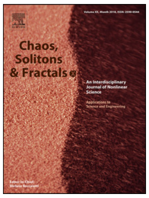
Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos



Maximum likelihood-based extended Kalman filter for COVID-19 prediction

Jialu Song[a], Hujin Xie[a,*], Bingbing Gao[b], Yongmin Zhong[a], Chengfan Gu[c], Kup-Sze Choi[c]

[a] School of Engineering, RMIT University, Melbourne, VIC 3000, Australia
[b] School of Automatics, Northwestern Polytechnical University, China
[c] Centre for Smart Health, School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Article history:
Received 2 February 2021
Revised 6 March 2021
Accepted 22 March 2021
Available online 2 April 2021

Keywords:
COVID-19 modelling
Extended Kalman filter
SEIRD model
Maximum likelihood estimation
Time-dependent model parameters

ABSTRACT

Prediction of COVID-19 spread plays a significant role in the epidemiology study and government battles against the epidemic. However, the existing studies on COVID-19 prediction are dominated by constant model parameters, unable to reflect the actual situation of COVID-19 spread. This paper presents a new method for dynamic prediction of COVID-19 spread by considering time-dependent model parameters. This method discretises the susceptible-exposed-infected-recovered-dead (SEIRD) epidemiological model in time domain to construct the nonlinear state-space equation for dynamic estimation of COVID-19 spread. A maximum likelihood estimation theory is established to online estimate time-dependent model parameters. Subsequently, an extended Kalman filter is developed to estimate dynamic COVID-19 spread based on the online estimated model parameters. The proposed method is applied to simulate and analyse the COVID-19 pandemics in China and the United States based on daily reported cases, demonstrating its efficacy in modelling and prediction of COVID-19 spread.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

The COVID-19 disease was broken out in Wuhan, China in the end of 2019. In spite of a huge amount of control efforts by the Chinese government, this disease was evolved into an epidemic within a few months, affecting almost every country in the world [1–3]. On the 31st January 2020, the World Health Organization declared the COVID-19 pandemic as a Public Health Emergency of International Concern. As of 27th January 2021, the virus has infected more than 100 million individuals, caused more than 2 million deaths [4], and altered the life of billions of people [5].

A huge amount of efforts has been dedicated to battling the COVID-19 disease, especially by the affected countries. However, this battle is far from over, with new infections detected every day. The forecast plays a significant role in control of the COVID-19 epidemic to provide early warnings, monitor the virus spread, evaluate the effects of virus containment measures, predict the future trend, and allocate the resources to counteract the pandemic. Thus, it is imperative to develop epidemiological modelling to predict and analyse the evolving trend of the COVID-19 pandemic.

Research endeavours have been devoted to prediction and forecast of COVID-19 spread, leading to various epidemiological models for characterisation of the COVID-19 transmission process. The Susceptible-Infected-Removed (SIR) model is commonly used for epidemiological modelling, where people who are susceptible to infection will be easily infected, people who are infected will either be cured or die, the susceptible population will gradually decrease over time, and the infectious disease will disappear eventually [6]. The removed (R) in the SIR model includes both individuals who are dead due to the epidemic and who are recovered. The Susceptible-Infected-Removed-Dead (SIRD) model improves the SIR model by dividing the removed (R) into the recovered (R) and dead (D) [7]. Most of the diseases have an incubation period between 5 to 14 days, during which the infected population is asymptomatically increased. However, the SIR model does not consider the population exposed to an epidemic. The Susceptible-Exposed-Infected-Recovered (SEIR) model adds the exposed population in the SIR model to take into account the effect of the incubation period of an infectious disease [8]. Similar to the SIRD model, the susceptible-exposed-infected-recovered-dead (SEIRD) model improves the SEIR model by considering the dead population [9].

Different from other infectious diseases, the COVID-19 virus has a strong characteristic of dynamic propagation, i.e., its epidemiological process strongly evolves with time. In order to capture this characteristic and reflect the actual behaviours of dynamic COVID-19 spread, the model parameters such as the infection, death and recovery rates as well as basic reproduction number must be time-

[*] Corresponding author.
E-mail address: s3463392@student.rmit.edu.au (H. Xie).

https://doi.org/10.1016/j.chaos.2021.110922
0960-0779/© 2021 Elsevier Ltd. All rights reserved.

dependent in COVID-19 modelling. However, the existing studies on COVID-19 modelling are dominated by fixed model parameters [1,2,10], while those based on time-dependent model parameters are still limited and mainly focused on the SIR and SEIR models [11].

In addition to an epidemiological model, a real-time estimation method is also required for prediction and analysis of the dynamic transmission process of COVID-19. The recursive least square (RLS) is the commonly used method for online parameter estimation in epidemic modelling [12]. It conducts online parameter estimation by minimizing a linear least-square cost function related to system observations. Chen et al. used the ridge regression, which is an improved RLS method to track the infection rate and recovery rate based on the SIR model [11]. As an extension of RLS, the Kalman filter (KF) uses linear Gaussian system state equations to optimally estimate system state through system observations. In addition to the update process of estimated system state with observations, which is similar to RLS, KF also involves a prediction process to predict the dynamic evolution of system state. Comparing to RLS, KF can achieve the estimation in the accuracy of minimum mean-square error even without using observations. Vaid et al. studied the COVID-19 pandemics in the United States (US), Canada and Sweden using KF based on Markov chain Monte Carlo (MCMC) sampling of the SIR model [13]. Zeng and Ghanem also developed a KF by switching between different linear Gaussian models based on observation data for modelling of the COVID-19 pandemic in the United States (US) [14]. In general, similar to RLS, KF can be applied to linear systems only, while the existing epidemiological models for COVID-19 prediction are nonlinear.

The nonlinear RLS is an extension of RLS to nonlinear systems. Piccolomini et al. used the nonlinear RLS to predetermine the SEIRD model parameters for COVID-19 [15]. However, due to the involvement of expensive computations, the nonlinear RLS can only be conducted offline, unsuitable for online determination of model parameters for COVID-19 forecast.

Ensemble KF (EnKF) extends the traditional KF to nonlinear systems by approximating the distribution of system state using a random ensemble of system state to calculate the state error covariance from ensemble members. Nkwayep et al. developed an EnKF based on the SIR model for prediction of the COVID-19 pandemic in Cameroon [16]. However, the size of ensemble members is critical for epidemic modelling. The use of a small ensemble size will lead to long-range spurious correlations in the error covariance and further the filtering divergence, while the use of a large ensemble size will lead to expensive computations [17]. In addition, EnKF cannot handle sharp coherent features such as the travelling waves found in epidemics [10]. The ensemble adjustment KF (EAKF) improves EnKF by adjusting each ensemble member towards the ensemble mean to prevent the filtering divergence [18]. However, this improvement is achieved at the cost of computational efficiency. Further, EAKF also suffers from the problem of expensive computations in case of large ensemble size [19].

