

A Study of the Insurance Industry in Extreme Weather Based on ARIMA Models and EWM-TOPSIS

Chen Liu^{1,*}, Xianghan Wang², Siyuan Wei³

¹ International College of Zhengzhou University, Zhengzhou University, Zhengzhou, China

² School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, China

³ School of Civil Engineering, Zhengzhou University, Zhengzhou, China

*Corresponding Author: liuc9654@163.com

ABSTRACT

This paper evaluates the underwriting approach of property insurance by comparing the total premium per person and affordability per person, using the ARIMA model to predict the underwriting decisions in Japan and Turkey for the next five years, combining factors such as population density and economic development on a smaller regional scale, and establishing a second model, the FRA. Applying the theory of catastrophic factors, the Fuzzy Risk Assessment Model calculates the causal factor intensity risk values and vulnerability values to derive a more accurate integrated risk assessment value. Afterwards, the URA model is applied to the problem of community siting in Australia and the C-GIS model is established. By analysing factors such as terrain data, population distribution and average temperature to determine suitable locations for development, the URA model is added to analyse the risk level under extreme weather and generate satellite maps for real estate siting. Finally, buildings of cultural or community importance in the area are considered, and a DPB model is built to solve the risk scoring and protection method using EWM-TOPSIS, which is classified into three levels: relocation protection, restricted protection and normal protection.

KEYWORDS

Extreme-weather events; URA model; FRA model; C-GIS model

1. INTRODUCTION

Property insurance plays a crucial role in the face of extreme weather and natural disasters, providing essential risk coverage and protection measures for the real estate industry. However, when selecting appropriate underwriting methods and protection strategies, various factors need to be considered, such as population density, economic development, and topographical data. This study aims to provide a more comprehensive and scientific guide to property insurance decision-making by comparing different underwriting methods and site selection options, combined with risk assessment models and protection value considerations. By building models such as ARIMA, URA and DPB, this study will delve into the coping strategies of property insurance in different scenarios, providing useful references and suggestions to improve the underwriting decisions and risk management capabilities of insurance companies [1].

2. RISK EVALUATION MODEL (URA MODEL)

2.1. URA Modeling

We modeled URA based on the projected number of claims paid by insurance companies and the average amount residents could spend for such insurance.

First, we define the expected annual per capita benefit. The expected annual per capita benefit is the average per capita property damage expected to result from an extreme weather event over a given period of time (usually one year). It can be expressed as:

$$E(r) = \frac{\sum_{i=1}^n P_i L_i}{N} \quad (1)$$

Where $E(r)$ refers to the expected annual payout per capita, $P(i)$ refers to the annual probability of the occurrence of the i th extreme weather, $L(i)$ is the average amount of property damage caused by the i th extreme weather, and n refers to the number of types of extreme weather events [2].

Next, we define gross premium G . Gross premium G consists of net premium P and surcharges. Net premiums are determined based on risk assessment, while surcharges usually include management fees, marketing fees, and profits. We use a fixed ratio approach to establish the relationship between gross and net premiums. The relationship between gross and net premiums can be expressed as:

$$G = \frac{P}{1-e} \quad (2)$$

$$P = E(r) \quad (3)$$

e refers to the ratio of additional premiums to total premiums, generally taken as 40%. We assume that all insurance periods are for one year, and P refers to the annual net premium.

Finally, we conduct risk assessment. It is assumed that the average annual income of the residents of a region is denoted as symbol I . In general, the proportion of investment in consumer insurance ranges from 0.8% to 1% of the average annual income. Based on a combination of specific risk assessments and income levels of residents in multiple regions, we find a balance between providing residents with adequate risk protection and avoiding excessive burdens, i.e., 0.0084. Therefore, we take the investment in property to be 0.84 percent of average annual income.

If $G > 0.84\%I$, the insurer is not recommended to cover the area.

If $G \leq 0.84\%I$, the insurer is recommended to cover the area.

