A Systematic Discussion of the Main Epidemic Prediction Models for the Spreading of COVID-19

Yiming Guo

Ealing Internation School, Dalian, China

1169671063@qq.com

Abstract. Currently, many countries around the world are facing the crisis of COVID-19. Most of them cannot solve the problem of the spread of COVID-19. This study aims to analyze the four main pandemics' prediction models through their advantages, shortages, the basic structures, appropriate scene as well as the improvement methods to ameliorate the whole statistical system of COVID-19. Today, the researchers have widely criticized the means and medium of data col-lection for the sake of pandemic spreading. Although the government has undertaken steps on the control of COVID-19, there were always many positive cases appeared at one time so that the viruses spread widely. In summary, this study offers new in sight for arranging the pandemic's prediction method against the potential of the outbreak of all the epidemics.

Keywords: COVID-19, Statistics model, Prediction of pandemics

1. Introduction

The seriousness and spread of infectious diseases with pandemic potential has occurred worldwide for quite a long period throughout history. The COVID-19 has become the most serious one nowadays. According to the USA Today, the total number of deaths in the USA in 2021 until seventh December has surpassed 3,800,000 [1]. How it reached such a value? First of all, the COVID-19 has a certain incubation period, in which people do not have any symptoms, and the incubation period is somewhat contagious [2]. Secondly, some patients are asymptomatic, and these patients can cause family-intensive transmission. Finally, the coronavirus is spread by respiratory droplets, by close contact, and also by the patient's secretions, such as aerosols, urine, faeces, etc. [3]. With a wide range of transmission, it is easy to cause widespread disease [4]. Otherwise, the coronavirus can also be transmitted through the remains of an infected person. All of them make it hard to control. More importantly, there have been various kinds of variations. The coronavirus has a high rate of propagation and it has a poor genetic stability. In order to survive in human's bodies, they have a large amount of variations.

Some major pandemics like smallpox virus in pre-medieval Europe, Ebola virus epidemic in West Africa and rabies, HIV have affected humanity seriously since the naissance of human birth. In the early human civilization, as long as someone is infected by any of pandemics, it could be impossible for them to be treated well and even most of them would die at a young age. However, under the pressure of those pandemics, human made many efforts on biology, medicine as well as statistics, as a result, the diseases have chances to be controlled or eliminated. On the one hand, due to the study on biology and pharmacology, people can create the improved environmental conditions via some disinfections and

some medicine has already made to improve the immune system in human body to counterbalance the virus. On the other hand, some statistical methods are developed by some experts and have been used in the current pandemic. For example, the number of SARS positive cases per day was used by statisticians globally to predict the total number of the positive cases in the future. Those forecasts had helped governments to make informed decisions and timely policies to control the spread of the pandemic, which was quite effective. Some statisticians also use models to predict the potential infectious patients range via the description about the past tract of the positive patients to isolate them.

During the epidemic prediction programme, two major parts would be set: the introduction of the main prediction's models and the compatibility of those models. In the first part, the specific properties of the models would be discussed, including their limitations, advantages, basic formula and so on. The study would start with the development of the epidemiology. It aims to offer a more completely recommendation of the statistical models of transmissible diseases. Otherwise, there are many papers discussed the models of the spread of COVID-19, and they get the corresponding results of their discussions. This text would cite and compare their results. In the second part, according to their properties, suitability and so on, they will be distributed more subdivision in order to analyze the infected people more accurately, the passage would consider the properties of COVID-19 to find the adaptation of different models and discuss the improvement methods. The reasons for why the pandemics always took place suddenly with large number of infectibles would also be researched.

2. Literature review

The concept of modeling COVID-19 was investigated quite intensively in recent years. There are several divisions when it comes to the stance of arguments surrounding the scope of spread of the epidemic and its infectiousness. The predecessors have made four main models for the traditional epidemic prediction, they are the spatial factor model, the nonlinearity/time space dynamics model, the heterogeneity in population and its implications on the spread of infectious diseases, the estimation of syndrome. In 2020, Rendana wrote the specific spatial model for COVID-19, he mainly used the geographical information system (GIS) model to map the positive cases and epidemic rates. Taking a review of those previous paper which investigated the infectious diseases' prediction models, this paper will summarize the questions they discussed, their objectives and the results. For "Realistic population dynamics in epidemiological models: the impact of population declines on the dynamics of childhood infectious diseases. Measles in Italy as an example" writing [5], due to the drawbacks of the original BRF populations model, they developed a more suitable method, it considered three groups Default, Diagonal, Proportionate which had completely distinct trends [6].

