Transactions on Economics, Business and Management Research
ISSN: 2960-1789, eISSN: 2960-2254 | Volume 11
EMFT 2024

Research on the insurance underwriting model based on catastrophic risk

Runyu Zhong [1, *, #], Kaixin Li [2, #]

[1] College of insurance, University of International Business and Economics, Beijing, China, 100105

[2] College of information, University of International Business and Economics, Beijing, China, 100105

* Corresponding Author Email: 202230083@uibe.edu.cn

[#]These authors contributed equally.

Abstract. The frequent occurrence of extreme weather events has led to significant property damage, and at the same time, the amount of natural disaster claims received by insurance companies has also increased dramatically. In order to more accurately assess insurance companies' underwriting capacity and coverage of catastrophe risks, since catastrophe risks in Los Angeles and Guangzhou are more frequent and more representative, this paper takes these two places as examples and establishes catastrophe risk assessment index and catastrophe risk severity scale based on the time series method. Next, We comprehensively consider the time value of money, and based on the insurer's surplus model, we aim to assist insurance companies in underwriting judgment and risk sharing decisions such as reinsurance, so as to provide a reasonable measurement standard for insurance companies.

Keywords: Catastrophe Risk, Time Sequence Model, Insurer Earnings Model.

1. Introduction

The increasingly frequent natural disasters caused by global climate change and population growth pose great challenges to the social economy and insurance industry. Due to its low frequency and high loss characteristics, catastrophe risk makes it difficult for traditional actuarial methods to effectively estimate potential losses, which highlights the urgency of establishing catastrophe insurance model. The catastrophe insurance model combines computer technology, mathematical method and geographic information system (GIS) to make probability prediction and loss simulation of potential disasters, providing a quantitative system for evaluating potential losses [1-2].

In the research field of catastrophe insurance model, many scholars have put forward different theories and models from their own professional perspectives. Qian Xiaotao and others focused on the probability of bankruptcy and assessed the risk solvency of insurance companies in the face of catastrophe [3]. Chao Wen et al. studied the cumulative catastrophe loss process [4] from the perspective of continuous time. Li Zhaopeng used the VaR model to construct a risk dispersion framework to realize the hierarchical risk management of [5]. Yang Zhiwei introduced the Copula function and the weather index, which provides an innovative method for the premium determination of crop insurance [6]. In the context of catastrophe swap, Shang Qin et al. explored the insurability boundary function [7]. Masako Ikefuji et al. combine stochastic economy and climate change to solve the general dynamic finite horizon economy-climate model (SDICE *) [8] with potential heavy-tail uncertainties and general utility functions. Andreas Lang et al. applied a large set of initial conditions for a high-resolution single model, effectively updating the hazard component [9] of its risk model. However, many traditional models still have limitations, unable to dynamically consider temporal changes and environmental factors, and difficult to capture the nonlinear relationship of natural disasters [10].

Facing these challenges, future research need to further improve the dynamic adaptability and precision of the model to better serve the insurance industry and social development. This involves not only the mathematical and statistical method innovations of the model, but also a deep understanding of environmental changes and the impacts of human activities. Based on the thorough analysis of the existing catastrophe insurance model, this paper studies the extreme weather events in Los Angeles and Guangzhou by constructing the disaster risk assessment index and time series based risk severity analysis. At the same time, considering the time value of capital, combined with the surplus model of insurance companies, to provide accurate risk assessment and capital allocation decision support for insurance companies. These methods aim to improve the accuracy of catastrophe risk and optimize risk management strategies.

2. Model building

2.1. Time-series analysis based on ARIMA

The time series model is a statistical model that synthesizes the principles of analogy (autoregression) and inertia (moving average). It is capable of forecasting future data points based on given historical data and time-varying trends, making it a comprehensive and accurate approach. Hence, this paper adopts the time series model, utilizing nearly 20 years of historical data from the Los Angeles and Guangzhou regions to forecast precipitation, temperature, and GDP for these areas, aiming to make a rational assessment and prediction of future meteorological data.

