

## *Retraction*

# **Retracted: Risk Assessment and Prediction of Rainstorm and Flood Disaster Based on Henan Province, China**

### **Mathematical Problems in Engineering**

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] G. Deng, H. Chen, and S. Wang, "Risk Assessment and Prediction of Rainstorm and Flood Disaster Based on Henan Province, China," *Mathematical Problems in Engineering*, vol. 2022, Article ID 5310920, 17 pages, 2022.

## Research Article

# Risk Assessment and Prediction of Rainstorm and Flood Disaster Based on Henan Province, China

Guoqu Deng , Hu Chen , and Siqi Wang

Henan University of Science and Technology, Luoyang 471023, China

Correspondence should be addressed to Guoqu Deng; dengguoqu@haust.edu.cn

Received 10 October 2021; Revised 2 December 2021; Accepted 11 December 2021; Published 18 February 2022

Academic Editor: Lazim Abdullah

Copyright © 2022 Guoqu Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To reasonably evaluate and predict the loss of rainstorm and flood disaster, this study is based on the rainfall data and rainstorm and flood disaster data of 18 cities in Henan Province from 2010 to 2020, using GIS technology and weighted comprehensive evaluation method to analyze the risk of rainstorm and flood disaster factors in various regions. The four risk factors of hazard risk, hazard-pregnant environment sensitivity, hazard-bearing body vulnerability, and disaster resilience were analyzed in compartment analysis. At the same time, a new rainstorm and flood disaster prediction model was constructed in combination with the hybrid PSO-SVR algorithm. The research results show that there are many rivers in Henan Province, the terrain tends to be higher in the west and lower in the east, and most areas are low plains, making most cities in Henan Province at a moderate risk level. For the more developed cities such as Zhengzhou, Luoyang, and Nanyang, the hazard risk, sensitivity, vulnerability, and disaster resistance are high, and they are prone to heavy rains and floods. For the economically underdeveloped, the terrain is high or hills, such as Sanmenxia City; Xinyang City and other places have low hazard risk and are not prone to rainstorms and floods. By constructing a hybrid PSO-SVR model, selecting two representative cities of Zhengzhou and Luoyang, and predicting the daily rainfall, the number of disasters, and the direct economic loss, the calculated RMSE and MAPE values are both less than GA-SVR, the traditional SVR, and BPNN models, which have verified the superiority of the model proposed in this study and the practical value it brings. To further verify the prediction accuracy of the hybrid model, the average value of RMSE and MAPE of other 16 cities are calculated, and the result is still smaller than other three models, and the study can provide some decision-making references for the urban rainstorm and flood management.

## 1. Introduction

The rainstorm disaster risk assessment is based on some obtained disaster evaluation index values, which is a comprehensive evaluation of the risk of hazard factors, sensitivity of hazard-pregnant environment, vulnerability of hazard-bearing body, and disaster resilience in the study area, providing reference for regional rainstorm risk management, flood prevention, and disaster mitigation planning [1–4]. It also provides a reasonable reference for long-term social and economic development planning [5–11]. With the deepening of high-intensity human engineering activities on a global scale, the global climate warming trend is becoming more and more significant, and extreme weather events occur frequently and cause a series of disasters. According to the global risk report released by World Economic Forum (WEF), extreme

weather emergencies from 2017 to 2020 have ranked first in terms of the probability of occurrence of the top ten global risks for four consecutive years. Affected by factors such as global temperature rise, sea level rise, and surface subsidence in some areas, the frequent and widespread problems of storm and flood disasters have become increasingly prominent [12–15]. Rainstorm and flood disasters have become one of the most common and serious natural disasters in many large cities in the world. In recent years, cities in the social and economic development centers of developing countries have suffered heavy rain disasters [16–18]. In this context, it is particularly important to conduct a scientific and reasonable risk assessment of storms and floods.

Zandvoort and Vlist [19] used a multilayer safety analysis method to evaluate rainstorm disasters to design a more robust and effective avoidance plan to improve the disaster

resistance in the area and reduce post-disaster losses. Otar et al. [20] constructed a model of the relationship between hazards in Georgia and used the geographic information system (GIS) analysis to obtain a zoning map of storm and flood risk. Benito et al. [21] integrated geological conditions, water conservancy construction, and historical statistical disaster records based on long-term extreme storm and flood risk assessment and constructed a multidisciplinary storm disaster assessment model. Alfa et al. [22] developed a storm flood risk assessment system including height and slope based on the Ofu River Basin in Nigeria. Weerasinghe et al. [23] used indicators such as hazard risk and carrier vulnerability to assess the flood risk level and constructed a rainstorm risk assessment system in the western provinces of Sri Lanka. Mahootchi and Golmohammadi [24] extended the two-stage stochastic optimization mathematical model to reduce the losses caused by the rainstorm disaster and further control the emergency cost and developed a rainstorm disaster model with higher matching degree. Alsubaie et al. [25] constructed a systematic disaster response planning platform based on Supervisory Control and Data Acquisition (SCADA).

With the rapid development of China's urbanization process, the frequent extreme rainstorms in major cities in China have caused huge losses to the society and economy [26–28]. Zhang and Cai [29] relied on fuzzy mathematics theory and used quantitative analysis to measure the ambiguity of risk factors and then constructed a fuzzy comprehensive index evaluation system. Peng et al. [30] simulated different causes of the degree of flood risk in the Maozhou River Basin under the dangerous situation of disaster factors, which are based on the volume submergence algorithm. However, the above methods cannot evaluate composite systems with multiple uncertainties. In recent years, with the rapid development of artificial intelligence technology, many scholars tend to apply intelligent algorithms to various evaluation tasks. Feng and Chen [31] used support vector machines (SVM) for the first time to establish a regression inference rainfall forecast model in the Sichuan Basin. The test results show that the SVM inference model has good forecasting capabilities. Xie et al. [32] established a support vector regression (SVR) rainfall prediction model based on the high-dimensional nonlinearity and periodicity of rainfall data, using seasonal autoregressive input feature selection method and grid parameter optimization method. In addition, some scholars have used algorithms such as backpropagation neural network (BPNN) and extreme gradient boosting (XGBoost) in the rainstorm disaster assessment model [33–36], but due to the uncertainty of model parameter settings and research, the incompleteness of the construction of the index system in the process caused the accuracy of model evaluation to be greatly reduced.

In summary, based on disaster risk theory, comprehensive consideration of the natural environment, and social economic conditions, to further improve the rainstorm and flood disaster risk assessment indicators and methods [37–40], this research will first use the logistic binary regression analysis to affect the occurrence of rainstorm and flood disaster. Then, with the help of the weighted

comprehensive evaluation method and GIS-related technology, the compartment analysis of rainstorm and flood disaster in Henan Province was carried out. Finally, the particle swarm optimization (PSO) algorithm was combined with SVR to establish a hybrid PSO-SVR rainstorm and flood disaster model [41, 42], which is expected to provide a scientific basis for the comprehensive evaluation of storm and flood disasters.

## 2. Research Data and Index System Construction

**2.1. Overview of Study Area.** Henan Province is located in the middle and lower reaches of the Yellow River in east-central China. It is bounded by  $31^{\circ}23'$ – $36^{\circ}22'$  north latitude and  $110^{\circ}21'$ – $116^{\circ}39'$  east longitude. As for the neighboring provinces, Henan borders Anhui and Shandong to the east, Hebei and Shanxi to the north, Shanxi to the west, and Hubei to the south. Of the total area of approximately  $167,000 \text{ km}^2$ , there are  $74,000 \text{ km}^2$  of mountains and hills (44.3% of the province's total area) and  $93,000 \text{ km}^2$  of plains and basins (55.7% of the province's total area). The terrain of Henan is high in the west and low in the east, with the altitude difference between the highest point and the lowest point as high as 2390.6 meters. Plains are widely distributed in the province, among which the central, eastern, and northern plains are formed by the alluvial accumulation of the Yellow River, Huaihe River, and Haihe River, also known as the Huanghuaihai Plain [43]. Due to its location in the warm temperate zone and subtropical zone, it belongs to a humid-semihumid monsoon climate with large seasonal differences. The annual average temperature is between  $12^{\circ}\text{C}$  and  $16^{\circ}\text{C}$ , and the annual average precipitation is about  $500 \text{ mm}$ – $900 \text{ mm}$ . There are many mountains in the south and west. Especially, the precipitation of Dabie Mountains can reach more than  $1100 \text{ mm}$ . Annual precipitation is concentrated in summer, often accompanied by torrential rain disasters [44]. This paper lists the historical rainstorm process information in Henan Province (see Table 1). Data come from the China Meteorological Administration.

According to data about July 2021 released by the Henan Meteorological Bureau, rainstorm occurred in many areas in Henan. Within the 13 days from July 10 to July 23, there were 11 days of rainfall in Zhengzhou. It should be noted that the rainfall from 16:00 to 17:00 on July 20 reached  $201.9 \text{ mm}$  per hour, breaking through the historical extreme value in mainland China. The rainstorm and flood disaster has caused great damage to people's lives in Henan province, causing 302 deaths and 50 missing. Among them, 292 people were killed and 47 people were missing in Zhengzhou City; 7 people were killed and 3 people were missing in Xinxiang City; 2 people were killed in Pingdingshan City; 1 person was killed in Luohe City. In view of this, the construction of rainstorm and flood disaster risk assessment model and the implementation of comprehensive evaluation of the disaster in this region can provide strong technical support for the research on storm flood management and risk warning services.

TABLE 1: Historical rainstorm and precipitation process information in Henan Province.

| Time (year/<br>month/day) | Main distribution area of<br>rainfall (city) | Maximum accumulated<br>rainfall (mm) | Maximum daily rainfall (mm)<br>(month/day) | The average rainfall of the<br>province (mm) |
|---------------------------|--|--------------------------------------|--|--|
| 1963/8/2-8/8              | Xinxiang                                     | 699                                  | 253 (8/3)                                  | 199  |
| 1975/8/4-8/8              | Zhumadian                                    | 1631                                 | 755 (8/7)                                  | 163  |
| 1982/7/28-8/5             | Luoyang                                      | 666                                  | 265 (8/1)                                  | 169  |
| 1996/7/28-8/6             | Xinyang Zhumadian                            | 438                                  | 249 (8/3)                                  | 135  |
| 2005/6/25-26              | Kaifeng Zhumadian                            | 496                                  | 160 (6/26)                                 | 107  |
| 2007/8/2-8/3              | Zhengzhou Nanyang                            | 374                                  | 87 (8/2)                                   | 132  |
| 2008/7/22-7/23            | Nanyang                                      | 542                                  | 229 (7/22)                                 | 142  |
| 2010/7/18-7/20            | Zhengzhou Luoyang                            | 596                                  | 182 (7/18)                                 | 150  |
| 2012/8/19-8/20            | Zhengzhou                                    | 295                                  | 78 (8/20)                                  | 114  |
| 2016/7/18-7/20            | Anyang                                       | 732                                  | 703 (7/19)                                 | 80   |
| 2021/7/16-7/20            | Zhengzhou                                    | 617                                  | 362 (7/20)                                 | 88   |

**2.2. Construction of Indices System.** The evaluation of rainstorm and flood disaster is a comprehensive evaluation process based on the selection of reasonable indicators. Therefore, the reliability of the evaluation results is inseparable from the correctness of the index selection. In this research, based on the theory of disaster risk system, referring to existing research results and considering the availability of data for rainstorms and floods, index system of rainstorm and flood disaster loss prediction is constructed from the aspects of hazard factors, hazard-pregnant environment, hazard-bearing body, disaster resilience, and disaster loss (see Table 2).

**2.2.1. Hazard Factors.** The distribution of rainstorms directly affects the possibility of flooding and the degree of loss of disasters. According to relevant studies, the duration of the rainstorm and the maximum rainfall in a short time are closely related to the occurrence of rainstorm and flood disaster. Based on this, considering the availability of data, this research selects the bulletin data issued by the urban meteorological stations corresponding to 18 cities in Henan Province as the benchmark and selects indicators such as process accumulated rainfall, continuous rainfall days, and accumulated rainfall in 12 h and 24 h.

