Analysis of Rainfall Characteristics Caused by Urban Geological Disasters: Evidence from Wuhan

Ming Li[1]**, Linyi Li**[2]
[1] Wuhan Meteorological Observatory, Wuhan 430040, Hubei, China
[2] Cardiff University, Cardiff, UK
DOI: 10.32629/aes.v4i4.1384

Abstract: The research indicates that geological disasters are primarily influenced by weather and climatic conditions, in addition to geological stability factors . Typically, the prediction and forecasting of geological disasters are based on geological conditions, combined with rainfall forecasts. Geological conditions serve as the foundation for the occurrence of geological hazards, often necessitating geological hazard zoning within the corresponding region through disaster investigations and associated calculations. This zoning is further analyzed to predict the risk of geological disaster occurrence based on factors such as prior rainfall frequency, rainfall intensity, and rain type characteristics.
Keywords**:** machine learning method, geological hazard, analysis

1. Introduction

Situated in the heart of central China, Wuhan is located at the confluence of the Yangtze River and the Han River. It marks the transition zone between the Jianghan Plain and the Dabie Mountains' hilly and mountainous terrain. Wuhan is considered a region with relatively stable crust. However, due to the distribution of ancient river channels, ancient lakes, and modern land reclamation projects, the geological structure of Wuhan is quite complex. Wuhan's geological conditions give rise to potential urban geological issues, including faults, karst underground rivers, landslides, and floods. Wuhan's geological department surveys indicate that with the implementation of various construction projects such as road upgrades, rail transit, cross-river passages, and underground space development in recent years, human engineering activities have become increasingly intense. This has led to the alteration and disruption of the natural geological environment, resulting in continuous geological disasters. These disasters not only impact urban planning, construction, and management but also pose a threat to urban safety.

This article aims to uncover the patterns of the impact of heavy rainfall on geological disasters in Wuhan and attempts to enhance meteorological risk forecasting methods, thereby providing technical support for fine-grained geological disaster forecasting.

2. Data Sources and Calculation Methods

2.1 Data Sources

Historical geological disaster data were obtained from the formal survey results of the 1:50,000 environmental geological survey conducted in the Wuhan Urban Development Area since 2012. Due to variations in survey methods, areas of interest, and technological constraints, early survey results may not have comprehensive and complete records of geological disaster occurrence times, locations, and magnitudes. Therefore, this study utilized data from the Wuhan Geological Department's summary of emergency response to geological disasters in Wuhan from 2016 to 2021 and geological disaster statistics. These disaster records were obtained through emergency investigations and on-site surveys by relevant authorities and were subject to review and compilation.[1]

Rainfall actual data were collected from 19 national meteorological observatories and 510 regional automatic weather stations within the range of 113.6° to 115.2°E and 29.8° to 31.4°N, as presented in monthly observation data reports. Precipitation data were extracted for daily rainfall within the period from 20:00 (BTC) of the previous day to 20:00 of the current day and have undergone quality control.

2.2 Computational Methods

Let D_h represent the date of a geological disaster at a specific location, and compile a time series of daily rainfall data at the nearest grid point, denoted as $R = \{r\_1, r\_2, ..., r\_i, ..., r\_n\}$, with $i \in n$, for determining the characteristics of D values over the past 6 years. Identify the dates of heavy rainfall or more significant precipitation at neighboring grid points before the geological disaster, referred to as D_r. Calculate $D = D_h - D_r$.[2]

Ridge regression, also known as L2 regularization or Tikhonov regularization, is a commonly used regularization method in machine learning for performing regression analysis on ill-posed problems. When you have M samples with N features, and $N > M$, using a linear regression model can lead to an issue where (X^T X)^(-1) is not a full-rank matrix and thus not invertible, preventing you from calculating regression coefficients. Ridge regression introduces an additional term to the least squares estimate, known as the ridge regression estimate:

$$w = (X^T X + \lambda I)^{-1} X^T y \tag{1}$$

In this equation:w represents the ridge regression coefficients.[3] λ is the ridge coefficient, a non-negative value. In practice, λ can be determined based on the ridge trace plot. I is the identity matrix (with all diagonal elements equal to 1 and all other elements equal to 0).When $\lambda = 0$, you obtain the least squares solution. As λ increases, the ridge regression coefficients, w, approach 0. Ridge regression improves upon least squares estimation by adding an L2 norm penalty term to the linear model, ensuring that (X^T X + λI) is full rank and thus invertible. By sacrificing unbiasedness in favor of reducing some feature information and lowering model precision, ridge regression provides more practical regression coefficients.[4]The key aspect in ridge regression estimation is the selection of the ridge coefficient λ. Currently, there is no universally accepted formula for determining the optimal value of λ. Generally, λ is chosen by plotting a ridge trace based on the sample data.

3. Geological Hazard Overview in Wuhan

Based on data collected since 2012, geological disasters in Wuhan are mainly categorized into two types. One type comprises slope-related geological disasters, such as landslides and collapses. The other type includes subsidence geological disasters, which commonly manifest as ground subsidence and collapse in karst regions.[5] A total of 68 potential geological hazard points were identified and verified throughout the city, primarily involving five types of disasters, including landslides, collapses, ground subsidence, and others . Slope-related geological disasters, like landslides, collapses, and mudslides, constitute the majority at 73.83%.

