Nat Hazards
https://doi.org/10.1007/s11069-018-3279-y

ORIGINAL PAPER

Spatial distribution patterns of global natural disasters based on biclustering

Shi Shen[1,2,3,4] · Changxiu Cheng[1,2,3,4] · Changqing Song[2,3,4] · Jing Yang[1,2,3,4] · Shanli Yang[1,2,3,4] · Kai Su[1,2,3,4] · Lihua Yuan[2,3,4] · Xiaoqiang Chen[2,3,4]

Received: 9 November 2017 / Accepted: 15 March 2018

Abstract Understanding the spatial distribution patterns (SDPs) of natural disasters plays an essential role in reducing and minimizing natural disaster risks. An integrated discussion on the SDPs of multiple global disasters is still lacking. In addition, due to their high quantity and complexity, natural disasters constitute high-dimensional data that represent a challenge for an analysis of SDPs. This paper analyzed the SDPs of global disasters from 1980 to 2016 through biclustering. The results indicate that the SDPs of fatality rates are more uneven than those of occurrence rates. Based on the occurrence rates, the selected countries were clustered into four classes. (1) The major disasters along the northern Pacific and in the Caribbean Sea and Madagascar are storms, followed by floods. (2) Most of Africa is mainly affected by floods, epidemics, and droughts. (3) The primary disaster types in the Alpine-Himalayan belt and the western Andes are floods and earthquakes. (4) Europe, America, Oceania, and South and Southeast Asia are predominantly influenced by floods. In addition, according to the fatality rates, the selected countries were clustered into eight classes. (1) Extreme high temperatures mostly result in high fatality rates (HFRs) in developed countries. (2) Epidemics lead to HFRs in parts of Africa. (3) Droughts produce HFRs in East Africa. (4) Earthquakes result in HFRs along the eastern Pacific coastline and the Alpine-Himalayan belt. (5) Tsunamis mainly cause HFRs in Thailand, Indonesia, and Japan. (6) Storms result in scattered but distinct HFRs along the coastal regions of the Pacific and Indian Oceans. (7) Floods cause concentrated HFRs in South Asia and northeastern South America. (8) Finally, volcanoes cause HFRs in Colombia, while extreme low temperatures cause HFRs in Ukraine and Poland.

✉ Changxiu Cheng
chengcx@bnu.edu.cn

[1] Key Laboratory of Environmental Change and Natural Disaster, Beijing Normal University, Beijing 100875, China

[2] State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

[3] Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[4] Center for Geodata and Analysis, Beijing Normal University, Beijing 100875, China

Published online: 24 March 2018

Keywords Natural disasters · Biclustering · Spatial distribution · Multiple disasters

1 Introduction

It is of great significance to identify and understand the spatial distribution patterns (SDPs) of natural disasters, especially any spatial heterogeneity in those patterns. This is essential for the enhancement of disaster resilience, reduction of natural disasters, and sustainable economic development (Dilley 2006; Shi et al. 2010; Council 2012). Particularly, the spatial pattern of global natural disasters can assist multinational corporations and organizations in minimizing the loss of life and property (Bathrellos et al. 2017). A contributing factor to this end is to enhance our knowledge regarding the global SDPs of multiple natural disasters from an integrated perspective.

Generally, many previous studies focused on the spatial distribution of disasters in local reference areas (Borden and Cutter 2008; Han et al. 2016; Li et al. 2016; Peng et al. 2017), although a few recent studies have focused on mapping single disasters by analyzing the global distribution pattern (Cvetkovic 2013; Sasa and Cvetkovic 2014). However, the spatial patterns of individual or local regional-scale disasters may confuse decision makers because of the variance in the available measurements and the limitations of the investigated spatial scales (Borden and Cutter 2008). Moreover, natural disasters have numerous interconnections and interactions (Tweed and Walker 2011; Yang et al. 2011), and they are prone to occur across national boundaries (Oh and Oetzel 2011). Hence, simultaneously considering all types of disasters and all possible countries is a vital approach for synthetically analyzing the global spatial pattern of natural disasters.

In addition, the analysis of multiple global natural disasters poses a challenge regarding the high dimension of the data caused by treating countries and disasters as variables and sample sizes, respectively (Ferraty 2010). Typical high-dimensional data have particular characteristics, such as the existence of a large number of irrelevant attributes, a sparse data distribution, and a large amount of local information, which are not easily solved using ordinary methods.

Alternatively, the biclustering method is an adequate approach to address this challenge. Different from traditional clustering methods, the biclustering method is able to simultaneously cluster objects and their attributes to discover subsets of similar properties in the dataset (Madeira and Oliveira 2004; Padilha and Campello 2017). This method has been widely used to analyze bioinformatics (Cheng and Church 2000; Pontes et al. 2015), juvenile crime (Izenman et al. 2011), and natural disasters (Shen et al. 2016).

