Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Establishment of epidemic early warning index system and optimization of infectious disease model: Analysis on monitoring data of public health emergencies

Li Xiong, Peiyang Hu*, Houcai Wang

School of management, Shanghai University, Shanghai, China

Abstract

The ability to mitigate the damages caused by emergencies is an important symbol of the modernization of an emergency capability. When responding to emergencies, government agencies and decision makers need more information sources to estimate the possible evolution of the disaster in a more efficient manner. In this paper, an optimization model for predicting the dynamic evolution of COVID-19 is presented by combining the propagation algorithm of system dynamics with the warning indicators. By adding new parameters and taking the country as the research object, the epidemic situation in countries such as China, Japan, Korea, the United States and the United Kingdom was simulated and predicted, the impact of prevention and control measures such as effective contact coefficient on the epidemic situation was analyzed, and the effective contact coefficient of the country was analyzed. The paper strives to provide early warning of emergencies scientifically and effectively through the combination of these two technologies, and put forward feasible references for the implementation of various countermeasures. Judging from the conclusion, this study reaffirmed the importance of responding quickly to public health emergencies and formulating prevention and control policies to reduce population exposure and prevent the spread of the pandemic.

1. Introduction

Public health emergency is an unexpected emergency event in the risk society, which causes great damage to all aspects of social life. Different kinds of public emergencies are characterized by uncertainty, complexity, diffusion and cross-domain [1]. Among them, the most typical is a major infectious disease event. Infectious diseases have been a disease that severely threatens human life and health and seriously affects the development of social economy [2,3]. According to a COVID-19 real-time statistical data report by World Health Organization (WHO) in August 1st, 2021, revealed that the global cumulative numbers to 198.5 million cases and over 4.2 million deaths cases since the start of the pandemic. The number of deaths from the COVID-19 pandemic [4]. According to the characteristics of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), such as diversity of causes, fast transmission speed, wide range of influence, great harm and complex monitoring, prevention and control, it is a great challenge for emergency agencies, and also puts forward higher requirements for improving the ability and level of responding to public health emergencies [5]. It warns government agencies and researchers that they should explore and summarize this public health emergency in depth, and use more efficient methods to make practical management plans.

Information technology innovation is increasing globally during the COVID-19 pandemic [6]. Modern information technology marked by digitization and networking has gradually become the basis for the modernization and intelligence of public health emergencies [7]. It is required to strengthen the technical support of emergency management equipment, optimize and integrate all types of science and technology resources, and promote science and technology independent innovation in emergency management, relying on science and technology to improve the level of intelligence and perfect emergency management. To adapt to the trend of the development of "infodemic" progresses [8], it is indispensable to promote the modernization of emergency management with information method, and improve the abilities of monitoring and early warning, supervision and law enforcement, assist command and decision-making, rescue and actual combat, and social mobilization.

* Corresponding author. School of Management, Shanghai University, 99 Shangda Road, BaoShan District, Shanghai, PR China.
E-mail address: peiyang_hu@163.com (P. Hu).

https://doi.org/10.1016/j.ijdrr.2021.102547
Received 12 May 2021; Received in revised form 29 August 2021; Accepted 30 August 2021
Available online 2 September 2021
2212-4209/© 2021 Elsevier Ltd. All rights reserved.
At the time of a large outbreak, relevant decision-makers may need to develop more scientific and rational solutions based on the future evolution of the epidemic. For this reason, we choose Susceptible-Exposed-Infected-Recovered(SEIR) model, which is an infectious disease dynamic model, to simulate and predict the future evolution trend of the epidemic. The infectious disease dynamic model, represented by SEIR, considers the occurrence and spread of disease in the population and the related social environment. It can reflect the intrinsic dynamic characteristics of the spread of infectious diseases, better reproduce the development process of diseases, reveal the flow law, and predict the change trend. However, it also indicates that the SEIR model is strongly influenced by other factors such as the environment, so if the model is used directly to predict, the results may differ from the actual situation. In practical application, the more detailed the classification of parameters and features, the more accurate the trend prediction of the model. The causes of uncertain factors will lead to large fluctuations in the accuracy of prediction. So the conclusion that we come to is that factors related to the COVID-19 must be taken into account.

In order to improve the SEIR model and make it more in line with the current status of COVID-19, we mine the key information of the occurrence of public health emergencies, and establish a set of scientific, reasonable and sensitive index system [9]. Predict the situation trend of infectious diseases based on huge amounts of epidemic data and other data that have a serious impact on the pandemic. With the assistance of large data and artificial intelligence, it is vital to quantify the influencing factors of the epidemic situation and to establish the whole process evaluation index system on the basis of forming an effective epidemic data structure. It is crucial to combine the theoretical prediction algorithm with the actual early warning index [10].

According to the above research ideas, the rest of this paper is organized as follows: the second part describes the existing relevant research, which is the cornerstone of the follow-up research. The third and fourth parts respectively discussed formulate preventive indicator system and prediction model based on SEIR separately. The fifth part is a synthesis of the experiments carried out in the first two parts. In the end of the paper, it contains the conclusions and discussions.

2. Relevant research

2.1. Research on the COVID-19 pandemic data compilation and warning indicators

Since 2020, COVID-19 has brought several decades of unprecedented emergency management pressure to government agencies around the world, but also motivating scholars to investigate public health emergencies [11,12]. After the occurrence of the epidemic, various pandemic-related data cover all aspects of the epidemic, as well as key information on the occurrence, development and outcome of the epidemic. It also includes some unproven or fake data information [13]. According to the characteristics of this outbreak, selecting the influencing factors of the outbreak from the massive data, selecting the emergencies from different angles at multiple levels and all levels, which can help us collect data and information more effectively, thus ensuring the credibility of prediction and reducing the waste of resources.

In the early stages of COVID-19 outbreak, the identification of clustered epidemics can be used for spatial tracing [14]. The purpose of spatial tracing is to find the first case and the initial source of infection. It is also the key to eradicate the epidemic fundamentally. It mainly depends on spatial analysis after collecting data from government agencies, social media, crowdsourcing system and other methods of early cases [15]. Huang et al. [16] prospectively collected and analyzed data from 2019 laboratory-confirmed patients with COVID-19 pandemic. Data is derived from electronic medical records using standardized data collection forms shared by the International Federation of Serious Acute Respiratory Tracts and Emerging Infections. Around 66% of the cases were found to have a history of exposure to the seafood market in Southern China, suggesting that the seafood market in South China may be an early outbreak site of the epidemic in Wuhan. Although previous research data provided great support to prevent virus transmission, spatial tracing has been the focus and difficulty of epidemic prevention and control, because the uncertainty of virus micro-transmission cannot be eliminated by inferring from field investigation data. After the epidemic began to spread widely, Jia et al. [17] developed a spatio-temporal risk tracing model after collecting more than 10 million personal mobile phone data in and out of Wuhan to predict the distribution of the epidemic using population movement data with a view to population migration data. Through the dynamic population flow between regions, we can identify the direction of the spread of the epidemic and analyze the spatial transmission path of the epidemic, so as to provide support for effective prevention and control decisions. During public health emergencies, real-time updated data about the pandemic may become a useful solution to emergency management. The analysis of pandemic-related data through natural language processing has also become a hot spot for scholars. This kind of research provides a simple choice for practical application and a good starting point for future research [18]. For instance, Zheng et al. [19] considered the effects of prevention and control measures and public awareness of prevention. They fused pandemic-related data and embedded natural language processing module and artificial neural network for analysis, and finally pointed out that in addition to the common prevention and control measures, public awareness of the epidemic also has a significant impact on the spread of COVID-19.