2.2. Model Demo

In order to simplify the number of calculations, we restrict the extreme weather in Japan to include only earthquakes, typhoons and storm surges. Turkey is located in the seismic zone between the Eurasian and African plates, and earthquakes are frequent. In addition, Turkey is located in a high mountainous plateau with complex topography and concentrated rainfall, and is affected by floods and mudslides every year. Especially in spring and summer, the melting of snow and glaciers in mountainous areas, together with heavy rainfall and lightning, tend to cause floods and mudslide disasters. In addition, Turkey is geologically complex, mountainous and rocky, and landslides and mudslides are common extremes. Combining the data, earthquakes, floods, mudslides and landslides are the most dominant weather extremes that Turkey faces. Therefore, in order to simplify the amount

of calculations, we limit the weather extremes occurring in Turkey to earthquakes, floods, mudslides and landslides.

By using the number of extreme weather occurrences in Japan and Turkey in the 30 -year period 1992-2022, we counted the percentage of extreme weather types in the two countries (Figure 1).

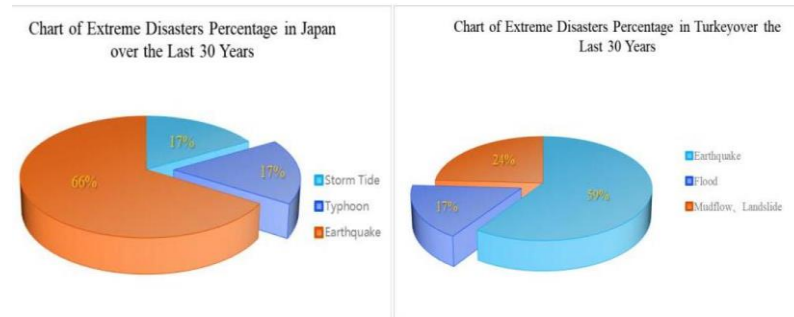


Figure 1. Percentage of extreme weather types in Japan and Turkey, 1992-2022

Using the total amount of property damage from 30 years of extreme weather, we derived the average annual property damage for the two countries (Figure 2).

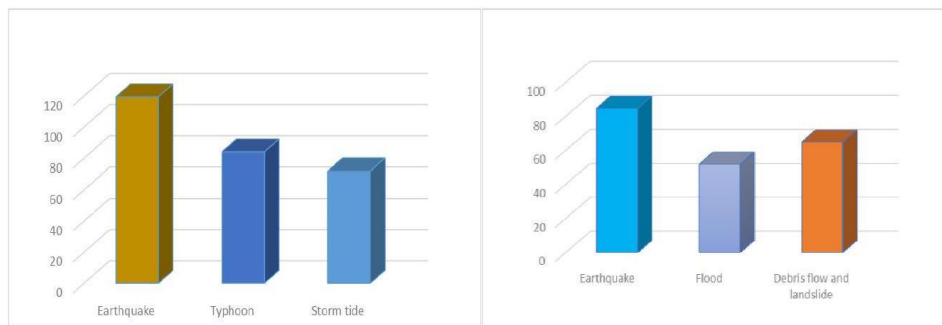


Figure 2. Average annual property damage caused by extreme weather in Japan and Turkey, 1992-2022

Combining the probability of occurrence of extreme weather in Japan and Turkey from 1992 to 2022 we predicted the probability of occurrence of each extreme event in the five-year period 2023-2028 by using MATLAB linear regression prediction method (as shown in Figure 3) [3].

