Financial Engineering and Risk Management (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/ferm.2024.070420
ISSN 2523-2576 Vol. 7 Num. 4

Research on Extreme Weather Risk Assessment Model Based on Entropy Weight Method and Cluster Analysis Method

Jiayu Zou, Yankun Guo, Liming Liu

School of Energy, Power and Mechanical Engineering, No. 2 Campus of Baoding Subordinate to North China Electric Power University, Baoding, 071003, China

Keywords: Insurance, Conservation, K-Means, Risk, Entropy Weight Meth

Abstract: In recent years, the number of extreme weather has gradually increased, and the real estate industry has suffered setbacks. Meanwhile, the premium of property insurance has correspondingly increased, and the real estate industry and the insurance industry have encountered difficulties. Aiming at underwriting risks for insurance companies, this paper establishes an Extreme Weather Risk Assessment Model (EWA) based on K-Means, which is used to evaluate whether to write insurance policies in an area where the frequency of extreme weather is increasing. Two regions, Texas and London, were selected for practical analysis. Aiming at the investment Location problem of Real Estate companies, this paper establishes the optimization Model of Real Estate Location Evaluation Model (REE) by using Entropy Weight Method and Cluster Analysis Method on the basis of EMA model, which enables real estate companies to evaluate whether to invest in certain locations.

1. Introduction

The property insurance and real estate investment industries face significant sustainability and decision-making issues in the face of climate change and extreme weather events. Industry often uses catastrophe models to assess the risks posed by extreme climate hazards, which are designed to analyze the risk of low-frequency, high-damage hazards. Because extreme disasters are rare, only a small number of actual observations are recorded in historical data. Using the model to make projections that rely on historical, incomplete data and potentially outdated models is likely to underestimate risk and premium pricing; Especially for low-frequency, high-loss risks, this approach has the potential to cause risk assessments to lag decades behind the current risk landscape[1][2][3][4]. Based on entropy weight method and K-means clustering analysis method, this paper establishes an extreme weather risk assessment model with higher reliability through comprehensive analysis of historical data, which can be used to help insurance companies analyze which insurance model should be adopted for specific areas

2. The Extreme Weather Risk Assessment Model

With the increase of natural disasters, in order to better deploy property insurance, paper propose a regional Extreme Weather Risk Assessment Model (EWA). The model is used by insurance companies to assess whether an area is worth covering. Based on the characteristics of extreme weather in an area over the course of a year, Climate Model is created to describe the probability of extreme weather in an area over the course of a year. And using K-means Cluster Analysis Model[5], the risk of a region is divided. Then, based on the data of total assets and premium rate of insurance companies, the EWA is established by using the time value theory of funds[6]. According to the model, whether the insurance company should bear the corresponding risk is evaluated. According to the two models established above, the measures that owners should take are summarized, and the relevant data of Texas and London are brought into the model for verification calculation.

2.1 The Establishment of the Climate Model

2.1.1 The Extreme Weather Risk Assessment Model

In order to describe the probability of extreme weather in a region in a year, paper define an indicator as the frequency of extreme weather events P_w, and the expression for calculating P_w is as follows:

$$P_w = \alpha \frac{\sum_{i=1}^{n} x_i}{365} \tag{1}$$

Where x represents the number of extreme weather events in a year, i and n represent a certain type and number of extreme weather events, respectively. α is the climate disturbance factor, which 0.9 is taken here.

Paper bring in the extreme weather data of the United States in 2021, you can know the probability of four kinds of extreme weather in each state in the United States such as Wildfires, Floods, Storms, Extreme temperatures. The Distribution of extreme weather in the 50 states of the United States in 2021 is shown in Fig 1. At the same time, the average cost of a claim across all 50 states is shown in Fig 2.

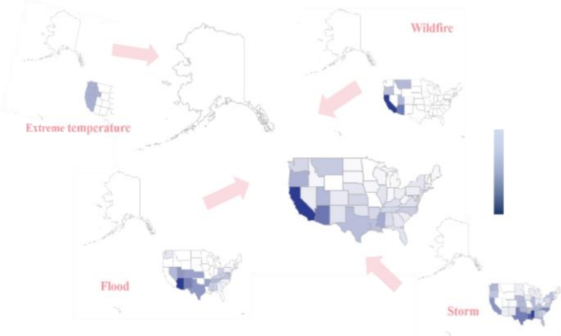

Figure 1: Distribution of extreme weather in the 50 states of the United States in 2021