**2.2.2. Hazard-Pregnant Environment.** The occurrence of rainstorm and flood disaster is closely related to the terrain. Generally speaking, the greater the terrain undulation, the less likely the occurrence of flood disasters; the smaller the terrain undulation, the greater the possibility of flood disasters. The degree of vegetation coverage directly affects the water conservation capacity to a certain extent. At the same time, the density of the river network in the study area can indirectly reflect the relative magnitude of the risk of rainstorms and floods to a certain extent. Therefore, places with high river network density are more likely to encounter flood disasters.

**2.2.3. Hazard-Bearing Body.** It mainly refers to the object of the impact of rainstorms and floods. At the same time, the spatial distribution of population, roads, and houses are inseparable from the extent of flood damage. On the one

hand, the GDP of per person of city's residents in the area reflects the development status of social and economic construction. On the other hand, it can also reflect the residents' ability to withstand rainstorm and flood disaster.

**2.2.4. Disaster Resilience.** It refers to the ability of a region to defend against rainstorms and floods. This article mainly considers the regional GDP and the average output value. On the one hand, it not only reflects the local fiscal revenue, but more importantly, it reflects the comprehensive flood resilience capability of the region in response to rainstorms and floods. Third, from the perspective of drainage capacity, the capacity of the drainage system directly affects the occurrence of flood disasters.

**2.2.5. Rainstorm and Flood Disaster Loss.** Based on the reliability and availability of disaster loss data caused by historical rainstorms and floods, this article selects accumulated rainfall in 24 h (daily rainfall), number of rainstorm and flood disasters, and direct economic loss as the target variables for the hybrid PSO-SVR model prediction.

**2.3. Data Sources.** This paper adopts a total of 2,700 pieces of data selected from various regions in Henan Province from 2010 to 2020.

**2.3.1. Hazard Factor Data.** Relevant data are collected from China Meteorological Network, China Meteorological Administration, Henan Meteorological Administration, and Henan Meteorological Observation Center, as well as the data of 120 meteorological stations corresponding to 18 cities in Henan Province. Considering that June to August is the rainy season every year, the data of 15 days with rainfall greater than 60 mm in each city from June to August are selected as the data of the current year for analysis and processing.

**2.3.2. Hazard-Pregnant Environment and Hazard-Bearing Body Data.** Data of the digital elevation model (DEM) come from geographic space data cloud (ASTER G-DEM 30M resolution digital elevation data), and data of

TABLE 2: Indices system of rainstorm and flood disaster loss prediction.

| Variable attributes                               | Disaster elements                    | Primary indices   | Secondary indices  | Reference source                                      |
|---|--------------------------------------|---|--|---|
| Characteristic variable                           | Hazard factors                       | Rainfall data   | Process accumulated rainfall ( $J1$ : mm)                | Gong et al. [45], Li et al. [46], Huang et al. [47]   |
|   |                                      |   | Continuous rainfall days ( $J2$ : h)                     |   |
|   |                                      |   | Accumulated rainfall in 12 h ( $J3$ : mm)                |   |
|   |                                      |   | Accumulated rainfall in 24 h ( $J4$ : mm)                |   |
|   | Hazard-pregnant environment          | Natural environment   | Digital elevation model (DEM) ( $Z1$ : m)                | Wang and Deng [48], Hu et al. [49]                    |
|   |                                      |   | Relief amplitude ( $Z2$ : m)                             |   |
|   |                                      |   | Vegetation coverage ( $Z3$ : %)                          |   |
|   |                                      |   | Density of the river network ( $Z4$ : $m^2/m^2$ )        |   |
|   | Hazard-bearing body                  | Social economic data  | Density of population ( $S1$ : thousand people/ $km^2$ ) | Gong et al. [50], Pan et al. [51], and Li et al. [46] |
|   |                                      |   | GDP of per person ( $S2$ : people/ten thousand yuan)     |   |
| Proportion of garden green space area ( $S3$ : %) |                                      |   |  |   |
| Proportion of house area ( $S4$ : %)              |                                      |   |  |   |
| Disaster resilience                               | Disaster resilience and construction | GDP of every city ( $K1$ : ten million yuan)                    | Zhang et al. [52] and Xu et al. [53]                     |   |
|   |                                      | GDP of per square kilometer ( $K2$ : ten million yuan/ $km^2$ ) |  |   |
|   |                                      | Length of urban drainage pipeline ( $K3$ : ten thousand km)     |  |   |
|   |                                      |   |  |   |
| Target variable                                   | Rainstorm and flood disaster loss    | Disaster loss situation   | Number of rainstorm and flood disasters ( $Y1$ : time)   | Zhang et al. [54] and Wang et al. [55]                |
|   |                                      |   | Direct economic loss ( $Y2$ : hundred million yuan)      |   |

topographic undulation (select the DEM standard deviation within the city to which you belong) and river network density come from the National Basic Geographic Information Center. Considering the uncertainty of vegetation coverage, vegetation coverage is selected as the evaluation standard of water conservation capacity. Vegetation coverage data come from “China Forestry Statistical Yearbook” of the database in China Forestry and Economic and Social Development Statistics. The carrier data come from the “China Statistical Yearbook” of 2010–2020 National Bureau of Statistics and “Henan Statistical Yearbook” of Henan Statistics Bureau.

**2.3.3. Disaster Resilience and Rainstorm and Flood Disaster Loss.** Relevant data come from “China Statistical Yearbook” of 2010–2020 National Bureau of Statistics, “National Water Regime Annual Report” of Ministry of Water Resources of the People’s Republic of China, and the “Statistical Bulletin” of the Ministry of Civil Affairs of the People’s Republic of China. To facilitate the calculation, the disaster loss caused by heavy rains and floods in each year is selected as the calculation data. Considering the availability of data released by meteorological stations, elevation distribution maps of meteorological stations in various regions of Henan Province are shown in Figure 1.

### 3. Risk Compartment Analysis of Rainstorm and Flood Disaster

After the construction of the indices system, the logistic regression model is used for correlation analysis to reduce

the impact of indices data on the error of the hybrid PSO-SVR prediction model. Since the model is a regression analysis for the binary variable of dependent variable, the data of each evaluation index are used as independent variables in the evaluation of flood disaster. In addition, the occurrence of rainstorm and flood disasters is represented by 0 (the number of disasters is 0) and 1 (the number of disasters is not 0) as the dependent variables of two classifications [56]. In this research, risk compartment analysis of rainstorm and flood disaster includes four parts of analysis. The risk analysis of hazard factors is mainly analyzing the influencing factors of risk sources. The sensitivity analysis of hazard-pregnant environment reflects the impact of natural geographical environment on rainstorm and flood disaster. The vulnerability analysis of hazard-bearing body mainly analyzes the influence of different rainstorm and flood intensity disasters for the distribution of population and the condition of regional economic and infrastructure. The analysis of disaster resilience reflects the level of the ability of disaster prevention and recovery [57–63]. This paper firstly presents the results of regression analysis (see Table 3).

According to the significance at the 95% level, the significance value greater than 0.05 indicates that the correlation of variables is low. Therefore, the impact of  $Z1$ ,  $S2$ , and  $K2$  will not be considered in this paper.

**3.1. Risk Analysis of Hazard Factors.** The disasters caused by rainstorm are mainly manifested as strong rains and high intensity. The rainstorm and flood disaster occurs when a short period of accumulated rainfall is too large to drain water. In this paper,  $J1$ ,  $J3$ ,  $J4$ , and  $J2$  are selected to

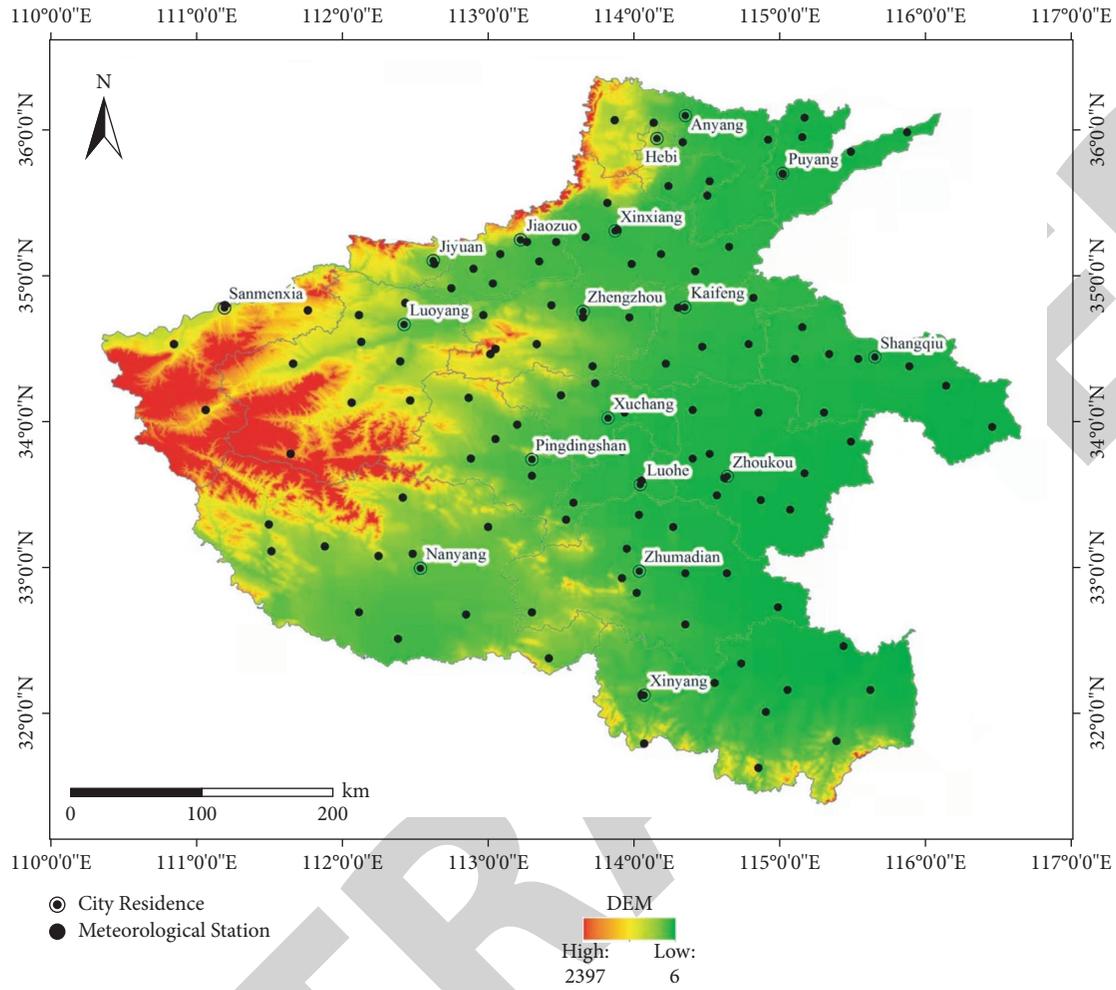


FIGURE 1: Spatial distribution of automatic meteorological observation stations in Henan Province.

