WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Research on Risk Assessment and Underwriting Decision Making Based on ARIMA Model

Guanyu Liu *

Department of Economics, University of Macau, Macao Special Administrative Region, China, 519000

* Corresponding author: tonyliugy011@gmail.com

Abstract. Addressing the critical challenge of extreme weather events, this study introduces a comprehensive model for enhancing insurance underwriting strategies through predictive analytics and risk assessment methodologies. Utilizing an integrated approach that combines Autoregressive Integrated Moving Average (ARIMA) for forecasting extreme weather occurrences, Fuzzy Comprehensive Evaluation for assessing regional payment capabilities, and Monte Carlo simulations for detailed risk quantification, the re-search aims to refine the insurance industry's capacity for anticipating and mitigating financial losses. A systematic clustering analysis further segments regions based on their risk profiles, allowing for tailored insurance premium settings and underwriting decisions. Findings from the application of this model across diverse geographical areas underscore the potential for significantly im-proved risk management and economic resilience in the insurance sector. Through detailed statistical analysis and predictive modeling, the study demonstrates the importance of advanced analytical frameworks in adapting to the evolving dynamics of climate risk.

Keywords: Risk Assessment Model; ARIMA Forecasting; Fuzzy Comprehensive Evaluation; Monte Carlo Simulation; System Clustering.

1. Introduction

In recent years, the escalation of extreme weather events has posed unprecedented challenges to the insurance industry, necessitating the development of robust risk assessment models to sustain financial stability and resilience. Academics have studied the pricing of property insurance, such as Guo and Gao have used ENN and ERF models to price rice insurance in Yunnan Province, China, and Tao has assessed and priced the risk of agricultural insurance; there have also been studies on extreme weather, such as Wang, Xia & Ren have proposed methods for predicting extreme weather, like using Observing System Simulation Experiments (OSSE), fitted with Gumbel distribution, Weibull distribution, normal distribution [1-4]. However, previous studies lacked models for assessing extreme weather risks and deciding whether insurance companies would underwrite them. The intricacy of forecasting these events and their financial repercussions requires an integrative approach that not only anticipates the occurrence of such events but also evaluates their potential impact on insurance underwriting and premium pricing. This paper delves into the construction of a comprehensive model that leverages Autoregressive Integrated Moving Average (ARIMA) forecasting, Fuzzy Comprehensive Evaluation for payment capacity analysis, and Monte Carlo simulations for risk quantification. By systematically clustering regions based on their vulnerability to extreme weather events, this study proposes a nuanced strategy for insurance companies to adjust their underwriting practices and premium settings, thereby enhancing their adaptability and economic resilience in the face of climate change. Through the lens of this research, we aim to bridge the gap between predictive analytics and practical underwriting decisions, offering a novel perspective on managing climate-related risks in the insurance sector (data from https://www1.ncdc.noaa.gov/pub/data/swdi/stormevents/csvfiles/).

2. Insurance integrated modelling methodology

2.1. Forecasting natural disaster frequency and loss with ARIMA model

Drawing upon the Li et al. (2023)'s insurance pricing approach and model, the paper establishes a connection between the spot and futures markets by considering the risk-free interest rate (RF), employing an "insurance + futures" model as the pricing mechanism [5]. This involves converting the expected loss (L) due to natural disasters in a country into its present value with time(t). This value is then multiplied by the ratio of a region's population (e) to the total population (E), which in turn is multiplied by the penetration rate (F_1) of insurance purchases among residents of each region. The product is further multiplied by the probability of insurance payouts, calculated as the probability of natural disasters occurring ($P(x)$) divided by the frequency of such events within a year (r), to finally derive the insurance pricing (P). (Note: The penetration rate F_1 refers to the penetration rate of prior insurance purchases, allowing insurance companies to set prices based on historical penetration rates.)

$$P = \frac{L}{(1+RF)^t} \cdot \frac{e}{E} \cdot F_1 \cdot \frac{P(x)}{r} \tag{1}$$

To calculate the total revenue (TR) of an insurance company within a region, it is equal to the product of the insurance pricing in that region and the number of insurances sold (Q) there. The number of insurances sold in a region is calculated as the product of the region's population (e) and the penetration rate (F_2). (Note: The penetration rate F_2) refers to the current penetration rate or the penetration rate for the current year, serving as a measure of the present time. F_1 and F_2 have the same meaning but are applied on different temporal scales.)

