

Research on extreme weather insurance and real estate decision model based on AHP-EWM-CVM and K-Means clustering technology

Aojing Han ^{1,*}, Hongwei Lin ²

¹ School of Public health, Wuhan University, Wuhan, China, 430071

² School of Electronic Information, Wuhan University, Wuhan, China, 430072

* Corresponding Author Email: aj.han@whu.edu.cn

Abstract. Aiming at the challenge of extreme weather to insurance industries and real estate industries, this paper constructs an insurance model and a real estate decision model. The insurance model uses analytic hierarchy Process (AHP), entropy weight method (EWM), coefficient of variation method (CVM) and K-Means clustering algorithm to evaluate the risk level in different regions and predict the values of risk indicators up to the year 2035, and then calculate future insurance premiums. At the same time, multiple constraints are used to determine the future risk safety interval, which provides the basis for insurance companies to make decisions. In the case application, this paper finds that China is currently medium risk region and will be a medium risk region in the future, and obtains the time point when insurance companies can underwrite policies in this region. In addition, this paper also constructs a real estate decision model, which comprehensively considers three factors: risk, development potential and risk cost, and provides decision making basis for real estate developers through GE matrix. In the case application, this paper predicts that in the next nearly ten years, China is a medium willingness and medium risk area, suggesting that real estate developers should be cautious in deciding whether to develop. This study provides decision support for insurance industries and real estate industries to deal with extreme weather, which is of theoretical and practical significance.

Keywords: Insurance model, Analytic hierarchy process, Coefficient of variation method, K-Means clustering algorithm.

1. Introduction

1.1. Problem background

Climate change has become one of the major issues affecting the development of human society, among which extreme climate events have significant implications for the ecological environment, social life, and human safety [1]. For instance, the super typhoon "Haiyan" that made landfall in the Philippines in 2013 affected 11.236 million people, caused 6,009 deaths, and damaged 1.181 million houses, resulting in heavy losses. Since the 1980s, global warming has become the most important climate background, greatly affecting the regional hydrothermal structure, thereby increasing the frequency and intensity of extreme climate events. Extreme climate events have significant suddenness and unpredictability, and their occurrence will undoubtedly have a considerable negative impact on regional sustainable development.

Nowadays, "disaster management" of extreme weather events has shifted from "disaster management" prior to the early 1990s to "disaster risk management", which now tends to dominate discussions and actions. Disaster risk management encompasses policies and strategies to reduce disaster risk, strengthen resilience and reduce disaster losses. The insurance industry, as the world's largest aggregator of risk, has been identified as an integral part of disaster risk management for extreme weather events. Climate risk insurance integrates financial disaster reduction strategies to address loss and damage caused by climate hazards. However, the increased frequency and severity of extreme weather events also pose a significant threat to the insurance industry [2]. As more claims need to be paid, the cost of insurance rises, including the cost of maintaining and repairing damaged

property. Insurers must reassess risks to reflect new weather phenomena, which can lead to higher premiums in high-risk areas or, in extreme cases, insurers choosing not to provide coverage. The reinsurance market is also under pressure as it must cope with increasing claims. The sustainability and profitability of insurers are affected, as they need to balance the need to provide sufficient coverage to attract customers while ensuring that they do not go out of business as a result of paying too many claims.

Therefore, the property and casualty insurance industry must actively look for innovative ways to address climate change challenges, including developing weather data analysis tools, establishing risk-sharing programs, and promoting sustainable development practices to ensure sustainability and long-term health. This is a major and complex issue that needs to be a significant priority now and for a long time to come and one that requires the joint efforts by all sectors of society to address [3].

1.2. Restatement of the problem

(1) Develop a model for insurance companies to help them determine the areas and conditions in which they should write policies to ensure future solvency and the long-term health of the company, and apply the model in the area experiencing extreme weather events to demonstrate its effectiveness.

(2) By adjusting and optimizing insurance models, we help real estate developers choose the right place and way to develop their properties.

