# Research on Natural Disaster Risk Based on Garch Volatility and Monte Carlo Algorithm

Mingjie Sun [1, #, *], Yulu Dai [2, #], Jingyi Wang [3, #]

[1] School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China, 200237

[2] School of Mathematics, East China University of Science and Technology, Shanghai, China, 200237

[3] School of Business, East China University of Science and Technology, Shanghai, China, 200237

* Corresponding Author Email: smj9839894454@gmail

[#]These authors contributed equally.

**Abstract.** The purpose of this paper is to study the optimal decision problem of different agents. Firstly, the expression of the profit process of the insurance company is established, and then the loss caused by the natural disaster is fitted into the insurance formula using the double cross-verified natural disaster model. The first is the prediction model of random disaster loss based on Garch volatility test. The other is ridge regression, which combines the possible relevant factors and then makes the forecast. The two results were cross-verified by Garch volatility test again, and then risk assessment and scheme decision were made based on Monte Carlo algorithm and time value theory, and 50,000 simulations were carried out. By estimating the probability of frequency, the probability of bankruptcy under different premiums is obtained. This paper makes an empirical study of New Mexico and Henan in China. For the United States, the paper concludes that an average unit price of $34-35 per policy is appropriate. For Henan Province, this paper focuses on the implementation of natural disaster insurance, and carries out model simulation. The result shows that the premium price of 15% bankruptcy rate is 8 ~ 10 yuan, which is lower than the current policy price, proving that the policy has advantages.

**Keywords:** Double Cross Verification, Kriging, Monte Carlo Algorithm.

## 1. Introduction

With global warming, extreme weather is becoming more frequent. The occurrence of extreme weather will not only lead to property damage for the owners but also make insurance companies face more frequent claims [1]. If the insurance company implements too few policies, it will lose customers and become unable to operate; If insurance companies ignore the risks of extreme weather, they will face huge payouts [2]. How to make the insurance company does not lose customers at the same time while paying less amount is the problem that the insurance company must face. How to adapt to the frequent occurrence of extreme weather, how to choose habitable sites, and how to protect cultural relics in extreme weather have become urgent issues. In the article "Research on Current Situation and Problems of Medical liability Insurance Policy Implementation based on Horn-Mitter Model -- taking S City of Shandong Province as an example", literature review, interview and questionnaire survey are used to establish and implement insurance policies [3]. However, these methods are time-consuming and costly, and are not suitable for natural disaster insurance, because the occurrence of natural disasters is uncertain. From the perspective of law perfection, the paper "Research on the Legal System of Catastrophe Insurance in China" holds that catastrophe insurance has three obvious characteristics: marketization of operation mode, social security purpose and policy orientation [4]. However, there are no data studies and predictions, which cannot provide strong support for the claim. In this paper, the average range of insurance unit price is obtained by double cross-validation of existing data, and a more appropriate insurance policy is formulated.

## 2. Risk assessment and decision-making model for natural disaster insurance

### 2.1. Mathematical Simulation of the Surplus Process of an Insurance Company at Time

The mathematical simulation of the surplus process of an insurance company at time *t* is as follows:

$$U_{(t)} = S + I_n - k\sum_{i}^{M(t)} O_i \tag{1}$$

S represents the initial cost in the formula, O represents the loss incurred to the insurance formula under the occurrence of natural disasters, which is an uncertain fluctuation amount, and $I_n$ represents the total premium in that area, which is a standard that needs to be artificially determined.[5]

The expression for the probability of bankruptcy is as follows:

$$\varphi_{(t)} = P\{\exists t\epsilon[0,T], U_{(t)} < 0 \mid U_{(0)} = s\} \tag{2}$$

### 2.2. A Natural Disaster Model with Double Cross Validation

The algorithm is divided into two stages. Firstly, Monte Carlo cross-validation is used to detect abnormal samples, and the abnormal samples are initially screened to improve the stability. Secondly, based on the theory of Mahalanobis distance square approximately following chi-square distribution, the optimal threshold is determined adaptively, and the abnormal samples are re-identified in the remaining data set [6].

