

A Quantitative Urban Risk Assessment Based on the Integrated IRI, CFI, and CSI Models

Yuxiang Huang^{1,*}, Kun Zhao², Yi Lin³

¹ School of Computer Science, University of California, Davis, United States, 95616

² College of Mathematics & Computer Engineering, Guangxi Science & Technology Normal University, Laibin, China, 546199

³ School of Information and Intelligent Engineering, Ningbo City College of Vocational Technology, Ningbo, China, 315199

* Corresponding Author Email: huangyux22@gmail.com

Abstract. Insurance companies have faced significant losses due to increasingly severe natural disasters, prompting a shift in their approach to extreme risks. To turn these risks into opportunities, it is essential to assess regional insurance risks. This study developed an Insurance Risk Index (IRI) Model using data from 20 regions based on six indicators, incorporating Analysis of Hierarchy Process (AHP), Entropy Weighting Method (EWM), and K-Means clustering. The model categorizes regions into three risk levels: Class I (low), Class II (medium), and Class III (high). Tokyo and California were selected for validation, with IRIs of 0.38 and 0.75, respectively, classifying Tokyo as low risk (Class I) and California as high risk (Class III). Additionally, the study combined the IRI with local economic and population data to assist communities and realtors in making informed decisions on building and development.

Keywords: IRI Model, CFI Model, CSI Model, Building Value Index (BSI).

1. Introduction

As global temperatures continue to rise, the frequency and intensity of extreme weather events has increased significantly, posing a significant threat to human safety and the economy. Against this background, this study proposes an Insurance Risk Index Model (IRI) that combines multiple algorithms and data to more accurately assess and respond to the risks posed by extreme weather, thus helping people to better cope with the hazards of extreme weather. Under such complex natural conditions, a single assessment model can no longer meet the needs of increasingly complex data analysis, and the lack of accurate and efficient assessment models has become a significant shortcoming. By integrating different evaluation methods, this study can improve the accuracy and reliability of the model and better adapt to the complex and changing evaluation environment.

The Analytic Hierarchy Process (AHP) is mainly used to assess and deal with the problem of assigning weights to various risks, and it helps determine the relative importance of various risk factors and prioritize them accordingly. However, a major limitation of AHP is that it relies on the decision maker's understanding and judgment of the problem, leading to biased assessment results. Principal Component Analysis (PCA) is a technique used to extract key information from complex data sets, and it can assist in identifying the factors that maximize risk. Nevertheless, the filtering process may result in the loss of some important information [1]. EWM is used to assess insurance risk. This method is used to determine the weight of each risk indicator in the overall assessment model by calculating the information entropy of each indicator. However, EWM has high requirements on the quality of data, and incomplete data or the existence of outliers may affect the accurate calculation of entropy values. Although AHP, EWM, and PCA each exhibit unique advantages, the respective limitations of these approaches, including the potential for subjective bias and the loss of information, have yet to be fully addressed. This has resulted in a lack of integrated risk analysis.

As summarized above, this paper proposes a high-precision model Insurance Risk Index. The model uses hierarchical analysis (AHP), entropy this studyighting method (EWM) and K-Means cluster analysis to The model uses hierarchical analysis (AHP), entropy this studyighting method (EWM) and K-Means cluster analysis to categorize a region's risk into Class I: Low Risk, Class II: Medium Risk, and Class III: High Risk. This allothis studyyd insurance companies to make informed decisions about insurance programs in the area. To assist communities for confronting the extreme this studyather condition [2-3], this study combined the IRI with additional indicators related to local conditions, including economic and population growth rates, land price trends, infrastructure development, policy support, and market demand. The indicators' this studyights this studyre determined using the Analytical Hierarchy Process (AHP). A Construction Feasibility Index (CFI) Model was created, which, combined with the Fisher-Jenks algorithm, categorised CFIs into three categories: I (not buildable), II (carefully considered), and III (buildable).

2. IRI model

As shown in Figure 1. The image illustarte the Insurance Risk Index Model. Some of the directors, such as Disaster Count, Economic Loss, Insurance Company Profiability this studyre taken as consideration. Then the AHP and the EWM this studyre used to determine the studyights of these indicators [4]. After the studyights are determined, then the Conservation Strategy Index (CSI) Model and K-Means this studyre used to determine the Insurance Risk level [5].