2. Insurance model

In the insurance model, we define three smaller models, which are Underwriting Policy Risk Assessment (UPRA) model, Insurance Premium Pricing (IPP) model, and Insurance Company Decision (ICD) model. The UPRA model is used to evaluate the risk of an area, the IPP model is used to predict future insurance premiums for an area, and the ICD model is used to formulate underwriting conditions for insurance company decision making. We will analyze the construction and role of each model in detail.

2.1. Underwriting policy risk assessment (UPRA) model

2.1.1. Establishment of evaluation indicators

Based on the relevant literature, we summarize three factors that influence the level of risk in the region: hazard factors, exposure factors and risk resistance factors, and use them as primary indicators in the evaluation system [4-6]. After further reflection and generalization, we concretized the indicators and refined eight specific indicators as the secondary indicators of the evaluation system. The specific risk assessment indicator system is shown in Table 1.

Table 1. Selected indicators

Primary indicators	Secondary indicators	description
Hazard (H)	Highest temperature (HT)	Annual Maximum Temperature(°C)
	Lowest temperature (LT)	Annual Minimum Temperature(°C)
	Monthly precipitation (MP)	Average monthly precipitation during the year
Exposure (E)	Population density (PD)	Population size per unit area
Resistance (R)	Life expectancy (LE)	Expected duration of survival of an individual's life
	Educational level (EL)	Proportion of GDP spent on education(%)
	GDP	Gross Domestic Product
	Disaster response capacity (DRC)	Score for the implementation of a national disaster risk management strategy

2.1.2. Determination of the weights for indicators

There are many methods to determine the weight of indicators. To make our model more accurate, we decide to use the combination weighting method to calculate the weight of all indicators. Our

combination weighting method combines Analytic Hierarchy Process (AHP) in the subjective weighting method and Entropy Weight Method (EWM) and the Coefficient of Variation Method (CVM) in the objective weighting method [7]. Because the AHP judgment is more subjective, it is easy to change by the subjective influence of the decision maker. At the same time, because of the high sensitivity of the data, it may cause errors due to the data itself. Therefore, our combination weighting method synthesizes these methods to help us reduce errors and improve accuracy.

➤ Analytic Hierarchy Process

We use the Analytic Hierarchy Process (AHP) method to determine the weights of the three primary indicators.

First, we construct judgment matrix for the primary indicators.

$$A = (a_{ij})_{n \times n} \quad (1)$$

Where a_{ij} represents the importance of i th indicator relative to j th indicator, while n represent the number of indicators.

Then, we perform the consistency test and the results show that the consistency of the matrix is acceptable. We use arithmetic mean, geometric mean and eigenvalue methods respectively to calculate the weights of each indicator and take the average value as the final weights. The weights of each indicator calculated by AHP method are denoted as w_{j1} .

➤ Entropy Weight Method

We use the Entropy Weight Method (EWM) method to determine the weights of the eight secondary indicators, which is a method that uses the law of data itself to assign weight objectively.

First, we have collected data on indicators for the major countries in the world (30 countries) to form the original matrix. Due to the different types of indicators, we need to standardize the original matrix to obtain standardized matrix X. Then, we normalize the matrix X to obtain the matrix Y. The specific calculation formulas are as follows.

Positive indicators:

$$x_{ij} = \frac{x_{ij} - \min\{x_{1j}, \dots, x_{nj}\}}{\max\{x_{ij}, \dots, x_{nj}\} - \min\{x_{1j} - x_{nj}\}} \quad (2)$$

Negative indicators:

$$x_{ij} = \frac{\max\{x_{1j}, \dots, x_{nj}\} - x_{ij}}{\max\{x_{ij}, \dots, x_{nj}\} - \min\{x_{1j} - x_{nj}\}} \quad (3)$$

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \quad (4)$$

Where x_{ij} represents the j th indicator of the i th country, p_{ij} is the weight of j th indicator of the i th country, and n is the number of countries.