2.2. Data preprocessing

Before proceeding with data analysis, it is essential to ensure the usability of the data. The ARIMA (Autoregressive Integrated Moving Average) model is used for forecasting time series data, and it should be guaranteed to be stable. Without stability, the intrinsic patterns within the data cannot be captured effectively. Therefore, we assess the stability of our data by examining the results of the Augmented Dickey-Fuller (ADF) test, specifically through the analysis of the t-value, to determine if we can significantly reject the null hypothesis of a non-stationary series ($P<0.05$). If the data is found to be non-stationary, differencing analysis can be employed to stabilize the data.

2.3. Three steps of the ARIMA model building

Autoregression (AR). Describe the relationship between current and historical values and use variable historical time data to predict itself. The order of the AR model is denoted as the p-value. The formula is as follows:

$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \xi_t \tag{1}$$

Difference (I). The goal of the difference is to transform a non-stationary sequence into a stationary sequence, with the order of the difference noted as the d value. The formula is as follows:

$$\text{d\_order\_y} = (1 - \mathrm{B})^{\mathrm{d}} \mathrm{y_t} \tag{2}$$

Moving average (MA). The moving-average model focuses on the cumulative effect of the error terms in the autoregressive model. The order of the MA model is denoted as the q-value. The formula is as follows:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{3}$$

Finally get

$$Y_t = \sum_{i=1}^{p} \varphi_i x_{t-i} + \alpha_t - \sum_{j=1}^{q} \theta_j \alpha_{t-j} \tag{4}$$

$$X_t \sim I(d) \tag{5}$$

$$Y_t = \Delta^d X_t \tag{6}$$

3. Establishment of indicators

3.1. Degree of precipitation injury

Precipitation Damage Extent is an assessment index centered around the annual total precipitation and the harm it causes to a region. Studies have demonstrated that the amount of precipitation directly influences the frequency and intensity of natural disasters such as floods, droughts, hurricanes, and storms, or indirectly affects geological disasters through means such as water infiltration and soil liquefaction. Therefore, based on historical data, we have constructed a time-series model.

Firstly, to ensure the data's usability, we have verified the stability of the data through the Augmented Dickey-Fuller (ADF) test.The results for Guangzhou are shown in Table 1 and the results for Los Angeles are shown in Table 2.

Table 1. The ADF test table for the precipitation Guangzhou.

ADF Check list							
variable	Differential order	t	P	AIC	critical value		
					1%	5%	10%
Split again	0	-3.875	0.002***	357.817	-3.833	-3.031	-2.656
	1	-4.061	0.001***	342.154	-3.889	-3.054	-2.667
	2	-3.726	0.004***	328.615	-3.924	-3.068	-2.674

Note: * * *, * * and * represent the significance levels of 1%, 5% and 10%, respectively

Table 2. The ADF test table for the precipitation Los Angeles.

ADF Check list							
variable	Differential order	t	P	AIC	critical value		
					1%	5%	10%
Split again	0	-6.638	0.000***	395.218	-3.7	-2.976	-2.628
	1	-3.638	0.005***	383.066	-3.809	-3.022	-2.651
	2	-4.533	0.000***	370.939	-3.833	-3.031	-2.656

Note: * * *, * * and * represent the significance levels of 1%, 5% and 10%, respectively

As can be seen, $P < 0.05$ means that the data is stable and time series prediction can be made.

Next, we performed the ACF and PACF tests, and the results are shown in the following figure:

It can be concluded that in the process of evaluating precipitation in the Los Angeles area, when we automatically search for optimal parameters based on the Akaike Information Criterion (AIC), the model result is an ARIMA model (2,0,1). The model exhibits a good fit with a coefficient of determination R^2 of 0.513, indicating satisfactory performance and fulfilling the requirements. In contrast, for the Guangzhou area, the model result is also an ARIMA model (2,0,1), but with a higher coefficient of determination R^2 of 0.746, demonstrating a better performance and meeting the required standards.