Similarly, the maximum amount an insurance company might pay out in a region, assumed as the total variable cost (TC), is calculated as the expected loss (L) due to natural disasters in the country multiplied by a proportional coefficient, representing the expected loss in a region, further multiplied by the probability of natural disasters occurring. Economists believe that rational individuals, when making decisions, do not consider sunk costs, meaning costs such as rent and wages are disregarded. Therefore, the economic profit (g) is equal to the total revenue (TR) minus the total variable costs (TC).

$$TR = P \cdot Q \tag{2}$$

$$Q = F_2 \cdot e \tag{3}$$

$$TC = \frac{e}{E} \cdot L \cdot P(x) \tag{4}$$

$$g = TR - TC \tag{5}$$

The principle of ARIMA (p, d, q) involves transforming a non-stationary time series into a stationary one, where the dependent variable is modeled based on its own lagged values as well as the current and lagged values of the random error term [6].

(1) The autoregressive model (AR) describes the relationship between current values and their historical values, using the variable's own historical data to predict its future values. The formula for a p-th order autoregressive process is given by:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i \, y_{t-i} + \varepsilon_t \tag{6}$$

In the formula: y_t represents the current value; y_{t-i} denotes the value from i days prior; μ is the constant term; p is the order; γ_i are the coefficients for the ith autoregressive term; ε_t is the error term for the current value [7].

(2) Differencing process (I): This refers to the process of subtracting the lagged value from the current value of a time series, typically employed in a first-order differencing process to achieve stationarity in the original data series. The order of differencing here corresponds to the value of model parameter d, which is commonly set to 1.

(3) Moving Average Model (MA): The moving average model focuses on the aggregation of error terms from the autoregressive model. The formula for a q-th order moving average process is:

$$y_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{7}$$

In the formula: $\theta_i$are the error coefficients of the i-th moving average term; $\varepsilon_{t-i}$is the error of the value from i periods ago.

(4) The Autoregressive Integrated Moving Average model (ARIMA)combines the autoregressive, differencing process, and moving average components, with the formula being:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i\, y_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{8}$$

2.2. Evaluating payment capacity: A fuzzy comprehensive approach

To measure the quantity of policies, it's easy to think of a positive correlation between the number of policies and a client's ability to pay for them. Therefore, the paper considers its three personal income, personal consumption, and family debt rate aspects of solvency to form the indicator Q. Then we establish the initial fuzzy composite judgment matrix [8].

$$R = \left(a_{ij}\right)_{3*n} \tag{9}$$

The a_{ij} indicates affiliation of the i^{th} indicator to the j^{th} program. The membership function can help us determine the score of a_{ij}, the function is shown below:

$$A(x) = \begin{cases} 1, x < a \\ \frac{b-x}{b-a}, a \le x \le b \\ 0, x > b \end{cases} \tag{10}$$

$$B(x) = \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \le x \le b \\ 1, x > b \end{cases} \tag{11}$$

The function $A(x)$ measures a small-scale fuzzy function, which in this context assesses the family debt rate. Conversely, $B(x)$ measures a large-scale fuzzy function, focusing on personal income and personal consumption.

Then, based on the weight of the three indicators, the paper obtains a fuzzy vector:

$$C = [\alpha, \beta, \theta] \tag{12}$$

Finally, the paper gets a comprehensive evaluation outcome. H represents a fuzzy vector, where each element corresponds to the degree of membership of the evaluated object to the evaluative the statement j [9].

$$H = C \cdot R \tag{13}$$

2.3. Measuring the number of policies through multiple linear regression

Referring to the Liu & Jiang (2024)'s thoughts, to measure the relationship between the number of policies Q and pricing P and the ability to pay H, the insurance model needs to develop a multiple linear regression model for Q with the following equation [10]:

$$Q = \beta_0 + \beta_1 H + \beta_2 P \tag{14}$$

Using this model, it is possible to find out the relationship between the number of policies Q and pricing P and the ability to pay H. Regarding the goodness of fit and adjusted goodness of fit, the more independent variables we introduce, the better the goodness of fit becomes. However, the paper tends to use the adjusted goodness of fit, which decreases instead if the newly introduced independent variables reduce the SSE by a particularly small amount.

$$R^2 = 1 - \frac{SSE}{SST} \tag{15}$$

$$SSE = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 \tag{16}$$

$$R_{ad}^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} \tag{17}$$

To study more precisely the important factors affecting the evaluation quantity (removing the influence of the scale), the paper considers using standardized regression coefficients.