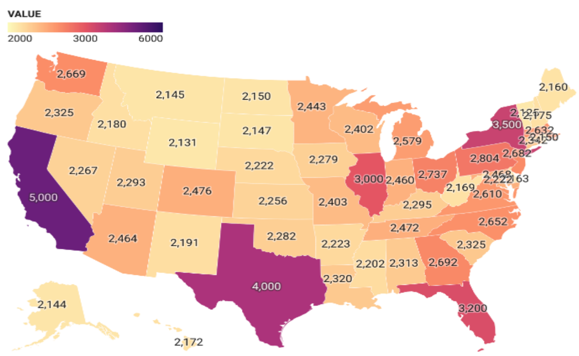

Figure 2: The average cost of a claim across all 50 states

The P_w values can be calculated separately for each state.

Based on state P_w values and the average claim cost C of insurance expenses in the known states, the coverage risk F_A can be calculated:

$$F_A = P_w \cdot C \tag{2}$$

2.1.2 Regional risk division

Calculated P_w values and insured risks and K-means clustering based on state risk results for 50 US states. The risks are classified into three risk levels which is shown in Table. 1.

Table 1: Regional classification based on risk

Risk	High	Middle	Low
Area quantity	3	28	19
Maximum risk	492.8	300	92
Minimal risk	300	92	0

Through statistical analysis of the data, paper can draw Fig 3 and Fig 4.

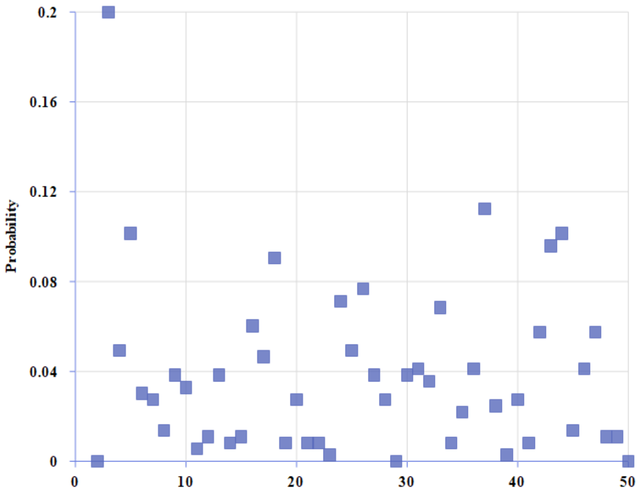

Figure 3: The P_w in all 50 U.S. states

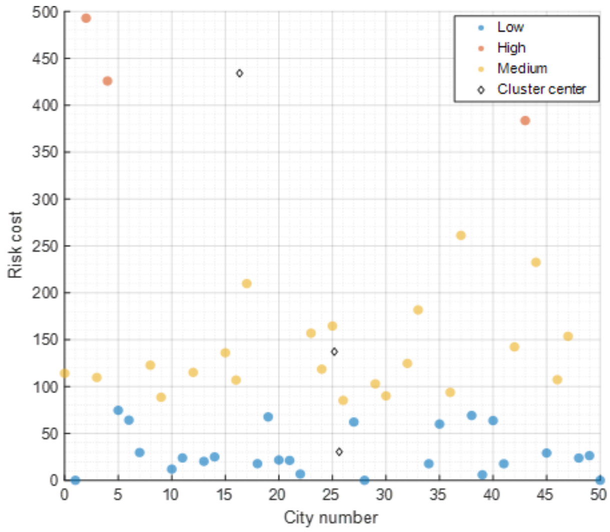

Figure 4: Value at risk for all 50 U.S. states

2.2 The Establishment of the EWA

To build an underwriting risk assessment model, paper consider the next three years and select the following metrics:

First, paper calculates F_A from the P_w derived from the climate model and the value of C for a selected area.

Obtain the annual insurance cost for the selected area F_I:

$$F_I = F_A \cdot P_I \tag{3}$$

$$P_A = \frac{F_A}{A} \times 100\% \tag{4}$$

Based on The Time Value Theory of Funds[4], the three-year premium for the region is calculated here:

$$F_{P,i} = \frac{F_I}{(1+P_A)^i} \tag{5}$$

$$F_P = \sum_{i=1}^{3} F_{P,i} \tag{6}$$

Compare the calculated three-year total premium F_P.If,

$$F_P < q \cdot A \tag{7}$$

It means that the insurance company can cover the area, otherwise it can not be covered.

