第51卷第9期 中国科学技术大学学报 Vol. 51, No. 9
2021年9月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Sep. 2021
Received: 2021-06-04; Revised: 2021-08-18 doi:10.52396/JUST-2021-0146

The transmission rate of COVID-19 pandemic under different mobility control in Europe

LI Ying[1], SUN Tianyi[1], JIN Baisuo[2], ZHANG Bo[2]*
1. School of Data Science, University of Science and Technology of China, Hefei 230026, China;
2. School of Management, University of Science and Technology of China, Hefei 230026, China
* Corresponding author. E-mail: wbchpmp@ustc.edu.cn

Abstract: COVID-19 pandemic captured the full attention of the world in 2020, and the government declared a series of non-pharmacological interventions (NPIs) to curb the influence of social movement on transmission. In different countries, different policies bring about different results. Quantifying the effect of the movement becomes a vital issue for evaluating the effectiveness of these actions. The transmission rate changes and is hard to computer after altering activity. Therefore, this research sets some European countries as the research objects, collects mobility data and daily cases during some periods, and proposes a mobility-susceptible-exposure-infectious-recovery (M-SEIR) model. Unlike the SEIR model, the movement change is quantified as a variable ($\sigma(t)$) and added in the M-SEIR model. With random sampling to get the number of people in different initial states, this research iterates the model. The iterative filtering ensemble adjustment Kalman filter (IF-EAKF) is used to adjust the subsequent iterative results. In the research, it receives the changing trend of parameters and the daily new estimation in the end. Set the first round as the fitting period and repeat the experiment 100 times in the fitting part. The result confirms the feasibility and robustness of the model. In addition, this study makes a reasonable forecast for European countries about the second round. By controlling the strength and the time point of applying non-pharmacological interventions, the research predicts the impact of these actions on the pandemic and provides some suggestions for the deployment of relevant policies in the future. Finally the study eliminates the external factors such as motion and temperature, and obtains an interesting discovery: Despite the daily case in the third round higher than that in the first round, the transmission parameter in the former appears lower than that in the latter. The further survey shows that it might be related to vaccination.

Keywords: M-SEIR model; mobility; contact matrix; transmission rate
CLC number: O213.9 **Document code**: A
2020 Mathematics Subject Classification: 62P99

1 Introduction

COVID-19 broke out at the end of 2019. After just a few months, it became a worldwide pandemic. Accordingly, governments adopted relevant public health measures to contain the large-scale spread. For the pandemic, China made strict non-pharmacological interventions (NPIs), such as declaring a "state of emergency", "closing school and workplace", and "prohibiting gathering activities" to limit social movement in workplaces, schools, and other places. These emergency responses positively affect the confirmed cases, which shows an apparent decrease in number. However, when European countries face the same situation, the response is not the same. Although the NPIs in these countries achieved a small spread at the beginning of the pandemic, most countries in Europe had a more severe condition in the latter half of 2020. Thus, quite a few countries reannounces to carry out a slice of NPIs. Nevertheless, NPIs played a limited role in curbing the spread when facing the second round. Many countries had to accept the third round of pandemic again after ending the second round. To some extent, quantifying the effect of NPIs on the pandemic is still a problem to be solved. Also, the daily transmission rate needs special attention.

Considerable studies have been provided on the research of the transmission system of COVID-19 from

Citation: Li Ying, Sun Tianyi, Jin Baisuo, et al. The transmission rate of COVID-19 pandemic under different mobility control in Europe. J. Univ. Sci. Tech. China, 2021, 51(9): 654-670,716.

mobility. Reference [1] believed that the flow data from Wuhan to other cities in China was related to the number of confirmed cases in other cities (although the predictive power has declined over time). It captured population movement data over the past period, added GDP and the local population as control variables, and computed the effectiveness of control measures. Reference [2] provided a new susceptible-exposure-infectious-recovery (SEIR) model, which integrated fine-grained and dynamic mobility networks. The study concluded that some "superspreader" points would account for a large majority of the infections in America. Reference [3] performed a sensitivity test on a slice of pandemic parameters and inferred reproductive number R_t to search the effect of the NPIs during the control period. This research adopted different ways to establish a model for the development of the pandemic. However, it could not explain the quantitative relationship between the rate of transmission and mobility change brought by NPIs. Reference [4] presented an overview of the policy responses in some European countries from Feb. to Jun. 2020, and this study offered a comparative policy analysis.

Based on the classical susceptible-exposure-infectious-recovery (SEIR) model, we collect the mobility data and propose a mobility-susceptible-exposure-infectious-recovery (M-SEIR) model. The model has an excellent performance to fit the impact of mobility on transmission. Section 2 defines the movement trend by computing the contact matrix and the mobility change data ($\sigma(t)$). In the third section, the model applies the iterative filtering (IF) algorithm. The technique used in the filtering process is known as the ensemble adjustment Kalman filter (EAKF). In the empirical analysis (Section 4), the study first carries out the correlation test between $\sigma(t)$ and daily case to search for the delay effect of mobility. The test manifests that the delayed time was three days. After iterative filtering, we achieve a comparison between $\beta(t)$ and $\sigma(t)$. The difference between $\beta(t)$ and $\sigma(t)$ turns out to be highly correlated with temperature and humidity. The adjustment for difference would improve the model. Iterate the model and algorithm can get the daily observation. Finally, we make a reasonable forecast of the pandemic under several different control intensities. The forecast contains information about the mobility control effect and delay-effect in quite a few countries. The predictions under different conditions are very similar to the real situation. Besides, we compare the estimated parameters in three rounds. Interestingly, the reproductive number of European countries in the third round is lower than that in the first round. This phenomenon pushes us to think about whether the decline of the reproductive number is related to personal immunity, especially after vaccination. Section 5 describes the conclusions.

2 Theoretical basis

2.1 Model introduction

We establish a mobility-susceptible-exposure-infectious-recovery (M-SEIR) model to compute the spread of the pandemic. The model considers the specific characteristics of infection and incorporates human movement into the model. The population is divided into five categories, and different categories denote different possible conditions related to diseases:

· Susceptible S: A healthy group of characters who never contact an infected person but lack immunity and are vulnerable to infection after contacting an infectious person.

· Exposure E: A group of characters who contact an infectious person can not temporarily infect other people.

· Recorded infectious I^r: A group of characters that can infect others. They have some transmission symptoms and are confirmed by the hospital.

· Unrecorded infectious I^u: A group of characters who can spread disease. They have transmission symptoms but are not confirmed by the hospital.

· Recovery R: People who are isolated, recovered, or died.

The model consists of the following ordinary differential equations (ODEs), in which the time variable $\beta(t)$ contains information about mobility and contact in a different area:

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = -\frac{\beta(t)SI^r}{N} - \frac{\mu\beta(t)SI^u}{N} \tag{1a}$$

$$\frac{\mathrm{d}E_t}{\mathrm{d}t} = \frac{\beta(t)SI^r}{N} + \frac{\mu\beta(t)SI^u}{N} - \frac{E}{Z} \tag{1b}$$

$$\frac{\mathrm{d}I^r}{\mathrm{d}t} = \alpha\frac{E}{Z} - \frac{I^r}{D} \tag{1c}$$

$$\frac{\mathrm{d}I^u}{\mathrm{d}t} = (1-\alpha)\frac{E}{Z} - \frac{I^u}{D} \tag{1d}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \frac{I^r}{D} \tag{1e}$$

In the above formulae, susceptible individual S tends to decrease after he contacts the confirmed group I^r or the unconfirmed group I^u. Driven by transmission, the trend of infection increases. In a general environment, the transmission rate β equals the transmission probability times the average number of contacts per person. Other parameters contain the speed of the individual from one group to the next group. One can see the pace in the flowchart (Figure 1).

Specifically, the parameter Z means the average duration period from group E to I^r or I^u, and $\frac{1}{Z}$

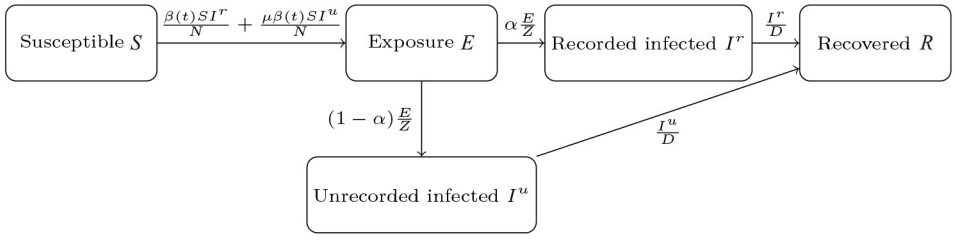


Figure 1. The speed that from the present condition to the next condition.

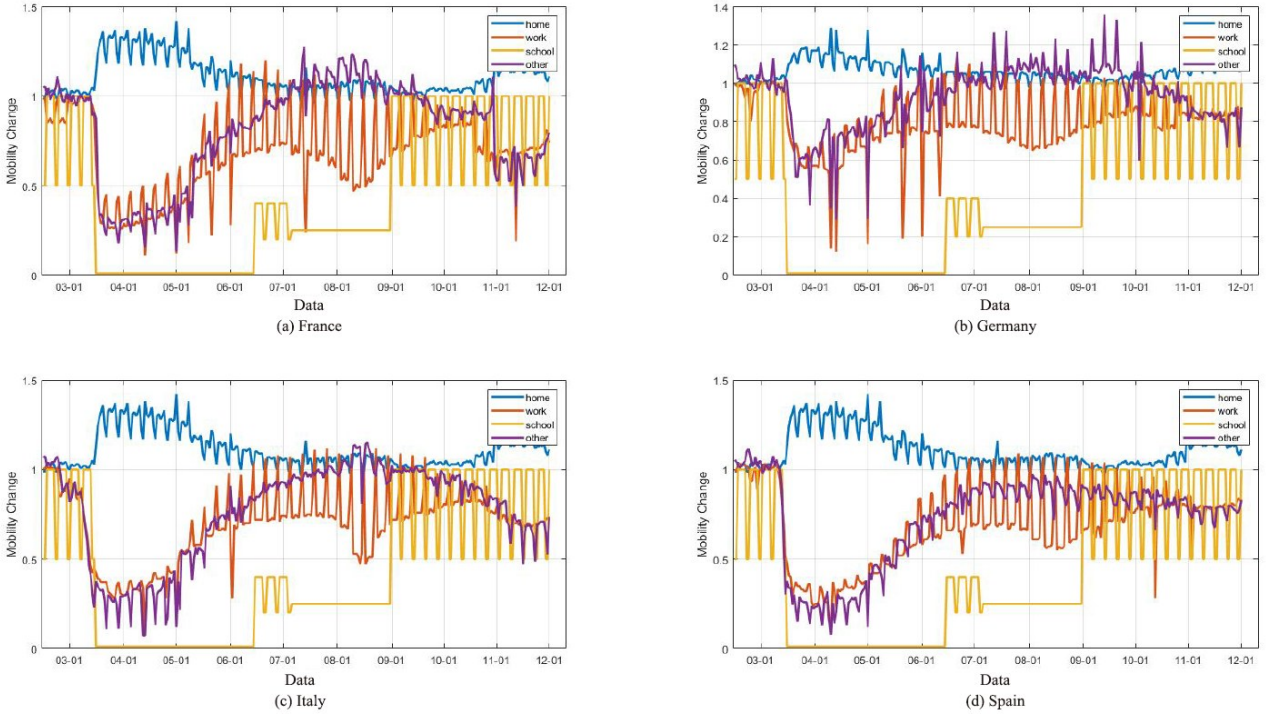


Figure 2. Mobility. The curve shows mobility change in different environments (home, school, work, other) from Feb. 15 to Nov. 30, 2020

indicates the varying speed of the exposed group E to the confirmed infection group I^r or I^r. The parameter D provides the average infection period of I^r or I^u, and $\frac{1}{D}$ suggests the speed of changing from the infectious to recovered or dead. The parameter μ contains the transmission rate reduction factor of the undocumented infection group, which exists in the interval of $[0,1]$. The parameter α denotes the ratio of the confirmed infection group I^u in the whole infectious group.