To standardize the data, the original data is subtracted from its mean and then divided by the standard deviation of the change to calculate the value of the new variable, the regression equation formed by the new variable is called the standardized regression equation, and the standardized regression coefficient can be obtained accordingly after the regression.

However, cross-sectional data are prone to heteroskedasticity problems, which is when the paper uses the optimization methods of OLS and robust standard errors.

The hypothesis test for heteroskedasticity is as follows:

Assume the regression function is: $Q = B_0 + B_1 + B_2$, test the following original hypnosis:

$$H_0: E\left(\varepsilon_i^2 \middle| H, P\right) = \sigma^2 \tag{18}$$

If H_0 is not valid, then the conditional variance $E\left(\varepsilon_i^2 \middle| H, P\right)$ is the function of (H, P), called "Conditional variance function". The BP test assumes that this conditional variance function is linear:

$$\varepsilon_i^2 = \delta_1 + \delta_2 x_{i_2} + \delta_3 x_{i_3} + u_i \tag{19}$$

If the heteroskedasticity is related to only some of the explanatory variables, only some of the explanatory variables can be used. Other variables can also be added, such as the fitted value $\hat{y}$, or variables not in the regression equation z. According to the equation (previous one), the original hypothesis simplifies to:

$$H_0\colon \delta_2 = \delta_3 = 0 \tag{20}$$

Since the perturbation term ε_i is not observable, so the residual square is used to perform an auxiliary regression on the explanatory variables:

$$e_i^2 = \delta, +\delta_2 H + \delta_3 P + error \tag{21}$$

Still using nR^2:

$$\mathrm{nR}^2 \rightarrow \chi^2(\mathrm{k}-1) \tag{22}$$

Where $R^2$is the auxiliary regression of R^2. The difference between the BP test and White's test is that the latter also includes a squared term and a cross term. Therefore, the BP test can be viewed as a special column of White's test. The advantage of the BP test is that it is constructive, it can help to identify the specific form of heteroscedasticity.

2.4. Quantifying Risk (Monte Carlo Model)

From the ARIMA model, the frequency of natural disasters*(r)* can be obtained, and the probability of natural disasters occurring in each year can be calculated using the following formula*:* Drawing on the Li et al. (2023)'s ideas, the probability obtained by averaging the frequencies over the next three years (p) ensures that the probabilistic predictions are robust and reduce volatility [11]:

$$p = (r_1 + r_2 + r_3)/(365 * 3) \tag{23}$$

Since risk is defined as the uncertainty of deviation from expected returns or expected outcomes, according to Sun (2023) the paper quantifies the risk by modelling the uncertainty on the probabilities we derived through Monte Carlo methods as follows [12]:

The paper defines p' as the uncertainty through the Monte Carlo method in the$[\mathrm{p}-\mathrm{p}',\ \mathrm{p}+\mathrm{p}']$.

Select N random values within the interval, and subsequently incorporate these N probability values into the profit assessment model to derive N "uncertain" profits. Recalling our definition that risk is the uncertainty of deviation from expected returns, the paper characterizes risk as the variance of the values of these N probabilities [11], as follows:

$$R = \sigma\left(C_1, C_{2,} C_{3\ldots} C_N\right) \tag{24}$$

2.5. Coverage decision through systematic clustering

The paper modeled the question of whether or not coverage is covered as a question with two types of output values, true and false. The data in the dataset is categorized by combining similar objects using a systematic clustering model. Referring to Tian et al. (2023)'s mind, the paper hopes to determine whether to cover the risk based on the output by categorizing the insurance objects on their own terms [13].

3. Data analysis: the United States as an example

First, this paper assesses the ability to pay (H) of U.S. residents to facilitate the paper's subsequent regression on the number of policies (Q). Here, this paper assesses the ability to pay of U.S. residents. A fuzzy composite evaluation model was used for the 50 U.S. states.

Create an affiliation function for three factors: revenue, consumption, and household debt-to-income ratio for U.S. residents:

Based on research, the top 25% of consumers in the United States have a monthly expenditure of $2,600, while the bottom 25% spend $1,500 per month, hence the membership function for consumption has parameters a = 1,500 and b = 2,600 (see Equation (11)). Similarly, the annual income for the upper-middle-income group in the United States is $90,000, whereas it is $40,000 for the lower-middle-income group, resulting in the membership function for income with parameters a = 40,000 and b = 90,000 (see Equation (11)). The high debt-to-income ratio for US residents is 1.8, with a low ratio at 1, thus in the membership function for the debt-to-income ratio, a = 1 and b = 1.8 (see Equation (10)).

