

# Trajectory Tracking of COVID-19 Epidemic Risk Using Self-organizing Feature Map

CHEN Ning<sup>1</sup>, CHEN An<sup>2,3\*</sup>, YAO Xiaohui<sup>1</sup>

<sup>1</sup> Institute of Urban Safety and Environmental Science, Beijing Academy of Science and Technology, P. R. China

<sup>2</sup> Institutes of Science and Development, Chinese Academy of Sciences, Beijing, P. R. China

<sup>3</sup> University of Chinese Academy of Sciences, Beijing, P. R. China

## Abstract

The ongoing COVID-19 has become a worldwide pandemic with increasing confirmed cases and deaths across the globe. By July 2022, the number of cumulative confirmed cases reported to the World Health Organization (WHO) has risen to 550 million, with more than 6 million deaths in total. The analysis of its epidemic risk remains the focus of attention all over the world for a long time. The Self-organizing feature map (SOM), a vector quantization method, offers a data mapping approach to tracking the response of time series data on a well-trained map. This study aims at a trajectory tracking of COVID-19 epidemic risk in 237 countries measured by the number of new confirmed cases and deaths per day for over one year. A hybrid clustering method uses SOM and K-means to generate a risk map and then displays the trajectory of daily risk on the map. The experimental results demonstrate the promising functionality of SOM for trajectory tracking and give experts insights into the dynamic changes of COVID-19 risk.

## Keywords

Trajectory tracking, Self-organizing map, Visualization, Clustering, Epidemic risk

## Cite this article as

Chen, N., Chen, A. and Yao, XH. (2022) Trajectory Tracking of COVID-19 Epidemic Risk using Self-organizing Feature Map. *Bulletin of the Chinese Academy of Sciences*, 36(2), 91–100.

DOI: <https://doi.org/10.1051/bcas/2022003>

---

\* To whom correspondence may be addressed. Email: [anchen@casisd.cn](mailto:anchen@casisd.cn)

## 1. Introduction

The 2019 novel coronavirus disease (COVID-19) has emerged as a worldwide pandemic rapidly since the outbreak in December 2019. In January 2020, the United States announced the first confirmed case of COVID-19, and United Kingdom reported the first new death due to the epidemic. From March 2020, COVID-19 continued to spread across continents with the first confirmed cases appearing in Latin America and Caribbean region including Brazil, Mexico, Ecuador, Dominica, Argentina and Chile. In December 2020, a new variant of coronavirus with an estimated 40~70% increase in infectivity was reported in the United Kingdom through virus gene sequencing. A new round of strong infectious diseases subsequently broke out in Europe and the United States, and then spread violently in winter. The reported number of confirmed cases and deaths grew at an unprecedented rate (Chatterjee, P., Nagi, N., & Agarwal, A., *et al.*, 2020). Till July 2022 the number of confirmed cases reported to World Health Organization (WHO) has risen to above 550 million in total across the globe with more than 6 million deaths.

COVID-19 has a great negative impact on the world economy, leading to a sharp drop in economic growth, raised unemployment rates, reductions in trade and cross-border investments, and fluctuation of commodity price (Kraus, S., Clau, T., & Breier, M., *et al.*, 2020). In terms of global GDP losses, COVID-19 caused the worst recession since the end of the 2nd World War, twice as much as the global economic crisis in 2008. As a large-scale infectious disease, the impact of COVID-19 over the world is difficult to eliminate for a long time. It is estimated that the whole world will probably bear the aftermath of COVID-19 through the next decade.

The COVID-19 crisis also posed crucial challenges to emergency management. Due to the insufficient investment in public health, many countries failed to respond in time in the face of the COVID-19 outbreak. Later the inconsistent cooperation among countries in the prevention and control of the pandemic allowed it to spread over the world. The emergency management of COVID-19 crisis receives great attention in academia, with the focuses placed on medical diagnosis support, risk modeling and prediction, public opinion management, drilling and assistant decision support, *etc.*

A growing body of research has reported applying information technologies, such as surveillance system, machine learning, computational intelligence, remote sensing, and GIS, in guiding localized planning and decision-making against COVID-19 (Asadzadeh, A., Pakkhuo, S., & Saeidabad, M., *et al.*, 2020). A variety of machine learning methods were found to be applicable in analyzing the COVID-19 data for prediction, classification and clustering. The prediction of new confirmed cases and required equipment is a focus of critical concern in the prevention and control of COVID-19 outbreak. Fantazzini forecasted the number of new daily cases and deaths of COVID-19 using Google Trends data of 158 countries (Fantazzini, D. 2020). Qin *et al.* predicted the number of new suspected or confirmed cases of COVID-19 based on social media search indexes (SMSI) (such as dry cough, fever, chest distress, coronavirus, and pneumonia) using regression models (Qin, L., Sun, Q., & Wang, Y., *et al.* 2020). Warda *et al.* predicted the number of infections and deaths due to COVID-19 using a GIS mapping model in Lahore District, Pakistan (Warda, R., Song, W., & Kaif, G., *et al.*, 2020). Booton *et al.* employed a mathematical model to estimate the required facilities such as hospitals, beds and ventilators to meet the need of COVID-19 in South West England (Booton, R. D., Macgregor, L., & Vass, L. *et al.*, 2021). S.K. Tamang *et al.* used artificial neural network and curve fitting techniques to forecast the future cases of COVID-19 and achieved satisfactory results in India, the USA, France, the UK, China and South Korea (Tamang, S. K., Singh, P. D., & Datta, B., 2020).