Next, we calculate the information entropy of each indicator.

$$e_j = - \frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln(p_{ij}) \quad (5)$$

Where, e_j is the information entropy of the j th indicator, and n is the number of countries.

Finally, we calculate the information utility value and normalize to get the weight, which are denoted as w_{j2} .

$$d_j = 1 - e_j$$

$$w_{j2} = \frac{d_j}{\sum_{j=1}^m d_j} \quad (6)$$

Where d_j is the information utility value of the j th indicator, w_{j2} is the corresponding weight of j th indicator, and m is the number of indicators.

➤ Coefficient of Variation Method

Furthermore, we apply Coefficient of Variation Method (CVM) to weight eight secondary indicators.

We use matrix X to calculate the mean and standard deviation of each indicator for the 30 countries to derive the coefficient of variation. Then, the coefficients of variation are normalized to get the weights of indicators, which are denoted as w_{j3} .

$$CV_j = \frac{\sigma_j}{\mu_j} \quad (7)$$

$$w_{j3} = \frac{CV_j}{\sum_{j=1}^m CV_j} \quad (8)$$

➤ Combination weight

By the above three method, we calculate the weights of the three primary indicators (w_{j1}) and eight secondary indicators (w_{j2} , w_{j3}). For the secondary indicators, we choose the average value of w_{j2} and w_{j3} as the final weight for each indicator to minimize the error.

2.1.3. The establishment of UPRA model

In this model, we introduce Risk Score (RS) to quantitatively describe the risk level of underwriting policies. In addition, we divided the values of RS into three intervals by K-means clustering algorithm. In other words, the underwriting policy risk level is divided into three levels.

➤ Calculation of RS

First, we construct the formulas to calculate three primary indicators respectively.

$$H = w_{HT} \cdot HT + w_{LT} \cdot LT + w_{MP} \cdot MP$$

$$E = w_{PD} \cdot PD \quad (9)$$

$$R = w_{LE} \cdot LE + w_{EL} \cdot EL + w_{GDP} \cdot GDP + w_{DRC} \cdot DRC$$

Then, we construct the formula for RS. The weights corresponding to three primary indicators are regarded as the degree of their influence on RS. The greater the weight, the greater the influence of the corresponding indicator on RS. Accordingly, we construct the RS calculation formula reasonably.

➤ Determination of underwriting policy risk level

$$RS = w_H \cdot H + w_E \cdot E + w_R \cdot R \quad (10)$$

By using K-means clustering algorithm, RS values 30 countries are divided into three categories, thus classifying the underwriting policy risk into three classes, defined as: high risk, medium risk, and low risk. The data are discrete, but we need continuous intervals in the RS division. Therefore, we make the following minor adjustments according to the above categories (Table 2).

Table 2. Level of underwriting policy risk

Underwriting policy risk level	Low risk	Medium risk	High risk
RS	≤ 0.30	$0.30 < RS \leq 0.42$	$0.42 <$

● **High risk:** Avoiding the concentration of underwriting in high risk areas can reduce a company's exposure to catastrophic losses. If an insurer underwrites too much in high risk areas, natural disasters

or other catastrophic events could result in a large number of claims, which could have a serious impact on the company's financial position.

- **Medium risk:** In medium risk areas, although the risk increases, so do the number of insured people. In medium risk areas, it is easier for insurers to make accurate risk assessments. This means that they can better predict the potential losses that policyholders are likely to suffer and price insurance products accordingly. This can improve a company's profitability.

- **Low risk:** Low risk means a low probability of claims, which helps insurers maintain a more stable financial position and reduces the risk of adverse claims.

In short, we recommend that insurers underwrite in medium and low risk areas.