After determining the affiliation function, referring to Gao (2023), the paper collects the average monthly income (in dollar) [14], the average annual consumption (in dollar), and the average debt-to-income ratio of the residents of the 50 states in 2021, and then the calculations of equations (9)-(13) can be used to obtain the ability to pay (H) of the residents of the 50 states, utilizing jupyter notebook, as shown in Table 1.

Subsequently, by gathering data on the frequency (r), loss (L), and penetration rates (F_1) of extreme weather natural disasters in the United States from 1980 to 2021, predictions for the frequency and resultant losses of extreme natural disasters from 2022 to 2030 can be made using equations (1)-(8), as illustrated in Figures 1 and 2. Insurance companies can then price property insurance premiums for a region, calculate the expected number of policies (Q), total revenue (TR), and total variable costs (TC) for a region, ultimately determining the expected profit (g) and deciding whether to underwrite insurance in that area.

Insurance companies can subsequently set the pricing for property insurance premiums within a region. By designating the number of policies (Q) as the dependent variable, and both the payment capacity (H) and the insurance premiums (P) as independent variables, multivariate linear regression is applied to historical data across varying time and regional scales through Equations (14)-(21), utilizing Stata software with results illustrated in Figure 3.It can be seen that the R- squared is 0.8312, which indicates that the correlation between P and H and Q is good and the P- value is almost zero, which indicates that the regression is significant. The final function is $\mathrm{Q} = 28.17279\mathrm{P} + 216243\mathrm{H} + \beta_0$.Incorporating the payment ability derived from fuzzy comprehensive evaluation and insurance premium pricing forecasted by the ARIMA model enables the prediction of the number of policies an insurance company can sell in different regions in the forthcoming years. This allows for the calculation of the total revenue (TR) and total variable costs (TC) for a region, ultimately facilitating the determination of the expected profit (g) and assessing the feasibility of underwriting insurance in that area, as shown in Table 1.

For this simulation, the paper took N=200, p'=0.02, and performed a Monte Carlo simulation, using equations (22) and (23), which gave us 50 sets of variances, which is the value of the risk of the insurance company's expected profits [11]. Finally, utilizing a systematic clustering model (with k=5 in this example), if an insurance company has a high-risk tolerance, it can opt to underwrite insurance in regions within categories 3, 4, and 5 where g is greater than zero. Conversely, if the company has a low risk tolerance, it can choose to underwrite insurance in regions within categories 1 and 2 where g is greater than zero. Results obtained using SPSS software are presented in Table 1.