In medical diagnosis, artificial intelligence techniques, particularly classification have achieved a variety of innovations and advancements by analyzing the overwhelming volume of patient data (Shaikh, F., Dehmeshki, J., & Bisdas, S., *et al.*, 2021). Random forest (Casiraghi, E., Malchiodi, D., & Trucco, G., *et al.*, 2020) and multivariable logistic regression (Woo, S. H., Rios-Diaz, A.J., & Kubey, A. A., *et al.*, 2020) were applied to quickly predict the risk of COVID-19 regarding in- or out-patient care. Deep learning and principal component analysis were used for the fast and accurate detection of COVID-19 epidemic by analyzing the X-ray and CT images in clinical diagnoses (Hu, S., Gao, Y., & Niu, Z., *et al.*, 2020; Zturk, A., Zkaya, U., & Barstuan, M., 2020).

For a novel disease being studied, clustering has the potential to reveal the differences between COVID-19 and other viral pneumonia (Householder, J., Householder, A., & Gomez-Reed, J. P., *et al.*, 2020). Alqurneh *et al.* examined the safety of co-medication for COVID-19 using partition clustering (Alqurneh, A., Mustapha, A. & Sharef, N. M., 2020). Zarikas compared the COVID time series of 30 countries with respect to active cases, active cases per population, and active cases per population and per area using a hierarchical clustering approach (Zarikas, V., Pouloupoulos, S. G., & Gareiou, *et al.*, 2020).

It is noted that self-organizing feature map (SOM) facilitates the visualization and property exploration of data by clustering the data and simultaneously preserving the topological properties. SOM has been widely applied as an analytical tool in a diversity of real world applications such as fault diagnosis (Li, Z., Fang, H., & Huang, M., *et al.*, 2018), crop evapotranspiration (Adeloye, A. J., Rustum, R., & Kariyama, I. D., 2011), clinical voice analysis (Huang, D.W., Gentili, R.J., & Katz, G. E., *et al.*, 2016), video summarization (Rani, S., Kumar, M., 2020), ice shape roughness assessment (Neubauer, T., Hassler, W., & Puffing, R., 2020). For time-series data, SOM offers an easily understandable manner of trajectory tracking, so that the dynamic changes of observations might be detected (Chen, N., Ribeiro, B., & Vieira, A., *et al.*, 2013; Dewan, P., Ganti, R., & Srivatsa, M., 2017; Ling, C., Delmelle, E. C. 2016; Qi, J., Liu, H., & Liu, X., *et al.*, 2019).

In this paper, the COVID-19 risk of 237 countries is measured in terms of two reliable barometers of the severity of the pandemic, namely the number of new confirmed cases and that of new deaths per day over one year (the first day is 3, January 2020, while the last day is 10, January 2021), and then projected on a well-trained SOM with the risk grades defined by K-means. Then the temporal trajectories of risk are constructed across this space and displayed on the component planes, making it possible to keep track of the dynamic changes of daily risk. The contributions of the paper lie in *a*) defining the risk grades of COVID-19 epidemic by clustering the new confirmed cases and deaths of countries and *b*) offering a visual way for decision-makers to understand the temporal trajectory of COVID-19 risk. Compared with the existing studies

mostly focused on the prediction of epidemic risk using some quantitative models, the presented work sheds light on the trajectory tracking and comparative studies of epidemic risk among countries, with the help of SOM's capability in both clustering and visualization. The findings are of great value to government departments in designing control measures, improving public health management, and further exploring the shortcomings and problems in public health system.

The rest of this paper is organized as follows. Section 2 outlines the methodology of trajectory tracking for COVID-19 time series data. Section 3 is devoted to some analytic results of the proposed approach. Finally, Section 4 concludes the paper along with some suggestions and highlights the future directions.

## 2. Methodology and Framework

In this section, the framework of trajectory tracking of COVID-19 risk is schematically described in Figure 1. The publicly available COVID-19 data released by WHO includes the official daily counts of new confirmed cases and deaths from 3, January 2020 to 10, January 2021 (374 days totally) in 237 countries. The data contains negative values probably due to input mistakes or corrections of previous data on both new confirmed cases (0.08% of all cases) and new deaths (0.03% of all cases). Because it is difficult to distinguish these two problems, negative values are directly ignored in preprocessing. Considering the low proportion of negative values in all data, it will not have a significant impact on the clustering results. The data is firstly cleaned by removing the negative values, and transformed to a 2-d matrix where the first dimension is the number of new confirmed cases per day and the second is the number of new deaths. The data is sorted first by the countries and then by the days. Afterwards the absolute values are normalized to the unity range for subsequent processing. Secondly, the data is explored by a clustering process composed of SOM and K-means. The former facilitates the trajectory visualization of time series data on the map, and the latter defines the risk grades by partitioning the neurons into a few clusters. Finally the trajectory of COVID-19 risk is visualized on the map for easy tracking.