The particle filter (PF) conducts nonlinear state estimation by using a sequence of independent random samples distributed according to certain conditional probability distributions to approximate the system state. Calvetti et al developed a PF based on the SEIR model for dynamic estimation of the reproduction number of the COVID-19 epidemic in US [20]. However, PF suffers from the particle degeneracy phenomenon and its accuracy largely depends on the choice of the importance sampling density function and resampling scheme. Despite various improvements such as the use of MCMC sampling [21] and Metropolis-Hastings (M-H) rules [22] for improvement of resampling, PF still suffers from expensive computations in the case of large sample size.

The extended Kalman filter (EKF) is an extension of the traditional KF to nonlinear systems by constructing the slope of the nonlinear system model in the mean. Although EKF has been used to predict and analyse various epidemic diseases [23,24], these studies mainly consider fixed model parameters, while the research on using EKF for modelling of the COVID-19 disease by considering time-dependent epidemiological model parameters is still limited. Just recently, Younes and Hasan also developed an EKF based on the Lotka–Volterra model to online estimate the dynamic behaviours of COVID-19 spread [25]. However, since the state prediction is achieved based on historical data, this method is unable to reflect the impact of external interventions on COVID-19 spread. Further, it still considers fixed rather than time-dependent model parameters for COVID-19 modelling.

This paper presents a novel method for prediction and analysis of dynamic COVID-19 spread based on the SEIRD epidemiological model with time-dependent model parameters. By discretising the SEIRD epidemiological model in time domain, the nonlinear state-space equation for describing COVID-19 dynamics is constructed to simultaneously estimate the time-dependent model parameters and transmission state. The time-varying model parameters are estimated according to the maximum likelihood principle to account for the dynamic effects of the infection, death and recovery rates on COVID-19 spread. Subsequently, an EKF is developed to dynamically estimate the transmission state based on the online estimated model parameters for COVID-19 forecast. Simulations and analysis on the COVID-19 pandemics in China and US have been conducted based on daily reported cases to evaluate the performance of the proposed method for COVID-19 modelling.

2. SEIRD model

The SIR model is commonly used to estimate emerging infectious diseases such as the COVID-19 disease. It is defined as

$$\begin{aligned}\frac{dS}{dt}&=-\frac{\beta SI}{N}\\ \frac{dI}{dt}&=\frac{\beta SI}{N}-\gamma I\\ \frac{dR}{dt}&=\gamma I\end{aligned}\tag{1}$$

where S, I and R denote the susceptible, infected and recovered numbers, N denotes the total population and $N=S+I+R$ under the assumption that the birth rate is equal to the death rate, $-\beta SI/N$ denotes the rate of decrease in susceptible individuals, $\beta SI/N-\gamma I$ represents the growth rate of infected individuals, γI represents the growth rate of recovered individuals, β is the infection rate, and γ is the recovery rate.

The COVID-19 disease has an asymptomatic infection trend, which cannot be described by the SIR model. It also has an incubation period of 14 days, where healthy people who have been in contact with the infected but without getting sick immediately will become virus carriers. These people are classified by the exposed cohort (E), which is in relation to the incubation rate α. To account for the exposed (E), the SEIR model adds the following equation in the SIR model

$$\frac{dE}{dt}=-\alpha E+\frac{\beta SI}{N}\tag{2}$$

Further, the SEIRD model adds the death component (D) with the death rate μ in the SEIR model to describe the disease spread dynamics

$$\begin{aligned}\frac{dS}{dt}&=-\frac{\beta SI}{N}\\ \frac{dE}{dt}&=-\alpha E+\frac{\beta SI}{N}\end{aligned}$$

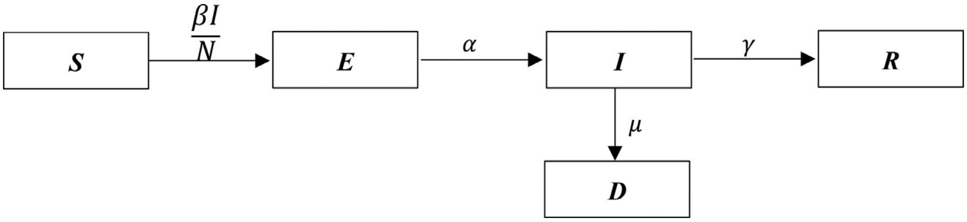


Fig. 1. The SEIRD model.

$$\frac{dI}{dt} = \alpha E - (\gamma + \mu)I \tag{3}$$

$$\frac{dR}{dt} = \gamma I$$

$$\frac{dD}{dt} = \mu I$$

Fig. 1 illustrates the structure of the SEIRD model. The susceptible cohort (S) will be infected to the exposed cohort (E) with the infection rate β. The exposed cohort (E) will then be transferred to the infected (I) with the incubation rate α. Finally, people who are infected will be cured with the recovery rate γ or die with the death rate μ.

Among the parameters of the SEIRD model, the incubation rate α is the inverse of the average latent time from showing symptoms to being infected. Since the average latent time from showing symptoms to being infected is a constant, the incubation rate α is commonly considered as a constant [7,12]. However, the infection rate β, recovery rate γ and death rate μ are not constant in the real-world situation. The contact with symptomatic infected persons is the main source to affect the infection rate. Other factors such as government control measures taken at various stages of the epidemic and aerosol transmissions also affect the infection rate. The recovery and death rates are related to both medical and population health levels in a country and how the government handles and controls the COVID-19 epidemic. Since they are variable in the dynamic environment of COVID-19 spread, these model parameters must be considered time-dependent in the modelling process. Subsequently, by discretizing the SEIRD model in time domain, the discrete system equation for COVID-19 spread is obtained as

$$S_{t+1} = S_t - \frac{\beta_t S_t I_t}{N} \tag{4}$$

$$E_{t+1} = E_t - \alpha E_t + \frac{\beta_t S_t I_t}{N} \tag{5}$$

$$I_{t+1} = I_t + \alpha E_t - (\gamma_t + \mu_t)I_t \tag{6}$$

$$R_{t+1} = R_t + \gamma_t I_t \tag{7}$$

$$D_{t+1} = D_t + \mu_t I_t \tag{8}$$

$$N = S_t + E_t + I_t + R_t + D_t \tag{9}$$

where β_t, γ_t and μ_t denote the infection rate, recovery rate and death rate at time point t.

The basic reproduction number R_0 is an important index in epidemiology and is calculated as

$$R_0 = \frac{S_t}{N}\left(\frac{\beta_t}{(\gamma_t + \mu_t)}\right) \tag{10}$$

If $R_0 < 1$, the infectious disease will gradually decrease over time. If $R_0 > 1$, the infectious disease will spread exponentially and become an epidemic. However, the epidemic will not last forever, because the population that may be infected will slowly decrease, some portion of the population may die from the infectious disease, and some may develop immunity after recovery. If $R_0 = 1$, the infectious disease will become endemic in the population. The larger R_0 is, the more difficult the control of the infectious disease will be.

