

Assessing Extreme Weather Risks for Insurance Underwriting with ARIMA and Entropy Weight Method: Insights from Brazil and India

Yiheng Xu^{1,*}, Leyi Wang², Jiankun Liang³

¹School of Finance, Nankai University, Tianjin, China, 300350

²School of Agricultural Economics and Rural Development, Renmin University of China, Beijing, China, 100872

³College of Economics and Management, China Agricultural University, Beijing, China, 100083

*Corresponding author: 2113605@mail.nankai.edu.cn

Abstract. As the frequency of extreme weather events increases around the world, it has had a great impact on the insurance industry, especially in Brazil and India, which suffer frequent natural disasters. Insurance plays a significant role in risk sharing and post-disaster reconstruction, but due to the uncertainty of profitability, whether insurance companies should underwrite insurance policies in these areas has become an important issue facing them. Given this, this paper establishes an insurance assessment model based on extreme weather based on ARIMA and entropy weight method and combined with TOPSIS. By evaluating the potential loss and profit of the insurance company after the insurance, to judge whether the insurance company should be insured in the field of extreme weather insurance in a given area. Last, this paper applied our model to Brazil and India, countries that are experiencing severe meteorological disasters, and concluded that it is wise for insurance companies to underwrite policies there.

Keywords: ARIMA, Entropy Weight Method, TOPSIS, Insurance Assessment Model.

1. Introduction

Living in the global village, we should be blessed for what we have gained from the planet. However, the health of Mother Earth is not that positive. From drought, flood, and storm to wildfire and extreme temperatures, weather events are sweeping across the globe. According to WMO, the worldwide economic annual losses due to hydrological, meteorological, and climatological hazards have grown sevenfold during 1970-2019^[1]. The growing threat posed by extreme weather events has exaggerated the calls for property preservation, which fall on insurance. If correctly designed and well-functioning, insurance limits financial uncertainty following disasters and may incentivize the uptake of property-level resilience measures, particularly when insurance premiums are risk-based^[2]. However, more and more insurance companies find it harder to decide whether they should underwrite policies in certain areas. They lack a practical method to judge whether they can benefit from this field because the insurance companies will only consider whether they are profitable to judge whether they are insured, and do not care about the objective loss. Thus, exploring a practical method to help insurance companies make better decisions under the threat and uncertainty brought by extreme weather events is a sobering call to action.

This study established an insurance assessment model based on ARIMA and the entropy weight method. Here, it used the ARIMA model to predict the loss of each disaster, and the annual premium income situation. Such an approach is feasible. Ahmed et al (2022) used this method to evaluate the losses in the oil insurance sector^[3]. In addition, Yibo Feng (2021) established an ARIMA prediction model and used this model to predict the life insurance premium income^[4]. At the same time, it used the entropy weight method to weighted and sum the loss of each disaster and calculate the total loss. Of course, this method is also desirable. Liu et al (2020) evaluated the risk of flood disasters in Bangladesh, India, and Myanmar through this method^[5].

Subsequently, this study predicted the specific number of occurrences of these disasters in the target year according to the ARIMA model after collecting historical data on these disasters. After predicting these values, it collected the amount of losses caused by each single disaster and determined the weights of the five disasters using the entropy weight method. Then, the total loss of natural disasters in the target year can be obtained by linear summation of the product of the frequency and weight of each disaster in the target year. In insurance company income, namely insurance premium respect, this study obtained data on the total global premiums per year from 2009 to 2023 from the Global Insurance Market Report, then used the ARIMA model to predict premium status for the target year. After obtaining the loss value and premium for the target year, the study used the min-max normalization in the TOPSIS method combined with the unit Length Normalization programming method to process them. After the establishment of this insurance assessment model, this study conducted a relevant discussion in two countries, Brazil and India, for they are experiencing severe meteorological disasters.

