# Research on Natural Disaster Risk Assessment and Real Estate Site Selection Strategies Based on EWM-RSR Algorithm and Arcgis

Yifan Zhang †,*, Tianxiang Huang †, Shipu Wei †

School of Electrical and Information Engineering, Anhui University of Technology, Ma'anshan, China, 243032

* Corresponding Author Email: 15656536629@163.com

† These authors contributed equally.

**Abstract.** As global climate change leads to a high incidence of extreme weather with significant socio-economic impacts, developing accurate disaster risk assessment and management models is critical to reduce the effects and optimize resources. Starting from the development of strategies for insurance practitioners around the world, this study uses research to help them find ways how to manage their underwriting and respond to the challenges of climate change. First, this study developed a catastrophe underwriting risk assessment model to predict extreme weather risks in different regions by building an accurate catastrophe risk prediction model, which insurers can use to make risk assessments. Based on the results of the risk assessment, and in conjunction with their profitability, they decide whether or not to underwrite the policy. Looking ahead, communities and real estate developers must determine how and where to build and develop. and where to build and develop. Adjustments to the insurance landscape and strategy will have a profound impact on real estate development. This paper have constructed an Arcgis-based site selection model to help communities and real estate developers select development sites with low underwriting risk and high underwriting acreage. And by using actual data to find out that the insurance company is suitable to underwrite in the Wandsworth area and not in the Lahore area.

**Keywords:** Extreme weather, Entropy weight-rank-sum ratio model, Risk assessment, Arcgis site selection.

## 1. Introduction

In recent years, with the intensification of global climate change, there has been a significant increase in the frequency and intensity of extreme weather events, and natural disasters, ranging from floods and droughts to typhoons, have occurred frequently, with enormous socio-economic impacts. [1] These disasters not only threaten human lives and properties but also pose unprecedented challenges to global economic development and social stability. Therefore, the establishment of effective disaster risk management and assessment models is of great significance in reducing the impact of disasters and optimizing resource allocation. In the insurance industry, how to accurately assess disaster risks and rationally formulate insurance products is the key to the sustainable and sound development of insurance companies.

Based on the reference of related literature and research results, this study selects a total of 10 indicators of social, ecological, and climatic factors, combines the EWM-RSR algorithm to obtain an objective catastrophe risk factor cr, and constructs a break-even algorithm to evaluate the profitability of insurance companies to underwrite or not to underwrite in terms of underwriting profit and loss. Finally, this study chose two districts, Wandsworth and Lahore, on different continents to demonstrate our model. In the future, this study predicts that communities and real estate developers will be more inclined to choose locations with low risk of natural disasters for development and investment. By analyzing the team's relevant geographic data, appropriately adjusting the influencing factors in the catastrophe risk factor, and combining with the Arcgis site selection method, this paper comprehensively select a location that meets all the requirements as a location for real estate development, providing real estate developers with a reliable method of selecting site.

## 2. Indicator selection and algorithm application for disaster risk assessment

### 2.1. Establishment of catastrophe risk assessment factors

Catastrophe risk assessment in different regions is influenced by three main dimensions: social, ecological climatic impacts. [2] This study got it from www.comap.com/contests/mcm-icm. Then, we take these three main aspects as first-level indicators. For these three primary indicators, this method finally retained the 10 most representative secondary indicators to construct our evaluation model by reviewing relevant information and combining relevant literature, the nature of the indicators is shown in Table 1.

**Table 1.** Selection of indicators

| Level 1 indicators | Level 2 indicators | Description | Nature of the indicator |
|---|---|---|---|
| Society | GRP | Gross region production | - |
| | FAI | Fix asset investment | - |
| | PRP | Permanent resident population | - |
| Ecology | DPE | Disaster prevention expenditure | - |
| | FCP | Forest coverage rate | - |
| | DOD | Degree of grought | + |
| | SER | Soil erosion rate | + |
| Climate | AP | Amount of precipitation | + |
| | WS | Wind speed | + |
| | NND | Number of natural disasters | + |

This study consider three main aspects: social impact, ecological impact and climate impact, which are described in the following sections.