At present, some countries have established a unified monitoring, warning and reporting system for public health emergencies, like Electronic Surveillance System for Early Notification of Community-based Epidemics (ESSENCE) in the US. However, the system did not respond in time during the COVID-19 pandemic, which may be due to the single source of monitoring and early warning data indicators in the existing early warning system. The value of a large number of spatio-temporal big data has not been mined and reflected, and the monitoring data still mainly comes from medical and health institutions; The data information is one-sided and lacks other information of great significance for early monitoring and early warning, such as specific symptoms, contact history, life history, traffic history, etc. It is impossible to accurately carry out infectious disease early warning only by relying on the data of clinical diagnosis results.

To sum up, a handful of studies have been the research basis for systematic research based on the data compilation of COVID-19 and related indicators. However, the research and practical application of COVID-19 are mostly based on static case data, and there is a lack of evaluation indicators to extract and summarize various data. There are many factors influencing the outbreak of infectious diseases and they cover a wide range of areas, the comprehensive support of multiple data is increasingly needed to improve the monitoring and early warning capabilities. If we can set up a monitoring and early warning index system based on multivariate data for infectious disease epidemic and take multivariate data into account, we can further improve the accuracy and sensitivity of early warning, so as to play a certain supporting role in the study of early warning decision-making of new emergency think tanks as well as strengthen and highlight their risk early warning capabilities. Therefore, it is necessary to analyze and sort out the data of each link in detail to form an index system for epidemic assessment.

2.2. Research on prediction model for dynamic evolution of COVID-19

As an important entry point for COVID-19, the Internet has become the main platform for information sources [20]. In this epidemic prevention war, can we rely on modern technology, integrate various algorithms in the field of modern big data, build intelligent epidemic identification model, and bring into play the value in fighting epidemics? Based on Susceptible-Infected-Recovered(SIR) model, Inman et al. [21] predicted the prevalence of COVID-19. The key parameter of
SIR differential equation is Q-value. In the outbreak of COVID-19, this value was very high, which indicates that the virus spreads quickly. Maintaining social distance helps to reduce the likelihood of contact with people infected with the epidemic, which helps to reduce Q-values and control transmission. Wibowo and Wihayati [22] chose the most common supervised learning algorithm in machine learning, a logistic regression model, and used the least squares criterion and gradient descent method to non-linear regression of data, to find the non-linear relationship between days and the number of confirmed patients, and then used mathematical modeling to predict the distribution of patients with COVID-19 and the number of patients recovered in Indonesia. In addition, Radulescu and Cavanagh [23] created a general SEIR model, focusing on describing the transmission mechanism of COVID-19 in susceptible populations. Combined with the actual mode of transmission, the average latency, the average infection period and the proportion of atypical patients were used to simulate the pandemic in major prevalent countries. Noam Barda et al. [24] considered that personal data may not be of much value at the beginning of the epidemic, but they also need to establish an overall risk prediction model to support prevention and decision-making. By developing a risk predictor for severe respiratory infections, the epidemic risk can be predicted on the basis of calibrated predictions. By describing the latest advances and applications of current Artificial Intelligence (AI) and Big Data technologies in various fields, Quoc-Viet Pham et al. [25] have provided researchers with new insights on the use of new technologies such as AI and Big Data to study the epidemic situation of new coronary pneumonia, and promoted the use of emerging technologies for epidemic prediction and control.

Jumping out of the background of COVID-19, many scholars have conducted machine learning-based identification model analysis and research on various epidemics before. For example, in view of the lag of traditional infectious disease surveillance, a hidden Markov model for epidemic identification and surveillance is proposed, which combines large data from Internet search engine and WHO infectious disease surveillance data. Using the time characteristics of the spread of infectious diseases in the same country or region, this method can achieve a high accuracy and real-time identification and monitoring of infectious disease outbreaks in a single country or region. Compared with traditional methods, the lag is much reduced [26]. Combining basic data with other factors, Pan et al. [27] proposed to use Long Short-Term Memory (LSTM) loop neural network to predict the incidence of infectious diseases, in order to effectively improve the accuracy of the identification model.

Indeed, many researchers are starting to carry out data-based epidemic prediction studies using various algorithms to supplement the deficiencies of existing prediction systems. In these studies, large data, such as Internet search queries, are being used abroad to monitor the occurrence of infectious diseases. Internet search data can be collected and processed at a near real-time speed. Sherry et al. [28] have found that searching for data over the Internet can create infectious disease identification models faster than traditional monitoring systems. In addition, some scholars such as Huang et al. [29] have attempted to use the generalized additive model (GAM) to identify and predict hand, foot and mouth disease, which includes the best results obtained by searching and querying data to obtain new tools for identifying and monitoring large data, which have the advantage of easy access to the incidence of infectious diseases and can identify infectious disease trends in both epidemic and official organizations.

After a comprehensive analysis of the above research progress, it is found that the dynamic prediction research based on the infectious disease dynamics model needs to involve multiple factors and must be comprehensively applied. When using SEIR and other dynamic models for research, it is necessary to analyze the relationship between social factors such as personnel mobility, population density and specific prevention and control measures. Different parameters have a great impact on the prediction results, so it is necessary to accurately introduce relevant parameters and complete the prediction. This model has strong pertinence, but its generalization ability needs to be further improved. The dynamic prediction research based on machine learning does not need to consider too many influencing factors, but its results are directly affected by the data correlation. For short-term prediction, small errors can be obtained by using relevant machine learning methods, but with the passage of time, the errors will gradually accumulate. Without other intervention measures, the prediction may deviate from the actual results seriously, and the sustainability is not strong [30–32]. The next part in this paper will dig several kinds of data to understand the evolution status after the outbreak, and gain insight into the risk level of the current epidemic gap. On this basis, determine the influencing factors of the epidemic situation and quantify them, establish a scientific and reasonable index system, establish a corresponding system dynamic model, and predict the model based on the index system. To analyze the impact of national emergency management policies on the trend of COVID-19.