SOM is a vector quantization method designed for

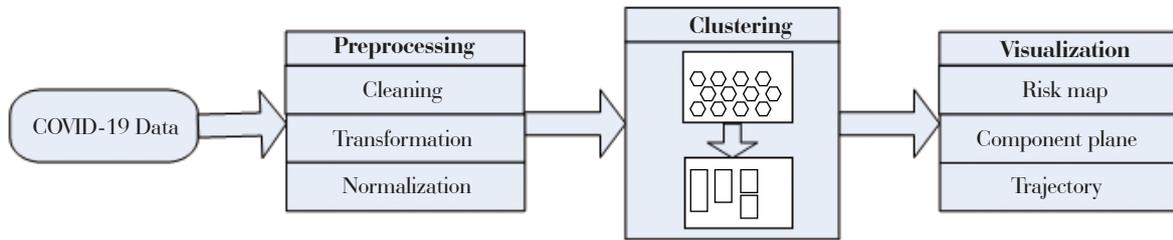


Figure 1. Framework of trajectory tracking of COVID-19 risk

unsupervised learning (Kohonen, T., Schroeder, M. R., & Huang, T. S., 2001). Each neuron is associated with a prototype representing a cluster of similar input. The prototypes are placed on a regular low-dimensional grid in an ordered fashion. During the training, each input is projected to the best-matching unit (BMU) on the map through a competition mode. Then the prototype of both BMU and the adjacent neurons on the map is updated by the gradient descent according to the neighborhood kernel function and learning rate function. The termination condition of iterative learning is usually the maximal number of iterations or the convergence of prototypes. In this manner the neurons are topologically ordered on the grid gradually so that similar inputs are likely projected to the same or adjacent neuron in the map grid space. The neighborhood kernel function  $h_{ij}$  indicates the inter-connectivity strength between neurons  $i$  and  $j$ , defined as:  $h_{ij}(t) = \exp \frac{-\|c_i - c_j\|^2}{2\delta(t)^2}$ , where  $\delta(t)$  is the neighborhood radius function, and  $c_i$  ( $c_j$ ) is the position of neuron  $i$  ( $j$ ) on the map lattice. Let  $\alpha_i$  be the initial learning rate,  $\alpha_e$  the ending learning rate, and  $t_{max}$  the total number of iterations. The learning rate defined as  $\alpha(t) = \alpha_i (\alpha_e / \alpha_i)^{t/t_{max}}$  is a function monotonically decreasing with learning epoch  $t$ .

SOM groups the data to a number of small and compact clusters with respect to BMU taking the neurons as the cluster centers. In the post-clustering, the initial clusters are aggregated to a few clusters using K-means (Vesanto, J., Alhoniemi, E., 2000). Given a set of  $n$  objects, K-means aims to partition the objects (namely the neurons in this study) into  $k$  ( $k \leq n$ ) sets so as to minimize the within-cluster sum of squares. To find the optimal value of  $k$ , Davies-Bouldin (DB) index is commonly used which evaluates the quality of clustering in terms of both intra-cluster compactness and inter-cluster separation. Finally the data is assigned

with the cluster number of the corresponding BMU on the map. The two-step clustering process is described as follows (Chen, N., Chen, L., & Ma, Y., *et al.*, 2018; Chen, N., Ma, Y., & Tang, C., *et al.* (2020).

**SOM clustering step**

$N$ : number of training instances,  $M$ : number of neurons,  $x_i (i=1, \dots, N)$  the input vector,  $m_i (i=1, \dots, M)$  the prototype of neuron  $i$ .

- (1) For  $p=1, \dots, M$ , initialize  $m_p$  of map neurons;
- (2) For  $i=1, \dots, N$ , input  $x_i$  to the map one at a time;
- (3) For  $p=1, \dots, M$ , calculate the pairwise distance  $\|x_i, m_p\|$ ;
- (4) Project  $x_i$  to BMU  $c$ , where  $c = \text{argmin} \|x_i, m_p\|$ ;  $1 \leq p \leq M$
- (5) For  $p=1, \dots, M$ , update the prototype:  $m_p = m_p + \alpha(t)h_{c,p}(t)(x - m_p)$ ;
- (6) Repeat from (2) a few iterations until the termination condition is satisfied.

**Post-clustering step**

maxClusters: maximal number of clusters,  $\{C_1, C_2, \dots, C_k\}$ : clusters,  $\mu_i$ : center of cluster  $i$ .

- (1) For  $k=2$  to maxClusters
- (2) Perform K-means clustering at a fixed  $k$ ;
- (3) Calculate intra-cluster compactness  $S_i$  and inter-cluster separation

$$M_{ij}:$$

$$S_i = \|x, \mu_i\| / |C_i| \quad (i=1, \dots, k)$$

$$x \in C_i$$

$$M_{ij} = \|u_i, u_j\| \quad (i, j=1, \dots, k)$$

- (4) Calculate DB-index:  $DB = \sum_{i=1}^k \max_{i \neq j} (S_i + S_j) / M_{ij}$
- (5) Repeat from (1) until maxClusters is reached;
- (6) Select the optimal  $k$  that minimizes DB-index.

