Analysis of Influencing Factors of Geological Disaster Risk Based on Big Data

Xin Lu[1](✉), Huanghuang Lu[1], Xun Wang[2], and Jiao Yuan[2]

[1] School of Software, University of Electronic Science and Technology of China, Chengdu, China
un365@qq.com
[2] China Railway Eryuan Engineering Group Co. Ltd., Chengdu, China

Abstract. Geological disaster risk is affected by many internal and external factors. It is very necessary to understand the influence relationship of these internal and external factors for the prevention and control of geological disasters. This paper combines Apriori correlation analysis method with data statistical analysis method to analyze the internal and external incentives of geological disaster risk, so as to solve the problems of limited analysis data, less disaster samples and narrow adaptability of analysis conclusions in the existing research. This subject collects the data related to geological disasters in China in recent ten years as analysis samples for big data analysis experiments. The analysis data not only includes the elevation, slope, slope direction, terrain type, stratum lithology, geological structure, surface coverage type, vegetation coverage and other internal factors of geological disaster areas, but also covers the rainfall, rainfall and vegetation in geological disaster areas External factors such as surface temperature and soil humidity. The results show that geological disaster risk is related to internal and external factors, but their correlation and influence are different, which has certain application significance for geological disaster prevention and control.

Keywords: geological disasters · risk factors · big data analysis · relevance

1 Introduction

Geological disaster refers to a kind of surface change phenomenon under the action of stress inside and outside the earth or human activities. It usually causes damage to the surrounding environment, loss or threat to human life and property. The types of geological disasters mainly include landslide, settlement, debris flow, collapse, ground fissure, etc. [1]. China is a country with frequent geological disasters, among which landslide, debris flow and collapse are the three most frequent geological disasters. The annual human and economic losses caused by geological disasters are very huge [2, 3]. Therefore, many experts and scholars are committed to the study of geological disaster risk, analyze the correlation between geological disaster risk and internal and external influencing factors, and put forward the strategies of preventing and controlling geological disasters.

G. Guan et al. (Eds.): ICBDSS 2022, AHCS 8, pp. 1062–1074, 2023.
https://doi.org/10.2991/978-94-6463-064-0_111

LV Meina used qualitative and semi quantitative statistical analysis methods to analyze the relationship between collapse geological disasters and formation lithology, elevation, slope, rainfall and other influencing factors by making statistics on the occurrence of collapse geological disasters in Guangzhou in recent years, and obtained the strong and weak relationship of collapse geological disasters in Guangzhou affected by the above factors [4]. Taking Shanxi Province as an example, Huang Hucheng analyzed the information data of major geological disasters in Shanxi Province by the end of 2015 by using the data statistical method, and obtained the relationship between geological disasters in Shanxi Province and landform, rock and soil mass, rainfall and human activities [5]. Li Renwei and others analyzed the geological disaster information data of Yan'an City in 2018 by using the coupling analysis method, and combined with the spatial distribution characteristics of geological disasters, gave the relationship between geological disaster risk and elevation, slope, formation lithology, average annual rainfall, water system, road, population density and vegetation type in Yan'an City [6]. Nie Bingqi et al. Combined with grey correlation method and orthogonal design, taking the wading slope of a reservoir area as an example, analyzed the sensitivity between various influencing factors and slope geological disaster risk [7].

From the above academic achievements, it can be seen that most scholars will choose a single analysis method to analyze the influencing factors of geological disaster risk in specific areas, and the analysis conclusions are only applicable to specific areas. In addition, the influencing factors analyzed by some researchers are not comprehensive enough, resulting in the analysis conclusion can not support complex application scenarios. In order to solve the above problems, this paper compiles the collection program to obtain the domestic geological disaster information in recent ten years from the website system of the Ministry of natural resources, and also collect the big data of internal and external incentives in the geological disaster area from other external systems. The relationship between influencing factors and geological disaster risk is analyzed by using Apriori correlation analysis algorithm and descriptive statistical analysis method.

2 Related Works

The analysis of influencing factors of geological disasters involves the analysis of internal factors and external incentives of geological conditions. It also involves data analysis algorithms and data statistical processing methods.