2.2 The definition of $\beta(t)$

In Section 2.1, the model defines the transmission rate $\beta(t)$ as the product of the transmission probability $\beta_0(t)$ and the average number of contact per infectious $\sigma(t)$. $\beta(t)$ is a multiple and means the infected number after the infectious individual contacts the susceptible group on time t. $\beta_0(t)$ is determined by the activity of the virus so that it would keep the same during a short period. $\sigma(t)$ means the frequency of contact between each infectious and susceptible. When $\beta(t)$ in Equation (2) is multiplied by the number of infectious at this time point, and then times the proportion of the susceptible group in the total population, the M-SEIR model can calculate the daily number from the susceptible group to exposed group on time t.

The average contact number remains unchanged early before governments take non-pharmaceutical intervention against the pandemic, so the transmission rate keeps a constant value. However, mobility $m_k(t)$ (k=home, school, workplace and other) decreases and affects the average number of contact $\sigma(t)$ in turn after taking corresponding interventions. Contact $\dot{C}(k)$ records the frequency of contact between each infectious person and susceptible in environment k. Define

$$\beta(t) = \beta_0(t)\sigma(t) \tag{2}$$

where

$$\sigma(t) = \sum_{k=1}^{4} m_k(t)\dot{C}(k) \tag{3}$$

It is not easy to measure $\beta_0(t)$, so we do not

quantify it directly. Now, let us learn $\sigma(t)$ firstly. Two definitions need to be introduced in the step, contact matrix C and mobility data in several environments.

We introduce the mobility data in the beginning. The data plays a critical role in incorporating movement information into the contagious process. Google collects anonymous mobility data, which reports the changes relative to the baseline value (Figure 2). The data reflects the decline of population movement after the outbreak. To accurately measure these changes in overall movement, we categorize the variables of Google data into "home", "workplace", "schools", and "others", i.e. $m_k(t)$ $(k=1,2,3,4)$.

Secondly, we present the contact matrix C. Contact matrix $C=C_{i\times j}$ is proposed by Reference [5]. When environment k (home, school, work, and others) is set, the age-specific contacts can be calculated through the hierarchical model of POLYMOD contact data. The POLYMOD project possesses a seminal study by Mossong et al.[6] in which the social contact structure of 100000 contacts across eight European countries is collected. Reference [7] supports the project and illuminates the strong associativity of social contacts with age driving the early period of an pandemic. When $k=1$ (home), c_{ij} shows the frequency of contact between the i-age and j-age group, see Figure 3 ($i,j=1-10$, 11−20, 21−30, 31−40, 41−50, 51−60, 61−70, over 71).

In the initial period, the model adopts E_i to record the number of the i-age exposed group. Suppose that the age structure for every country classifies the initial exposure. In a short period, transmission probability $\beta_0(t)=\beta_0$ maintains a constant value. And the change of mobility $m_k(t)$ brought by NPIs would influence contact frequency $\sigma(t)$.

Here needs an additional factor. For individuals in different age groups, the probability that an individual from S turns into E[8,9] manifests a difference which affected by his health and other factors. The probability or the attack rate h_i for i-age is estimated from the confirmed case (Figure 4). Taking the country as a whole, the attack rate of i-age is measured by the proportion of the hospitalization number of i-age in the entire number of hospitalizations.

Unlike the SEIR model, the M-SEIR model divides age into eight groups. Furthermore, it calculates the decreasing number of the susceptible group S_i in the i-age group separately after S_i is exposed to the infectious I_j^r or I_j^u in the j-age group. The following ODEs (Equation (4)) describes the decrease rate of total susceptible $S=\sum_{i=1}^{8} S_i$. Suppose that the initial number of j-age infectious I_j^r and I_j^u obeys the age structure w_j in the country, i.e., $I_j^r=I^r w_j$, $I_j^u=I^u w_j$. $N_j=Nw_j$ denotes the j-age population of the country. Suppose

$$C_i=\sum_{j=1}^{8} C_{ij}.$$

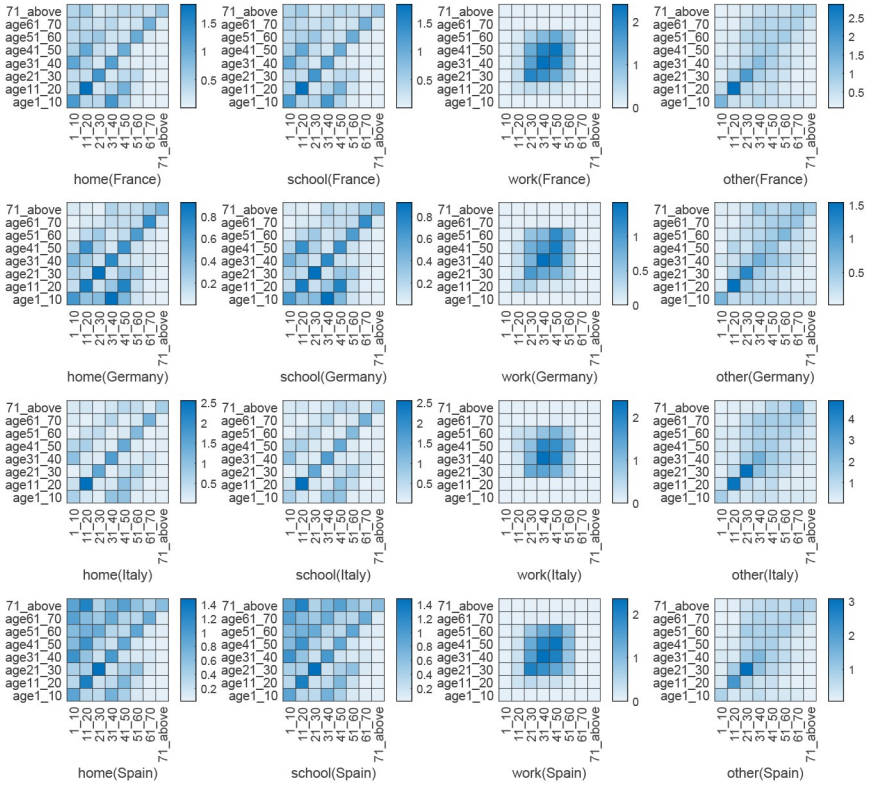


Figure 3. The contact matrix. In different environments, the frequency of contact between the i-age group and j-age exists varying. After quantification, the differences between other countries can be observed.

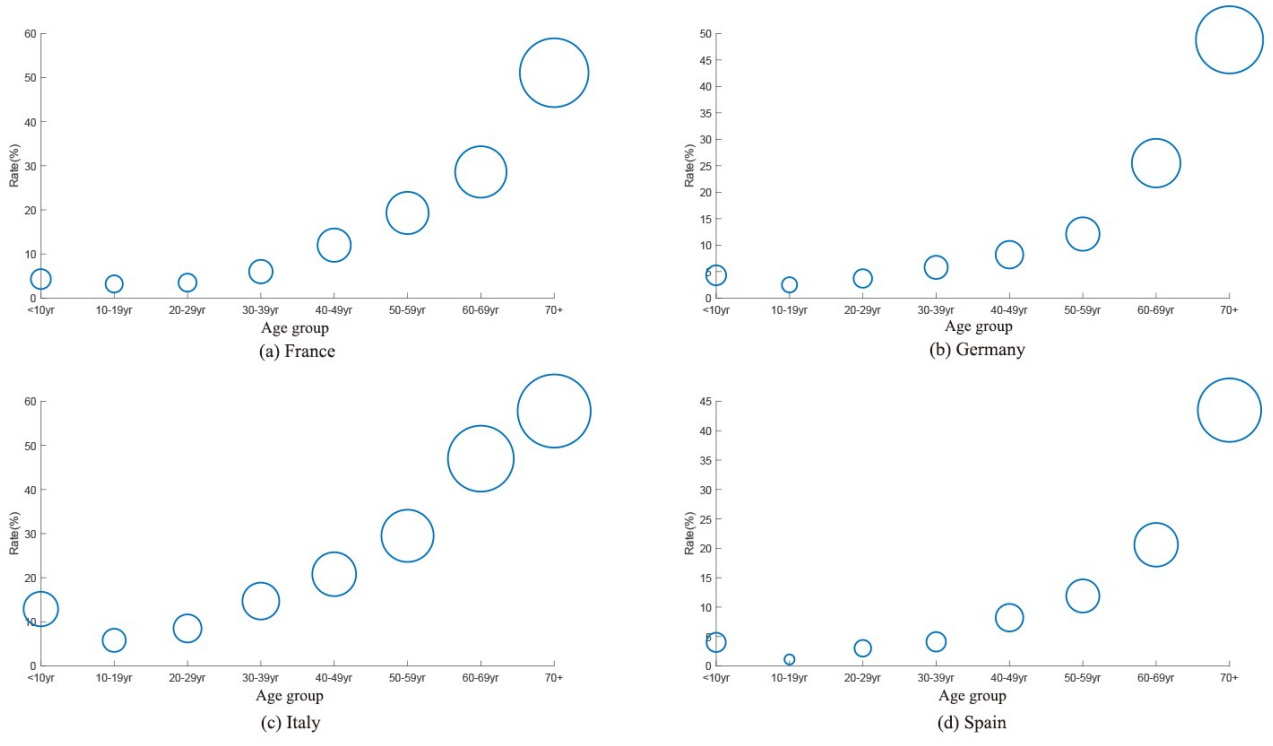


Figure 4. Attack rate. Takes country as a whole, the attack rate of i-age is measured by the proportion of the hospitalization number of i-age in the whole number of hospitalizations.

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= -\sum_{i=1}^{8}\beta_0 h_i S_i \sum_{j=1}^{8}\frac{C_{ij}I_j^r}{N_j} - \sum_{i=1}^{8}\mu\beta_0 h_i S_i \sum_{j=1}^{8}\frac{C_{ij}I_j^u}{N_j} = \\ &-\sum_{i=1}^{8}\beta_0 h_i S_i \sum_{j=1}^{8}\frac{C_{ij}I^r w_j}{Nw_j} - \sum_{i=1}^{8}\mu\beta_0 h_i S_i \sum_{j=1}^{8}\frac{C_{ij}I^u w_j}{Nw_j} = \\ &-\sum_{i=1}^{8}\beta_0 h_i S_i \sum_{j=1}^{8}C_{ij}\frac{I^r}{N} - \sum_{i=1}^{8}\mu\beta_0 h_i S_i \sum_{j=1}^{8}C_{ij}\frac{I^u}{N} = \\ &-\sum_{i=1}^{8}\beta_0 h_i S_i C_i \frac{I^r}{N} - \sum_{i=1}^{8}\mu\beta_0 h_i S_i C_i \frac{I^u}{N} = \\ &-\sum_{i=1}^{8}\beta_0 h_i S w_i C_i \frac{I^r}{N} - \sum_{i=1}^{8}\mu\beta_0 h_i S w_i C_i \frac{I^u}{N} = \\ &-\frac{SI^r}{N}\beta_0\sum_{i=1}^{8}w_i C_i h_i - \frac{\mu SI^u}{N}\beta_0\sum_{i=1}^{8}w_i C_i h_i \end{aligned} \tag{4}$$

At time t, C_i means the cumulative frequency of contact between i-age susceptible group and all infectious group. Define $C_i(t) = \sum_{k=1}^{4} m_k(t) C_{ik}$ and C_{ik} contains the average daily contact frequency of the i-age group in environment k. $\dot{C}(k) = \sum_{i=1}^{8} w_i h_i C_{ik}$ describes the average contact frequency in environment k (Table 1).

$$\begin{aligned}\frac{\mathrm{d}S_t}{\mathrm{d}t} &= -\frac{SI^r}{N}\beta_0\sum_{i=1}^{8}w_i h_i\sum_{k=1}^{4}m_k(t)C_{ik} - \\ &\frac{\mu SI^u}{N}\beta_0\sum_{i=1}^{8}w_i h_i\sum_{k=1}^{4}m_k(t)C_{ik} = \\ &-\frac{SI^r}{N}\beta_0\sum_{k=1}^{4}m_k(t)\dot{C}(k) - \frac{\mu SI^u}{N}\beta_0\sum_{k=1}^{4}m_k(t)\dot{C}(k) = \\ &-\frac{SI^r}{N}\beta_0\sigma(t) - \frac{\mu SI^u}{N}\beta_0\sigma(t) \end{aligned} \tag{5}$$

The model clarifies the impact of mobility after each country carries out NPIs, i.e.