Given a time series data  $x = \{x_1, \dots, x_n\}$ , the trajectory of  $x$  on the map is defined as the vector  $\{BMU(x_1), \dots, BMU(x_n)\}$ , where  $BMU(x_i)$  is the best-matching unit of  $x_i$  through the competition mode. In this manner, the trajectory can be calculated for each country over the observed period.

**3. Results and Discussions**

In this section, the COVID-19 data is analyzed using the proposed method for risk visualization and trajectory tracking. The risk caused by COVID-19 is simply measured by two barometers, *i.e.*, the number of new confirmed cases (NC) and the number of deaths (ND) daily.

Figure 2 shows the number of countries reported to WHO per day during the period, indicating that COVID-19 has become a worldwide pandemic since April 2020, with new confirmed cases coming from more than 140 countries and deaths from over 80 countries daily. Later the spread of COVID-19 remains relatively stable with small-scale fluctuations. Figure 3 outlines the distribution of NC and ND with respect to days and countries respectively. There is a significant difference in the evolution of COVID-19 cases among the 237 countries. Compared with the initial stage, the distinction of NC among countries is more apparent in the later stage. The second round of COVID-19 outbreaks due to the new variant of coronavirus led to the dramatic increase in both confirmed cases and deaths. The sharp increases of new deaths appeared during both the first and the second round of epidemic outbreak.

Currently, the risk levels of COVID-19 epidemic are defined differently among countries. The US Centers for Disease Control and Prevention (CDC) has launched an epidemic rating system that grades a country according to the number of people in a country and the number of confirmed cases over the past 28 days. China divides the risk of COVID-19 into high, medium and low levels according to the cumulative cases and new confirmed cases within 14 days. It is known that the number of deaths is a critical index on assessing the harm of public health events. In this study, the number of new confirmed cases and the number of deaths are used jointly to evaluate a country's epidemic risk.

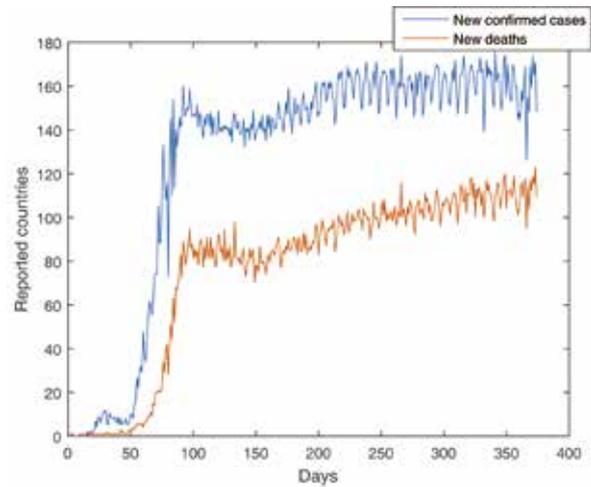


Figure 2. Number of reported countries per day starting from 3 January 2020

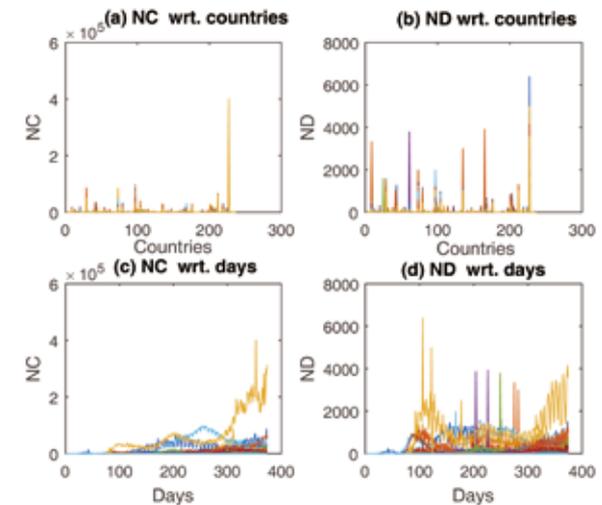


Figure 3. Distribution of new confirmed cases and new deaths from 3 January 2020

The preprocessed data is fed to SOM for initial clustering and then K-means for post-clustering grouping. The final number of clusters is determined by the DB-index, which is 4 as disclosed in Figure 4 (a). By sorting the center of clusters in descending order, the grades of epidemic risk can be defined as red (major risk), blue (high risk), yellow (general risk) and green (low risk). Figure 4 (b) visualizes the risk map in four colors representing different risk grades. The risk map reveals that the upper neurons in green have low risk and the neurons in red on the left bottom corner are characterized by major risk. The other neurons are divided into two risk grades, *i.e.*, high risk and general

risk. From the component planes shown in Figure 5, the neurons marked as major and high risks are characterized by high values of both new confirmed cases and deaths. On the contrary the neurons marked as low risk have relatively low values of both new confirmed cases and deaths. In this sense, the joint index of risk grades reasonably represents the harm of the epidemic. The corresponding epidemic early warning rules can be formulated according to the risk level of SOM nodes. When new data is observed, it is first mapped to a node according to the nearest principle. If the risk level of the node is major or high, the epidemic warning information can be released.