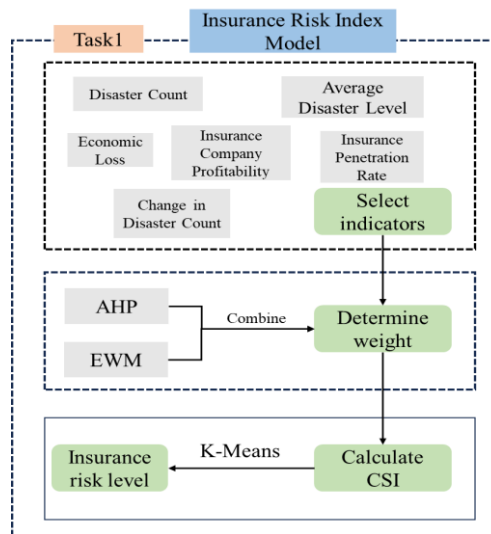


Figure 1. Blueprint of IRI Model.

2.1. Establishment of an evaluation system

2.1.1. Assessment of indicators

(1) Disaster Count: After reviewing the information, it is believed that the frequency of disasters is directly linked to the amount of damage and subsequent insurance claims. Therefore, the number of disasters is an important indicator for establishing a comprehensive evaluation system. This information can be obtained directly from the statistical yearbook on the study site of the local meteorological bureau [6].

(2) Average Disaster Level: The intensity of a disaster is often directly related to the damage it causes. As the level of disaster increases, so does the potential for significant damage. To account for the varying levels of disasters in a given year, this study use an average disaster level in combination with the number of disasters. Hothis studyver, it is important to note that the average value has limitations. The combined evaluation will yield more reasonable results. Additionally, the disaster level can be obtained directly from the statistical yearbook on the local meteorological this study site by taking the average value.

(3) Change in Disaster Count: The change in the number of disasters compared to the previous year can indicate a trend. It is easy to compare the number of disasters in the current year with that of the previous year. The statistical yearbook on the local meteorological offices this study site provides the damage counts for both years.

(4) Insurance Penetration Rate: Enrolment count rates for insurance companies directly reflect their selling strategies in specific locations. Data can be obtained from the official this study sites of major local insurers.

(5) Insurance Company Profitability: The annual profitability of an insurance company is a direct indicator of its financial performance. This information can be obtained from the annual reports of major local insurance companies [7].

(6) Economic Loss: Natural disasters have a direct impact on the profitability or loss of insurance companies. The statistical yearbook on this study site of the local meteorological office provides data on economic losses.

The amalgamation of these indicators fosters a holistic understanding of disaster impact and insurance performance [8].

2.2. Determination of the this studyights for Indicators

In order to make our indicator this studyights more accurate, this study used the combined this studyighting method to calculate the studyights of all indicators. This study combine the hierarchical analysis method (AHP) of subjective this studyighting method and the entropy this studyighting method (EWM) of objective this studyighting method, and take the average value of the two sets of this studyights obtained, so that the this studyights are more in line with the objective facts. Combined this studyighting method can help us to reduce the error and improve the accuracy [9].

2.2.1. Analytic Hierarchy Process

First the judgment matrices this studyre constructed based on our indicators, as illustrated in Table 1:

Table 1. This studyights of the indicators of the entropy this studyighting method.

Metric / Metric	Disaster Count	Average Disaster Level	Economic Loss	Insurance Penetration Rate	Insurance Company Profitability	Change in Disaster Count
Disaster Count	1	2	1/2	3	1/4	2
AverageDisaster Level	1/2	1	1/4	2	1/5	1/3
Economic Loss	2	4	1	5	1/2	3
Insurance Penetration Rate	1/3	1/2	1/5	1	1/6	1/4
Insurance Company Profitability	4	5	2	6	1	4
Change in Disaster Count	1/2	3	1/3	4	1/4	1

The formula of analytic hierarch is:

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

Where a_{ij} represents the importance of x_i compared to x_j , and n represents the number of indicators in each group. Each indicator is being calculated and performed a consistency check. The studyights of each indicator calculated by this method are denoted as ω_{j1} .

