

Research on the Extreme Weather Prediction Model Based on Linear Regression and Its Application in Insurance Evaluation

Nanlong Sun *

School of Agricultural Engineering, Jiangsu University, Zhenjiang, China

* Corresponding author: 2901495645@qq.com

Abstract. This study presents a novel approach to predicting extreme weather events using a linear regression model and explores its application in the insurance evaluation process. The predictive model is developed by analyzing historical data on the frequency and intensity of extreme weather occurrences, such as floods, droughts, and hurricanes. The model employs cumulative data analysis to establish a linear relationship between the occurrence of extreme weather events and the passage of time. The results demonstrate a high correlation coefficient, indicating a reliable prediction capability. The study then extends the application of this model to the insurance sector by constructing an insurance evaluation model. This model assesses the feasibility of providing insurance services in areas prone to extreme weather by comparing insurance premiums with potential claims. The evaluation model incorporates the impact of extreme weather on three major industries: agriculture, manufacturing, and services, and assigns weights accordingly. Case studies of Japan and Florida, USA, are conducted to validate the model's effectiveness in real-world scenarios.

Keywords: Extreme Weather Prediction, Linear Regression, Insurance Evaluation, Cumulative Data Analysis, probability occurrence.

1. Introduction

With the increasing frequency of extreme weather, the exposure of property insurance in people's daily life is so increasing. To protect the property after the irresistible extreme weather, the purchase rate of property insurance has shown a rapid rising trend, but at the same time, a large number of claims have led to the profit and loss state of insurance companies is not optimistic, so how to achieve the sustainable development of property insurance has become an urgent issue to address. [1, 2].

To better mitigate the damage caused by extreme weather, it is imperative to construct models that predict the likelihood of such events occurring. This paper primarily analyzes the probabilities of three types of extreme weather events: floods, droughts, and hurricanes, and explores how these predictive models can be utilized to assess insurance risks and provide decision support for insurance companies.

2. Extreme Weather Prediction Model

2.1. Model Preparation

In this paper, the data we used mainly include the number of extreme weather events, the impact of extreme weather on the economy data and total population of the region over the years. These data were primarily sourced from the National Oceanic and Atmospheric Administration (NOAA) and Our World in Data. By conducting cumulative analysis of this data, we established a linear regression model to forecast the frequency of future extreme weather events.

Worldwide, the frequency and intensity of extreme precipitation and flood disasters continue to increase, posing a severe threat to the stability of urban agglomerations [4]. In order to better resist the damage caused by extreme weather, we need to build models to predict the probability of extreme weather. In this paper, we mainly analyze the probability of flood, drought and hurricane. When we analyze the prediction models for general situations, we do not need to consider the impact of volcanic eruptions, so the following data collection process, we collect the data of floods, droughts and

hurricanes over the past years. Since there are no direct statistics on the number of hurricanes worldwide over the past year, the percentage of hurricanes reported is twice the percentage of landslides reported, so we can infer the number of hurricanes over the past year based on the number of landslides [5]. All the extreme weather events are summarized in the Fig. 1.