2. Establish the Risk-loss Model based on the ARIMA Model

To evaluate the potential loss caused by extreme weather events, it is necessary to identify the total occurrence times and the amount of damage caused by a single event.

2.1. Data Pre-processing

This paper collected data on global disasters from EM-DAT. Among the aggregate 14 types of disasters, it picked out six indicators that belong to extreme weather events and added up their occurrence times.

Among these six types of extreme weather events, it further excluded the “glacial lake outburst flood” that only happened four times in 24 years, which can be considered a black swan event. Thus, the study screened out five kinds of disasters and counted the total occurrence times, which are shown in Table 1.

Table 1. The total number of disasters

Name of the disaster	Total number
drought	399
extreme temperature	488
flood	3999
glacial lake outburst flood	4
storm	2523
wildfire	298

2.2. ARIMA Model Used to Predict Occurrence Times

As the time series model holds that the development trend of things has continuity, it uses this continuity to predict the future trend. Given the past annual data, it can make a proper projection. Here, X_{it} represents how many times a disaster has happened in year t while 1-5 represents the five extreme weather events (drought / extreme temperature/flood/storm/wildfire).

Initially, this study tests the stationarity of these five series to ensure that it can implement the ARIMA model. This paper specifies that when $p < 0.005$, it can be regarded as a stationary time series. According to Eviews statistics, the results are shown in Table 2.

Table 2. The original sequence statistics

Statistics	Drought	Extreme temperature	Flood	Storm	Wildfire
The t-statistics	-1.1207	-5.5954	-3.5291	-4.7968	-2.2804
p-value	0.6875	0.0001	0.0169	0.0009	0.1864
stationarity	NO	YES	YES	YES	NO

It can be seen from the results that the drought series and the wildfire series are not stationary. For the general time series, this study uses the difference method to transform it into a stationary time series. Results are shown in the following Table 3.

Table 3. First-order differenced statistics

Statistics	D1.drought	D1.extreme temperature	D1.flood	D1.storm	D1.wildfire
The t-statistics	-8.2180	-9.0110	-4.8541	-7.3193	-11.7730
p-value	0.0000	0.0000	0.0010	0.0000	0.0000
stationarity	YES	YES	YES	YES	YES

As the time series are all stationary after the first-order differential processing, it can construct the ARIMA model to predict its future performance. Here, we choose the third variable “flood” as an example to illustrate its method:

$$X_{3t} = c + \varphi_1 X_{3,t-1} + \dots + \varphi_p X_{3,t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (1)$$

In which:

- ε_t represents the white noise with a variance of σ_ε^2 .
- $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_p]^T$ represents the autoregressive coefficient.
- $\theta = [\theta_1, \theta_2, \dots, \theta_q]^T$ represents the moving average coefficient.

Then, this paper determines the range of p and q by analyzing the autocorrelation function and partial autocorrelation function graphs, as shown in Figure 1. The autocorrelation graph usually decays abruptly in the time series, and if all cases in the graph that lag behind q_0 are close to 0 within the confidence interval, the order of the MA part can be considered less than or equal to q_0 . Similarly, a partial autocorrelation function image usually drops to 0 abruptly at a certain order, and if all the lags after lagging p_0 are close to 0 within the confidence interval, the order of the AR parts can be considered to be p_0 .