From 2016 to 2021, Wuhan experienced 167 days with 254 geological disasters, with a concentration of these events in densely populated areas of the northern mountainous regions, central urban areas, surrounding new urban districts, and the southern plains and lake areas. Comparing this with the existing potential geological hazard points, it is evident that geological disasters mainly occur at or near these identified locations.

4. Preliminary Rainfall Characteristics of Geological Hazards

4.1 Preliminary Rainfall Statistics

Between 2016 and 2021, Wuhan experienced a total of 26 instances of D_r, occurring within the timeframe of April 11, 2017, to August 23, 2021, with a maximum 24-hour rainfall of 264.7 mm. The average $\bar{D}$ amounted to 4.26 (days), and in 25 of these cases, D was less than or equal to 13 days. This implies that geological disasters occurred within 13 days following D_r, with a total of 64 incidents. Considering the continuity of rainfall and building upon previous research, this study adopted n=15 days to statistically analyze the preliminary rainfall characteristics of geological hazards.

Between 2016 and 2021, Wuhan city experienced a total of 254 geological disasters. However, one of these cases was excluded due to an unknown cause (ground heating) and the difficulty in classification. Out of the remaining 253 geological disasters, 64 of them were preceded by heavy rainfall within the time frame of [-15 to 0 days], with a probability of occurrence amounting to only 25.29%. The statistical peak occurred at -1 and 0 days, indicating that geological hazards were most likely to occur within the same day and up to 48 hours after heavy rainfall.[6]

Due to the differing influencing factors for the development and formation of slope-related and subsidence collapse geological hazards, the study conducted statistical analyses for the occurrences of these two types of geological disasters, daily average rainfall within the 15 days prior, and the frequency of heavy rainfall within that period for the 254 geological disasters that occurred between 2016 and 2021.

Classifying the 253 geological disasters that occurred between 2016 and 2021 into slope-related and subsidence-related geological hazards, there were a total of 52 subsidence-related and 201 slope-related incidents.

4.2 Feature Analysis

For the rainfall time series Rd ($d \in$ [-15,0]), although each Rd is treated as an independent variable, there is often a certain degree of correlation between neighboring Rd due to the climatic characteristics of heavy rainfall in Wuhan and the constraints of daily rainfall statistics. Calculating the correlation coefficients between adjacent columns of the matrix composed of Rd time series for the 253 geological disasters, the maximum correlation coefficient reaches 0.51. For the 26 sequences with heavy rainfall, the highest correlation coefficient between adjacent columns reaches 0.94, indicating significant multicollinearity among the variables. This justifies the use of ridge regression analysis.

A machine learning method based on ridge regression was employed to analyze the data. The analysis determined that $\lambda=1$ is suitable for the ridge regression model. This approach was used to train a classification model. The results of various regression models obtained after cross-validation are shown in Figure 1.

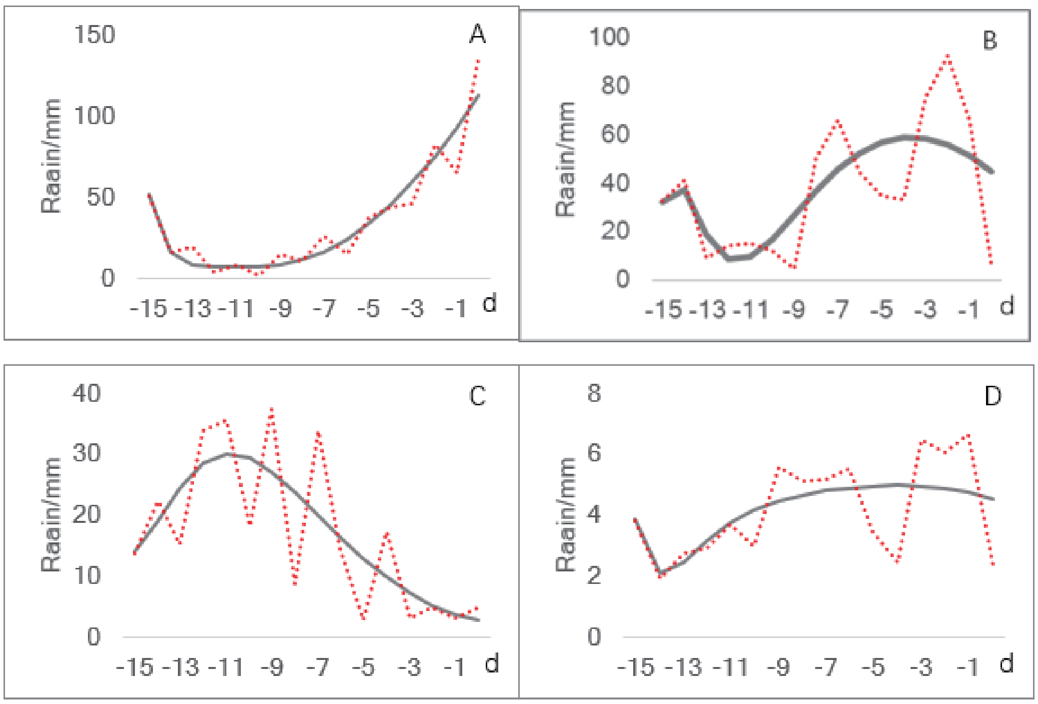

Figure 1. Rainfall average and ridge regression results of 0~15 days before geological disaster (A: d=0;B: d ∈ [-1,-3];C: d ∈ [-4,-15];D: No heavy rainfall)