For each time point, the response on the map is found by projecting the data to the best-matching unit. The trajectory of time series data is generated by connecting the responses in sequence. The number of days along the trajectory in each risk grade is calculated correspondingly. In summary, Figure 6 displays the histogram chart of four risks of all reported countries spanning over the observed period. The x-axis represents the number of days in the corresponding risk grade, and the y-axis the number of countries in each container. It is noticed that most countries are categorized to the grade of low risk for COVID-19 on most days, indicating the epidemic is basically under control except in a small number of countries.

The countries with high risk are of crucial concern for risk management. Consequently, we select seven countries whose trajectory is marked with major risk over at least two days. The countries are located in Asia (India), Europe (United Kingdom, Germany, Italy, France), and America (United States, Brazil). The frequency of days along the trajectory located in each risk grade is summarized in Table 1. Particularly the first day in red risk can be found. The risk trajectories of these counties are shown on the component planes in Figure 7 regarding NC and ND respectively. For easy tracking we sampled the responses every 10 days by marking the starting point and end point on the map. The trajectory makes it possible to get insights into the dynamic change of risk. All trajectories started from the low risk (NC=ND=0 at the beginning), and evolved in different ways showing the distinction of dynamic risk among countries. It is noticeable that the risk of the United States increased quickly and remained in

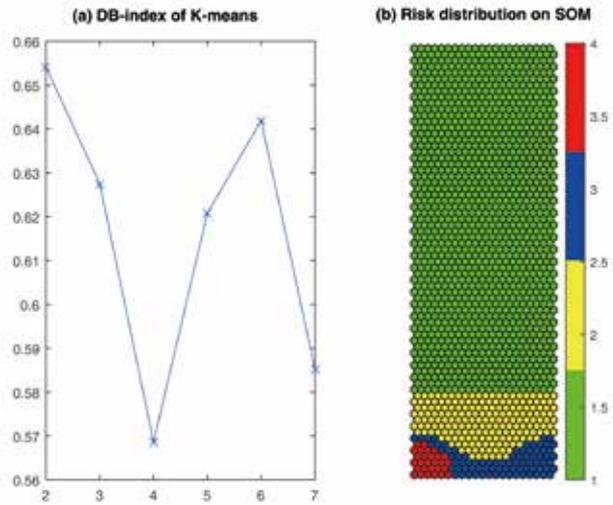


Figure 4. Risk distribution using K-means

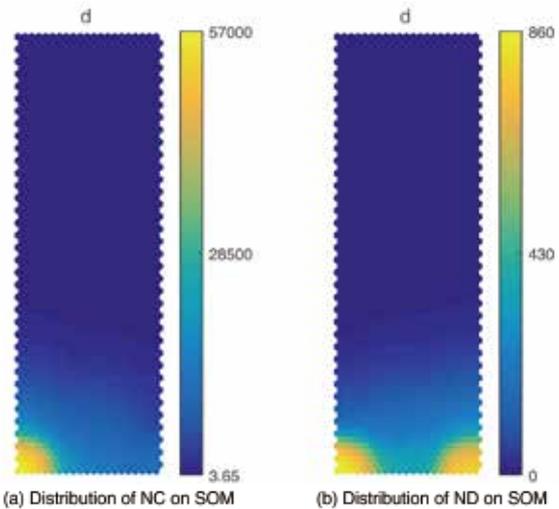


Figure 5. Component planes of SOM

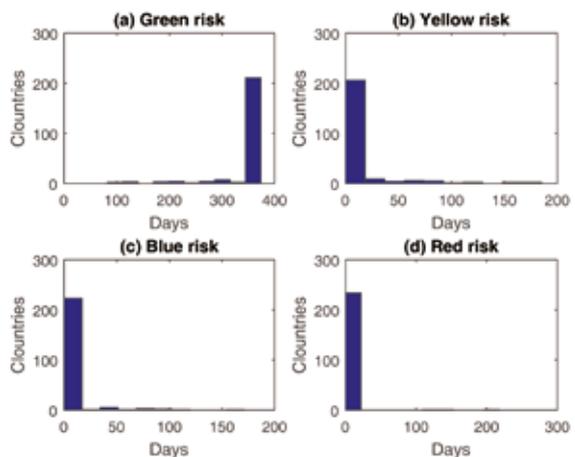


Figure 6. Risk distribution of countries

**Table 1. Risk distribution of 7 high-risk countries in 374 days**

	Green risk	Yellow risk	Blue risk	Red risk	First day in Red risk
Brazil	104	47	101	122	June, 2020
France	210	95	59	10	October, 2020
Germany	262	78	28	6	December, 2020
India	134	51	48	141	July, 2020
Italy	214	59	86	15	November, 2020
United Kingdom	204	73	80	17	November, 2020
United States	82	7	71	214	April, 2020

blue or red grade for a long time, almost spanning over the two rounds of epidemic outbreaks. Since March 2020, COVID-19 has spread quickly and widely across the country. The risk was marked as red firstly in April, 2020. In fact, the United States, the country with the highest number of confirmed cases, accounted for approximately 24% of officially reported global cases until January 2021. In Latin America, the novel coronavirus continued to slowly spread since March 2020, where Brazil was the most affected with a gradual increase in both confirmed cases and deaths.

