Research on Prediction of Insurance Compensation Amount after Fire Disaster

Shixuan Han*

Dalian University of Technology, Leicester International College, Panjin, Liaoning, 124221

*Corresponding author: dd1844004931@163.com

Abstract. Fire is one of the major disasters that seriously threaten public safety and social development, so it is particularly essential to predict the amount of insurance compensation after a fire. This paper predicts the number of fire occurrences from various aspects such as region, housing, and car fires based on the LSTM model. A model for predicting insurance compensation after a fire is established based on the LSTM-Poisson distribution model and factors such as average house prices. Taking the number of fire occurrences in four regions of the UK from 2001 to 2021 as a case, the fire insurance compensation model is used to predict based on the average house prices and car prices in the UK, providing a reference for insurance companies in terms of fire insurance compensation.

Keywords: Number Of Fires; Insurance Compensation Amount; LSTM; Poisson Distribution.

1. Introduction

Fire has a significant impact on social order and personal and property safety. It is one of the main emergencies that threaten the living environment and public safety. With the rapid development of society and economy, the risk and probability of fire accidents have correspondingly increased, resulting in huge property losses. Insurance companies pay huge compensation funds. Therefore, accurately predicting the number of fires and the compensation amount of insurance companies is of great significance for enhancing disaster prevention capabilities, effectively allocating insurance resources, and promoting sustainable development of social enterprises.

Scholars have conducted extensive research on fire occurrence prediction. Ma Shuqi proposed a fire frequency prediction method based on seasonal autoregressive moving average model, providing theoretical support for fire prevention and control work [1]. Wang Jiachen predicted the occurrence of forest fires in Anhui Province from 2015 to 2019 by establishing meta-linear regression and GM(1,1) models [2]. Zhuang J integrated spatial and temporal factors of fire occurrences to establish a comprehensive fire occurrence prediction model [3]. Ziyu Li used 2019 and 2020 data as test samples to test the trained neural network, and then used the neural network to predict the basic pension insurance income and expenditure data for the next five years [4]. Xueyan Li designed a merge average algorithm to calculate the size of elderly disabled population and long-term care insurance expenditure [5]. Rui Geng et-al. used the grey Deng's correlation degree and the equal dimensional grey system GM (1,1) model to explore the influencing factors of basic medical insurance fund expenditure and predict them [6]. Vasilakos Christos' optimal fuzzy clustering based on the DTW similarity of ND.VI time series data, combined with the use of LSTM networks and clustering data, can provide useful results for monitoring vegetation dynamics in fragmented Mediterranean ecosystems [7]. Hu Jun constructed a quantitative determination model for building fire rate and used the Latin hypercube sampling method to conduct stratified sampling of the combustion loss rate at different stages of building fire. [8]. Gu Jinkai analyzed the issues discovered during the research process, such as different mediation and insurance compensation projects, different compensation calculation standards, and the nature of household registration affecting the compensation amount, and proposed solutions [9]. Guiying constructed a life insurance density prediction model based on LSTM neural network in Hunan Province, predicting the average absolute relative error (MAPE) [10].

In summary, scholars have conducted extensive research on predicting the number of fires and applied the research results to specific practices, which has achieved certain results in exploring the laws of fire occurrence and improving emergency response capabilities. However, scholars have insufficient research on post fire insurance compensation. Therefore, this paper uses the LSTM model to predict the number of fires and the LSTM with Poisson distribution model to estimate the post fire insurance compensation amount, in order to maximize the prediction accuracy of the insurance amount while adapting to nonlinear relationships. In terms of innovation, paper has designed a unique post fire insurance compensation model, introducing a proportional coefficient p to ensure the accuracy of the prediction results while minimizing the prediction process.

2. Preliminary

2.1. LSTM model

LSTM network is an improved Recurrent Neural Network (RNN) that introduces memory cells and gate mechanisms to store and control information flows. Information can be ensured to be screened and retained for a long time in its unique neural structure, effectively avoiding the problem of information loss when predicting insurance compensation amounts after a fire disaster.

The structure of the LSTM model is shown in Figure 1, σ represents the sigma function, W and b are the weights and bias parameters of the model, while $\otimes$ represents element by element multiplication, tanh represents a hyperbolic tangent function. These formulas describe the calculation process of input gates, forgetting gates, output gates, and memory units, as well as how to update hidden states.