Table 3. Pridictions of precipitat.

year	Los Angeles	Guangzhou
2021	6810.00	5985.44
2022	5679.10	4826.97
2023	9989.48	3926.75
2024	9780.83	3528.96
2025	8346.25	2644.34
2026	8716.47	1880.41
2027	9238.61	1101.60

Finally, for Guangzhou, China and Los Angeles, Table2 gives the forecast values for the next 7 periods,The relevant results are shown in Table 3.

3.2. Temperature damage index

The Temperature Damage Index is an evaluative indicator focused on the harm caused by the annual average temperature to a region. Research indicates that higher temperatures lead to an increase in extreme weather events such as heatwaves, drought, and severe storms, or accelerate the melting of glaciers and thermal expansion of oceans, increasing the risk of floods and storms in coastal areas; it can also exacerbate the urban heat island effect, increasing energy consumption and urban vulnerability. Therefore, in conjunction with historical data, we provide forecast values for the next seven periods (the forecasting method refers to [1], which is not elaborated here). Specific values are shown in Table 4 as follows:

Table 4. Average temperature prediction Fig.

year	Los Angeles	Guangzhou
2021	24.03	28.23
2022	28.13	29.16
2023	29.04	27.82
2024	28.07	26.44
2025	26.03	27.68
2026	28.03	25.56
2027	27.00	27.65

3.3. GDP assessment index

GDP is an indicator measuring the total volume of economic activity of a country or region, and it is also an important measure to assess the impact of catastrophic events on the economy. After a catastrophic event, a reduction or decline in GDP may indicate an interruption or decrease in economic output, production, and service activities; conversely, capital investment, infrastructure repair, and human resource activities involved in reconstruction and recovery efforts can have a positive impact on the economy and contribute to an increase in GDP. Therefore, observing changes in GDP is also a key factor in measuring catastrophic risk and insurability. This question adopts a time series approach, combined with historical data, to provide forecast values for the next seven periods (forecasting method refers to [1], details not repeated here). Specific values are shown in Table 5as follows:

Table 5. GDP assessment index

year	Los Angeles	Guangzhou
2021	77515.80	114409.53
2022	78493.68	117760.70
2023	79158.62	121205.09
2024	79610.77	124742.72
2025	79918.22	128373.58
2026	80127.28	132097.67
2027	80269.44	135915.00

3.4. The amount of catastrophe insurance claims

Catastrophic insurance claim amounts refer to the compensation paid by insurance companies to the insured or beneficiaries under an insurance contract due to losses caused by catastrophic events. This metric reflects the actual occurrence of catastrophic risks and is a key determinant in calculating the profits of insurance companies. In this study, we combine historical data to establish a BP (Back Propagation) neural network model:

Firstly, we determine the number of nodes in the hidden layer.

$$h = a + \sqrt{m + n} \tag{7}$$

Herein: h represents the number of nodes in the hidden layer, m denotes the number of nodes in the input layer, n signifies the number of nodes in the output layer, and a is a regulating constant between 1 and 10.

Subsequently, we apply the Delta Rule for general learning:

$$\Delta W = -\vartheta E' \tag{8}$$

$$\frac{\delta E'}{\delta \omega} = \frac{\delta \frac{1}{2}\left[t - f(\omega^T x)\right]^2}{\delta \omega} \tag{9}$$

Thereinto $\Delta W = -\vartheta E' = \vartheta \delta x, \delta = (t - y) f'(\omega^T x)$ Can be claimed by the insurance company Its operating mechanism is as shown in the following figure:

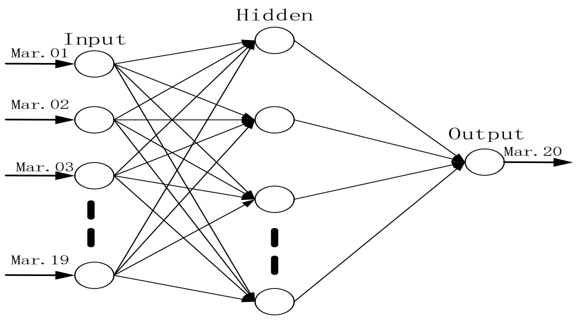

Figure 1. The BP neural network operating mechanism.