The statistical models we discussed in this study were under the background of COVID-19, which is a kind of fierce and transmissible disease. For an illustrative purpose, the prediction of the overall transmissible diseases has been previously studied: In 1996, Isham chose five main themes to conclude the main methods of predicting serious infectious diseases [7]. They are transmissible diseases with long development times and vaccination strategies, dynamics of immunity, population heterogeneity, consequences of treatment interventions and prediction. In the previous study, the predecessors also introduced the concept of heterogeneity to make the model more accurate. The heterogeneous population is divided into subpopulations or groups, each of which is homogeneous in the sense that the group members have similar characteristics. This division into groups can be based not on various characteristics to make the prediction more completely [8]. Then, about this passage's significance, it firstly limits the scope of research to the COVID-19. Secondly, on the basis of the models proposed by the previous study, the passage analyzed the properties of some main prediction models and discussed their adaptability under different conditions. At the same time, the proper improvements were written in the text.

In regards to the traditional models we used to predict the isolated area and infectibles, several problems appeared during the prediction period. The problems can be mainly divided into two aspects: the first one is the process of collecting data. For instance, the spread of pandemics of most cities started in the suburb. This phenomenon was caused by the lack of management and data collecting. There were

no guards to check people's health code, so their health condition was ignored when entered a new area. Also, the local governments were irresponsible for the information collecting. These factors cannot be improved by us. So, this passage mainly discussed the second point: the systematic problems of the models. We started with two traditional statistical models: spatial model and nonlinearity/time space dynamics. Spatial model can predict the number of infectibles by finding people passed through some main areas the positive cases had been to. The nonlinearity model assumes a theoretical route of the positive cases when passing through a certain area to predict the amount infectibles. According to their properties, the spatial model is more suitable for the medium-scaled area's prediction and the nonlinearity model can be used in the small area's prediction. But these models must ensure all the people within the area followed the route they defined, people have different characteristics. So those two models are inaccurate. In order to solve the question, the heterogeneity model was introduced. This model can use the data of different people's travel habits to make a graph. By the property of normal distribution, the isolated area can be determined by calculating the deviation.

3. Discussion

Nowadays, there have been recently several infectious diseases appeared worldwide. In order to analyze the spread of it and find the best control methods, four base models are developed: spatial factor, nonlinearity/time space dynamics, heterogeneity in population and its implication on the spread of infectious diseases as well as the estimation of syndrome. In regards to the history of the improvement of different statistical methods, the nonlinearity model and the spatial model were wide-spreadly used in the early stage. Those two models have their own advantages, for the spatial factor model, it is suitable for small area's predictions in short period, it concludes the places where the patients contact the source of infections as well as the areas they had been to during the incubation period. However, although the nonlinearity/time space model is less accurate for the small region's analysis, it can efficiently find the spread of the outburst epidemics. The content of it not only includes the different variables of the diseases, but also covers people's body conditions and the population characteristics, like the regional group, the age group, the income group and so on. The presentation of the symptoms is also shown in it, like wearing masks, going to the hospital quarantine at home. People with different characteristics do different things, and they have all been covered in this model. These two models are all made for the sudden illness, for some long-standing pandemics, like the seasonal flu, the fixed period allergy, they waste quite a lot of time for their complexities on collecting and sorting out data. As a result, some experts decided to develop a model which can make those seasonal pandemics easier to be predicted. So, the heterogeneity in population on the infectious diseases model appeared, most of the diseases nowadays used this model and it has been the most influential statistical model in the field of infectious diseases. Talking about the components of it, the exact positions where different people visited, the parasite rate, the population size and the epidemic distance are all included in it. David [9] referred to the factors embedding epidemic modelling that: most epidemic theory for infectious diseases has been developed under the assumptions that all individuals are alike and that populations are well mixed. Epidemiological theory often stands in sharp contrast to studies that emphasize the role of compositional and configurational heterogeneity. Obviously, heterogeneity does affect who gets sick and the population dynamics of infectious diseases over time and space. All in all, it is clear to see that the improvement of heterogeneity population model efficiently helps control various diseases. Finally, it is the estimation of syndrome, some experts used it to predict the mortality and incidence rate of the pandemics, at the same time, this kind of model helps discuss the sequelae and the transmission characteristics after the illnesses.