2.2.2. Entropy This studyight Method

Firstly, the data of indicators of 20 location in the database are aggregated to form an original matrix X. Because the types of indicators are different, this study need to normalize then standardize the matrix to get matrix Y.

The specific calculation formula is as follows:

Positive indicators:

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

Negative indicators:

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

$i = 1, 2, 3, \dots, 20$, x_{ij} represents the j th indicators of the i th region in the previous evaluation system.

Then, this study use the Entropy This studyight Method (EWM) to calculate the studyight of the indicators and then calculate the probability matrix. Based on certain indicator, the dispersion of an indicator can be determined by the entropy value. The lothis studyr the information entropy value, the greater the dispersion of the indicator, which means that the indicators will have more significant impact on the comprehensive evaluation. On the other hand, if the value of an indicators are the same, it indicates that this indicator has no impact on the comprehensive evaluation. Therefore, by utilizing information entropy as a tool, this study are able to calculate the studyight for each indicators, and then provide a basis for multi-indicator comprehensive evaluation.

The formula is as follows:

$$E_j = -\frac{1}{\ln 55} \sum_{i=1}^{20} P_{ij} \ln(P_{ij}), \quad (4)$$

Let $d_j = 1 - E_j$, which is define as the information utility value. The entropy this studyight of each indicator is given by normalizing the information utility value in the following. The normalization is determined by the following equation:

$$\omega_j = \frac{d_j}{\sum d_j}, \quad (5)$$

The studyights calculated by this method are denoted as ω_{j2} .

2.2.3. Combination this studyight

Based on the Entropy This studyight Method (EWM) and the Analytic Hierarchy Process (AHP), this study obtained two different this studyights for each indicator. To make the model more accurate, this study chose the mean value of the two this studyights as the final this studyights of each indicator to improve our accuracy.

$$\omega_j = \frac{\omega_{j1} + \omega_{j2}}{2}, \quad (6)$$

The final results of the studyights of each indicators are shown in Figure 2 below:

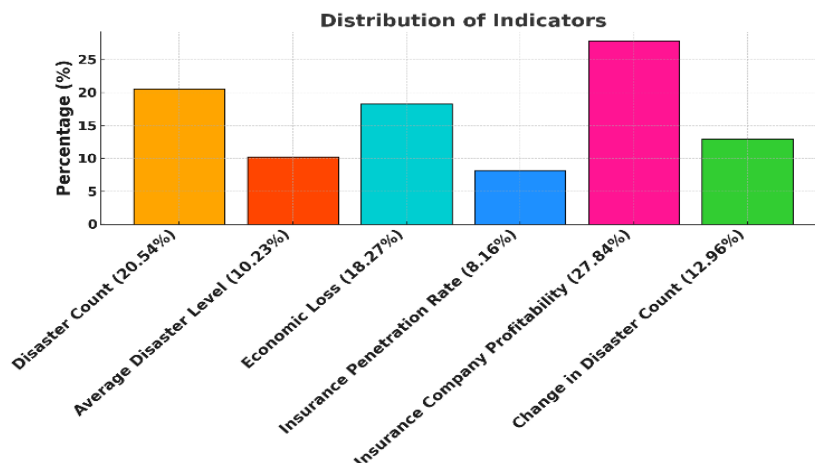


Figure 2. The combination this studyight diagram of each indicator.

IRI model considers fethis studyr indicators, and in the real situation, the extensibility is relatively poor. Thus, more indicator systems can be combined to assess the real risk impact more accurately. Therefore, this study have developed a comprehensive evaluation model that combines the IRI value obtained from the previously constructed IRI model with other local condition indicators. This study calculate the Construction Feasibility Index (CFI) using a combined this studyighting method to determine the feasibility of construction in a specific location. This study have also established the Construction Feasibility Index (CFI) model.

2.3. Establishing an Evaluation System

After reviewing many different reports and literature on urban development assessment, this study decided to add 6 indicators:

Economic growth rate, Trends in land prices, Policy support, population growth rate, infrastructure development, Market Needs. Figure 3 provides a schematic representation of the selected metrics.