2.2. Insurance premium pricing (IPP) model

Since an increase in risk in an area leads to higher insurance premiums, we assume the initial insurance premium is IP_0 , α is the risk-related premium adjustment factor. Then the premium IP_t in year t is:

$$IP_t = IP_0 \cdot (1 + \alpha_t)$$
$$\alpha_t = \frac{RS_t}{RS_0} \quad (11)$$

Since many indicators used to evaluate RS change over time, we respectively employ Grey Prediction Model, Autoregressive Integrated Moving Average (ARIMA) model, and Logistic Model to forecast the values of the indicators up to the year 2035 in order to assess future risk and thus price insurance premiums.

2.3. Insurance company decision (ICD) model

We assume that the insurance company's compensation expense is CE , and use Grey Prediction Model to predict CE_t in year t .

With the increase in extreme weather events and regional risk, insurers face a double risk in the future. On the one hand, if insurers underwrite policies that are too risky, they could result in paying too many claims. On the other hand, if insurers do not underwrite in these risky areas, they may go out of business with too few customers. In order to provide insurers with the ability to pay future claims while ensuring the long-term health of the insurer, we propose an Insurance company decision model to ensure that insurance companies do not operate at a loss.

$$IP_t - CE_t \geq 0 \quad (12)$$

We also consider that citizens' insurance premiums are based on their income level, which means that they use a portion of their income to purchase insurance. In order to ensure that the normal life of citizens is not affected so that they can insure themselves over a period of time, we set, in accordance with the "double-ten" principle, that the premiums paid by an individual should not exceed 10% of his or her annual income.

$$IP_t < I_t \times 10\% \quad (13)$$

I_t is the annual income in year t , forecasting by ARIMA.

Combined equations (11) (12) (13), ICD model is demonstrated as below.

$$IP_0 \times \left(1 + \frac{RS_t}{RS_0}\right) - CE_t \geq 0$$
$$IP_0 \left(1 + \frac{RS_t}{RS_0}\right) < I \times 10\% \quad (14)$$

2.4. The application of insurance model

2.4.1. Selection of the research area

We choose China, where extreme weather events occur, for our study. We collect data on relevant indicators and apply established insurance model to determine the risk level, predict future insurance premiums, and formulate underwriting conditions to inform insurance company decisions. Temperature and precipitation data are from the CRU TS dataset (<https://crudata.uea.ac.uk/cru/data/hrg/>), CHINA STATISTICAL YEARBOOK ON ENVIRONMENT and China Meteorological Administration, while other indicator data are taken from the website Our World in Data (<https://ourworldindata.org/>).

2.4.2. Risk level determination based on UPRA model

We bring the relevant indicator data from China into the UPRA model, and get an RS value of 0.414, which is classified as medium risk.

2.4.3. Future insurance premium forecast based on IPP model

First, we forecast the eight indicators for China separately. Then, we take the predicted values of the indicators into the UPRA model to calculate the future RS, and take the RS values into the IPP model to derive the future premium pricing in China.

The results indicate that future RS value is slowly decreasing and China will change from a high-risk area to a medium risk area (Figure 1). As a result of lower risk, premiums are also on a downward trend (Figure 2).