3. Estimation algorithm

3.1. EKF

Consider the COVID-19 transmission process as a dynamic system. The system state vector is defined as

$$\boldsymbol{x}_t = \begin{bmatrix} S_t \\ E_t \\ I_t \\ R_t \\ D_t \end{bmatrix} \tag{11}$$

The system state equation is described as

$$\boldsymbol{x}_{t+1} = f\left(\boldsymbol{x}_t, \boldsymbol{\theta}_t\right) + \boldsymbol{\omega}_{t+1} \tag{12}$$

where $\boldsymbol{\theta}_t = [\beta_t \ \gamma_t \ \mu_t]$ collects the time dependent model parameters and $\boldsymbol{\omega}_{t+1}$ is the process noise at time t+1 which is assumed as a Gaussian white noise with zero mean and covariance $\mathbf{Q}$, i.e.,

$$\boldsymbol{\omega}_{t+1} \sim \ N(0, \mathbf{Q}) \tag{13}$$

$f(\boldsymbol{x}_t, \boldsymbol{\theta}_t)$ is the nonlinear system function formed from (4)-(8), i.e.,

$$f\left(\boldsymbol{x}_t, \boldsymbol{\theta}_t\right) \begin{bmatrix} S_t - \frac{\beta_t S_t I_t}{N} \\ E_t - \alpha E_t + \frac{\beta_t S_t I_t}{N} \\ I_t + \alpha E_t - (\gamma_t + \mu_t) I_t \\ R_t + \gamma_t I_t \\ D_t + \mu_t I_t \end{bmatrix} \tag{14}$$

Linearizing the nonlinear system function using the first-order Taylor expansion yields the Jacobian matrix $\boldsymbol{F}_t$

$$\frac{\partial f\left(\boldsymbol{x}_t, \boldsymbol{\theta}_t\right)}{\partial \boldsymbol{x}} |_{\boldsymbol{x}_t = \boldsymbol{x}_{t|t+1}} = \boldsymbol{F}_t \boldsymbol{x}_t \tag{15}$$

where

$$[\boldsymbol{F}_t] = \begin{bmatrix} 1 - \frac{\beta_t I_t}{N} & 0 & -\frac{\beta_t S_t}{N} & 0 & 0 \\ -\frac{\beta_t I_t}{N} & 1 - \alpha & \frac{\beta_t S_t}{N} & 0 & 0 \\ 0 & \alpha & 1 - (\gamma_t + \mu_t) & 0 & 0 \\ 0 & 0 & \gamma_t & 1 & 0 \\ 0 & 0 & \mu_t & 0 & 1 \end{bmatrix} \tag{16}$$

The observation equation is expressed as

$$\boldsymbol{y}_{t+1} = \boldsymbol{H}\boldsymbol{x}_{t+1} + \boldsymbol{v}_{t+1} \tag{17}$$

where $\boldsymbol{H}$ is the observation function to define the relation between the state vector $\boldsymbol{x}_{t+1}$ and measurement vector $\boldsymbol{y}_{t+1}$, and $\boldsymbol{v}_{t+1}$ is the observation noise which is assumed as a zero-mean Gaussian white noise of covariance $\boldsymbol{R}$, i.e.,

$$\boldsymbol{\upsilon}_{t+1} \sim N(0, \boldsymbol{R}) \tag{18}$$

The classic EKF procedure can be described as the following steps:

(i) Set the initial estimation state vector of $\hat{\boldsymbol{x}}_0$ and error covariance $\boldsymbol{P}_0$

$$\hat{\boldsymbol{x}}_0 = E[\boldsymbol{x}_0] \tag{19}$$

$$\boldsymbol{P}_0 = E\left[\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0)\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right)^{\mathrm{T}}\right] \tag{20}$$

(ii) State prediction

$$\hat{\boldsymbol{x}}_{t+1}^{-} = f\left(\hat{\boldsymbol{x}}_t, \boldsymbol{\theta}_t\right) + \boldsymbol{\omega}_{t+1} \tag{21}$$

$$\boldsymbol{P}_{t+1}^{-} = \boldsymbol{F}_t \boldsymbol{P}_t \boldsymbol{F}_t^{\mathrm{T}} + \boldsymbol{Q} \tag{22}$$

(iii) Measurement update

$$\boldsymbol{K}_{t+1} = \boldsymbol{P}_{t+1}^{-} \boldsymbol{H}^{\mathrm{T}} \left(\boldsymbol{H} \boldsymbol{P}_{t+1}^{-} \boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}\ \right)^{-1} \tag{23}$$

$$\hat{\boldsymbol{x}}_{t+1} = \hat{\boldsymbol{x}}_{t+1}^{-} + \boldsymbol{K}_{t+1}\left(\boldsymbol{y}_{t+1} - \boldsymbol{H}\hat{\boldsymbol{x}}_{t+1}^{-}\right) \tag{24}$$

$$\boldsymbol{P}_{t+1} = \boldsymbol{P}_{t+1}^{-} - \boldsymbol{K}_{t+1} \boldsymbol{H} \boldsymbol{P}_{t+1}^{-} \tag{25}$$

(iv) Repeat (ii) and (iii) for the next time step until all samples are processed.

3.2. Model parameter estimation

The infection rate β_t, death rate μ_t and recovery rate γ_t are the time-varying model parameters, which need to be estimated during the modelling process. The maximum-likelihood estimation is a statistical method to estimate parameters by maximizing their posteriori probability densities based on the deep-rooted Bayesian formalism [26,27]. It simplifies the problem of parameter estimation as a problem of maximizing a log-likelihood function. In this paper, we will adopt the maximum-likelihood method to estimate the model parameters from the available measurement data $\boldsymbol{y}_{1:k} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k)$, where $k \geq 1$ is an integer.

Define the innovation vector $\tilde{\boldsymbol{y}}_j$ as

$$\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \boldsymbol{H}\hat{\boldsymbol{x}}_j^{-} \quad (j = 1, 2, \cdots, k) \tag{26}$$

For the nonlinear Gaussian system described by (21)-(24), we have

$$\tilde{\boldsymbol{y}}_j \sim N\left(0, \boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}\right) \tag{27}$$

where

$$\boldsymbol{P}_{\tilde{\boldsymbol{y}}_j} = E\left[\tilde{\boldsymbol{y}}_j \tilde{\boldsymbol{y}}_{\boldsymbol{j}}^{\mathrm{T}}\right] = \boldsymbol{H}\boldsymbol{P}_j^{-}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R} \tag{28}$$

Assuming the available measurements are independent each other, according to the multiplication theorem of conditional probability, we can obtain

$$p\left(\boldsymbol{y}_{1:k}\boldsymbol{\theta}_t\right) \approx \prod_{j=1}^{k} \frac{1}{\sqrt{(2\pi)^m \left|\boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}\right|}} e^{-\frac{1}{2}\tilde{\boldsymbol{y}}_{\boldsymbol{j}}^{\mathrm{T}} \boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}^{-1} \tilde{\boldsymbol{y}}_j} \tag{29}$$

where m is the dimension of the measurement vector.