2.1 Influencing Factors of Geological Disaster Risk

The geological disaster risk of an area is mainly affected by the geological conditions, internal factors and external incentives of the area [8, 9]. The internal factors of geological conditions largely determine whether geological disasters occur in the area, and the factors mainly include the following aspects:

1) Terrain. It reflects the surface shape of the region formed under the action of internal and external stresses of the earth, that is, the surface fluctuation degree of the region. Terrain type, elevation and slope of mountain have a certain impact on geological

disaster risk. The slope of the mountain has the most obvious influence. The greater the slope of the mountain, the greater the impact on the stability of the slope.

2) Formation lithology. It reflects the stability of surface rock and soil. Generally, mountains with complete geological structure and hard rock and soil have good stability, while mountains composed of soft and loose rock and soil often have poor stability.
3) Geological structure. The strata of different geological structure types can bear different stress in the earth. If the geological structure of an area is complex and fragile, the stratum in the area can only bear small internal and external forces. On the contrary, the more solid and stable the geological structure, the stratum in this area can bear greater internal and external forces.
4) Vegetation cover. Plants have a certain solidification effect on the land, and they can stabilize the formation rock and soil in a region. Plants can also slow down the scouring of rainfall on rock and soil and make the rock and soil bear less impact of rain. Therefore, areas with good surface vegetation coverage have better geotechnical stability.
5) Surface coverage. Surface coverage reflects the coverage of the land surface in an area, and it can also reflect the degree of human use of the land in the area. Woodland and grassland are two types of land cover with low degree of human land use, and they have good surface stability. Artificial surface and cultivated land are two types of surface coverage with high degree of human utilization, and their surface stability is relatively poor.

Geological disasters are affected not only by the internal factors of the regional geological conditions, but also by the external incentives of the region. The main external incentives are as follows:

1) Meteorology. Meteorology is the main external inducing factors of geological disasters, including rainfall, soil humidity and surface temperature. Among them, rainfall is the most influential external inducing factor [10].
2) Seismicity. Seismic activity will directly affect the stability of geotechnical structure in an area, resulting in the loosening of geotechnical structure in the area. Once there is rainfall or aftershocks, there is a great possibility of secondary geological disasters.
3) Human activities. Human activities are also one of the most direct factors affecting the occurrence of geological disasters. Engineering activities such as oil and gas field exploitation, mine development and reservoir construction will change the geological conditions of the region to a certain extent, have a certain impact on the surface stability of the region, and then increase the risk of geological disasters in the region.

2.2 Apriori Association Analysis Algorithm

Apriori association analysis algorithm is a kind of data analysis algorithm to find the association rules between data items from the data set. This paper uses Apriori correlation analysis algorithm to analyze the relationship between geological disaster risk and internal factors of geological conditions. It involves the following basic concepts [11].

1) Itemset: a collection of data items. An itemset with three items is called a three item set. For example, {mountain, fault and loose rock formation} is a three item set.
2) Support: the probability that itemsets A and B occur simultaneously. For rule A → B, the calculation formula of its support is shown in Eq. 1.

$$\text{Support}_{A\rightarrow B} = \frac{N(A \cup B)}{N} \tag{1}$$

In the above formula, N ($A \cup B$) refers to the number of data samples containing itemset {A, B}. N is the total number of data samples. For example, if the itemset {mountain, occurrence of geological disasters} has 80 and the total number of geological disasters has 100, the support of {mountain, occurrence of geological disasters} is 80%. Minimum support is the minimum threshold value of support set according to the application.
3) Frequent itemsets: frequent itemsets are itemsets that meet the minimum support. For example, if the set minimum support is 50%, and the support of {mountain, geological disaster occurrence} is 80%, which is greater than 50%, then {mountain, geological disaster occurrence} is a frequent itemset.
4) Confidence: the probability that if itemset A occurs, itemset B will also occur. For rule A → B, the calculation formula of confidence is shown in Eq. 2.

$$\text{Confidence}_{A\rightarrow B} = \frac{N(A \cup B)}{N(A)} = \frac{\text{Support}_{A\rightarrow B}}{\text{Support}_{A}} \tag{2}$$

In the above formula, N ($A \cup B$) refers to the number of data samples containing itemset {A, B}. N (A) refers to the number of data samples in which item A occurs. For example, if the number of {mountain, geological disaster occurrence} is 80 and the number of N (mountain) is 80, the confidence of {mountain, geological disaster occurrence} is 1. Confidence is essentially the conditional probability that item B also occurs on the premise of item A. Like the minimum support, the minimum confidence also sets the minimum threshold according to the application. If there is a rule that satisfies both minimum support and minimum confidence, then we call this rule strong rule.