From Figure 1, we can observe that the average rainfall for near-term periods is 135mm, while the regression value is 111mm. The maximum value of short-term rainfall averages at 92.5mm, occurring at -2d, with a peak in the regression curve at 75.4mm on -3d. For the mid-term rainfall, the peak values for both the average curve appear at -7d, -9d, and -11d, while the peak in the regression curve appears at -11d. All of these peaks are below 50mm. The decreasing trends in the rainfall values are due to variations in the number of days in each interval, which cause the data to be dispersed. Therefore, it is recommended to analyze and assess the maximum values or cumulative rainfall for the [-1,-3] and [-4,-15] intervals separately.

Based on the criteria in Table 1, the 253 geological disasters are categorized into four types: "Near-term" type, where heavy rainfall occurs on the same day as the geological disaster (d=0); "Short-term" type, where heavy rainfall occurs in the 1st to 3rd day prior to the geological disaster (d ∈ [-1,-3]); "Mid-term" type, where heavy rainfall occurs in the 4th to 15th day prior to the geological disaster (*d* ∈ [-4,-15]); and "No heavy rainfall" type, where there is no heavy rainfall in the 15 days before the geological disaster.

Traditionally, the calculation of effective rainfall mainly considers the influence of near-term rainfall and partially accounts for short-term rainfall. However, the estimation for mid-term rainfall impact is often insufficient. In situations without heavy rainfall, there is generally continuous rainfall within 0-15 days before the disaster occurrence (Figure 1D).

Table 1. Statistics of effective rainfall in the early period of geological disasters in Wuhan from 2016 to 2021

Rc	<10mm	10-25mm	25-50mm	50-100mm	100-200mm	>=200mm
Frequency	63	62	36	32	24	37
Proportion (%)	24.80	24.41	14.17	12.60	9.45	14.57

5. Conclusion

(1) Geological disasters in Wuhan are primarily categorized into two types: slope-type geological disasters, including landslides and collapses, and subsidence-type geological disasters, such as ground subsidence and collapses. The influencing factors for these two types of geological disasters differ significantly. Based on the locations and types of disasters, geological disasters in Wuhan predominantly occur in mountainous areas, hilly regions, and lake areas, primarily due to the geological characteristics of each region.

(2) From 2016 to 2021, Wuhan experienced a total of 26 days with heavy rainfall or above, and 25 of these days were followed by geological disasters occurring within 0 to 13 days. During this period, a total of 254 geological disasters occurred over 167 days. The analysis of effective rainfall calculated from historical geological disaster data revealed that the probability of strong rainfall occurring in the 0 to 13 days prior to geological disasters ranged from 9.38% to 23.44%, with an average of 13.96%. Strong rainfall events were most likely to lead to geological disasters within the first 48 hours after their occurrence

Author Profile: Li Ming, male, born in Wuhan, Hubei Province, senior engineer, Wuhan Meteorological Observatory, 430040, mainly engaged in weather forecasting and early warning and meteorological service technology research in big cities.

Acknowledgments

This paper was supported by the following projects: Wuhan Knowledge Innovation Project (2022022101015009); Key Project of Wuhan Meteorological Bureau (WHZ202201).

References

[1] Chen Yobin, Yang Wenfa, Chen Li, et al. Application of a method for estimating critical rainfall in watershed based on Mahalanobis distance identification [J]. Water Resources and Hydropower Technology,2016,47(1):21-25.

[2] Gao Xinbo. Fuzzy Cluster Analysis and its Application [M]. Xidian University Press,2004.

[3] Huang Chuhui, Li Guoping, Zhang Fangli, et al. Evolution characteristics of heavy rain events in Sichuan mountains under the influence of climate change in recent 10 years [J]. Heavy Rain Hazards, 2019,39(04):335-343.

[4] Huang Runqiu, Xiang Xiqiong, Ju Neng Pan. Current situation and problems of regional geological hazard evaluation in China [J]. Geological Bulletin of China, 2004(11):1078-1082.

[5] Jin Guodong, LIU Yancong, NIU Wenjie. Comparison between distance-weighted inverse ratio interpolation method and Kriging interpolation method [J]. Journal of Changchun University of Technology (Natural Science Edition),2003,024(003):53-57.

[6] Liu Peiting, Liu Min, Chen Wei, et al. Analysis on the characteristics of annual extreme hourly precipitation in Central China [J]. Rainstorm Hazards, 2019,39(5):508-515. (in Chinese)