As owners, they can take the necessary measures to reduce the risk in the community based on the model established above. Assuming that the insurance company's insurance rate for a region P_I stable at a value, the owner can take measures to reduce the cost of risk claims C, reduce the F_I, and then reduce the premium.

Reducing the cost of risk claims can be done by strengthening the building's resilience, complying with local building codes, and participating in climate education and early warning systems.

2.3 Related calculation of indicators

According to the different regions, paper will use data from two regions: Texas, USA, and London, UK. And bring them into the EMA for analysis and test to demonstrate our model.

Paper assume that the total assets of the insurance company is $5 million.And anther Raw data for both regions is shown in Table 2.

Table 2: Raw data for both regions

Date	Texas	London
P_w	0.0959	0.0411
C	4000	4410
A	5,000,000	5,000,000
q	50	50
P_I	3429	3931

The corresponding F_P for London is $4.27 million, for Texas $7.9 million.

The London F_P is less than the preset value of '$q \cdot A$', indicating that the London region can bear the corresponding risk. It indicates that the London area can take the corresponding risks and undertake the insurance. And Texas does not recommend insurance coverage.

3. The Real Estate Location Evaluation Model

Unlike insurance companies, communities and real estate developers assess the risk of a place not only for climate factors, but also for the population and future development of the area. Therefore, paper need a site assessment model for real estate companies to make the risk assessment of a site more flexible. Paper improve the underwriting risk calculation formula of the first question and redefine the Real Estate Siting Risk Index (RSI). Paper list the relevant indicators that affect RSI in Fig 5.

3.1 Selection of indicators

First, paper apply The McKinsey Logical Tree[7] to identify all the necessary indicators that can be used to assess the risk of a house. Paper divided the factors affecting RSI into three categories, and selected corresponding indicators respectively.

Paper use the Insurance Coverage Gap (ICG) calculated in the EWA to reflect the impact of insurance model on site selection. The smaller the insurance coverage gap, the smaller the risk index.

Since the insurance coverage gap has taken into account the impact of extreme weather, paper only need to select the Earthquake Hazard Rating (EHI) of the region to analyze the environmental impact[8]. The higher the earthquake disaster level, the greater the real estate loss caused by the earthquake. It is worth mentioning that areas with high earthquake hazards are not necessarily at high earthquake risk. Earthquake risk is defined as the damage that can be caused by exposure to an earthquake disaster, and it depends not only on the disaster level, but also on the number of people and property exposed to the disaster and the degree of vulnerability of people and property to the disaster[9].

The Building Earthquake Resistance Rating (BSR) can reflect the degree of resistance of a real estate to earthquake disasters. Therefore, paper believe that the higher the earthquake resistance level of the building, the smaller the risk of real estate investment in the region.

Another building design indicator is The Proportion of Weather-resistant Materials and Renewable Energy Use (PRU). The use of weather-resistant materials can make the house withstand the test of climate, such as the comprehensive damage caused by light, cold and hot factors, so that its tolerance is stronger. The use of renewable energy is conducive to the long-term development of real estate companies in the region. Therefore, paper believe that the greater the proportion of weather-resistant materials and renewable energy use, the smaller the risk of real estate investment.

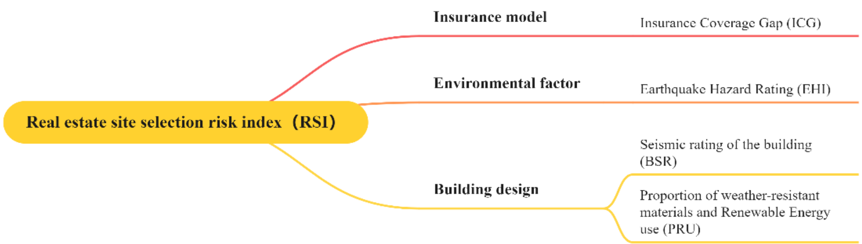

Figure 5: Indicators that affect RSI

Paper have collected relevant indicator data as shown in the following Fig 6.