TABLE 3: Results of logistic correlation regression analysis.

| Indices       | B      | S.E.  | Wald    | df | Significance level | Exp (B) | Confidence interval 95% exp (B) |             |
|---------------|--------|-------|---------|----|--------------------|---------|---------------------------------|-------------|
|               |        |       |         |    |                    |         | Lower limit                     | Upper limit |
| J1            | 5.216  | 2.767 | 2.314   | 1  | 0.000              | 1.355   | 0.627                           | 2.326       |
| J2            | 4.507  | 1.936 | 2.132   | 1  | 0.000              | 5.236   | 3.022                           | 7.658       |
| J3            | -2.064 | 0.216 | 18.125  | 1  | 0.024              | 0.235   | 0.029                           | 2.157       |
| J4            | 1.672  | 0.652 | 2.921   | 1  | 0.036              | 1.229   | 0.158                           | 2.192       |
| Z1            | -0.056 | 0.108 | 0.725   | 1  | 0.213              | 0.002   | —                               | —           |
| Z2            | -1.249 | 8.254 | 0.452   | 1  | 0.000              | 0.025   | 0.012                           | 0.704       |
| Z3            | -2.353 | 0.832 | 19.831  | 1  | 0.000              | 0.037   | 0.005                           | 0.236       |
| Z4            | 0.237  | 1.602 | 1.374   | 1  | 0.035              | 1.368   | 0.529                           | 2.871       |
| S1            | 1.203  | 0.362 | 0.257   | 1  | 0.016              | 0.765   | 0.233                           | 1.986       |
| S2            | 0.075  | 1.206 | 0.034   | 1  | 0.107              | 0.013   | —                               | —           |
| S3            | -1.215 | 3.371 | 1.562   | 1  | 0.047              | 1.103   | 0.716                           | 4.965       |
| S4            | 2.127  | 0.762 | 5.261   | 1  | 0.022              | 0.125   | 0.062                           | 2.185       |
| K1            | 1.862  | 0.935 | 2.317   | 1  | 0.028              | 0.757   | 0.214                           | 2.166       |
| K2            | -0.051 | 0.236 | 0.965   | 1  | 0.154              | 0.006   | —                               | —           |
| K3            | -1.239 | 7.362 | 1.322   | 1  | 0.012              | 1.667   | 0.971                           | 3.492       |
| Constant term | 3.653  | 0.368 | 197.532 | 1  | 0.000              | 0.042   | —                               | —           |

Note. B represents the logistic regression coefficient; Exp(B) is the odds ratio, which reflects the influence of each evaluation factor on the risk; SE is the standard error, which indicates the average error of the estimated value; Wald is a statistic used to test whether the independent variable has an impact on the dependent variable; df is the degree of freedom; “—” means no data.

represent rainfall intensity and rainfall frequency, respectively. To facilitate the assessment of hazard factors, accumulated rainfall in 24 h at 95th, 90th, 80th, 70th, and 60th percentiles is used as the critical disaster-causing rainfall for rainfall classification, respectively. The specific grading standards are as follows. The rainfall in the digits of 60%–70% is level 1. The rainfall in the digits of 70%–80% is level 2. The rainfall in the digits of 80%–90% is level 3. The rainfall in the digits of 90%–95% is level 4. The rainfall in the digits above 95% is level 5. The higher the classification level, the greater the effect of inducing flood formation. Therefore, the weights for levels 1 to 5 are set to 1/15, 2/15, 3/15, 4/15, and 5/15. According to the criterion of critical rainfall, the number of rainstorm intensity occurrences in each city within 15 days is counted. In addition, the sum of products of weights of precipitation intensity and frequency of different grades after normalization is calculated. The hazard factors that characterize each city is assigned to the GIS as the attribute value for rasterization. With the GIS built in the natural segment point classification method, the hazard factors are divided into lowest risk area, lower risk area, moderate risk area, higher risk area, and highest risk area.

### 3.2. Sensitivity Analysis of Hazard-Pregnant Environment.

Based on the analysis of the formation mechanism of rainstorm and flood, the hazard-pregnant environment mainly considers factors such as relief amplitude, vegetation coverage, and density of the river network. The greater the difference in terrain undulation is, the less likely it is to cause flood disasters. The greater the vegetation coverage, the greater the water conservation capacity and thus the lower the probability of flooding. The higher the density of the river network, the closer it is to the water source and the higher the risk of flooding is. Based on the research experience, the relief amplitude level 1 to 5 standards are set as 1.5, 1.2, 0.85, 0.5, and 0.2. The vegetation coverage level 1 to 5 standards are set as 0.8, 0.7, 0.6, 0.45, and 0.2. The density of river network level 1 to 5 standards is set as 0.055, 0.04, 0.03, 0.02, and 0.01. After standardizing these factors, the corresponding weight index is calculated according to the degree of influence of each factor on the rainstorm and flood. Based on GIS, the hazard-pregnancy environment sensitivity can be divided into lowest sensitivity area, lower sensitivity area, moderate sensitivity area, higher sensitivity area, and highest sensitivity area.

### 3.3. Vulnerability Analysis of Hazard-Bearing Body and Analysis of Disaster Resilience.

The degree of risk caused by rainstorm and flood disaster is related to the hazard-bearing body that bears the rainstorm and flood disaster. The greater the population density in the area, the greater the number of people affected by the disaster. To a certain extent, the regional GDP of per person reflects the individual's bearing strength after the disaster. Garden green space has the function of blocking and weakening rainstorm and flood; the larger the proportion of garden green

space area, the greater the ability to resist flood. The large proportion of house area will increase the possibility of rainstorm and flood disaster. From the results of logistic correlation regression analysis, the corresponding significance value of S2 is greater than 0.05, which indicates the level of S2 was not significant. Thus, this article selects S1, S3, and S4 as the evaluation indices of vulnerability. Because the relative bearing capacity of each region in the province to rainstorm and flood disaster is different, the weight of each region should be considered when calculating the vulnerability of the hazard-bearing body. Therefore, the vulnerability index of the hazard-bearing body will be obtained according to the weighted comprehensive evaluation method, and the vulnerability of the hazard-bearing body is divided into lowest vulnerability area, lower vulnerability area, moderate vulnerability area, higher vulnerability area, and highest vulnerability area by using GIS.

For all regions, the ability to withstand disasters can be measured from K1 and K3 during rainstorm. Due to the gap in economic development and social infrastructure among regions, the weighted comprehensive evaluation method is still used to calculate the disaster resilience index, and the disaster resilience capacity is divided into lowest disaster resilience area, lower disaster resilience area, moderate disaster resilience area, higher disaster resilience area, and highest disaster resilience area.

### 3.4. Compartment Analysis of Disaster Factors.

Using the weighted comprehensive evaluation method and GIS-related technologies, combined with the data standardization formula (1), and the weighted comprehensive evaluation method formula (2), calculate the weight of relevant factors and establish a risk evaluation set for the four factors related to rainstorm and flood disaster compartment analysis, as shown in Table 4. The compartment analysis of rainstorm disaster-causing risk of hazard factors, sensitivity of hazard-pregnant environment, vulnerability of hazard-bearing body, and disaster resilience in various regions of Henan Province is shown:

$$Y_{ij} = \frac{X_{ij} - X_{\min_i}}{X_{\max_i} - X_{\min_i}}, \quad (1)$$

where  $Y_{ij}$  represents the standardized value of the  $i$ th index in the  $j$ th area,  $X_{ij}$  represents the  $i$ th index value in the  $j$ th area, and  $X_{\max_i}$  and  $X_{\min_i}$  represent the minimum and maximum values of the  $i$ th index:

$$V = \sum_{i=1}^n W_i Y_i, \quad (2)$$

where  $V$  represents the value of the evaluation factor,  $W_i$  represents the weight of index  $i$ ,  $Y_i$  is the standardized value of the  $i$ th index, and  $n$  represents the number of evaluation indexes.

It can be seen from the compartment analysis map of disaster risk elements of rainstorm and flood in Henan Province (Figure 2); the risk has certain regional differences:

TABLE 4: Risk assessment criteria for rainstorm and flood disaster.

| Disaster risk elements (weight)    | Indices | Weight | Risk assessment criteria |                      |                        |                      |                       |
|------------------------------------|---------|--------|--------------------------|----------------------|------------------------|----------------------|-----------------------|
|                                    |         |        | First level (lowest)     | Second level (lower) | Third level (moderate) | Forth level (higher) | Fifth level (highest) |
| Hazard factors (0.48)              | J1      | 0.16   | 150                      | 200                  | 250                    | 350                  | 450                   |
|                                    | J2      | 0.24   | 40                       | 45                   | 50                     | 55                   | 60                    |
|                                    | J3      | 0.18   | 30                       | 50                   | 70                     | 100                  | 150                   |
|                                    | J4      | 0.42   | 50                       | 100                  | 150                    | 200                  | 250                   |
| Hazard-pregnant environment (0.23) | Z2      | 0.51   | 1.50                     | 1.20                 | 0.85                   | 0.50                 | 0.20                  |
|                                    | Z3      | 0.35   | 0.80                     | 0.70                 | 0.60                   | 0.45                 | 0.20                  |
|                                    | Z4      | 0.14   | 0.055                    | 0.041                | 0.03                   | 0.02                 | 0.01                  |
| Hazard-bearing body (0.17)         | S1      | 0.39   | 0.5                      | 1.0                  | 2.0                    | 3.0                  | 4.0                   |
|                                    | S3      | 0.36   | 0.75                     | 0.65                 | 0.55                   | 0.45                 | 0.35                  |
|                                    | S4      | 0.25   | 0.4                      | 0.45                 | 0.5                    | 0.55                 | 0.60                  |
| Disaster resilience (0.12)         | K1      | 0.78   | 5.5                      | 4.5                  | 3.5                    | 2.5                  | 1.5                   |
|                                    | K3      | 0.22   | 4.0                      | 3.0                  | 2.5                    | 2.0                  | 1.0                   |

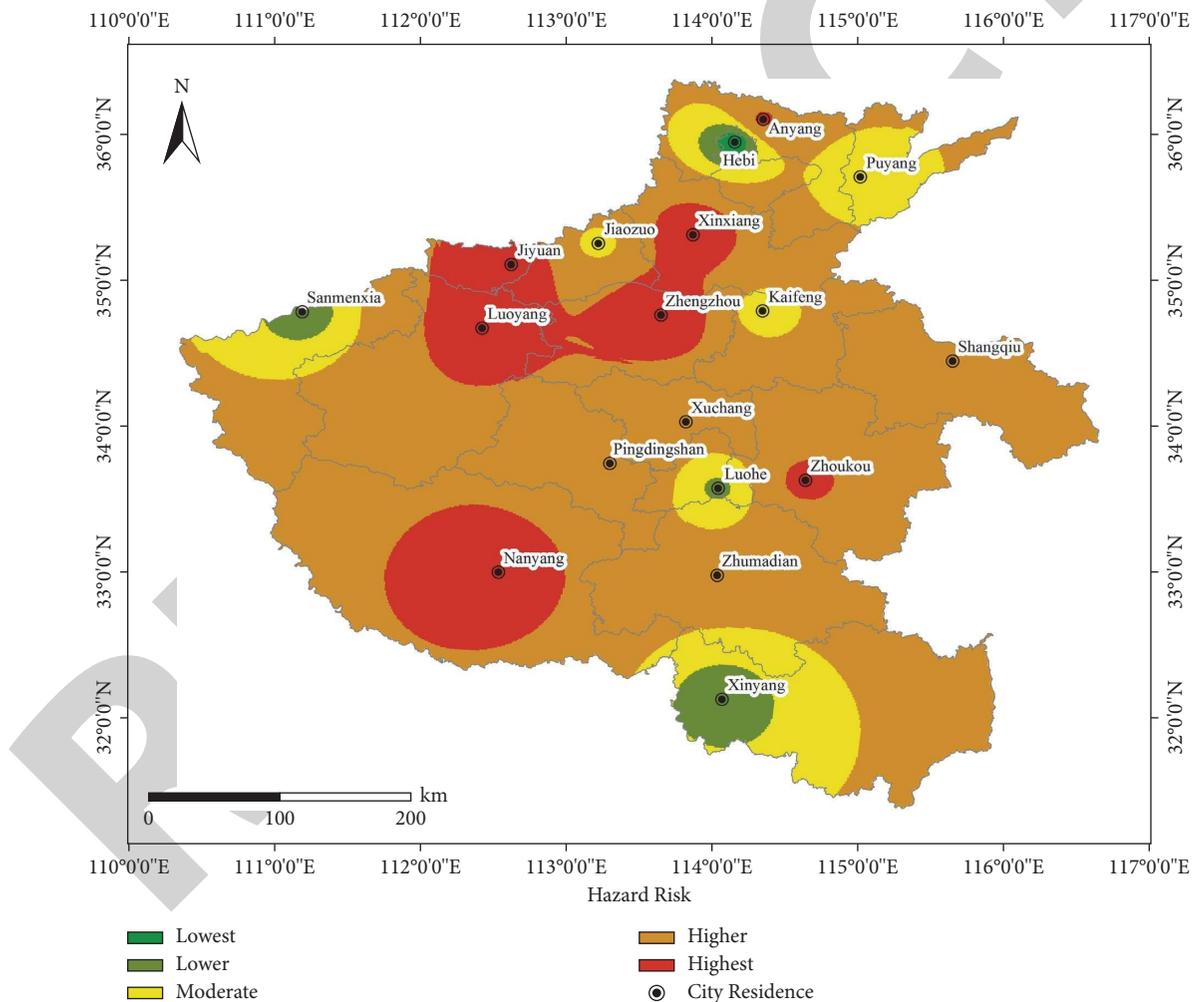


FIGURE 2: Assessment map of hazard risk of rainstorm and flood disaster in Henan Province.

