

A Modeling Study of Insurance and Real Estate Risk Assessment in the Context of Global Climate Change

Lianshu Peng[#], Yida Wang[#], Zhensheng Pan^{*, #}

International school, Beijing University of Posts and Telecommunications, Beijing, China, 100876

*Corresponding author: 617446218@bupt.edu.cn

[#]These authors contributed equally.

Abstract. This paper focuses on the problem of economic losses caused by extreme weather events in the context of global climate change, and aims to develop a risk assessment model applicable to the insurance and real estate industries. The study first uses the Spearman correlation coefficient to conduct sensitivity analysis, determines the premium as the optimization variable, and combines the idea of present value of compound interest to construct a risk rating evaluation system, which classifies the risk into A, B, and C to guide the decision to insure. Subsequently, with the help of historical climate data, the ARIMA algorithm is used to predict future climate risk, and empirical evaluation is conducted for the United States and Australia to predict the future loss trend of the two countries. In order to improve the science of real estate siting decision-making, the study introduces a GIS-based decision support system, combines the gradient boosting tree algorithm to predict the risk factors, and constructs a visual GIS model to assess the suitability of siting. In addition, the study quantifies the cultural, historical, economic, and community values of the building facilities, and ranks and weights the factors by entropy weighting method to emphasize the protection of high-value factors. This study provides powerful tools and strategic recommendations for the insurance and real estate industries to address climate risk.

Keywords: Compound Interest Present Value Assessment, GIS, Gradient Lifting Tree, ARIMA.

1. Introduction

Extreme weather events are increasingly becoming a crisis for property owners and insurers [1-2]. As the effects of global climate change become more pronounced, the frequency and intensity of extreme weather events are increasing, posing significant challenges to the insurance industry and creating an urgent need for sustainable development [3-4]. Losses from extreme weather events are beyond the capacity of many insurers. Catastrophic losses from extreme weather events such as hurricanes, floods and heavy rainfall are on the rise, putting the insurance industry's capital reserves and claims-paying capacity to a severe test. Insurers need to develop more inclusive and flexible insurance products to better cope with increasingly complex and diverse risks. Climate change poses a serious challenge to the long-term robustness of the insurance industry [5]. Continued extreme weather events have exacerbated the complexity of risk assessment and management, and traditional risk modeling and pricing methodologies may not be able to effectively address future risks [6]. Therefore, insurers need to strengthen their monitoring and early warning capabilities for climate-related risks and actively adopt new technologies and data analysis in order to improve the accuracy and flexibility of risk forecasting. In summary, extreme weather events pose a serious challenge to the insurance industry, but also provide new opportunities for its development. Insurance companies need to actively respond to the challenges posed by climate change and strengthen their awareness of innovation and sustainable development in order to cope with the risks of future uncertainty and achieve sound long-term business development.

2. Construction of risk assessment model

The main sources of data in this paper are shown in Table 1.

Table 1: Main Data Description and Data Source

Data Description	Data Source
Losses from extreme weather as a percentage of GDP	https://ourworldindata.org
World GDP statistics	https://www.macrotrends.net
Other data	document literature

2.1. Estimated Losses and Premiums

Based on the frequencies of the indicators, the determinants of product insurance risk were derived as shown in Figure 1 below:



Figure 1: Regional property risk

We established a correlation analysis model by introducing Spearman's correlation coefficient to ascertain the relationship between the indicators. A statistical technique for determining the correlation between two variables is the Spearman's rank correlation coefficient. It may be used to any sample size, doesn't require assumptions about the distributional features of the data, and is derived by ranking the raw data.

With X_i and Y_i corresponding to the corresponding i_{th} ($1 < i < n$) element, X and Y are arranged in the same ascending or descending order to obtain a new sequence of variables x and y . Accordingly, X_i and Y_i correspond to the i_{th} element, respectively, and with $d_i = x_i - y_i$ the set of elemental differences, the formula for calculating the Spearman's correlation coefficient between the random variables X and Y is as follows:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^2 (n^2 - 1)} \tag{1}$$

The numerator, which represents the correlation difference between the variables, is the sum of the sequence errors, while the denominator is a constant associated with the length of the variable sequence. The direction and degree of the trend of change between the two variables are reflected by the value of ρ , which ranges from $[-1, 1]$. A value of 0 indicates no correlation between the variables, a value of negative indicates a negative correlation, and a value of positive indicates a positive correlation. The stronger the correlation between the two variables, the larger the absolute value of the value.