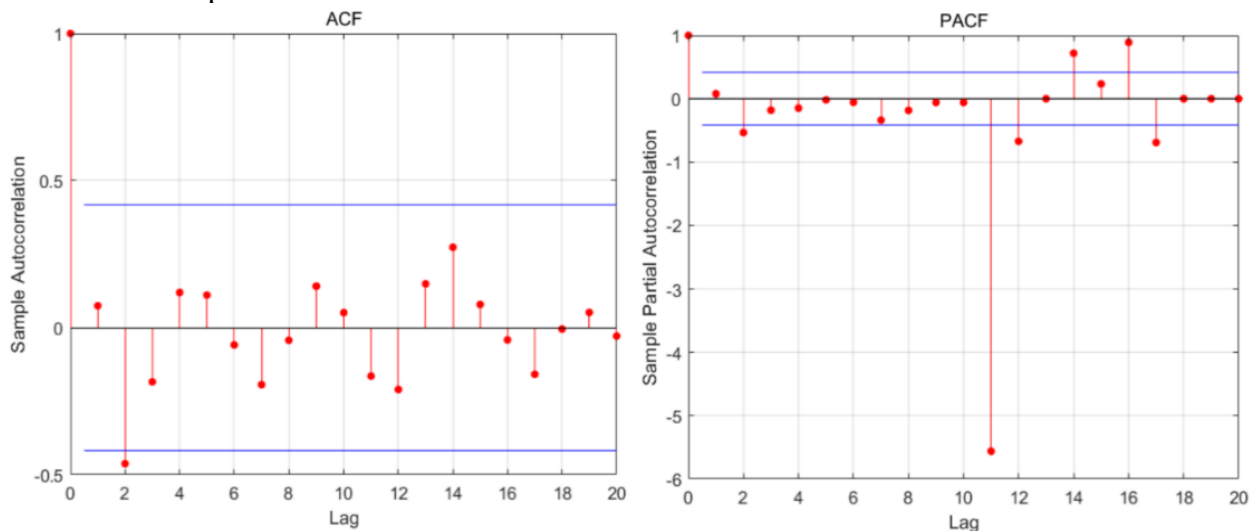


Figure 1. The ACF and PACF Images

To determine the values of p and q more accurately, this paper sets both p_0 and q_0 to 3. In this ARIMA model, it takes 3 for the lag value p of the observed value and q for the error term while the difference order d is 1.

Bringing in the data from 2000 to 2023, it can get the projections from 2019 to 2023 under the ARIMA model, as shown in Figure 2.

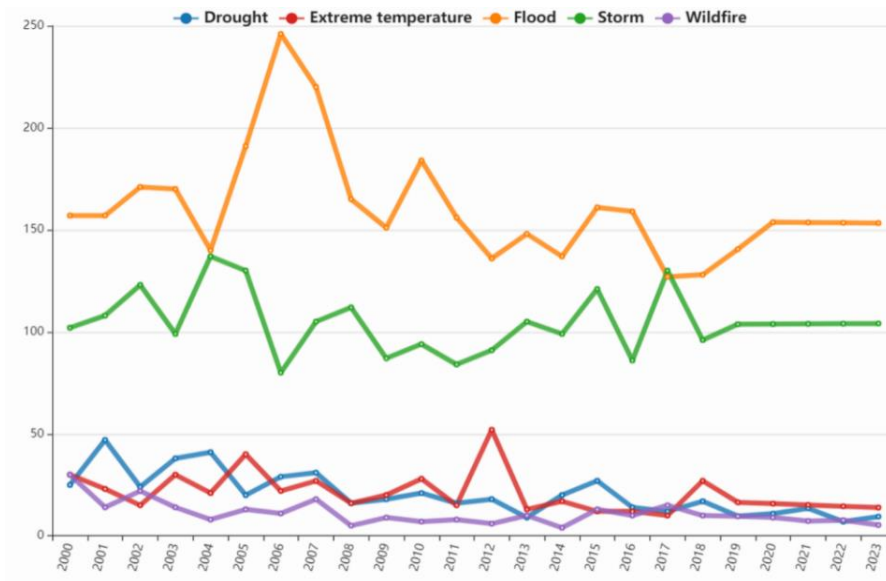


Figure 2. Annual Occurrence Times of Each Disaster

2.3. Determine the Weights by EWM

Since has screened out 5 disasters in the previous section and their occurrence times in year t are X_{1t} to X_{5t} , this paper chooses a linear regression model to combine these five variables in a bid to predict Y_t , which represents the extent of damages due to extreme weather events in year t .

$$Y_t = w_1X_{1t} + w_2X_{2t} + w_3X_{3t} + w_4X_{4t} + w_5X_{5t} \tag{2}$$

Considering that a disaster with a larger destructive force should have a greater impact on the risk-loss model, this paper uses the average single damage of each disaster combined with the entropy weight method to determine the coefficients in the linear regression model.