generally, most areas of Henan are prone to rainstorm and flood disaster due to the geographical location of most plains and the influence of the Yellow River and Huaihe River basins. Among them, Zhengzhou, Luoyang, Jiyuan, and Xinxiang in the northwest of Henan Province, Nanyang in

the southwest, and Zhoukou in the Middle East are highest risk areas of rainstorm and flood disaster. Anyang, Shangqiu, Xuchang, Pingdingshan, and Zhumadian are higher risk areas of rainstorm and flood disaster. Puyang, Jiaozuo, and Kaifeng are moderate risk areas of rainstorm and flood

disaster. Sanmenxia and Luohe are lower risk areas of rainstorm and flood disaster. Xinyang is the most southern city in Henan Province, located in the upper reaches of the Huaihe River, bordering Hubei Province and Anhui Province. Xinyang is located in the transition area from subtropical zone to warm temperate zone, with obvious seasonal climate and abundant rainfall, but Xinyang is high in the South and low in the north; it is a ladder landform with alternating hills and rivers and diverse forms, and its forest vegetation is dense, which is not easy to cause rainstorm and flood disasters. Thus, it shows lower risk of rainstorm and flood disaster. Hebi has a warm temperate climate with dry climate and little rain, and its geographical location is high in the West and low in the East. The climate and geographical conditions make the risk of rainstorm and flood disaster in this area the lowest.

As can be seen from the compartment analysis map of sensitivity of rainstorm and flood disaster in Henan Province (Figure 3), Zhengzhou, as the intersection of hilly land and plain flood, is a part of the plain in North China, and the terrain is relatively flat. Moreover, the Yellow River and the Huaihe River pass through the region. Among them, the Yellow River system has many river networks such as the main stream of the Yellow River, Yiluo River, and Sishui River. Therefore, Zhengzhou has become a highest sensitivity environment prone to rainstorm and flood disaster in Henan Province. Luoyang, Xinxiang, Anyang, and Nanyang in the southwest are higher sensitivity areas of rainstorm and flood disaster. Puyang, Kaifeng, Xuchang, Pingdingshan, Zhumadian, and Xinyang are moderate sensitivity areas of rainstorm and flood disaster. Hebi, Jiaozuo, Jiyuan, Shangqiu, Zhoukou, and Luohe are lower sensitivity areas of rainstorm and flood disaster. Sanmenxia is the lowest risk area of rainstorm and flood disaster. The landforms of Sanmenxia are mainly mountains, hills, and plateau, and its altitude is between 300 and 1500 meters. The unique geomorphic characteristics make Sanmenxia the lower sensitivity area of rainstorm and flood disaster.

It can be seen from the compartment analysis map of the vulnerability of the rainstorm and flood disaster in Henan Province (Figure 4) that Zhengzhou, Luoyang, and Nanyang are the highest vulnerability cities with high population density and large proportion of garden green space area. Xinxiang, Jiaozuo, Xuchang, and Zhoukou are higher vulnerability areas of rainstorm and flood disaster. Anyang, Kaifeng, Shangqiu, Zhumadian, and Shangqiu are moderate vulnerability areas of rainstorm and flood disaster. Puyang, Sanmenxia, and Pingdingshan are lower vulnerability areas of rainstorm and flood disaster. Due to the influence of the density of population, the proportion of garden green space area and the proportion of house area, the vulnerability of rainstorm, and flood disaster in Hebi, Jiyuan, and Luohe are lowest.

In addition, it can be seen from the compartment analysis map of the disaster resilience of the rainstorm and flood disaster in Henan Province (Figure 5) that cities with larger GDP in regions such as Zhengzhou and Luoyang have highest disaster resilience. Nanyang, Xinxiang, Xuchang, and Zhoukou are higher disaster resilience areas of

rainstorm and flood disaster. Jiaozuo and Xinyang are moderate disaster resilience areas of rainstorm and flood disaster. Anyang, Kaifeng, Shangqiu, Pingdingshan, and Zhumadian are lower disaster resilience areas of rainstorm and flood disaster. Due to the limitations of economic and backward infrastructure construction, Hebi, Puyang, Jiyuan, and Luohe are lowest disaster resilience areas of rainstorm and flood disaster. At the same time, Sanmenxia, a mountainous area with higher altitudes, the northeastern region where economic development is relatively backward, has the lowest resilience.

#### 4. Unequal Weight Clustering Hybrid PSO-SVR Algorithm

The unequal weight clustering hybrid PSO-SVR algorithm is an integrated machine learning algorithm. The main idea is to perform stepwise regression and dimensionality reduction processing on the data first and then perform clustering after unequal weight processing. Small sample data can better reflect the superiority of the PSO-SVR algorithm.

**4.1. SVR Algorithm Principle.** SVM refers to a common discriminated method. In adherence to the SRM principle, it shows unique advantages in handling small samples and high-dimensional feature space problems. SVM is first used to solve the model recognition problems, but recently, it has also been applied to address nonlinear regression estimation problems through introducing the insensitive loss function  $\varepsilon$ . When being used to tackle the regression problems, SVM is referred to as support vector regression (SVR), and the main thinking of SVR is to map the dataset  $x_i (i = 1, \dots, n)$  to a high-dimensional feature space through nonlinear function. The specific relation involved can be expressed as

$$f(x) = \omega^T \phi(x) + b, \quad (3)$$

where  $f(x)$  is the output value,  $\omega$  and  $b$  are the coefficients, and  $\phi(x)$  is the nonlinear mapping function which can convert the input value to the high-dimensional feature space. The regulatory value of  $\omega$  and  $b$  is indicated by

$$\begin{aligned} \text{Min}_{\omega, b, \xi_i^*, \xi} R_\varepsilon(\omega, \xi_i^*, \xi) &= \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n (\xi_i^* + \xi), \\ \begin{cases} y_i - \omega^T \phi(x_i) - b \leq \varepsilon + \xi_i^*, & i = 1, 2, 3, \dots, n, \\ -y_i + \omega^T \phi(x_i) - b \leq \varepsilon + \xi_i, & i = 1, 2, 3, \dots, n, \\ \xi_i^*, \xi_i \geq 0, & i = 1, 2, 3, \dots, n, \end{cases} \end{aligned} \quad (4)$$

where  $R_\varepsilon()$  is the empirical risk [64, 65],  $C$  is the regularization parameter,  $\xi_i^*$  is the error greater than  $\varepsilon$ , and  $\xi_i$  is the error less than  $-\varepsilon$ . The above function denotes a quadratic optimization problem which can be converted to the dual problem. Given this, the final equation of SVR is:

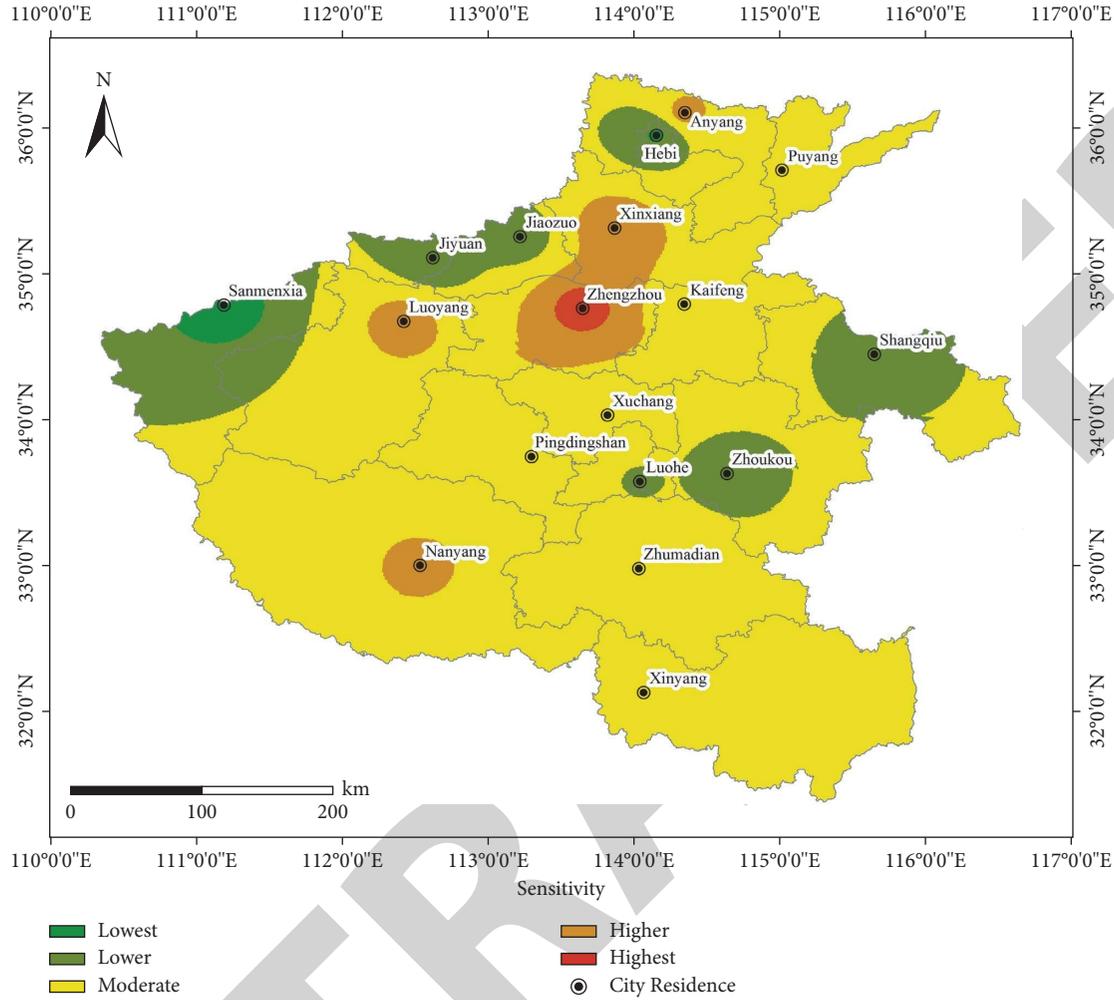


FIGURE 3: Assessment map of sensitivity of rainstorm and flood disaster in Henan Province.

$$f(x) = \sum_{i=1}^n (\beta_i^* - \beta_i)K(x_i, x_j) + b, \quad (5)$$

where  $\beta_i^*$  and  $\beta_i$  are the Lagrangian coefficients, and  $K(x_i, x_j)$  is the kernel function of SVR which stands for the inner product of two vectors. The kernel function of vectors  $x_i$  and  $x_j$  can be defined as

$$K(x_i, x_j) = \phi(x_i)\phi(x_j). \quad (6)$$

There are several types of kernel function in the existing research, including both linear kernel function and Gaussian kernel function. Gaussian kernel function, as one of the most commonly used kernel functions, is also referred to as the radial basis function (RBF). This function is able to map the data to infinite dimensions and has relatively lower computational complexity. Therefore, the research uses RBF as the kernel function of SVR, and the function can be defined as

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\gamma^2}\right), \quad (7)$$

where  $\gamma$  is the Gaussian parameter. SVR parameter combination is the key to realize high-precision prediction. Accordingly, the paper adopts the PSO algorithm to determine parameters including  $\gamma$ .