In this paper, the sample data set of geological hazards is formed by collecting the data of geological hazards and their influencing factors. Apriori association analysis algorithm is used, and the minimum support and minimum confidence are set to find out the frequent itemsets. According to the obtained frequent itemsets, the correlation between geological disaster risk and internal factors of geological conditions is analyzed.

2.3 Algorithm Descriptive Statistical Analysis Method

Descriptive statistical analysis method is a kind of method to statistically describe data sets by means of data statistics, data classification, statistical charts and so on. Descriptive statistical analysis methods mainly include data trend analysis, data dispersion analysis, data distribution analysis and statistical chart analysis [12].

1) Data trend analysis. It is used to reflect the trend level of the overall data. It usually uses indicators such as median, mode and average to describe the data trend.
2) Data dispersion analysis. It is used to reflect the degree of difference between data. Standard deviation and variance indicators are usually used to reflect the degree of difference between data.
3) Data distribution analysis. In many data analysis, it is necessary to understand the overall distribution of data samples, such as using kurtosis and skewness indicators to describe the normal distribution of sample data.
4) Statistical chart analysis. With the help of software tools, the data set is visually analyzed and processed, such as histogram, pie chart and broken line chart to describe the data relationship. Statistical chart analysis is intuitive and concise, and the data analysis conclusion is easy to understand.

This paper will use statistical chart analysis to present the data relationship between geological disaster risk and external incentives. It can more clearly and intuitively analyze the data relationship between influencing factors and geological disaster risk.

3 Data Collection of Influencing Factors of Geological Hazards

This paper analyzes and deals with the relationship between domestic geological disasters and their internal and external influencing factors in recent ten years. These analysis data can not be obtained directly. This paper develops a crawling program to collect the required data from the external system. The collected data of internal influencing factors of geological disasters include terrain type, elevation, slope, slope direction, formation lithology, geological structure, surface coverage type, vegetation coverage, etc. the data of external influencing factors include rainfall, soil humidity and surface temperature.

1) Geological disaster point event data: in the geological disaster service application website of the Ministry of natural resources through Python crawler program (http://geohazard.geol.ckcest.cn/) collect domestic geological hazard point event data.
2) Terrain type data: from osgeo China Center website (https://www.osgeo.cn/data/) obtain terrain type data of geological hazard area. According to the longitude and latitude of the geological disaster area, the terrain type data of the area (plain, basin, hill, mountain, plateau, etc.) are obtained and stored in the local disaster sample database.
3) Terrain data: from NASA website system (https://urs.earthdata.nasa.gov) get DEM digital elevation data file of disaster area. The data file is extracted, calculated and processed to obtain the elevation, slope and aspect data of the disaster area, and stored in the local disaster sample database.
4) Geological condition data: from the website system of national geological data museum (http://www.ngac.org.cn/DataSpecial/geomap.html) obtain the geological map of the disaster area, extract the formation lithology, geological structure and other geological condition data of the area, and store them in the local disaster sample database.

5) Surface coverage type data: from the global geographic information public goods website (http://www.globallandcover.com) collect the surface coverage type data of geological disaster area and store it in the local disaster sample database.
6) Vegetation cover data: from the European Center for medium range weather forecasting website system (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5) collect the leaf area index data of geological disaster areas and store them in the local disaster sample database.
7) Meteorological data: from the website of the National Environmental Information Center (https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs) obtain the meteorological data of the disaster area (rainfall, soil humidity, surface temperature, etc.) and store it in the local disaster sample database.

4 Analysis on Influencing Factors of Geological Disaster Risk

In this paper, Apriori correlation analysis algorithm and descriptive statistical analysis method are used to analyze the data of geological disaster risk and internal and external influencing factors, with the goal of obtaining more comprehensive and accurate analysis results.