Based on research and considering the insurance gap in the occurrence of natural disasters, the following formula can be derived:

$$o_i = D_v \times \sigma \tag{3}$$

$$\sigma = 1 - \delta \tag{4}$$

$O_i$ represents the amount that the insurance company needs to compensate, $D_v$ represents the number of losses caused by natural disasters, $\sigma$ represents the insurance gap caused by this natural disaster, and $\delta$ represents the coverage rate of the insurance.

### 2.3. Improved Random Disaster Loss Prediction Model based on GARCH Volatility Test

The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model is a model used to capture the volatility of time series data. Since the traditional estimation methods of variance (Monte Carlo algorithm, historical simulation method, etc.) usually assume that the series is normally distributed, however, the return series is often characterized by a sharp peak and a thick tail, which means that the VaR estimated by the traditional method will underestimate the tail risk, that is, it will ignore the extreme risk caused by extreme events, resulting in inaccurate risk estimation results. The GARCH family volatility model can filter the autocorrelation and conditional heteroscedasticity of the return series well [7]. The principle is as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_p\sigma_{t-p}^2 \tag{5}$$

P represents the auto regressive order of volatility, while q represents the moving average order of volatility. The paper collected the loss values for various years in the target location for prediction. Based on the parameter estimation of the GARCH model, the paper obtained the upper and lower threshold values for the conditional variance (volatility) series and assessed the situation accordingly. The generated volatility is assumed to follow a standard normal distribution. Using this assumption, the paper generates realistic random volatility scenarios. By combining these generated volatility scenarios with the original values, the paper can forecast potential future loss values. It is important to note that this forecasting process can be iterated continuously, allowing for ongoing refinement and improvement of the predictions.

The formula for the normal distribution is as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{6}$$

### 2.4. Factor fitting model based on ridge regression

The paper believes that the number of losses caused by natural disasters, although a fluctuating and uncertain quantity each year, is closely related to certain factors. These factors include the local GDP, the number of automobiles, the area of agricultural land, the concentration of carbon dioxide, the local topography (such as hills, plains, mountains, etc.), and the frequency and intensity of natural disasters.

By adding L2 norm penalty term to the coefficient vector in the loss function, ridge regression is used to limit the growth of parameters that contribute little to the prediction and increase the complexity of the model. Therefore, ridge regression model has the advantage of restricting model complexity and reducing model overfitting risk [8].

The paper establish the relationship between the amount of losses caused by natural disasters and potential influencing factors using ridge regression.The formula for ridge regression is as follows:

$$\beta^{baririgse} = argmin\left\{\left(\sum_{=1}^{N}\left(y_i - \beta_0 - \sum_{j=1}^{p}\beta_\beta x_i\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right\} \tag{7}$$

After the expression of the relationship between disaster loss and independent variable is obtained, the disaster loss can be predicted by predicting the value of relevant independent variable. Although we have directly predicted the amount of damage caused by natural disasters, this paper argues that indirect prediction through other factors may be more accurate than direct prediction of the amount of damage caused by natural disasters. For different types of data, this paper will adopt different prediction methods according to their characteristics. For variables with small changes, such as agricultural land area and local topography (such as hills, plains, mountains, etc.), as well as variables that usually show an overall growth trend, such as GDP and car ownership, this paper uses the ridge regression model to forecast. On the other hand, for the occurrence frequency of natural disasters that fluctuates with time, this paper adopts GARCH model and volatility prediction to select the appropriate prediction model according to the characteristics of the variables, which can predict the disaster loss more accurately.

### 2.5. Risk assessment and decision-making model based on Monte Carlo algorithm and time value theory

Monte Carlo algorithm is based on the principle of probability statistics and repeatedly approximates the real value of complex system through a large number of experiments. In the risk assessment model, Monte Carlo algorithm can process the uncertainty factors such as natural disasters as random variables and predict the risk assessment results by simulating these random variables [9].

$$I_{n+1} = I_n \cdot (1 + i) \tag{8}$$

$I_{n+1}$ represents the equivalent value of the premium for the n th year.

By utilizing the Monte Carlo algorithm combined with the Cox-Ingersoll-Ross (CIR) model, the paper generate the interest rate I based on the time value theory.