In this paper, disaster data are derived from the EM-DAT (D. Guha-Sapir) database and are generalized into 11 types. The biclustering method is used to obtain unique insights regarding the SDPs of global natural disasters. The spatial distribution characteristics, including the patterns in the disaster occurrence and fatality rates, are proposed, analyzed, and then aggregated at the country level, after which they are compared for an evaluation of the spatial heterogeneity.

The rest of this paper is organized as follows. Section 2 introduces the data resources and the biclustering method. In Sect. 3.1, the biclustering results are presented and evaluated. The SDPs of the occurrence and fatality rates are identified and analyzed in Sects. 3.2 and 3.3, respectively. Finally, the discussion and conclusions of this paper are given in Sect. 4.

2 Data resources and methods

A total of 25,042 natural disaster records from 1980 to 2016 were derived from the OFDA/CRED International Disaster Database (EM-DAT), which includes significant disaster events that satisfy at least one of the following criteria: At least ten people are killed; at least 100 people are affected; a state of emergency is declared by a governmental entity; and a call for international help is issued. The EM-DAT also provides information on the countries and fatalities of disaster events. In addition, both population data and land area data were obtained from the World Bank. The occurrence and fatality rates of each disaster type were calculated for the corresponding countries.

2.1 Biclustering method

The biclustering method used in this paper following Zhao and Karypis (2003) was executed in gCLUTO 1.0. After constructing the country-disaster matrix, the repeated bisection algorithm was applied to cluster the data matrix. The repeated bisection algorithm is an extension of the basic *k*-means clustering method. It first splits all of the objects into two clusters via two-means clustering and then selects one of those clusters for splitting and other operations. This process repeats $k - 1$ times until k clusters are produced. The details for the repeated bisection algorithm are given by Algorithm 1.

Algorithm 1: The Repeated Bisection Algorithm

1: **Initialize** the list of a cluster consisting of all objects

2: **Repeat**

3: Select a cluster from the list of clusters

4: Bisect the selected cluster using a 2-way clustering algorithm

5: Add the resultant two clusters to the list of clusters to replace the selected cluster

6: **Until** the list of clusters contains k clusters

The similarity in the repeated bisection algorithm can be expressed by the cosine similarity, the Euclidean distance, or the Jaccard similarity coefficient. In this paper, the cosine similarity was chosen to determine the similarity between each cluster object and its centroid, and it is calculated as follows:

$$\mathrm{sim}(\mathbf{X}, \mathbf{Y}) = \cos(\mathbf{X}, \mathbf{Y}) = \frac{\sum_{i=1}^{n} x_i y_i}{\|\mathbf{X}\|_2 \|\mathbf{Y}\|_2}, \quad (1)$$

where the vectors $\mathbf{X}$ and $\mathbf{Y}$ represent two clustering objects in the attribute space. Suppose each clustering object has n properties; each property is represented by an n-dimensional vector $[(x_1, x_2, \ldots, x_n)]$. x_i indicates the ith attribute of the clustered object, and $\|\mathbf{X}\|_2$ is the 2-norm of the vector $\mathbf{X}$.

Meanwhile, the clustering results were evaluated using the average internal similarity (ISim), the average external similarity (ESim), the average standard deviation of the internal similarity (ISdev), and the average standard deviation of the external similarity (ESdev). Generally, better clustering results have higher average internal similarities and lower average internal standard deviations of those similarities; in addition, their average

external similarities are lower, and the average external standard deviations of those similarities are higher.

2.2 Definitions

This paper integrated 11 types of natural disasters from the EM-DAT database: droughts (DR), floods (FL, including coastal floods, riverine floods, and flash floods), earthquakes (EQ), storms (ST, including cyclones, tropical cyclones, and tornadoes), epidemics (EP), landslides (LS, including debris flows), wildfires (WF), extreme high temperatures (EH), extreme low temperatures (EC), volcanoes (VO), and tsunamis (TS).