3. Formulate Preventive Indicator System for COVID-19 based on feature engineering

3.1. Overview of feature engineering

Feature engineering is a term used in the field of machine learning [33]. In the era of big data, data has become so complex that it needs to accurately capture its most essential characteristics. In other words, the feature is to extract useful information from the dataset to predict the results. It is also the main basis for predicting samples of machine learning algorithm models and an important factor to determine the upper performance limit of algorithm models. In datasets with different structures, there are different forms of features [34]. In a tabular (structured) dataset consisting of variables or attributes (columns) and instances (rows), variables or attributes can be considered characteristics. In a nutshell, features are the information that is processed from the original data to describe the individual samples. Before features are actually used for prediction, there is no guarantee that the original attributes or extracted features will have a positive effect on the result prediction [35]. To sum up, in order to get a practical and effective prediction model, this paper first needs to make an index system to ensure that the extracted data features are effective, which is called feature engineering. The flow diagram is shown in Fig. 1 below.

Formally, feature engineering uses professional background knowledge or skills to process the original data, generating some information describing the characteristics of the sample, replacing the original data as input to the model, and making the features play a better role in the algorithm [36]. In the field of computer science, there is the saying that “feature engineering determines the upper limit of generalization ability”. Thus, feature engineering is important throughout the model process. Owing to a difficult and time-consuming process, feature engineering requires a variety of possibilities tests. In the actual data mining work, it takes a lot of time to complete feature extraction and organization, and then only a simple model is needed to obtain good prediction results [37].

3.2. Formulate the epidemic warning indicator system

For early warning, a set of index system must be established to analyze and evaluate the situation of an infectious disease, judge the probability and severity of the crisis, and decide whether to send a crisis alarm. Early warning indicators are indicators with potential early warning value and can give an early warning in the early stage of crisis events [38]. For the epidemic situation of new infectious diseases, the outbreak trend can be reflected in the fluctuation range of early warning indicators. Once the fluctuation range of early warning indicators exceeds the warning line, an alarm can be issued and corresponding epidemic investigation or intervention can be carried out. This section
will analyze the factors that affect the development of epidemic situation from the perspective of feature engineering, quantify the influencing factors of epidemic situation on the basis of forming an effective data structure of epidemic monitoring, and establish an epidemic situation evaluation index system to provide a basis for epidemic prediction and prevention and control.

As the most important feature extraction, early warning indicators need to have the characteristics of timeliness, accuracy, scientificity and practicability [39]. The sensitivity of infectious disease early warning system mainly depends on which node the monitoring data used comes from the occurrence and development of infectious diseases. For example, the use of data on social and natural factors affecting the occurrence and development of infectious diseases, infectious disease related symptoms or suspected cases helps to improve the sensitivity of infectious disease monitoring and early warning [40–42]. Based on the three stages of the occurrence and development of infectious diseases, the early warning indicators are divided into three categories: pre-outbreak indicators, early-symptom-period indicators and specific-syndrome-period indicators, as shown in Fig. 2.

(1) Pre-outbreak indicators. These indicators refers to various source factors leading to the crisis, which belong to the indicators in the early stage of epidemic outbreak or epidemic, and are mainly related to the nature of infectious diseases, such as the exposure of pathogens, the change of transmission route of infectious sources, or the potential risk factors found in the external environment conducive to the outbreak or epidemic of infectious diseases.

(2) Early-symptom-period indicators. These indicators refers to that when the alarm source indicators change abnormally, there are always some omens, mainly manifested in the increase in the sales of all kinds of over-the-counter drugs that may be related to infectious diseases in the hospital and the increase in patients

![Flow diagram of feature engineering](image1)

![Infectious disease occurrence and development stage model](image2)
with similar clinical symptoms with infectious diseases, which can be regarded as the warning indicators of atypical symptoms.

(3) Specific-syndrome-period indicators. These indicators refer to the risks that have a certain impact on the society and have been shown by the epidemic situation. They are indicators in the period of typical symptoms of infectious diseases, such as the number of reported cases and deaths of infectious diseases.

According to the idea of feature engineering, the construction of index system needs to have a clear logic, each index is independent and related to each other to form an organic whole. Based on Fig. 3, this paper intends to build a more comprehensive and systematic epidemic early warning indicator system from five aspects: indicators of the nature of infectious diseases, indicators of potential risk factors, health-related indicators, indicators of epidemic identification and indicators of network public opinion, as shown in Fig. 3 [43–45].

(1) Indicators of the nature of infectious diseases. These indicators primarily monitor the properties of infectious disease agents, and the role of such indicators for the predictive early warning of infectious diseases cannot be discounted. Take the example of SARS-COV-2, which is characterized by a wide range of hosts [46] and by a highly infectious and highly covert transmission profile [47]. Only a full understanding of the pathogen characteristics of infectious diseases can be taken at the crux and a good response can be made.

(2) Indicators of potential risk factors. The main transmission routes of SARS-COV-2 are airway droplets and close contact transmission [48,49]. The patients infected by the new coronavirus and asymmetric infected persons are the main sources of transmission. In the process of epidemic prevention and control, grasping the transmission route is the precondition for effective prevention and avoidance of infectious diseases.

(3) Health-related indicators. These indicators are the precursors of an outbreak or pre-epidemic of infectious diseases, including medical-related behavioral indicators, indicators related to the use of medical services, indicators of consultation information, etc.

(4) Epidemic identification indicators. These indicators are mainly real-time statistical data of monitoring report of new crown pneumonia epidemic and are the most commonly used indicators for predicting infectious diseases in theoretical research and practical work. The monitoring and reporting network information such as the number of confirmed cases and the number of suspected cases can be used for the forecasting and early warning research of epidemic situation [50,51].

(5) Public opinion indicators. In the era of big data, social media platforms not only play a role of information dissemination in public events, but also have the ability to depict public behavior and emotional characteristics [52]. The accompanying network public opinion information often has the characteristics of fast dissemination, strong outbreak, etc. It also interacts with the public events themselves and constantly affects the security risks of public events. Therefore, public opinion indicators also play a supplementary role.

Based on the basic structure of infectious disease epidemic warning indicators system, this paper combines the process of the occurrence and development of new coronavirus pneumonia with the basic principles of early warning theory, collects and sorts out various possible data collections, and proposes a three-level basic framework of COVID-19 warning indicator system (Table 1).

The construction of this indicator system covers the whole process of the occurrence and development of infectious diseases, and comprehensively considers the multifaceted data such as pathogen surveillance data, natural outbreak data, hospital reporting data, social and network media, which can fully mine the value of big data and improve the timeliness and accuracy of infectious disease information warning. Meanwhile, based on the classification of early warning indicators, this index system is divided into three levels to form a whole, which has a certain theoretical basis. Among the foregoing warning indicators, some can be obtained by publishing data sources; some can be obtained unconventionally, but can also be obtained through some efforts in the current situation; and a few others are difficult to obtain under current conditions, but remain in the indicator system due to the importance of indicators for early warning work.

Because geographical variations, economic differences, and other factors, different regions, different infectious diseases have their own epidemic characteristics, so for all regions, a uniform system of early warning indicators should be established for all diseases, and it is inappropriate and impossible to set a fixed cordon. The indicator system in this paper is built for infectious diseases, especially COVID-19. There is an overall and individual relationship when the specific application, and the early warning indicators should be rounded and revised according to the characteristics of specific diseases and specific regions.