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = -\frac{\beta(t)SI^r}{N} - \frac{\mu\beta(t)SI^u}{N} \tag{6}$$

It is similar to get

$$\frac{\mathrm{d}E_t}{\mathrm{d}t} = \frac{\beta(t)SI^r}{N} + \frac{\mu\beta(t)SI^u}{N} - \frac{E}{Z} \tag{7}$$

2.3 Data description

We collect daily counts of COVID-19 for 22 countries from European Centre for Disease Prevention and Control①. The data is collected up from Feb. 15, 2020 to Dec. 1, 2020, and contains most countries in Europe. The population mobility data at the country level comes from Google②. It provides social mobility from Feb. 15, 2020 to Dec. 1, 2020.

To facilitate meticulous analysis of the impact of mobility, we chose France, Germany, Italy, Spain as the primary analysis objects which encounter severe attacks and adopt relevant measures. Other countries are fitting objects.

① COVID-19 situation updates. Available at https://www.ecdc.europa.eu/en.

② COVID-19 community mobility reports. Available at https://www.google.com/covid19/mobility.

Table 1. $\dot{C}(k)$. The average contact times of all countries under different environment.

Country	home	work	school	other
Austria	2.99	3.25	2.53	4.99
Belarus	2.27	2.38	2.27	3.27
Belgium	2.48	2.02	3.48	4.48
Croatia	2.31	2.26	1.31	4.31
Czechia	2.95	3.14	1.95	3.95
Denmark	1.35	2.70	2.35	3.35
France	2.77	2.36	3.77	3.77
Germany	2.70	2.00	2.70	3.70
Greece	2.20	2.16	3.20	3.20
Hungary	2.38	2.10	3.38	3.38
Ireland	2.01	1.37	1.01	4.01
Italy	3.25	2.27	3.25	4.25
Netherlands	3.48	2.02	3.48	4.48
Norway	2.67	2.16	2.67	4.67
Poland	3.31	2.26	3.31	4.31
Portugal	2.95	2.14	2.95	4.95
Romania	3.35	2.70	3.35	4.35
Serbia	3.18	2.61	2.18	4.18
Spain	2.73	2.61	1.73	3.73
Sweden	2.20	2.16	1.20	3.20
Switzerland	2.38	2.10	2.38	3.38
UK	2.01	2.37	4.01	4.01

3 Algorithm

Theoretically, the M-SEIR model can accurately estimate the daily case. Its result owns a perfect curve. Empirically, the actual curve illustrates volatility affected by various uncontrollable factors. This volatility may influence the fitting. Just the M-SEIR model cannot explain the existence of volatility. Meanwhile, there is an unknown variable $\beta_0(t)$ in the M-SEIR model. $\beta_0(t)$ is subject to the time limit. When the model directly adopts the maximum likelihood to determine the optimal parameters, it must set the segmentation points in advance and fit each stage separately if we want to control $\beta_0(t)$. So we use the iterative filtering ensemble adjustment Kalman filter (IF-EAKF) to deal with the volatility and avoid segmentation.

Determine whether its system obeys linear and the Gaussian distribution noise occupies the first position for estimation problems. If the motion equation and the observation equation are linear, the Gaussian distribution maintains Gaussian after linear transformation in a linear system with Gaussian noise. We can describe its distribution as long as we calculate its first and second moments. Therefore, the Bayes rule calculates the posterior probability distribution of sample x for the system. Kalman filter (KF) provides an unbiased optimal estimation in the recursive form of the linear system (estimating x_k from x_{k-1}).

However, the model in the paper obeys a non-linear, non-Gaussian system. The usual methods for the system include extended Kalman filter (EKF), unscented Kalman filter (UKF), central differential Kalman filter (CDKF), cubature Kalman filter (CKF) and non-linear optimization[10]. EKF converts the non-linear equation into a linear equation by the first-order expansion of Taylor and then applies the same method as KF to process it. Nevertheless, the loss of accuracy also appears severe. Unscented Kalman filter (UKF), central differential Kalman filter (CDKF), and volume Kalman filter (CKF) totally adopt some fixed sampling points (σ points) to calculate the propagation of these σ points in a non-linear function. The updated value of σ is obtained through the non-linear state equation. This method does not directly linearize the non-linear equation but introduces the weighted σ into the non-linear equation to approximate the actual result. Besides, if we discard the Gaussian assumption, we can adopt enough sampling points to express the output distribution. This Monte Carlo method equals the idea of particle filtering (PF). Furthermore, non-linear optimization chooses to discard the filter idea. That is, the next moment does not depend on the value of the last moment. Then the process regards all the states as variables, and the motion equations and observation equations are constraints between variables. Constructing an error function and then minimizing the error can help us obtain the solution. However, non-linear optimization also requires continuous gradients of the error function and iterates by the gradient direction, so local linearization becomes inevitable. Therefore, we consider IF-EAKF as an adjustment method for estimating such non-linear models. The idea of IF-EAKF is similar to PF but more accurate. It randomly collects some points to form a set of states and then brings them into the non-linear equation. Next, IF-EAKF adjusts the gap between the actual value and the estimated value in the iterative process. After determining the initial state, we obtain the estimated value in the next period by the M-SEIR model. For this point, IF-EAKF can adjust the value. Then the algorithm brings the modified estimator into the M-SEIR model again and continues to iterate. Adjustment by EAKF happens in each period. The estimated results of the model are more consistent with the actual value.

The main idea of the algorithm is as follows: randomly extracts the initial values of the parameters and variables. Then the model iterates and the algorithm adjusts the estimator to obtain the changing trends of the parameters and variables.

3.1 Setting initial value

According to the present situation, the average incubation period of COVID-19 has one week. So the initial fitting date is one week before the first reported date in the country. If there exists a time lag between the first reported date and the subsequent reports, we choose the point in the following reports to replace the first reported date (Figure 5). The first reported day in France is February 9th, and the second reported day is February 19th. The time interval is more than one

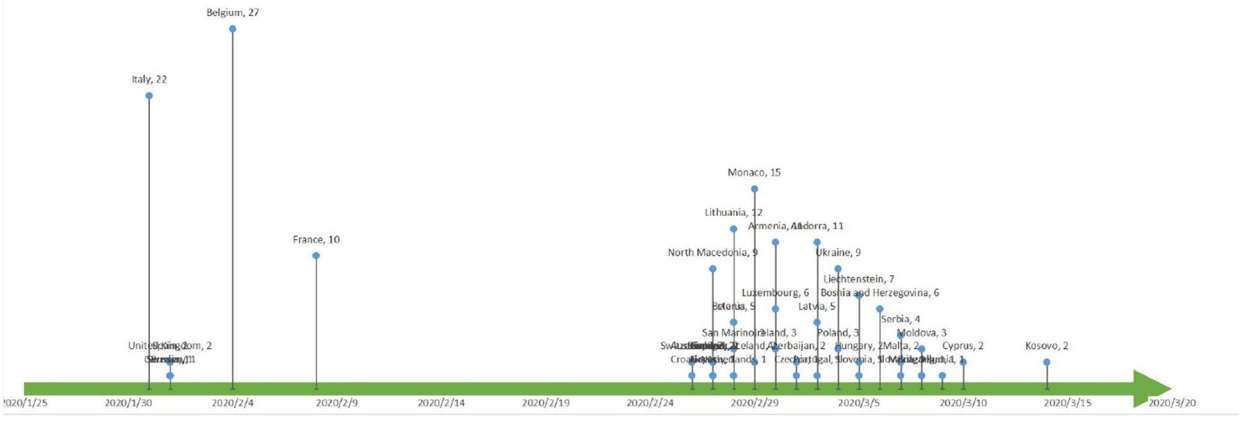


Figure 5. The date when every country first reported domestic cases. The numbers on the axis indicate the time difference between the first and second reports.

week, so the initial date chooses February 12th, a week before February 19th.

The fitting needs the initial number of individuals in different categories and the initial values of the parameters. For the initial unit of different groups, the number of the suspected group S_0 equals the total population of the country, and the number of confirmed cases I_0^r is 0. The number of I_0^u comes from the interval $[0,50]$, and the corresponding initial exposure E_0 is between 0.5 and 3 times of I_0^u.

Then we should choose the initial unit of parameters. The prior range of the parameter can be shown in Table 2[11] and samples the initial parameters unit from the range. The initial unit contains the above two parts and then iterates the unit by Runge-Kutta 4th method.

Table 2. The prior ranges of parameters.

	β	μ	α	Z	D
The lower value	0.01	0.1	0.01	2	2
The upper value	2	1	1	5	5

3.2 Method

The algorithm applies the IF-EAKF to obtain sequences of the state. The principle of IF follows[12]:

① Iterate the initial unit through a specific function and get the next unit.

② Adjust the obtained unit by filtering techniques, and each adjustment causes the variance of variables to decrease.

③ Bring the adjusted unit into the iterative function again to get the next iteration.