2.2. Sensitive analysis

A sensitivity analysis on the five parameters mentioned above—expected claim rates, average cost per policy, average amount per claim, time value of money discount rate, and fixed costs—must be carried out in order to evaluate the effects of various factors on insurance. Linear Profit and Loss Analysis and Compound Present Value are Used to Assess the Cost of Risk. We can begin by figuring out what the entire cost of production is. A combination of fixed and variable costs determines it:

$$C = C_f + C_v \times Q \tag{2}$$

Next, the total amount of the claim is calculated, where Q represents the insurance number and rc is the claim rate, depending on the disaster occurrence factor rz and the disaster loss factor rs :

$$C_c = rc \times Q \times Ca \times \frac{1}{(1+i)^t} \tag{3}$$

$$rc = rz \times rs \tag{4}$$

Operating income:

$$B = (P - T_b - T_v) \times Q \tag{5}$$

Lower limit of capital preservation yield:

$$Q^* = \frac{C_f}{P - T_b - T_v - C_v} \tag{6}$$

Break-even sales revenue:

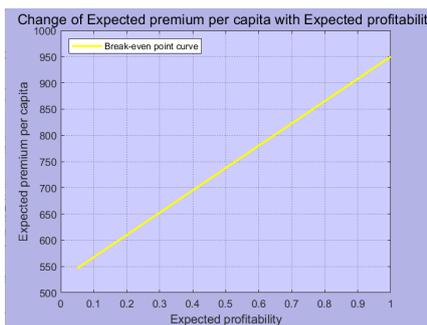
$$Q^* = \frac{C_f}{P - T_b - T_v - C_v} \tag{7}$$

The lower the value of Q is, the stronger the ability of the program to adapt to market changes, that is, the stronger the ability to resist risks.

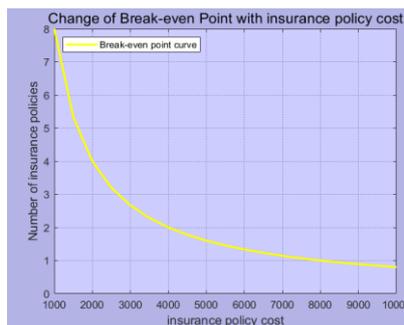
The default value of our mode is shown in Table 2.

Table 2: Insurance program bench marking factors

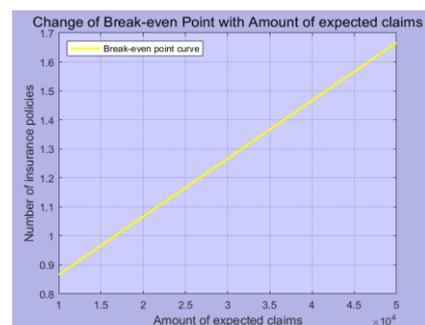
Norm	reference value
Expected claim rate	0.06
Average cost per policy	3000
Average amount per claim	100000
Discount rate of time value of funds	8%
Fixed value	2000



(a) Expected loss ratio



(b) Average cost per policy



(c) Average amount per claim

Figure 2: Part of insurance basis factors affect the change

As shown in Figure 3, the independent variable of each indicator changes by the same multiple and the slope of the change is taken to plot the sensitivity analysis as follows:

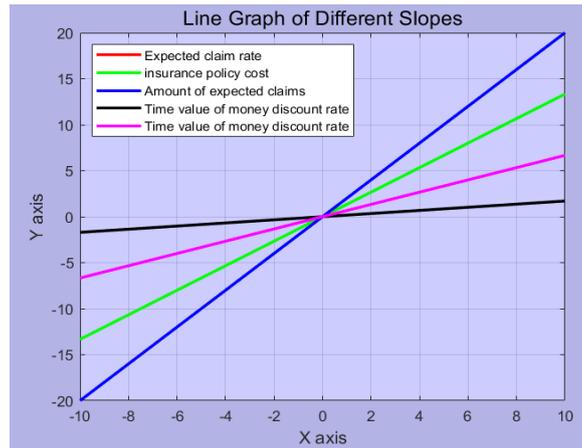
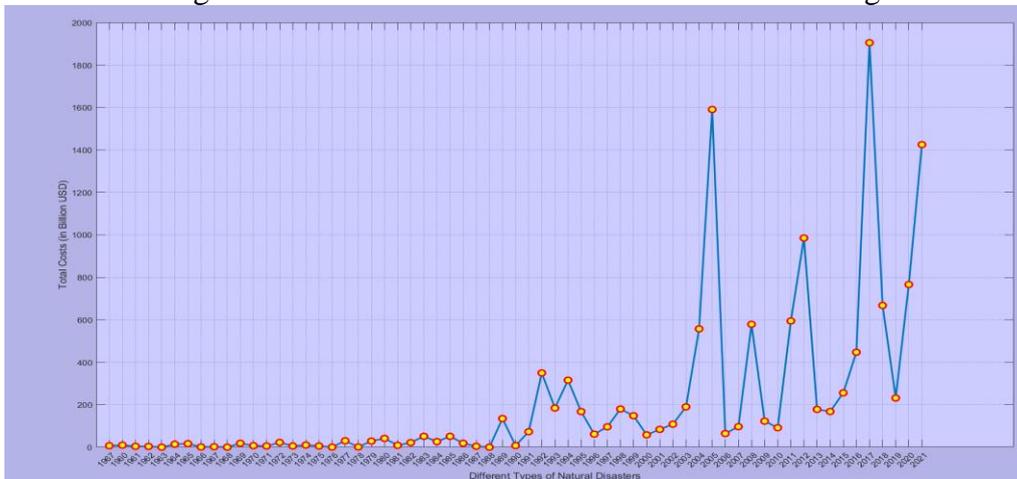


Figure 3: Sensitivity Analysis Chart

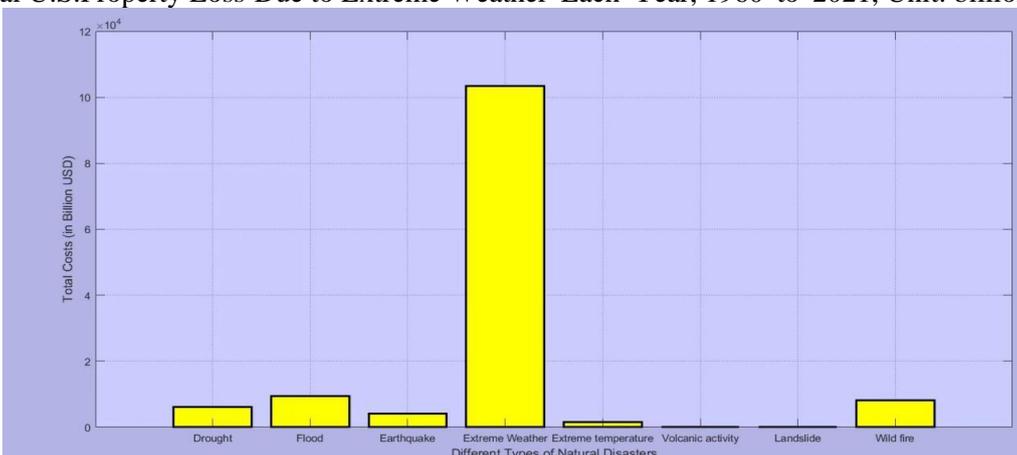
This paper can only observe four straight lines because line: Amount of expected claims and line: Expected claim rate overlap. Their horizontal axes intersect at zero, As shown in Figure 3, Expected claim rate have the greatest impact on the break even revenue balance, which is identified as the model regression criterion.