4.1 Analysis of Internal Influencing Factors of Geological Disaster Risk Based on Apriori Algorithm

The internal factors of a region's geological conditions determine whether there is the possibility of geological disasters in the region. In this paper, an Apriori algorithm is proposed to conduct association analysis on the collected geological disaster data set to find out the correlation between geological disaster risk and internal influencing factors in the region. The pseudo code of the algorithm is shown in Table 1.

After several groups of experiments, the reasonable minimum support of this data set is 30%. This data set is the data set of geological disasters that have occurred, and its confidence is 100%. After the calculation and processing of the algorithm program, the output association rule results are shown in Table 2.

From the association rules given in Table 2, the following conclusions can be drawn:

1) In the existing geological hazard areas, the hazard points with a slope range of 0–20° have 72% support. The analysis results are not completely consistent with the expected situation that the greater the slope is, the greater the risk of geological disasters is. The reasons are as follows: a) when the gradient of a slope is small, the slope is easy to carry out human activities, the greater the influence of external incentives, and the greater the possibility of landslide, debris flow and other geological disasters. However, when the slope reaches a certain degree, the impact of human activities will be reduced accordingly, and the risk of geological disasters will be reduced. b) The steeper the mountain, the more difficult it is for the loose rock and soil on its surface to be fixed on the slope, and the less loose material on the surface. The rock and soil on the surface of steep mountains are less and less in the long-term scattering process, which is difficult to provide sufficient material

Table 1. Apriori algorithm pseudo code for association analysis of geological hazard data

```
Input: geological hazard dataset data, minimum confidence minC,
minimum support minS

Output: association rule table apriList
---
Process 1:
//Gets the length of the Data. Find a list of non repeating elements from
the Data. Find all element combinations in the list and store the resulting
combination list combinationList.
for Combination in combinationList:
   count[Combination] = 0
   if Combination in Data[i]
      count[Combination] += 1
   end if
end for
Process 2:
//Comparing each different combination method, the element satisfying
the minimum confidence is selected as the association rule.
For Combination in CombinationList:
  if Combination's support > minC
    apriList.add(Combination)
  end if
end for
Process 3: Repeat process 2
Output association rule table apriList
```

Table 2. Association rules output by Apriori algorithm

Preceding Item	Subsequent Item	Support
Slope less than 10°	Occurrence of geological disasters	39.5%
Slope is between 10°–20°	Occurrence of geological disasters	32.5%
Vegetation cover index is between 1–2	Occurrence of geological disasters	48.1%
Vegetation cover index is between 2–3	Occurrence of geological disasters	33.4%
Elevation less than 500 m	Occurrence of geological disasters	42.3%
Mountainous areas	Occurrence of geological disasters	72.5%
Fracture zone	Occurrence of geological disasters	63.5%
Loose rock formation	Occurrence of geological disasters	53.1%
Forest land	Occurrence of geological disasters	46.8%
Farmland	Occurrence of geological disasters	30.4%

conditions for geological disasters. To sum up, geological hazard risk has a certain correlation with surface slope. Especially in the slope range of 0–20°, the geological hazard risk increases with the increase of slope. When the slope gradient is greater than 20°, the geological hazard risk will not increase all the time.

2) The support of leaf area index in the range of 1–2 is the largest, followed by 2–3. This result is different from the expected situation that the smaller the leaf area index is, the greater the support is. The reason is that the area with leaf area index less than 1 belongs to the area with sparse vegetation, and the surface layer of this area

may be hard rock and soil, resulting in sparse vegetation. Hard rock and soil areas are not prone to geological disasters. Therefore, the support of leaf area index less than 1 is small.