Note that in the filtering process, the adjustment technique adopted here is the EAKF. The robustness of the estimated result in Reference [11] proves the superiority of the method. The fitting only obtains an initial unit consisting of five state variables and six parameter variables. State variables contain the number of five groups in the model, and the most critical variable adopts the observed variable o_t, i. e. daily reported case. Algorithm infers other state variables x_t from the non-linear relationship which relates to the variable and the observed variable. Three hundred state units constitute an integrated member, which plays an essential role in helping the algorithm to break through the linear constraints of the Kalman algorithm[13]. Based on that, although ODES has uncertain parameters and a non-linear relationship, EAKF can make a right prediction for iteration trend of the next state variables by the Bayesian method and the following Equation (8).

$$o_{t,\text{post}}^k = \frac{\sigma_{t,\text{obs}}^2}{\sigma_{t,\text{obs}}^2 + \sigma_{t,\text{prior}}^2}\bar{o}_{t,\text{prior}} + \frac{\sigma_{t,\text{prior}}^2}{\sigma_{t,\text{obs}}^2 + \sigma_{t,\text{prior}}^2}o_t + \sqrt{\frac{\sigma_{t,\text{obs}}^2}{\sigma_{t,\text{obs}}^2 + \sigma_{t,\text{prior}}^2}}(o_{t,\text{prior}}^k - \bar{o}_{t,\text{prior}}) \tag{8}$$

The state variable $o_{t,\text{post}}^k$ implies the posterior estimate of daily reported case at the kth state unit. o_t means true daily reported case. We apply $o_{t-1,\text{post}}^k$ to predict the state variable $o_{t,\text{prior}}^k$ in the next iteration and get the prior of daily reported case $o_{t,\text{prior}}^k$. Mean $o_{t,\text{prior}}^k$ in 300 state units can get $\bar{o}_{t,\text{prior}}$. Then we calculate the variance of $o_{t,\text{prior}}$ and the true value o_t respectively. Two variances are recorded as $\sigma_{t,\text{prior}}^2$ and $\sigma_{t,\text{obs}}^2$. After adjusting iteration, algorithm can obtain the $o_{t,\text{post}}$. Then it can receive other state variables $x_{t,\text{post}}$ from the covariance between the variable and the observed variable (Equation (9)). With the obtained trend of state variables, we adopt Runge-Kutta 4th and the M-SEIR model to get the next tendency of parameters.

$$x_{t,\text{post}}^i = x_{t,\text{prior}}^i + \frac{\sigma(\{x_{t,\text{prior}}\}_{n'}\{o_{t,\text{prior}}\}_n)}{\sigma_{t,\text{prior}}^2}(o_{t,\text{post}}^i - o_{t,\text{prior}}^i) \tag{9}$$

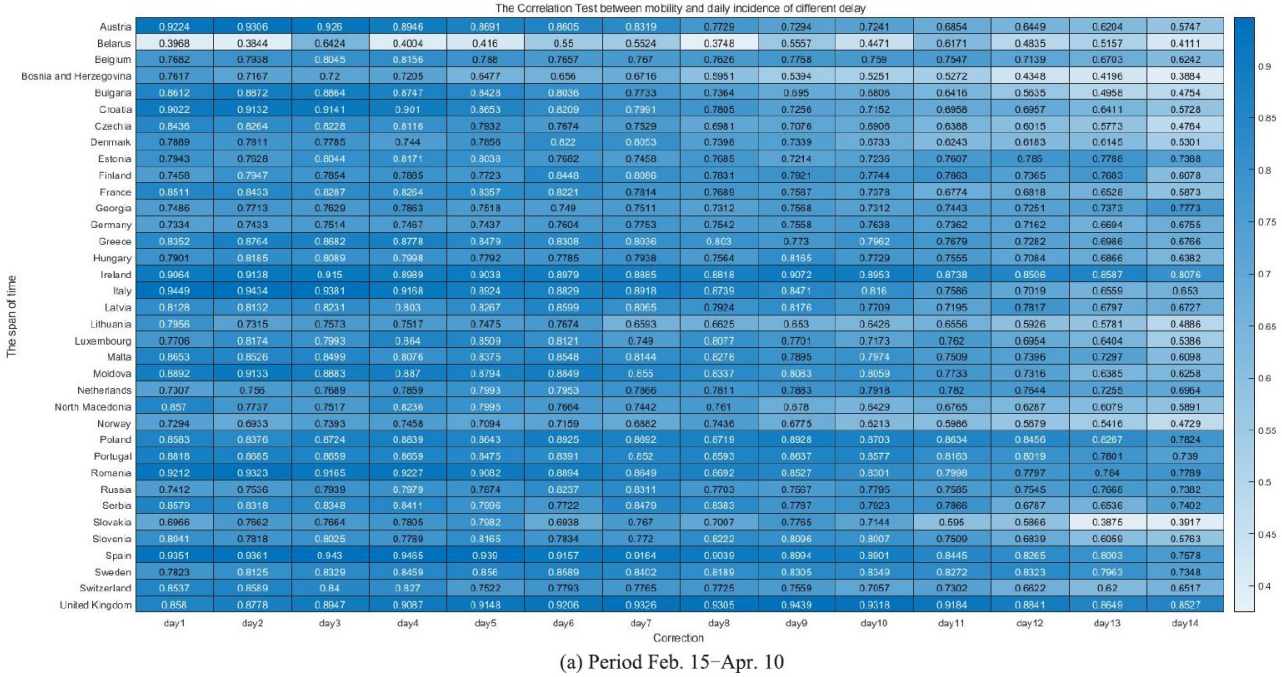
In a system with uncertain factors, IF-EAKF can help the iterative value approach the actual value. Also, set the integrated unit can reduce the random error caused by the initial state sampling. As the number of iterations increases, the fluctuating trend of the estimated parameter decreases in the subsequent iteration process. This work would help us to receive the trend of the transmission rate.

4 Empirical analysis

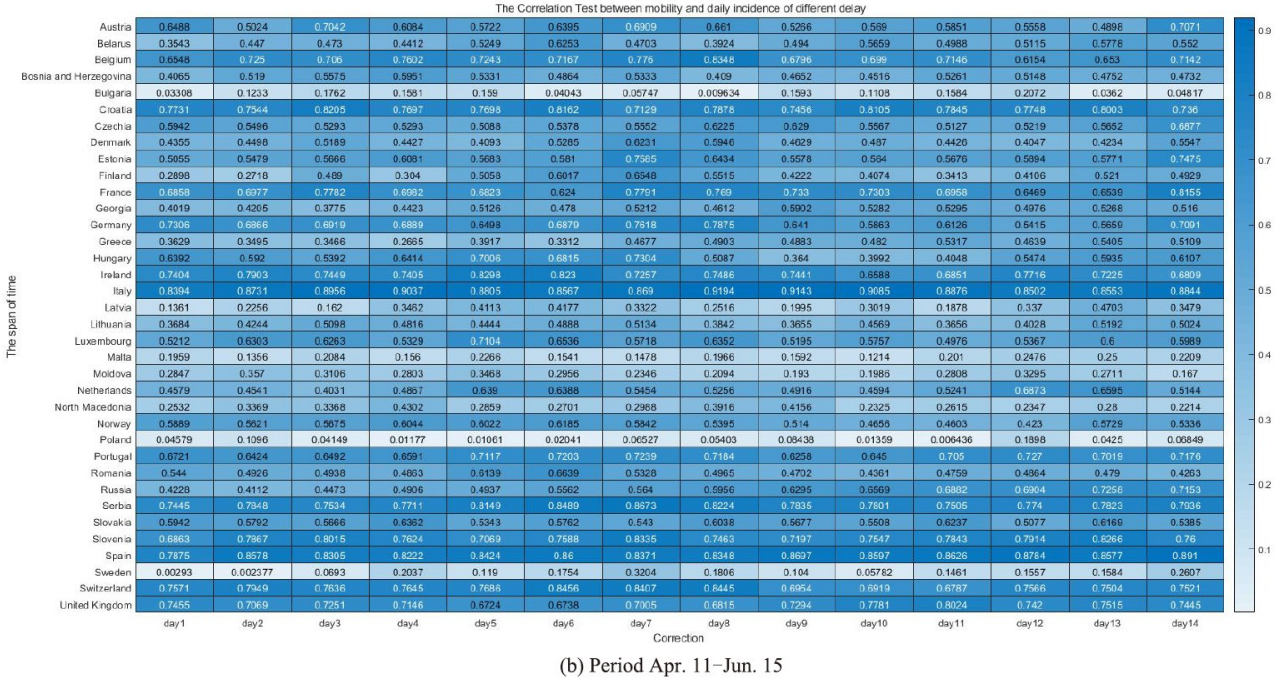
4.1 Correlation test

Notice that the effect of mobility may not be reflected timely in cases reported on the same day. Therefore, the fitting needs to consider the delayed impact of mobility on daily cases, i. e., add a delay factor to mobility sequence $\sigma(t)$. The section tests the correlation between $\sigma(t)$ and the daily case during the first round of the pandemic. The test calculates Spearman rank correlation coefficient between $\sigma(t)$ and the daily cases under different time-delay to observe the time lag.

The correlation applies segmentation to make the test. Figure 6(a) presents the correlation value between $\sigma(t)$ and the daily reports under different delays. In actuality, $\sigma(t)$ is correlated negatively with the daily case. Because the value appears negative, the test depends on the absolute value. The period in Figure 6(b) starts from April to June 2020. Considering the mobility delay-effect (previous 0 − 13 days), the



(a) Period Feb. 15–Apr. 10



(b) Period Apr. 11–Jun. 15

Figure 6. The test of the relationship between mobility and daily case((a)Feb. 15, 2020−Apr. 10, 2020; (b)Apr. 11, 2020−Jun. 15, 2020). The abscissa shows the delay days, the ordinate represents the country.

correlation in mobility under three-day delay appears the best performance. After May, the population movement gradually rebounds, and the number of daily cases increases in the country. Nevertheless, the spread of the virus is significantly suppressed for most countries due to some external factors such as temperature factors. The correlation also declines slightly (Figure 6(b)).

4.2 The fitting for the first round

Sections 2.1 and 2.2 construct the model framework of the research. The IF-EAKF algorithm described in Section 3 can continuously adjust the difference between the actual and estimated value. Then we can optimize the M-SEIR model and get the tendency of parameters in the model.

If we apply the M-SEIR model to fit the daily case, it must adopt piecewise fitting and determine the parameters of each segment by maximum local likelihood. The selection of segment points and the choice of initial values significantly influence the fitting effect. So we should optimize the M-SEIR model at first.

The idea of the paper proves excellent by fitting the first round of the pandemic. The experiment subjects include France, Germany, Italy and Spain. After the iteration in the SEIR model, the result contains the transmission sequence $\beta(t)$ in the first period (Figure 7). When the mobility effect $\sigma(t)$ doesn't appear in the SEIR model, we obtain the varying trend of transmission $\beta(t)$ by IF-EAKF.

Their changing inclinations are highly consistent compared with $\sigma(t)$ obtained in Section 2. However, the estimated daily case fluctuation appears too noticeable, and this fluctuation is not conducive to the model fitting. So we put $\sigma(t)$ in the SEIR model for iteration. The iteration can acquire a better tendency of the transmission rate $\beta(t)$ (Figure 8). At this time, the fluctuation range of the estimated daily case is further compressed and more consistent with the fluctuation of the actual value.

Even though the tendency of $\sigma(t)$ and $\beta(t)$ by the M-SEIR model owns strongly consistent, the transmission decline in the former looks higher than in the latter. Practically, the relaxation of control does not immediately rebound at the end of the first round (in June). These appearances also explain the need to consider the transmission probability of the virus. Section 2.2 emphasizes the importance of a reasonable period to ensure that the transmission probability remains stable. Therefore, we observe $\beta(t)$ and $\sigma(t)$, and then calculate the trend of the virus transmission probability $\beta_0(t)$ from Equation (2) (Figure 8). External factors lead to a slice of gaps in the transmission probability and change the transmission rate in turn. After referring to other relevant materials[14-16], the temperature and humidity may be indispensable in discussing the transmission probability. To further discuss the $\beta_0(t)$, the variable is divided into monthly sections and regresses with the monthly local temperature, humidity (Figure 9). Table 3 shows its regression result.