2.3. Data Visualization

Combining the data looked for into two csv files, we visualized the data collected from the United States and Australia. We analyzed regional disaster events in the United States and Australia separately. Total losses from different types of natural disasters and the total value of losses per year are visualized for both regions from 1960 to 2021. The result are shown in Figure 4 and Figure 5.

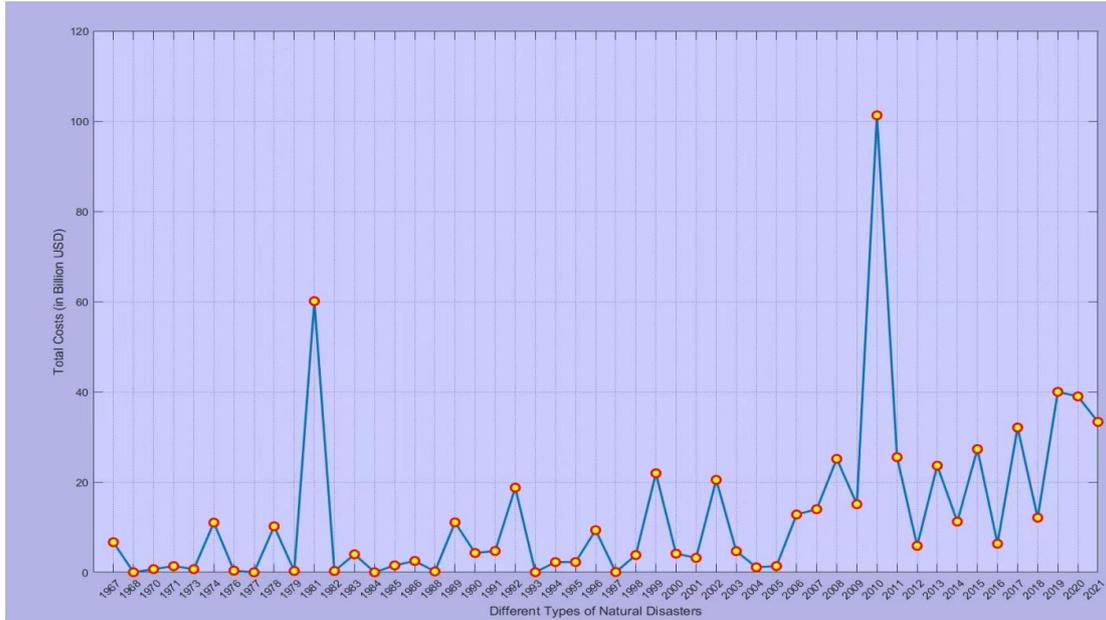


(a) Total U.S. Property Loss Due to Extreme Weather Each Year, 1960 to 2021, Unit: billion dollars

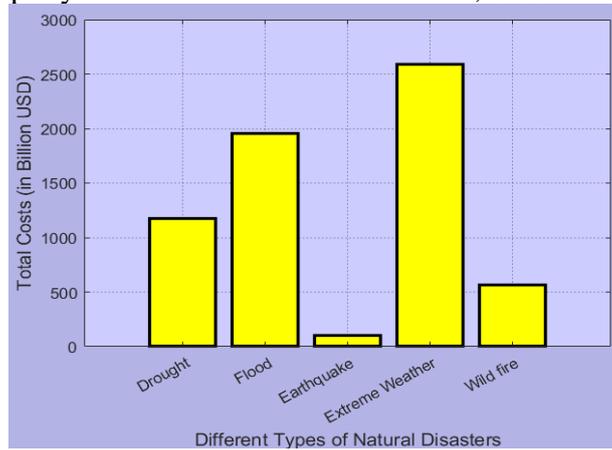


(b) Total property damage from different extreme weather events in the United States from 1960 to 2021

Figure 4: USA



(a) Total Australia Property Loss Due to Extreme Weather, 1960 to 2021, Unit: billion dollars



(b) Total property damage from different extreme weather events in Australia from 1960 to 2021