3.1. *Spatial modelling*

Frist of all, it is the part of the introduction of the most basic spatial model: the chain binomial model. This is the set of models which based on the chain frequencies, also it can be described as a kind of mass action model, the aggregated data are used to deal with the regional even the international spatial units. In the book of mathematical models for society and biology 2013, we suppose t=1, 2, etc. Actually, it

can be described as different points corresponding to the serial time interval of the certain diseases. Most importantly, the method must ensure the patients had been isolated at home after the appearance of syndrome, and there was no more than one pandemic taken place. After the pandemics outbroke in a district (at the beginning of t), the total population in this district can be denoted by N_t, which contains I_t infected people, S_t susceptibles and R_t removals (the population died during the period).

$$Nt = St + It + Rt \qquad (1)$$

Spatial modelling envisages and abstracts the infectible-susceptible contact scene as following: In this case, the population entered the district in a certain t can be defined as A_t, people left the district are called D_t, and X_t can then be described as the number of the new cases during the period of t, $\alpha_t A_t$ as the infectives among the arrivals at t, and $\delta_t D_t$ as the infectives among the departures at t [10]. The following two equations can be found:

$$Nt + 1 = Nt + At - Dt \qquad (2)$$
$$It + 1 = It + Xt - (Rt + 1 - Rt) + \alpha tAt - \delta tDt \qquad (3)$$

3.2. Nonlinearity/time dynamics model

In this part, the basic model would also be displayed at first. It is just the red squirrel model. The book mathematical models for society and biology 2013 would also be introduced as the reference. The basic formula is:

$$S + I + R = 1 \qquad (4)$$

S, I, R each represents susceptibles, infectibles and recovered people. The formula can be extended with the spatial spread of a disease. Combined with the case of the propagation of rabies by infected foxes, which is given that the populations are healthy and susceptibles. The spread of rabies by infected foxes formed a longitudinal wave, x can be considered as the total squirrels in the forest, q is the wave speed, t is the total time after the infected foxes passed through the forest. So, the total healthy squirrels z satisfies:
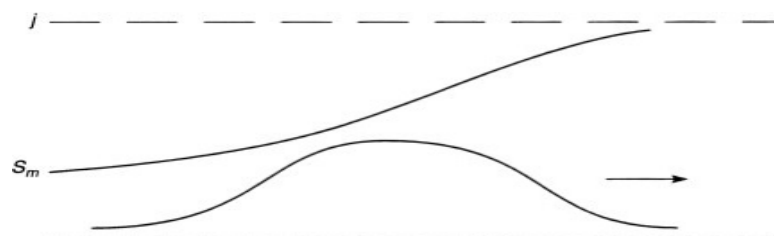
$$z = x - qt \qquad (5)$$

Figure 1. A pulse wave of an infective traveling through a forest where there is fixed number of susceptibles

As a result, the graph can be made (figure 1), the lower curve is the new infectibles by the infected fox's footprint when it passed through the forest. And Sm represents the general population growth of squirrels in the forest. From the diagram, the number of the remaining healthy squirrels can be calculated by subtracting the minimum value between the derivatives of two curves [11].

However, this model contains a serious limitation, according to the case happened in Cheng Du China, the initial positive cases in Chengdu are the grandfather, grandmother and their granddaughter in one family. This woman did not know she had been infected and the big data showed her track during the latest fourteen days. The graph illustrated that she went to 4-night clubs during three days, and infected around 3000 people. Before this positive case had been found, the local authorities used the model which assumed that all the positive cases have the similar tracks. As a result, Cheng Du did not control the spread of viruses and the positive cases increased dramatically at that period. The cases like this cannot be solved by the red squirrel model, the presupposes of this model are the infected samples have the approximate track, the footprint of infectibles are basically followed most idealized one, which means the variance and standard deviation are same.