First, it uses $average_1 \dots average_5$ to represent the average damage of each disaster from 2000 to 2023. The sorted data are listed in Table 4.

Table 4. Average damage

Disaster	Notation	Damage
drought	$average_1$	1966987
Extreme temperature	$average_2$	1611376
flood	$average_3$	812502
storm	$average_4$	1594770
wildfire	$average_5$	1152592

Then, it uses the entropy weight method to determine the weight of the previous five variables, which are $w_1 \dots w_5$. Since these 5 variables must not have any correlation, we can use the entropy weight method. AVG_i is the standardized form of $average_i$.

$$AVG_i = 0.998 \frac{average_i - \min\{average_1, \dots, average_5\}}{\max\{average_1, \dots, average_5\} - \min\{average_1, \dots, average_5\}} + 0.002 \tag{3}$$

$$P_i = \frac{AVG_i}{\sum_{j=1}^5 AVG_j} \rightarrow e_i = -P_i \ln(P_i) / \ln(5) \rightarrow g_i = 1 - e_i \rightarrow w_i = \frac{g_i}{\sum_{j=1}^5 g_j} \tag{4}$$

The results obtained from the mathematical derivation procedure described above are shown in Table 5.

Table 5. Weights determined by damage

w_1	w_2	w_3	w_4	w_5
0.1936	0.1965	0.2102	0.1967	0.2030

Therefore, it comes to the equation:

$$Y_t = 0.1936X_{1t} + 0.1965X_{2t} + 0.2102X_{3t} + 0.1967X_{4t} + 0.2030X_{5t} \quad (5)$$

3. Develop the Insurance Model based on the ARIMA Model

This paper obtained data on the total global premiums per year from 2009 to 2023 from the Global Insurance Market Report and then used the ARIMA model to predict future premiums Z_t .

Firstly, it carried out the stationarity tests on each of these indicators to ensure the applicability of the ARIMA model. Relevant results are shown in Table 6.

Table 6. The stationarity test

Statistics	Z_t	$D1.Z_t$	$D2.Z_t$
The t-statistics	-0.4151	-1.627	-3.6390
P-value	0.8794	0.4418	0.0222
Stationarity	NO	NO	YES

As can be seen from Table 6, the series is stationary in the second order. After testing the second-order differential stationarity of the series, it builds another ARIMA model :

$$Z_t = c_0 + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \theta_1 \varepsilon'_{t-1} + \dots + \theta_q \varepsilon'_{t-q} + \varepsilon'_t \quad (6)$$

The process for analyzing autocorrelation and partial autocorrelation images is the same as above, as shown in Figure 3. Here, this paper sets the lag value p of the observed value and the lag value q of the error term to 3, while the difference order d is 2.