Figure 5: Australia

2.4. Establishment of evaluation system

2.4.1 Risk assessment models

For the correlation of correlation coefficients of different regions, the risk assessment model of the region is constructed by applying including multiple regression. The risk assessment model for each region is expressed as:

$$P (\text{ExtremeWeatherEvent} | \text{region}) = \sigma(\beta_0 + \beta_1 \times \text{HistoricalData} + \beta_2 \times \text{MeteorologicalConditions} + \beta_3 \times \text{GeographiFeatures} + \dots) \quad (8)$$

Categorized into three risk levels, A, B and C, based on prediction results. Insured if style rating is A or B, not insured if C.

2.4.2 Premium calculation

Depending on the risk class, different Base Premium Rates (BPR) are set to represent the base cost per unit of face amount. Based on the risk class, the Base Premium Rate is adjusted by a factor to reflect risk differences. Thus the premium (P) can be expressed as

$$P = \text{Base Premium Rate} \times \text{Adjustment Factor} \quad (9)$$

2.4.3 Adjustment Factor calculation

Based on regional disaster differences, the calculation of the adjustment factor varies slightly from region to region, with the following formulas:

$$AdjustmentFactor = \alpha + \gamma \times p(extremeweatherevent/ area) + \gamma z \times factors \quad (10)$$

3. ARIMA model

This paper chose the ARIMA model because the value-at-risk of insurance fluctuates with a general trend over time ((Figure 6), where the trend is constituted by the influence of historical labels, the fluctuation are constituted by the influence of contingencies over a period of time, and the general trend itself is not necessarily stable [7-8]. By using auto-correlation and difference of the data, the time series patterns hidden behind the data are extracted and the future data is predicted, the basic formula of the ARIMA model is as follows:

$$Y_t = c + \phi_1 \times Y_{t-1} + \phi_2 \times Y_{t-2} + \dots + \phi_p \times Y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \times \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (11)$$



Figure 6: Arima Flowchart

3.1. Stability Analysis

The ADF (Augmented Dickey-Fuller) time series data stability test was first performed on the U.S.data to detect the presence of a unit root in the time series [9-10]. This paper can observe that P is less than 0.05, which means the data is unstable.

3.2. Difference and Normalisation

We convert non-smooth data to smooth by difference. Below are the results of the data after first order difference compared to the original data it can be seen that the stability of the data has been improved somewhat (Figure 7).

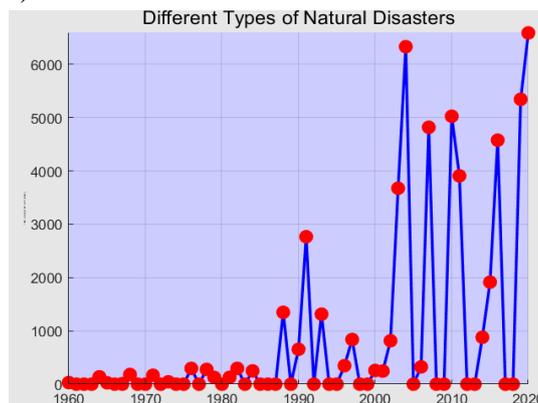


Figure 7: Total U.S. Property Loss Due to Extreme Weather Each Year after first order difference, 1960 to 2021, Unit: billion dollars

3.3. Model parameter determination

Autocorrelation coefficient plots and partial auto-correlation coefficient plots are used to determine the two parameters, p and q, in the ARIMA model, while the order of the difference determines the ARIMA parameters. After establishing these parameters, we can fit the data with the ARIMA model and evaluate the model fit by looking at the residual plots.

It is possible to calculate the values of P and Q since the post-truncated tail orders of the PACF and ACF are P and Q, respectively. The relevant parameters are as below Figure8.