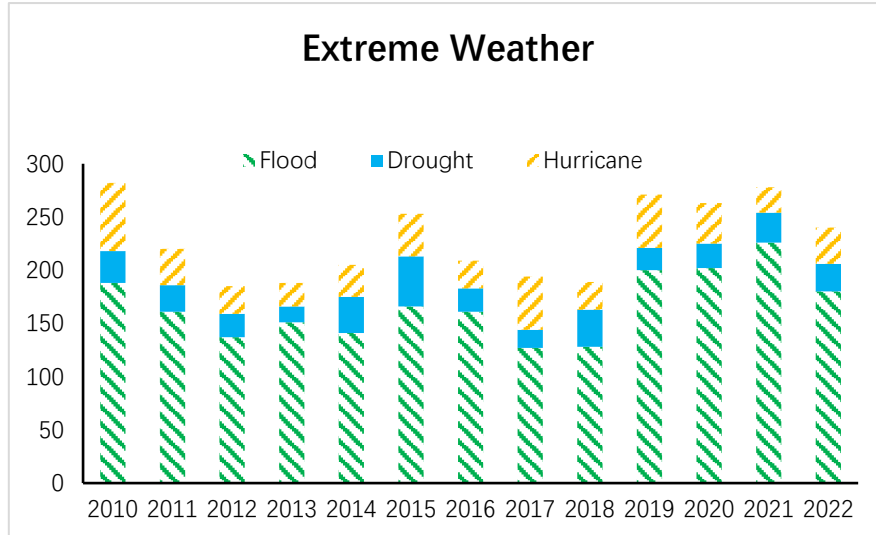


Figure 1. Extreme weather events

2.2. Parameter Determination

From the Fig. 1, we can clearly find the annual occurrence times of each extreme weather. Because only observing the data of each year cannot find the fitting curve with high correlation, we use accumulation to generate a new sequence.

$$y_k = \sum_{i=1}^k x_{ij} \quad (1)$$

Where x_{ij} is the number of the j th extreme weather events in year i ; y_k is total number of extreme weather events in the previous k years. Next, we carry out linear fitting for each group of data, and the fitting function relationship follows the principle that the closer the correlation coefficient is to one, the better the fitting effect will be. Fig. 2 is linear fits of the frequency of the three kinds of extreme weather.

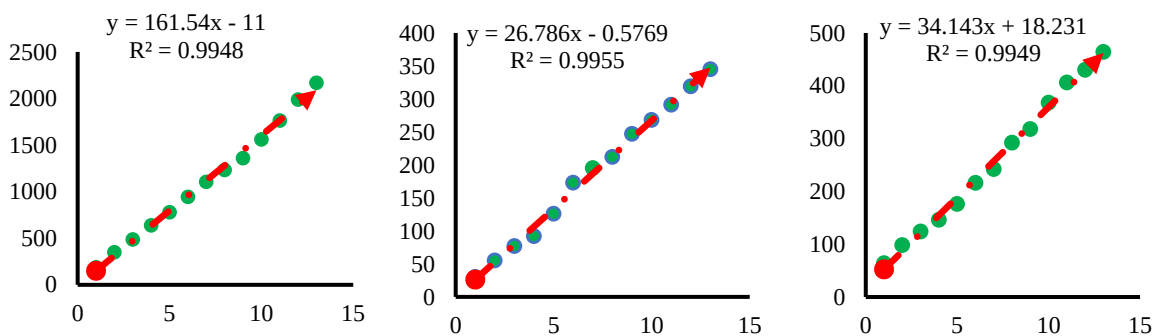


Figure 2. Linear fitting results (Floods, Drought and Hurricane)

In the figure above, the horizontal coordinate is the cumulative years, and the vertical coordinate is the number of global reports. It is not difficult to see from the linear fitting results in the figure above that the correlation coefficients are all greater than 0.99, and it can be approximated that the linear relationship is completely consistent.

The forecast model for the number of annual occurrences of the three types of extreme weather in the future is shown below.

$$ef_n = f_n - f_{n-1}. \quad (2)$$

$$ed_n = d_n - d_{n-1}. \quad (3)$$

$$eh_n = h_n - h_{n-1}. \quad (4)$$

Where ef_n , ed_n , eh_n respectively are the number of floods, the number of droughts, the number of hurricanes in year n .

According to the prediction model of the future occurrence frequency of the three kinds of extreme weather determined by (2), (3) and (4), when we need to calculate the occurrence frequency of extreme weather in the next year, we only need to bring in the data of the previous year to obtain the required calculation results.

3. Insurance Evaluation Model

When evaluating whether a region can be insured, it is necessary to compare the insurance premium with the insurance loss rate. We will start from the three major industry types to evaluate the cost. The impact of extreme weather on the insurance cost of property insurance purchased by the three major industries is considered separately. This phenomenon is common in agricultural insurance, for example, when Hazel modified his rainfall insurance in 1986 to reflect the prevailing conditions.

In order to make the coverage of the model as wide as possible and avoid some influencing factors that have little influence on the result, we roughly divide the buyers of property insurance into three major industries according to the industrial proportions of most countries.

$$W_a:W_m:W_s = 5\%:31\%:64\%. \quad (5)$$

Where W_a , W_m , W_s respectively are the proportion of agriculture output to GDP, the proportion of manufacturing output to GDP, the share of services output to GDP.