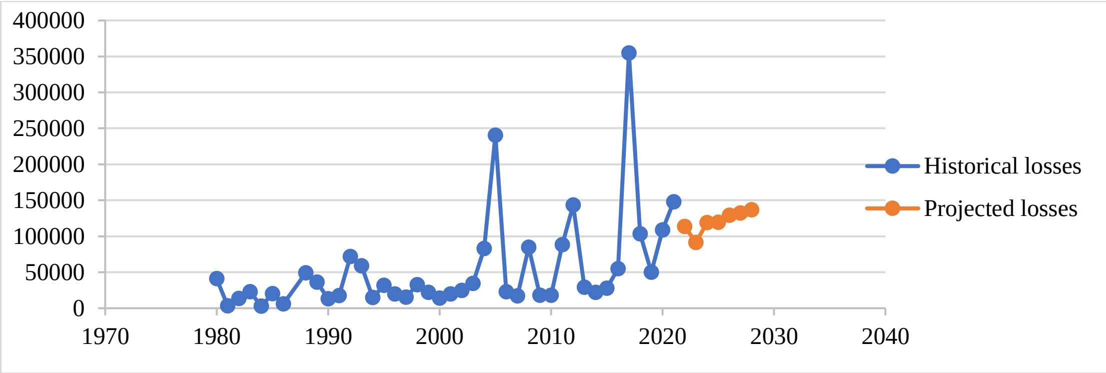

Figure 1. ARMIA model predicts extreme natural disaster losses in the US

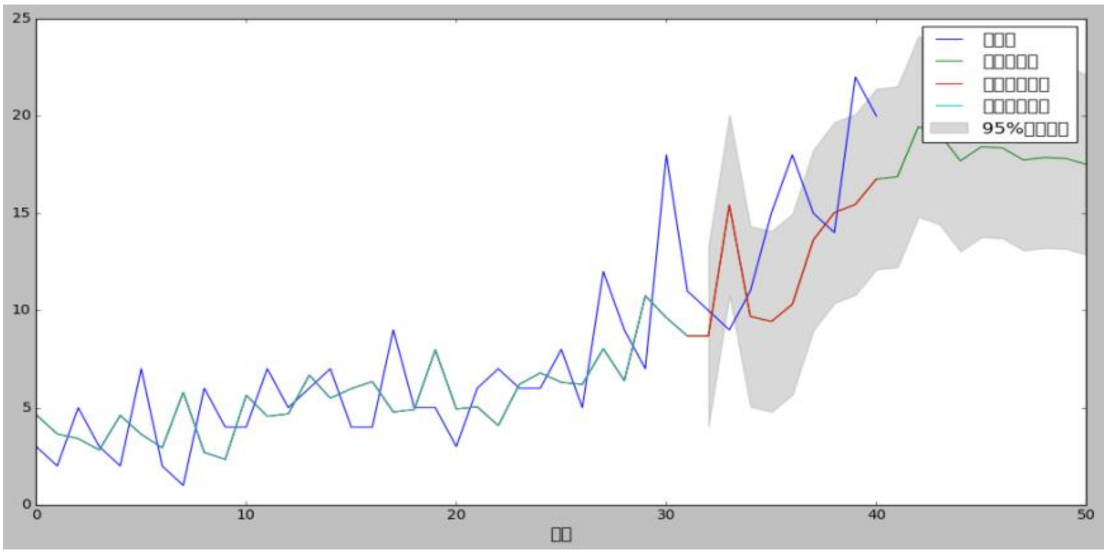

Figure 2. ARMIA model predicts frequency of extreme natural disasters in the US

Source	SS	df	MS			
				Number of obs	=	50
				F(2, 47)	=	115.70
Model	1.3057e+12	2	6.5283e+11	Prob > F	=	0.0000
Residual	2.6519e+11	47	5.6424e+09	R-squared	=	0.8312
				Adj R-squared	=	0.8240
Total	1.5709e+12	49	3.2058e+10	Root MSE	=	75116

Q_Number_of_insur~s	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
P_insurance_premium	28.17279	1.950354	14.44	0.000	24.24918	32.09639
H_ability_to_pay	216243	61961.56	3.49	0.001	91592.38	340893.5
_cons	-93947.63	33807.11	-2.78	0.008	-161958.8	-25936.5