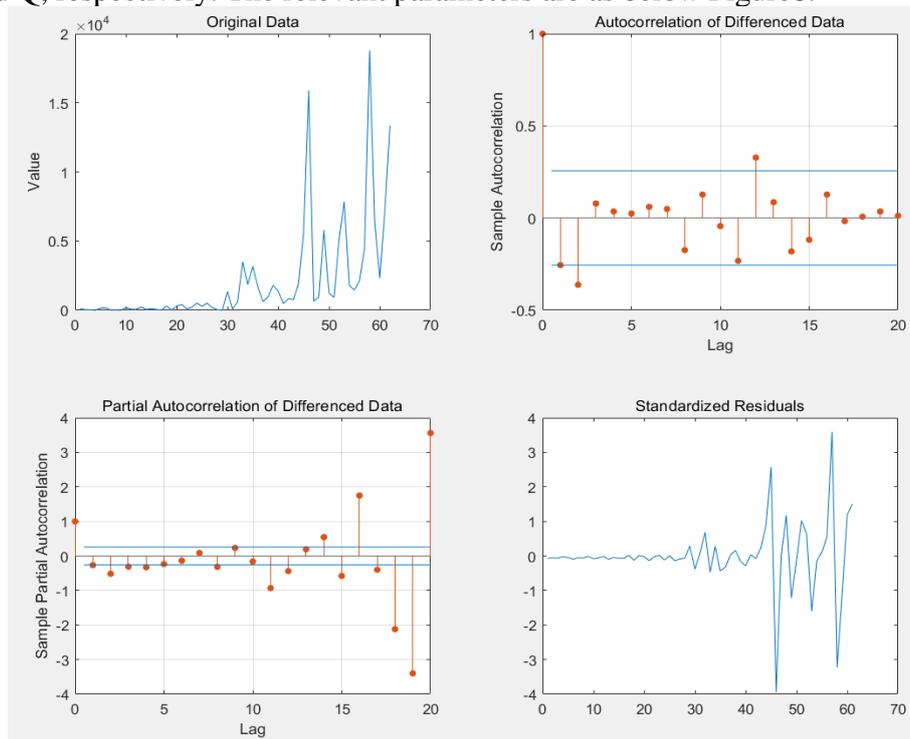


Figure 8: Parameters related to the ARIMA model

3.4. Model Predict Result

We can see an upward trend in insured losses due to natural disasters every year. (Table 3) The total expected compensation cost (N) is calculated first:

$$N = (total.national.losses / national.population) * M \tag{12}$$

To reach the expected profit margin P needs to be guaranteed while covering losses:

$$M * C = M * K + N(1 + P) \tag{13}$$

Thus, the cost of participation is calculated:

$$C = K + \frac{N}{M} (1 + p) \tag{14}$$

The cost of enrollment per person is C, the number of enrollees is M, the total premium income is CxM, and the base premium is K.

Table 3: Projected losses over the next ten years

Year	American Estimated loss	Australian Estimated loss
2022	13569	338.9490
2023	13786	344.3911
2024	14003	349.8331
2025	14221	355.2751
2026	14438	360.7172
2027	14655	366.1592
2028	14873	371.6013
2029	15090	377.0344
2030	15370	382.4853
2031	15524	387.9274

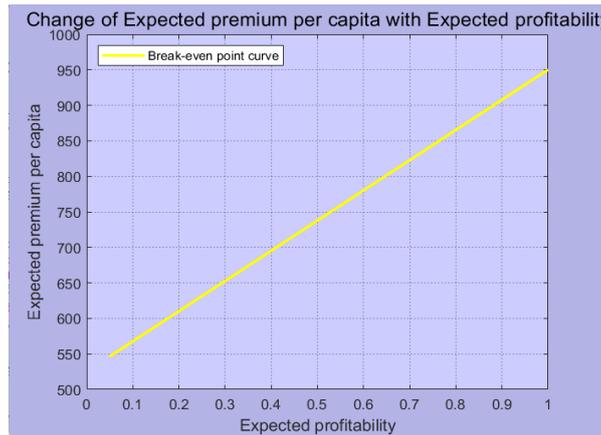


Figure 9: Change of Expected premium per capita with Expected profitability

The change in enrollment costs with expected profitability is plotted in Figure 9. As an example, projecting the U.S. to 2024 and assuming that the base premium industry standard is being \$800K. Expected profit margin P of 20%. We can calculate the U.S. premium to be \$5,892. Similarly, the annual premium for Australia is calculated to be \$2,415.