In Fig. 3, we will list the basic insurance types of the three industries and the property insurance evaluation objects we selected for each industry type.

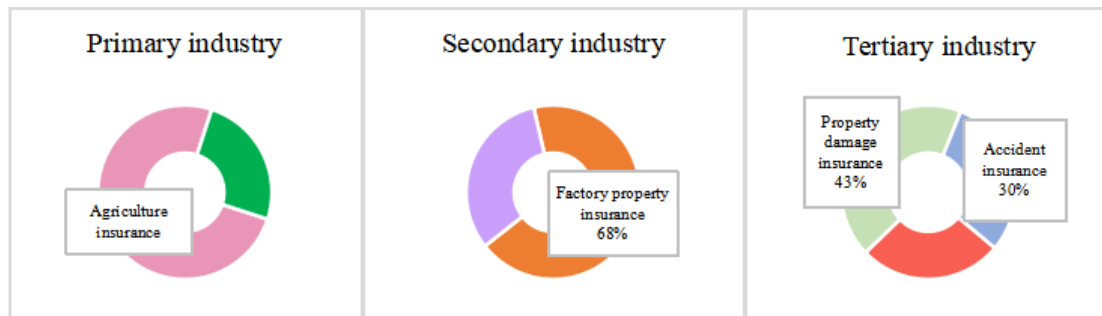


Figure 3. The proportion of primary insurance

From the figure above, we can see that the main types of property insurance for the three industries are agriculture insurance, factory property insurance and property damage insurance.

3.1. Primary Industry

The main property insurance of the primary industry is agriculture insurance, and the main payouts of agriculture insurance are compensated through the following ways.

$$Pa_n = fa_n \times l. \quad (6)$$

$$l = \frac{S_f}{S_o}. \quad (7)$$

Where Pa_n is the agriculture insurance payout. fa_n is the premium for agriculture insurance in year n . l is the relative rate of loss after experiencing extreme weather. S_f is the area destroyed after extreme weather. S_0 is the initial area.

Agriculture insurance premiums fluctuate with the number of extreme weather events, so the impact of extreme weather needs to be taken into account when calculating the annual premium for agriculture insurance. In most areas, agriculture is mainly affected by flood and drought, and the relative impact of other extreme weather on agriculture is relatively small, so only two extreme weather types of floods and drought are considered in the calculation of insurance premiums.

Before we calculate the premium, we firstly calculate the extent to which extreme weather affects the premium.

$$Pa_n = fa_n \times l. \quad (8)$$

$$Pa_n = fa_n \times l. \quad (9)$$

Where qa is the extreme weather impact factor. Za_n is the frequency of extreme weather in year n . α_1 , α_2 respectively are the weights of floods, droughts. According to the above expression, we can calculate the change the change of insurance premium with the number of extreme weather events per year.

$$Fa_n = Fa_{n-1} \times (1 + qa). \quad (10)$$

Where Fa is the cost of agriculture insurance.

3.2. Secondary Industry and Tertiary Industry

The secondary industry and the tertiary industry are the main components of the national economy. When these industries are affected by extreme weather, the national economy will have obvious fluctuations. Therefore, the impact of extreme weather on insurance premiums can be measured through the change of GDP. In Fig. 4 we fit the economic effects of floods and hurricanes respectively.