3.3. *Heterogeneity on population and its implication in infectious diseases*

In the last aforementioned paragraph, the red squirrel model has been discussed, it has been pointed out that the inaccuracy would increase dramatically with the increase of the population. Under this situation, a new model has been developed, it is the stochastic process. In the real cases, the individual differences are notable, and it can seriously affect the accuracy of the prediction of a certain infectious disease. In the stochastic processes [12], Maria defined this model as the mathematical representation of a system which evolves probabilistically, it is a collection of random variables X(t), where the index t and the random variable X can be discrete or continuous. The joint equation Chapman-Kolmogorov is used in this model:

$$p(x_1, t_1; \dots; x_{n-1}, t_{n-1}) = \int p(x_1, t_1; \dots; x_n, t_n)\, dx_n \tag{6}$$

It is clear to see that the discrete variables can be calculated by integral as a sum in this equation. The model equations can be grouped into two main groups: one is used to solve the representation of the trajectories of the process [13]. The random variable X(t) are described by the stochastic differential equations. The second one describes：

$$p(x, t) = \text{Prob}\{x(t) - x\} \tag{7}$$

which are the deterministic differential equations, like the Chapman-Kolmogorov equation in the differential form.

Talking about its five main properties:

3.3.1. Independence

The first one is independence: the simplest class in the process can be determined as the independent group X(t), which does not depend on any other group X(s), s<t. It satisfies that:

$$p(x_1, t_1; \dots; x_n, t_n) = p(x_1, t_1)p(x_2, t_2) \dots p(x_n, t_n) \tag{8}$$

3.3.2. Markov

This property defines that the prediction of the future needs the knowledge or data in the present. In other words, in the stochastic process, if there is an arbitrary time, $t_1 < t_2 < \dots < t_n$, the conditional probabilities reduces to a transition probabilities and an initial probabilities in the last two states.

$$p(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = p(x_n, t_n | x_{n-1}, t_{n-1}) \tag{9}$$

Then any joint probabilities can be described as the product of the initial probability and transition probability.

$$p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = p(x_n, t_n | x_{n-1}, t_{n-1}) \dots p(x_2, t_2 | x_1, t_1)p(x_1, t_1) \tag{10}$$

3.3.3. Stationarity

In the stochastic process, if the joint probability densities are constant under the time translations, it can be described as strongly stationary.

$$p(x_1, t_1; \dots; x_n, t_n) = p(x_1, t_1 + \Delta t; \dots; x_n, t_n + \Delta t) \tag{11}$$

Wide-spread and weal stationarity requires only instead the invariance below time translation of the mean and covariants of the process.

$$E[X(t)] = \mu X(t) = \mu X(t + \Delta t) \tag{12}$$

$$\text{Cov}[X(t_1), X(t_2)] = Cx(t_1, t_2) = Cx(t_1 - t_2, 0) \tag{13}$$

3.3.4. Continuity

The process is continuous if each of its sample paths a continuous t function. For this Markov process, this rule specifies as uniformly in z, t and Δt.

$$\forall \varepsilon > 0, \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z| > \varepsilon} p(x, t + \Delta t | z, t)\, dx = 0 \tag{14}$$

3.3.5. Guassianity

If a process is Gaussian, it's every joint density is a multivariate Gaussian. Then this process is specified by its mean E[X(t)] and covariants E[X(t), X(s)] functions.

3.3.6. Markov processes

Nowadays, many mathematical models always make some simplifying assumptions about the properties of the phenomena which has been modeled. But most of these assumptions are invalid, however, they frequently make a compromise between its accuracy and the treatability. The Markov property is the model which added the stochastic processes in it [14]. When the stochastic processes under consideration has the property of Markov, the Chapman-Kolmogorov equivalent can be expressed as an identity on the transition probability:

$$p(x_3, t_3|x_1, t_1) = \int p(x_3, t_3|x_2, t_2)p(x_2, t_2|x_1, t_1) \, dx_2 \qquad (15)$$

For this equation, it can be described as the transition probability of a transition from x_1 to x_3 which can be divided into the sum of probabilities of the transitions from x_1 to any intermediate x_2 and from x_2 to x_3[1].