Considering differences in the land areas and populations of different countries, two metrics, namely the OPA (occurrences per unit area) and FPP (fatalities per unit people), were defined to estimate the occurrence and fatality rates of disasters among different countries. The formal definitions of the OPA and FPP are as follows:

$$\mathrm{OPA}_{i,j} = \frac{\sum \left(\mathrm{Occur}_{i,j,k}\right)}{\mathrm{Area}_j}, \tag{2}$$

$$\mathrm{FPP}_{i,j} = \frac{\sum \left(\mathrm{Fatality}_{i,j,k}\right)}{\overline{\mathrm{Pop}_j}}, \tag{3}$$

where $\mathrm{OPA}_{i,\,j}$ indicate the number of occurrences per 10,000 km^2 of the ith type of disaster in the country j over a period of time, $\mathrm{Occur}_{i,\,j,\,k}$ represents the number of occurrences of the ith type of disaster in the country j during the year k, and Area_j is the land area of the country j. Similarly, $\mathrm{FPP}_{i,\,j}$ indicate the number of fatalities per 10,000 people of the ith type of disaster in the country j over a period of time, $\mathrm{Fatality}_{i,\,j,\,k}$ represents the number of fatalities of the ith type of disaster in the country j during the year k, and $\overline{\mathrm{Pop}_j}$ is the average population of the country j during the investigated period.

In this study, the rows and columns in the data matrix represent the different countries and natural disaster types, respectively. The numerical values in the data matrix indicate the OPA and FPP values from 1980 to 2016. The country-disaster-occurrence matrix and the country-disaster-fatality matrix were constituted according to the OPA and FPP values.

In addition, in order to meet statistical requirements of the analytical results, countries with less than 30 disaster events were not selected for clustering in Sect. 3.2, while countries with less than 1000 deaths due to disasters in the past 27 years were not included in the clustering samples in Sect. 3.3. Therefore, there are 96 countries and 11 types of disasters in the country-disaster-occurrence matrix. There are 84 countries and 11 types of disasters in the country-disaster-fatality matrix.

3 Analysis of the spatial pattern of global natural disasters

3.1 Biclustering results of global natural disasters

Through the global occurrence rates and fatality rates of natural disasters from 1980 to 2016, the biclustering results for all of the countries were generated using gCLUTO 1.0 (Fig. 1). The peaks are labeled with the serial numbers of the clusters. The heights and volumes of the peaks represent the average internal similarity and the number of members

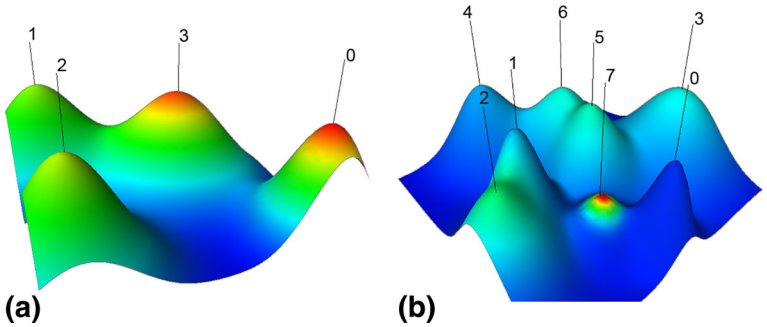

Fig. 1 Mountain visualization of four OPA clusters (**a**) and eight FPP clusters (**b**)

within the cluster, respectively, and colors of the mountain peaks represent the average standard deviation of internal similarity in the cluster. A redder peak indicates a lower average standard deviation of the internal similarity in the corresponding cluster.

According to the OPA, four country clusters were identified (Fig. 1a). As shown in Fig. 1a, the OPA biclustering results show good separability because of the four high and scattered mountain OPA peaks. In addition, their peaks are reddish, indicating that the four clusters meet the characteristic features (i.e., high ISim, low ISdev, low ESim, and high ESdev) of good biclustering results.

Figure 1b shows that eight country clusters were identified based on the FPP. Although the mountain cluster peaks in Fig. 1b are not as scattered as those in Fig. 1a, they are still clearly distinguishable. Except for class 7, the other cluster peaks are high. Hence, the FPP biclustering results are relatively acceptable.

The detailed biclustering performance of fatalities data of the country-disaster-occurrence matrix is shown in Table 1. The cluster represents the number of classes. The size shows the number of countries in each class. For OPA clusters, all four clusters have higher ISim than ESim. At the same time, their ISdevs are lower than ESdevs. Due to this performance, the four OPA clusters are very fine.

The detailed biclustering performance of fatalities data of the country-disaster-fatality matrix is shown in Table 2. For FPP clusters, clusters 0–6 have significantly high ISim and low ISdev. Meanwhile, their ESims are much lower than ISims and their ESdevs are higher than ISdevs. Although the ISim of cluster 7 is not as much high as the other seven clusters, it is still much higher than its ISdev and ESim. Hence, the eight FPP clusters result is reasonable.