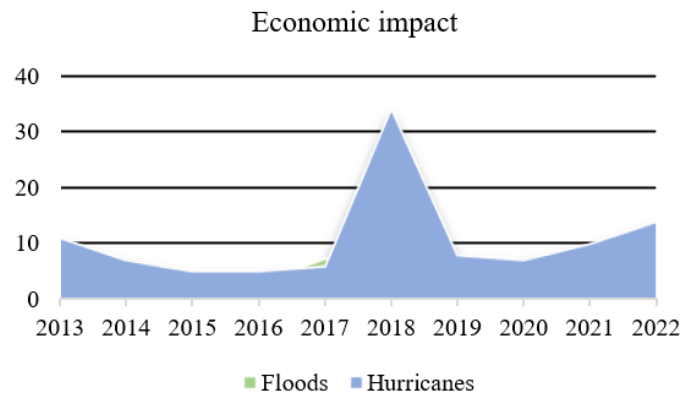


Figure 4. Economic effects of floods and hurricanes

Since the fitting relationship cannot be found from the data of the economic affected percentage over the past years alone, we still use the cumulative method for linear fitting. However, it can be seen from the results of linear fitting that not only the correlation coefficient is small, but also the fitting function relationship is relatively complex. Therefore, we still use the extreme weather prediction model to represent the occurrence times of extreme weather in the following evaluation process of property insurance in the secondary and tertiary industries. In the process of analysis, we still consider the influence factors of extreme weather.

$$qm = \frac{zm_n}{zm_{n-1}}. \quad (11)$$

$$qs = \frac{Zs_n}{Zs_{n-1}}. \quad (12)$$

$$Zm_n = Zs_n = \beta_1 \times ef_n + \beta_2 \times eh_n. \quad (13)$$

Where qs , qm respectively are the influencing factors of floods and hurricanes. Zm_n , Zs_n respectively are the frequency of floods and the frequency of hurricanes in year n . β_1 , β_2 respectively are the weights of floods and hurricanes. Since the main types of property insurance in the secondary and tertiary industries are not the same, we need to discuss them separately when calculating premiums and claims.

3.2.1. Factory Property Insurance

In the processing industry, the amount and premium of general claims are higher than those in agriculture, so the calculation of the compensation of the factory property insurance is simpler, the calculation is as (14).

$$Pm_n = L_m. \quad (14)$$

Where Pm_n is the factory property insurance payout; L_m is the actual amount of damage. It should be emphasized that the above formula is only valid when the actual loss amount is less than the rated amount of insurance, and this situation is valid when most property insurance claims are paid, so we use this formula to calculate in the modeling process. Based on the previous extreme weather prediction model we can list the premiums for factory property insurance.

$$Fm_n = Fm_{n-1} \times (1 + qm). \quad (15)$$

Where Fm_n is the cost of factory property insurance.

3.2.2. Property Damage Insurance

The tertiary industry is mainly dominated by the service industry, and there are many kinds of insurance in the service industry. However, in the category of property insurance, property damage insurance is the most common insurance, so the analysis of the tertiary industry mainly takes property damage insurance as the research object. The principle amounts payable for property damage insurance are calculated as (16).

$$Ps_n = \frac{Fs_n \times L_s}{L_a}. \quad (16)$$

Where Ps_n is the property damage insurance payout; L_s is the actual amount of damage in the service industry; L_a is the actual value of all items. The change law of property damage insurance premiums of the tertiary industry affected by the extreme weather is consistent with the change law introduced above, and the specific expression is as follows.

$$Fs_n = Fs_{n-1} \times (1 + qs). \quad (17)$$

Where F_n is the cost of property damage insurance.

Next, we need to make a comprehensive evaluation of the scores of the three industries. In the above analysis, we solved all the expressions of all the variables, and then we need to superimpose all the variables according to the weight distribution of the three industries to get the final extreme weather evaluation model. Our model compares the cost of insurance with the payouts to determine whether areas affected by extreme weather are suitable for providing insurance services.

$$IEM = \frac{F_n}{P_n}. \quad (18)$$

$$F_n = 5\% \times Fa_n + 31\% \times Fm_n + 64\% \times Fs_n. \quad (19)$$

$$P_n = 5\% \times Pa_n \times pa_n + 31\% \times Pm_n \times pm_n + 64\% \times Ps_n \times ps_n. \quad (20)$$

Where F_n is the sum of all insurance costs. P_n is the sum of all insurance payouts. IEM is index value of extreme weather evaluation model. Where pa_n , Pm_n , ps_n respectively are the corresponding occurrence of extreme weather.

Through the above evaluation model, the size relationship between the insurance premium and the insurance payout can be compared. If the value of the model is greater than one, insurance services can be provided in the region, but if the value of the model is less than one, insurance services cannot be provided in the region.

3.3. Application of Insurance Evaluation Model

To test the validity of our model, we will select two regions affected by extreme weather on different continents and use the results of our model to determine whether these regions are suitable for providing insurance services.