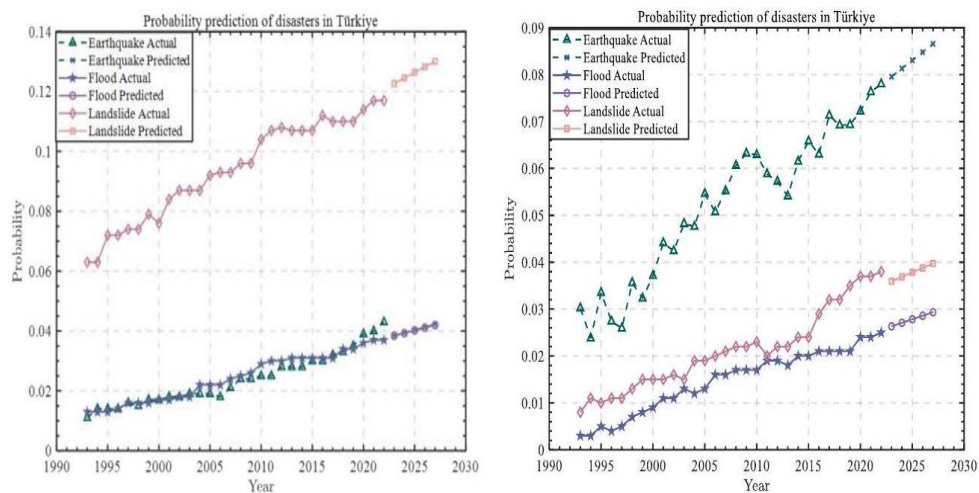


Figure 3. Probability of occurrence of each extreme event in the next five years

By using the per capita income of the residents of Japan and Turkey from 2013 to 2022, we forecast the per capita income of the residents of Japan and Turkey in the five-year period from 2023 to 2028 by using MATLAB linear regression forecasting method (as shown in Figure 4).

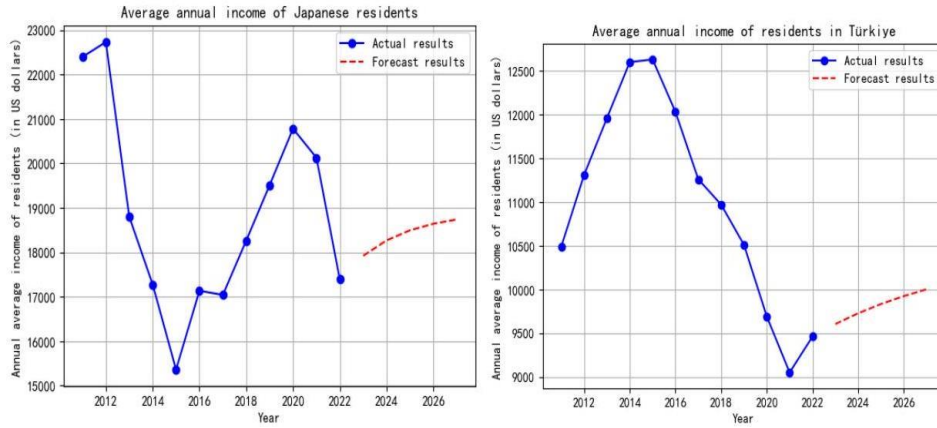


Figure 4. Forecasts of per capita income of residents of Japan and Turkey in the next five years

Through the above demonstration for two regions, Japan and Turkey, we show the applicability and flexibility of the URA model. The model can select the most representative and threatening extreme weather events to the region as the study object according to the characteristics and actual situation of different regions, in order to provide a decision-making basis to insurance companies. The model demonstrates its applicability and flexibility by selecting appropriate extreme weather events in different regions for study, providing insurance companies with a strong decision-making basis to deal with extreme weather risks in specific regions.

2.3. Model Improvement - Fuzzy Risk Assessment Model

2.3.1. Reasons for improvement

Although the URA model can select the most representative and threatening extreme weather events based on different regions to provide a decision-making basis for insurance companies, there are some limitations in considering only the factors of a single extreme weather event. Therefore, we introduced more reference variables to try to further increase the accuracy and usefulness of decision-making in a small area.

2.3.2. Model building

Prior to modeling, a Hazard Risk Assessment System (HRA) diagram was created to facilitate the assessment of extreme weather risks (e.g., Figure 5)