Human social activities often lead indirectly to large-scale natural disasters. For example, human beings engage in large-scale mining of mineral resources, groundwater extraction, and other activities for the sake of economic development, and these actions may lead to geological disasters such as ground subsidence and earthquakes. At the same time, in more socio-economically and in other developed areas, human utilization of natural resources tends to be higher, so it is necessary to consider these social indicators.

The impact of ecosystems on natural disasters is far-reaching and complex. Forest destruction, including excessive deforestation and forest degradation, reduces the water storage capacity of the ground and increases the rate of runoff, leading to soil erosion. This not only increases the frequency and intensity of floods but can also lead to disasters such as mudslides; over-exploitation of underground resources (e.g. minerals, groundwater) can lead to geologic disasters such as ground subsidence and ground cracks. In addition, activities such as the construction of large reservoirs and the over-pumping of groundwater may alter the water table, triggering or exacerbating phenomena such as earthquakes and ground subsidence. Geological damage not only threatens the safety of human habitation but may also destroy infrastructure and cause significant socio-economic losses.

There is a strong link between climate change and natural disasters, and the trend toward an increase in the frequency and intensity of certain types of natural disasters poses a great challenge to human societies. Global warming is causing sea levels to rise as a result of the melting of the polar ice caps and the thermal expansion and contraction of seawater. This not only threatens the safety of coastal cities and low-lying countries but may also lead to saltwater intrusion into freshwater systems, affecting agriculture and drinking water resources; the intensity of tropical storms (including typhoons and hurricanes) is likely to increase as sea surface temperatures rise. This means that storms bring higher wind speeds and higher rainfall, thus causing more damage to human society and the natural environment.

In the establishment of this model, this model use the entropy weight method (EWM) to find the objective weights, which can avoid the bias caused by subjective factors, and the specific operation steps are as follows:

(1) Data normalization: This model normalize the data so that the data is restricted to [0, 1] to reduce the adverse effects caused by singular sample data.

Construct the original matrix. n evaluation indicators from m samples form the original matrix X=$\left(x_{ij}\right)_{m\times n}$, included among these, $x_{ij}$ is the jth indicator in the ith sample, $1 \le i \le m$, $1 \le j \le n$.

In positive indicators:

$$x'_{ij} = \frac{x_{ij} - min(x_{ij})}{max(x_{ij}) - min(x_{ij})}, i = 1,2 \dots, m; j = 1,2, \dots, n \tag{1}$$

Among the negative indicators:

$$x'_{ij} = \frac{max(x_{ij}) - x_{ij}}{max(x_{ij}) - min(x_{ij})}, i = 1,2 \dots, m; j = 1,2, \dots, n \tag{2}$$

In the formula: $\min\left(\mathrm{x_{ij}}\right)$ is the smallest value in the original matrix X; $max\left(x_{ij}\right)$ is the maximum value in the original matrix X.

(2) Calculate the information entropy: P represents the proportion of relevant elements in the feature column, which is calculated as follows, where the formula represents the value of each element, and M represents the summation of all elements in the column where the element is located. Then use the information entropy formula to calculate the information entropy of each indicator, which is calculated as follows:

$$\mathrm{H} = -\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{p_i} \log_2 \mathrm{p_i} \tag{3}$$

$$e_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} \left[p_{ij} \cdot \ln\left(p_{ij}\right)\right] \tag{4}$$

After calculating the weights of the indicators from the above method, this method use the rank-summation comprehensive evaluation model to assess the risk of regional catastrophe occurrence. The principle of rank-sum-ratio comprehensive evaluation method is to obtain the dimensionless statistic RSR through rank transformation in a matrix with n rows and m columns; the RSR value is used to directly rank the advantages and disadvantages of the evaluation objects or to rank them by grades, to make a comprehensive evaluation of the evaluation objects. The specific steps are as follows:

(3) Organize and rank. Arrange the m indicators of the n evaluation objects into a raw data table with m rows and n columns. Compile the rank of each evaluation object for each indicator, in which the efficiency indicators are compiled from small to large, the cost indicators are compiled from large to small, and the average rank is compiled for the same indicator with the same data. The obtained rank matrix is denoted as $\mathrm{R} = \left(\mathrm{R_{ij}}\right)_{\mathrm{m\times n}}$.