3.3.1. Research Object I: Japan

As we all know, Japan has a high frequency of extreme weather due to its special geographic location, and the three major insurance companies in Japan are also committed to solving the problem of the coverage of property insurance, so it is more meaningful to use our model to analyze whether to provide insurance services. Some of the major extreme weather events in Japan and their prevalence are shown in Fig. 5.

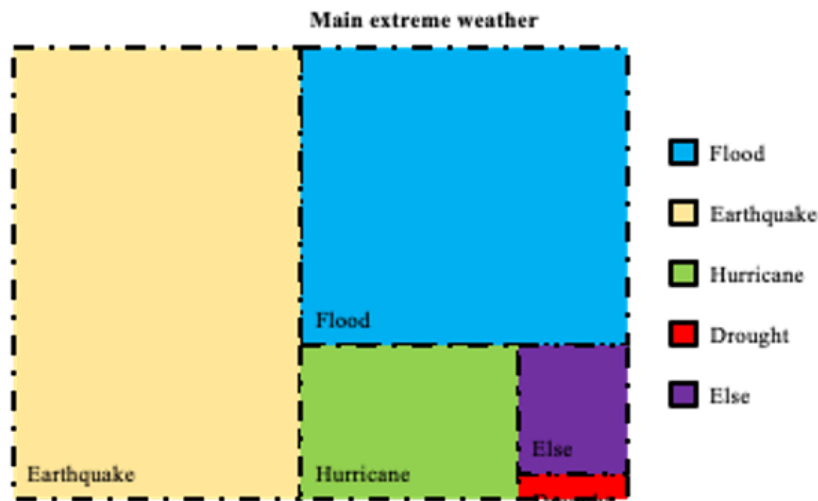


Figure 5. Main extreme weather in Japan

As you can see from the figure above, the major kinds of extreme weather are floods and hurricanes. Although, the volcanoes in Japan are active, the probability of eruption is very low, so the impact of volcanic eruption can be ignored for the time being. We are looking at the coastal area of Japan, which are the most vulnerable to extreme weather. In the figure below, we show the statistical results of the number of extreme weather events in Japan every ten years.

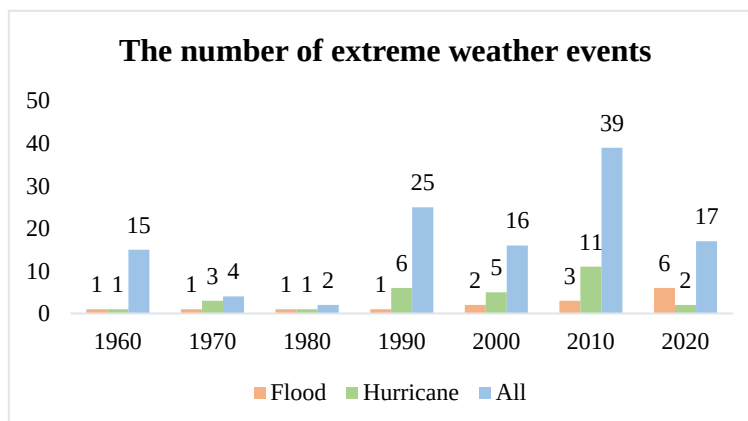


Figure 6. The number of extreme weather events in Japan

Since these data did not confirm to any fitting relationship, we still adopted the summation method to find the linear relationship of the data, and combined with the extreme weather prediction model, calculated the probability of flood and hurricane occurrence in Japan as (21).

$$pj_f = 5, pj_h = 13. \quad (21)$$

Where pj_f is the number of floods in Japan per year. pj_h is the number of hurricanes in Japan per year. On Japan's eastern seaboard, there is a high density of urban clusters, so the effects of extreme weather can be considered omni – directional. In particular, when the annual probability of extreme weather is greater than one, the total value of the insurance cost will be less than the total value of the insurance payout, so the coastal areas of Japan cannot provide insurance services. The forecast is in line with the situation presented by Tokyo Marine Nippon Motor Insurance Company.