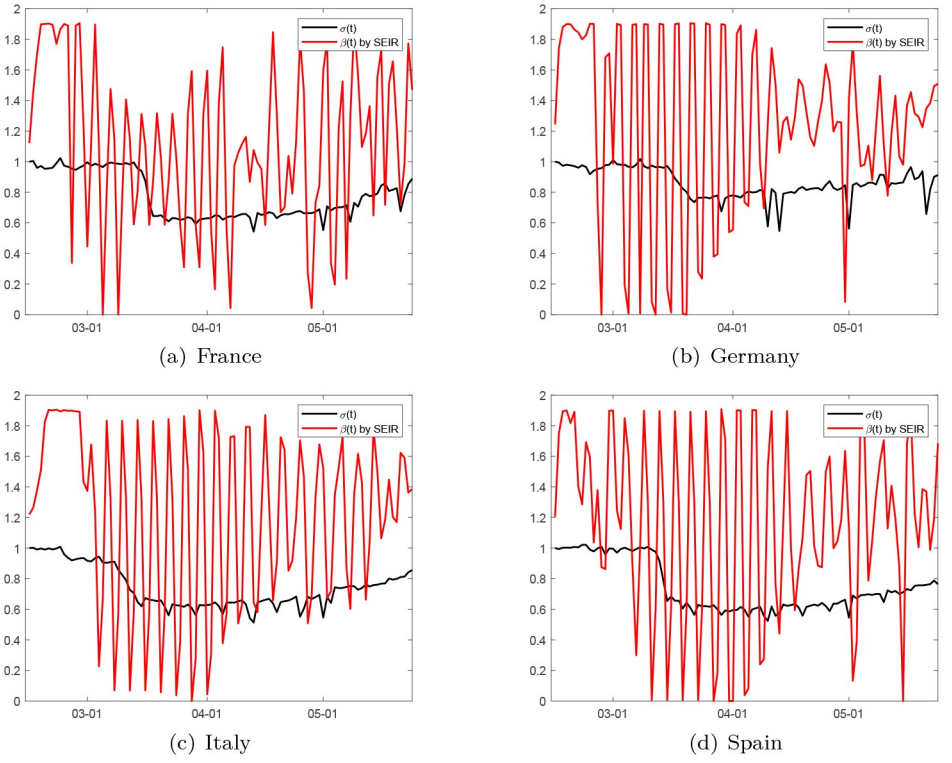


Figure 7. The comparison between mobility $\sigma(t)$ and the transmission rate $\beta(t)$ obtained by the SEIR model(Period: Feb. 15, 2020 -May 31, 2020).

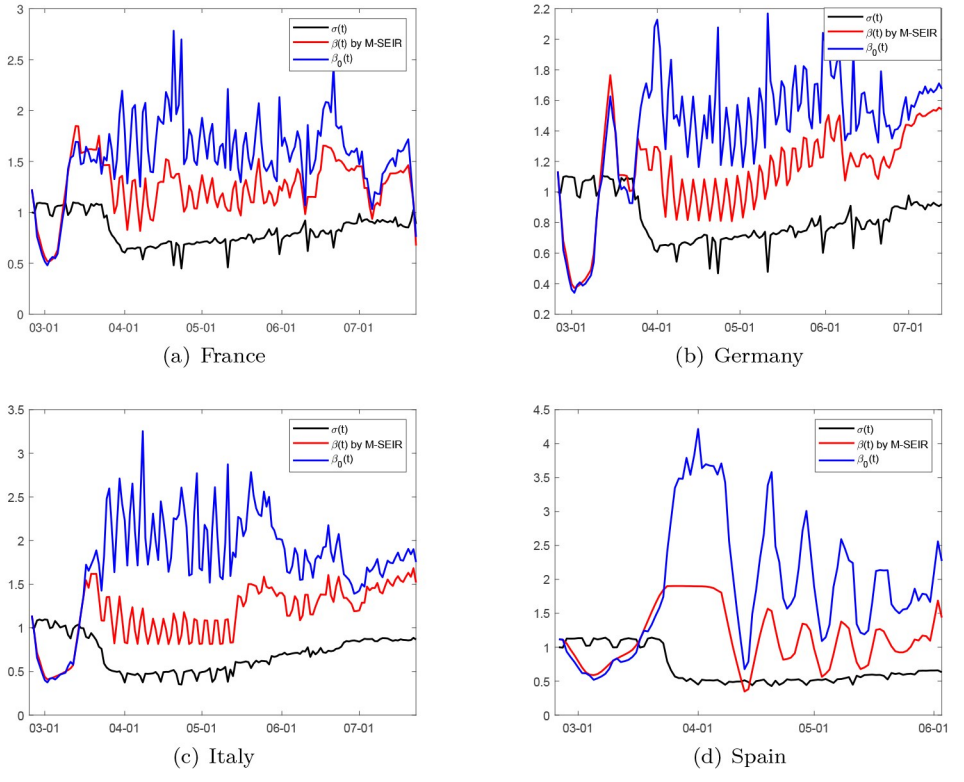


Figure 8. The figure observes the trend changes of $\beta(t)$ and $\sigma(t)$, and calculate the trend of the virus transmission probability $\beta_0(t)$ from Equation (2) (Period: Feb. 25, 2020–Jul. 15, 2020).

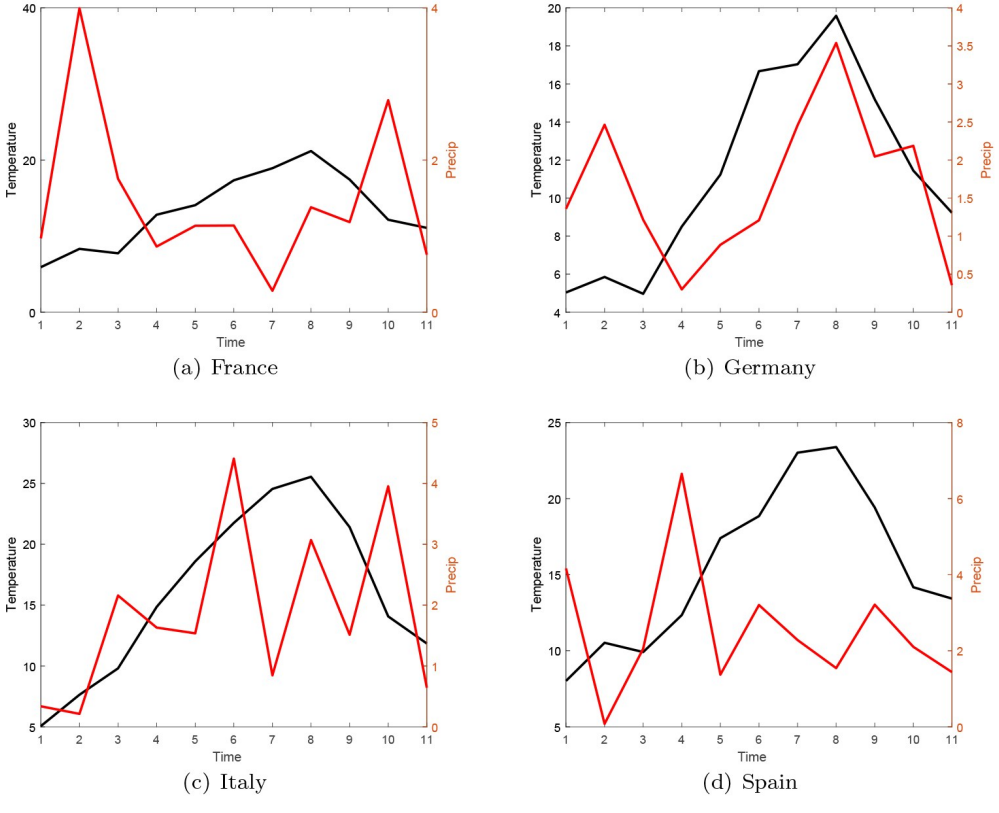


Figure 9. The temperature and humidity (Period: Jan. 2020 – Nov. 2020). Local temperature and humidity are divided into monthly sections and regress with $\beta_0(t)$.

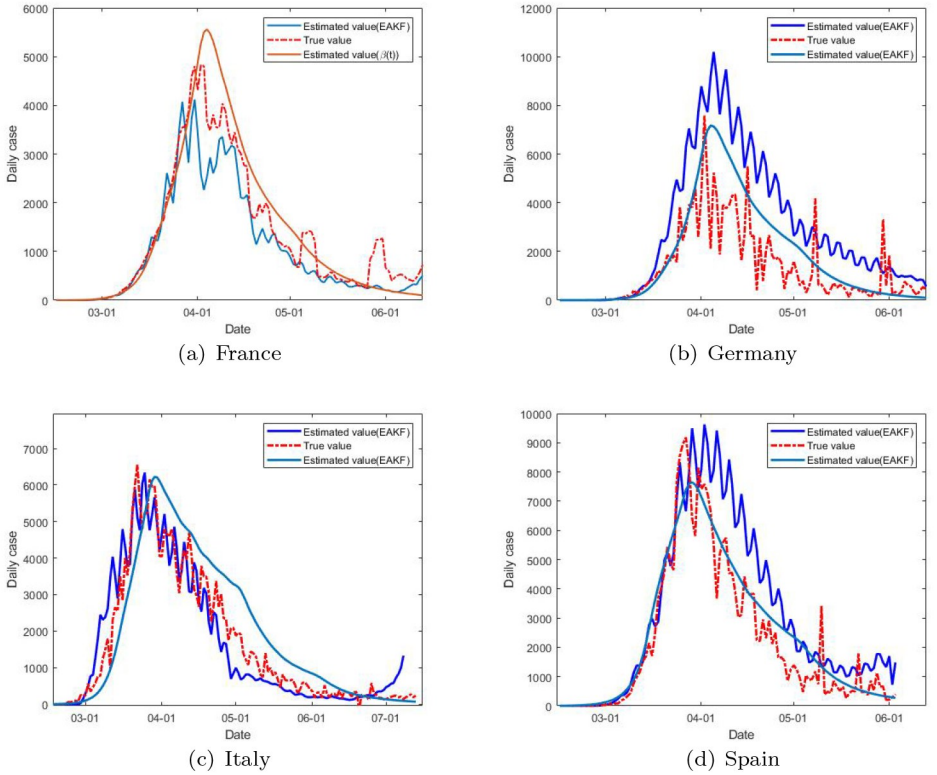


Figure 10. The comparison of fitting daily case by adopting IF-FAKF and adjusted the M-SEIR model(Period: Feb. 15, 2020–Jun. 15, 2020).

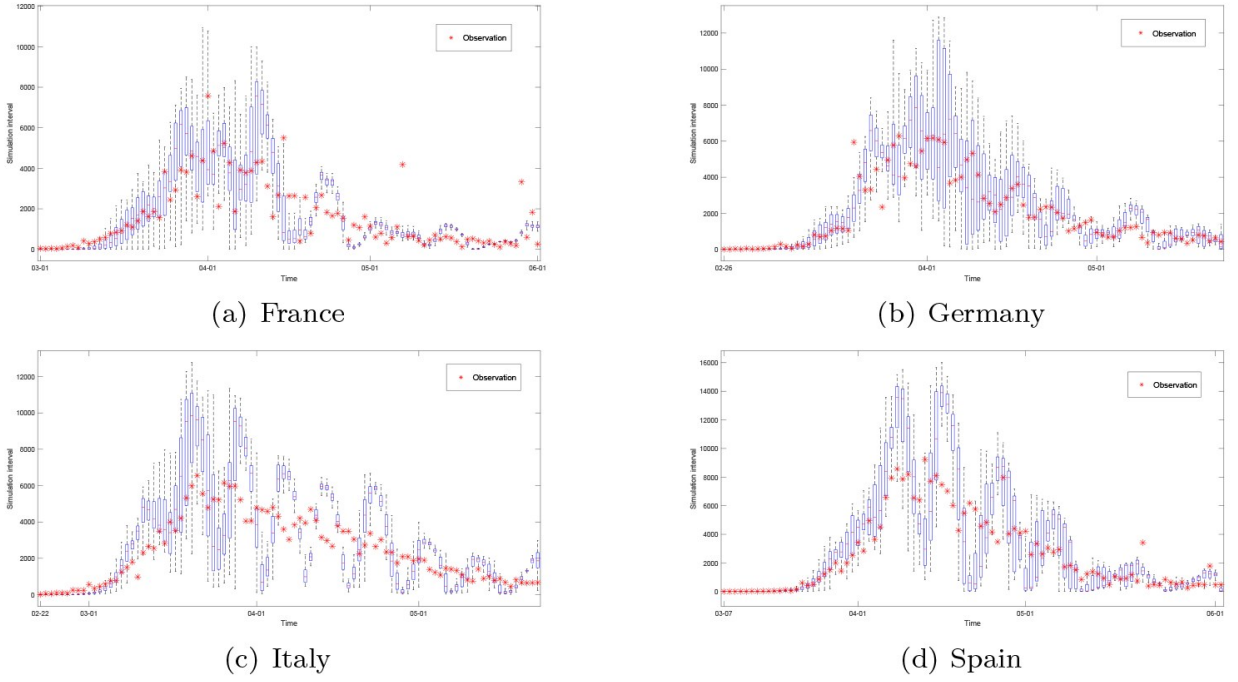


Figure 11. The confidence interval of the daily data and the observation. Show 100 experiment results by plotting a box line map, and * means the observation(Period: Feb. 2020–Jun. 2020).