The trajectory of risk grades depicts the risk changes in a direct way, as shown in Figure 8. The four European countries, namely France, Germany, Italy and the United Kingdom, characterized by low or general risk mainly during the first round of epidemics, show an abrupt risk rise since the second round of outbreaks at the end of 2020. In Brazil and the United States, the epidemic began to deteriorate in the second quarter of 2020 and remained in the high or major risk grade for a long time. In Asia, India is the worst affected country in terms of new confirmed cases and deaths reported daily. It almost stays in high risk since the middle of 2020, as indicated by the trajectory produced from the dataset, after eliminating some errors of reported data. So far, more than 10 million cases and 150 thousand deaths have been reported officially.

#### **4. Conclusions and Future Work**

COVID-19 pandemic has become a severe public health emergency of international concern, with a profound impact on the medical resources, the economy, and the society. The world has continuously taken

actions to diminish the prospects of its proliferation. The trajectory that documents the stabilization or reversal of epidemics provides critical information support for risk assessment and management. In this study, the dynamic changes of COVID-19 epidemic are characterized by first defining the risk grades on an organized map with respect to the number of new confirmed cases and deaths, and then tracking the temporal trajectory of risks on the map over the observed period.

The trajectory visually displays the dynamics of risk change. From the results, we make a number of interesting findings. Firstly, the second round COVID-19 epidemic is much more severe than the first one regarding the number of confirmed cases and deaths. It is found that the new variant increased the transmissibility in the process of evolution, which has resulted in a rise in mortality. Secondly, the risk of COVID-19 epidemic shows significant differences among countries. From the perspective of emergency management, it is crucial to control the spread of the epidemic at the beginning of an outbreak. The novel coronavirus pneumonia is mainly transmitted through respiratory droplets and contact, so the close contacts with confirmed cases have a high infection risk. It is necessary to perform strict centralized isolation for medical observation, so as to achieve early detection, early isolation, and early treatment. Additionally, the COVID-19 crisis reflects the vulnerability of the public health system of many countries in the face of a large-scale epidemic, for example, a shortage of prevention materials at the beginning of the epidemic, the inconsistent cooperation among countries. It is urgent to increase the investment to public health and strengthen the joint prevention and control of public health emergencies.