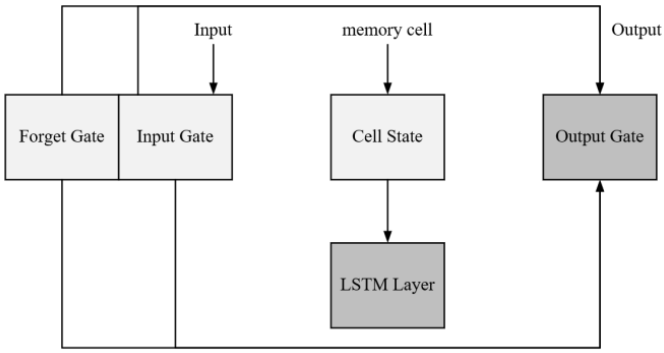


Figure 1. LSTM Structure Diagram

The calculation process of LSTM is roughly represented by equations (1) - (6). When the time step is t, the input vector of the hidden layer of LSTM is x_t, the output vector is h_t, and the memory unit is c_t.

The input gate controls how much input data x_t currently flows into the memory unit of the network, that is, how much effective information c_t store, as shown in equation (1):

$$\text{Oblivion Gate：}\ i_t = \sigma(W_{xi} + W_{hi}h_{t-1} + b_i) \tag{1}$$

Forgetting gate is the core component of LSTM unit, which can forget some information and retain some information. The forgetting gate controls the self connecting unit, which can determine which parts of historical information will be discarded, that is, the impact of the information in the previous memory unit c_{t-1} on the current memory unit c_t, as shown in equations (2) and (3):

$$\text{Amnesia Gate：}\ f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{2}$$

$$\text{Input Gate：}\ c_t = \sigma(f_t \otimes c_{t-1} + i_t \otimes \widetilde{c}_t) \tag{3}$$

The output gate controls the impact of memory unit c_t on the current output value h_t, that is, some information in c_t will be output in time step . The output h_t of the output gate and LSTM unit at time t is shown in equations (4) and (5):

$$\text{Output Gate}: o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{4}$$

$$\text{Hidden State}: h_t = o_t \otimes \tanh(c_t) \tag{5}$$

The role of candidate cell state $\tilde{c}_t$ in LSTM is to generate a candidate memory state for updating the content of the memory unit, as shown in equation (6):

$$\text{Candidate Cell State}: \tilde{c}_t = \tanh(W_{xc} + W_{hc}h_{t-1} + b_c) \tag{6}$$

2.2. Poisson distribution

Poisson distribution is a discrete probability distribution. Due to its independence, which means that the occurrence of past events does not affect the probability of future events, it is commonly used to describe the number of times a random event occurs per unit time (or unit area, unit length, etc.). The independence of Poisson distribution can effectively predict the probability quality function of the insurance compensation amount after a fire disaster. In addition, the Poisson distribution is suitable for counting the number of occurrences of low probability events, and the shape of the distribution can be adjusted by the average number of occurrences. The probability mass function of the Poisson distribution can be expressed as equation (7):

$$P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}(k = 0,1,...) \tag{7}$$

In the formula, X represents the number of times a random event occurs; k represents the specific number of times; λ represents the average number of occurrences of random events per unit time (or unit area, unit length, etc.); e is the base of the natural logarithm; The factorial represented by $k!$.

2.3. A Model of Insurance Compensation Amount after Fire Disaster

According to the established model, the claim amount c caused by a fire can be calculated based on the house price h. Assuming that the claim amount caused by a fire is directly proportional to the house price, the proportional coefficient is p .To further calculate the value of p , two additional factors are introduced, namely the average housing price ah in the UK in 2021 and the average post fire insurance compensation amount ac in the UK in 2021, namely:

$$c = p \times h \tag{8}$$

$$p = \text{ac} / \text{ah} \tag{9}$$

It was found that the average fire claim amount in the UK in 2021 was approximately 340000 pounds, while the average house price in 2021 was 2 million pounds, so p -value was determined to be 0.17. This means that for every 1 unit increase in housing price, the claim amount will increase by 0.05 units.