Consider $\beta_0(t)$ affected by temperature and humidity, and we can iterate the case directly on the M-SEIR model by maximum likelihood instead of using IF-EAKF. Now we do not need to worry segmented. After making relevant adjustments involving the above steps, we set ① only the M-SEIR model, ② the M-SEIR model by IF-EAKF to fit the data. The results are compared in Figure 10. The goodness of fit test R^2

(Table 4) proves that the model has an excellent performance.

The iterative process needs to set up the initial value in advance and the difference in the initial value may affect the fitting of the model. Therefore, we change the initial values of the relevant variables and parameters, and then iterate the model 100 times. Figure 11 shows the confidence interval of the daily data by IF-EKAF. Moreover, the confidence interval of the estimated parameter appears in Table 5. Figure 12(a) and Table 6 illustrate the situation of only using the M-SEIR model. The result shows that IF-EKAF can break through the initial value constraints and converge to the optimal value in the iterative process to achieve the robustness of the results.

The fitting effect of only the M-SEIR model looks better than using IF-EAKF in the M-SEIR model. So we apply the model without IF-EAKF in other countries. The estimated value is close to the actual data for these countries (Figure 12(a)). Practically, the goodness of fit test R^2 seems to own a good fit (Table 7). The subsequent Q-Q plot also demonstrates that the residual between the estimated and the valid values are in an acceptable range (Figure 12(b)).

4.3 Predict and discuss

Until May 2021, Europe had faced three rounds of pandemics. The first rebound started in March 2020 and ended at the end of May 2020. The boundary between the second and the third time appears not clear. So we set that the period from August to December includes the second round. The third time is from January to May 2021.

The prediction regards the second round as their predicted object. In October 2020, France and Germany redeclared the state of emergency and sealed off all countries. To determine this measure can play a similar role in curbing the spread of the pandemic, we make a prediction based on the following assumption. Considering the level of subsequent government intervention: ①Enter the direct control state of April. The government immediately adopts multi-faceted control, i.e., 100% control. ②On the existing basis, reduced to 50% control. ③Reduced to 80% control. ④ Reduced to 90% control. ⑤No limit.

Table 3. Regression of temperature and humidity on $\beta_0(t)$.

Value	Coefficients	Standard error	T-statistic	P-value
Intercep	0.8218	0.1257	6.5365	2.01E−07
Humidity	0.0181	0.0293	0.6203	0.05393
Temperature	−0.0268	0.0072	−3.7257	0.0007

Table 4. The comparison of fitting effect after adjusting the model(Period: Feb. 15, 2020−Jun. 30, 2020).

		MSE	RMSE	MAE	R^2
France	M-SEIR	205891.3470	453.7525	297.4607	0.8743
	IF-EAKF	282397.9405	531.4113	311.4429	0.8276
Germany	M-SEIR	811758.4386	900.9764	816.8417	0.7324
	IF-EAKF	1480926.5225	1216.9332	907.3515	0.5119
Italy	M-SEIR	794695.5253	891.4570	782.8266	0.7492
	IF-EAKF	1169596.6770	1081.4789	777.8924	0.6308
Spain	M-SEIR	467428.2965	683.6873	488.2388	0.9250
	IF-EAKF	1964072.8691	1401.4538	931.2192	0.6849

Table 5. The confidence interval of the estimated parameter by IF-EAKF (CI:95%)(Period: Feb. 15, 2020−Jun. 30, 2020).

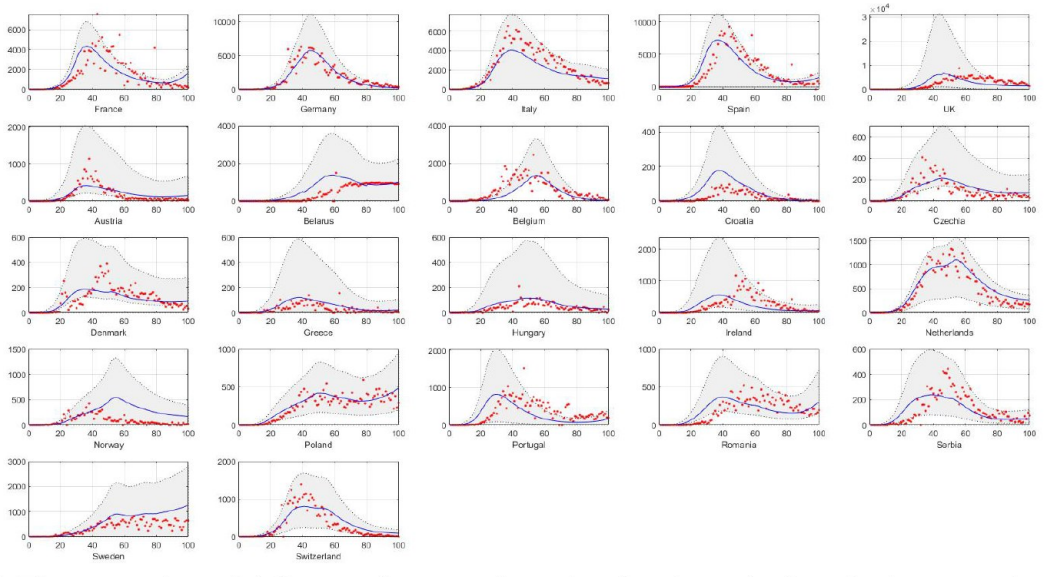
	β	μ	α	Z	D
France	0.9051(0.8970,0.9132)	0.5748(0.5723,0.5774)	3.3921(3.3815,3.4027)	0.4676(0.4642,0.4711)	3.3851(3.3742,3.3959)
Germany	0.8966(0.8922,0.9010)	0.5694(0.5680,0.5707)	3.4005(3.3954,3.4056)	0.4604(0.4586,0.4623)	3.3921(3.3871,3.3972)
Italy	0.9825(0.9792,0.9858)	0.6098(0.6083,0.6112)	3.4798(3.4753,3.4842)	0.5106(0.5087,0.5124)	3.4761(3.4718,3.4805)
Spain	0.9573(0.9544,0.9601)	0.5959(0.5947,0.5972)	3.4653(3.4610,3.4696)	0.4937(0.4922,0.4952)	3.4581(3.4537,3.4624)

Table 6. The confidence interval of the estimated parameter by the M-SEIR model(CI:95%)(Period: Feb. 15, 2020−Jun. 30, 2020).

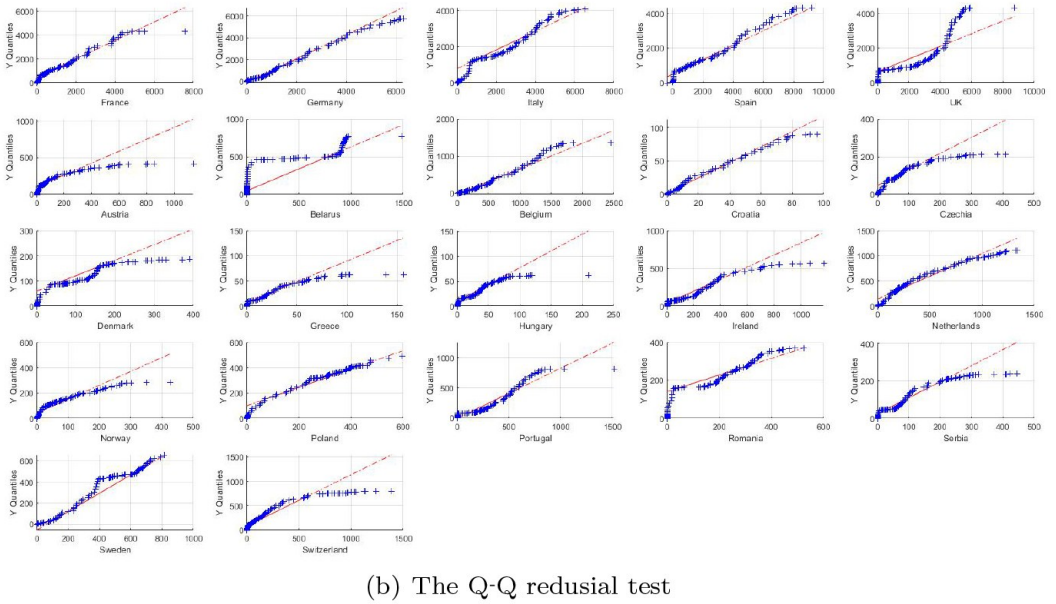
	β	μ	α	Z	D
France	0.9207(0.8971,0.9233)	0.5790(0.5719,0.5790)	3.3333(3.3114,3.4026)	0.4839(0.4645,0.4913)	3.3324(3.3135,3.3943)
Germany	0.8860(0.8731,0.9122)	0.5736(0.5682,0.5826)	3.3114(3.2964,3.3536)	0.4755(0.4686,0.4790)	3.3107(3.2835,3.3562)
Italy	1.0026(0.9792,1.0855)	0.5719(0.5583,0.60334)	3.3491(3.3025,3.4046)	0.4758(0.4337,0.5122)	3.3471(3.2714,3.4002)
Spain	0.9075(0.8944,0.9201)	0.5704(0.5243,0.5961)	3.3535(3.3110,3.3826)	0.4732(0.4234,0.4972)	3.3515(3.2847,3.4139)

Table 7. The estimated parameters of other countries in Europe and the fitting effect evaluation(Period: Feb. 15, 2020–Jun. 30, 2020).

Country	β	μ	Z	α	D	R^2
Austria	1.0486	0.4725	3.0630	0.3508	3.0644	0.6369
Belarus	1.0545	0.4814	3.0931	0.3623	3.0919	0.5261
Belgium	1.1571	0.6113	3.5497	0.5248	3.5480	0.4137
Croatia	1.0150	0.4299	2.9140	0.2967	2.9147	0.5887
Czechia	1.0145	0.4303	2.9068	0.2981	2.9071	0.5406
Denmark	1.0627	0.4912	3.1339	0.3748	3.1320	0.4361
Greece	1.0842	0.5184	3.2140	0.4075	3.2153	0.4935
Hungary	1.0457	0.4697	3.0497	0.3467	3.0486	0.5139
Ireland	1.0668	0.4957	3.1416	0.3801	3.1414	0.7685
Netherlands	1.0678	0.4977	3.1367	0.3815	3.1382	0.8781
Norway	1.0892	0.5247	3.2399	0.4154	3.2419	0.4630
Poland	1.0841	0.5182	3.2133	0.4081	3.2125	0.8078
Portugal	1.1046	0.5448	3.3182	0.4410	3.3168	0.0857
Romania	1.1821	0.6432	3.6792	0.5655	3.6788	0.3903
Serbia	1.1020	0.5417	3.3043	0.4378	3.3045	0.6720
Sweden	1.0918	0.5279	3.2512	0.4202	3.2537	0.7061
Switzerland	1.0425	0.5256	3.3112	0.4236	3.1122	0.6321
United Kingdom	1.1454	0.5967	3.5077	0.5069	3.5013	0.7123



(a) The comparison of daily cases between the real and estimated value of other countries



(b) The Q-Q redusial test

Figure 12. The empirical test.