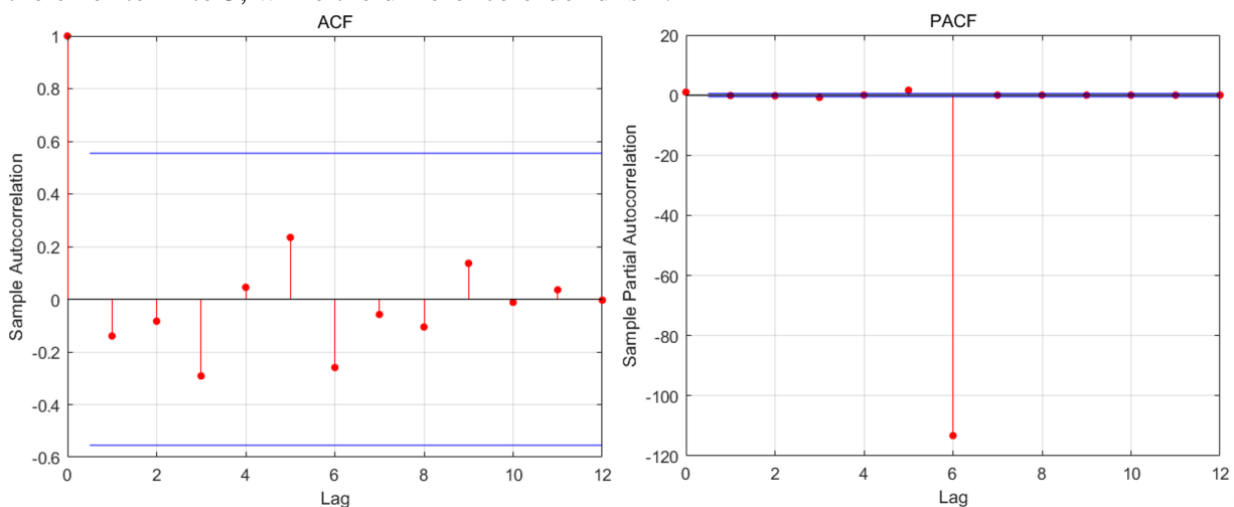


Figure 3. The ACF and PACF Images

Bringing in the past 15 years' data can predict future premiums, as shown in Figure 4.

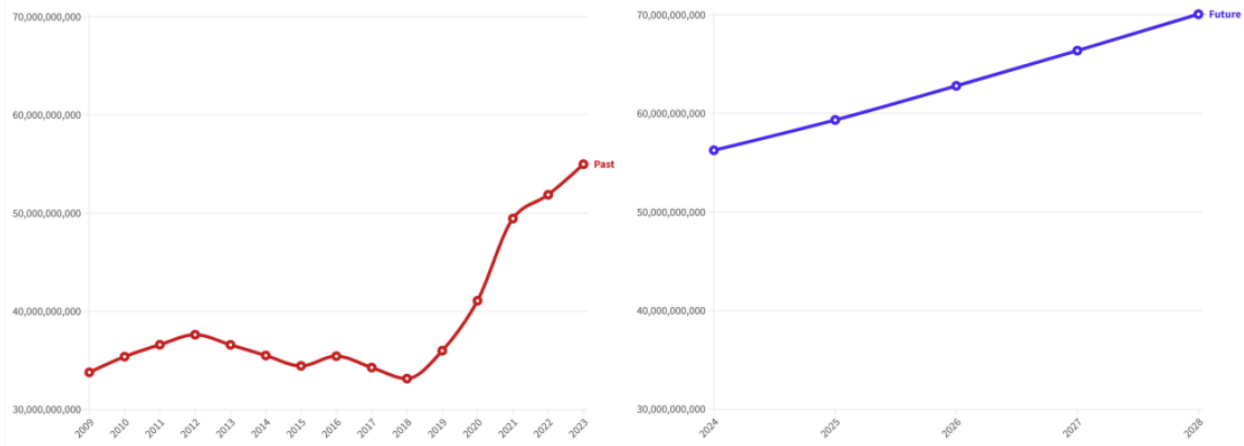


Figure 4. The Past and Future premium from 2009 to 2028

4. Assessing Insurance Gains and Losses through TOPSIS

Whether an insurance company should underwrite a policy depends on two factors: compensation loss and premium income. It's obvious that if the income is larger than the compensation, an insurance company should take action as it is profitable. Thus, this paper uses the risk loss model established above to measure the level of potential loss and the premium model to measure the level of income.

Since the dimensions of the two functions are different, this study uses the min-max normalization in the TOPSIS method combined with the unit Length Normalization programming method to process the variables.