In the process of epidemic development, a large number of data produced by the epidemic show the characteristics of multi-source, massive, rapid growth and constant updating. Therefore, the construction of early warning index system by integrating multi-party
Table 1 COVID-19 warning indicator system.

| Level 1 indicators | Level 2 indicators | Level 3 indicators | The source of the data |
|-------------------|--------------------|--------------------|------------------------|
| A Pre-outbreak Indicators | B1 Indicators of the nature of infectious diseases | C1 Variations of conventional popular advantage strains | National Genomics Science Data Center |
| | | C2 Changes in drug resistance, toxicity, antigen resistance, etc. of new coronavirus | National Genomics Science Data Center |
| | | C3 Focus on monitoring the carrying rate of new coronavirus in the population | Monitoring data from the local health board |
| | | C4 A major laboratory safety incident occurred | The latest notice of the National Health and Construction Commission |
| | | C5 Network direct coverage | Direct reporting system of the National Infectious Diseases Network Migration data |
| | | C6 Natural or social factors that are suitable for the epidemic/outbreak of infectious diseases | |
| A Early-symptom-period Indicators | B1 Health-related indicators | C6 Sales of various types of over-the-counter drugs (e.g. cough control, fever medicine, etc.) in hospitals that may be related to the outbreak | Medical institutions - medical-related supplies sales monitoring |
| | | C7 Absence rates and absences due to illness in densely populated units | Health monitoring of relevant organizations or employment units at the grass-roots level |
| | | C8 Morning check-ups in densely populated units revealed abnormal body temperature | Health monitoring of relevant organizations or employment units at the grass-roots level |
| | | C9 Number of fever emergency patients | Diagnosis and reporting by medical institutions |
| | | C10 Track the number of quarantined persons (excluding suspected cases) | Diagnosis and reporting by medical institutions |
| | | C11 Number of unexplained pneumonias | Diagnosis and reporting by medical institutions |
| | | C12 Number of non-specific and similar clinical syndromes that may be associated with neo-coronary pneumonia | |
| A Specific-symptom-period Indicators | B1 Epidemic identification indicators | C13 Number of confirmed cases | Real-time big data reporting of outbreaks |
| | | C14 Number of suspected cases | Real-time big data reporting of outbreaks |
| | | C15 Number of deaths | Real-time big data reporting of outbreaks |

4. Prediction model of dynamic evolution of COVID-19 based on system dynamics

The prediction of infectious diseases can early detect the trend of disease development, lay the foundation for the early warning of disease, and provide a theoretical basis for formulating prevention strategies and measures. When responding to the COVID-19 pandemic, it is of great significance to establish an appropriate prediction model and improve the accuracy of prediction.

4.1. Overview of system dynamics

System dynamics, also known as industrial dynamics, mainly refers to the use of system simulation to analyze and solve management problems [52]. The study of system dynamics on system problems is based on the close dependence of the internal behavior mode and system structure of the system. By establishing a mathematical model, the causal relationship in changing forms can be gradually discovered [54]. As the study of system dynamics has been applied to the study of various fields. By combining various research methods, more accurate mathematical simulation results can be obtained, which provides a basis for decision-making. At present, system dynamics is increasingly being used in socioeconomic research to explore economic and social operating patterns through system design and simulation, thus providing managers with more reliable recommendations [55]. System dynamics methods have also been applied to the study of major contagious diseases, through system construction and simulation, the relationship between indicators in the system can be analyzed, and the future trend of indicators can be predicted. System dynamics models consist of variables, parameters, and functional relationships, which can be used to analyze the relationship between factors and dynamic change trends.

4.2. Overview of SEIR model

SEIR model, also known as a kind of compartment model, is an infectious disease model invented in the early last century, and also a classical system dynamics model. This model can roughly show the monitoring data is undoubtedly a further promotion and supplement to the existing early warning system and existing research. Its analysis process and results also provide a certain theoretical basis for the next part of the research on epidemic prediction using SEIR model.
process of an infectious disease from latent to onset and then to the end [56]. It is suitable for the type of malignant infectious disease with latent period [57,58].

According to the traditional SEIR model, the population within the epidemic scope of infectious diseases is divided into four categories: S, namely susceptible population, refers to the population who has not yet acquired the disease, but lacks immunity and is susceptible to infection after contacting the susceptible population; E, namely exposed population, refers to the susceptible population who initially experiences the incubation period and will have symptoms after a period of time; I refers to the population who is infected with infectious diseases, and the infectious disease can be transmitted Class s members; and R, people who are removed from the infectious disease model, refers to the people who are isolated or immune to the disease [59]. Assuming that the total number of people in the sample area is constant, \( N(t) = \text{constant} \), in unit time \( t \), the number of susceptible people that a patient can infect is directly proportional to the total number of susceptible people \( S(t) \) in this environment, and the proportional coefficient is \( \sigma \), so the number of people infected by all patients in time \( t \) is \( \sigma S(t)E(t) \) [60]. The number of latent persons who can be transformed into infected persons is directly proportional to the total number of latent persons \( E(t) \) in this environment, and the proportion coefficient is \( \beta \), so the number of persons infected by all patients in time \( t \) is \( \beta E(t)I(t) \). In addition, the disease can be cured or lead to death, resulting in \( R(t) \) population. At time \( t \), the number of patients removed per unit time is directly proportional to the number of patients. The proportion is recorded as \( \gamma \), and the number of patients removed per unit time is \( \gamma R(t) \) [61].

Therefore, the flow chart of traditional SEIR model is shown in Fig. 4 below:

It can be seen from Fig. 4 that this model is a one-way model, the number of susceptible people is continuously input to the number of infected people, and at the same time, the number of infected people is also input to the number of recovered people in one-way, so the number of susceptible people and the number of infected people will eventually drop to 0, and at the same time, all people will become the number of recovered people [62].

The differential equation of traditional SEIR model is as follows [63]:

\[
\begin{align*}
\frac{dS(t)}{dt} &= -\beta S(t)I(t) \\
\frac{dE(t)}{dt} &= \beta S(t)I(t) - (\sigma + \gamma)E(t) \\
\frac{dI(t)}{dt} &= \sigma E(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t)
\end{align*}
\]

(1)

4.3. Improvement of SEIR model based on epidemic warning indicator system of feature engineering

As mentioned earlier, the basic SEIR model shows the temporal variation of the four groups of people under relatively relaxed conditions mainly through differential variance. However, it may not be directly used in the prediction of COVID-19. In the first step of modifying the model, we first need to analyze the characteristics of COVID-19. Unlike some viruses that caused infection in the past, such as SARS, which show infectivity only during the outbreak period, COVID-19 is infectious in the latent period. Based on these characteristics, we have derived a new model flowchart as shown in Fig. 5 below:

The above flowchart shows the evolution of COVID-19 without human intervention. Accordingly, it is necessary to determine the infection probability of the \( \beta_2 \) model, the propagation probability of the model is modified, so that the susceptible individuals is transformed into the latent, and the number of healthy susceptible contacted by the latent every day is \( S_2(t) \).