Table 1 Biclustering performance of the country-disaster-occurrence matrix

Cluster	Size	ISim	ISdev	ESim	ESdev
0	13	0.886	0.046	0.544	0.170
1	30	0.923	0.030	0.636	0.105
2	17	0.871	0.032	0.684	0.078
3	36	0.884	0.043	0.712	0.055

Table 2 Biclustering performance of the country-disaster-fatality matrix

Cluster	Size	ISim	ISdev	ESim	ESdev
0	4	0.987	0.009	0.066	0.059
1	5	0.955	0.017	0.074	0.045
2	13	0.920	0.065	0.050	0.058
3	26	0.942	0.049	0.135	0.048
4	12	0.924	0.030	0.149	0.072
5	9	0.912	0.047	0.139	0.087
6	12	0.866	0.044	0.250	0.072
7	3	0.533	0.187	0.072	0.051

3.2 The spatial distribution pattern of global natural disaster occurrence rates

A spatial distribution map of the OPA was generated through ArcGIS 9.3 and MATLAB; it is displayed in Fig. 2, which shows a clear spatial heterogeneity at the global scale and a uniform spatial distribution at the regional scale. The selected countries were clustered into four classes; those classes are filled with the same colors in the map. The boxplots detail the statistical distribution characteristics of the OPA for each type of disaster among the corresponding classes. For instance, among the countries in class 0, the OPA of ST have the highest average and variance. Beneath the boxplots, bars of red blocks indicate the average cosine similarities among the OPA of the corresponding natural disasters and

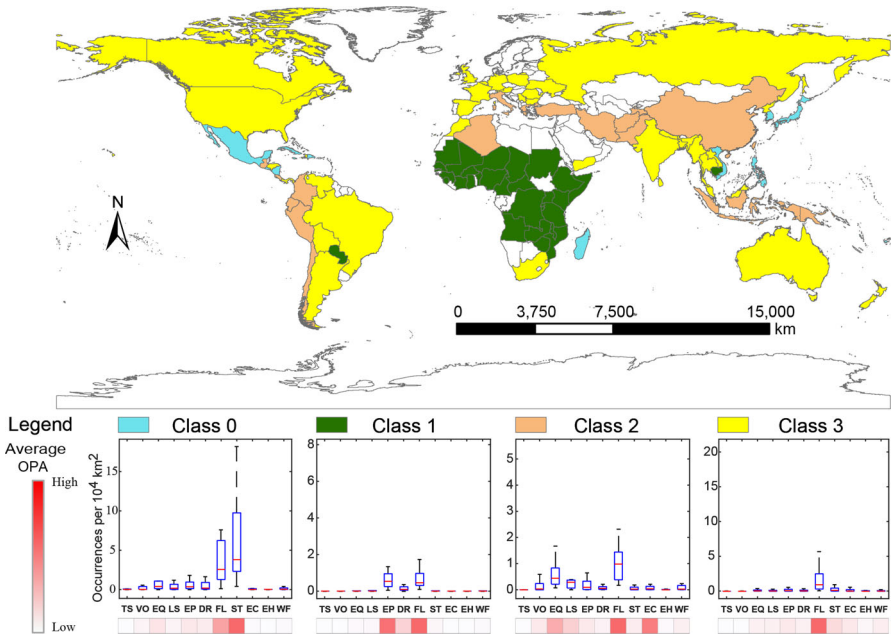

Fig. 2 Spatial pattern of global disaster clusters based on the OPA

classes. A redder block indicates a higher average OPA. For example, the OPA of the class 0 countries have a higher OPA for ST and FL than for the other disaster types.

As shown in Fig. 2, the class 0 region distributes along the North Pacific Coast as well as in Madagascar and the Caribbean Sea. Storms are the major type of disaster, followed by floods. A total number of 13 countries were identified. The maximum OPA of ST exceed 18 occurrences per 10^4 km^2. This spatial pattern coincides with the cyclone tracks (Knapp et al. 2010).

Meanwhile, most countries of the class 1 region concentrate in Africa and are mainly affected by floods and epidemics, followed by droughts. There are 30 countries in the class 1 region in total. The OPA of FL and EP have the similar average. This result indicates that floods are one of the most common disasters in Africa and implies that epidemics are correlated with floods in these areas (Li et al. 2016).

Figure 2 shows that the class 2 region locates within the Alpine-Himalayan orogenic belt and the western Andes where the major disaster type is floods, followed by earthquakes. A total of 17 countries were identified in the class 2 region. These two regions are also located within the Mediterranean-Himalayan seismic belt and the Circum-Pacific seismic belt. Moreover, Indonesia is located at the intersection of these two seismic belts. As a result, following FL, EQ are the second-most major disaster in this region. The OPA of EQ range from 0 to 2, but the maximum OPA of FL are slightly larger than 2.