4. Catastrophe Risk Assessment (LPRA) model

4.1. The principle of EMW

In the EWM (Exponential Weighting Method) calculation, the weights of various indicators are determined by the information entropy of the indicators, which reflects the degree of variability of the indicators. The higher the information entropy, the stronger the significance of the indicator in the evaluation. Therefore, using EWM to determine the weights of each indicator in the calculation of the comprehensive catastrophe risk index is an objective method. After preprocessing the data, we obtained data from two regions over the last 20 years. Since these indicators are all positive, we have standardized the data:

$$X_i^{new} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \qquad i=1,2,3...... \tag{10}$$

In X_i^{new}Represents the data after homing,X_iRepresents the initial data,N Represents the number of data.

4.2. Calculation of Weights for Each Indicator in the Model Based on the Entropy Weight Method

The entropy weight method is an approach that utilizes the objective rules of the data itself to determine the importance weights of various evaluation indicators. In order to measure and assess the catastrophe risk level of a region, we take into account its risk to the entire socio-economic framework. Consequently, we need to calculate the information entropy for each of the four sub-indicators of the standardized EWM model: the Precipitation Damage Degree Index, the Temperature Damage Index, GDP, and Catastrophic Insurance Claims amount.

$$e_j = -\frac{1}{\ln n}\sum_{i=1}^{n} p_{ij}\ln(p_{ij}) \tag{11}$$

Where e_j is the j th index and the th information entropy, $p_{ij}$is the i th data weight in the j th index, and n is the number of data

Subsequently, calculate the information utility value and execute normalization to acquire the weights.

$$d_j = 1 - e_j \tag{12}$$

$$W_j = \frac{d_j}{\sum_{j=1}^{m} d_j} \tag{13}$$

Where d_j is the information utility value of the j th indicator,W_j is the weight corresponding to the j indicators, and m is the number of indicators.

Upon completing the aforementioned processing, we obtained the "Precipitation Damage Degree Index" search index at 19.9, "Temperature Damage Index" at 18.4, "GDP" at 45, and "Catastrophic Insurance Claims Amount" at 16.7.

4.3. Catastrophic Risk Level Severity Scale

We categorized the severity levels of disaster occurrences in two regions over the past 20 years into minor, moderate, and severe categories. By employing the K-means algorithm to cluster the dataset into three groups, we divided the dataset into the three clusters presented in Table 6.

Table 6. Catastrophe Risk Level Severity Scale.

Clustering categories	frequency	percentage%
trifling	11	36.667
moderate	11	36.667
serious	8	26.666
summation	30	100.0

5. Insurer's surplus model

5.1. Time Value of Money

The time value of money (TVM) refers to the concept that a certain amount of money in hand today is worth more than the same amount in the future due to its potential earning capacity. Given the time lapse between underwriting and claim payments, the value of currency will undergo corresponding changes. To more accurately establish models and evaluate underwriting risks, this paper takes into consideration the time value of money.

$$FV = PV \times (1 + i)^n \tag{14}$$

The term "PV" denotes the present value at the time of underwriting, while "FV" signifies the future value at the time of claim payments. The variable "i" represents the average interest rate of the investment.

5.2. Model building

$$U_t = u + rt - S_t - W \tag{15}$$

Assuming that event N_t follows a Poisson process with an intensity of $\lambda(\lambda > 0)$, the probability distribution function for event $N_t$can be denoted as $E(N_t) = \lambda t$. Here, λrepresents the frequency of occurrence, and for event $\lambda *$, the expected number of disasters occurring annually is 1.