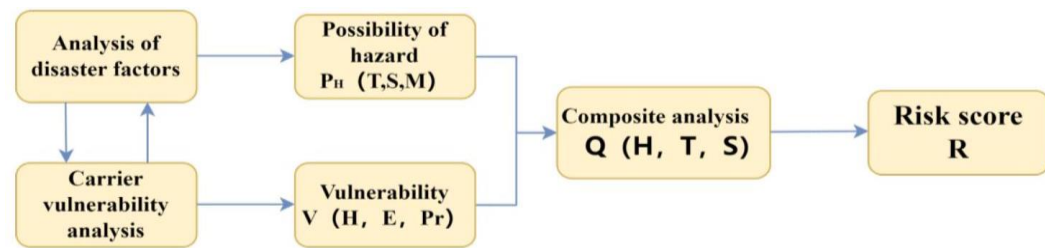


Figure 5. Disaster risk assessment system map

For the same region, it may be jointly affected by multiple weather extremes. Therefore, to ensure a more accurate decision model, we propose the concept of overall value at risk. Its calculation formula is as follows:

$$C = \sum_{i=1}^m W_i R_i \quad (4)$$

$$W_i = \frac{p_i}{\sum_{i=1}^m p_i} \quad (5)$$

Where C is the overall risk value of the area under decision; m refers to the category of extreme weather (volcanoes, earthquakes, typhoons, etc.) in the area under decision; W_i means the weight of the impact of each category of extreme weather on the total risk score; R_i means the risk value of each category of extreme weather; p_i means the probability of occurrence of each type of extreme weather.

2.3.3. Calculating the risk of extreme weather events

$P_H(T, S, M)$ where T refers to time, in the modeling assumptions we have a one-year insurance period, S refers to the location where the decision is made to insure; and M refers to the intensity of the occurrence of the catastrophic event (intensity of the causative factor).

In $V(H, E, Pr)$, V refers to potential vulnerability, H refers to the affected population, E refers to the level of economic development of the region, and Pr refers to the affected area of the real estate sector.

Considering that disaster analysis should be considered from the perspectives of time, space, and intensity of the causative factors, the stability of the affected environment, and the vulnerability of the disaster-bearing body. In the following, we will conduct a fuzzy risk assessment from these aspects.

For the disaster-affected factors, to make the model more tractable. We consider extreme weather events as an overall disaster-causing factor. This allows for a more intuitive measure of the impact of extreme weather events on the affected area, without the need to consider the impact of each detailed factor in too much detail.

We analyze this by using the International Disaster Scale (IDS), which is derived from the data $(p_1^y, p_2^y \dots p_m^y)$ that p_i^y denotes the probability of an extreme weather event of level i within the decision area.

Calculation of the intensity of the causative factor.

$$\begin{aligned} P_i^y &= \sum_{j=1}^m P_j^y \\ M &= N \cdot P_i^y \end{aligned} \quad (6)$$

N refers to the average annual number of extreme weather occurrences; P_i^y denotes the probability value of exceeding level i . P_H is a function of time T , region S , and the intensity of the disaster-causing factor M . Through the above equation, we can calculate the intensity of the disaster-causing factor in the decision-making region P_H .

The intensity of the causal factor is normalized as follows.

$$R_y = 0.5 + 0.5 \cdot \frac{P_H - P_{H_{\min}}}{P_{H_{\max}} - P_{H_{\min}}} \quad (7)$$

Where R_y is the risk value of the causal factor.

2.3.4. Vulnerability modeling

By dividing the affected area of a certain place into i sub-areas at a certain time, the following data are counted for each of the i sub-areas.

$$PPV = 0.5 + 0.5 \cdot \frac{\rho - \rho_{\min}}{\rho_{\max} - \rho_{\min}} \quad (8)$$

PPV means Potential Population Vulnerability Value, ρ means the actual population density of the decision area, ρ_{\max} means the maximum value of population density in the decision area, ρ_{\min} Minimum value of population density in the decision area.

Similarly, we obtain the potential vulnerability value *EPV* for economic development and *BPV* for real estate in the decision-making region.

In the weighting analysis, through the data analysis of the disaster, the three variables of population density, the level of regional economic development and the affected area of the real estate industry and the value of economic losses can be analyzed with partial correlation coefficients, and then through the partial correlation coefficients to get the weights W_H, W_E, W_{Pr} , so as to get the value of the comprehensive vulnerability risk. R_v The risk value of comprehensive vulnerability is obtained.

$$R_v = W_H \cdot PPV + W_E \cdot EPV + W_{Pr} \cdot BPV \quad (9)$$

Considering the above factors, through the statistical analysis of the disaster data and combining the experience of experts, we assigned the weights of 0.7,0.3 to the risk values of the disaster-causing factor and the potential vulnerability, respectively, and finally we arrived at the comprehensive risk value.

$$R = 0.7R_y + 0.3R_v \quad (10)$$

Where R_y refers to the risk value of the causative factor, R_v means the value of risk of potential vulnerability.