Taking the logarithm of (29) yields

$$\log\left[p\left(\boldsymbol{y}_{1:k}\boldsymbol{\theta}_t\right)\right] = -\frac{1}{2}\left\{m * \log(2\pi) + \sum_{j=1}^{k}\left(\tilde{\boldsymbol{y}}_{\boldsymbol{j}}^{\mathrm{T}} \boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}^{-1} \tilde{\boldsymbol{y}}_j + \log\left(\left|\boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}\right|\right)\ \right)\right\} \tag{30}$$

By ignoring the constant term, we define the cost function as

$$\boldsymbol{J}\left(\boldsymbol{\theta}_t\right) = \sum_{j=1}^{k}\left[\tilde{\boldsymbol{y}}_{\boldsymbol{j}}^{\mathrm{T}} \boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}^{-1} \tilde{\boldsymbol{y}}_j + \log\left(\left|\boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}\right|\right)\ \right] \tag{31}$$

such that the maximum likelihood estimates $\hat{\boldsymbol{\theta}}_t$ of parameter $\boldsymbol{\theta}_t$ satisfies the following condition

$$\hat{\boldsymbol{\theta}}_t = \arg\min_{\boldsymbol{\theta}}\left\{\sum_{j=1}^{k}\left[\tilde{\boldsymbol{y}}_{\boldsymbol{j}}^{\mathrm{T}} \boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}^{-1} \tilde{\boldsymbol{y}}_j + \log\left(\left|\boldsymbol{P}_{\tilde{\boldsymbol{y}}_j}\right|\right)\right]\right\} \tag{32}$$

It can be seen from (32) the problem of estimating parameter $\boldsymbol{\theta}_t$ is converted to a minimization problem. Solving (32) is equivalent to solving

$$\frac{\partial \boldsymbol{J}\left(\boldsymbol{\theta}_t\right)}{\partial\left(\boldsymbol{\theta}_t^l\right)} = 0 \tag{33}$$

where l=1,2,3 is the index corresponding to β_t, γ_t and μ_t respectively.

Let $\boldsymbol{J}^{(1)}(\boldsymbol{\theta}_t)$ represent the 3 × 1 vector of partials as follow

$$\boldsymbol{J}^{(1)}\left(\boldsymbol{\theta}_t\right) = \left[\frac{\partial \boldsymbol{J}\left(\boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t^1}, \frac{\partial \boldsymbol{J}\left(\boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t^2}, \frac{\partial \boldsymbol{J}\left(\boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t^3}\right]^{\mathrm{T}} \tag{34}$$

In order to calculate $\boldsymbol{J}^{(1)}(\hat{\boldsymbol{\theta}}_t) = 0$, we denote $\boldsymbol{J}^{(2)}(\boldsymbol{\theta}_t)$ as the 3 × 3 matrix of second-order partials, where

$$\boldsymbol{J}^{(2)}\left(\boldsymbol{\theta}_t\right) = \left[\frac{\partial \boldsymbol{J}^2\left(\boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t^l \partial \boldsymbol{\theta}_t^n}\right]_{l,n=1,2,3}^{\mathrm{T}} \tag{35}$$

Assuming that $\boldsymbol{J}^{(2)}(\boldsymbol{\theta}_t)$ is nonsingular and denoting $\boldsymbol{\theta}_t^{(i)}$ be the estimate of $\boldsymbol{\theta}_t$ at the ith iteration, we can approximate $\boldsymbol{J}^{(1)}(\boldsymbol{\theta}_t^{(i+1)})$ using the Taylor expansion

$$\boldsymbol{J}^{(1)}\left(\boldsymbol{\theta}_t^{(i+1)}\right) \approx \boldsymbol{J}^{(1)}\left(\boldsymbol{\theta}_t^{(i)}\right) - \boldsymbol{J}^{(2)}\left(\boldsymbol{\theta}_t^{(i)}\right)\left[\boldsymbol{\theta}_t^{(i+1)} - \boldsymbol{\theta}_t^{(i)}\right] \tag{36}$$

Letting the right-hand side of (36) be equal to zero and solving for $\boldsymbol{\theta}_t^{(i+1)}$, it is obtained

$$\boldsymbol{\theta}_t^{(i+1)} = \boldsymbol{\theta}_t^{(i)} + \left[\boldsymbol{J}^{(2)}\left(\hat{\boldsymbol{\theta}}_t\right)\right]^{-1} \boldsymbol{J}^{(1)}\left(\boldsymbol{\theta}_t^{(i)}\right) \tag{37}$$

Iterating (37) until the process is converged, we will readily have the maximum likelihood estimation $\hat{\boldsymbol{\theta}}_t$. Subsequently, the obtained maximum likelihood estimation $\hat{\boldsymbol{\theta}}_t$ will be fed back to the EKF filtering process to calculate the soft tissue deformation state. Fig. 2 illustrates the framework of the maximum likelihood-based EKF.

4. Performance evaluation

Simulation analyses were conducted based on two actual cases to comprehensively evaluate the performance of the proposed EKF for COVID-19 modelling. One is the COVID-19 pandemic in Wuhan, China, which is the first outbreak in the world. The other is the COVID-19 pandemic in US, which has not yet reached the peak and involves the largest number of confirmed cases. Since the true values for the actual cases are unknown, the reported data were taken as reference for calculation of estimation error. The initial state, error covariance and model parameters were set based on the observation data on the first day of each simulation analysis. The transmission states of COVID-19 estimated by EKF were also compared with those numerical solutions calculated from the discrete SEIRD model (4)-(8) based on parameter identification via the offline constrained least-squares algorithm [15,34].