Therefore, on the basis of formula (1), it is necessary to add latent \( E \) and contact number \( S_2(t) \) to S infection probability \( \beta_2 \) and contact number \( S_2(t) \) of susceptible healthy people, which will make more normal people become latent, but also more in line with the actual situation of new coronary pneumonia virus infection. The formula of modified SEIR model is as follows:

\[
\begin{align*}
\frac{dS(t)}{dt} &= -\beta_2 S(t)I(t) - \beta_2 S_2(t)E(t) \\
\frac{dE(t)}{dt} &= \beta S(t)I(t) - (\sigma + \gamma)E(t) + \beta_2 S_2(t)E(t) \\
\frac{dI(t)}{dt} &= \sigma E(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t)
\end{align*}
\]

(2)

In addition to the characteristics of COVID-19, the intervention of human environment has a significant impact on the trend of the epidemic. In the third part of the study, we have mentioned the epidemic situation identification indicators, which is the main object of study in this part. In addition, we have summarized indicators of potential risk factors. The data represented in this section is likely to be an influential factor in the SEIR model, interfering with the direction of prediction. The main transmission routes of SARS-CoV-2 are respiratory foam transmission and close contact transmission, with contacts and infected persons becoming the main source of infection.

Therefore, based on the SEIR model flow chart, combined with the
potential risk factor indicators in the previous part, this paper obtains three new factors: the number of isolation susceptible subjects, the number of suspected cases and the number of confirmed cases, so as to represent the change of epidemic situation after intervention by human factors. The flow chart of the optimized model is shown in Fig. 6.

According to Fig. 7, the differential equation of the SEIR model is also modified. Assuming that the total number of persons managing a region is constant, \( N(t) = n \) (\( n \) is a constant). Let the isolation proportion be \( q \), the contact rate be \( c \), and the effective contact coefficient be \( \theta(1 \) is taken in the natural state), then \( \theta c \) be the effective contact rate. The conversion rate from susceptible to isolated susceptible was \( \theta(cq(1 - \beta)) \), from susceptible to isolated latent was \( \theta(cq) \), and from susceptible to latent was \( \theta(cq)(1 - q) \). Considering the impact of infected persons without isolation I and latent persons E on susceptible population S, and that SQ will become susceptible again after de-isolation, we choose \( \rho \) to indicate the rate of de-isolation, \( \rho = 1/\text{day of isolation} \). Record that the rate of latent transformation to the infected person is \( \sigma \), \( \sigma = 1/\text{latent days} \), the isolation rate of the infected person is \( \delta \), the mortality rate is \( \gamma_1 \), and the cure rate is \( \gamma_2 \). The optimized formula (3) is as follows:

\[
\begin{align*}
\frac{dS(t)}{dt} &= -[\theta c + \theta cq(1 - \beta)]S(t)[I(t) + \theta E(t)] + \rho S[t] \\
\frac{dE(t)}{dt} &= \theta c\beta(1 - q)S(t)[I(t) + \theta E(t)] - qE(t) \\
\frac{dI(t)}{dt} &= \sigma E(t) - (\delta + \gamma_1 + \gamma_2)I(t) \\
\frac{dR(t)}{dt} &= \gamma_1E(t) + \gamma_2I(t) \\
\frac{dS(t)}{dt} &= \theta c\beta(1 - \beta)S(t)[I(t) + \theta E(t)] - \rho S[t] \\
\frac{dE(t)}{dt} &= \theta c\beta qS(t)[I(t) + \theta E(t)] - \delta E(t) \\
\frac{dI(t)}{dt} &= \delta(t) + \delta E(t) - (\gamma_1 + \gamma_2)I(t) \\
\frac{dR(t)}{dt} &= \gamma_1E(t) + \gamma_2I(t)
\end{align*}
\]

In the above-mentioned model, part of the index data comes from the epidemic surveillance data certified by China National Health and Construction Commission and the official notification, which ensures the reliability of the data on the one hand, and the authenticity of the model on the other hand. Secondly, the weight of the new impact parameters is affected by the actual emergency management mechanism of each country. The policy constant is determined by the expert scoring method, and the index weight is calculated by the analytic hierarchy process to ensure the accuracy of the prediction results.

5. Experiment and results

5.1. Experimental data acquisition and selection

Based on the above COVID-19 epidemic early warning system, the basic data of SEIR model was obtained from COVID-19 spread in real-time published by the center for Systems Science and Engineering at Johns Hopkins University [64]. To ensure the accuracy, it is also checked through the daily announcement of National Health and Wellness Commission of China and the global data published by the WHO. After data screening, classification and other pre-processing work, combined with the data trend of various countries, the cumulative confirmed cases, cumulative suspected cases, newly diagnosed cases, cumulative death cases and cumulative number of cured cases from January 16 to May 15, 20 in China, cumulative confirmed cases and cumulative suspected cases from February 1 to April 30, 20 in Japan, Korea, and the United Kingdom (UK) were finally retained, new confirmed cases, cumulative deaths and cumulative number of cured cases, cumulative confirmed cases, cumulative suspected cases, new confirmed cases, cumulative deaths and cumulative number of cured cases from March 1 to May 29, 2020 in the United States (US), using 90-day new corona outbreak-related datasets from each of the five countries as training set samples, with time datasets of confirmed cases. As shown in Table 2.

5.2. Experiment and results

According to the latest population statistics of China [65], Japan
International Journal of Disaster Risk Reduction 65 (2021) 102547

9

Table 3
Initial data used in each country.

| Country | Population | Initial value of cumulative cases | Initial value of new cases | Initial value of cumulative deaths | Initial value of cumulative cured cases | Initial value of removers |
|---------|------------|-----------------------------------|---------------------------|-----------------------------------|----------------------------------------|--------------------------|
| China   | 1.40 billion | 6                                 | 8                         | 2                                 | 10                                     | 10                       |
| Japan   | 125.30 million | 20                                | 5                         | 0                                 | 1                                      | 1                        |
| Korea   | 52.23 million | 12                                | 1                         | 0                                 | 0                                      | 0                        |
| UK      | 67.08 million | 36                                | 13                        | 8                                 | 8                                      | 8                        |
| US      | 329.48 million | 36                                | 6                         | 1                                 | 7                                      | 8                        |

Fig. 7. Graphs of cumulative confirmed cases of new coronavirus within 90 days in each country.

[66], Korea [67], the United Kingdom (UK) [68], and the United States (US) [69], the current population of these countries is about 1.40 billion, 125.30 million, 52.23 million, 67.08 million and 329.48 million respectively. Based on the data set mentioned above, we have obtained the initial value s of cumulative confirmed cases, new confirmed cases, cumulative deaths and cumulative cured cases, which is China’s real-time statistics on January 16, 2020; Japan, Korea and the UK’s real-time statistics on February 1, 2020; the US’s real-time statistics on March 1, 2020, just as shown in Table 3. In the actual situation, since both the cases of cure and death represent that the patient no longer has the risk of infecting others in most cases, the number of cumulative deaths and cumulative deaths were added up to the number of removers from the model in the study.