**4.2. PSO Algorithm Principle.** PSO represents a population computing technology developed on the basis of iteration optimization. Its first step is to initialize a group of particles. Then, the rate and location of these particles in the following iteration can be updated by tracking two extremums (single extremum  $P_{ibest}$  and global extremum  $P_{gbest}$ ). When these two extremums are found, the PSO algorithm will be taken to recognize the rate and distance of each particle.

Suppose there are  $m$  particles in the  $d$ -dimensional search space. The  $i_{th}$  particle is indicated by  $x_i = (x_{i1}, \dots, x_{id})$ , where  $i = 1, \dots, m$ . In another word, the position of the  $i_{th}$  particle is  $x_i$ . The rate of the  $i_{th}$  particle is also a vector, expressed as  $v_i = (v_{i1}, \dots, v_{id})$ . The optimal location of the particle is  $p_i = (p_{i1}, \dots, p_{id})$ , and the optimal location of the whole population is  $p_g = (p_{g1}, \dots, p_{gd})$ . The standard PSO algorithm updates existing PSO algorithm, which is defined as

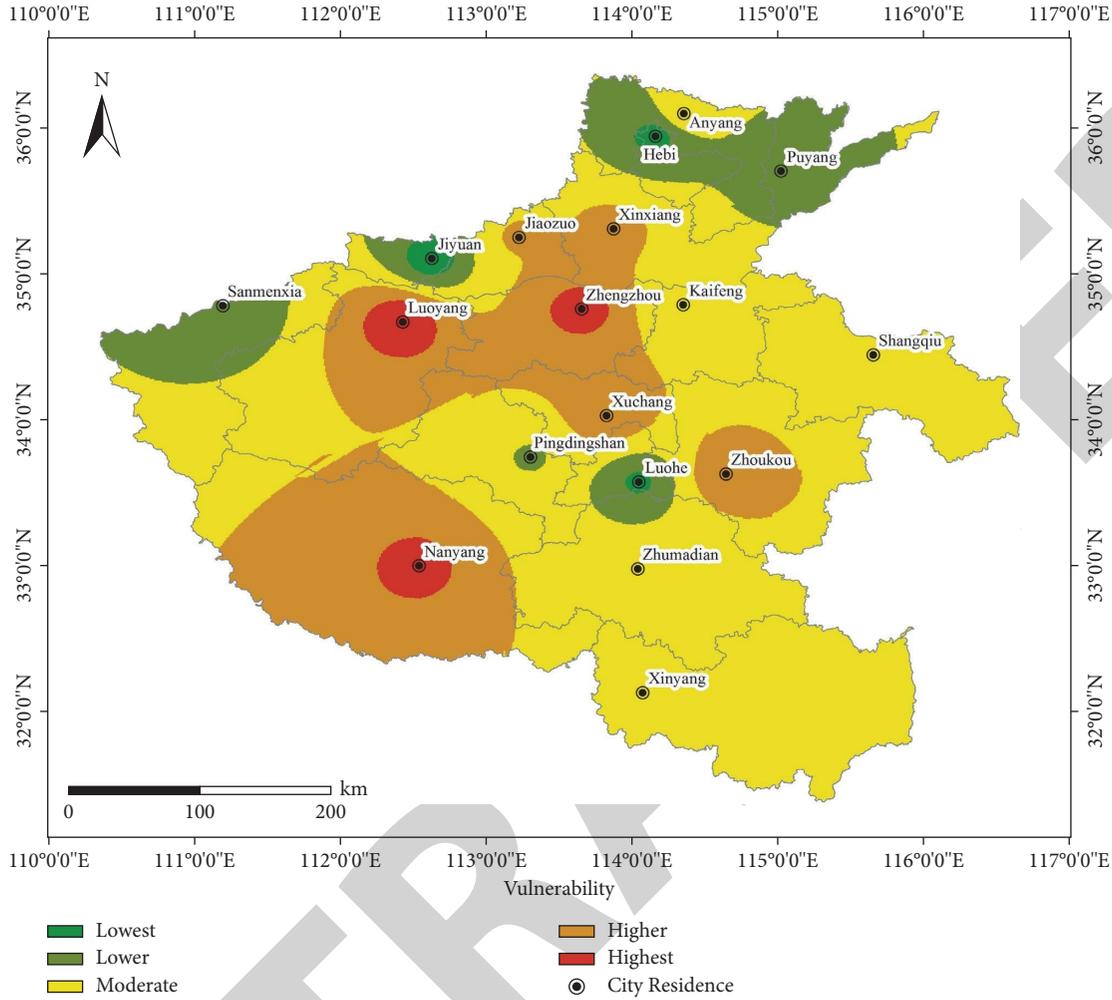


FIGURE 4: Assessment map of vulnerability of rainstorm and flood disaster in Henan Province.

$$\begin{aligned}
 v_{i,k+1}^d &= \bar{\omega}v_{i,k}^d + c_1r_1(P_{i,k}^d - x_{i,k}^d) + c_2r_2(P_{g,k}^d - x_{i,k}^d), x_{i,k+1}^d \\
 &= x_{i,k}^d + v_{i,k+1}^d.
 \end{aligned} \tag{8}$$

**4.3. Unequal Weight Clustering Hybrid PSO-SVR Model.** The model is mainly based on the PSO-SVR algorithm and mainly includes the following steps.

**4.3.1. Generating Data Matrix.** Before performing stepwise regression and dimensionality reduction, it is assumed that there exist  $c$  sample data and  $n$  independent variables (variable data after feature processing) in the experimental data. The set can be expressed as  $X = (X_1, \dots, X_n)$ , the dependent variable in this study is denoted by  $Y$ , and the model is written as follows:

$$Y = \beta_0 + \beta_i X_i + \varepsilon, \quad i = 1, \dots, n. \tag{9}$$

By calculating the regression coefficient of  $X_i$ , the  $F$ -test statistic values of the corresponding coefficients are  $(F'_1, \dots, F'_n)$  with  $F'_\tau$  being the maximum value. Under a

given significance level,  $\alpha = 0.05$ , the corresponding critical value is  $F^{(1)}$ . When  $F'_\tau \geq F^{(1)}$ ,  $X_\tau$  is added to the regression model and  $S_1$  is represented by the selected variable index set.

**4.3.2. Establish a Binary Regression Model.** Establish a binary regression model of dependent variable  $Y$  and a subset of independent variables  $\{X_{\tau_1}, X_1\}, \dots, \{X_{\tau_1}, X_n\}$ . There are a total of  $n - 1$  subsets. The regression coefficient and the corresponding  $F$ -test statistic are calculated as  $F''_k$  ( $k \notin S_1$ ).  $F''_\tau$  refers to the maximum value. With a given significance level  $\alpha = 0.05$ , the corresponding critical value is  $F^{(2)}$ . When  $F''_\tau \geq F^{(2)}$ , then  $X_{\tau_2}$  is added to the regression model. Otherwise, the variable introduction process is ended.

**4.3.3. Repeat the Above Operation.** Repeat the operation of Step 2 to obtain the final required equation model of this research as follows:

$$Y_\tau = \sum_{i=1}^k \beta_i X_{\tau_i}, \quad k \leq n. \tag{10}$$

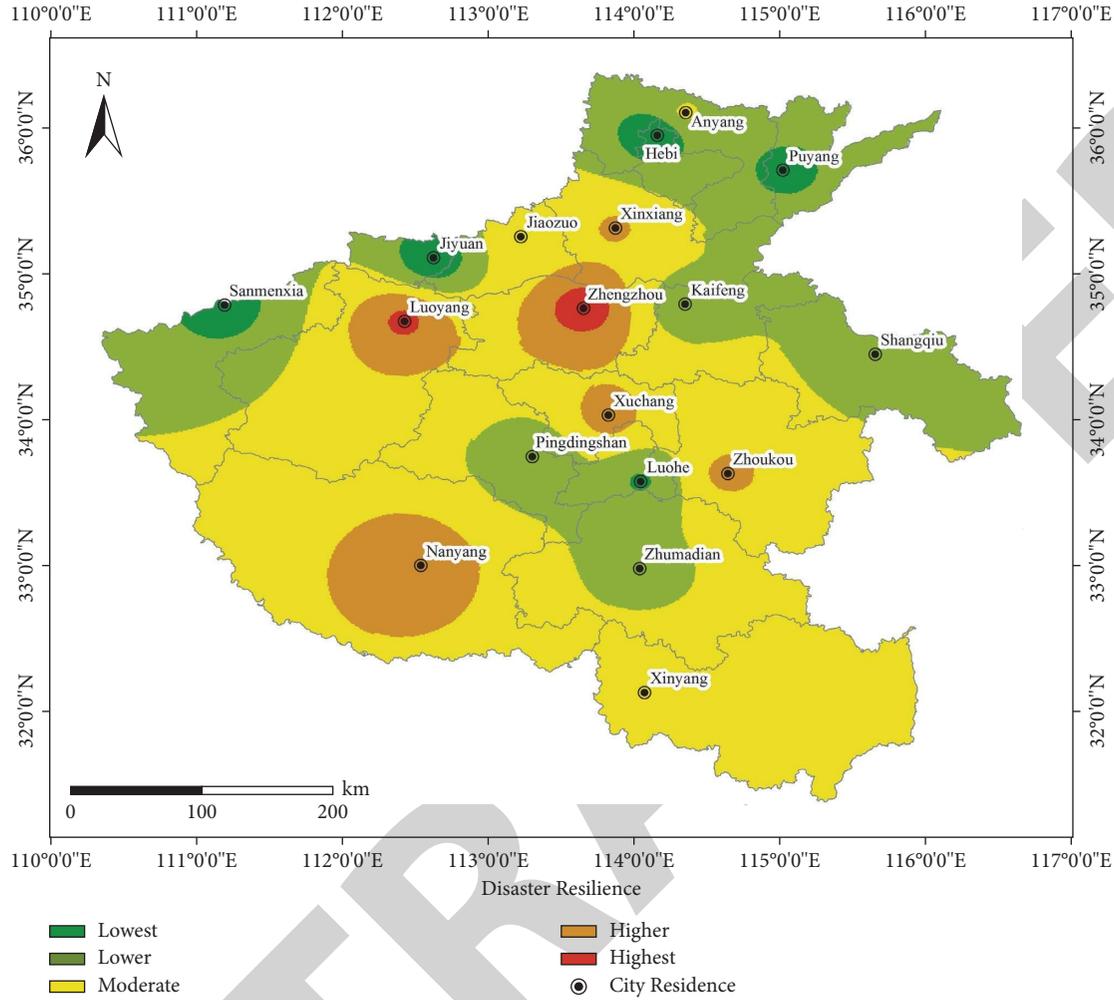


FIGURE 5: Assessment map of disaster resilience of rainstorm and flood disaster in Henan Province.

4.3.4. *Select Centroids.* Select  $l$  centroids, multiply the feature-encoded original data by the corresponding coefficient  $\beta_i$ , and input the new dataset  $Y_\tau$  into the  $K$ -means clustering algorithm to obtain  $l$  datasets. These are  $u_1, \dots, u_l \in R^n$ ,  $Y_{\tau_i} \in R^n, i = 1, \dots, c$ . The Euclidean distance from each sample in the dataset  $Y_\tau$  to the centroid  $u_j$  is calculated, the centroid of the collection is continuously updated, and it was classified into  $l$  collections. The specific calculation formula is written as follows:

$$S_i = \arg \min \left\| Y_{\tau_i} - \frac{\sum_{i=1}^c Y_{\tau_i} |S_i = j}{\sum_{i=1}^c 1 |S_i = j} \right\| \quad S_i = j = 1, \dots, l. \quad (11)$$

4.3.5. *Model Test.* Take 10% of each set  $S_i$  as the test set, and finally, input it into the PSO-SVR model to obtain  $S(S_i, P(c, \gamma, \epsilon))$  as the predicted result.