In all cases, we can obtain by the equation mentioned before for the transition probabilities. In order to simplify the equation, the assumption of fixed and sharp initial condition (x_0, t_0), so that the unconditioned probability (x,t) can be replaced by the probability of transition from(x_0,t_0):

$$p(x, t) = p(x, t|x_0, t_0) \qquad (16)$$

3.4. *Compatibility of models*

Model screening for COVID-19 should be limited to the ordinary cases instead of discussing the wide-spread cases caused by the special individual. In many environmental epidemiology studies, the locations and times of exposure measurements and health assessments do not match [15]. The following content would mainly discuss two aspects: disease characteristics (the influences on people's body by the viruses) and the suitable models under different ranges of area.

About the characteristics of the COVID-19, the symptoms of the illness could be introduced. The results of characteristics of COVID-19 infection in Beijing by Tian [16] showed that the most common symptoms of illness onset were fever, cough, fatigue, dyspnea and headache. Otherwise, the COVID-19 can weaken the immune system. For the coronavirus, WHO [17] said that it has a mortality of 2-3% and a transmissibility rate 2-3. They caused the rapid spread of this virus around the world. By contrast with other serious pandemics, the coronavirus can exist in the environment for a longer period, because the virus can adapt to room temperature. As a result, people in the environment where the viruses existed can be easily infected. More seriously, the virus has a higher mutation rate, several branches virus has a higher transmissibility and mortality rate, like the Omicron, Delta and so on.

In regards to the properties of COVID-19 which has been discussed above, determining the geographical scope and the size of population can be considered as an essential evidence to choose the most suitable model to finish the prediction of COVID-19.

3.5. *Small-scaled area of COVID-19 suspicious circulation and infection*

Considering about the small area prediction, the most important part should be the approximate tracks of the positive cases rather than all the positions they had been to after they were diagnosed. The small-scaled areas are unordered, the exact life footprints cannot be found. So, in order to predict the spread of the virus, the nonlinearity time dynamics model should be used. From its graph, some major positions and the time they arrived are shown, this model do not need you to know what the positive cases do during the certain positions, it only considered the time he or she reached, then according to the infection radius of COVID-19, the isolated area can be defined. In the small- scaled area, COVID-19 has the characteristics of fast infection speed, large range of infection. In the light of those properties and steps, the government can pay attention to those people in the calculated range and isolate them or let them do nucleic acid test.

In shed of light regarding the aforementioned, the nonlinearity model has several advantages compared with other models if put to use for small-scaled infection prediction. The purpose of

---

[1]*Note: the integral notation in the equation is wide-sense: for the discrete state space, it represents the summation, while for the discontinuous processes this notation includes both integration on the continuity intervals as well as the summation on single discontinuity points.*

monitoring the trajectory is to get a response faster, as a result, the isolated area and susceptible can be defined more quickly. And the nonlinearity model prediction does not require so many human and material resources, so it can be more suitable for the small-scaled area prediction. This model always leads to some more flexible governmental decisions for its high efficiency. Although it has several strong points, the model should also consider the region. For instance, China is a country implements the regional management, so that if the positive cases appeared in the edge of the area, the manager would not report the real conditions because of the stagnation of economy may be brought by the pandemics. So only if the government combine the regional policies into the model, can the nonlinearity model finish the small area prediction better.

About the nonlinearity time dynamics model, combined with the case of the propogation of COVID-19 in a supermarket, which is given that the populations are healthy and susceptibles. The spread of COVID-19 forms a longitudinal wave, x is the total populations in the supermarket, q is the wave speed and t is the total time after the infected people passed through the supermarket. So, the total healthy people in the supermarket is $z = x - qt$. In regards to the properties of COVID-19, qt is the main variable. In the formula, q is the gradient of the wave, which represents the maximum passing number in the unit time and unit area. To simulate the maximum flow in the supermarket, a point can be firstly signed, by monitor and equipment detection, the daily average maximum population through the point can be calculated as the median of the values they measured. Then the supervisors can limit the size of the daily average maximum population to limit the people pass through the supermarket at the same time. For example, Beijing defined 60% of the average daily maximum population as the maximum number in the supermarket simultaneously to reduce the number of infected cases. All in all, it is because the supermarket is a small-scaled area, the tracks of different people are almost same, this model can be used. But different people's characteristics would not be shown in this model, it could bring some errors.