The CIR (Cox-Ingersoll-Ross) theory is a concept in finance that suggests interest rates fluctuate around a mean value. If the interest rate deviates from the mean, it tends to revert back to the mean. The value of the interest rate at a later time is always dependent on its previous value. The Euler discretization formula for the CIR model is as follows. The paper has implemented the random generation of interest rates using Python.

$$\chi_t = x_{t-1} + k(\theta - x_{t-1})\Delta t + \sigma\sqrt{x_{t-1}}\sqrt{\Delta t}\epsilon_t \tag{9}$$

Some simulation results are shown below Fig.1.

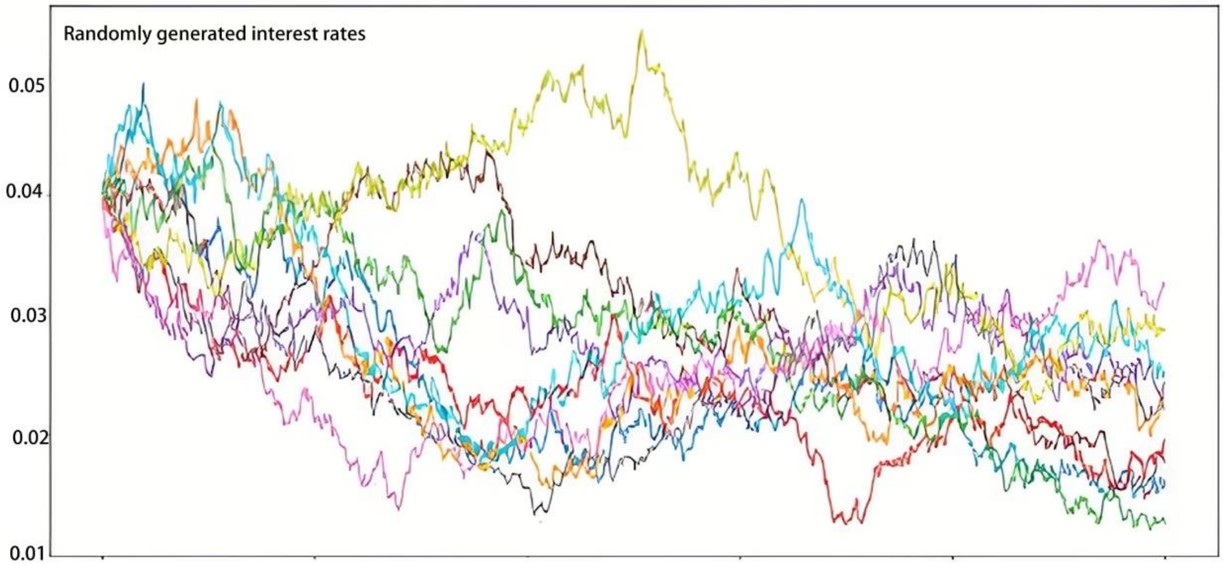


**Fig. 1** Random interest rate simulation

By setting different payout amounts and plugging them into the formula, the paper can calculate the bankruptcy probability for each payout amount. The paper considers a probability less than 15% as an indication that the plan is feasible.

$$I_n = \rho \cdot N \tag{10}$$

$I_n$ represents the premium for each policy, and $N$ represents the number of policyholders. The paper has ultimately determined the price per person per policy.

## 3. Improved Random Disaster Loss Prediction Model

### 3.1. Garch wave detection

As can be seen from Fig. 2, the volatility is around 60, between 50% and 70%.