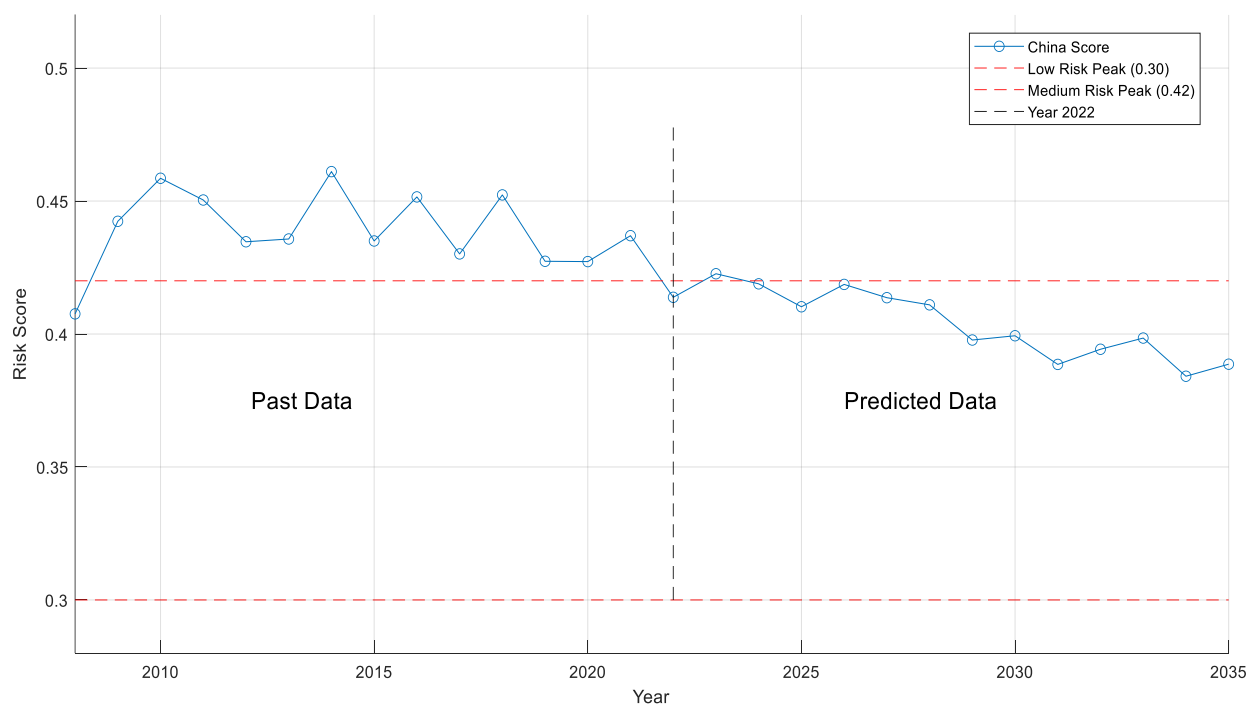


Figure 1. Future risk score in China

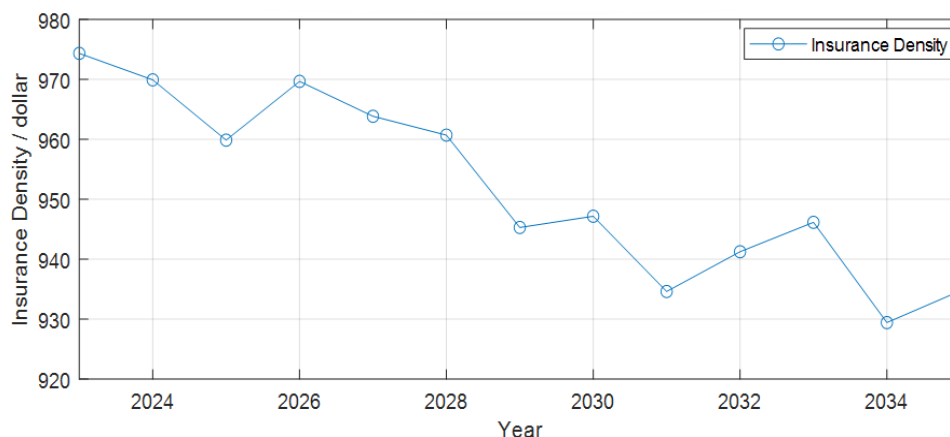


Figure 2. Future insurance premiums in China

2.4.4. Underwriting condition setting based on ICD model

Through calculations, we obtain a safe interval of RS value for China up to 2035. In other words, if the RS value predicted by the insurance company falls within this interval, the insurance company can underwrite policies and still make a profit. Based on the results, we believe that in China, insurers should not underwrite until 2032 and can underwrite in 2032 and beyond. In addition, the safety intervals become larger as time increases, which also suggests a decrease in risk (Figure 3).