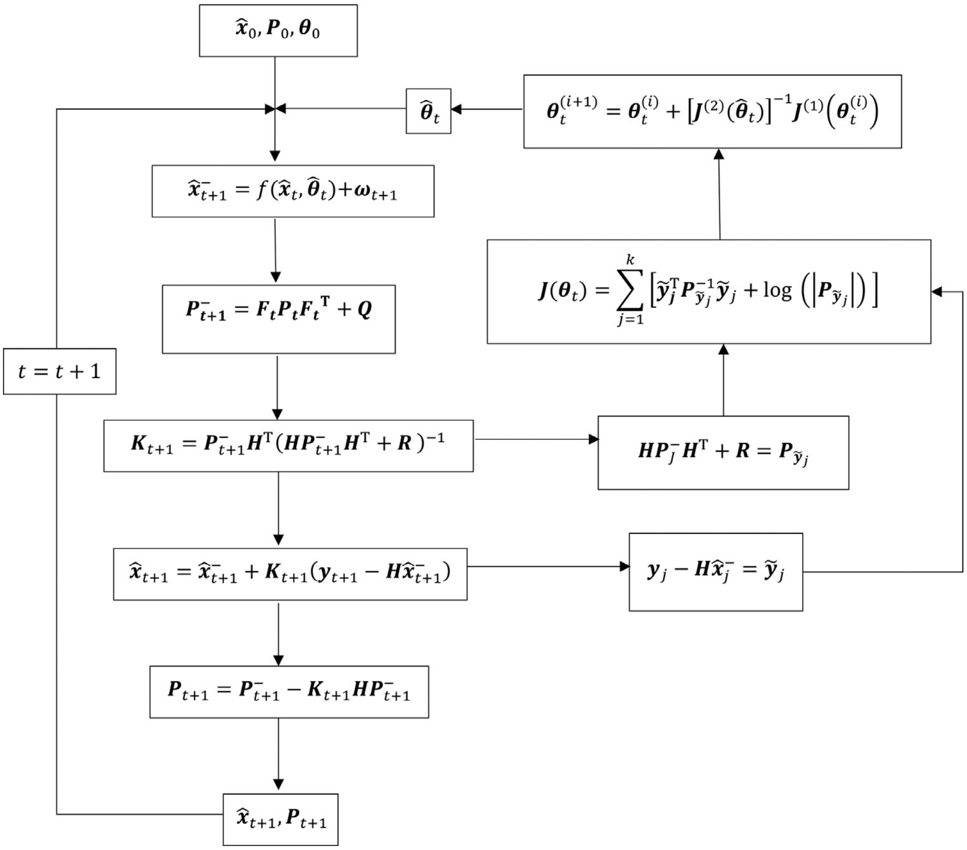

$\hat{x}_0, P_0, \theta_0$
$\hat{\theta}_t$
$\theta_t^{(i+1)} = \theta_t^{(i)} + \left[J^{(2)}(\hat{\theta}_t)\right]^{-1} J^{(1)}\left(\theta_t^{(i)}\right)$
$\hat{x}_{t+1}^{-} = f(\hat{x}_t, \hat{\theta}_t) + \omega_{t+1}$
$J(\theta_t) = \sum_{j=1}^{k} \left[\tilde{y}_j^{T} P_{\tilde{y}_j}^{-1} \tilde{y}_j + \log\left(\left|P_{\tilde{y}_j}\right|\right)\right]$
$P_{t+1}^{-} = F_t P_t F_t^{T} + Q$
$t = t + 1$
$K_{t+1} = P_{t+1}^{-} H^{T} (H P_{t+1}^{-} H^{T} + R)^{-1}$
$H P_j^{-} H^{T} + R = P_{\tilde{y}_j}$
$\hat{x}_{t+1} = \hat{x}_{t+1}^{-} + K_{t+1}(y_{t+1} - H\hat{x}_{t+1}^{-})$
$y_j - H\hat{x}_j^{-} = \tilde{y}_j$
$P_{t+1} = P_{t+1}^{-} - K_{t+1} H P_{t+1}^{-}$
$\hat{x}_{t+1}, P_{t+1}$

Fig. 2. The proposed EKF algorithm based on model parameter estimation.

The mean error and RMSE (Root mean square error) are used to measure the estimation accuracy. They are defined as

$$Mean\ error = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{x}_i - x_i^{ref}\right) \tag{38}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\hat{x}_i - x_i^{ref}\right)^2} \tag{39}$$

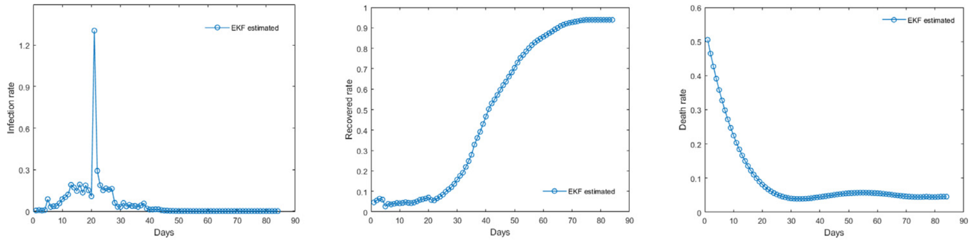
4.1. COVID-19 spread in Wuhan of China

The COVID-19 was first broken out in Wuhan, China. No special intervention measures for COVID-19 were taken before 23rd January 2020. On 23rd January 2020, the Chinese government locked down the Wuhan city, suspended all public transportations and services in the city, and imposed quarantine to households in the entire city [28]. The Chinese government also adopted a series of strict measures, including travel bans, suspension of schools and extension of the Spring Festival holiday for the entire country to control the virus spread. These strong controls prohibited the virus spread and saved many lives at a huge economic cost. The number of daily new cases decreased to zero on 18th March 2020 in Wuhan. The epidemic in Wuhan almost vanished in the end of April 2020.

Simulation trials were conducted by focusing on the COVID-19 spread during the lockdown period of the Wuhan city from 23rd January (Day 0) to 15th April 2020 (Day 83), where the model parameters, i.e., the infection rate, recovery rate and death rate, were mainly affected by government control measures. The constant incubation rate α was set to 0.33 [29,30]. The time from 10th to 22nd January 2020 was just before the Chinese Spring Festival, during which large-scale population movements were occurred in the entire country. Before the city lockdown on 23rd January 2020, there were about 5 million people moved out of the Wuhan city. This large-scale population migration may cause the epidemic to spread widely in Wuhan and to other places in the country [31]. Therefore, the total population used in the simulation was 5 million less than the total population of Wuhan in 2019. The daily reported data on the COVID-19 spread in the Wuhan city [32] were taken as the observation data for the simulation analysis.

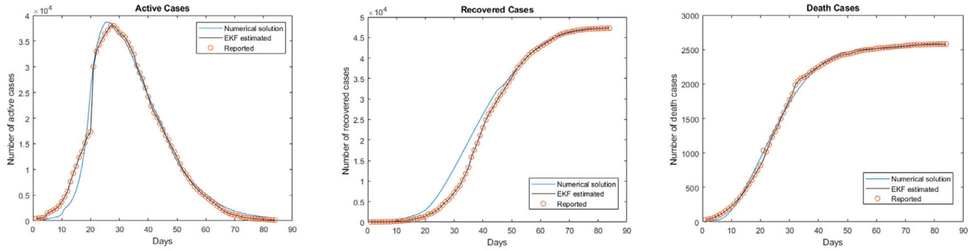
Fig. 4 illustrates the virus transmission state estimated by the proposed EKF. It can be seen that the estimated numbers of active, recovery and death cases are very close to their reported data. As shown in Fig. 4(a), the reported number of active cases suddenly increases to 13,436 on the 20th day of the lockdown, leading to the largest increase on one single day. This is because the medical diagnosis became more reliable, leading more symptomatic cases to be classified as confirmed cases [33]. The EKF estimation captures the largest single-day increase. As shown in Figs. 3(a) and 4(a), the estimated infection rate increases to the maximum and the estimated number of active cases also involves a large increase on that day (13th February 2020). After the peak of reported active cases, the reported number of active cases begins to drop sharply due to the lockdown effect. Both estimated infection rate and estimated active cases also involve such dramatic drops.