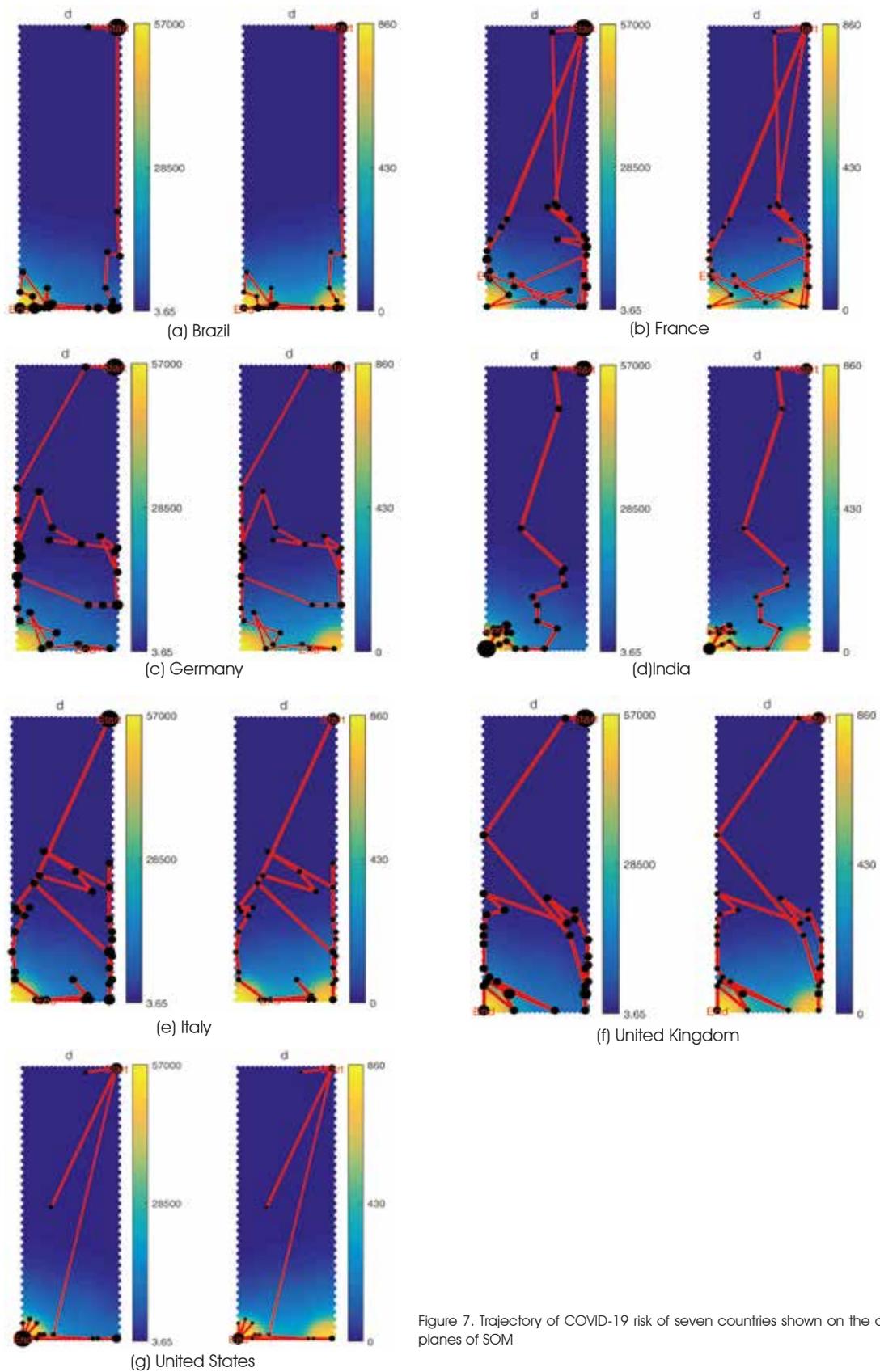


Figure 7. Trajectory of COVID-19 risk of seven countries shown on the component planes of SOM

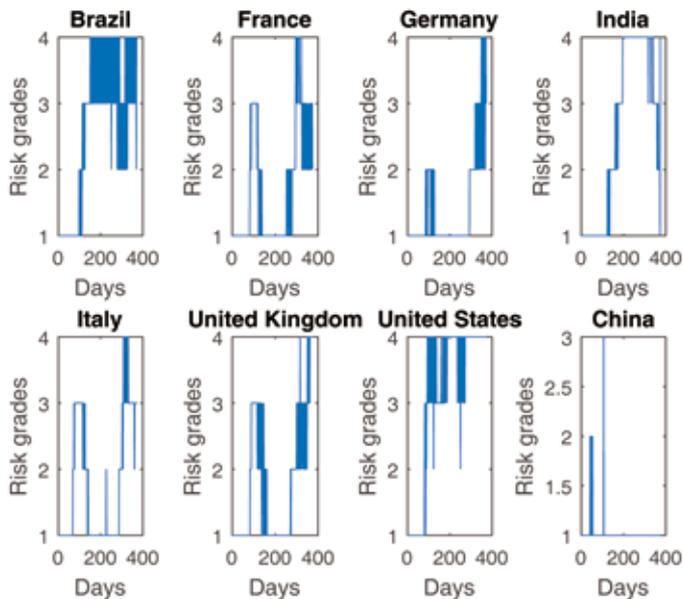


Figure 8. Risk trajectory of seven countries from 3 January 2020

Although the study is performed on data during a fixed period, the proposed method can be applied to time-series data for long-term trajectory tracking. In fact, with the development of vaccines and the improvement of global immunity, the COVID-19 epidemic risk will change dramatically and will continually remain the focus of future research. The evolution patterns of COVID-19 risk will be investigated by clustering the trajectories for further risk comparison and analysis. Moreover, a predictive model will be constructed for risk forecast as an interesting direction of practical concern.

### Acknowledgements

This work was supported by National Office of Philosophy and Social Sciences (19AZD019) and National Ethnic Affairs Commission (2020-GMB-015).