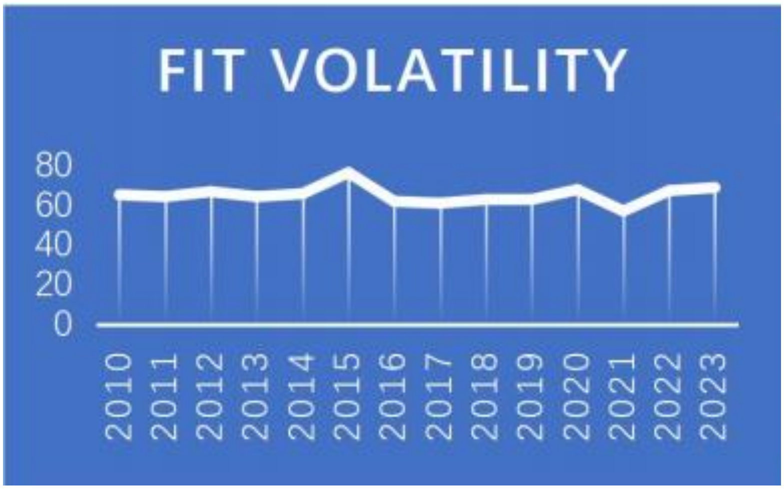


**Fig. 2:** Result of garch wave detection

### 3.2. Prediction of the amount of natural disasters

This article predicts that natural disasters will cause losses of approximately 45000 units. From the distribution map, it can be observed that there are some outliers with extremely high values, while most of the values are concentrated in the middle as shown in Fig.3.

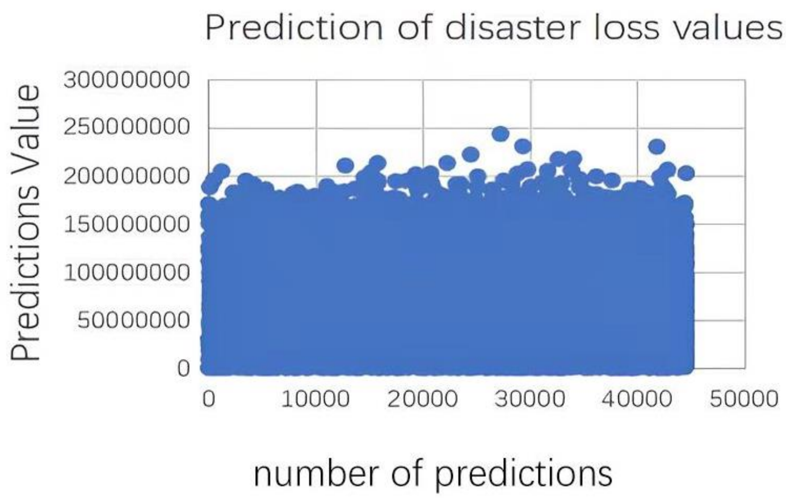


**Fig. 3:** Prediction of disaster loss value

### 3.3. Factor fitting model based on ridge regression

Ridge plot is shown in Fig. 4:

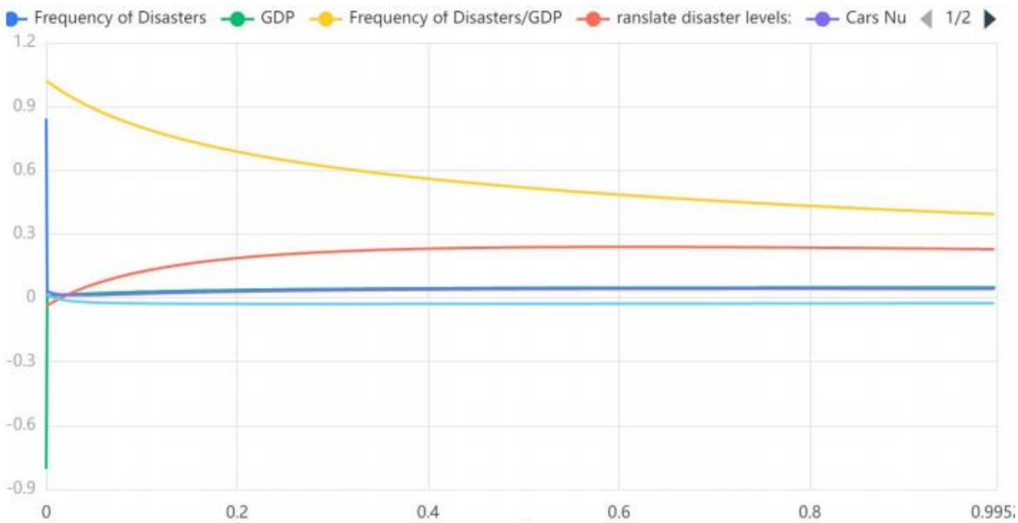


**Fig. 4:** Ridge trace map

According to variance enlargement factor method, K=0.109 was determined.