3. Model feasibility testing

3.1. Model Assumptions

For the prediction model of post fire insurance amount, it is necessary to exclude some factors that may affect the model:

(1) Eliminate the influence of weather and climate;

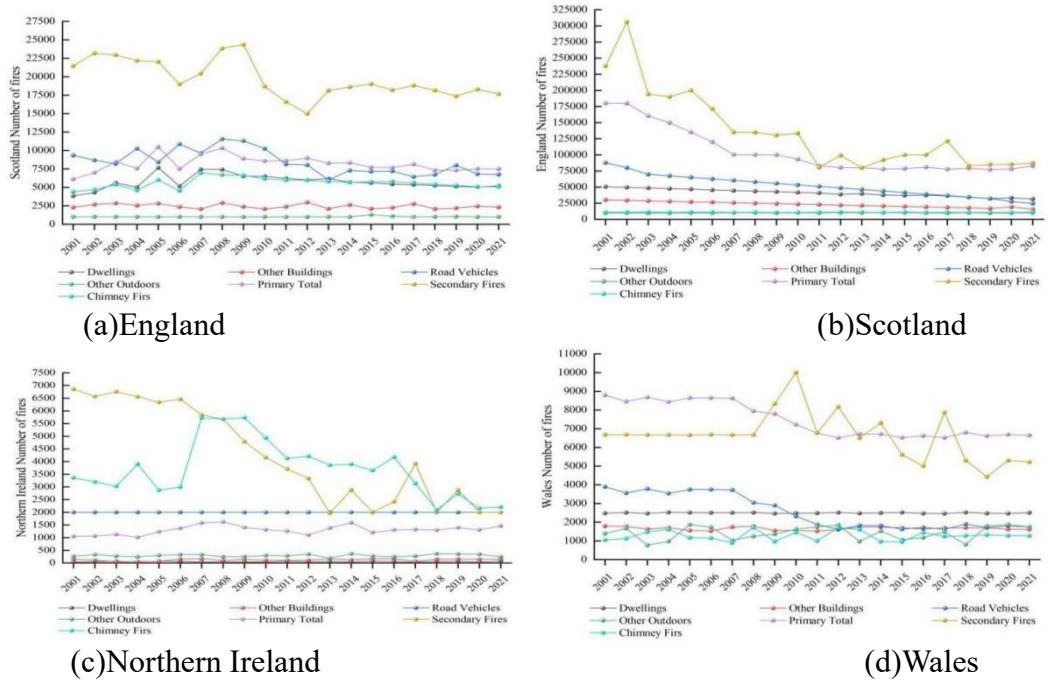
(2) Excluding insurance companies from compensating for losses caused by false alarms;

(3) After eliminating the impact of annual fluctuations in housing prices in different regions on insurance limits, the insurance limits for different types of fires can be calculated using housing prices from the past 5 years;

(4)For non building fires, after excluding the influence of car brands, choose the local average car price to calculate the insurance limit for different types of fires.

3.2. Data visualization

Due to the UK's ability to provide accurate and reliable data, as well as its long-term, diverse, and representative nature, it facilitates policy reference and international comparative research, so the paper selecting the total number of fires in the UK from 2001 to 2021 as the research object, the fires were subdivided into 7 categories: Dwellings, Other Buildings, Road Vehicles, Other Outdoor, Secondary Fires, Chimney Firs, and Primary Total. The LSTM model was used to predict the number of fires in 2022. Use data visualization methods to create a line chart of four regions to determine whether there are outliers in the database used for prediction, in order to visually observe the distribution and trend of the data. The results are shown in Figure 2.



(a)England (b)Scotland

(c)Northern Ireland (d)Wales

Figure 2. Number of regional fires in the UK, 2001-2021

It can be seen that the database selected for prediction has no noise, i.e. no outliers. Moreover, the comparison error between the predicted data and the actual data from 2001 to 2021 is relatively small, and the feasibility of the model has been verified.

3.3. Feasibility testing

Fire quantity data from 2001 to 2011 were first used for training, and the model was used to predict fire quantity prediction data from 2012 to 2021 to demonstrate the feasibility of the LSTM model. The accuracy and reliability of the model can be evaluated. This verification process will help determine whether the model can effectively predict future data and provide a basis for further prediction and decision-making using the model. The predicted results of the number of fires from 2012 to 2021 are shown in Table 1, and the comparison between the predicted and true values of the total number of fires in the UK from 2012 to 2021 is shown in Figure 3.