Figs. 3(b) and 4(b) show that both estimated recovery rate and recovery cases have a similar trend and the estimated recovery cases follows closely to the reported ones. Both estimated recovery rate and cases are small in the early stage of the COVID-19 pandemic, because there is always a lack of effective medical treatment for a new epidemic in the early stage. However, after the 30th day of the lockdown, with the improvement of medical treatment, the estimated recovery rate as well as the numbers of estimated and reported recovery cases greatly increase and eventually



(a) (b) (c)

Fig. 3. The model parameters estimated by the proposed EKF for the COVID-19 pandemic in Wuhan: (a) the infection rate; (b) the recovery rate; and (c) the death rate.



(a) (b) (c)

Fig. 4. The transmission states estimated by the proposed EKF and calculated from the discrete SEIRD model based on parameter identification via the constrained least-squares algorithm for the COVID-19 pandemic in Wuhan: (a) the number of active cases; (b) the number of recovery cases; and (c) the number of death cases.

Table 1
Statistical estimation errors of the proposed EKF and numerical solution for the COVID-19 pandemic in Wuhan.

State Variables	Mean error		RMSE	
	EKF	Numerical solution	EKF	Numerical solution
I	581.25	1062	242.43	1670
R	561.76	1489	813.97	2512
D	23.79	31.88	41.24	44.37

remain at a constant value. As shown in Fig. 3(c), the estimated death rate has an opposite trend with the estimated recovery rate. It is at maximum in the early stage and will then greatly decrease due to the improvement of medical treatment. After the 30th day of the lockdown, the estimated death rate involves small variations and eventually remains at a constant value. Correspondingly, as shown in Fig. 4(c), both estimated and reported numbers of death cases greatly increase before the 30th day of the lockdown. After that, both estimated and reported numbers of death cases gradually increase and eventually remain at a constant value.

On the 30th day of the lockdown, because the infected are either cured or died, the infection rate drops to the lowest within the first 30 days of the lockdown, and both estimated and reported active cases also begin to decline. This indicates the government control measures started taking into effect to control the epidemic. After the 70th day of the lockdown, the estimated number of active cases is gradually close to zero, while both estimated numbers of death and recovery cases are gradually close to a constant value, indicating that the COVID-19 epidemic in Wuhan is almost over.

Further, the transmission state was also calculated from the discrete SEIRD model (4)-(8) based on parameter identification via the offline constrained least-squares algorithm [15,34] for the COVID-19 spread in Wuhan, and was further compared with that estimated by EKF. As shown in Fig. 4, the transmission state estimated by EKF approximates the reported data more closely than the numerical solution. Table 1 lists the statistical errors of both EKF estimation and numerical solution, demonstrating that the EKF estimation has much higher accuracy than the numerical solution for the COVID-19 spread in Wuhan.

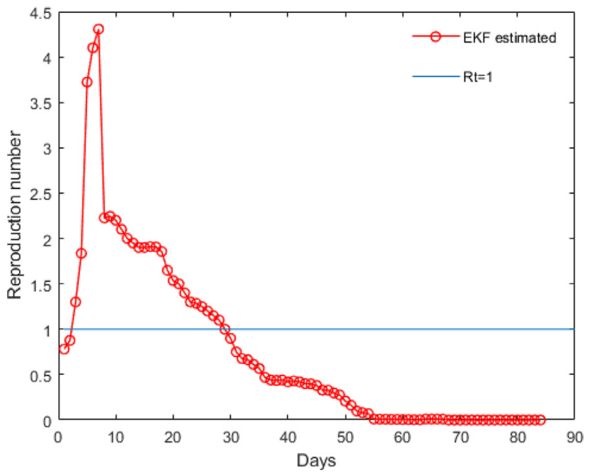
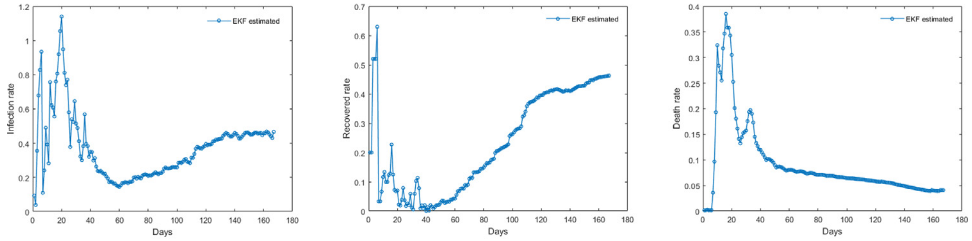


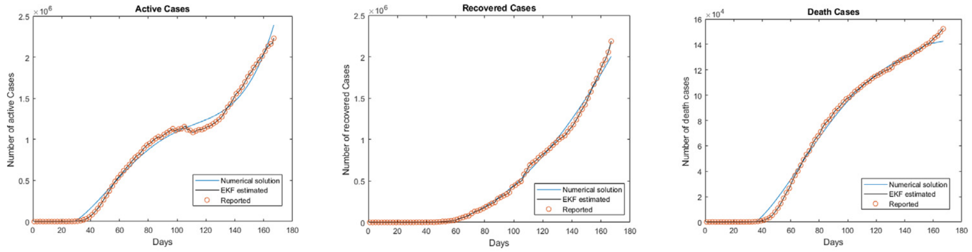
Fig. 5. The basic reproduction number estimated by the proposed EKF for the COVID-19 pandemic in Wuhan.

The performance of the proposed EKF was also evaluated in terms of the basic reproduction number R_0. Fig. 5 illustrates the development trend of R_0 for the COVID-19 spread in Wuhan, China during the lockdown period. The estimated R_0 increases from the initial 0.75 to the peak on the 8th day of the lockdown. After the peak, it drops rapidly and becomes smaller than 1 on the 30th day of the lockdown, indicating that the COVID-19 epidemic starts diminishing. This development trend is also reflected



(a) (b) (c)

Fig. 6. The model parameters estimated by the proposed EKF for the COVID-19 pandemic in US: (a) the infection rate; (b) the recovery rate; and (c) the death rate.



(a) (b) (c)

Fig. 7. The transmission states estimated by the proposed EKF and calculated from the discrete SEIRD model based on constrained least-square parameter identification for the COVID-19 pandemic in US: (a) the number of active cases; (b) the number of recovery cases; and (c) the number of death cases.

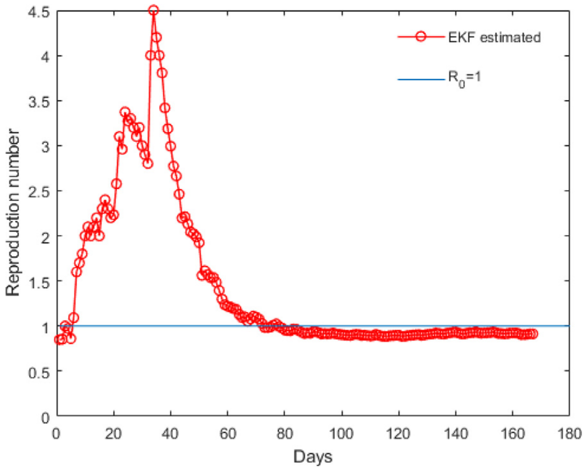


Fig. 8. The estimated basic reproduction number by the proposed EKF for the COVID-19 pandemic in US.

in Fig. 4(a), where both estimated and reported numbers of active cases rapidly decrease after the 30th day of the lockdown. After the 30th day of the lockdown, the estimated R_0 continues dropping until the 55th day of the lockdown, after which the estimated R_0 remains close to zero. This indicates that the COVID-19 epidemic has been contained and is almost over, due to the government control measures. It should be noted that the 55th day of the lockdown on which R_0 drops to zero for the first time also matches the observation data on that day where the number of daily new infected cases in Wuhan is zero for the first time.