Figure 3. Multiple linear regression of policy counts via Stata

Table 1. Multiple linear regression of policy counts via Stata

Name of the state	Debt-to Income Rate	Consumption per Month ($)	Income per Year ($)	Affiliation of Debt-to-Income Rate	Affiliation of Income per month	Affiliation of Consumption per Year	H	g (profit)	R (risk)	Risk ranked	Category
Alabama	1.36	1655	56929	0.55	0.1409090	0.33858	0.3427047	56572008	14386217	29	1
Alaska	1.55	2484	81133	0.3125	0.8945454	0.82266	0.6911776	-2351258	1056385	45	2
Arizona	1.78	1829	70821	0.025	0.2990909	0.61642	0.3437952	717345938	82982950	16	1
Arkansas	1.28	1534	50784	0.65	0.0309090	0.21568	0.2905447	-6358786	4690835	35	1
California	1.55	2529	81575	0.3125	0.9354545	0.8315	0.7069863	32774129907	3276492968	1	3
Colorado	1.78	2147	84954	0.025	0.58818181	0.89908	0.54358654	373302901	46740477	19	2
Connecticut	1.36	2311	80958	0.55	0.73727272	0.81916	0.71384581	-6442259	5700573	34	2
Delaware	1.55	1904	68687	0.3125	0.36727272	0.57374	0.43342781	-10695955	715033	48	1
Florida	1.66	1943	59734	0.175	0.40272727	0.39468	0.33119018	5433930228	570062801	3	4
Georgia	1.45	1833	61497	0.4375	0.30272727	0.42994	0.39404418	1040426965	120789285	11	1
Hawaii	2.06	2731	82199	1	1	0.84398	0.937592	13706631	3872690	37	5
Idaho	1.95	1698	76918	1	0.18	0.73836	0.649344	9218062	4240004	36	2
Illinois	1.175	1934	79253	0.78125	0.39454545	0.78506	0.66676263	3478461236	362685913	4	3
Indiana	1.175	1597	70190	0.78125	0.08818181	0.6038	0.50234954	652331271	75794127	18	2
Iowa	1.175	1789	72429	0.78125	0.26272727	0.64858	0.57262518	154071671	20685686	26	2
Kansas	1.175	1727	75979	0.78125	0.20636363	0.71958	0.58411609	118936745	16793145	28	2
Kentucky	1.175	1622	55629	0.78125	0.11090909	0.31258	0.39267972	-11373642	6804919	33	1
Louisiana	1.45	1832	57206	0.4375	0.30181818	0.34412	0.35944345	96879466	17600941	27	1
Maine	1.55	1900	71139	0.3125	0.36363636	0.62278	0.45195290	10658625	3452688	38	1
Maryland	1.95	2356	97332	1	0.77818181	1	0.93345454	792407679	88378087	14	5
Massachusetts	1.175	2464	86566	0.78125	0.87636363	0.93132	0.86981209	1303619685	139849393	7	5
Michigan	1.175	1734	64488	0.78125	0.21272727	0.48976	0.49409718	1028933296	118349662	13	2
Minnesota	1.28	2007	80441	0.65	0.46090909	0.80882	0.65680072	780235414	86383045	15	2
Mississippi	1.45	1567	46637	0.4375	0.06090909	0.13274	0.20261872	-39895029	1275339	44	1
Missouri	1.28	1612	63594	0.65	0.10181818	0.47188	0.41429745	116413405	22224081	25	1

Montana	1.78	1835	64999	0.025	0.30454545	0.49998	0.29885563	-3763935	1570616	39	1
Nebraska	1.175	1663	78109	0.78125	0.14818181	0.76218	0.58370154	47153417	8062943	31	2
Nevada	1.78	1928	64340	0.025	0.38909090	0.4868	0.31894727	52156554	10625088	30	1
New Hampshire	1.55	2188	88841	0.3125	0.62545454	0.97682	0.67211436	-10747143	1387114	41	2
New Jersey	1.55	2562	88559	0.3125	0.96545454	0.97118	0.77185836	1188024074	132543766	9	2
New Mexico	1.55	1608	53463	0.3125	0.09818181	0.26926	0.23090854	-24156788	1351054	42	1
New York	0.755	2252	72920	0.017834	0.68363636	0.6584	0.47380110	3142108105	342347186	5	2
North Carolina	1.45	1768	62891	0.4375	0.24363636	0.45782	0.38746890	1152268701	131299032	10	1
North Dakota	1.175	1835	68882	0.78125	0.30454545	0.57764	0.55679463	-4772472	893652	46	2
Ohio	1.175	1666	62689	0.78125	0.24272727	0.45378	0.46115972	1779811385	194874212	6	2
Oklahoma	1.28	1619	60096	0.65	0.10818181	0.40192	0.38822254	-556947	6945639	32	1
Oregon	1.66	2071	81855	0.175	0.51909090	0.8371	0.54306727	171468241	24237382	24	2
Pennsylvania	1.175	1767	72627	0.78125	0.24272727	0.65254	0.56820918	985835642	1192470791	12	2
Rhode Island	1.55	2147	74982	0.3125	0.58818181	0.69964	0.55006054	-6019940	1334613	43	2
South Carolina	1.78	1690	62542	0.025	0.17272727	0.45084	0.23965418	311734252	39623716	20	1
South Dakota	1.28	1691	73893	0.65	0.17363636	0.67786	0.51823490	-1759767	1399964	40	2
Tennessee	1.36	1663	62166	0.55	0.14818181	0.44332	0.38678254	188219945	30668835	21	1
Texas	1.175	1883	67404	0.78125	0.34818181	0.54808	0.55806154	7311361732	767409156	2	4
Utah	1.78	1844	87649	0.025	0.31272727	0.95298	0.48251018	221418485	27531324	23	2
Vermont	1.36	2073	76079	0.55	0.52090909	0.72158	0.60990472	-3152612	825012	47	2
Virginia	1.78	2101	80268	0.025	0.54636363	0.80536	0.49355309	1212344758	133826838	8	2
Washington	1.55	2199	87648	0.3125	0.63545454	0.95296	0.66557036	690294494	81147320	17	2
West Virginia	1.28	1485	46836	0.65	0	0.13672	0.249688	-26268370	559823	49	1
Wisconsin	1.175	1923	69943	0.78125	0.38454545	0.59886	0.58928263	207337088	30644661	22	2
Wyoming	1.55	2107	71052	0.3125	0.55181818	0.62104	0.50771145	-5299158	497694	50	2