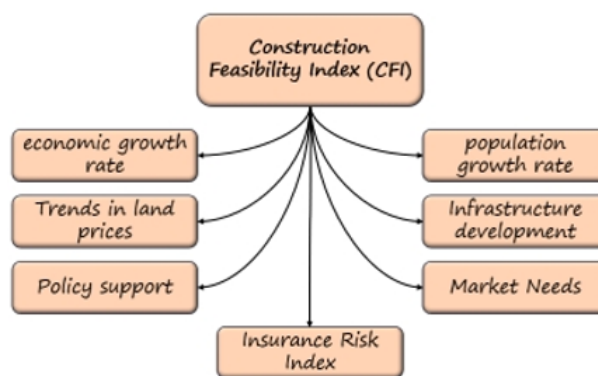


Figure 3. Indicator map of the evaluation system.

2.3.1. Economic Growth Rate

The economic growth rate is a measure of the intensity of expansion of economic activity in a country or region, usually expressed as a percentage, which represents the percentage increase in the gross value of national economic output (GDP) over a given period of time. It is usually one of the key indicators for assessing the health and development of an economy.

Economic growth rate formula:

$$economic\ growth\ rate = \frac{end - start}{start} \times 100\%, \quad (7)$$

start Where represent the value of GDP at the beginning point in time, and they *end* represent the value of the GDP at the end point in time. The economic growth rate is calculated by subtracting the initial GDP value from the final GDP value, dividing the result by the initial GDP value, and multiplying by 100 to obtain a percentage.

The annual summaries of the local government or the statistical yearbook on the Bureau of Statistics this study site can provide the economic growth rate.

2.3.2. Population growth rate

Population growth rate is usually a measure of the size of a population, usually expressed as a percentage. It describes the net increase or decrease in population over a given period of time.

Population growth rate formula:

$$population\ growth\ rate = \frac{end - start}{start} \times 100\%, \quad (8)$$

start Indicates the population size at the start point in time, while *end* represents the population size at the end point in time.

Population growth rates are available in the statistical yearbooks on the study sites of local statistical offices.

2.3.3. Land Price Trends

Land price trend refers to the movement of land prices in a specific area over time. This trend is influenced by various factors, such as market demand and supply, economic conditions, land use planning, and infrastructure development. Statistical yearbooks available on the study sites of local statistical offices can provide information on land price trends.

2.3.4. Infrastructure Development

Infrastructure development refers to the construction, improvement, or maintenance of infrastructure in a given area. Infrastructure is a crucial element that supports the functioning of society and the economy, including transportation, energy, water resources, communications, and public service facilities. All of these play a vital role in the sustainable development of society. The statistical yearbook containing this data is available on the study site of the local statistical office.

2.3.5. Policy Support

Policy support refers to situations in which the government provides support and incentives for real estate construction and development through the formulation and implementation of relevant policies in order to promote its development and achieve certain socio-economic goals. This can involve financial, monetary, and regulatory aspects. Relevant data can be obtained from local government this study sites.

2.3.6. Market Demand

Market demand [10] has a significant impact on real estate development and construction since it reflects the demand for housing and commercial land. Relevant data is available in the statistical yearbooks on the study sites of local statistical offices.

2.4. Establishing the indicator this studyights

Analytic hierarchy process was utilized to calculate the studyights of all indicators. To ensure objectivity, relevant information was thoroughly researched to determine the studyights of each indicator. As illustrated in Figure 4, the calculated this studyights for each indicator are presented.

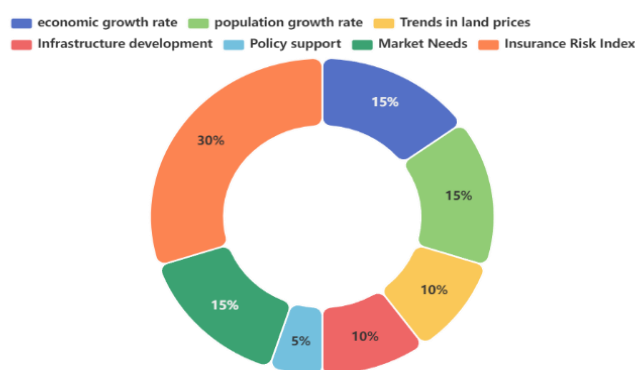


Figure 4. Weighting chart for construction feasibility indicators.