It normalize Y_t s into y_t and Z_t into z_t similarly, detailed equations as follows.

$$y_t = \frac{a_t}{\sqrt{a_{t-2}^2 + a_{t-1}^2 + a_t^2}}, \quad a_t = \frac{Y_t - \min\{Y_{t-2}, Y_{t-1}, Y_t\}}{\max\{Y_{t-2}, Y_{t-1}, Y_t\} - \min\{Y_{t-2}, Y_{t-1}, Y_t\}} \quad (7)$$

$$z_t = \frac{b_t}{\sqrt{b_{t-2}^2 + b_{t-1}^2 + b_t^2}}, \quad b_t = \frac{Z_t - \min\{Z_{t-2}, Z_{t-1}, Z_t\}}{\max\{Z_{t-2}, Z_{t-1}, Z_t\} - \min\{Z_{t-2}, Z_{t-1}, Z_t\}} \quad (8)$$

After the normalization, the study induces s_t which can be calculated by $s_t = z_t - y_t$:

If it is a positive number, then the insurance company should underwrite policies in that area as the probability of making a profit is greater.

Now, it has successfully established an extreme-weather-based insurance model.

5. Model Application in Brazil and India

By collecting global data, it shows that Brazil and India have more extreme-weather events compared to other countries and they both show an upward trend. People in these two areas may claim wider insurance coverage and be willing to pay more premiums. However, due to the higher frequency of these disasters, insurance companies will also face a larger risk of insurance compensation for damages or loss caused by extreme weather events. Thus, applying our insurance model in Brazil and India is of great significance.

Firstly, the study counted how many times the five extreme-weather events happened in Brazil and India from 2019 to 2023, as shown in Figure 5 and Figure 6.

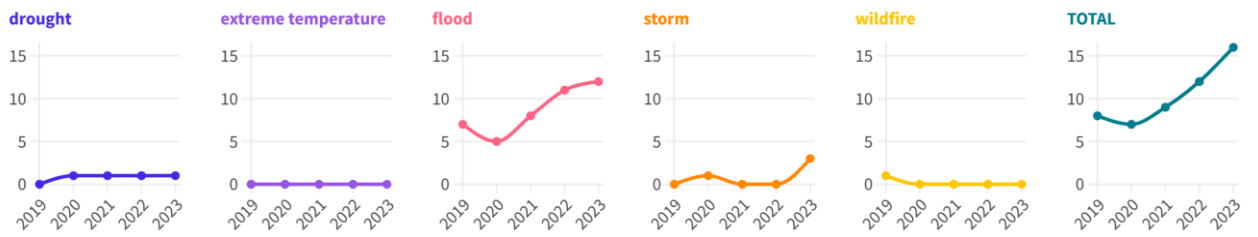


Figure 5. Extreme-weather events in Brazil from 2019 to 2023

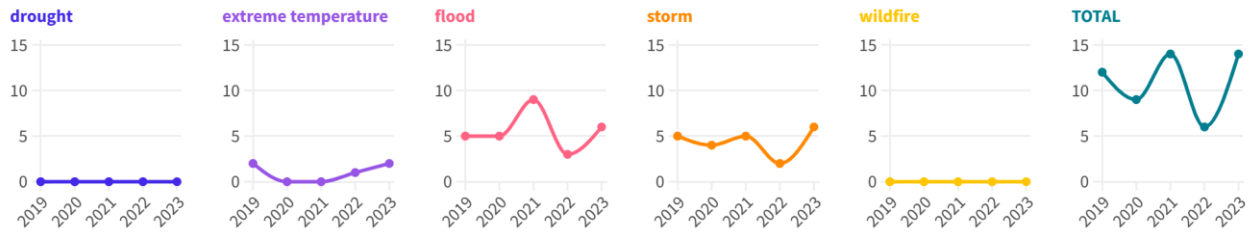


Figure 6. Extreme-weather events in India from 2019 to 2023

Then it brings the previous data into the ARIMA model. Here, the paper selects coefficients based on the ARIMA model the establish the Risk-loss model mentioned earlier, without selecting new coefficients:

$$X_{i,t} = c + \varphi_1 X_{i,t-1} + \dots + \varphi_p X_{i,t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (t=2024, p=q=3) \quad (9)$$

Having got the projected extreme-weather occurrence times, it can obtain the projected loss by the function we established before:

$$Y_t = 0.1936X_{1t} + 0.1965X_{2t} + 0.2102X_{3t} + 0.1967X_{4t} + 0.2030X_{5t} \quad (10)$$

After calculation, the result in Brazil is 3.5096, while it's 3.9027 in India.