(4) Ralculate the rank sum ratio (RSR). According to the formula:

$$RSR_i = \frac{1}{mn} \sum_{j=1}^{m} R_{ij}, i = 1,2 \dots, n \tag{5}$$

When the weights of the evaluation indicators are different, this model calculate the weighted rank-sum ratio (WRSR), which is given by the formula:

$$WRSR_i = \frac{1}{mn} \sum_{j=1}^{m} w_j R_{ij}, i = 1,2 \dots, n \tag{6}$$

In the formula: $\mathrm{w_j}$ is the weight of the jth evaluation indicator.

(5) Graded sorting. According to the value $RSR_i$ of $WRSR_i$(i=1, 2, …, n). The evaluation objects are ranked in grades, and the larger the value, the better their evaluation.

Finally, according to the Probit critical value substituting into the regression model to calculate the RSR critical (fitted value), RSR critical (fitted value) for interval comparison, and then get the staging grade level. The RSR critical (fitted value) is used as the catastrophe risk coefficient (CR), combined with the normal distribution of the relevant division method for grading, and then the grading level.

According to the entropy weighting method mentioned above, this paper find the weights occupied by each indicator as shown in Table 2.

**Table 2.** Indicator weights

| norm | information entropy (physics) | information utility valued | Weights (%) |
|---|---|---|---|
| degree of aridity | 0.808 | 0.192 | 18.736 |
| soil erosion rate (%) | 0.917 | 0.083 | 8.103 |
| Precipitation (mm) | 0.842 | 0.058 | 5.664 |
| Wind speed (m/s) | 0.926 | 0.074 | 7.243 |
| Number of earthquake disasters | 0.56 | 0.44 | 42.992 |
| Gross regional product (billions of dollars) | 0.97 | 0.03 | 2.927 |
| Region-wide investment in fixed assets (billions of dollars) | 0.952 | 0.048 | 4.642 |
| Resident population | 0.968 | 0.032 | 3.08 |
| Government expenditure on disaster prevention and mitigation ($ million) | 0.984 | 0.016 | 1.519 |
| Forest and grass cover (%) | 0.948 | 0.052 | 5.093 |

After the rank-sum ratio evaluation model, this study obtained the following bins for the catastrophe risk coefficient (CR):

At 0<CR<0.15, Low risk of catastrophe in the region;

At 0.15<CR<0.45, there is a medium risk of a catastrophe occurring in the area;

At CR>0.45, the region is at high risk of catastrophe.

### 2.2. Underwriting profit-loss model

When an insurance company intends to underwrite in a certain region, it needs to consider its profitability to determine whether it has the strength to underwrite in this region, as follows: Let the insurance company's underwriting income be P, the number of policies is m, and the premium be n. Then:

$$P = m \cdot n \tag{7}$$

Its insurance fixed costs are $w_1$, the unit variable cost is$w_1$, the total cost is W, then:

$$W = w_1 + \sum_1^m w_2 \tag{8}$$

Let the total amount of its compensation be H, the payout rate be u, and the present value coefficient of compound interest be l, then:

$$H = u \cdot P \cdot I \tag{9}$$

Where the coefficients of the present value of compound interest l, t are the number of years of coverage, and the formula for the present value of compound interest is:

$$I = \frac{1}{(1+i)^t} \tag{10}$$

Let q be the probability of the area being affected, u be the odds ratio, and CR be the catastrophe risk coefficient that has been derived:

$$u = CR \cdot q \tag{11}$$

Let the profit made by the insurance company be X, then:

$$X = P - W - H \tag{12}$$

In formula (12), the company is profitable when X>0; when X<0, the company loses money.

After establishing the catastrophe risk assessment model and underwriting break-even analysis model, this research selected Wandsworth with a catastrophe risk factor of 0.05 and Lahore with a catastrophe risk factor of 0.62 for model demonstration to determine whether we should underwrite in these two areas. The following is a description of the data and the results for the two regions.

By querying the region's insurance companies for relevant data, this research came up with the following indicators. As shown in Table 3 and Table 4.