3) Many geological disasters are mainly concentrated in areas with an elevation of less than 500 m. Only areas with an elevation of less than 500 m exceed the minimum support. The elevation of an area can reflect the strength of human activities to a certain extent. Generally, human activities will be more frequent and intense in low altitude areas, and the risk of geological disasters in this area is greater.
4) Most geological disasters occur in mountainous areas, and their support degree reaches 72.5%. This is because the mountainous terrain area has many steep slopes, which provides internal conditions for the occurrence of geological disasters. From the data in Table 2, it can be concluded that there is an obvious correlation between terrain type and geological disaster risk. The risk of geological disasters in mountainous areas is the most likely, and the risk of geological disasters in other terrain types is significantly lower than that in mountainous areas.
5) Geological disasters are easy to occur in the area of fracture zone structure, and its support degree reaches 63.5%. The main reason is that the internal stress of the earth in the structural area of the fracture zone is large, and the structural activity is relatively intense, so the risk of geological disasters in this kind of geological structural area is high. At the same time, the structural area of the fracture zone is often the area with high earthquake incidence, which is another important reason for the frequent occurrence of geological disasters in the structural area of the fault zone. The earth's internal stress in other structural areas is less than that in fault zone structural areas, and the risk of geological disasters is also lower than that in fracture zone structural areas.
6) Geological disasters often occur in areas with loose rock and soil, with a support degree of 53.1%. Because the surface of the loose rock group area is covered with loose geotechnical structure, the water absorption reaction is also very obvious, which is very easy to provide material conditions for the occurrence of geological disasters. On the contrary, areas with stable geotechnical structure and hard formation lithology are not easy to provide material conditions for the occurrence of geological disasters, and the risk of geological disasters is small.
7) Forest land occupies a high degree of support in the surface coverage types of geological hazard points. According to the analysis of point (4), about 72.5% of geological disasters occur in mountainous areas. Most of the surface cover types in mountainous areas are forest land. Although this kind of surface cover can solidify water and soil loss, geological disasters may still occur in forest land areas due to the excitation of some external incentives (such as earthquake, heavy precipitation, etc.). In addition to forest land, cultivated land also accounts for 30% of the support. Cultivated land is the most common surface type in human development. Human activities will affect the stability of the surface, thus increasing the risk of geological disasters in the region.
8) In the association rule table, there is no slope aspect type item with support greater than 30%, which indicates that slope aspect is not directly related to geological disaster risk.

4.2 Analysis of External Influencing Factors of Geological Disaster Risk Based on Descriptive Statistics

This paper uses descriptive statistical analysis method to analyze the relationship between geological disaster risk and external influencing factors. For example, in the analysis of the relationship between geological disasters and rainfall, soil humidity and surface temperature, not only the data on the day of geological disasters, but also the data on the ten days before geological disasters are used to analyze the relationship between the dynamic changes of these external incentives and the risk of geological disasters.

1) **Statistical Analysis of Geological Disaster Risk and Rainfall.**

As one of the main external inducing factors of geological disasters, the impact of rainfall on geological disaster risk is reflected in the following two aspects: a) When rainfall occurs, with the continuous infiltration of rain, the water content of soil increases, and the softening degree of rock and soil layer also increases, which reduces the stability and shear strength of regional rock and soil. b) The continuous scouring of rainfall also brings external forces to the regional surface. At the same time, the weight of rock and soil layer is also increasing, which increases the possibility of displacement of regional surface rock and soil and increases the risk of geological disasters.

In order to further quantify the impact of rainfall on geological disaster risk. Based on the information data of Rainfall Type geological disasters in China in recent ten years, this paper counts the total rainfall on the day of geological disasters.

Rainfall (in mm) is classified as follows: $<$ 10, 10–25, 25–50, 50–100, 100–250, $>$ 250. The total rainfall on the day of geological disaster is grouped according to the rainfall level, and the statistical data are shown in Fig. 1.

As can be seen from Fig. 1, the proportion of disasters with total rainfall in the range of 10–100 mm is about more than 80%. The rainfall is between 0–50 mm, and the proportion of geological disasters increases with the increase of rainfall. The proportion of land disasters with rainfall in the range of 50–100 mm is smaller than that in the range of 20–50 m. The proportion of geological disasters with rainfall greater than 100 mm is smaller. The reason is that when the daily rainfall is in the range of 10–50 mm, it

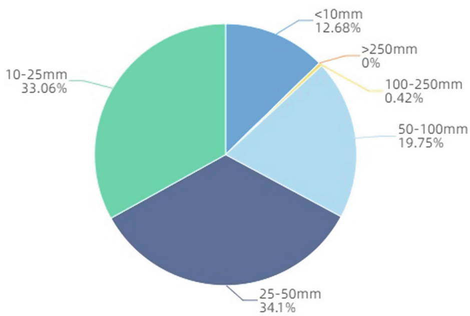

Fig. 1. Statistics of total rainfall on the day when the disaster occurs

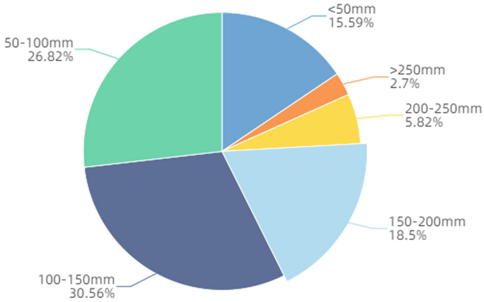

Fig. 2. Statistics of total rainfall ten days before the disaster

belongs to moderate rain and heavy rain, which has provided enough external force for geological disasters.