2.3.5. Decision criteria

When $R \geq 0.75$, it is recommended that insurers underwrite the area; conversely when $R < 0.75$, it is not recommended that insurers underwrite the decision area.

3. URA MODEL APPLICATION

3.1. URA Evaluation Model Indicators

In response to future real estate development that is more sustainable and adaptable, our URA model will incorporate the following five metrics and factors to analyze where, how, and how feasible certain sites can be built to provide a strong case for building better neighborhoods and properties.

- 1) Population Density
- 2) Average Temperature
- 3) Annual Rainfall
- 4) Transportation accessibility
- 5) GDP Level

3.2. URA Model Application

In this modeling application, we use Australia as an example (the population density heat map is shown in Figure 6) to assess where, how and how feasible it can be built in MATLAB.

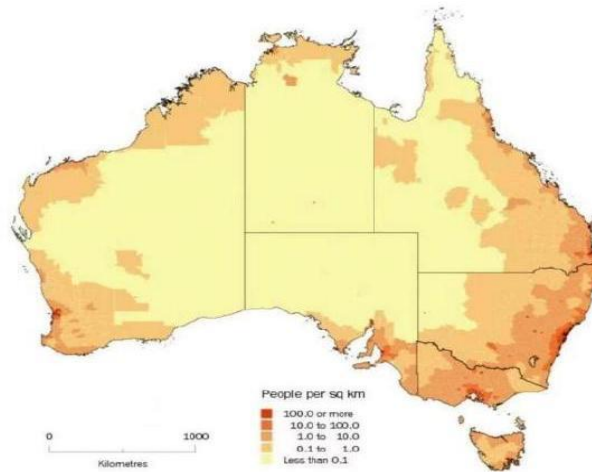


Figure 6. Heat map of population density in Australia

The K-means clustering algorithm was used to classify the areas on the map into different categories or clusters to better understand the spatial distribution and population density on the map; also, for high-dimensional geospatial datasets, the data was reduced to a lower dimensional space for better visualisation and analysis of the data. Finally, each location was scored using these metrics to identify more appropriate real estate sites [4].

In the modelling process, we used various GIS functions and tools provided by MATLAB. In order to assess the impact of different factors such as population density, average temperature, annual precipitation, etc. on the site selection more comprehensively and accurately, and to give appropriate weights to the different factors to improve the quality of decision-making and reduce subjective bias, we also used weighted averaging and multifactorial evaluation to make a comprehensive assessment of the potential sites [5]. Through these analyses, we derived scores and rankings for each location and visualised them in the form of scatter plots. Meanwhile, based on the extreme weather data of the past few years, we analysed the risk level under extreme weather using the URA model and used this as a constraint to generate a scatter plot for the assessment of real estate development under the insurance model (e.g. Figure 7).

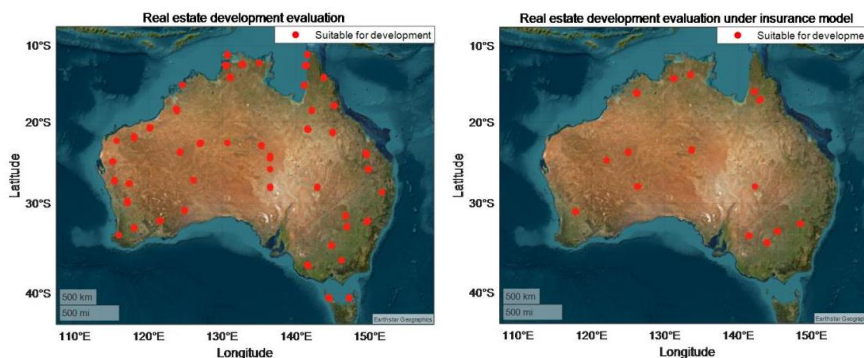


Figure 7. Scatterplot of real estate development evaluation before and after constraints

4. DECISION MAKING MODEL (DPB MODEL)

4.1. Decision model evaluation indicators

According to our URA model, there are some areas where coverage is not recommended. However, there are some areas where buildings are worthy of coverage given their cultural, historical, economic or community significance.