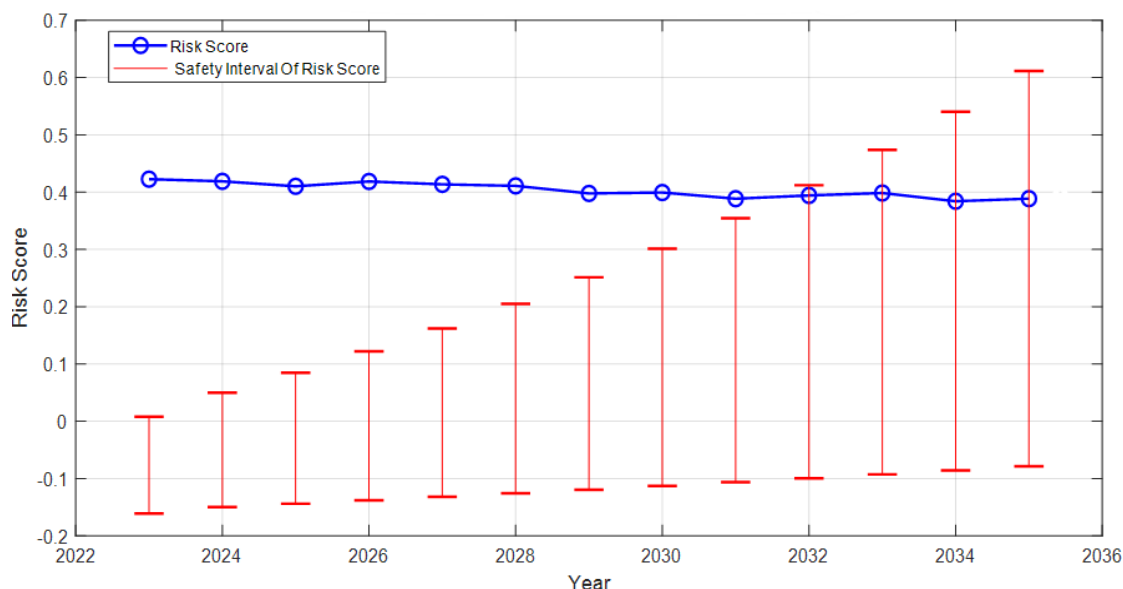


Figure 3. Safety interval for future risk in China

2.4.5. What can property owners do?

Although China is medium risk area, insurers will not underwrite policies when the predicted RS values are outside the safe intervals. In order to change an insurer's underwriting decision, property owners can take these actions to reduce the predicted RS values.

- Awareness-raising: Raise awareness of extreme weather events and knowledge of self-help and first aid. Develop disaster response plans and conduct regular drills to improve resilience.
- Preventive measures: Install waterproofing to reduce the risk of flooding and water damage, use fire-resistant building materials to reduce the occurrence and spread of fires, use earthquake-resistant building materials and install earthquake-resistant bracing systems to improve the seismic capacity of buildings, install wind-resistant windows and reinforce roof structures in wind-prone areas, and install disaster warning devices.
- Regular inspection and maintenance of building structures and infrastructures: Regular inspections of housing structures, including walls, roofs, and other parts of the building, are conducted to ensure that the building structures are strong and can withstand natural disasters such as

earthquakes. In addition, the electrical system, plumbing system, natural gas system, etc. are regularly inspected and maintained to ensure that the facilities and equipment are functioning properly and to reduce the occurrence of accidents.

3. Real estate decision model

3.1. Real estate willingness (REW) model

3.1.1. Establishment of evaluation Indicator

Based on the relevant literature [8-10], we summarize three factors affecting the willingness to develop real estate: risk, development potential and risk cost, and use them as the primary indicators in the evaluation system. After further reflection and generalization, we concretize the indicators and refine eleven specific indicators as the secondary indicators of the evaluation system. The specific risk evaluation indicator system is shown in Table 3.