**Table 3.** Wandsworth's norm

| Norm | Reference value |
|---|---|
| Average cost per policy | 1000 |
| Total underwriting costs | 100000 |
| Average individual policy claims | 2000 |
| Time value of money discount rate | 0.0525 |
| Claims rate | 0.003 |

Profit and loss analysis shows that profit is greater than zero, so the insurance company can make a profit by underwriting in this region, and the consolidated solvency ratio is 253.56%, which is sufficient to pay all the claims of the policies, which is conducive to the healthy development of the insurance company in the long term. Therefore, the insurance company can underwrite in this region.

**Table 4.** Lahore's norm

| Norm | Reference value |
|---|---|
| Average cost per policy | 10 |
| Total underwriting costs | 1200 |
| Average individual policy claims | 90 |
| Time value of money discount rate | 0.07 |
| Claims rate | 0.124 |

The profit and loss analysis shows that the profit is less than zero, so the insurance company cannot make a profit by underwriting in this region, and the consolidated solvency ratio is 78.92%, which is not enough to pay all the claims of the policies, which is not conducive to the long-term healthy development of the insurance company. Therefore, the insurance company should not underwrite insurance in this region.

## 3. Geographic Data Analysis for Optimizing Real Estate Site Selection

When looking to the future, communities and real estate developers need to determine how and where to build and grow. [3] Changes in the insurance landscape have significant implications for real estate decisions. Real estate developers will inevitably prefer to develop in areas where the underwriting risk is low, i.e. where the land they develop can be insured. This study make appropriate adjustments to the social, ecological, and climatic influences in the catastrophe risk model in the previous question to provide an Arcgis-based site selection model for real estate developers.

### 3.1. Site Selection Requirements

#### 3.1.1 Social Impact

When considering the location of a real estate project, social factors are an important dimension to consider, which directly affects the project's market attractiveness, future value-added potential, and social benefits. [4]

Economic situation:

Economic factors such as a community's economic level, employment rate, and the average income of its residents can also affect the value and attractiveness of real estate. Areas with good economic conditions tend to attract more homebuyers and investors.

Population structure:

The size of the population, age distribution, and family structure in the area can affect the demand for real estate. For example, young families may be more inclined to look for homes close to schools and children's activity centers, while retirees may prefer quieter neighborhoods with good support services.

### 3.1.2 Ecological Impact

Ecological factors are a very important aspect when considering the location of a real estate project, not only for the sustainability of the project but also for the quality of the living environment and the health of the ecosystem. Areas with better ecological environments have a lower probability of extreme weather and thus a lower probability of natural disasters destroying the development site.

Conservation of natural resources:

Site selection should avoid occupying or destroying precious natural resources, such as water sources, forests, and wetlands. These areas usually have important ecological value and play an irreplaceable role in maintaining biodiversity, regulating climate, and protecting water resources.

Soil and terrain considerations:

Factors such as soil type and topographical features need to be considered when evaluating potential real estate development sites. For example, real estate projects should be avoided in areas prone to frequent natural disasters such as mudslides, landslides, and flooding. Also, consider the soil's carrying capacity and permeability to ensure the stability of the building and its ability to prevent flooding and drainage.

### 3.1.3 Climate Impact

Consideration of climatic factors when selecting a site for real estate is critical, as climatic conditions directly affect the design, construction, and longevity of a building, as well as the comfort and health of its occupants. [5]

Temperature:

The degree is a key factor in real estate site selection. Areas with high or low temperatures require special building materials and designs to ensure comfortable indoor temperatures. For example, in hot areas, more ventilation and shading may be needed, while in colder areas, good insulation and heating systems are required.

Measured quantity of rain:

Precipitation conditions such as rainfall and snowfall are equally important for real estate siting. In rainy or snowy areas, the roof of a building needs to be designed to withstand the weight of precipitation, and the drainage system needs to be specially designed to prevent water and moisture accumulation.