### 3.4. Volatility comparison

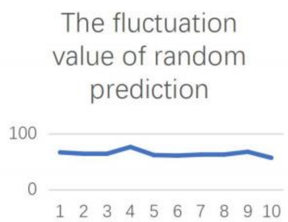


**Fig. 5:** Random prediction

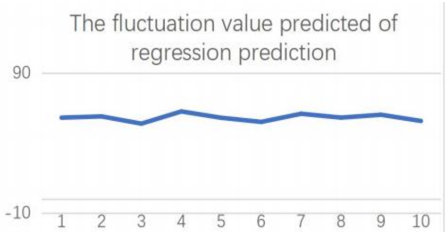


**Fig. 6:** Predicted of regression prediction

From Fig 5 and 6,the interpolation test is carried out again, and it is found that there is little difference between the two fluctuations, which indicates that the predicted data is reasonable.

### 3.5. Risk assessment and decision-making model

The paper generate 5000 random rates according to the CIR formula, and the paper run 100 simulations, each with a step size of 50. The results are shown in Fig. 7 and Fig. 8.

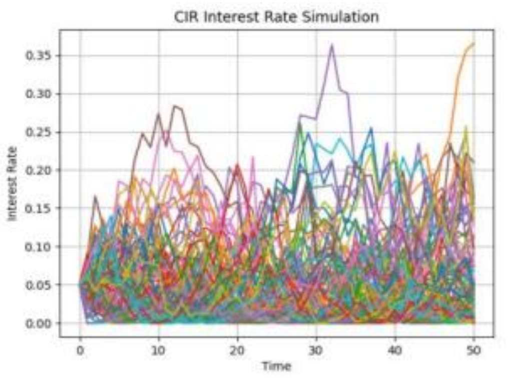


**Fig. 7:** CIR Interest Rate Simulation

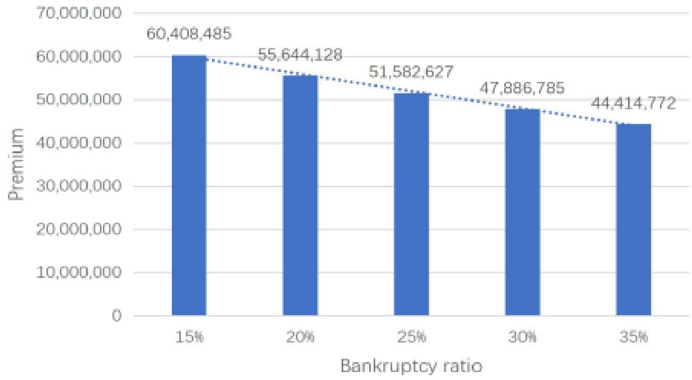


**Fig. 8:** The risk of bankruptcy

Since the Monte Carlo algorithm estimates the probability by frequency, the paper make three repeated predictions and get a conclusion: when the total premium income is 58037680.85 to 60103084.9, the bankruptcy profile is about 15%. According to the information, the insurance coverage rate in the United States is about 80%, and the paper set a safety factor of 30%:

$$I_n = \rho \cdot N \tag{11}$$

The total number of people in New Mexico is 2,117,522, the number of people who buy insurance =2,117,522*50%, and the unit price of insurance is better between \$54.8 and \$56.6.

### 3.6. Insurance decision scheme

From Fig. 9 can see that as premiums go up, the probability of bankruptcy goes down. The risk can be reduced by increasing the premium, but this practice will lead to the continuous loss of insurance business of the insurance company and cannot fundamentally reduce the repayment risk of the insurance company. Therefore, the insurance company should not over-rely on increasing the premium. The other is to increase the initial capital, which can be increased by expanding the scale of insurance, or by issuing bonds in the capital market[10]. It is recommended to establish a split decision, to establish different payouts based on the intensity of the disaster, and to ensure that the expected value is between \$54.8 and \$56.6.