The above results demonstrate that the proposed EKF can effectively predict the evolving trend of the COVID-19 spread. They also indicate that since the outbreak of the COVD-19 epidemic, the lockdown measures implemented by the Chinese government effectively controlled the pandemic spread. In other words, the government's early intervention measures played a decisive role in controlling the COVID-19 epidemic.

4.2. COVID-19 spread in US

The COVID-19 spread in US started in New York and California, and then severely spread in New York. On 10th March 2020, the number of confirmed cases in US was just about 700. However, on 26th March 2020, the number of confirmed cases was rapidly increased to over 80,000 and the number of death cases was also rapidly increased from 30 to about 1600. This shows that the epidemic was already in a state of community spread in US. As of 24th October, 2020, the virus has infected more than 8 million individuals and caused more than 223,000 deaths in US [4].

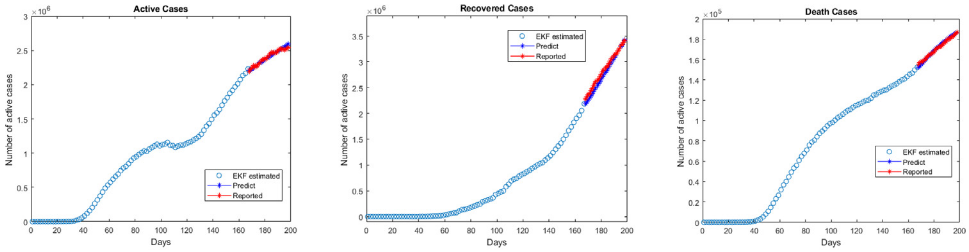
Simulation trials were conducted by focusing on the COVID-19 spread in US from 15th February to 30th July 2020 based on the daily reported data [4]. The constant incubation rate α was set to 0.2 [35]. Currently, the number of confirmed cases in US is still increasing without any sign to reach the peak. The development trend of the COVID-19 epidemic is still not clear. Fig. 5 shows the estimated model parameters. The estimated virus transmission state is shown in Fig. 6, and the associated estimation errors are listed in Table 2.

It can be seen from Fig. 6(a) that in the early stage of the COVID-19 epidemic, the estimated infection rate, recovery rate and death rate fluctuate greatly. This is because the detection mechanism in US in the early stage of the epidemic was not reliable. Not every suspected case with symptoms was tested, while only the people who went abroad or had contact with people who returned to US were tested. On 13th March 2020, the US government declared a state of emergency in response to the COVID-19 disease and began to implement control measures such as city lockdown and expanded testing for the COVID-19 disease. This also led to a sharp rise in the reported active cases. As shown in Figs. 6(a) and 7(a), from the 40th day of the simulation analysis, the estimated infection rate begins to increase. The estimated number of active cases closely follows the reported number and also begins to increase from that day, showing that the epidemic enters a period of outbreak growth.

As shown in Fig. 7(a), around the 100th day of the simulation analysis, both estimated and reported numbers of active cases involve small variations and even begin to decline. This indicates that

Table 2
Statistical estimation errors of the proposed EKF and numerical solution for the COVID-19 pandemic in US.

State Variables	Mean error		RMSE	
	EKF	Numerical solutions	EKF	Numerical solutions
I	33,200	41,857	38,010	45,273
R	13,715	32,110	20,907	36,407
D	1014	4451	1245	7110



(a) (b) (c)

Fig. 9. The 30-days prediction for the COVID-19 pandemic in US: (a) the number of active cases; (b) the number of recovery cases; and (c) the number of death cases.

despite the continuously increased active cases, the government control measures started taking into effect to control the COVID-19 epidemic. However, the US government began to reduce the level of control measures on 1st May 2020 for the economy recovery. This action makes the reported number of active cases begin to rise sharply after 10th June 2020 (i.e., 110th of the simulation analysis). As shown in Fig. 7(a), both reported and estimated numbers of active cases begin to increase after the 110th day of the simulation analysis. The above trend is also reflected in the estimated infection rate. As shown in Fig. 6(a), the estimated infection rate involves small variations around the 100th day and begins to increase after the 110th day.

As shown in Fig. 6(b), after the 60th day of the simulation analysis, the estimated recovery rate begins to increase due to the improvement of medical treatment. This trend is also reflected in both estimated and reported numbers of recovery cases. As shown in Fig. 7(b), both estimated and reported number of recovery cases begin to increase after the 60th day. As shown in Fig. 6(c), the estimated death rate has the highest value in the early stage and will then decrease after the 40th day due to the improvement of medical treatment. Figs. 7(b) and 5(c) show that the numbers of active and death cases are still increasing without reaching their peaks, implying the uncertainty of the future development of the COVID-19 epidemic in US. Similarly, the transmission state estimated by EKF was also compared with that calculated from the discrete SEIRD model (4)–(8) based on parameter identification via the offline constrained least-squares algorithm [15,34] for the COVID-19 spread in US. As shown in Fig. 7, the transmission state estimated by EKF approximates the reported data more closely than the numerical solution. Table 2 lists the statistical errors of both EKF estimation and numerical solution, demonstrating that the EKF estimation has much higher accuracy than the numerical solution for the COVID-19 spread in US.

The above results clearly indicate that in the absence of effective vaccine, if US government does not adopt more strict control measures to control the spread of the epidemic in the future, the number of confirmed cases will continue increase.

Fig. 8 illustrates the basic reproduction number R_0 estimated by the proposed EKF for the COVID-19 pandemic in US. The estimated R_0 reaches the peak on about the 40th day of the simulation analysis. This is also reflected in Fig. 7(a), where both estimated and reported numbers of active cases begin to rapidly increase on that day. After the peak, the estimated R_0 dramatically decreases and becomes smaller than 1 on the 80th of the simulation analysis, implying that the epidemic starts diminishing. After the 80th day, the estimated R_0 remains close to about 1, indicating that the spread of the COVID-19 epidemic has been contained to some extent but is still uncertain and may develop into an endemic epidemic, thus requiring the government to strengthen control measures to curb the spread of the COVID-19 epidemic.

Based on the control measures and reported data from 15th February to 30th July 2020 in US, simulation trials were also conducted to predict the COVID-19 spread for the next 30 days since 30th July 2020. As shown in Fig. 9, the active cases, recovery cases and death cases predicted by the proposed EKF are very close to the actually reported cases, respectively. It can also be seen from Fig. 7(a) that despite the continuous increase of the active cases without any sign to reach the peak at the end of the simulation day (30th July 2020), the growth of the active cases exhibits a slowing down trend.