Next, the study predicts the future premium based on the premium model developed earlier, as shown in Figure 7.

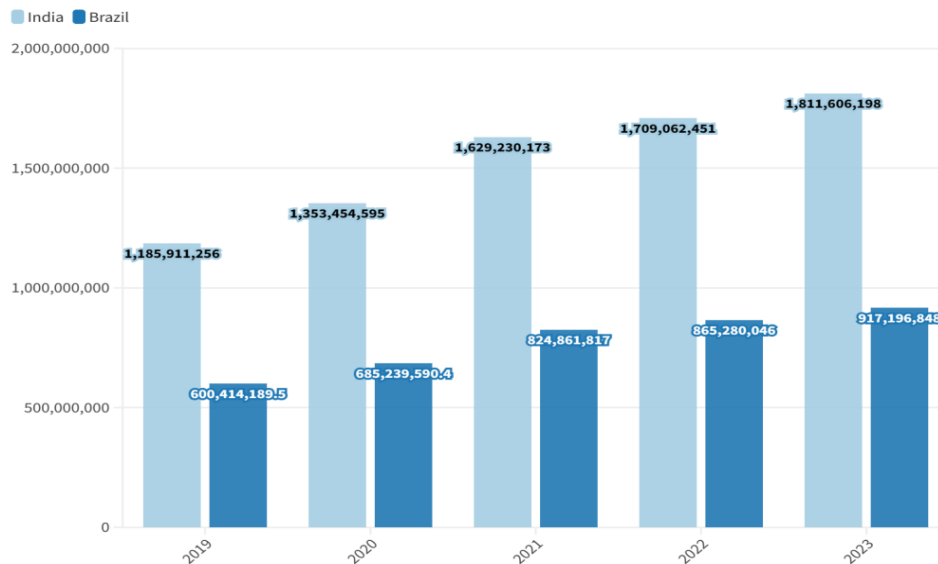


Figure 7. Gross Premium in Brazil and India (unit: US dollar)

Given the premium data from 2019 to 2023, it turns to the premium model. Here, the study also selects coefficients based on the ARIMA model in the insurance model mentioned earlier, without selecting new coefficients:

$$Z_t = c_0 + \varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p} + \theta_1 \varepsilon'_{t-1} + \dots + \theta_q \varepsilon'_{t-q} + \varepsilon'_t \quad (t=2024, p=q=3) \quad (11)$$

The paper predicts that premiums in Brazil will be 1133071385.6 dollars in 2024 and 2013454779 dollars in India.

Last, it normalizes the two variables in a bid to conclude:

$$y_t = \frac{a_t}{\sqrt{a_{t-2}^2 + a_{t-1}^2 + a_t^2}}, \quad a_t = \frac{Y_t - \min\{Y_{t-2}, Y_{t-1}, Y_t\}}{\max\{Y_{t-2}, Y_{t-1}, Y_t\} - \min\{Y_{t-2}, Y_{t-1}, Y_t\}} \quad (12)$$

$$z_t = \frac{b_t}{\sqrt{b_{t-2}^2 + b_{t-1}^2 + b_t^2}}, \quad b_t = \frac{Z_t - \min\{Z_{t-2}, Z_{t-1}, Z_t\}}{\max\{Z_{t-2}, Z_{t-1}, Z_t\} - \min\{Z_{t-2}, Z_{t-1}, Z_t\}} \quad (13)$$

For Brazil, $s_{2024} = z_{2024} - y_{2024} = 0.98172 - 0.91915 > 0$ it is wise for insurance companies to write policies in this area.