## 3.2. Demonstration of site selection model


**Figure 1.** Distribution of available resources

The catastrophe risk coefficient CR of Hangzhou is 0.07, so the area is a low-risk underwriting area and suitable for real estate development. This paper selected part of the urban area from 31.32°N to 31.35°N and 120.59°E to 120.61°E as a demonstration of the site selection model. For a real estate developer to select a site for development, it is important to know the distribution of land and roads in the area and to have a general understanding of the area in which the development will take place. The figure shows the distribution of available land, urban roads, and rivers in the area. [6] The results are shown in Figure 1.

Preliminary Screening: Avoiding the saturated number of commercial building and neighborhoods, a preliminary screening of areas available for real estate development was conducted. The preliminary screening areas are shown in the Figure 2.

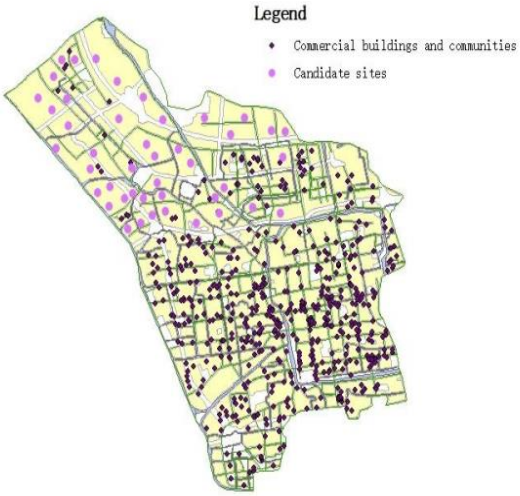


**Figure 2.** Initial screening area

Further screening is required for candidate areas that have been initially screened. [7] The site of a new property must be on an unallocated site. The boxed unallocated sites are shown in Figure 3.

**Figure 3.** Undistributed resource areas

In addition to meeting the above requirements, the site should also meet the requirement of being within 200 meters of a road, which is conducive to the convenience of transportation. Considering the risk of heavy rainfall, the site should be located within 150 meters of a river to effectively protect against the possible effects of flooding. The results are displayed in Figure 4.

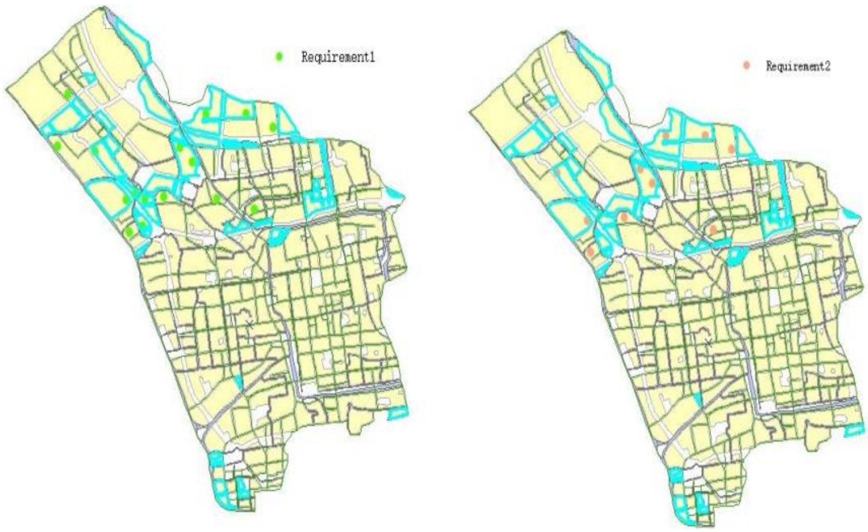


**Figure 4.** Areas that can withstand heavy rainfall

Finally, the overlapping locations were selected among the locations that met all the requirements and were identified as the final site locations. The identified site locations are shown in the figure 5. [8]