Table 3. Selected indicators

Primary indicators	Secondary indicators	description
Hazard	Highest Temperature	Maximum daily highest temperature during the year
	Lowest Temperature	Minimum daily lowest temperature during the year
	Monthly precipitation	Average monthly precipitation during the year
Development Potential	Population density	Population size per unit area
	Educational Level	Proportion of GDP spent on education(%)
	Freshwater Resource	Percentage of freshwater
	Medical Resource	Hospital beds (per 1,000 people)
	Transport Situation	Rail lines (total route-km)
	Renewable Energy	Renewable energy consumption (% of total final energy consumption)
Risk Cost	Insurance Premium	-
	Cost of Building Risk	Costs based on different levels of risk

3.1.2. Determination of the weights for indicators

We draw on the combination weighting method used in constructing the insurance model to determine the weights of the indicators. By using AHP method, we determine the weights of the three primary indicators. By using EWM method and CVM method, and taking the average of the weights obtained by these two methods, we determine the weights of the eleven secondary indicators.

3.1.3. The establishment of REW model

In this model, we introduce Willingness Score (WS) to quantitatively describe the willingness to develop real estate. In addition, we divide the values of WS into three classes by K-means clustering algorithm. In other words, the willingness to develop real estate is divided into three levels: high willingness, medium willingness, and low willingness (Table 4).

Table 4. Level of real estate willingness

Real estate willingness level	Low willingness	Medium willingness	High willingness
WS	≤ 0.32	$0.32 < RS \leq 0.42$	$0.42 <$

3.2. Decision-making analysis based on UPRA model and REW model

In order to making the appropriate programme, we use the GE matrix (McKinsey Matrix). The relevant results are shown in Table 5.

Table 5. GE matrix of decision-making programme

		WS		
		High	Medium	Low
RS	Low	I	II	III
	Medium	IV	V	VI
	High	VII	VIII	IX

✧ Programme 1: For areas I, II and IV, real estate development can be carried out.

✧ Programme 2: For areas III, V and VII, it is cautious to decide whether or not to develop them. If developed, good water and electricity supply, transportation network, medical and educational facilities should be available nearby to maintain the sustainability and attractiveness of the construction project. At the same time, strengthen the building structure, optimise the building layout and use good building materials to improve the disaster resistance and safety of the building.

✧ Programme 3: For areas VI, VIII, and IX, real estate development is not recommended.

3.3. The application of real Estate decision model

We choose China as the research object to apply our real estate decision model. Data for indicators are mainly taken from the website Our World in Data (<https://ourworldindata.org/>), China Economic and Social Big Data Research Platform (<https://data.cnki.net/internationalData>), China Meteorological Administration, the CRU TS dataset (<https://crudata.uea.ac.uk/cru/data/hrg/>) and so on.

Initially, we forecast the future RS and WS values. The results show that the future WS generally demonstrates a downward trend, indicating a decrease in the willingness for real estate development. Except for 2026, the WS value is within the medium willingness level. Meanwhile, the RS value declines slowly, but also remains within the medium risk level (Figure 4). That is, we predict that in the next nearly 10 years, China is a medium willingness and medium risk area, and according to the GE matrix, property developers should adopt Programme 2, which means that they should be cautious in deciding whether to develop or not. If developed, it should be in the vicinity of good water and electricity supplies, transportation networks, medical and educational facilities, and so on.



Figure 4. Future risk score and willingness score in China

4. Sensitivity analysis

Since our assessment model contains many different indicators, a change in an indicator in one region should not have a significant impact on the RS value for that region. To verify the robustness of the assessment model, we adjust the values of each indicator in the UPRA model by 10%. We analyze the data for China. The results show that the magnitude of change in RS is small and that changes in individual indicators do not lead to excessive changes in RS (Figure 5).