In addition, the class 3 region, distributing in Europe, the eastern Andes, Oceania, South and Southeast Asia, the USA, and Canada, is primarily affected by FL. A total of 36 countries were identified in the class 3 region. In the matrix, it is clear that FL are the most prominent type of natural disaster. The OPA of FL vary from 0 to 5. These areas cover tropical, temperate, and semiarid climates, indicating that global factors are involved in floods.

3.3 The spatial distribution pattern of global natural disaster fatality rates

Figure 3 displays the spatial distribution map of global natural disaster fatality rates according to the FPP using ArcGIS 9.3. All of the selected countries were clustered into eight classes, and the same classes are filled with identical colors in the map. Similarly, the bars of red blocks beneath the map represent the average FPP among the corresponding natural disasters and classes. Redder blocks indicate a higher average FPP. Meanwhile, the boxplots in Fig. 4 provide detailed the statistical distribution characteristics of the FPP for each type of disaster in the corresponding class. For example, the FPP of the class 0 countries have the highest, average, and variance for DR.

Comparing Fig. 2 with Fig. 3, it is noteworthy that several colored countries in Fig. 2, such as Canada, Argentina, and Bolivia, are not shown on the map of global natural disaster fatality rates. This is because the fatalities in each of these countries over the last 27 years are less than 1000. This indicates that these countries have a strong resilience to natural disasters.

Figure 3 clearly shows that extreme high temperatures often result in high fatality rates in developed countries distributed throughout medium- and high-latitude zones. The 13 countries in the class 2 region include the USA, Russia, and Australia in addition to others located throughout Western Europe; these countries have the largest average FPP of EH. This shows that extreme high temperatures are the main lethal disasters in these countries.

As shown in Fig. 3, high fatality rates in the class 3 region are predominantly caused by epidemics. A total of 26 countries, including Peru, Malaysia, and Laos in addition to

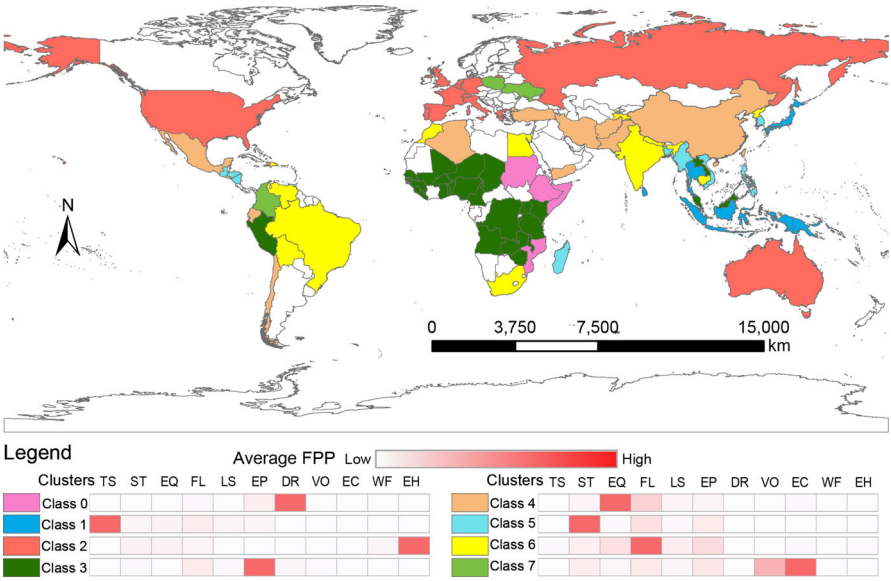

Fig. 3 Spatial pattern of global disaster clusters based on the FPP

numerous African countries, were identified in this region. These countries have the largest average FPP of EP. This indicates that EP constitute the main lethal disaster in the class 3 region.

It is apparent from Fig. 3 that droughts are responsible for the high fatality rates in parts of East Africa. There are only four African countries in the class 0 region, namely Sudan, Ethiopia, Somalia, and Mozambique. The corresponding block bar and boxplot show that these countries have the largest average FPP of DR. Consequently, the most common major disaster that causes high fatality rates in these countries is DR.

According to Fig. 3, earthquakes result in high fatality rates in countries located along the eastern Pacific coast and the Alpine-Himalayan belt. In particular, there are 12 countries in the class 4 region distributed throughout the Mediterranean-Himalayan seismic belt and the Circum-Pacific seismic belt. Since EQ have the largest average FPP, the predominant type of disaster in this region is EQ.