In addition, this paper also makes grouping statistics on the total rainfall within ten days before the occurrence of geological disaster at an interval of 50 mm, so as to analyze the long-term impact of rainfall on geological disaster risk. The statistical data are shown in Fig. 2.

Compared with Fig. 1 and Fig. 2, it can be found that they have roughly the same distribution. The range of 0–200 mm accounts for the largest proportion of the total rainfall in the ten days before the disaster, while in the range of 0–150 mm, the proportion of the disaster increases with the increase of rainfall. When the rainfall is greater than 150 mm, the number of geological disasters decreases with the increase of rainfall. The reason for the above situation is consistent with the total rainfall of the day.

2) **Statistical Analysis of Geological Disasters and Soil Moisture.**

Soil moisture can reflect the content of soil moisture in a region. The most common way to increase soil moisture is rainfall. This paper makes a statistical comparison between the average soil moisture on the day of geological disaster and the first ten days. The proportion that the average soil humidity on the day of geological disaster is greater than that in the previous ten days is 93.56%. It can be seen that before the occurrence of geological disasters, the water in the soil has a process of accumulation and growth, and the soil moisture will rise. Here, calculate the increase value of the soil humidity value on the day of geological disaster minus the average soil humidity value in the previous ten days, and count the proportion of different increase ranges. The statistical data are shown in Fig. 3.

It can be seen from Fig. 3 that the average soil humidity of the disaster site on the day of the disaster has an obvious increase compared with the average soil humidity of the previous ten days. Among them, the proportion of geological disasters with the increase of soil moisture in the range of 5–10% is 35.78%, which accounts for the largest proportion. The increase of soil moisture in the range of 10–15% also accounts for 21.78%. The above statistical results fully show that there will be an obvious process of soil water accumulation and increase before the occurrence of geological disasters. When the soil moisture in an area increases continuously, the stability of rock and soil

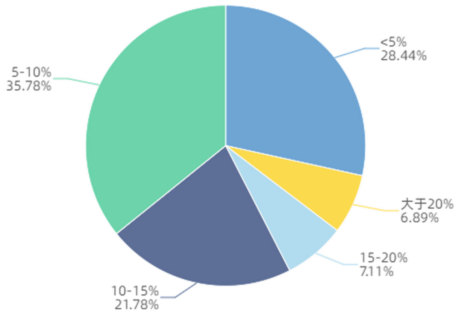

Fig. 3. Statistical of average humidity increase on the day of disaster compared with the previous ten days

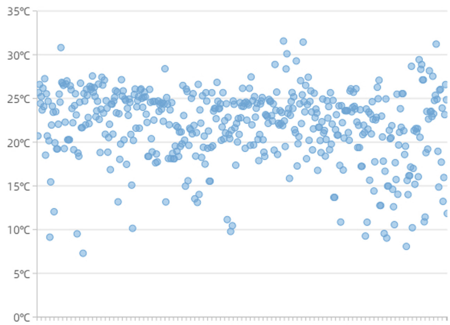

Fig. 4. Scatter diagram of average surface temperature distribution on the day of geological disaster

layers in the area will decrease, and the bearing capacity of internal and external forces will become smaller and smaller. After the soil humidity rises to a certain extent, once there is rainfall scouring or other external forces, geological disasters will occur in this area.

3) **Statistical Analysis of Occurrence of Geological Disasters and Surface Temperature**

The surface temperature is mainly affected by seasonal changes. To a certain extent, it will also affect the evaporation of water in the soil, resulting in the change of soil humidity. In order to specifically analyze the relationship between surface temperature and geological disaster risk, the scatter data relationship of average surface temperature on the day of geological disaster is given here, as shown in Fig. 4.