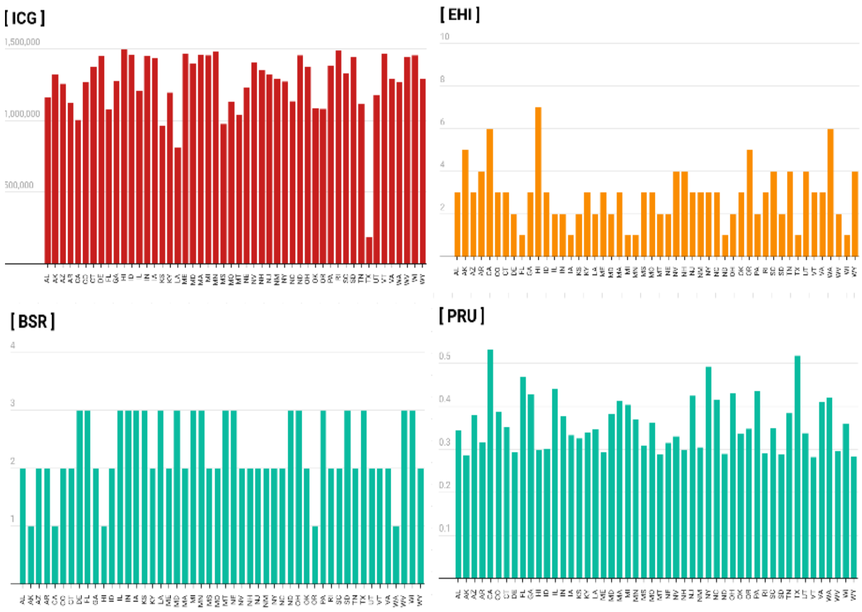

Figure 6: The value of each indicator

3.2 Related calculation of indicators

First, paper use the Entropy Weight Method[10] to calculate the weight of each index.

Here, paper obtained relevant data for 50 regions. They can be divided into positive and negative indicators. The increase of the positive indicators represents the increase of the development index, but not the negative indicators.

To normalize all data equally, paper first process all negative indicators,

$$X_{ij} = max(X_i) - X_{ij} \tag{8}$$

And all the data are normalized,

$$Y_{ij} = \frac{X_{ij} - min(X_i)}{max(X_i) - min(X_i)} \tag{9}$$

Where X_{ij} Represents the raw data of the i th index value of the national j, and $\max(X_i)$ and $\min(X_i)$ represent the minimum and maximum data of the i th indicator of national j.

Calculate national information entropy E_j,

$$P_{ij} = \frac{Y_{ij}}{\sum_{i=1}^{n} Y_i} \tag{10}$$

$$E_j = -\frac{1}{\ln n} \sum_{i=1}^{n} P_{ij} \ln P_{ij} \tag{11}$$

If $P_{ij} = 0$, substitute the (12) into the (11):

$$\lim_{P_{ij} \to 0} P_{ij} \ln P_{ij} = 0 \tag{12}$$

Based on information entropy $\mathrm{E_j}$, calculate the weight of each indicator:

$$W_i = \frac{1 - E_i}{k - \sum E_i} (i = 1,2, \ldots, k) \tag{13}$$

The results are as follows Table 3. Paper obtained the analysis of the weight results of relevant indicators affecting RSI and RSI values by states, as shown in Figure 7 and Figure 8.

Table 3: The weight of each indicator

	ICG	EHI	BSR	PRU
Weight	0.4079	0.0690	0.1692	0.3539

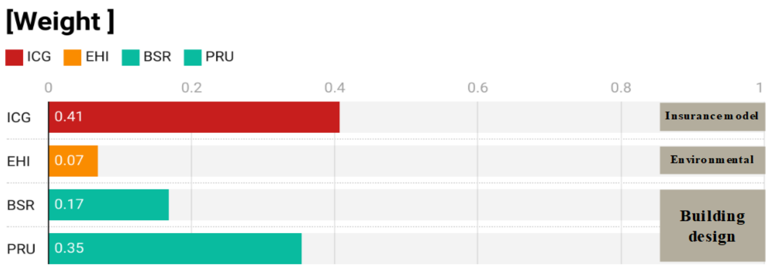

Figure 7: Weight result analysis

According to the calculated weights, the weighted average method is used to calculate the RSI.

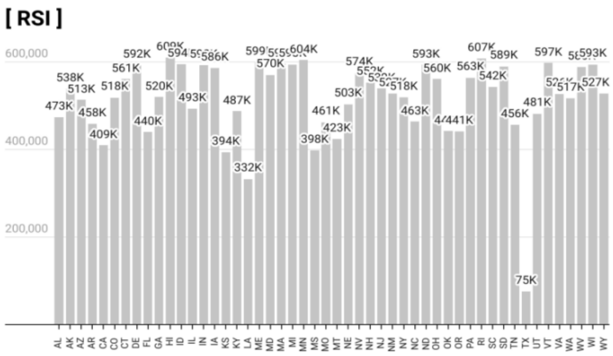

Figure 8: RSI values by states

Then, according to K-means, regions are divided into three categories according to RSI to evaluate whether real estate companies are suitable for investment and construction in a certain region.