For India, $s_{2024} = z_{2024} - y_{2024} = 0.94767 - 0.80874 > 0$ thus the suggestion for insurance companies is the same.

As can be seen from the above examples, the insurance model does not require overly complex derivation, yet provides a good measure of the factors that have a great influence on decisions, which can help insurance companies make better policies.

6. Conclusions

The paper built an insurance assessment model to help insurance companies decide whether they should underwrite policies in certain areas. It divided the companies' net profit into two parts: loss and profit. To evaluate the potential loss, this study used the ARIMA model to project the occurrence times of each disaster and weighted them based on the Entropy Weight Method. To evaluate the profit, it collected data on global average premiums and projected its future value by the ARIMA model. Then the net profit can be quantified by the difference between loss and value after TOPSIS.

In addition, the model constructed in this paper is applied to India and Brazil, which are suffering from more serious meteorological disasters. Through the model constructed in this study, insurance companies have correct underwriting policies in both countries.

This paper provides a research idea and framework to apply to the insurance assessment in the field of meteorological disasters. Practical research shows that this assessment model can be well applied in two countries: Brazil and India. Thus, this model proves that it can well solve the problem of whether insurance companies should cover meteorological disaster insurance in a certain area.

References

- [1] WMO. (2021). Weather-related disasters increase over past 50 years, causing more damage but fewer deaths. Retrived from: <https://wmo.int/news/media>.
- [2] Miriam Breckner, Florian Englmaier, Till Stowasser, Uwe Sunde. Resilience to natural disasters— Insurance penetration, institutions, and disaster types [J]. *Economics Letters*, Volume 148, 2016.
- [3] Khalil Ahmed A., Liu Zaiming, Ali Attia A. Using an adaptive network-based fuzzy inference system model to predict the loss ratio of petroleum insurance in Egypt [J]. *Risk Management and Insurance Review*. Volume 25, Issue 1. 2022. PP 5-18.
- [4] Yibo Feng. Prediction of Life Insurance Premium Income Based on ARIMA and LSTM Combination Model [J]. *Scientific and Technological Innovation*, 2021, (33): 166-168.
- [5] Yuanyuan Liu, Shaoqiang Wang, Xiaobo Wang, Dong Jiang, N H Ravindranath, Atiq Rahman, Nyo Mar Htwe, Tartirose Vijitpan. Flood disaster risk assessment in Mumbai area based on AHP _ entropy weight method [J]. *Geographical Research*, 2020, 39 (08): 1892-1906.
- [6] Cui Guang-Hai; Wang Zhen; Li Jun-Li; Jin Xing; Zhang Zhi-Wang. Influence of precaution and dynamic post-indemnity based insurance policy on controlling the propagation of epidemic security risks in networks [J]. *Applied Mathematics and Computation*. Volume 392, Issue . 2021. PP 125720-.

- [7] DEHM J. Climate change, Slow Violence and the Indefinite Deferral of Responsibility for Loss and Damage [J]. Griffith Law Review, 2020 (29):18.
- [8]]NORLANDER L, PILL M, ROMERA B M. Insurance Schemes for Loss and Damage: Fools' Gold[J]. Climate Policy, 2020 (6): 707-715.
- [9] Lemoine D. The climate risk premium: how uncertainty affects the social cost of carbon [J]. Journal of the Association of Environmental and Resource Economists, 2021, 8(01): 27- 57.
- [10] FARAHANI R Z, LOTFI M M, Baghaian A, et al. Mass casualty management in disaster scene: A systematic review of OR&MS research in humanitarian operations[J]. European Journal of Operational Research, 2020, 287.
- [11] SAWALHA I H. Views on business continuity and disaster recovery[J]. International Journal of Emergency Services, 2021, 10 (3): 351-365.