## 5. Result Analysis and Discussion

5.1. *Evaluation Methods.* The logistic binary regression built in SPSS 26.0 is used to preliminarily screen the relevant

factors affecting the occurrence of rainstorm and flood disaster, and then, the compartment analysis map of the four factors clearly shows the characteristics of each region. To further predict the rainstorm disasters, this research selects the accumulated rainfall in 24 h (daily rainfall), the number of rainstorm and flood disasters (selecting the daily rainfall exceeding 60 mm per month), and direct economic loss (average monthly rainstorm and flood disaster economic loss) in each region of Henan Province from 2010 to 2020 which are used as target variables to participate in the construction of the hybrid PSO-SVR model. The data from 2010 to 2019 are selected as the training samples, and the data from 2020 are used to verify the prediction accuracy of the model. To more objectively analyze the accuracy of the hybrid PSO-SVR model in rainstorm and flood disaster prediction, the SVR model without parameter optimization algorithm and the GA-SVR model and artificial neural network BPNN model are constructed from the same experimental samples for comparative verification.

The SVR model without parameter optimization algorithm is selected for comparison, which is mainly used to highlight the impact of parameter optimization on the prediction results. The GA-SVR model is chosen to compare

and highlight that PSO is more applicable to this model than GA algorithm due to its better optimization parameter effects. The BPNN model in artificial neural network (ANN) is chosen mainly because the model can still guarantee the sound prediction effects through establishing a relatively stable generalized regression neural network (GRNN) via radial basis neurons and linear neurons even when the model has limited experimental data. The research chooses the root mean square error (RMSE) and mean absolute percentage error (MAPE) to test the proposed hybrid model:

$$\text{RMSE} = \frac{1}{n} \sqrt{\sum_{i=1}^n (x_i^* - x_i)^2}. \quad (12)$$

RMSE is the mean square root of the square sum of the errors of the corresponding points of the predicted result and the real value. The smaller the value of RMSE is, the better the accuracy of the prediction model is:

$$\text{MAPE} = \sum_{i=1}^n \left| \frac{x_i^* - x_i}{x_i} \right| \times \frac{100}{n}. \quad (13)$$

MAPE is often used as a statistical index to measure the accuracy of prediction. The value of MAPE is smaller, the accuracy of the prediction model is better and the deviation from the real value is smaller, where  $n$  represents the number of time instances and  $x_i^*$  and  $x_i$  represent the predicted results and the real value in formulas (12) and (13).

**5.2. Result Analysis.** According to the compartment analysis of rainstorm and flood disaster, to explain the result, this research selects the data of Zhengzhou City and Luoyang City in Henan Province for result analysis. The performance comparison of each model is shown (see Table 5), and the prediction results of each model are shown in Figure 6. Note: the RMSE and MAPE calculate the average daily rainfall in 15 days.

The hybrid PSO-SVR model is run by Python compiler, and the RMSE and MAPE of each model are calculated (see Table 5). It can be seen that the four models perform well in the prediction of daily rainfall, with RMSE controlled within 24 and MAPE controlled within 31%. Among them, the hybrid PSO-SVR model performs best, with RMSE lower than 10 and MAPE lower than 15%. The effect of the GA-SVR model with parameter optimization is lower than that of the hybrid PSO-SVR model, followed by the BPNN model and the original SVR model, respectively. At the same time, the prediction of hybrid PSO-SVR for the number of disasters and disaster economic loss is still more accurate than other models.

Collecting the real values and the predicted results and using the gray correlation analysis method [66] to calculate the correlation between the two values, the calculated average correlations of the hybrid PSO-SVR, GA-SVR, SVR, and BPNN models to the predicted results and real values of the target variables are 0.696, 0.667, 0.639, and 0.625; the average correlations are greater than 0.625, respectively. It can be seen that the predicted results of the hybrid PSO-SVR is closest to the real value, which shows that the model has

better learning ability and generalization ability for the data of rainstorm and flood disaster.

At the same time, it can be seen from the predicted trend graph in Figure 6 that, for daily rainfall (Figures 6(a) and 6(b)), number of disasters (Figures 6(c) and 6(d)), and direct economic loss (Figures 6(e) and 6(f)), the prediction effect of hybrid PSO-SVR is better than other algorithm models; especially, at peak values of rainfall, the fitting effect is the best of all. It is further verified that the model proposed in this research has a strong learning ability to a certain extent on the complicated problem of rainstorm and flood disaster, and it also shows the better generalization ability of SVR to deal with small samples of high-dimensional data.

To explain and verify the practicability of the hybrid model proposed in this study, the hybrid PSO-SVR model is predicted for other 16 cities in Henan Province. The average value of RMSE and MAPE corresponding to the prediction results is shown in Figure 7.

The average RMSE of 16 cities is less than 18 (Figure 7(a)), and the average MAPE of 16 cities is less than 21% (Figure 7(b)). The experimental results show that the hybrid PSO-SVR model is better than the other three models. When there are many characteristic dimensions of experimental data, it will affect the accuracy of prediction and increase the computational complexity. Firstly, regression analysis is selected to reduce the dimension of rainstorm and flood disaster index system in the research, so as to obtain a better combination of variables and reduce the complexity of data processing and the time required for prediction; By using the PSO optimization algorithm, this research achieves the automatic selection of parameters and overcomes the premature convergence problem of SVR. When the experimental data are complex, the processing efficiency and performance of GA algorithm are not as good as PSO algorithm. The comparison of RMSE and MAPE shows that the PSO algorithm is more suitable to optimize the parameters of SVR than GA algorithm. Due to the average daily rainfall, the number of disasters and disaster economic loss belong to small sample data; the performance of SVR prediction is better than BP neural network. Therefore, it is worthwhile to use the hybrid PSO-SVR model to predict the data of rainstorm and flood disaster.

Through the evaluation and analysis of rainstorm and flood disasters in Henan Province, a reasonable and effective evaluation mechanism is found (Figure 8). To a certain extent, we hope that the mechanism can help the government in disaster prevention and reduction. We will continue to conduct in-depth research and explore and establish a better evaluation mechanism in the future. Combined with historical disaster information, relevant analysis, and disaster risk zoning, this research puts forward the following suggestions:

- (1) For cities with backward economic development and low population density, we should increase the construction of infrastructure, so as to improve the resistance and response ability to rainstorm and flood disasters.

TABLE 5: Model performance comparison.

| Target variable        | PSO-SVR (RMSE/MAPE) |                 | GA-SVR (RMSE/MAPE) |                 | SVR (RMSE/MAPE) |                 | BPNN (RMSE/MAPE) |                 |
|------------------------|---------------------|-----------------|--------------------|-----------------|-----------------|-----------------|------------------|-----------------|
|                        | Zhengzhou           | Luoyang         | Zhengzhou          | Luoyang         | Zhengzhou       | Luoyang         | Zhengzhou        | Luoyang         |
| Average daily rainfall | 7.25<br>11.36%      | 8.16<br>14.95%  | 9.18<br>17.78%     | 11.58<br>18.82% | 18.96<br>30.25% | 23.75<br>30.80% | 14.38<br>20.18%  | 21.73<br>27.68% |
| Number of disasters    | 8.57<br>14.58%      | 7.62<br>13.11%  | 12.19<br>15.77%    | 16.15<br>20.39% | 55.23<br>38.33% | 61.95<br>43.81% | 41.20<br>36.59%  | 26.35<br>33.76% |
| Disaster economic loss | 9.61<br>18.65%      | 10.36<br>16.74% | 13.31<br>20.36%    | 15.77<br>23.18% | 35.61<br>41.93% | 52.73<br>46.52% | 30.84<br>29.71%  | 29.97<br>32.63% |

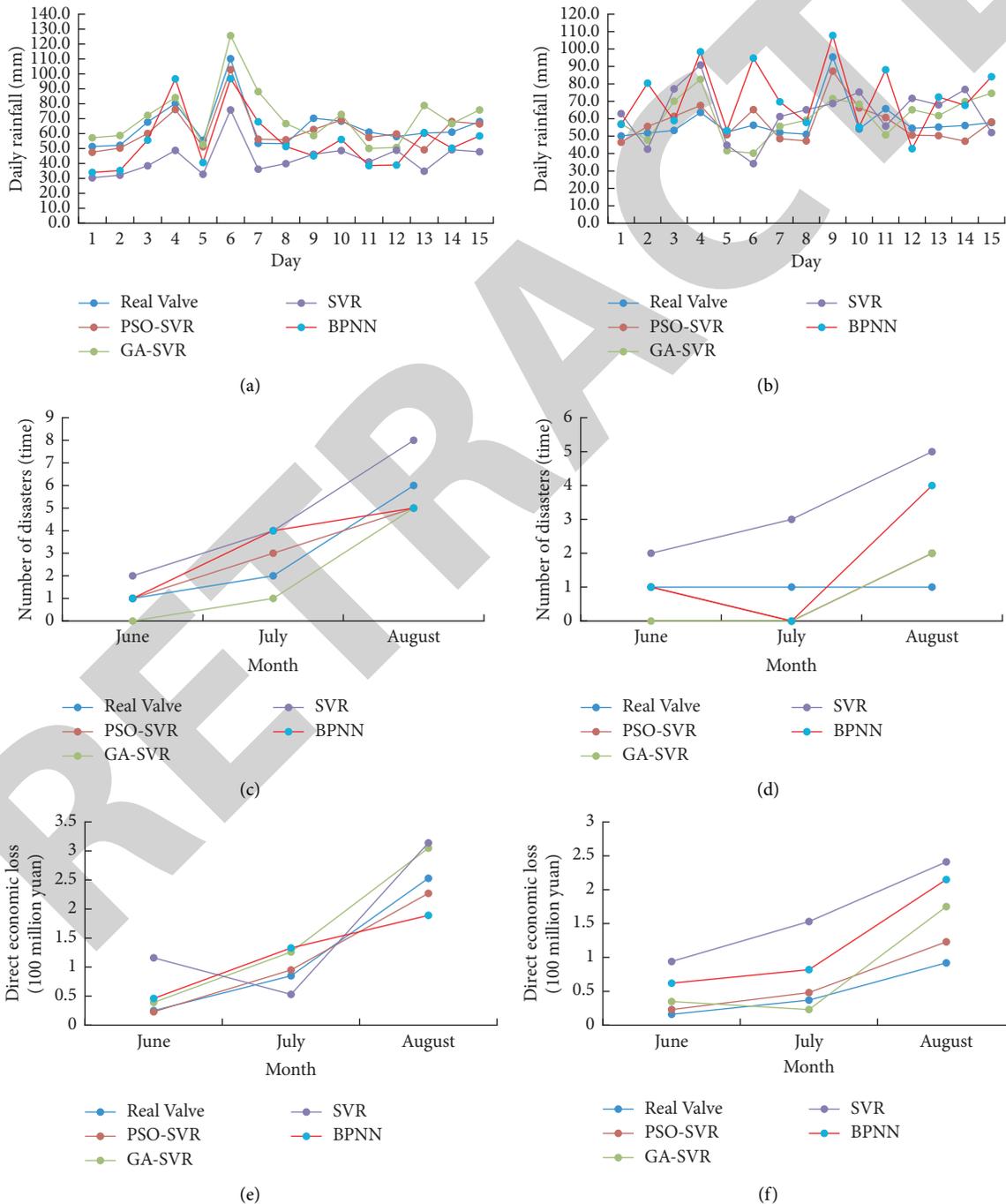


FIGURE 6: Assessment comparison of prediction results of various indicators in Zhengzhou City and Luoyang City in Henan Province: (a) and (b) rainstorm forecast; (c) and (d) number of disasters; (e) and (f) direct economic loss.