5. Conclusions

This paper presents a novel method for prediction and analysis of dynamic COVID-19 spread. It converts the epidemiological modelling of COVID-19 into a problem of filtering identification to estimate the dynamic behaviours of COVID-19 spread. The nonlinear state-space equation is established by discretising the SEIRD epidemiological model in time domain. A maximum likelihood estimation theory is established to estimate the time-varying model parameters. Based on above, an EKF is developed for online prediction and analysis of dynamic COVID-19 behaviours based on the estimated model parameters. Simulation analyses demonstrate that the proposed method can track and predict the evolving trend of COVID-19 spread based on daily reported cases. The results are consistent with the success in controlling the COVID-19 epidemic in Wuhan, which was attributed to the early interventions on public health, and the difficulty experienced in the initial stage of the outbreak in US that prompts for the needs of effective control measures.

Future research work will focus on improving the proposed method by considering the errors involved in epidemiological modelling. It is expected that adaptive filtering algorithms will be developed to enable the proposed method to accommodate epidemiological modelling errors.

Declaration of Competing Interest

The authors declare no conflict of interest for this paper.

CRediT authorship contribution statement

Jialu Song: Conceptualization, Methodology, Software, Writing - original draft. **Hujin Xie:** Conceptualization, Methodology, Writing - original draft. **Bingbing Gao:** Methodology, Writing - review & editing. **Yongmin Zhong:** Investigation, Methodology, Writing - review & editing. **Chengfan Gu:** Writing - review & editing. **Kup-Sze Choi:** Writing - review & editing.

References

[1] Anastassopoulou C, Russo L, Tsakris A, Siettos C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PloS One 2020;15(3):e0230405.

[2] Kucharski AJ, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis; 2020.

[3] Yang Z, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. (in eng). J Thorac Dis 2020;12(3):165–74. doi:10.21037/jtd.2020.02.64.

[4] Coronavirus Resource Center, Johns Hopkins University, https://coronavirus.jhu.edu/

[5] Li R, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science 2020;368(6490):489–93.

[6] Zhao S, Chen H. Modeling the epidemic dynamics and control of COVID-19 outbreak in China. Quant Biol 2020:1–9.

[7] Fanelli D, Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos, Solitons Fractals 2020;134:109761.

[8] He S, Peng Y, Sun K. SEIR modeling of the COVID-19 and its dynamics. Nonlinear Dyn 2020:1–14.

[9] Rajagopal K, Hasanzadeh N, Parastesh F, Hamarash II, Jafari S, Hussain I. A fractional-order model for the novel coronavirus (COVID-19) outbreak. Nonlinear Dyn 2020;101(1):711–18.

[10] Mandel J, Beezley JD, Cobb L, Krishnamurthy A. Data driven computing by the morphing fast Fourier transform ensemble Kalman filter in epidemic spread simulations. Procedia Comput Sci 2010;1(1):1221–9.

[11] Chen Y-C, Lu P-E, Chang C-S, Liu T-H. A Time-dependent SIR model for COVID-19 with undetectable infected persons. IEEE Trans Netw Sci Eng 2020.

[12] Sameni R. Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus; 2020. arXiv preprint arXiv:2003.11371.

[13] Vaid S, McAdie A, Kremer R, Khanduja V, Bhandari M. Risk of a second wave of Covid-19 infections: using artificial intelligence to investigate stringency of physical distancing policies in North America. Int Orthop 2020.

[14] Zeng X, Ghanem R. Dynamics identification and forecasting of COVID-19 by switching Kalman filters. Comput Mech 2020;66(5):1179–93.

[15] Loli Piccolomini E, Zama F. Monitoring Italian COVID-19 spread by a forced SEIRD model. PloS One 2020;15(8):e0237417.

[16] Nkwayep CH, Bowong S, Tewa J, Kurths J. Short-term forecasts of the COVID-19 pandemic: study case of Cameroon. Chaos, Solitons Fractals 2020:110106.

[17] Bani Younes A, Hasan Z. COVID-19: modeling, prediction, and control. Appl Sci 2020;10(11) (2076-3417).

[18] Huang W, Provan G. An improved state filter algorithm for SIR epidemic forecasting. In: Proceedings of the twenty-second European conference on artificial intelligence; 2016. p. 524–32.

[19] Kang B, Yang H, Lee K, Choe J. Ensemble Kalman filter with principal component analysis assisted sampling for channelized reservoir characterization. J Energy Resour Technol 2017;139(3).

[20] Calvetti D, Hoover A, Rose J, Somersalo E. Bayesian dynamical estimation of the parameters of an SE (A) IR COVID-19 spread model; 2020. arXiv preprint arXiv:2005.04365.

[21] Endo A, van Leeuwen E, Baguelin M. Introduction to particle Markov-chain Monte Carlo for disease dynamics modellers. Epidemics 2019;29:100363.

[22] Yang W, Karspeck A, Shaman J. Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. PLoS Comput Biol 2014;10(4):e1003583.

[23] Cazelles B, Chau N. Adaptive dynamic modelling of HIV/AIDS epidemic using extended Kalman filter. J Biol Syst 1995;3(03):759–68.

[24] Ndanguza D, Mbalawata IS, Haario H, Tchuenche JM. Analysis of bias in an Ebola epidemic model by extended Kalman filter approach. Math Comput Simul 2017;142:113–29.

[25] Younes AB, Hasan Z. COVID-19: modeling, prediction, and control. Appl Sci 2020;10(11):3666.

[26] Gao B, Gao S, Hu G, Zhong Y, Gu C. Maximum likelihood principle and moving horizon estimation based adaptive unscented Kalman filter. Aerosp Sci Technol 2018;73:184–96.

[27] Gao Z, Mu D, Gao S, Zhong Y, Gu C. Adaptive unscented Kalman filter based on maximum posterior and random weighting. Aerosp Sci Technol 2017;71:12–24.

[28] Prem K, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health 2020;5(5):e261–70. doi:10.1016/s2468-2667(20)30073-6.

[29] Lin Q, et al. A conceptual model for the outbreak of Coronavirus disease 2019 (COVID-19) in Wuhan, China with individual reaction and governmental action. Int J Infect Dis 2020.

[30] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 2020;395(10225):689–97.

[31] Yang Q, et al. Short-term forecasts and long-term mitigation evaluations for the COVID-19 epidemic in Hubei Province, China; 2020. medRxiv.

[32] "Outbreak notification," Jan 2020. [Online]. Available: http://www.nhc.gov.cn/xcs/yqtb/list gzbd.shtml.

[33] Sun G-Q, et al. Transmission dynamics of COVID-19 in Wuhan, China: effects of lockdown and medical resources. Nonlinear Dyn 2020:1–13.

[34] Fonseca i Casas P, García i Carrasco V, Garcia i Subirana J. SEIRD COVID-19 formal characterization and model comparison validation. Appl Sci 2020;10(15):5162.

[35] Centers for Disease Control and Prevention, https://www.cdc.gov/.