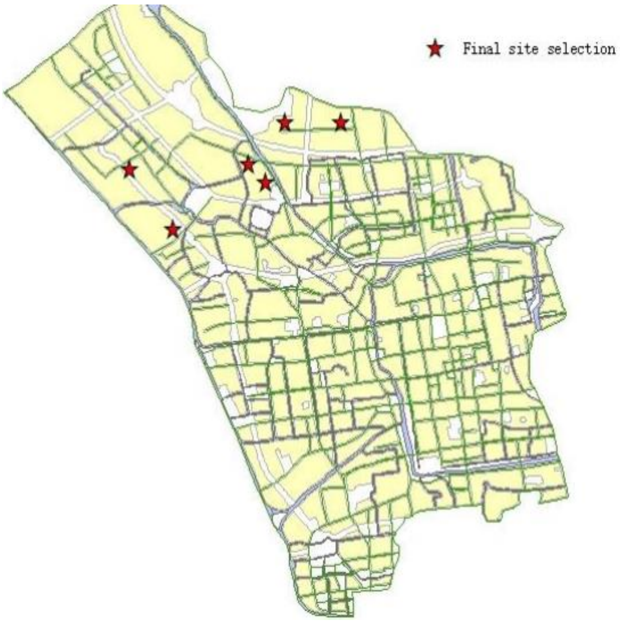


**Figure 5.** Final Site Map

## 4. Conclusion

In recent years, with the increasing human activities, the global climate situation is facing a great test. Floods, droughts, typhoons etc. Natural disasters triggered by extreme weather have caused immeasurable economic losses to mankind. With climate change and extreme weather events on the rise, property owners and insurance companies are facing an unprecedented crisis. For insurance practitioners, this study selected a total of 10 indicators of social, ecological, and climatic factors, combined with the EWM-RSR algorithm to obtain an objective catastrophe risk coefficient, CR, and constructed a breakeven algorithm to evaluate the profitability of insurance companies in terms of underwriting profit and loss to underwrite or not to underwrite. Finally, two districts, Wandsworth and Lahore, were selected for the validation demonstration in this study, and the results show that Wandsworth district can be underwritten while Lahore district cannot be underwritten. [9]

Going forward, community and real estate development will inevitably choose locations with low risk of natural disasters. For real estate developers, this study combines Arcgis's site selection methodology to comprehensively select a location that meets all requirements as a reliable site selection method for real estate development. This paper provides a research idea and framework applied to the field of insurance assessment related to the field of real estate site selection, proving the feasibility of the entropy weight-rank-sum ratio algorithm and the Arcgis site selection method. [10]

## References

[1] Fengtiao Tian. The rank-sum ratio method and its application [M]. Beijing, China Statistics Press, 1993.

[2] Baum D S. Assessing the risk of takeover catastrophe from large language models. [J] Risk analysis: an official publication of the Society for Risk Analysis, 2024.

[3] Marlina H, Ruslanjari D, Hakim B I. Disaster risk financing and insurance for earthquake-prone state buildings in Indonesia [J]. Jàmbá: Journal of Disaster Risk Studies, 2024, 16 (1): e1-e8.

[4] Wang S. A Study of Disaster Insurance in Extreme Weather Based on Logistic Regression and ARIMA Models [J]. Advances in Computer, Signals and Systems, 2024, 8 (4).

[5] Yingjie F, Frank S, Stefan V, et al. Catastrophe insurance and flexible planning for supply chain disruption management: a stochastic simulation case study [J]. International Journal of Production Research, 2024, 62 (4): 1108-1125.

[6] Ghoreishi M, Lindenschmidt E K. Unlocking effective ice-jam risk management: Insights from agent-based modeling and comparative analysis of social theories in Fort McMurray, Canada [J]. Environmental Science and Policy, 2024, 157103731-.

[7] Li H, Wang Y, Chen W, et al. Adaptability Evaluation of Human Settlements in Chengdu Based on ArcGIS [J]. Sustainability, 2024, 16 (5).

[8] Liu C, Pan H, Wei Y. Spatial Distribution Characteristics and Influential Factors of Major Towns in Guizhou Province Analyzed with ArcGIS [J]. Sustainability, 2023, 15 (14).

[9] Liu C, Pan H, Wei Y. Spatial Distribution Characteristics and Influential Factors of Major Towns in Guizhou Province Analyzed with ArcGIS [J]. Sustainability, 2023, 15 (14).

[10] Sun X, Gu H, Tian Y. Urban Building Thermal Comfort Research Based on ArcGIS and Building Parameters: Analyzing and Improving Measures [J]. International Journal of Heat and Technology, 2023, 41 (2).