For the claim amount of the n insurance claim, we assume that variable $x_n$follows an exponential distribution with parameter $\gamma(\gamma > 0)$. Thus, the density function for $x_n$can be expressed as:

$$f(x) = \begin{cases} \gamma e^{-\gamma x}, x \in [0, +\infty) \\ 0, x \in (-\infty, 0) \end{cases} \tag{16}$$

$$E(x) = \int_0^\infty x \gamma e^{-\gamma x} dx = 1/\gamma \tag{17}$$

Simplifying the model based on the given information, we have a situation where insurer $(r - c)t - S_t > 0$ sees a positive profit when condition r$> \lambda * \frac{1}{\gamma}$+c is met. This occurs when the annual premium collected exceeds the sum of the product of the number of disasters occurring per year and the expected annual claim amount, plus the annual cost of underwriting profit.

Research indicates that for general residential properties, catastrophic insurance premiums typically range from 0.1% to 2% of the property value. However, in high-risk areas prone to floods or earthquakes, the premium rates may be higher. For the sake of simplified calculations, in our pricing strategy, we will uniformly use 2% of the property value to calculate the insurance premium revenue.

6. Conclusions

Catastrophe risk is characterized by high loss, low probability and difficult to predict, and often involves natural disasters or major accidents. This risk can lead to a substantial increase in the compensation amount of insurance companies, affecting the financial stability and solvency of the company. At the same time, accurate catastrophe risk assessment can also help insurance companies to carry out more detailed risk management and reinsurance arrangements, so as to cope with the huge potential compensation pressure. In the above study, we established the catastrophe risk assessment index and the insurer surplus model in Guangzhou and Los Angeles respectively. In the end,for the sake of simplified calculations, in our pricing strategy, we will uniformly use 2% of the property value to calculate the insurance premium revenue.

References

[1] Tan Lezhi. Technology Empowers Catastrophe Model Construction [J]. China Banking and Insurance News, 2024, June 25, p. 008.

[2] Tan Lezhi. Catastrophe Protection System: Perfecting the Four "Puzzle Pieces" [N]. China Banking and Insurance News, 2024, July 2, p. 004.

[3] Chao W., Qian Xiaotao. Analysis of solvency for catastrophe insurance payments from the perspective of ruin probability—Based on the mixed distribution model [J]. Journal of Insurance Professional College, 2023, 37(5): 51-57.

[4] Chao W. Indifference pricing of catastrophe bonds based on a continuous-time dynamic model [J]. Journal of Beijing University of Chemical Technology (Social Sciences Edition), 2022, (1): 17-20.

[5] Li, Z. Research on the Catastrophe Risk Dispersion Mechanism of Agricultural Insurance in Guizhou Province [D]. Guizhou University of Finance and Economics, 2023.

[6] Yang, Z. Weather Index Insurance Design Based on Copula Method: A Case Study of Jiamusi [D]. Northwest A&F University, 2023.

[7] Shang Qin, Qin Yida. The Impact of Catastrophe Swaps on the Boundaries of Insurability for Catastrophic Risks: A Case Study of Super Typhoon "Hato" [J]. Operations Research and Management Science, 2024, 33(1): 132-137.

[8] Ikefuji M, Laeven RJA, Magnus JR, Muris C. Expected utility and catastrophic risk in a stochastic economy-climate model [J]. Journal of Econometrics, 2020, 214(1): 110-129.

[9] Lang A, Poschlod B. Updating catastrophe models to today's climate - An application of a large ensemble approach to extreme rainfall [J]. Climate Risk Management, 2024, 44: 100594.

[10] Biardeau LT, Sahli MD. Investigating the non-linear impacts of seven types of natural disasters on inbound tourism: Insights from the EM-DAT database [J]. Tourism Economics, 2024, 30(4): 900-923.