The development of the epidemic is closely related to the selection of values. In the SEIR model, β represents the conversion rate from exposure to disease, which is usually determined by the nature of the disease itself. Based on analysis by medical experts from actual COVID-19 clinical cases [70,71], the incubation period of the disease is around 7 days, so β=1/7. In addition, other parameters, such as probability of infection σ, isolation ratio q, contact rate c, recovery rate γ, are closely related to the specific social environment and prevention and control policies of each country. In view of this, we use the data set of the first ten days of each country for parameter fitting and error testing, and generates the best historical simulation state under a series of parameter sequences through continuous trial and error, so as to alleviate the problem of inaccurate prediction model caused by parameters. Among them, we set at the beginning that if there is no protection policy, the contact rate in the area is 1, and the isolation ratio does not exist. In addition, we adopted the view of other experts that “the infection probability β of susceptible persons after contacting patients is about 0.19–0.24, which may be slightly different due to the social habits and age structure of different populations” [72]. Then, we adjusted the value of contact rate c according to the actual prevention and control policies of different countries. Take China as an example. From the lockdown of Wuhan on January 23, 2020, to the semi-closed state of major epidemic prevention cities, people across the country stayed at home and rested in line with national policies. Enterprises across the country were suspended, schools were suspended, and the entire society is in a state of shutdown, and the contact rate will decrease accordingly. Finally, the values of q and γ are related to the current medical conditions of various countries. Based on the indicators sorted out in the third part, we set different values under different actual conditions in various countries, and through continuous trial and error, we adjusted the parameters in order to make more realistic, reasonable and standardized forecasts.

When the number of new diagnoses decreases, the inflection point is reached. When the number of new cases increases to zero, the cumulative number of infected cases reaches the maximum, which is the peak of the cumulative number of infected persons. In this paper, the optimized SEIR model equation is converted into a program, and according to the model, real data from different countries are introduced to test the prediction results, which are integrated as shown in Table 4.

According to Table 4, the pandemic in China, Korea, and Japan can be effectively controlled before April 30, 2020, and new cases will gradually show a decreasing trend. Among them, China’s inflection point has the earliest appearance date, with a turn after a large outbreak of about one and a half months. Additional projections show that the United States is the country with the largest number of patients in the world, and even in general model projections, the highest peak in the
number of patients in the United States is not seen and can only be further speculated on the basis of projections toward. Apart from that, the situation in the UK with respect to the pandemic situation trend is also somewhat less optimistic, with the number of patients predicted to reach its peak in Mid-June 2020, approximately 333 000 cases being diagnosed cumulatively, at a time when nearly four months have passed since the pandemic just started. Based on this, the United Kingdom and the United States still need severe prevention and control, and even more proper emergency management measures are needed to prevent further epidemic spread and warn citizens of every effort to avoid travel to countries with similar conditions such as the United Kingdom and the United States. Citizens in these areas should also avoid out aggregation to countries with similar conditions such as the United Kingdom and the United States. Citizens in these areas should also avoid out aggregation to countries with similar conditions such as the United Kingdom and the United States. Citizens in these areas should also avoid out aggregation to countries with similar conditions such as the United Kingdom and the United States. Citizens in these areas should also avoid out aggregation to countries with similar conditions such as the United Kingdom and the United States.

6. Conclusions

As of the end of August 2021, the global impact of COVID-19 has not yet ended, and even due to the sudden mutation of the virus, SARS-CoV-2, it has become more difficult to cope with the epidemic. Based on the previous analysis, we believe that the conclusions of this study are important in order to bring warnings to the policies of various countries. The main content of this paper includes the following two aspects: Firstly, based on the level of early warning indicators, five analysis dimensions of the indicators of the nature of infectious diseases, indicators of potential risk factors, health-related indicators, indicators of epidemic identification and indicators of network public opinion, are extracted, thereby an early warning indicator system for the new crown pneumonia epidemic has been established. Second, it is based on the SEIR model in system dynamics, adjusting parameters according to the policies of various countries, fitting research on the dynamic progress, and simulating and predicting the future evolution trend of COVID-19 epidemic. In the experiment, this paper focuses on sorting out the value logic of the impact factors based on the indicator system, and through continuous trial and error, to obtain a predictive model with a higher degree of fit, and analyzes COVID-19 under the premise of different emergency measures in various countries. The parameters of the prediction model are set in combination with various indicators. The prediction trend can be adjusted through a variety of parameters to obtain results with higher credibility than traditional models. Therefore, when choosing an infectious disease prediction model, we should consider the characteristics and transmission routes of the epidemic disease in many ways, as well as the intervention measures of the prevention and control department.

Therefore, the early warning index system and optimization of infectious disease model proposed in this paper is an effective method, which can predict the peak and inflection point of this round of epidemic situation and the corresponding total number of infections many days in advance. By adjusting the parameters, it can also judge the approximate date of the infection peak. This will help the government to better plan their emergency management means and prevention and control measures.

7. Discussions and limitations

Based on the research in this paper, we believe that the forecast and early warning of major infectious diseases represented by COVID-19 need to include two aspects of research simultaneously. First of all, it is necessary for us to find sensitive and effective early warning indicators as well as establish a complete early warning monitoring system. On the one hand, an indicator system can be used to collect and filter useful and relevant data from the existing massive and responsible information to ensure that the data is timely, complete and accurate. On the other hand, we also hope that through the guidance of an indicator system, we will have the opportunity to discover some new valuable data, which is valuable for early detection of infectious disease outbreaks or epidemics, or for more accurate prediction and warning. In addition, at least at this stage, there is no mathematical model that can be directly applied to the prediction and early warning of various infectious diseases. Although many models have a certain degree of generalization. However, in practical applications, different infectious viruses have their own characteristics, and the prediction model must also be modified accordingly to enhance the ability of prediction and early warning.

The paper strives to provide early warning of emergencies scientifically and effectively through the combination of these two technologies, and put forward feasible references for the implementation of various countermeasures. Judging from the conclusion, the paper has made certain contributions to the prediction and early warning research of infectious disease epidemics on both theoretical and practical levels.

Firstly, in terms of theory, the paper provides new research perspectives and methods for epidemic forecasting and early warning research. Monitoring and early warning of infectious diseases based on multi-source data analysis is an inevitable trend in the development of the information age, and most of the existing studies in this field have relatively single data sources. Since the epidemic of infectious diseases is characterized by population size, time and spatial distribution, it is a concentrated expression of the interaction of natural geographical, environmental factors, social and economic development conditions, and the flow of people in a specific time and space. During the development of the epidemic, its occurrence a large amount of data presents the characteristics of multi-source, massive, rapid growth, and constantly updated. Therefore, the paper constructs an early warning indicator system by integrating multi-party monitoring data and applies it to the construction of predictive models. Through the establishment of more accurate models, the overall utilization of epidemic data resources in the whole society is realized, and a new theoretical idea is provided for epidemic early warning and prevention and control.