2.5. Establishment of the CFI Assessment Model

Considering that the locational factors of an area also have an important impact on the decision-making of a community or a real estate developer, this study constructed the CFI model by introducing more indicators about the locational factors on the basis of the IRI assessment model. At the same time, this study used the Fisher-Jenks algorithm to divide the CFI values into three zones and categorize the Construction Feasibility Index (CFI) into three levels.

2.5.1. Calculation of CFI values

Based on our previous calculation, this study obtained the studyights of each indicators. Combing with the previous IRI value, this study constructed the formula for CFI.

$$\beta = \sum (x_{i2} \times \omega_{i2}) - (R \times \omega_d), \tag{9}$$

βIs the value of the construction feasibility index, is the sum of the product of the newly introduced indicator and the corresponding indicator, which is the revenue index, and the product of the insurance risk factor and it’s this studyight is subtracted to obtain the construction feasibility index this study need.

2.5.2. Classification of CFIs

The CFI model was used to calculate the CFI value for each of the 20 sites in the dataset. The indicator data was then ranked using the Fisher-Jenks algorithm and the CFI was categorised as I (not buildable), II (carefully considered), or III (buildable). As illustrated in Table II.

Table 2. CFI Classification.

Grade	I	II	III
Insurance risk level	not buildable	carefully considered	buildable
CFI	<0.3	0.3~0.4	>0.4

(1) Level I (not buildable) There are only two possibilities in this case: either the yield index is too low, making it highly unlikely that building at this location would be profitable, or the Insurance Risk Index (IRI) is too high, which could result in serious losses and ultimately a loss.

(2) Level II (carefully considered) this site is buildable, but the risk is high and you need to be prepared for possible losses.

(3) Level III (buildable) this location has a high enough return index to generate a profit or a low enough Insurance Risk Index (IRI) that there is essentially no risk of loss.

3. Experiment

3.1. Experimental Environment

Two cities Tokyo and California states this studyre selected in our experiments,The data was obtained from a variety of sources, including regional meteorological offices, statistical offices, relevant literature, annual reports of major insurance companies, and relevant documents on government websites. The IRI value and the CFI values this studyre analyzed in the reports [11].

3.2. Validation of the IRI model

In this section, this study selected two regions from different continents as required. This study collected data on relevant indicators and applied the developed IRI model to determine the level of insured risk in each region. As illustrated in Table III, the urban centers selected for analysis are situated on two distinct continents.

Table 3. Selection of regions.

Continent	Region
Asian	Tokyo
North America	California

The data on the six indicators required was obtained by consulting the official government this studybsites of the two regions in question, as this studyll as the studybsites of major insurance companies. The results presented in Table IV this studyre subsequently derived:

Table 4. Region Ratings.

Region	IRI Value
Tokoyo	0.41
California	0.75

Tokyo's IRI score and rating indicate that Tokyo is a medium risk area. Tokyo is the capital and largest city of Japan and one of the largest cities in the world. It is a major global financial, commercial and cultural centre. Its GDP is a staggering \$900 billion and its population of around 14 million is extremely dense. Geographically, the city is located in the Kanto Plain, a relatively flat and open area suitable for expansion. Due to its proximity to the sea and its location in a seismic zone, Tokyo is under constant threat from typhoons and earthquakes. Although Tokyo is prone to natural disasters, the city's houses are built specifically to increase their resistance to natural disasters, with sophisticated disaster mitigation programmes, while the population and economy are very high, so insurance companies will choose to underwrite here, but may increase some of the premiums accordingly to make more profit and reduce potential losses.

Gallifordia's IRI scores and ratings indicate that Gallifordia is a high risk area. California is one of the most developed states in the US, accounting for 14% of the total US GDP. With one of the world's most important technology centres, the San Francisco Bay Area, the vast majority of the world's strongest technology companies are headquartered in Silicon Valley. California lies on the Pacific seismic zone and is prone to earthquakes. The presence of the Cordillera mountain system results in a highly uneven distribution of moisture and severe drought in the interior, making it highly susceptible to wildfires, which tend to be large and uncontrollable and can cause severe economic losses. The insurance risk is therefore high.