Observing the spatial distribution map of global disasters, it is evident that tsunamis often cause the highest fatality rates in the western Pacific and along the Indian Ocean coastlines. The five countries in the class 1 region are Japan, Indonesia, Thailand, Sri Lanka, and Papua New Guinea. TS show the largest average FPP in this region. Therefore, TS constitute the major disaster type that causes high fatality rates in these countries. Many documentaries have confirmed that three catastrophic tsunamis have caused enormous casualties in these countries (Stone and Kerr 2005; Tappin et al. 2008; Aoyama et al. 2014; McCall 2017). Papua New Guinea was struck by a tsunami in 1998, while Indonesia, Thailand, and Sri Lanka were heavily damaged by the 2004 Indian Ocean tsunami. In addition, Japan was badly hit by the 2011 Tohoku tsunami.

Figure 3 also shows that storms cause high fatality rates in Madagascar, Bangladesh, Myanmar, South Korea, Vietnam, the Philippines, and in the central American states. A total of nine countries were identified within the class 5 region. Because ST have the

largest average FPP in these countries, it is the most destructive natural disaster in this region. This is because these countries are distributed along the coastal areas of the North Pacific, the western Atlantic, and the northern and western Indian Ocean, which are all high-risk typhoon areas and have weak resilience to typhoons.

Although floods often lead to high fatality rates in India, Cambodia, North Korea, Brazil, Venezuela, Bolivia, and South Africa, etc., in contrast to the corresponding occurrence rate pattern, the spatial extent of FL as the main fatal disaster is dramatically reduced. There are 12 countries in the class 6 region. Due to the largest average FPP of FL within this region, FL represent the primary disaster that causes the fatality rates therein. This phenomenon suggests that most countries are effectively capable of addressing floods.

Moreover, there are three countries, Colombia, Ukraine, and Poland in the class 7 region. However, there are two main types of natural disasters in this region. First, VO have the second-highest average and the maximum FPP in Colombia. Second, EC have the highest similarity and the largest average FPP in Poland and Ukraine. In addition, Columbia has been struck by multiple volcanic disasters. VO and EC are aggregated together because they are uncommon fatal disasters.

Figure 4 displays the data ranges, averages, and variances of the FPP among the natural disasters for the eight classes shown in Fig. 3 using MATLAB. DR in East Africa have the highest average FPP among all of the disaster types. This implies that droughts, especially agriculture droughts, easily generate famine and consequently lead to many deaths (Masih et al. 2014). The maximum FPP of ST in the class 1 region approach 20. The maximum FPP of TS in the class 5 region reach 30. EP are also a serious issue in class 3 with a maximum of more than 10. The maximum FPP of EH in the class 2 region approach 4 with an average greater than 1. Because EH occur in developed countries with fatality rates comparable to FL in the class 6 region, EH should be appropriately valued. Finally, VO are found in Columbia, and the maximum FPP of VO approach 6. The average FPP of EC in both Poland and Ukraine exceed 0.

4 Discussion and conclusions

In this paper, the global SDPs of multiple natural disasters were analyzed. The findings demonstrate that the global occurrence rates display a relatively even SDP on a large scale. The selected countries were clustered into four regions. Floods are the most widely distributed disaster, followed by storms, earthquakes, landslides, droughts, and epidemics. The occurrence rates of the natural disasters are closely related to their geographical distributions and the environments where the corresponding main natural hazards are located. For example, the countries in the North Pacific, including Japan, Mexico, and the Philippines, have a higher frequency of storms. The occurrence rates of earthquake disasters in Italy, Turkey, China, and Chile, which are situated within active seismic belts, are more pronounced than in other countries.

The spatial heterogeneity of the fatality rates of natural disasters is more significant than that of the occurrence rates. An investigation into the fatality rates of the different types of disasters found that the selected countries could be clustered into eight regions. Countries with HFRs caused by extreme high temperatures are developed countries distributed throughout medium- and high-latitude zones. Countries with HFRs due to epidemics are distributed throughout most of Africa, while those with HFRs resulting from droughts are

