[8] Peng L, Yang W, Zhang D. Epidemic analysis of COVID-19 in China by dynamical modeling [J]. Medical Letter on the CDC & FDA, 2020. (2002):6563-6581.

[9] Kentaro Iwata, Chisato Miyakoshi. A Simulation on Potential Secondary Spread of Novel Coronavirus in an Exported Country Using a Stochastic Epidemic SEIR Model [J]. Journal of Clinical Medicine, 2020, 9(4): 944-951.

[10] Biao Tang, Xia Wang, Qian Li, et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions [J]. Journal of Clinical Medicine, 2020, 9(2):462-475.

[11] Huwen Wang, Zezhou Wang, Yinqiao Dong, et al. Estimating the Number of 2019 Novel Coronavirus Cases in Chinese Mainland[J]. SSRN, 2020:1-15.

[12] M Small, D Cavanagh. Modeling Strong Control Measures for Epidemic Propagation With Networks—A COVID-19 Case Study [J].IEEE access, 2020,8(8):109719-109731.

[13] Y Zhao, Y He, X Zhao. COVID-19 Outbreak Prediction Based on SEIQR Model[C] // 2020 39th Chinese Control Conference . Shenyang, China, 2020:1133-1137.

[14] Yang Zi-feng, Zeng Zhi-qi, Wang Ke, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions [J]. Journal of thoracic disease, 2020, 12(3):165-174.

[15] Y Yang, W Yu, D Chen. Prediction of COVID-19 spread via LSTM and the deterministic SEIR model[C] // 2020 39th Chinese Control Conference. Shenyang, China: IEEE, 2020: 822-825.

[16] Zheng Nan-ning, Du Shao-yi, Wang Jian-ji, et al. Predicting COVID-19 in China Using Hybrid AI Model[J]. IEEE transactions on cybernetics, 2020, 50(7):2891-2904.

[17] Romualdo P S, Alessandro V. Epidemic spreading in scale-free networks [J]. Phys Rev Lett, 2001, 86(1): 3200-3203.

[18] Kennedy J, Eberhart R C. Particle Swarm Optimization[C] // IEEE International Conference on Neural Networks. Piscataway, NJ: IEEE Service Center, 1995:1942-1948.

[19] Shi Y, Eberhart R C. A Modified Particle Swarm Optimizer[C] // The 1998 Conference of Evolutionary Computation. Piscataway, NJ: IEEE Press, 1998:69-73.