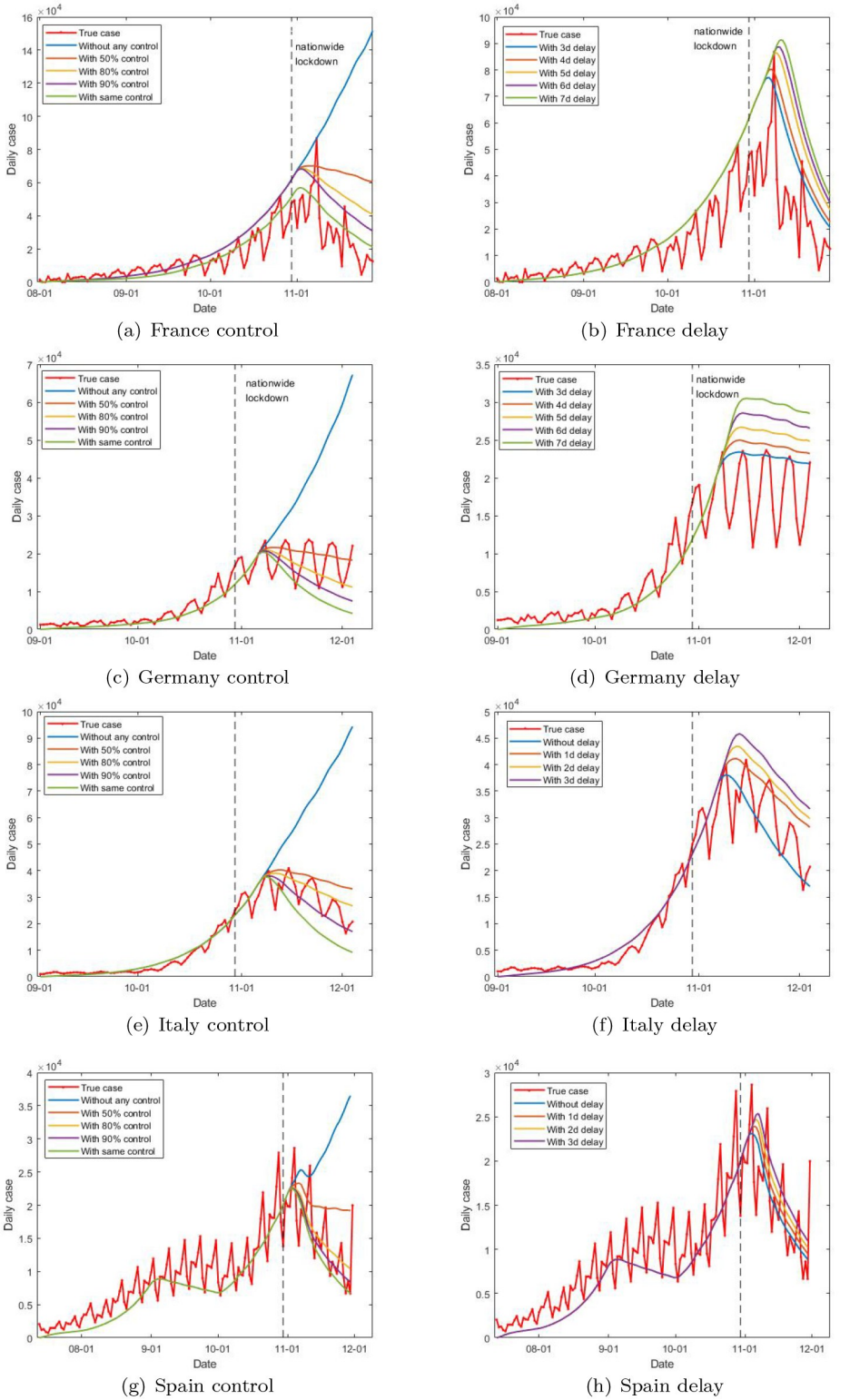


Figure 13. The prediction (Period: Sep. 1, 2020 – Dec. 1, 2020). In (a),(c),(e),(g), the model predicts the daily tendency after the country takes different control. (b),(d),(f),(g) imply the delay-effect under the measure that is most consistent with the actual situation.

Under the above assumptions, the situation of each country stays not the same (Figure 13). In France, the second implementation at least achieves 80%-control of the first. It owns a significant effect on the rapid spread of the disease. In Germany, the second implementation of the NPIs achieves only about 50% of the first. In the implementation stage, the daily case decreases slowly. The policy causes a more severe rebound in numbers

during the later period. Comparing Spain and Italy in an emergency at the same moment, the implementation of NPI reaches more than 75% of that in the first, which possesses a significant effect on curbing the rapid spread of the disease.

On the other hand, there exists a time lag between the point when the government announces the policy and when it plays a corresponding role. Therefore, the government should consider the delayed factors of the policy (Figure 13). If we set the announcement point as October 30th, 2020, when France and Germany imposed the lockdown. Differences occur in peak times under different delay days. In France, the NPIs come into effect four days after the government announces it. In Germany, the delayed time has three days. Unfortunately, the measures are not strong enough to reduce transmission. For the sake of comparison, suppose Italy and Spain redeclare a state of emergency at the same time. NPIs in two countries respectively achieve their desired effect with one delay-day and two delay-day. The work manifests that for Italy and Spain, the relevant actions are always well implemented in emergencies. This prediction can help policymakers analyze the time lag of the policy when facing the outbreak of infectious diseases.

Meanwhile, when we compare the various parameters of three rounds of pandemics (Table 8), the average infection speed of the second round looks significantly faster than the ones during the first and third rounds. Moreover, the number of reproductives R_t also confirms that the second round is higher than other rounds. For some countries, such as Austria, Germany, Italy, the number of reproductives rises again in the latter period of the second round. It can be approximately regarded the second and third rounds in these countries as a continuous process. $R_e(t)$ reflects the above result in Figure 14.

Table 8. The parameter in three rounds of the pandemic. The number implies the time of round.

		β		μ		Z		α		D		R_e		Ratio Correlation
		SEIR	M-SEIR	SEIR	M-SEIR	SEIR	M-SEIR	SEIR	M-SEIR	SEIR	M-SEIR	SEIR	M-SEIR	
	1	0.9843	1.0779	0.6084	0.6465	3.5432	3.6761	0.5153	0.5635	3.5429	3.6732	0.7975	0.9116	1.0000	
Austria	2	1.0018	1.0167	0.6029	0.6136	3.5060	2.6044	0.5092	0.7079	3.5060	2.5736	0.8066	0.9019	0.9894	
	3	0.9145	0.9844	0.5695	0.6027	3.3850	3.4611	0.4670	0.5091	3.3861	3.4601	0.7047	0.7924	0.8693	−0.7453
	1	0.7806	0.8740	0.5287	0.5653	3.2344	3.3634	0.4151	0.4613	3.2301	3.3606	0.5654	0.6693	1.0000	
Bulgaria	2	1.0480	1.0755	0.6281	0.6388	3.5845	3.6224	0.5404	0.5540	3.5857	3.6218	0.8689	0.9022	1.3479	
	3	0.9720	0.8434	0.6045	0.5813	3.4885	3.4298	0.5108	0.4821	3.4899	3.4304	0.7840	0.6388	0.9543	−0.8465
	1	0.9403	0.8893	0.5951	0.5695	3.4094	3.3349	0.4981	0.4661	3.4060	3.3341	0.7493	0.6849	1.0000	
Belarus	2	0.8657	0.8815	0.5815	0.5957	3.2852	3.3037	0.4711	0.4845	3.2848	3.3049	0.6741	0.6978	1.0188	
	3	0.8947	0.8715	0.5719	0.5570	3.3402	3.2534	0.4693	0.4571	3.3403	3.2534	0.6914	0.6619	0.9664	−0.6432
	1	0.9934	0.9676	0.6172	0.6012	3.5551	3.4627	0.5281	0.5144	3.5562	3.4638	0.8140	0.7802	1.0000	
Czechia	2	1.0320	1.0052	0.6188	0.6027	3.5578	3.4653	0.5298	0.5160	3.5581	3.4656	0.8471	0.8119	1.0407	
	3	0.9467	0.9221	0.5826	0.5674	3.4328	3.3436	0.4835	0.4709	3.4342	3.3449	0.7426	0.7111	0.9114	−0.7962
	1	0.8722	0.8495	0.5688	0.5541	3.3717	3.2840	0.4663	0.4542	3.3724	3.2847	0.6715	0.6427	1.0000	
Denmark	2	0.9480	0.9234	0.6028	0.5871	3.4089	3.3203	0.5020	0.4889	3.4100	3.3214	0.7605	0.7285	1.1335	
	3	1.1127	0.8136	0.6927	0.6747	3.7819	3.6835	0.6228	0.6066	3.7854	3.6870	0.9837	0.6401	0.9959	−0.5743
	1	0.9453	0.9207	0.5945	0.5790	3.4222	3.3333	0.4969	0.4839	3.4213	3.3324	0.7524	0.7207	1.0000	
France	2	0.9670	0.9419	0.5893	0.5740	3.4595	3.3696	0.4920	0.4792	3.4579	3.3680	0.7653	0.7329	1.0170	
	3	0.9648	0.9085	0.5892	0.5739	3.4572	3.3673	0.4916	0.4788	3.4568	3.3669	0.7633	0.6711	0.9311	−0.6789
	1	0.9097	0.8860	0.5889	0.5736	3.3998	3.3114	0.4882	0.4755	3.3991	3.3107	0.7183	0.6879	1.0000	
Germany	2	1.0588	1.0313	0.6252	0.6090	3.5869	3.4936	0.5377	0.5237	3.5875	3.4943	0.8753	0.8392	1.2200	
	3	1.0544	0.8752	0.6234	0.6071	3.5805	3.4874	0.5353	0.5214	3.5798	3.4867	0.8699	0.6339	0.9216	−0.7542
	1	0.9387	0.9143	0.5872	0.5719	3.4385	3.3491	0.4885	0.4758	3.4365	3.3471	0.7405	0.7091	1.0000	
Italy	2	1.0293	1.0026	0.6140	0.5981	3.5481	3.4559	0.5232	0.5096	3.5468	3.4546	0.8399	0.8049	1.1351	
	3	1.0192	0.8653	0.6102	0.5943	3.5342	3.4423	0.5188	0.5053	3.5333	3.4414	0.8280	0.6935	0.9779	−0.7578
	1	0.9317	0.9075	0.5856	0.5704	3.4430	3.3535	0.4858	0.4732	3.4410	3.3515	0.7332	0.7021	1.0000	
Spain	2	1.0675	1.0397	0.6288	0.6125	3.5998	3.5062	0.5420	0.5279	3.6000	3.5064	0.8860	0.8495	1.2100	
	3	1.0668	0.8961	0.6283	0.6120	3.5969	3.5034	0.5417	0.5277	3.5970	3.5034	0.8851	0.7486	1.0663	−0.8246
	1	0.9627	0.9377	0.5922	0.5768	3.4566	3.3667	0.4953	0.4825	3.4596	3.3697	0.7646	0.7323	1.0000	
UK	2	1.0556	1.0282	0.6240	0.6078	3.5832	3.4900	0.5361	0.5222	3.5817	3.4886	0.8715	0.8355	1.1409	
	3	0.8846	0.8616	0.5582	0.5437	3.3457	3.2587	0.4527	0.4409	3.3473	3.2603	0.6708	0.6419	0.8765	−0.5435