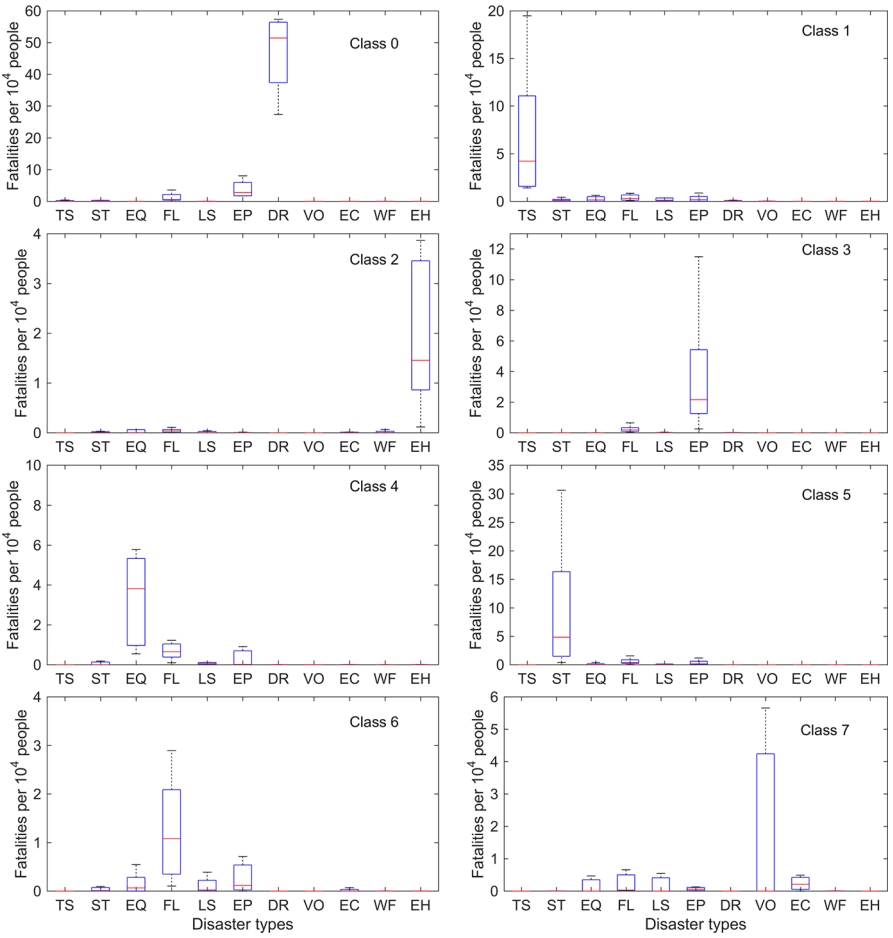

Fig. 4 Boxplots of the FPP for global natural disaster fatality rates. The boxplots detail the statistical distribution characteristics of the FPP for the disaster types among the corresponding classes

located in East Africa. Countries with earthquakes as the primary cause of their HFRs are situated along the Pacific Rim and the Alpine-Himalayan belt. Countries with HFRs attributable to tsunamis are concentrated along the Pacific and Indian Ocean coasts. Countries with floods as the main cause of their HFRs are India, Cambodia, North Korea, Brazil, Venezuela, and Bolivia in addition to the South African states. Colombia is the sole country with volcanoes as the cause of HFRs. Meanwhile, the HFRs in Ukraine and Poland are due to extreme low temperatures.

There are similarities between the SDPs of the occurrence rates and fatality rates. First, epidemics have significant occurrence rates and significant fatality rates in most African countries. In addition, earthquakes exhibit high occurrence rates and high fatality rates in the regions within the Mediterranean-Himalayan seismic belt and the Circum-Pacific seismic belt. According to the disaster system theory (Shi 1996, 2002), the losses caused by a disaster are synthetically determined by the type of the hazard, the distribution of exposure, the economic development capacity, and the resilience, among other factors.

Therefore, the low socioeconomic level, poor health conditions, and high occurrence rates of epidemics among African countries have led to a high fatality rate of epidemics in those countries. Because of their destructive power and relatively high occurrence rates, losses due to earthquakes in countries with low resiliencies are often high and are related to the distribution of seismic belts and exposure.

The differences between the spatial distributions of global disaster occurrence rates and fatality rates can provide insights for disaster risk management. Japan experiences more typhoons and earthquake disasters, but the most fatalities were caused by the Tohoku tsunami (Aoyama et al. 2014). Although the frequencies of floods, earthquakes, and landslides in Indonesia are high, the highest death count was the result of the Indian Ocean tsunami. Based on the above theory, these differences can be explained as follows: The occurrence rates of some catastrophes may be low, but the resiliencies of countries to floods, storms, and earthquakes are stronger than those with respect to the occasional catastrophic tsunami. Resiliencies to more frequent disasters are more developed and sufficient; however, for sudden-onset catastrophes, the resilience of a particular country is likely insufficient, and thus, the fatality rate is higher. Hence, research on rare catastrophes should receive more attention.

Additionally, this study showed that the biclustering method is able to solve problems with regard to high-dimensional data involving types of disasters and numbers of countries. Even though spatial correlations were not used, the clustering results showed a significant SDP of natural disasters from an integrated perspective. In future research, the complex and nonlinear effects of natural disasters on climate change will require more research to understand the interactions and occurrences of multiple natural disasters, especially at the global scale.