### References

- Adeloye, A. J., Rustum, R., & Kariyama, I.D. (2011). Kohonen self-organizing map estimator for the reference crop evapotranspiration. *Water Resources Research*, 47(8), 192–198.
- Alqumeh, A., Mustapha, A., & Sharaf, N. M. (2020). A partitioning-based approach for clustering COVID-19 drugs and co-medication for safe use. *International Journal of Integrated Engineering*, 12(5). doi:10.30880/ijie.2020.12.05.029
- Asadzadeh, A., Pakkhou, S., & Saeidabad, M., et al. (2020). Information technology in emergency management of COVID-19 outbreak. *Informatics in Medicine Unlocked*, 21, 100475.
- Boon, R. D., Macgregor, L., & Vass, L., et al. (2021). Estimating the COVID-19 epidemic trajectory and hospital capacity requirements in South West England: A mathematical modelling framework. *BMJ Open*, 11(1), e041536.
- Casiraghi, E., Malchiodi, D., & Trucco, G., et al. (2020). Explainable machine learning for early assessment of COVID-19 risk prediction in emergency departments. *IEEE Access*, 8, 196299–196325.
- Chatterjee, P., Nagi, N., & Agarwal, A., et al. (2020). The 2019 novel coronavirus disease (COVID-19) pandemic: A review of the current evidence. *The Indian Journal of Medical Research*, 151(2–3).
- Chen, N., Ribeiro, B., Vieira, A., et al. (2013). Clustering and visualization of bankruptcy trajectory using self-organizing map. *Expert Systems with Application*, 40(1), 385–393.
- Chen, N., Chen, L., Ma, Y., et al. (2018). Regional disaster risk assessment of China based on self-organizing map: Clustering, visualization and ranking. *International Journal of Disaster Risk Reduction*, 33, 196–206.
- Chen, N., Ma, Y., & Tang, C., et al. (2020). Risk assessment and comparison of regional natural disasters in China using clustering. *Intelligent Decision Technologies*, 14(3), 349–357.
- Dewan, P., Ganti, R., & Srivatsa, M. (2017). SOM-TC: Self-organizing map for hierarchical trajectory clustering. *IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 1042–1052. doi: 10.1109/ICDCS.2017.244
- Fantazzini, D. (2020). Short-term forecasting of the COVID-19 pandemic using Google trends data: Evidence from 158 countries. *Applied Econometrics*, 59, 33–54.
- Householder, J., Householder, A., Gomez-Reed, J. P., et al. (2020). Clustering COVID-19 lung scans. *arXiv e-prints*, arXiv:2009.09899.
- Hu, S., Gao, Y., & Niu, Z., et al. (2020). Weakly supervised deep learning for COVID-19 infection detection and classification from CT images. *IEEE Access*, 1–1.
- Huang, D.W., Gentili, R. J., & Katz, G. E., et al. (2016). A limit-cycle self-organizing map architecture for stable arm control. *Neural networks: the official journal of the International Neural Network Society*, 85(C), 165–181.
- Kohonen, T., Schroeder, M. R., & Huang, T. S. (2001). *Self-organizing maps*. Springer Berlin Heidelberg.
- Kraus, S., Clau, T., Breier, M., et al. (2020). The economics of COVID-19: Initial empirical evidence on how family firms in five European countries cope with the corona crisis. *International Journal of Entrepreneurial Behaviour & Research*, 26(5), 1067–1092.
- Li, Z., Fang, H., Huang, M., et al. (2018). Data-driven bearing fault identification using improved hidden Markov model and self-organizing map. *Computers & Industrial Engineering*, 116, 37–46.
- Ling, C., Delmelle, E.C. (2016). Classifying multidimensional trajectories of neighborhood change: A self-organizing map and k-means approach. *Annals of GIS*, 22(3), 1–14.
- Neubauer, T., Hassler, W., & Puffing, R. (2020). *Ice shape roughness assessment based on a three-dimensional self-organizing map approach*. AIAA Aviation Forum. doi:10.2514/6.2020–2805



- Qi, J., Liu, H., Liu, X., *et al.* (2019). Spatiotemporal evolution analysis of time-series land use change using self-organizing map to examine the zoning and scale effects. *Computers, Environment and Urban Systems*, 76, 11–23.
- Qin, L., Sun, Q., & Wang, Y., *et al.* (2020). Prediction of the number of new cases of 2019 novel coronavirus (COVID-19) using a social media search index. Social Science Electronic Publishing, *SSRN Electronic Journal*. doi:10.2139/ssrn.3552829
- Rani, S., Kumar, M. (2020). Social media video summarization using multi-visual features and Kohonen's self organizing map. *Information Processing & Management*, 57(3):102190.1–102190.17
- Shaikh, F., Dehmeshki, J., & Bisdas, S., *et al.* (2021). Artificial intelligence-based clinical decision support systems using advanced medical imaging and radiomics. *Current Problems in Diagnostic Radiology*, 50(2), 262–267.
- Tamang, S. K., Singh, P. D., & Datta, B. (2020). Forecasting of COVID-19 cases based on prediction using artificial neural network curve fitting technique. *Global Journal of Environmental Science and Management*, 6 (Special Issue (COVID-19)), 53–64.
- Vesanto, J., Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 11(3), 586–600.
- Warda, R., Song, W., & Kaif, G., *et al.* (2020). COVID-19 spread prediction and its correlation with social distancing, available health facilities using GIS mapping data models in Lahore, Pakistan. *Technium Social Sciences Journal*, 10. doi:10.47577/tssj.v10i1.1291
- Woo, S. H., Rios-Diaz, A. J., & Kubey, A. A., *et al.* (2020). Development and validation of a Web-based severe COVID-19 risk prediction model. *Cold Spring Harbor Laboratory Press*, 2020(4). doi:10.1101/2020.07.16.20155739
- Zarikas, V., Pouloupoulos, S. G., & Gareiou, Z., *et al.* (2020). Clustering analysis of countries using the COVID-19 cases dataset. *Data in Brief*, 31:105787, <https://doi.org/10.1016/j.dib.2020.105787>
- Zturk, A., Zkaya, U., & Barstuan, M. (2020). Classification of Coronavirus (COVID) from X-ray and CT images using shrunken features. *International Journal of Imaging Systems and Technology*, 31(1):5–15. doi:10.1002/ima.22469