Table 4: Regional classification based on RSI

RSI	High	Middle	Low
Area quantity	24	24	2
Maximum RSI	609,414	530,000	350,000
Minimal RSI	530,000	350,000	75,366

Paper sorted out the data of 50 regions by K-Means and clustered them into three risk index groups according to the RSI value as High, Medium and Low which is shown in table 4.

Paper believe that low and medium risk areas are suitable for investment construction, and high risk areas are not suitable.

For example, paper analyzed Arizona and West Virginia respectively. From the above table, paper can see that the RSI values of Arizona and West Virginia are 393806.3837 and 598562.0815 respectively. According to the RSI grouping, it is not difficult to see that Arizona and West Virginia belong to the medium risk and high risk areas respectively. When it comes to choosing an area as the site for building a community, as a real estate developer, I will choose the medium-risk area of Arizona rather than the high-risk area of West Virginia.

4. Conclusions

Based on K-Means and Entropy Weight Method, paper evaluates the underwriting risks of insurance companies and the investment risks of real estate companies in view of the difficulties of the real estate industry and the insurance industry caused by the gradual increase in the number of extreme weather in recent years. Aiming at the underwriting risk of insurance companies, this paper establishes a regional extreme weather probability assessment model, integrates the data of total assets and premium rate of insurance companies, and classifies a region into underwriting high, medium and low risk areas by using the theory of time value of funds. Aiming at the investment risk problem of real estate companies, this paper classifies a region as a high, middle and low risk region based on the relevant indicators and weights comprehensively selected by McKinsey Logical Tree analysis and Entropy Weight Method. In this paper, subjective and objective weighting methods are used to solve the problem, so that the model can take into account the objectivity of science and the subjectivity of human intervention when evaluating each index. It solves the risk assessment problem of real estate companies and insurance companies in a certain area under extreme weather.

References

[1] Charu Singh, Sanjeev Kumar Singh, Prakash Chauhan, Sachin Budakoti. Simulation of an extreme dust episode using WRF-CHEM based on optimal ensemble approach [J]. Atmospheric Research. Volume 249, Issue. 2021. PP 105296

[2] Andrea Lang, Benjamin Poschlod. Updating catastrophe models to today's climate–An application of a large ensemble approach to extreme rainfall [J]. Climate Risk Management. Volume 44, Issue. 2024. PP 100594

[3] Du Yue, Pan Yanxi, Ma Tengfei, et al. Development of catastrophe model and construction of catastrophe risk guarantee system [J]. Insurance theory and practice, 2023, (12):18-31.

[4] Wang Xuyi. Research on multi-level catastrophe insurance equilibrium model in China [D]. Shanghai University of Finance and Economics, 2023.

[5] Qin Maogang, Long Genyuan, Li Haiyun, et al. Quantitative analysis of geological environment stability of Zhongsha atoll based on K-means clustering hierarchical analysis model [J]. Journal of Tropical Oceanography, 2019, 42(02):113-123.

[6] He Ping. Analysis on the specific application of time value of funds in asset valuation income method [J]. Modern Economic Information, 2018, (19):175-176.

[7] Xu Yi, Qiu Aijun, Xie Zhenyu. Research on measures to improve quality and efficiency of procurement of sporadic materials based on Logic tree [J]. Bidding and Procurement Management, 2022, (11):48-49.

[8] Xu Shanshan, Huang Wenxin. A Preliminary study on Risk Management of Financing Credit Guarantee Insurance from the perspective of Robustness Assessment -- Based on a city survey [C]// China Association for the Promotion of International Science and Technology Working Committee, Nanyang Academy of Sciences. Proceedings of the International Academic Forum on Economic Management Studies. Yichang Central Branch, People's Bank of China; 2022:3.

[9] Wang Kongsong. Internet insurance risk identification research [D]. Foreign economic and trade university, 2022.

[10] Stepinac M , Loureno P B , Atali J ,et al.Damage classification of residential buildings in historical downtown after the ML5.5 earthquake in Zagreb, Croatia in 2020[J].International Journal of Disaster Risk Reduction, 2021, 56(11):102140.DOI:10.1016/j.ijdrr.2021.102140.