Secondly, in terms of reality, the paper puts forward a new idea based on big data thinking to construct the forecast and early warning of infectious disease epidemics, which can be effectively applied to actual prevention and control. The early warning indicators in this paper are based on the analysis of comprehensive big data, which can be regarded as the data structure of sudden public health incidents. Through the collection and analysis of data on meteorological factors, population flow trends, and the incidence of various syndromes, etc. Study the external factors and individual internal factors that affect the occurrence and development of emergency management events. Although the SEIR model selected in this paper can better capture the historical changes of the epidemic in actual forecasting, and can reveal the autocorrelation of the development of the epidemic, it also shows that the model is affected by various factors such as the incubation period of the disease, prevention and control policies and measures. The impact of is relatively large, so if you use this model alone to make predictions, you need to carefully study the relevant factors, and the setting of parameters will affect the accuracy of the prediction. The combination of indicators and forecasts has also improved the actual operability of the two aspects to a certain extent.

This paper combines the forecasting ideas and comprehensively analyzes the data of COVID-19, and obtains more comprehensive results. However, there is still room for further improvement in research. First of all, the index system proposed in this paper is based on the various data and information that has attracted attention after the occurrence of COVID-19, and has been extracted using feature engineering methods, but there is still a lack of actual verification. The development and improvement of the index system must be experimental and verified in practice, so as to further remove the indicators with low early warning value or difficult to obtain in practice, classify and simplify the indicators with high relevance, improve the operability of the index system as much as possible and promote the application of the index system. Therefore, it is necessary to carry out pilot studies to gradually explore the application of early warning indicators. The evaluation and verification of indicators requires the accumulation of a certain amount of historical data. Due to the time limit of writing this paper, there is not
enough data to verify and evaluate the indicators. Second, early warning and forecasting are based on historical data and actual materials to predict the future, and provide suggestions for the management department to grasp the status quo and the future. Prediction is only an analysis of the future development trend of infectious diseases, but early warning is different from prediction. It needs to alert early abnormal situations and initiate emergency response. This thesis only stays in the forecasting stage. How to combine forecasting and early warning in depth is the direction and focus of further development. Finally, the occurrence and development of infectious diseases are affected by many factors. If we can observe the dynamic process of a group of (multi-dimensional) related early warning indicators at the same time, and study it as a whole, and use appropriate models to predict, try to integrate multiple indicators and multiple methods for analysis, it is bound to be able to more systematically and comprehensively reflect the internal regularity and future trends of dynamic phenomena. This will also be our future research thinking.

In conclusion, the combination of index system and prediction method for epidemic surveillance is an alternative research direction and path, which provides a feasible way to study the spread of infectious diseases.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We acknowledge the supports from National Social Science Major Program of China (No. 21ZDA105).