In 2024, the U.S. will have a GDP per capita of \$75,269 and a premium share of 7.8%, while Australia will have a GDP per capita of \$64,003 and a premium share of 3.7%. Given the same expected profitability and other factors, the Americans will have to invest more in premiums, which in turn will affect the number of enrollees, and also reflect the higher risk rating of the U.S.

The owner has invested in establishing appropriate protection around the site to minimize potential damage from weather events [6]. This may include levees, floodwalls, storm gates, etc.

4. Conclusions

This paper describes the purpose of establishing a risk assessment model for insurance and a real estate siting assessment model to cope with the problem of increased economic losses due to global climate change. For this purpose, three models are developed in this paper: the Insurance Risk Assessment Model (IRA), the Real Estate Geographic Decision Support System (RDSS), and the Preservation of Buildings of Special Value (PBS) model. The IRA model determines the cost of insurance through correlation and sensitivity analyses, and predicts the future climate risk using historical climate data and the ARIMA algorithm, and has been evaluated in practice in the U.S. and Australia. The RDSS model predicts housing risk factors using GIS and gradient boosting trees and evaluates property siting suitability through GIS modeling. The PBS model calculates scores by quantifying the value factors of building amenities and using entropy weighting to rank and weight the scores to focus on protection factors. Finally, the article provides the results of an assessment of the Woodbury Arts Center with recommendations and a time schedule.

However, the insurance assessment model in this paper was able to categorize risk levels but did not provide specific risk-based continuous premium amount recommendations, and future research could further refine the risk ratings to obtain more accurate premium predictions. In addition, the data obtained by the crawler has a missing problem, and although the Lagrangian interpolation method is used in this paper for processing, it may still have a certain impact on the model fitting accuracy. Therefore, subsequent studies can try to use other interpolation methods, such as linear interpolation, spline interpolation or kriging interpolation, to further improve the accuracy and reliability of the model.

In summary, the research in this paper provides strong theoretical support and practical guidance for the insurance industry and real estate industry to cope with climate change risks, but it still needs further in-depth research and improvement.

References

- [1] Liang Rong, Chen Bingzheng. Assessment of global GDP loss rate of extreme weather events in the context of climate change [J]. *Systems Engineering Theory and Practice*, 2019, 39(3): 12.
- [2] Liu Lan. Global extreme weather towards normalization [J]. *Ecological Economy*, 2021, 37(9): 4.
- [3] Wu Daming. Analysis of global extreme weather and climate events and their impacts in recent years[J]. 2021.
- [4] Zhang Linhao. Inventory of global extreme weather and climate events in the first half of 2023 [J]. *Life and Disasters*, 2023(7): 24-27.
- [5] Huang Ying. Research on financial risk evaluation and control strategy of insurance companies [J]. *Investment and Entrepreneurship*, 2023(10): 93-95.
- [6] JIA Luyu, ZHANG Zeyang, LI Xuan, et al. Construction of risk evaluation index system for environmental pollution liability insurance [J]. *Environmental Protection Science*, 2021, 47(1): 9.
- [7] Gao Long. Research on ARIMA-based small-lot material production demand forecasting model [J]. *Modern Information Technology*, 2023, 7(15): 97-101.
- [8] ZHANG Shulin, FENG Pengfei, XU Jiajia, et al. Construction of a city highway traffic accident prediction model based on ARIMA model [J]. *Anhui Journal of Preventive Medicine*, 2023, 29(5): 407-411.
- [9] Zhao X, Liu K, Hui Y, et al. WOPCA-EGNPE algorithm based on variable division for fault monitoring of batch process [J]. *Brazilian journal of chemical engineering*, 2023(3): 40.
- [10] Worthington A C, Higgs H. Weak-form market efficiency in European emerging and developed stock markets. Discussion Paper No. 159 [J]. 2022.