Table.1. Forecast Results of the Number of Fires in the UK from 2012 to 2021

Year	Total	Dwellings	Other Buildings	Road Vehicles	Other Outdoors	Primary Total	Secondary Fires	Chimney Fires
2012	427048	53032	31013	37760	11778	138945	209756	9481
2013	423169	51925	30131	36571	11601	137440	207692	9653
2014	419205	50837	29243	35380	11399	135862	205539	9815
2015	415152	49765	28351	34184	11175	134210	203298	9958
2016	411001	48704	27454	32981	10931	132483	200966	10096
2017	406746	47650	26551	31771	10669	130678	198545	10209
2018	402379	46599	25643	30551	10388	128794	196035	10220
2019	397891	45536	24729	29318	10091	126832	193434	10129
2020	393271	44465	23809	28068	9777	124793	190742	9934
2021	388511	43391	22882	26799	9456	122683	187960	9736

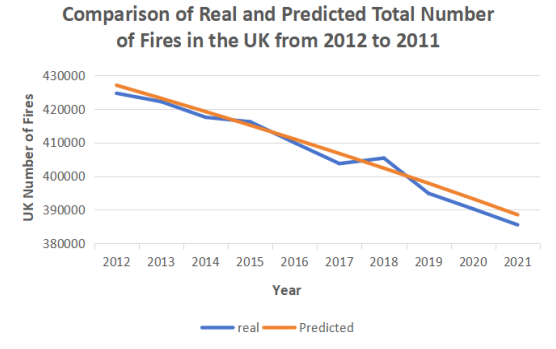


Figure 3. comparison of Real and Predicted Total Number of Fires in UK , 2001-2021

The observed difference between the predicted value and the actual value is small, and the error is within an acceptable range, further proving the feasibility of the LSTM prediction model. This means that when using past data for prediction, the model can produce accurate results and capture trends and patterns in the time series.

4. Example verification

4.1. Prediction of Fire quantity

Using the aforementioned LSTM model for fire count prediction, we forecast the number of fires in 2022. The model training process involves the combination of back propagation algorithm and Monte Carlo method. The initial learning rate is set to 0.001, dropout is set to 0.1 to prevent over-fitting, batch size is 8, and the model is trained for 1000 epochs. Eventually the predicted values with minimum standard deviation can be obtained and the results obtained are shown in Table 2.

Table.2. Prediction Results of Fire Quantity in 2022

	England	Scotland	Northern Ireland	Wales	Total
Dwellings	27643	5160	220	1747	34770
Other Buildings	14474	2372	356	1025	8227
Road Vehicles	19814	2211	206	1399	23630
Other Outdoors	5693	1142	42	408	7285
Primary Total	64535	18992	792	4160	88479
Secondary Fires	96207	18664	2476	6424	123771
Chimney Fires	3540	1152	2440	259	7391
Total	205937	42605	6927	15550	271064

According to the predicted data in the table above, in 2022, it is expected that 6927 fires will occur in Northern Ireland, 15550 in Wales, 42605 in Scotland, and England will be the region with the highest number of fires, with an estimated 205937. Overall, it is expected that there will be 271064 fire incidents in the entire UK.

4.2. Prediction of post fire insurance compensation amount

Based on the modeling of the amount of insurance compensation after a fire disaster, information on the price of housing and automobiles in each region of the UK is retrieved, as shown in Table 3.

Table.3. Average Prices of Housing and Cars in Four Regions of the UK

	housing prices	car prices
Northern Ireland	£130000.00	£13000.00
Wales	£150000.00	£25000.00
Scotland	£150000.00	£30000.00
England	£300000.00	£40000.00

Classify residential, other building, and chimney fires as building fires, and the estimated compensation amount is determined by the housing price. Classify road vehicles and other abnormal situations as non building fires, and the estimated compensation amount is determined by the vehicle price. In particular, since fires in secondary fires are relatively large and the losses associated with them increase, house prices were used to determine the amount of compensation for estimation. The resulting probabilities of various types of fires in different areas are shown in Figure 4.