3.3.2. Research Object I: America

As one of the largest countries in the world, the United States also has a wide variety of extreme weather. Therefore, we choose Florida as the research object in the analysis process, because the extreme weather phenomena in this state are more significant. Wildfires are a common type of extreme weather in the United States, so they need to be considered when calculating premiums and payouts for property insurance.

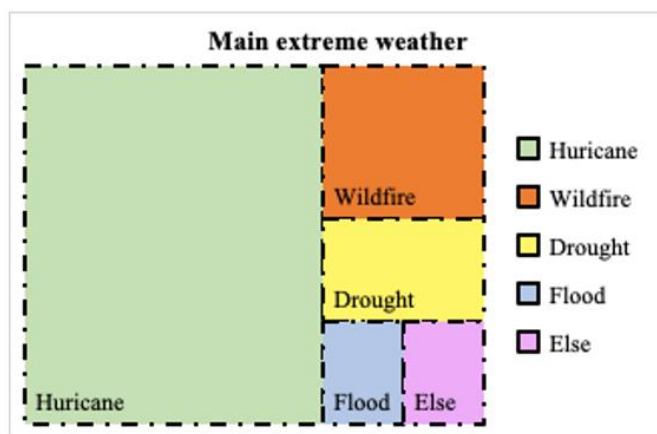


Figure 7. Main extreme weather in Florida, USA

Wildfires have an impact on all three industry types, so the impact on property insurance premiums can be seen in other types of extreme weather prediction models. The main types of extreme weather in Florida, USA, and their occurrence are shown in Fig. 7. It illustrates how often wildfires occur in the United States. The property insurance of the secondary industry and the tertiary industry are all affected by hurricanes, wildfires and floods. As can be seen from the above figure, the impact of

wildfires and hurricanes is greater than that of floods. Therefore, the following conclusion can be drawn even if the probability of flood is very low.

$$\Delta = \frac{Fm_n \times qm + Fs_n \times qs}{Pm_n \times pm_n + Ps_n \times ps_n} \quad (22)$$

Based on the above analysis, we can easily draw the conclusion that is less than one. The analysis of the primary industry is still the same because the extreme weather will have an impact on agriculture, and major extreme weather occurs more than once a year, so the insurance payout will be far greater than the insurance cost. Based on the above analysis, Florida is not suitable for providing property insurance services.

4. Summary

In this research, we have developed a comprehensive model for predicting extreme weather events and assessing their impact on the insurance industry. Our approach to model selection was guided by the need for broad coverage and ample data availability, ensuring that our model's predictions are more aligned with actual values, thus minimizing error. During the modeling process, we incorporated variables that consider the trends of extreme weather, establishing equivalent relationships through quantitative analysis, which enhances the sensitivity of our model. Our model's reasoning is progressive, with strong correlations between the models. The selection of multiple benchmarks, each backed by reputable universities, further strengthens the robustness of each model.

However, there are limitations to our model. Despite choosing factors with strong universality, we faced challenges in obtaining data for a small subset of regions, which negatively impacted the accuracy of our results. Future research will focus on addressing these limitations and exploring additional data sources to improve the model's performance.

References

- [1] Sun Zhengcheng, Huang Qiuyi. How Environmental Uncertainty Affects the Operating Performance of Insurance Companies. 2023 China International Conference on Insurance and Risk Management.
- [2] Berglez P, Al-Saqaf W. Extreme weather and climate change: social media results, 2008 – 2017 [J]. Environmental Hazards, 2021, 20 (4): 382 - 399.
- [3] Skees J R. Challenges for use of index-based weather insurance in lower income countries [J]. Agricultural Finance Review, 2008, 68 (1): 197.
- [4] Fang W, Xue Q, Shen L, et al. Survey on the application of deep learning in extreme weather prediction [J]. Atmosphere, 2021, 12 (6): 661.
- [5] Lubchenco J, Karl T R. Predicting and managing extreme weather events [J]. Physics today, 2012, 65 (3): 31 - 37.
- [6] Dwivedi R, Prasad K, Mandal N, et al. Performance evaluation of an insurance company using an integrated Balanced Scorecard (BSC) and Best-Worst Method (BWM) [J]. Decision Making: Applications in Management and Engineering, 2021, 4 (1): 33 - 50.