4.2. DPB model Building

4.2.1. Determining Building Indicator Weights - EWM

In order to process the obtained data, we try to evaluate the building indicators by entropy weight method (EWM). According to the definition of information entropy, the entropy value can be used to judge the dispersion degree of an indicator. The smaller the value of information entropy, the greater the degree of dispersion of the indicator, and the greater the influence or weight of the indicator on the comprehensive evaluation. Therefore, the tool of information entropy proof can be used to calculate the weight of each indicator of the building [6].

We form a matrix of the buildings under consideration with the four quantified evaluation metrics (j denotes community value, historical value, viewing value, and artistic value in that order) as follows:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{i1} & x_{i2} & x_{i3} & x_{i4} \end{pmatrix} \quad (11)$$

Since we will subsequently synthesize the TOPSIS evaluation method for decision making, we need multiple references as samples. Therefore, the process assumes that we consider i types of buildings, and then calculate the entropy and weights of the indicators by standardizing the evaluation matrix through Eqs. (13) and (14).

$$e_j = \frac{\sum_{i=1}^n P_{ij} \cdot \ln(P_{ij})}{\ln(n)} \quad (12)$$

$$W_j = \frac{d_j}{\sum_{j=1}^4 d_j} \quad (13)$$

Among them. $d_j = 1 - e_j$

The weight W_j of each indicator (successively protection value, historical value, viewing value and artistic value) is 0.0771,0.5207,0.3933,0.0089.

4.2.2. Decision Criteria Establishment-TOPSIS Method

Steps are as follows:

Step 1: Data acquisition.

Step 2: Evaluate the value of indicators:

Constructing the initial matrix

We find the normalization matrix with the help of Eqs. (13) and (14):

$$Z = \begin{pmatrix} 0.6255 & 0.9744 & 0.1801 & 0.4666 \\ 0.2085 & 0.1754 & 0.1561 & 0.5249 \\ 0.6255 & 0.0780 & 0.9604 & 0.5832 \\ 0.4170 & 0.1169 & 0.1441 & 0.4082 \end{pmatrix} \quad (14)$$

We combine the normalization matrix Z obtained in the above process with the weights of the indicators W_j by combining them according to equation (15) to obtain a new matrix A .

$$A = W_j \cdot Z \quad (15)$$

$$A = \begin{pmatrix} 0.0482 & 0.5074 & 0.0708 & 0.0041 \\ 0.0161 & 0.0913 & 0.0614 & 0.0047 \\ 0.0482 & 0.0406 & 0.3777 & 0.0052 \\ 0.0322 & 0.0609 & 0.0567 & 0.0036 \end{pmatrix} \quad (16)$$

Step 3: Determine the positive and negative ideal solutions.

When the assessed building is closer to the positive ideal solution, it indicates that the historic building is more important in the community's value significance and needs to be emphasized for preservation; when the assessed object is closer to the negative ideal solution, it indicates that the historic building is not as important in the community's value significance, and relocation for preservation can be considered. The specific indicators are as follows:

$$\begin{aligned} A^+ &= (\max\{a_{11}, \dots, a_{n1}\}, \dots, \max\{a_{14}, \dots, a_{n4}\}) \\ &= (0.0482 \quad 0.5074 \quad 0.3777 \quad 0.0052) \\ A^- &= (\min\{a_{11}, \dots, a_{n1}\}, \dots, \min\{a_{14}, \dots, a_{n4}\}) \\ &= (0.0161 \quad 0.0406 \quad 0.0567 \quad 0.0036) \end{aligned} \quad (17)$$

Calculate the distance of the i th ($i > 4$) protected building from the ideal solution.