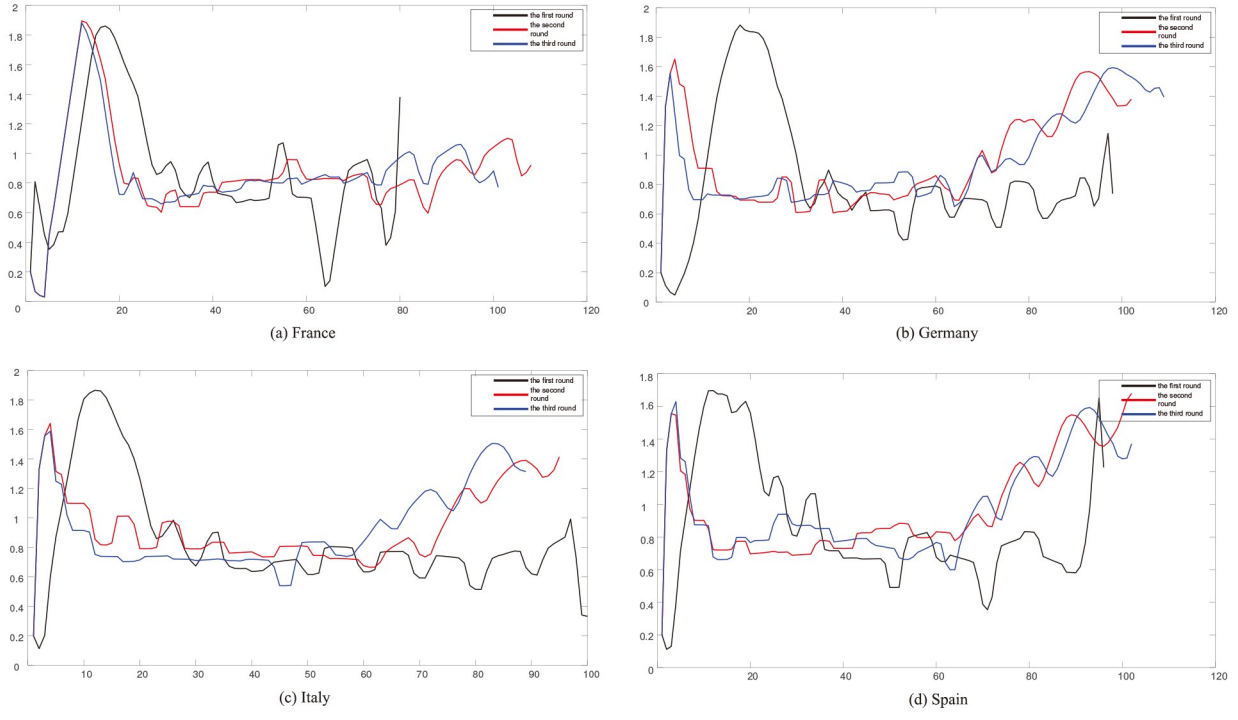


Figure 14. The comparison of R_e between three rounds. In M-SEIR, calculate the number of reproduction $R_e(t)$ in the three rounds by the equation $R_e(t)=\alpha\beta D+(1-\alpha)\mu\beta D$, and the results are compared. The above contains the situation in some countries.

When the experiment eliminates the factors of changes in movement, we observe a surprising result. The number of daily cases in the first round appears smaller than the one in the third round, and R_e owns the opposite result. We set R_e in the first round of outbreak as the base period and calculate its ratio to the other rounds. Compared with the third round, the value of ten countries is below one. External factors such as temperature, humidity can be approximately regarded as the same in two similar periods, so we speculate that the reduction of R_e is related to the level of personal immunity. After consulting the relevant information, we find that governments began to promote vaccination among the public in the third round (from early Jan. 2021 to May 2021). To verify this assumption, the downward trend of the infection rate ($\beta_0(t_i)$ means $\beta_0(t)$ in i-th round), which eliminates external factors, is compared with the vaccination rate in the corresponding period. The comparison shows that they are highly related.

5 Conclusions

To curb the rapid growth of the pandemic, governments have carried out a series of relevant NPIs to reduce the mobility and contacts in each environment. These policies involve work, schools, entertainment places, and other aspects. As shown in the first prevalent period, these actions proves effective in Europe. However, the control of crowd movement in the second and third rounds seems not to express a corresponding effect. Therefore, how to evaluate the effectiveness of mobility on the transmission rate exists many problems. We notice that the governments change the intensity of the contact by adjusting social mobility. So based on the classic SEIR model, we present a new epidemiological model for estimating the transmission rate of the pandemic. The M-SEIR model quantifies the transmission rate $\beta(t)$ influenced by a social movement. Next, adopt IF-EAKF to adjust the estimated value. After optimizing the model, it is confirmed to be effective in different spatial and temporal locations. Positive fitting results from European countries also tell the accuracy of the method. Prediction and discussion show the reference value of this model concerning the control of the infectious disease.

Nevertheless, it still has some limitation. Even though we consider the influence of age on transmission rate, it deserves to discuss in depth. European case frequently occurs in a resident who lives in long-term care facilities (LTCFs). The cumulative effect in a particular population requires us to stratify the patient by age and then compute them separately. Besides, some "superspreader" points should be responsible for a large majority of the infections. This aim will become the future direction of this research. The research keeps the

same interest in the role of vaccines in suppressing infection, which become another in-depth direction of the following examination.

Acknowledgments

We thank CHEN Mengyao, DAI Wei for their participation in discussions of the results. We are grateful to the reviewers for the careful reading of the manuscript and helpful remarks. This work is supported by the Double First-Class Initiative Research Fund (YD2040002004, YD2040002005), the Fundamental Research Funds for the Central Universities (WK2040000026, WK2040000027), and the National Natural Science Foundation of China (No. 12001517, No. 71873128).

Conflict of interest

The authors declare no conflict of interest.

Author information

LI Ying is currently a graduate student under the tutelage of Prof. JIN Baisuo at the University of Science and Technology of China. Her research focuses on the statistical analysis of epidemiological models and change point.

ZHANG Bo (corresponding author) is an Associate Professor at the University of Science and Technology of China (USTC). In 2017, he received his PhD degree from the Nanyang Technological University. From April 2017 to August 2019, he became a postdoctor at the Monash University. He joined the School of Management, USTC, in September 2019. Now he researches the high-dimensional random matrices, high-dimensional time series, network.

References

[1] Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 pandemic in China. Science,2020, 368(6491): 638-642.

[2] Chang S, Pierson E, Pang W K, et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature, 2020, 589(7840): 82-87.

[3] Kuhbandner C, Homburg S. Commentary: Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Frontiers in Medicine, 2020, 7: 580361.

[4] Desson Z, Lambertz L, Peters J W, et al. Europe's Covid-19 outliers: German, Austrian and Swiss policy responses during the early stages of the 2020 pandemic. Health Policy and Technology, 2020, 9(4):405-418.

[5] Kiesha P, Cook A R, Mark J, et al. Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLoS Computational Biology, 2017, 13(9):e1005697.

[6] Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Medicine, 2008, 5(3):e74.

[7] Glass L M, Glass R J. Social contact networks for the spread of pandemic influenza in children and teenagers. BMC Public Health, 2008, 8: Article number 61.

[8] Boehmer T K, Devies J, Caruso E, et al. Changing age distribution of the COVID-19 pandemic: United States, May-August 2020. Morbidity and Mortality Weekly Report (MMWR), 2020, 69(39): 1404-1409.

[9] Levin A T, Hanage W P, Owusu-Boaitey N, et al. Assessing the age specificity of infection fatality rates for COVID-19: Systematic review, meta-analysis, and public policy implications. European Journal of Epidemiology, 2020, 35(12): 1123-1138.

[10] Simon D. Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches. Hoboken, NJ: Wiley-Interscience, 2006.

[11] Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science, 2020, 368(6490): 489-493.

[12] Ionides E L, Breto C, King A A. Inference for nonlinear dynamical systems. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103(49): 18438-18443.

[13] Anderson J L. An ensemble adjustment Kalman filter for data assimilation. Monthly Weather Review, 2001, 129(12): 2884-2903.

[14] Qi H, Xiao S, Shi R, et al. COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. Science of the Total Environment, 2020, 728: 138778.

[15] Kodera S, Rashed E A, Hirata A. Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature and absolute humidity. International Journal of Environmental Research and Public Health, 2020, 17(15): 5477.

[16] Kissler S M, Tedijanto C, Goldstein E, et al. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science, 2020, 368(6493): 860-868.

(Continued on p. 716)

Table A2. Regression analyse on risk-sharing.

	(1)	(2)
	ΔlnPrice	
	Shanghai market	Hong Kong market
Connect	0.000(1.142)	0.001(1.578)
Difcov	0.047(1.171)	0.011(0.454)
Connect × Difcov	0.095*(1.818)	−0.105(−1.371)
Turnover	0.000(1.036)	−0.001***(−6.273)
Amihud	−0.000**(−2.056)	0.000(0.874)
Logsize	0.001***(7.724)	0.000(1.117)
ROA	−0.000(−0.848)	0.000(0.077)
Leverage	−0.000*(−1.866)	0.000(0.239)
Industry fixed effect	YES	YES
Constant	−0.017***(−7.046)	−0.004**(−2.339)
N	758	923
adj. R^2	0.375	0.162

[Note] This table reports the regression analysis on risk-sharing following the method of Chan and Kwok[11]. ΔlnPrice is the averagedaily log return after the SHSC program. Difcov presents the difference of stock return's covariance with local market return minus its covariance with other market returns. Fllowing Chan and Kwok[11], stock's log return is used to compute Difcov. *, ** and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

(Continued from p. 670)

在不同流动控制下欧洲 COVID-19 疫情的传播率

李影[1],孙天一[1],金百锁[2],张博[2*]

1. 中国科学技术大学大数据学院,安徽合肥 230026;
2. 中国科学技术大学管理学院,安徽合肥 230026

*通讯作者. E-mail:wbchpmp@ustc.edu.cn

摘要:在 2020 年,COVID-19 疫情引起全世界的关注,政府宣布了一系列非药物干预措施去遏制社会活动对传播的影响.各国不同力度的政策带来了相异的结果.为了评估这些行动的有效性,量化移动效应成为了关键问题.改变人群活动后,传播率是变化的且难以计算这种变化.因此,本文以一些欧洲国家为研究对象,收集各个国家在一些时期的人群移动情况以及每日的新增数据,并提出了流动-易感-暴露-感染-恢复(M-SEIR)模型。与 SEIR 模型不同,M-SEIR 模型中加入了一个量化控制措施影响的变量 $\sigma(t)$.采用随机抽样得到初始不同状态的人群数,对模型进行迭代.使用迭代-集成卡尔曼滤波技术(IF-EAKF)对后续的迭代结果进行调整,最后得到参数的变化趋势以及每日新增的估计值.在拟合部分,设置第一轮爆发为实验期,重复 100 次.它的拟合结果证实了模型的可行性和稳健性.此外,这项研究对受第二轮大流行影响的欧洲国家做出了合理的预测.通过调控政策的力度以及生效时间点,本文预测了非药物措施对流行病的影响,这为未来相关政策的部署提供了参考.最后,剔除人群移动、气温等外部因素后,研究得到了一个有趣的发现:尽管第三轮的每日报告远高于第一轮,但是第三轮的病毒传播参数要低于第一轮,进一步考察发现该下降与疫苗接种相关.

关键词:M-SEIR 模型;移动性;接触矩阵;传播率