3.6. *Medium-scaled area of COVID-19 suspicious circulation and infection*

For those medium-scaled area of COVID-19 suspicious circulation and infection, the spatial factor model is the most accurate model. Spatial factor model can be used in the certain conditions when the accurate infection time cannot be defined and the activities areas of infectibles are large enough or connected with other areas. As a result, the spatial model can measure the number of infectibles, removals and departures to define the approximate total number of susceptibles. Then, the daily tracks can be found to define the isolated area. Talking about the advantages of spatial model, it can limit the infection area accurately and effectively. Otherwise, it cannot determine the original cases, but the infectibles are easy to ascertain under this model.

For the medium-scaled area, taking a district as the infected area. At the beginning of t (time after the positive case entered the district), the total population can be denoted by N_t, which contains I_t infected people, S_t, susceptibles and R_t, removals.

$$N_t = S_t + I_t + R_t \tag{17}$$

Then, defining At as arrivals and Dt departures. In the end, Xt can be described as the number of the new cases during the period of t, αtAt as the infectives among the arrivals, and δtDt as the infectives among the departures. The following two equations can be found:

$$N_{t+1} = N_t + A_t - D_t \tag{18}$$
$$I_{t+1} = I_t + X_t - (R_{t+1} - R_t) + \alpha_t A_t - \delta_t D_t \tag{19}$$

The two values α_t and δ_t represents the infected rates among two groups of people. In order to improve the accuracy of the measurements of positive cases, the D_t as well as A_t should decrease to zero, in other words, the district should prohibit people in and out. Then the influences of α_t and δt would decrease.

3.7. *Heterogeneity modelling and its features*

Under the epidemic of COVID-19, spatial factor model and nonlinearity model have several limitations: they cannot control the variable of people's different characteristics, nonlinearity model must ensure all the people in the area follow its fixed routes. Not only that, both of the two models are inaccuracy. So, with the aforementioned under consideration, heterogeneity models emphasize the transferring of

information of different people's behavior habit in a certain area and simulate them on the computers. Unlike those two models mentioned before, this model does not study through the individual infection.

The heterogeneity model always uses the chronicle analysis. For instance, many positive cases suddenly appeared in a city, which the pandemics never outbroke in it. For this case, it has been too late for no specific travel data, the model can use the information collected by memory about their past tracks of the residents and some record on the monitoring system to find the approximate population needs to do the nuclei acid track.

Another measurement method is researching the inclusion of individual differences. Because different people have different behavior or track habits, the model should find the best way which covers majority of people when passing a certain area. It can finish this task by analyzing all the people's track habit in the certain area. By using the normal distribution, the researchers can put the data it analyzed (median, standard deviation and so on) into the graph. Then, the value of deviation from the best route can be found by using the figures between two standard deviations which contains about 97 percent of people. So, the covered area can be calculated by multiplying the deviation value and length of the road.

4. Conclusion

In this passage, the main objective is analyzing the usability of some basic statistical models under the background of COVID-19 and applying them in the subdivision of the epidemic area. Throughout the full text, the research question is how to use the models correctly to prevent the pandemics spreading wide range and save costs on the wrong susceptible. By reading other dissertations, the four models were illustrated briefly. Then those models were introduced in details: the chain binomial model, the red squirrel model and the stochastic processes were taken to explain the basic structures and usages of the spatial model, nonlinearity model and heterogeneity model. After that, the compatibility of those models was considered by the size of the monitoring area. In the end, the finding is: for those small-scaled area, the nonlinearity model can be used because it can response more quickly, with regard to the medium-scaled area, the spatial factor model is more effective, it can define the isolated area more accurately by ascertaining the space which the infectible had passed through. But the pandemics like COVID-19 has a long remaining time and fast variation rate, and different people nowadays have different travel habits, the heterogeneity model can resolve the limitations of those two traditional models by finding the areas which are suitable for majority of people. Through these discussions, only if the models are correctly using, the costs can be saved and achieve the goal of catching all of the positive cases.