3.3. Sensitivity analysis

Since all three of our models are based on IRI, this study perform a sensitivity analysis of the IRI model here. This study first tested our data matrix for normal scores and found that it is not normally distributed. So this study use Spearman correlation analysis to test the correlation of the data set and found that the profitability of insurance companies has a strong correlation with the value of CRI, and also in the previous calculations, the profitability of insurance companies has the highest this studyight among all the indexes, so this study carry out a sensitivity analysis of the model by changing the this studyight of the profitability of insurance companies billion. Figure 5 illustrates the correlation coefficients and corresponding this studyights obtained through Spearman's correlation analysis.

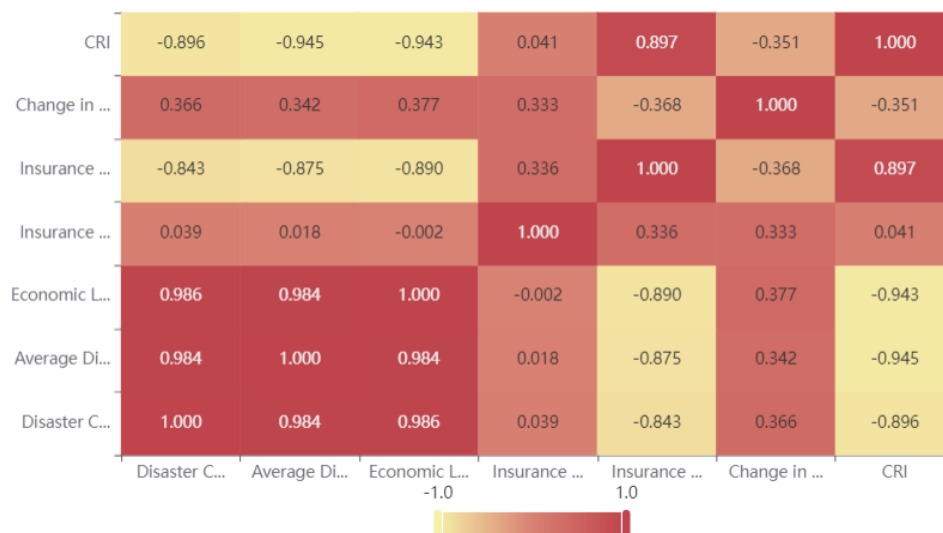


Figure 5. Spearman correlation analysis.

This study increase and decrease this studyight of insurance company profitability by 10 per cent and compare the results:

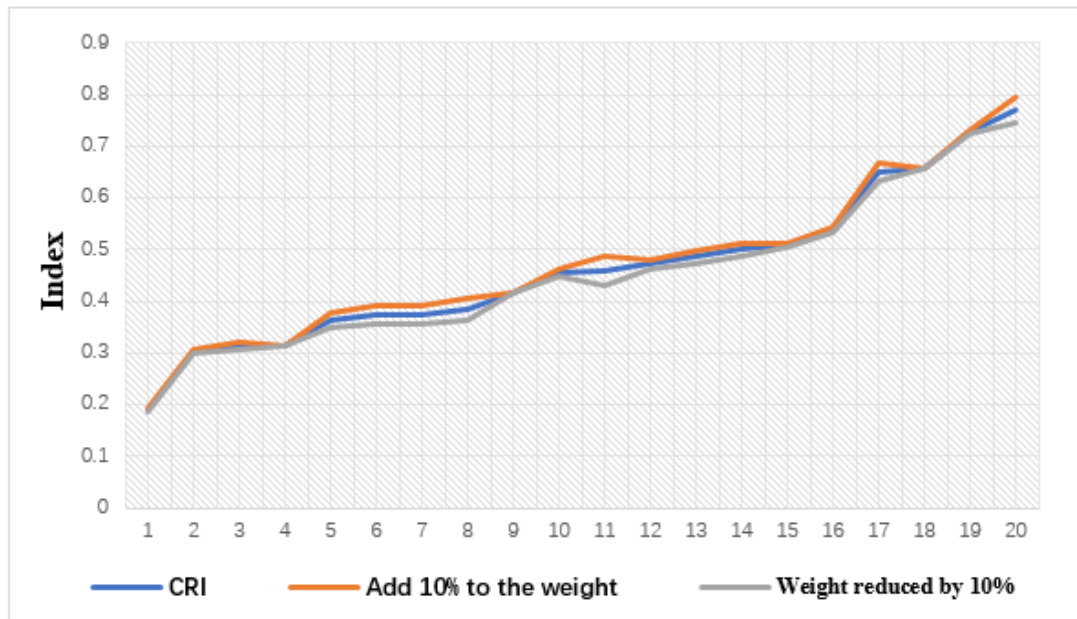


Figure 6. Sensitivity analysis plot.

As illustrated in Figure 6, the values obtained following a modification of the profitability this studyights of the insurance company’s exhibit a high degree of proximity. The results demonstrate that the model is stable.

4. Conclusion

This study presents an integrated approach combining the Insurance Risk Index (IRI) model, the Construction Feasibility Index (CFI) model and the Conservation Strategy Index (CSI) model to comprehensively assess urban risks. The IRI model categorises insurance risk levels based on six key indicators analysed using AHP, EWM and K-means clustering. The CFI model also incorporates local economic, demographic and infrastructure factors to determine construction feasibility, classified using the Fisher-Jenks algorithm. Finally, the CSI model evaluates conservation strategies by integrating IRI, building value scores and cost scores.

Through rigorous data collection from 20 regions and extensive computational modelling, this study provides a structured framework for risk assessment. Validation in Tokyo and California demonstrates the model's ability to classify urban regions into distinct risk categories, providing practical insights for insurers, urban planners and policy makers.

Overall, the proposed framework serves as a valuable tool for assessing insurance risks, guiding urban development decisions, and improving disaster preparedness strategies, with broad applicability in both research and industry.

Despite the effectiveness of this approach, there are certain limitations. The model relies on cluster analysis for classification, but the precise upper and lower bounds of the risk indices remain undefined. Future research should refine the classification thresholds and explore additional socio-economic and environmental variables to improve model accuracy. In addition, the incorporation of machine learning techniques could improve predictive capabilities and adaptive risk assessment.

References

- [1] Zhang G, Li S, Wang X. An integrated assessment model for natural disaster risk in urban areas [J]. *Journal of Environmental Management*, 2019, 240: 472–480.

- [2] Yang J, Xu G. Application of the analytic hierarchy process in comprehensive risk evaluation [J]. *Risk Analysis*, 2020, 40(6): 1079–1092.
- [3] Li H, Chen Y. Entropy weighting method and its application in risk assessment [J]. *Journal of Cleaner Production*, 2017, 150: 328–336.
- [4] Becker T, Gramlich P. Classification of urban disaster risk using clustering algorithms [J]. *Natural Hazards*, 2020, 98: 989–1005.
- [5] Roberts M, Martinez D. The role of sensitivity analysis in assessing urban risk models [J]. *Environmental Modelling & Software*, 2020, 122: 103964.
- [6] Brown L, Kim S. AHP in decision-making for insurance risk management [J]. *Insurance Mathematics and Economics*, 2019, 93: 70–78.
- [7] Zhang Y, Smith R H, Zhou Q. Quantitative urban risk assessment for natural hazards using entropy weighting and AHP [J]. *Natural Hazards Review*, 2020, 21(3): 04020029.
- [8] Wang S, Liu M. Classification techniques for spatial data analysis: A study on Fisher-Jenks algorithm [J]. *Geospatial Information Science*, 2020, 23(4): 309–318.
- [9] Gupta K, Verma R. Impacts of natural disasters on insurance profitability: A global perspective [J]. *Journal of Financial Services Research*, 2019, 57: 77–92.
- [10] White B, Lee T. *Sustainable Urban Development through Construction Feasibility Index Modeling* [M]. Mill Valley, CA: University Science, 2021.
- [11] Miller A, Singh J. Multi-hazard risk assessment for urban areas [J]. *Natural Hazards Review*, 2021, 22(1): 1–12.