Calculation of the distance to the maximum value

$$D^+ = \sqrt{\sum_{j=1}^4 (A_j^+ - a_{ij})^2} \quad (18)$$

Calculation of the distance from the minimum value

$$D^- = \sqrt{\sum_{j=1}^4 (A_j^- - a_{ij})^2} \quad (19)$$

Calculate the proximity of the i th ($i > 4$) building to the negative ideal solution. Calculate the score:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad 0 \leq S_i \leq 1 \quad (20)$$

4.3. Decision criteria

Firstly, when $S_i < 0.05$ the building is very close to the negative ideal solution (i.e., the safest position without external forces), there may be a danger because any slight disturbance may cause the building to collapse or be damaged. Therefore, the building needs to be relocated for protection, i.e., the building is moved to a safer location to ensure its stability and safety.

Secondly, when $0.05 \leq S_i \leq 0.5$ When the distance between the building object and the negative ideal solution is large, certain restrictive conservation measures are required to preserve its original character and value, while appropriate maintenance and renewal work is carried out to ensure its sustainability.

Finally, when $S_i > 0.5$ When the distance between the building and the negative ideal solution is farther, but still need to take conservation measures to protect its historical value and cultural heritage,

including strengthening the repair and maintenance work, rational utilization and development, and other normal conservation measures, in order to ensure its longterm preservation and heritage.

Therefore, we can get the criteria for decision making as follows: migration protection below 0.05, limiting protection from 0.05 to 0.5, and normal protection above 0.5. The equation is expressed as follows:

$$\begin{cases} S_i < 0.05 \\ 0.05 \leq S_i \leq 0.5 \\ S_i > 0.5 \end{cases} \quad (21)$$

5. CONCLUSIONS

Firstly, the selection of underwriting methods and coverage strategies for property insurance is crucial and requires comprehensive consideration of a variety of factors, including population density, economic development, and topographical data. Second, prediction and assessment using models such as ARIMA, URA and DPB can effectively guide insurers' underwriting decisions and improve their risk management and response capabilities. Finally, the methods and models in this study can provide more scientific and comprehensive decision support for the property insurance industry in the face of extreme weather and natural disasters, promoting the sustainable development of the industry and safeguarding the stable operation of the real estate industry. In future research, the model and methodology can be further improved and combined with more empirical data and cases to enhance the practicality and operability of the study.

REFERENCES

- [1] He Zijun. Exploration of cost and expense control in property insurance companies [J]. Finance and Economics, 2024, (11):60-62.DOI:10.19887/j.cnki.cn11-4098/f.2024.11.024.
- [2] Chen Chen, Feng Jiankui, Duan Jinping. How to build resilient cities when extreme weather is frequent? [N]. China Natural Resources News,2024-05-13(003). DOI: 10. 28291/ n. cnki. ngtzy. 2024.001083.
- [3] Chen Lusheng, Chen Qixiang, Chen Yulun, et al. Study on linear regression model for weight prediction of pig carcasses for online grading [J/OL]. Journal of Nanjing Agricultural University, 1-7[2024-05-28]. <http://kns.cnki.net/kcms/detail/32.1148.S.20240203.1935.002.html>.
- [4] GUAN Zing, YIN Yongqiang, ZHANG Xiaoxiang, et al. Assessment of flash flood vulnerability in small watersheds based on K-means clustering and integrated learning algorithm [J/OL]. Journal of Applied Science, 1-17[2024-05-28]. <http://kns.cnki.net/kcms/detail/31.1404.N.20240507.1417.002.html>.
- [5] Li Yanfang. Design of vector control system for permanent magnet synchronous motor based on MATLAB [J]. Value Engineering, 2024, 43(12):97-99.
- [6] LUO Qian, LI Yongmei, WANG Tenghua, et al. Constructing an evaluation system of rational medication use indicators in clinical departments based on the improved entropy weight method combined with TOPSIS method [J]. Clinical rational drug use, 2024, 17(15):170-172+ 177. DOI: 10. 15887/j.cnki.13-1389/r.2024.15.049.