4. Conclusion

This study has successfully developed and validated a comprehensive risk assessment model integrating ARIMA forecasting, Fuzzy Comprehensive Evaluation, and Monte Carlo simulations to predict and quantify the financial impact of extreme weather events on insurance underwriting. Our findings demonstrate the model's effectiveness in identifying regions with varying risk levels and suggest targeted underwriting strategies to optimize profitability and risk management. The systematic clustering approach further refines the decision-making process, enabling insurance companies to tailor their policies according to regional risk profiles and anticipated economic outcomes. Ultimately, this research provides a robust framework for the insurance industry to enhance resilience and strategic planning in the face of climate change-induced uncertainties. Future work will focus on refining predictive models and exploring their applicability in other sectors vulnerable to extreme weather risks.

References

[1] Guo, Z. Q., & Gao, S. H. (2024). Research on tax insurance pricing in China: An analysis of rice insurance pricing in Yunnan Province based on ENN and ERF models. Price Theory and Practice. https: //doi.org/10.19851/j.cnki.CN11-1010/F. 2024. 01. 015.

[2] Tao, R. (2024). Research on agricultural insurance risk assessment and pricing models. Economic and Technological Cooperation, (09), 59 - 61. https: //doi.org/10.13665/j. cnki. hzjjykj. 2024. 09. 052.

[3] Xia, D. D., & Ren, X. X. (2012). Prediction and management of extreme weather events in the United States. Marine Eggplant, (01), 73 - 74. https: //doi.org/CNKI: SUN: HYYB.0.2012-01-014.

[4] Wang, H. J. (2019). Construction of prediction methods for extreme weather and climate events. Journal of Atmospheric Sciences, (01), 1. https: //doi.org/CNKI: SUN: NJQX.0.2019-01 - 001.

[5] Li, H. Y., Tao, J. P., & He, L. (2023). Research on pricing of sugar cane price insurance based on the "insurance + futures" model. Journal of Huazhong Agricultural University (Social Sciences Edition), (06), 96 - 106.

[6] Cao, L. (2023). Research on insurance actuarial pricing model based on artificial intelligence. Finance & Information, (24), 165 - 167.

[7] Yang, Z. Y., Han, T., & Zhao, Z. G. (2022). Time series prediction of line loss rate based on ARIMA-GWO-SVR combination model. Power System and Clean Energy, (07), 58 - 63.

[8] Yu, Y. L., & Yang, D. Q. (2021). Research on the comprehensive fuzzy evaluation model of multi-index credit for cross-border e-commerce. Journal of Mudanjiang Normal University (Social Sciences Edition), (01), 11 - 22.

[9] Ma, D. B. (2019). Comprehensive evaluation model and application of audit risk based on dynamic fuzzy evaluation. Hebei Enterprise, (12), 15 - 16.

[10] Liu, X. M., & Jiang, Z. H. (2024). Analysis of factors affecting the demand for agricultural product logistics based on multiple linear regression: A case study of Sichuan Province. Logistics Technology, (03), 27 - 29.

[11] Li, X., Li, J., Li, K. M., Jiao, L., Zang, F., Mao, X. X., ... & Tai, X. S. (2023). Sediment heavy metal pollution characteristics and risk assessment of Lanzhou Yintan Wetland Park based on Monte Carlo simulation. Environmental Chemistry.

[12] Sun, Y. (2023). Risk identification and analysis during the construction phase based on Monte Carlo simulation. Popular Standardization, (20), 63 - 65.

[13] Tian, Y. M., Wu, S. F., & Zhao, X. (2023). Systematic clustering analysis of Traditional Chinese Medicine syndromes in patients with Type 2 diabetic nephropathy and its relationship with disease staging. Journal of Practical Medicine, (22), 2994 - 2999. doi: CNKI: SUN: SYYZ.0.2023-22 - 024.

[14] Gao, L. (2023). Research on small batch order production demand forecasting model based on ARIMA. Modern Information Technology, (15), 97 - 101.