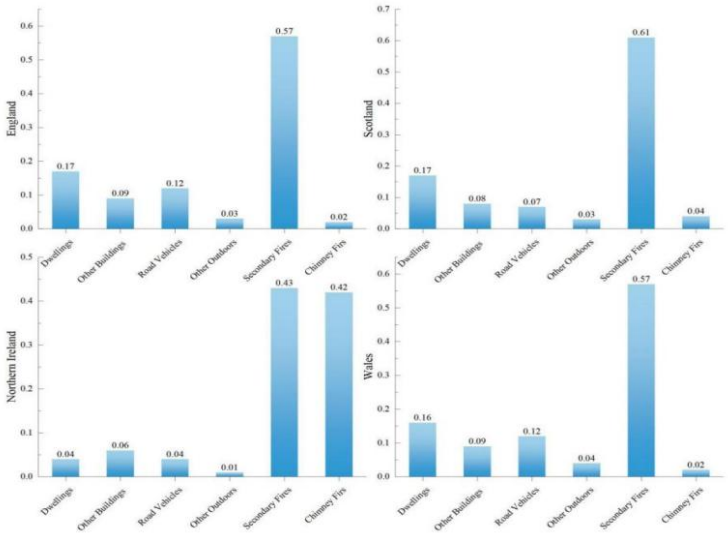


Figure 4. Probability of Various Types of Fire in Different Regions

Compensation amounts for different fires, and the number of insured individuals, the estimated total amount of post fire insurance compensation is obtained by using MATLAB software. According to statistical data, the fire claim rate is about 5%. Therefore, when 10000 people in each region are insured, the final post fire compensation amount is shown in Table 4.

Table.4. Prediction Results of Total Insurance Compensation Amount after Fire Disaster

Area	England	Scotland	Northern Ireland	Wales	Total
Amounts	£ 10497942	£ 11050000	£ 11730000	£ 22185000	£ 55462942

By comparing the predicted and true values of the total amount of insurance compensation after a fire disaster in 2022, it can be observed that the difference between the two is small and the error is within an acceptable range, proving the accuracy of the LSTM Poisson distribution prediction model.

5. Conclusion

Fires greatly impact the socio-economic development and normal lives of people, making post-fire insurance compensation work particularly important. In the context of predicting insurance compensation amounts after fires, this paper builds upon the LSTM model used to forecast the total number of fires in various regions of the UK in 2022. Based on the local average property prices and incorporating the LSTM-Poisson distribution model, the study establishes a model to predict the total insurance compensation amount for post-fire claims in 2022 for insurance companies. Paper designs a unique post fire insurance compensation model and introduces a proportional coefficient p, which not only ensures the accuracy of the prediction results but also greatly simplifies the prediction process. It focuses on solving fire insurance problems, which is a relatively novel argument. It accurately predicted the number of fires in the UK in 2022 and estimated the amount of post fire insurance compensation required by insurance companies.The research focuses on predicting fire compensation amounts from the perspective of insurance companies, ensuring their interests in the following year. The findings of this study can serve as a reference for insurance companies in terms of fire insurance compensation.

References

[1] Ma S., Liu Q., & Zhang Y. (2021). A prediction method of fire frequency: Based on the optimization of SARIMA model. Journal of Applied Statistics and Management, 40(8), 1481-1491. (In Chinese)

[2] Jiachen Wang, Yuxin Feng, Zhishang Hao et al. Prediction and Analysis of Forest Fires in Anhui Province [J] Volkswagen Standardization, 2021, (13): 198-200

[3] Zhuang, J., Liu, Y., Zhang, Y.et al. (2018). A Spatiotemporal Approach for Predicting Wildfire Occurrence in the United States. Journal of Applied Geography, 96, 15-25.

[4] Ziyu Li, Prediction of Basic Pension Insurance Revenue and Expenditure in Jilin Province Based on Neural Networks [D]. Jilin University, 2023

[5] Xueyan Li, Prediction and Policy Simulation of Long Term Care Insurance Expenditure in China [D]. Shandong University, 2021

[6] Rui Geng , Xiaoguang Fu , Yan Wang , Prediction Model and Empirical Study of Basic Medical Insurance Fund Expenditure [J] Statistics and Decision Making, 2022, 38 (01): 149-152

[7] Christos V,E. G T,Dimitris K. LSTM-Based Prediction of Mediterranean Vegetation Dynamics Using NDVI Time-Series Data[J]. Land,2022,11(6).

[8] Hu Jun, Shu Xueming, Xie Xuecai, Yan Jun, Zhang Lei Building fire insurance rates based on quantitative risk assessment [J] Journal of Tsinghua University (Natural Science Edition), 2023,63 (05): 775-782.

[9] Gu Jinkai How to Meet the Actual Compensation Needs of Employer Liability Mutual Insurance [J] Chinese Fisheries, 2019, (07): 27-28.

[10] Guiying Research on Life Insurance Demand Prediction in Hunan Province Based on LSTM Neural Network [D]. Hunan University, 2022.