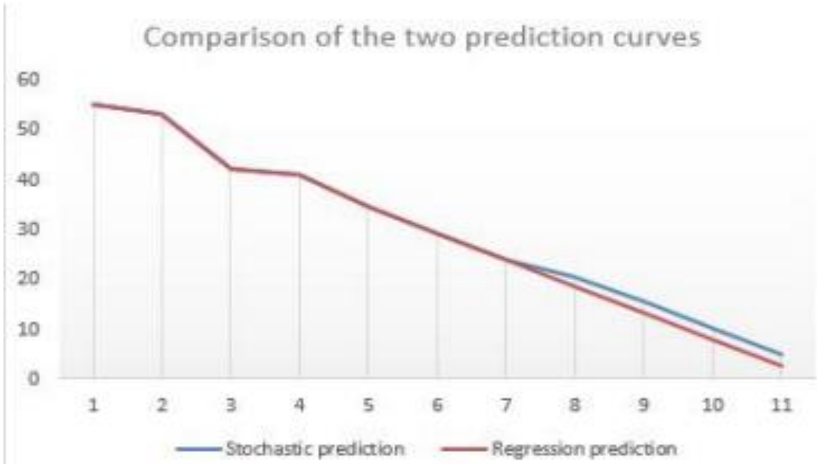


**Fig. 9** Comparison of two prediction curves

## 4. Adjustment of insurance model -- real estate decision model

### 4.1. The adjustment of insurance model

As can be seen from the insurance model in the previous question:

$$U_{(t)} = S + I_{\mathrm{n}} - k\sum_{\mathrm{i}}^{\mathrm{M(t)}} 0_{\mathrm{i}} \qquad \varphi_{(t)} = P\{\exists t \in [0,T], U_{(t)} < 0 \mid U_{(0)} = s\} \tag{12}$$

In the real estate decision model, S represents the initial scale of the real estate company, In represents the income from selling the real estate, Oi represents the cost required to build the real estate, and φ t represents the probability of loss of the real estate site under the following conditions.

For real estate, where the losses caused by natural disasters are greater, people are less willing to buy a house here, so the house will be sold at a lower price, resulting in a decrease in the income from selling real estate, In is decrease. Therefore, the amount of damage caused by natural disasters and In is closely related.

In order to withstand natural disasters and attract more people to buy our houses, the paper need to improve the disaster prevention ability of our houses. This will lead to Oi increase. According to architectural experience, the paper divides the disaster resistance of houses into the following three levels:

Low disaster rating: Relatively low disaster rating usually does not require additional seismic fortification requirements and measures, so the cost increase is small and negligible. Medium disaster rating: Medium disaster rating houses may require some additional seismic design and measures, such as strengthening the structure, adding earthquake-resistant components, etc. Depending on the specific requirements and the size of the building, the cost may increase by about 5 to 10 percent.

High disaster resistance rating: Buildings with high disaster resistance rating require more stringent seismic design and measures, such as the use of high strength building materials and complex structural forms [11]. This could lead to an increase in construction costs of about 10 to 20 percent or more, but at the same time it will attract more customers and improve In. Next, the paper analyzes the relationship between Dv and In and generates in by randomly generating Dv. The paper believes that in real estate will be related to GDP, disaster intensity, number, population, traffic (total kilometers). To estimate the needs of the growing population and life needs such as shopping trips, Arcgis is used to optimize the original model, and evaluate whether the area is easy to build houses by weighted layers. Then the paper compares it with randomly generated data to determine whether it is reasonable. The number of possible construction costs is obtained when the loss probability is 15%. If the cost can support us to build the corresponding grade of house, then the explanation can be built here.

### 4.2. Result of model

The paper continue to choose Henan, simulate this process, and predict whether each region in Henan Province is suitable for real estate development. The result is shown in the Fig. 10:

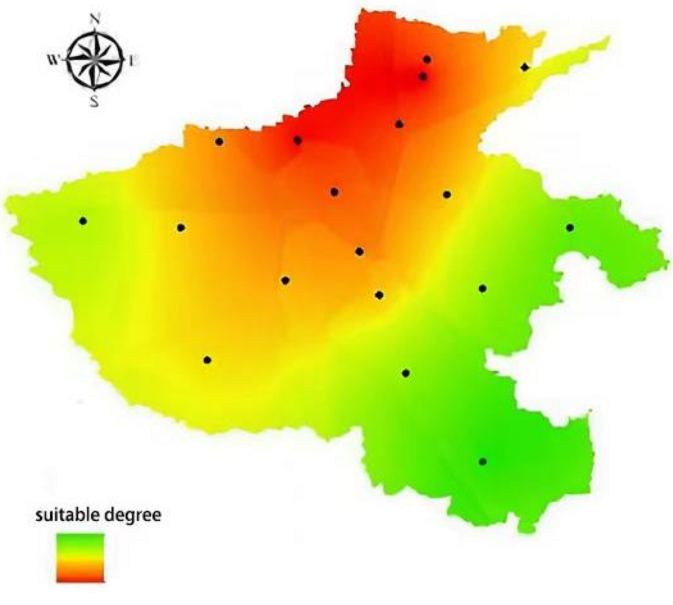


**Fig. 10:** Suitable area

The redder the color, the more suitable for real estate development, and the greener the less suitable.

## 5. Conclusions

Previous natural disaster models are generally catastrophe models, which only consider the occurrence of catastrophe, while this model considers the occurrence of natural disasters, which is more universal. The previous catastrophe model is a separate calculation of claims. This model considers the losses caused by auto insurance, house insurance, property insurance, life safety insurance and other insurance formulas through the statistics of natural disaster losses. Previous models held that the occurrence of disasters was subject to Possion distribution satisfying $\lambda > 0$. This model was double-cross-verified by its possible influencing factors, and the results were more accurate. In this model, ridge regression is adopted to reduce the negative impact of collinearity of influencing factors on prediction in data analysis and obtain more real and reliable regression coefficient. It's a policy that hasn't been practiced on a global scale, that could be hugely profitable and full of innovation.

## References

[1] Chao Wen, Qian Xiaotao. Simulation of catastrophe insurance bankruptcy probability based on Monte Carlo algorithm under random interest rate[J]. Journal of Minzu University of China (Natural Science Edition), 2021, 30(02): 47-51.

[2] Chao Wen, Qian Xiaotao. Solvency analysis of catastrophe insurance from the perspective of bankruptcy probability: Based on mixed distribution model [J]. Journal of Insurance Vocational College, 2023, 37(05): 51-57.

[3] Qu Tiantian. Research on current situation and problems of medical liability insurance policy implementation based on Horn-Mitter model [D]. Shandong university, 2022. DOI: 10.27272 /, dc nki. Gshdu. 2022.003537.

[4] ZONG Ning. Research on Catastrophe Insurance legal system in China [D]. Southwest University of Political Science and Law,2013.

[5] WUTHRICH M V. From ruin theory to solvency in non-life insurance[J]. Scandinavian Actuarial Journal, 2015(6): 516-526.

[6] Zhang Zhu-Shan-ying, Zhang Ruo-Jing, Gu Han-Wen, et al. Research on abnormal sample detection method of middle infrared spectrum double check [J]. Spectroscopy and Spectral Analysis, 2019, 44(06): 1546-1552.

[7] Chen Guodong, Wang Jiaqi. Research on tail risk measurement of Social Security Fund investment based on GARCH-POT-VaR model [J/OL]. China securities and futures, 1-16 [2024-08-29].

[8] Sun Chao. Fault Prediction of Metro Vehicle-mounted Equipment based on Ridge Regression [J]. Railway Communication Signal Engineering Technology, 2024, 21(08): 74-79.

[9] Ya Qi, Zhang Wen, Wang Jiahui, et al. Electric Power Engineering Cost calculation and Prediction Method based on Monte Carlo Algorithm [J]. Electrical Technology & Economics,2024, (07): 256-258.

[10] Chao Wen, Qian Xiaotao. Solvency analysis of catastrophe insurance from the perspective of bankruptcy probability: Based on mixed distribution model [J]. Journal of Insurance Vocational College, 2023, 37(05): 51-57.

[11] Zhang Xinru, Ma Shixia, Murui. Optimal reinsurance investment strategy with relative performance concern under stochastic volatility model and default risk [J]. Journal of Nankai University (Natural Science), 2022, 55(05): 30-43.