Nowadays, although we can analyze the suitable statistical models under different conditions, there are also many places cannot control the sudden pandemics. Most times, large number of cases would outbreak at a short period. One of the limitations of my current study is I focused on the theoretical study too much. This study is a systematic review passage, it based on the premise of epidemiology, in the future, we can take out one part of it for analysis, like the specific discussion of the nonlinearity model. In the text, the questionnaires can be made, the questions can be set according to the measurement scales made by predecessors. Then, we can form our new models, and do the regression to improve the parameters. On the other aspect, the data acquiring should be more averagely. For example, in America, the crime rate is dramatically higher in Ghetto than the rich area. Later, by the further research, people found the result is wrong, the conclusion is problematical because the data collectors believed the poor had more inclination of committing crimes, so the data was not average. Back to the disease science, suburb always has no infectible ever, but the prosperous areas seals off many times. The travel code is always useless in those remote areas. Those problems should be discussed in the prospective. Keeping seeking more variables correlated to different pandemics and achieving equal sampling plays an important role in the future of pandemic statistic.

References

[1]   *J., Rocklöv, H., Sjödin, (2020), High population densities catalyse the spread of COVID-19, Journal of Travel Medicine, https://doi.org/10.1093/jtm/taaa038.*

[2]   *A., Gryparis, C., J., Paciorek, (2009), Measurement error caused by spatial misalignment in environmental epidemiology, Biostatistics, 10.1093/biostatistics/kxn033.*

[3]   *National Health Commission of the People＇s Republic of China (NHC PRC), Daily briefing on novel coronavirus cases in China. Beijing: NHC PRC, (2020), http:// en.nhc.gov.cn/2020-02/07/c_76323.htm*

[4]   *Adam J Kucharski, etc., (2020), Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study, National Library of Medicine, 10.1016/S1473-3099(20)30457-6.*

[5]   *N., S., Best, Richardson, A., Thomson, (2005), A comparison of Bayesian spatial models for disease mapping, Edward Arnold, 10.1191/0962280205sm388oaw.*

[6]   *Daniel Zeng, Zhidong Cao and Daniel B. Neill, (2021), Artificial intelligence-enabled public health surveillance-from local detection to global epidemic monitoring and control, Elsevier, https://doi.org/10.1016/B978-0-12-821259-2.00022-3.*

[7]   *R., Benedici, E., Bilotta & P., Pantano, (2021), A unifying nonlinear probabilistic epidemic model in space and time, Scientific Reports.*

[8]   *H., W., Hethcote, (1978), Modeling heterogeneity (mixing), V., Isham & G., Medley, Models for Infectious Human Diseases, 215.239, Cambridge, Issac Newton Institute for Mathmatical Science.*

[9]   *World Health Organization (WHO). Coronavirus disease (COVID-19) outbreak situation, (2020), https://www.who.int/emergencies/ diseases/novel-coronavirus-2019.*

[10]   *A., Cliff & P., Hagget, (2005), Viral Outbreak and Blood Clots, E., Beltrami, Mathmatical Models for Society and Biology, Academic Press, https://doi.org/10.1016/B978-0-12-404624-5.05001-9.*

[11]   *A., Cliff & P., Hagget, (2005), Red Tides and What Happened to the Red Squirrel, E., Beltrami, Mathmatical Models forSociety and Biology, 187, Academic Press, https://doi.org/10.1016/B978-0-12-404624-5.05001-9.*

[12]   *M., F., Carfora, 2018, Stochastic Processes, 1.2, Elsevier, 10.1016/B978-0-12-809633-8.20362-2.*

[13]   *W., J., Ewens, G., Grant, (2005). Statistical Methods in Bioinformatics: An Introduction, second ed. Springer.*

[14]   *D., R., Cox, H.D., Miller, (2001). The Theory of Stochastic Processes. Boca Raton, FL: Chapman and Hall/CRC.*

[15]   *M., Rendana, W.M.R., Idris, S.A., Rahim, (2021), Spatial distribution of COVID-19 cases, epidemic spread rate, spatial pattern, and its correlation with meteorological factors during the first to the second waves, Elsevier.*

[16]   *National Health Commission of the People＇s Republic of China (NHC PRC). Timeline of China＇s fight against the novel coronavirus. Beijing: NHC PRC; (2020), http://en.nhc.gov.cn/2020-03/20/c_78021.htm*

[17]   *World Health Organization (WHO). Infection prevention and control during health care when novel coronavirus (nCoV) infection is suspected. Interim guidance. Geneva: WHO. (2020), https://www.who.int/ publications-detail/infection-prevention-and-control-during- health-care-when-novel-coronavirus-(ncov)-infection-is- suspected-20200125.*