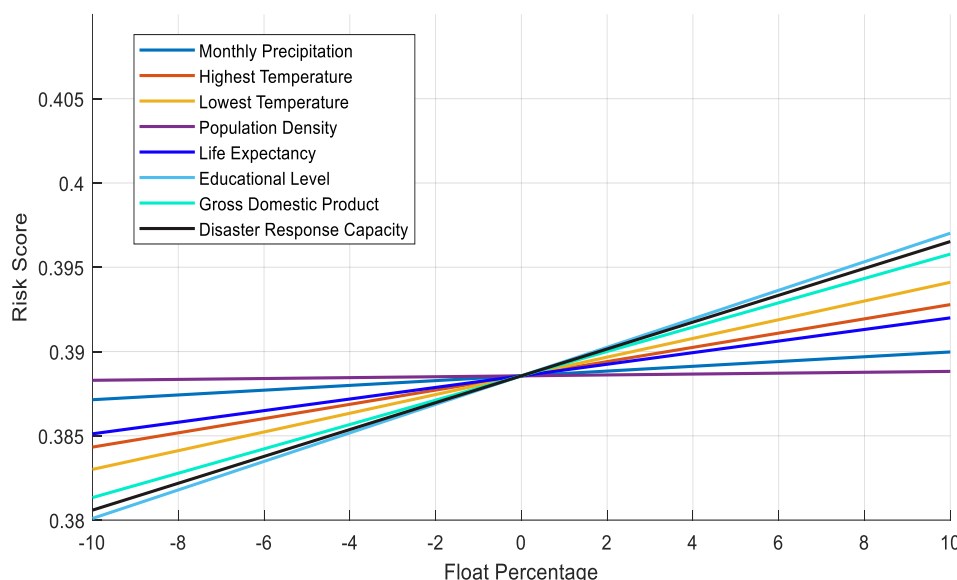


Figure 5. Sensitivity analysis to UPRA model

5. Conclusion

The insurance model and real estate decision model constructed in this study adopt various methods to assess risk or willingness. In the case study, this paper predicts that China will have lower risk in the future, which can provide a basis for insurance companies to determine the timing of underwriting policies. In addition, the study also finds that the willingness to develop real estate in China declines over the next nearly 10 years, suggesting that developers should be cautious in making decisions. The models and methods of this study have strong applicability and stability and can provide decision support for related fields, which have theoretical and practical significance.

References

- [1] Li, H.M., et al., Understanding systemic risk induced by climate change. *Advances in Climate Change Research*, 2021. 12 (3): p. 384 - 394.
- [2] Raymond, C., et al., Understanding and managing connected extreme events. *Nature Climate Change*, 2020. 10 (7): p. 611 - 621.
- [3] Kong, F. and Y. Wang, better understanding of climate catastrophe insurance in China: issues and opportunities, international insights, and directions for development. *Natural Hazards*, 2022. 114 (3): p. 2969 - 2990.
- [4] Bu, B., J. Wei, and Y.B. Ji, Study on Extreme Weather Risks in Beijing Based on Double-layer Indicators. *2nd International Conference on Applied Mathematics, Modelling, and Intelligent Computing (Cammic 2022)*, 2022. 12259.
- [5] Eltazarov, S., et al., Improving risk reduction potential of weather index insurance by spatially downscaling gridded climate data - a machine learning approach. *Big Earth Data*, 2023. 7 (4): p. 937 - 960.
- [6] Unterberger, C., How Flood Damages to Public Infrastructure Affect Municipal Budget Indicators. *Econ Disaster Clim Chang*, 2018. 2 (1): p. 5 - 20.

- [7] Sun, Y., J. Wu, and G. Li, Study on ecological impact of Saihanba Forest Farm Based on Evaluation Model. BCP Business & Management, 2022.
- [8] Anelli, D. and F. Tajani, Spatial decision support systems for effective ex-ante risk evaluation: An innovative model for improving the real estate redevelopment processes. Land Use Policy, 2023. 128: p. 106595.
- [9] Bergmann, P., et al. Enhanced Cash Flow Valuation in Real Estate Management by Integrating Innovative Materials and Risk Assessment. Sustainability, 2020. 12 (6): p. 2201.
- [10] Attoh, E., et al., Making physical climate risk assessments relevant to the financial sector - Lessons learned from real estate cases in the Netherlands. Climate Risk Management, 2022. 37: p. 19.