Acknowledgements We would like to thank the anonymous reviewers for their constructive comments on the manuscript. This work was supported by National Natural Science Foundation of China [Grant Number 41771537], National Key Research and Development Plan of China [Grant Number 2017YFB0504102], and the Fundamental Research Funds for the Central Universities.

References

Aoyama T, Shiratori N, Hagimoto K et al (2014) Lessons of the Great East Japan earthquake. IEEE Commun Mag 52:21–22

Bathrellos GD, Skilodimou HD, Chousianitis K et al (2017) Suitability estimation for urban development using multi-hazard assessment map. Sci Total Environ 575:119–134

Borden KA, Cutter SL (2008) Spatial patterns of natural hazards mortality in the United States. Int J Health Geogr 7:64

Cheng Y, Church GM (2000) Biclustering of expression data. Intell Syst Mol Biol 8:93

Council NR (2012) Disaster resilience: a national imperative. The National Academies Press, Washington, DC

Cvetkovic V (2013) Spatial and temporal distribution of geophysical disasters. J Geogr Inst Jovan Cvijic SASA 63:345–360

Dilley M (2006) Setting priorities: global patterns of disaster risk. Philos Trans R Soc A 364:2217–2229

Ferraty F (2010) High-dimensional data: a fascinating statistical challenge. J Multivar Anal 101:305–306

Guha-Sapir D EM-DAT: the CRED/OFDA international disaster database. http://www.emdat.be/

Han W, Liang C, Jiang B et al (2016) Major natural disasters in China, 1985–2014: occurrence and damages. Int J Environ Res Public Health 13:1118

Izenman AJ, Harris PW, Mennis J et al (2011) Local spatial biclustering and prediction of urban juvenile delinquency and recidivism. Stat Anal Data Min 4:259–275

Knapp KR, Kruk MC, Levinson DH et al (2010) The international best track archive for climate stewardship (IBTrACS). Bull Am Meteorol Soc 91:363–376

Li C, Chai Y, Yang L, Li H (2016) Spatio-temporal distribution of flood disasters and analysis of influencing factors in Africa. Nat Hazards 82:721–731
Madeira SC, Oliveira AL (2004) Biclustering algorithms for biological data analysis: a survey. IEEE/ACM Trans Comput Biol Bioinform 1:24–45
Masih I, Maskey S, Mussá FEF, Trambauer P (2014) A review of droughts on the African continent: a geospatial and long-term perspective. Hydrol Earth Syst Sci 18:3635–3649
McCall C (2017) Remembering the Indian Ocean tsunami. Lancet 384:2095–2098
Oh CH, Oetzel J (2011) Multinationals' response to major disasters: how does subsidiary investment vary in response to the type of disaster and the quality of country governance? Strateg Manag J 32:658–681
Padilha VA, Campello RJGB (2017) A systematic comparative evaluation of biclustering techniques. BMC Bioinform 18:55
Peng Y, Song J, Cui T, Cheng X (2017) Temporal-spatial variability of atmospheric and hydrological natural disasters during recent 500 years in Inner Mongolia, China. Nat Hazards 89:441–456
Pontes B, Giráldez R, Aguilar-Ruiz JS (2015) Biclustering on expression data: a review. J Biomed Inform 57:163–180
Sasa M, Cvetkovic VM (2014) Victimization of people by natural disasters: spatial and temporal distribution of consequences. Temida 17:19–42
Shen S, Cheng C, Su K et al (2016) Quantitative visualization about differences between scientists concerned nature disasters and historic events. In: 2016 International conference on behavioral, economic and socio-cultural computing (BESC), pp 1–6
Shi P (1996) Theory and practice of disaster study. J Nat Disasters 11:6–17
Shi P (2002) Theory on disaster science and disaster dynamics. J Nat Disasters 11:1–9
Shi P, Li N, Ye Q et al (2010) Research on integrated disaster risk governance in the context of global environmental change. Int J Disaster Risk Sci 1:17–23
Stone R, Kerr RA (2005) Girding for the next killer wave. Science 310:1602–1605
Tappin DR, Watts P, Grilli ST (2008) The Papua New Guinea tsunami of 17 July 1998: anatomy of a catastrophic event. Nat Hazards Earth Syst Sci 8:243–266
Tweed F, Walker G (2011) Some lessons for resilience from the 2011 multi-disaster in Japan. Local Environ 16:937–942
Yang T, Wang X, Zhao C et al (2011) Changes of climate extremes in a typical arid zone: observations and multimodel ensemble projections. J Geophys Res Atmos. https://doi.org/10.1029/2010JD015192
Zhao Y, Karypis G (2003) Clustering in life sciences. In: Brownstein MJ, Khodursky AB (eds) Functional genomics: methods and protocols. Humana Press, Totowa, pp 183–218