As can be seen from Fig. 4, the scattered points are mainly distributed between 15–30 °C, of which the number of points in the range of 20–25 °C is the most. This distribution is consistent with the fact that geological disasters mostly occur from May to September. China has the most rainfall in a year from May to September, and a large amount of rainfall has induced the occurrence of geological disasters. Simply according to the analysis of surface temperature, the higher the surface temperature, the more evaporation of soil moisture, which reduces the soil moisture. However, most areas in China have the climatic characteristics of the same period of rain and heat, that is, the negative impact of rainfall and the positive impact of surface temperature reach the maximum in the same period. There is a certain relationship between surface temperature and geological disaster risk, but its impact is obviously weaker than that of rainfall.

5 Conclusion

In this paper, Apriori correlation analysis algorithm and descriptive statistical analysis method are used to analyze the geological disaster risk and influencing factors of the collected geological disaster sample data set. The conclusions are as follows:

1) Geological disaster risk is related to many internal factors, such as terrain type, elevation, slope, geological structure, formation lithology, vegetation coverage, etc. It has stronger correlation with terrain type and geological structure, relatively weak correlation with vegetation cover, surface cover type, surface elevation and slope, and no obvious correlation with slope direction.
2) Geological disaster risk is also related to many external factors, such as regional rainfall, soil humidity and surface temperature range. The total rainfall of that day is within a certain range from that of the previous ten days, and the geological disaster risk increases with the increase of rainfall. The average soil humidity on the day of geological disaster increased significantly compared with the average soil humidity on the previous ten days. From May to September, when the surface temperature is high, the proportion of geological disasters is large.

Acknowledgment. Thanks to the China Railway Eryuan Engineering Group. Co. Ltd for providing research project support, as well as colleagues from the Academy of Science and Technology for their encouragement and help.

References

1. Li C Y, Meng H, Zhang R L, et al., "Risk assessment of geological hazards in county units in China", Journal of Hydrogeology and Engineering Geology, vol. 44, pp. 160–166, Mar. 2017.
2. Guan F J, Shen W Z, Zhang Z F, "Analysis, research and trend prediction of national geological disaster prevention and control", Journal of Chinese geological hazards and Prevention, vol. 29, pp. 1–2, Feb. 2018.
3. Wang Y, Lin Q G, Shi P J, "Spatial pattern and influencing factors of geological disaster casualties in China", Journal of Geography, vol.72, pp. 906–917, May. 2017.
4. Lv M N, "Study on Influencing Factors of collapse geological disaster in Guangzhou", Journal of Chinese geological hazards and Prevention, vol.31, pp. 127–133, Apr. 2020.
5. Huang H C, "Study on distribution law and influencing factors of main geological disasters in Shanxi", Journal of Energy and environmental protection, vol.42, pp. 87–90, Jan. 2020.

6. Li R W, Wang N Q, "Coupling analysis of geological hazards and influencing factors in Yan'an City", Journal of Science, technology and Engineering, vol.19, pp. 100–106, Jun. 2019.
7. Nie B Q, Tang M G, Shao S, et al., "Sensitivity analysis of influencing factors of wading slope stability based on grey correlation method", Journal of Changjiang Academy of Sciences, vol.36, pp.123–126, Jan. 2019.
8. Yi J J, "Discussion on Influencing Factors of environmental geological disasters", Journal of World nonferrous metals, pp.208–210, Aug. 2017.
9. Liu C Z, Chen C L, "Analysis on the causes of geological disasters in China", Journal of Geological comments, vol.66, pp.1334–1348, Sept. 2020.
10. Lin C Z, Wang N T, Wang Q, et al., "Research Progress on geological disaster early warning and prediction of rainfall landslide", Journal of Rainstorm disaster, vol.37, pp.393–400, Oct. 2018.
11. Luo Y F, He J, Zhao H J, et al., "Analysis of risk factors of stroke based on Apriori association rules", Journal of Chinese Digital Medicine, vol.12, pp.85–88, Nov. 2017.
12. Gu H Q, "Statistical analysis and chart of baseline information in clinical research", Journal of Chinese evidence based cardiovascular medicine, vol.10, pp.910–912, Aug. 2018.