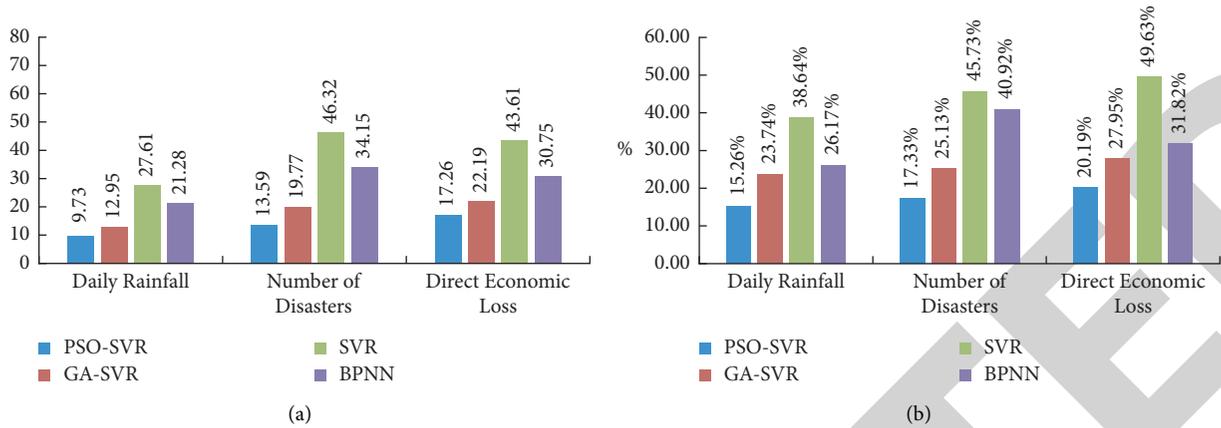


FIGURE 7: Index average values of 16 cities in Henan Province under different models: (a) RMSE. (b) MAPE.

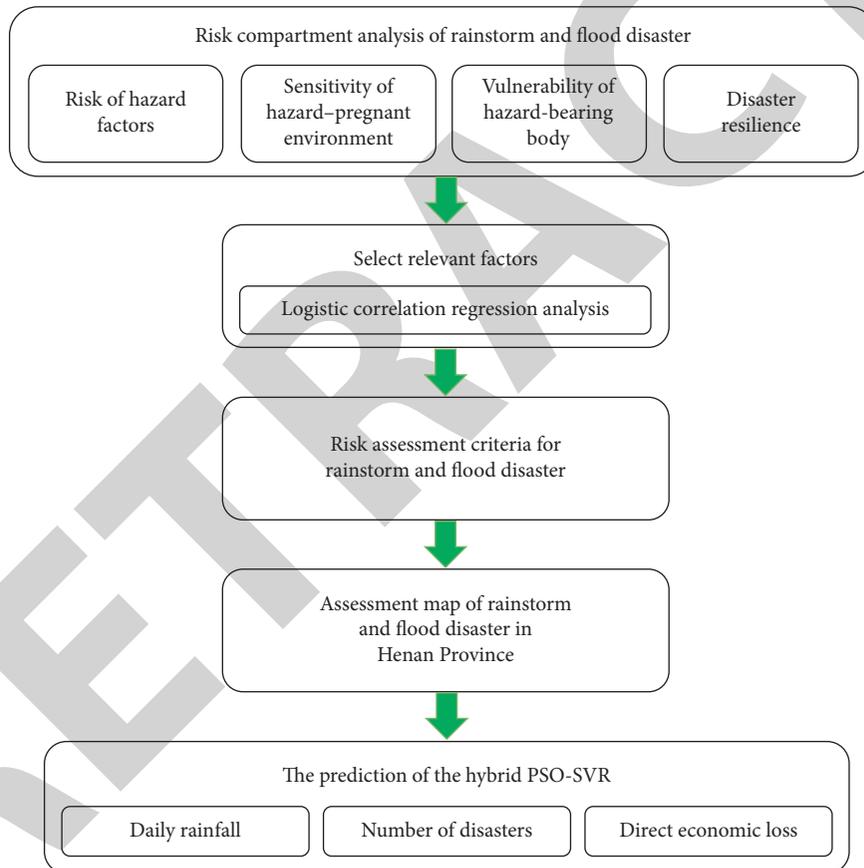


FIGURE 8: Risk evaluation mechanism of rainstorm and flood disaster.

(2) For cities with developed economy and concentrated population density, we should strengthen the cultivation of residents' awareness of disaster resistance, so as to improve residents' ability to face rainstorm and flood disasters. At the same time, the government should increase the area of garden green space, expand the area of urban vegetation, and continue to strengthen the construction of drainage capacity, so as to well resist the invasion of rainstorm and flood disasters.

With the rapid development of global climate change and urbanization, more and more cities are suffering from extreme rainstorm and flood disasters, which has caused huge losses to people's lives and social and economic construction. Therefore, it is very important to carry out risk assessment and prediction of rainstorm and flood disaster, which will help to improve the ability of regional emergency prediction, reduce losses caused by rainstorm and flood disasters. Then, to expand the applicable scope of the model, the experiments of the hybrid PSO-SVR model will be tested

in more cities, and better improvements will be made in the continuous experimental process, to strive to provide more accurate analysis and assessment and disaster loss prediction.

## 6. Conclusions

Based on the hybrid PSO-SVR machine learning algorithm, this article constructs a rainstorm and flood disaster assessment and prediction model in Henan Province. First, based on the existing relevant research, considering the availability of data, establish a reasonable index system and initially screen the relevant factors that affect the occurrence of rainstorms and floods through logistic binary regression; Then, using GIS technology and a weighted comprehensive evaluation method to analyze the various regions in Henan Province, and the hazard factors, hazard-pregnant environment, hazard-bearing body, and disaster resilience were analyzed. Research and analysis showed that the risk of hazard factors in most parts of Henan Province from June to August was high. There are many rivers distributed, and most areas are lower plains, which are prone to rainstorm and flood disaster. For economically developed areas, due to the influence of geographical location, the sensitivity of the hazard-pregnant environment is high; at the same time, the vulnerability and resilience are also high. Through the regional analysis of hazard factors, the compartment analysis characteristics of rainstorm and flood in various regions of Henan Province are clearly demonstrated. Finally, to solve the complex disaster loss prediction situation involving nonlinear multidimensional factors, a hybrid PSO-SVR rainstorm and flood disaster model was constructed. The research results show that the hybrid PSO-SVR rainstorm and flood disaster prediction model is better than the GA-SVR, SVR, and BPNN models. The four models involved in the experiment have a better prediction of daily rainfall than the number of disasters and direct economic loss. Hybrid PSO-SVR has the best fitting effect for high peaks values in rainstorm and flood disaster prediction involving complex multidimensional factors.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This paper was greatly supported by National Social Science Fund Key Project (15AGL013), Henan Provincial Department of Science and Technology Risk Management Innovation and Public Policy Soft Science Research Base, Henan Social Science Planning Project (2019BJJ030), Research on the Construction of Disaster Prevention and Reduction Support System in Large and Medium Sized Cities in Henan

Province (22240041001), and Henan Provincial Colleges and Universities Philosophy and Social Science Basic Research Major Project “Evaluation Research on Comprehensive Disaster Resilience Capacity of Chinese Communities” (2021JCZD04).

## References

- [1] G. F. Fan, Y. Zhang, Y. He, and K. Wang, “Risk assessment of drought in the Yangtze River Delta based on natural disaster risk theory,” *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 5682180, 7 pages, 2017.
- [2] J. Buckland and M. Rahman, “Community-based disaster management during the 1997 red river flood in Canada,” *Disasters*, vol. 23, no. 2, pp. 174–191, 2010.
- [3] K. M. De Bruijn, “Resilience and flood risk management,” *Water Policy*, vol. 6, no. 1, pp. 53–66, 2004.
- [4] L. M. Zhao, K. Wang, and P. H. Qiu, “Synthetic evaluation of disaster,” *Systems Engineering Theory & Practice*, vol. 17, no. 3, pp. 63–69, 1997.
- [5] E.-S. Rafiei, A. Azareh, B. Choubin, A. H. Mosavi, and J. J. Clague, “Evaluating urban flood risk using hybrid method of TOPSIS and machine learning,” *International Journal of Disaster Risk Reduction*, vol. 66, Article ID 102614, 2021.
- [6] P. Luo, Y. Mu, S. Wang et al., *Exploring Sustainable Solutions for the Water Environment in Chinese and Southeast Asian Cities*, AMBIO, Netherlands, 2021.
- [7] J. Guo, X. H. Wu, and G. Wei, “A new economic loss assessment system for urban severe rainfall and flooding disasters based on big data fusion,” *Environmental Research*, vol. 188, Article ID 109822, 2020.
- [8] D. R. Mu, P. P. Luo, J. Q. Lyu et al., “Impact of Temporal Rainfall Patterns on Flash Floods in Hue City, Vietnam,” *Journal of Flood Risk Management*, vol. 14, 2020.
- [9] H. Darabi, A. T. Haghighi, M. A. Mohamadi et al., “Urban flood risk mapping using data-driven geospatial techniques for a flood-prone case area in Iran,” *Hydrology Research*, vol. 51, no. 1, pp. 127–142, 2020.
- [10] G. Zhao, B. Pang, Z. Xu, D. Peng, and L. Xu, “Assessment of urban flood susceptibility using semi-supervised machine learning model,” *The Science of the Total Environment*, vol. 659, no. 1, pp. 940–949, 2019.
- [11] S. M. Liu, H. Wang, D. H. Yan, Q. F. Ren, D. Deng, and B. Y. Gong, “Analysis of spatiotemporal evolution of isolated rainstorm events in huai River Basin, China,” *Advances in Meteorology*, vol. 2017, Article ID 3010295, 17 pages, 2017.
- [12] N. Mori, T. Takemi, Y. Tachikawa et al., “Recent nationwide climate change impact assessments of natural hazards in Japan and East Asia,” *Weather and Climate Extremes*, vol. 32, Article ID 100309, 2021.
- [13] M. K. Buchanan, R. E. Kopp, M. Oppenheimer, and C. Tebaldi, “Allowances for evolving coastal flood risk under uncertain local sea-level rise,” *Climatic Change*, vol. 137, no. 3-4, pp. 347–362, 2016.
- [14] B. F. Christopher, B. Vicente, F. S. Thomas, and D. Qin, “Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change,” *Journal of Clinical Endocrinology & Metabolism*, vol. 18, no. 6, pp. 586–599, 2012.
- [15] H. C. Zhou and Z. Dan, “Assessment model of drought and flood disasters with variable fuzzy set theory,” *Transactions of the Chinese Society of Agricultural Engineering*, vol. 25, pp. 56–61, 2009.