References

[1] Y. Zhang, X. Cao, P. Wang, G. Wang, Q. Xiao, Emotional ‘inflection point’ in public health emergencies with the 2019 new coronavirus pneumonia (nCp) in China, J. Affect. Disord. 276 (2020) 797–803.
[2] M. Tortajada-Soler, L. Sánchez-Valdeón, M. Blanco-Nistal, J.A. Benitez-Andrades, C. Liébana-Presa, E. Bayón-Daristade, Prevalence of comorbidities in individuals diagnosed and undiagnosed with Alzheimer’s disease in León, Spain and for contingency procedures to follow in the case of emergencies involving people with Alzheimer’s disease, Int. J. Environ. Res. Public Health 17 (10) (2020) 3598.
[3] L. Zhang, W. Zhao, B. Sun, Y. Huang, G. Ginzel, How scientific research reacts to international public health emergencies: a global analysis of response patterns, Scientometrics 124 (2020) 747–772.
[4] M. Cevik, M. Tatar, O. Lloyd, A.E. Marano, A. Ho, Sars-cov-2, Sars-cov-1 and mers-cov viral load dynamics, duration of viral shedding and infectiousness - a living systematic review and meta-analysis, The Lancet Microbe 2 (1) (2021) e13–e22.
[5] S. Palmore, S. Martin, Public Health Emergencies and Human Rights: Problematic Jurisprudence Arising From the COVID-19 Pandemic, Social Science Electronic Publishing, 2020.
[6] K. Sulek, Innovation in the time of covid-19, Mech. Eng. 142 (8) (2020) 38–43.
[7] J.S. Shapiro, F. Mottashabi, G. Hirspasck, N. Soulakis, G. Kuperman, Using health information exchange to improve public health, Am. J. Public Health 101 (4) (2016) 616–623.
[8] T. Diseases, The covid-19 infodemic, Lancet Infect. Dis. 20 (8) (2020).
[9] X. Chen, R. Quan, A spatiotemporal analysis of urban resilience to the COVID-19 pandemic in the Yangtze River Delta, Nat Hazards 106 (2021) 829–854.
[10] A. Maa, A. Bp, M.A. Long, B. St, A. Mk, A. Pvm, Comparing the accuracy of several network-based covid-19 prediction algorithms - sciedirect, Int. J. Forecast. (2020).
[11] L. Lu, J. Peng, J. Wu, et al., Perceived impact of the Covid-19 crisis on SMEs in different industry sectors: Evidence from Sichuan, China, International Journal of Disaster Risk Reduction 55 (2021) 102085, 10223.
[12] V. Balachandar, B. Giri, COVID-19: a promising cure for the global panic, Sci. Total Environ. 725 (1) (2020) 138277.
[13] M. Shoji, S. Cato, T. Iida, K. Ishida, K.M. Mcelwain, The bright and dark sides of social media usage during the covid-19 pandemic; survey evidence from Japan, International Journal of Disaster Risk Reduction 54 (2) (2021).
[14] A. Xz, A. Qz, Z. Rui, E. Mzd, People-centered early warning systems in China: a bibliometric analysis of policy documents, International Journal of Disaster Risk Reduction 51 (2020).
[15] A. Lin, H. Wu, G.H. Liang, et al., A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster, Int. J. Disast. Risk Reduct. 49 (2020).
[16] C. Huang, Y. Wang, X. Li, L. Ren, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in wuhan, China, Lancet 395 (2020), 10223.
[17] J.S. Jia, X. Lu, Yu. Yuan, G. Xu, N.A. Christakis, Population flow drives spatiotemporal distribution of COVID-19 in China, Nature 582 (7801) (2020) 1–11.
[18] S. Behl, A. Rao, S. Aggarwal, S. Chaddha, H.S. Pannu, Twitter for disaster relief through sentiment analysis for covid-19 and natural hazard crises, International Journal of Disaster Risk Reduction (1) (2021) 102105.
[19] S. Da, J. Wang, H. Zhang, W. Cai, Z. Kang, T. Yang, et al., Predicting COVID-19 Using Hybrid AI Model, Social Science Electronic Publishing, 2020.
[20] P.C. Lins-Filho, M.M.S.D. Araújo, T.S.D. Macedo, A. Ferreira, A. Jr, Assessing the quality and readability of online information on covid-19, Research and Society Development 9 (12) (2020), e359120680.
[21] M. Imran, M.Y. Wu, E. Bee, M.J. Khan, Mathematical modelling of sir for COVID-19 forecasting, Rev. Argent. Clin. Psic. 30 (1) (2021) 218–226.
[22] F.W. Wilwoos, Whayati, Prediction modelling of covid-19 outbreak in Indonesia using a logistic regression model, J. Phys. Conf. 1803 (1) (2021) 8, 012015.
[23] A. Bđulescu, C. Williams, K. Cavanaug, Management strategies in a SER-type model of COVID-19 community spread, Scientific Rep. 10 (2020) 21256.
[24] N. Barda, R. Dan, A. Akriv, J. Levy, N. Dagan, Developing a COVID-19 mortality risk prediction model when individual-level data are not available, Nat. Commun. 11 (1) (2020) 4439.
[25] Q.V. Pham, D.C. Nguyen, Y. Fohnb-Thr, W.J. Hwang, P.N. Pathirana, Artificial intelligence (ai) and big data for coronavirus (COVID-19) pandemic: a survey on the state-of-the-arts, IEEE Access (99) (2020), 1–1.
[26] H. Cao, Workpiece quality prediction research based on multi-source heterogeneous industrial big data, in: International Conference on Big Data Analytics for Cyber-Physical-Systems, Springer, Singapore, 2020.
[27] Q. Pan, S. Wang, J. Zhang, Prediction of Alzheimer’s disease based on bidirectional lstm, J. Phys. Conf. 1187 (5) (2019), 052030.
[28] T. Sherry, A. Shehzad, B. Gilbert, H. Nadya, B. Shaha, B. Eilmar, et al., Mass media and the contagion of fear: the case of ebola in America, PloS One 10 (6) (2015), e0129179.
[29] Da-Cang, Jin-Feng Wang, Monitoring hand, foot and mouth disease by combining search engine query data and meteorological factors, Sci. Total Environ. 612 (2018) 1293–1295.
[30] A. R, M. A, P.B. B, Identifying the predictors of covid-19 infection outcomes and development of prediction models, Journal of Infection and Public Health (8) (2020).
[31] K. Leung, J.T. Wu, G.M. Leung, Real-time tracking and prediction of covid-19 infection using digital proxies of population mobility and mixing, Nat. Commun. 12 (15015) (2021) 1–8.
[32] A. Liu, L.H. Qi, Q.A. Rui, A. Yi, L.B. Yin, Prediction-based analysis on power consumption gap under long-term emergency: a case in China under covid-19, Appl. Energy 283 (2020) 116339.
[33] I. Ghosh, T.D. Chaudhuri, Feb-stacking and feb-dm models for stock trend prediction: a performance analysis for pre and post covid-19 periods, Decision Making Applications in Management and Engineering 4 (1) (2021) 51–86.
[34] K. Zhou, Y. Sun, L. Li, Z. Zang, J. Wang, J. Li, et al., Eleven routine clinical features predict covid-19 severity uncovered by machine learning of longitudinal data, The Lancet Digital Health 1 (2019) 3640–3648.
[35] F.F. Prado, L.A. Digiampietri, A Systematic Review of Automated Feature Selection: a Data Perspective, ACM Comput. Surv. 50 (6) (2016).
[36] S.K. Mettler, J. Kim, M.H. Maathuis, Diagnostic serial interval as a novel indicator for contingency procedures to follow in the case of emergencies involving people with alzheimer’s disease, Int. J. Environ. Res. Public Health 17 (10) (2020) 5398.
[37] C. Huang, Y. Wang, X. Li, L. Ren, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in wuhan, China, Lancet 395 (2020), 10223.
[38] J.A. Agyme, K.L. Chin, C.B. Landersdorfer, D. Lee, R. Ofer-Aizenso, Smell and taste dysfunction in patients with covid-19: a systematic review and meta-analysis, Mayo Clin. Proc. 95 (8) (2020).
[39] J. Grauer, H. Lwen, B. Liebchen, Strategic Spatiotemporal Vaccine Distribution Framework for the Survival of Infectious Disease Based on Covid 19, International Journal of Disaster Risk Reduction 51 (2020).
[40] M.Y. Kim, H. Cheung, H.S. Kim, Proposal of the Assembly Guideline for Infectious Diseases: Preparation for the Post-covid-19 Era, 2020.
A population-based cohort study of socio-demographic risk factors for covid-19 deaths in Sweden, Nat. Commun. 11 (2020) 5097.

Risk factors for mortality of coronavirus disease-2019 (covid-19) patients in two centers of Hubei province, China: a retrospective analysis, PloS One 16 (2021).

The pca-seq method applied to analyze of the dynamics of covid-19 epidemic indicators, J. Phys. Conf. 1715 (1) (2021) 6, 012025.

Positive externalities of an epidemic: the case of the coronavirus (covid-19) in China, J. Med. Virol. (2020) 1376–1379.

Introduction to Urban Dynamics, 1976.

Teaching strategic thinking using System Dynamics: lessons from a strategic development course, Syst. Dynam. Rev. 28 (1) (2012) 28–45.

The PCA-Seq method applied to analyze of the dynamics of covid-19 epidemic indicators, J. Phys. Conf. 1715 (1) (2021) 6, 012025.

Management strategies in a SEIR-type model of covid-19 community spread, arXiv: Physics and Society (2020).

Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, Lancet (London, England) 395 (2020) 689–697, 10225.

The effectiveness of quarantine of wuhan city against the corona virus disease 2019 (covid-19): a well-mixed seir model analysis, J. Med. Virol. 92 (7) (2020) 841–848.

Accounting for symptomatic and asymptomatic in a SEIR-type model of COVID-19, arXiv: Population and Evolution (2020).

A New Seaird Pandemic Prediction Model with Clinical and Epidemiological Data Analysis on Covid-19 Outbreak Isolated People in Incubation, Applied Intelligence, 2020, pp. 1–37. IP (IP).

Forecast and analysis of covid-19 epidemic based on improved seir model, J. Phys. Conf. 1802 (4) (2021), 042050.

Mathematical Modeling for Prediction Dynamics of the Coronavirus Disease 2019 (COVID-19) Pandemic, Quarantine Control Measures, 2020.

How detection ranges and usage stops impact digital contact tracing effectiveness for covid-19, Sci. Rep. 11 (9414) (2021) 1–11.

Forecasting the Worldwide Spread of COVID-19 Based on Logistic Model and SEIR Model, 2020.

Management strategies in a SEIR-type model of covid-19 community spread, arXiv: Physics and Society (2020).

How detection ranges and usage stops impact digital contact tracing effectiveness for covid-19, Sci. Rep. 11 (9414) (2021) 1–11.

Forecasting the Worldwide Spread of COVID-19 Based on Logistic Model and SEIR Model, 2020.

Management strategies in a SEIR-type model of covid-19 community spread, arXiv: Physics and Society (2020).

Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, Lancet (London, England) 395 (2020) 689–697, 10225.