- [16] J. L. Chen, G. R. Huang, and W. J. Chen, "Towards better flood risk management: assessing flood risk and investigating the potential mechanism based on machine learning models," *Journal of Environmental Management*, vol. 293, Article ID 112810, 2021.
- [17] Y. Gao, L. I. Bin, Z. Feng, and X. Zuo, "Global climate change and geological disaster response analysis," *Journal of Geomechanics*, vol. 23, no. 1, pp. 65–77, 2017.
- [18] R. S. Quan, "Rainstorm waterlogging risk assessment in central urban area of Shanghai based on multiple scenario simulation," *Scientia Geographica Sinica*, vol. 73, no. 3, pp. 1569–1585, 2011.
- [19] M. Zandvoort and M. Vlist, *The Multi-Layer Safety Approach and Geodesign: Exploring Exposure and Vulnerability to Flooding*, Springer International Publishing, NY, USA, 2014.
- [20] V. Otari, T. Nino, A. Avtandil et al., "Vulnerability, hazards and multiple risk assessment for Georgia," *Natural Hazards*, vol. 64, no. 3, pp. 2021–2056, 2012.
- [21] G. Benito, M. Lang, M. Barriendos, M. C. Llasat, F. Francés, and T. Ouarda, "Use of systematic, palaeo flood and historical data for the improvement of flood risk estimation, review of scientific methods," *Natural Hazards*, vol. 31, no. 3, pp. 623–643, 2004.
- [22] M. I. Alfa, M. A. Ajibike, M. A. Ajibike, and R. E. Daffi, "Application of analytic hierarchy process and geographic information system techniques in flood risk assessment: a Case of Ofu River Catchment in Nigeria," *Journal of Degraded and Mining Lands Management*, vol. 5, no. 4, pp. 1363–1372, 2018.
- [23] K. M. Weerasinghe, H. Gehrels, N. M. S. I. Arambepola, H. P. Vajja, J. M. K. Herath, and K. B. Atapattu, "Qualitative flood risk assessment for the Western Province of Sri Lanka," *Procedia Engineering*, vol. 212, pp. 503–510, 2018.
- [24] M. Mahootchi and S. Golmohammadi, "Developing a new stochastic model considering bi-directional relations in a natural disaster: a possible earthquake in Tehran (the Capital of Islamic Republic of Iran)," *Annals of Operations Research*, vol. 269, no. 1, pp. 439–473, 2018.
- [25] A. Alsubaie, A. D. Pietro, J. Marti et al., *A Platform for Disaster Response Planning with Interdependency Simulation Functionality*, Springer, Berlin, Heidelberg, 2013.
- [26] J. Feng, D. H. Yan, T. L. Qin, X. D. Zheng, and S. M. Geng, "Comprehensive response to flood disaster in inland river area of Northwest China under changing environment," *Water Resources Protection*, vol. 28, no. 4, pp. 17–23, 2012.
- [27] L. Weng, L. Ma, and S. F. Xu, "Study on risk assessment and defense countermeasures of urban rainstorm disaster: a case study of Nanjing City, Jiangsu Province," *Disaster Science*, vol. 30, no. 1, pp. 130–134, 2015.
- [28] F. Kong, "Perspective on urban rainstorm and waterlogging disasters in China under changing environment: situation, causes and policy suggestions," *Water Conservancy and Hydropower Technology*, vol. 50, no. 10, pp. 42–52, 2019.
- [29] Y. H. Zhang and T. Cai, *Study on the Construction of Fuzzy Comprehensive Evaluation System for Urban Rainstorm and Waterlogging Disaster Risk*, pp. 103–107, Water Conservancy Planning and Designing, 2019.
- [30] J. Peng, H. Wei, W. H. Wu, Y. X. Liu, and Y. L. Wang, "Risk assessment of urban rainstorm and flood disaster based on land use change scenario: a case study of Maozhou River Basin in Shenzhen," *Journal of Ecology*, vol. 38, no. 11, pp. 3741–3755, 2018.
- [31] H. Z. Feng and Y. Y. Chen, "A new method for dealing with nonlinear classification and regression problems (II): application of support vector machine method in weather forecasting," *Journal of Applied Meteorology*, vol. 3, pp. 355–365, 2004.
- [32] H. L. Xie, G. H. Peng, M. C. Guo, and H. C. Zheng, "Application of improved SVR model in annual precipitation prediction," *Mathematical Practice and Understanding*, vol. 47, no. 18, pp. 154–161, 2017.
- [33] J. H. Liu, Z. Zhang, and J. M. Wu, "Prediction model of urban flood disaster based on BP neural network," *Computer Engineering and Design*, vol. 26, no. 3, pp. 699–701, 2005.
- [34] X. Z. Cao and Q. Wang, "Application of neural network in meteorological observation data optimization," *Plateau Meteorology*, vol. 21, no. 1, pp. 96–97, 2002.
- [35] N. Liu, A. Y. Xiong, Q. Zhang, Y. J. Liu, Y. J. Zhan, and Y. M. Liu, "Development of basic dataset of severe convective weather for artificial intelligence training," *Journal of Applied Meteorology*, vol. 32, no. 5, pp. 530–541, 2021.
- [36] L. M. Mao, Y. Peng, J. Li, C. C. Guo, B. Kang, and Z. Cao, "Capacity prediction of terminal area under convective weather based on random forest," *System Engineering Theory and Practice*, vol. 41, no. 8, pp. 2125–2136, 2021.
- [37] P. J. Shi, "On the theory and practice of disaster research," *Journal of Natural Disasters*, vol. 5, no. 4, pp. 8–19, 1996.
- [38] R. A. Davidson and K. B. Lamber, "Comparing the hurricane disaster risk of U. S. coastal counties," *Natural Hazards Review*, vol. 8, no. 1, pp. 132–142, 2001.
- [39] E. Cavallo, S. Galiani, I. Noy, and J. Pantano, "Catastrophic natural disasters and economic growth," *The Review of Economics and Statistics*, vol. 95, no. 5, pp. 1549–1561, 2013.
- [40] B. Hu, J. Z. Yan, Y. Y. Ding, H. L. Huang, and W. J. Zhao, "Typhoon disaster risk zoning model," *Journal of Natural Disasters*, vol. 21, no. 5, pp. 152–158, 2012.
- [41] R. S. Patwal, N. Narang, and H. Garg, "A novel TVAC-PSO based mutation strategies algorithm for generation scheduling of pumped storage hydrothermal system incorporating solar units," *Energy*, vol. 142, no. 1, pp. 822–837, 2018.
- [42] W. L. Xiong and B. G. Xu, "Research on optimal selection method of SVR parameters based on PSO," *Journal of System Simulation*, vol. 1, no. 9, pp. 2442–2445, 2006.
- [43] S. Y. Ding, "Comment on the marginal effect of natural geography in the junction zone of Qinling Mountains and Huang Huai plain," *Journal of Geography*, vol. 61, no. 11, pp. 1230–1231, 2006.
- [44] H. Li, X. M. Wang, X. Zhang, L. Y. Lv, and W. M. Xu, "Precipitation characteristics and extreme analysis of rare heavy rain in Northern Henan on July 19," *Meteorology*, vol. 44, no. 9, pp. 1136–1147, 2018.
- [45] Y. B. Gong, L. Xiang, and G. F. Liu, "Prediction of flood disaster loss based on Gaussian process regression model – a case study of Chongqing," *Resources and Environment in the Yangtze River Basin*, vol. 28, no. 6, pp. 1502–1510, 2019.
- [46] H. H. Li, J. D. Wu, Q. Wang, C. Yang, and S. Pan, "Research on prediction model of rainstorm and waterlogging disaster in Shanghai based on machine learning method," *Journal of Natural Disasters*, vol. 30, no. 1, pp. 191–200, 2021.
- [47] D. Huang, C. Liu, H. Fang, and S. Peng, "Assessment of waterlogging risk in Lixiahe region of Jiangsu Province based on AVHRR and MODIS image," *Chinese Geographical Science*, vol. 18, no. 2, pp. 178–183, 2008.
- [48] P. Wang and H. W. Deng, "Research on flood disaster zoning based on GIS and logistic regression model," *Progress of Geoscience*, vol. 35, no. 10, pp. 1064–1072, 2020.
- [49] B. Hu, Y. Y. Ding, L. D. He, S. N. Gu, H. L. Huang, and W. J. Zhao, "Risk compartment analysis of rainstorm and

- flood disaster in Ningbo based on fuzzy comprehensive evaluation,” *Rainstorm Disaster*, vol. 33, no. 4, pp. 380–385, 2014.
- [50] Q. H. Gong, G. Q. Huang, M. Guo, and J. X. Zhang, “Flood disaster risk compartment analysis in Guangdong Province Based on GIS technology,” *Journal of Natural Disasters*, vol. 18, no. 1, pp. 58–63, 2009.
- [51] A. D. Pan, H. P. Liu, M. Z. Yang, T. W. Huang, and X. Li, “Research on flood risk assessment in Guangzhou,” *Regional Research and Development*, vol. 29, no. 5, pp. 104–109, 2010.
- [52] H. Zhang, J. Q. Zhang, and J. S. Han, “Study on flood disaster risk assessment and zoning based on GIS technology, taking the middle and lower reaches of Liaohe River as an example,” *Journal of Natural Disasters*, vol. 14, no. 6, pp. 141–146, 2005.
- [53] Z. X. Xu, H. Chen, M. F. Ren, and T. Cheng, “Research progress on disaster mechanism and risk assessment of urban floods in China,” *Progress in Water Science*, vol. 31, no. 5, pp. 713–724, 2020.
- [54] C. Zhang, X. M. Song, and W. Li, “Application of variable fuzzy evaluation method in flood disaster evaluation,” *Journal of Natural Disasters*, vol. 1, no. 5, pp. 34–39, 2008.
- [55] P. Wang, H. Huang, and W. Wu, “Prediction model of agricultural loss caused by flood and waterlogging in Hubei Province Based on support vector machine,” *Agricultural Sciences of Hubei Province*, vol. 53, no. 18, pp. 4437–4440, 2014.
- [56] Z. Y. Fan, X. F. Gou, M. Y. Qin, Q. Fan, J. L. Yu, and J. J. Zhao, “Evaluation of geological hazard susceptibility based on the coupling of information model and logistic regression model,” *Journal of Engineering Geology*, vol. 26, no. 2, pp. 340–347, 2018.
- [57] M. Capello, L. Cutroneo, G. Ferretti, S. Gallino, and G. Canepa, “Changes in the physical characteristics of the water column at the mouth of a torrent during an extreme rainfall event,” *Journal of Hydrology*, vol. 541, no. 1, pp. 146–157, 2016.
- [58] J. Birkmann, O. D. Cardona, M. L. Carreño et al., “Framing vulnerability, risk and societal responses: the MOVE framework,” *Natural Hazards*, vol. 67, no. 2, pp. 193–211, 2013.
- [59] R. Totschnig and S. Fuchs, “Mountain torrents: quantifying vulnerability and assessing uncertainties,” *Engineering Geology*, vol. 155, no. 1, pp. 31–44, 2013.
- [60] S.-K. Min, X. Zhang, F. W. Zwiers, and G. C. Hegerl, “Human contribution to more-intense precipitation extremes,” *Nature*, vol. 470, no. 7334, pp. 378–381, 2011.
- [61] X. Wu, D. Yu, Z. Chen, and R. L. Wilby, “An evaluation of the impacts of land surface modification, storm sewer development, and rainfall variation on waterlogging risk in Shanghai,” *Natural Hazards*, vol. 63, no. 2, pp. 305–323, 2012.
- [62] M. Jakob, D. Stein, and M. Ulmi, “Vulnerability of buildings to debris flow impact,” *Natural Hazards*, vol. 60, no. 2, pp. 241–261, 2012.
- [63] F. Guzzetti, C. P. Stark, and P. Salvati, “Evaluation of flood and landslide risk to the population of Italy,” *Environmental Management*, vol. 36, no. 1, pp. 15–36, 2005.
- [64] M. Barman and N. B. Dev Choudhury, “Season specific approach for short-term load forecasting based on hybrid FA-SVM and similarity concept,” *Energy*, vol. 174, pp. 886–896, 2019.
- [65] Z. Y. Xu, M. Y. Su, and Q. A. Yao, “Hybrid kernel SVM algorithm based on Improved PSO algorithm,” *Journal of Jilin University (Science Edition): Science Edition*, vol. 56, no. 3, pp. 625–630, 2018.
- [66] H. E. Fu, Y. J. Gao, Y. Y. Fen, J. Y. Huang, and Z. F. Liu, “Risk prediction of urban rainstorm and flood disaster based on GA-SVR-C, taking Shenzhen as an example,” *Yangtze River of People*, vol. 52, no. 8, pp. 16–21, 2021.