



Global Forecasting of Extreme Weather and Insurance Losses Using an LSTM-Based, Audit-Ready Framework

Hongbo Guo¹, Shuotian Li², Guojun Long³, Qiqi Liang⁴, Haochi Zhang⁵

¹School of Information Engineering, Yulin University, Shaanxi 719000, China

²Faculty of Computer Science & Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia

³School of Accounting, Guangdong Baiyun University, Guangzhou 510550, China

⁴School of Creative Design, Shenzhen Technology University, Shenzhen 518118, China

⁵School of Software and Microelectronics, Peking University, Beijing 102600, China

*Corresponding to: Shuotian Li (24051012@siswa.um.edu.my)

Abstract:

The property and casualty insurance industry increasingly relies on deep neural networks to quantify weather-driven risks. This study develops a forecasting framework based on long short-term memory networks to estimate the global impact of extreme weather on insurance claims by integrating authoritative meteorological and financial datasets. Specifically, we leverage globally consistent records of temperature, precipitation, and snowfall from NOAA's National Centers for Environmental Information (1995–2025) and insured-loss statistics from Swiss Re's global catastrophe reports, rather than focusing on local or regional case studies. The model captures long-range temporal dependencies without manual feature engineering and employs adaptive moment estimation to stabilize training and reduce prediction errors. A fully connected layer with rectified linear unit activation enhances nonlinear fitting, while post-hoc Shapley additive explanations clarify how weather variables and recent claims shape predicted losses. Benchmarks against classical baselines—random forest, support vector machine, and autoregressive integrated moving average—demonstrate consistent accuracy gains. Using three decades of data, including a decade reserved for out-of-sample evaluation, the framework delivers accurate forecasts with transparent attributions that support pricing, reinsurance planning, and catastrophe response under climate risk. This integrates extreme-weather signals with insurance losses, based on globally aggregated datasets, to provide reproducible, regulator-auditable global insights.

Keywords: Deep Neural Network Model, Long Short-Term Memory Network Model, Extreme Weather; Insurance Claims.

1 INTRODUCTION

The increasing impact of extreme weather events in the context of rapid societal development has made them a crucial research topic in the field of environmental science and risk management. For example, as shown in Fig. 1 below, it is an extreme weather event of lightning strikes in Bangkok, Thailand. Due to the intensification of global climate change, extreme weather events not only seriously affect the stability of the global economy, but also pose a major challenge to social development and public safety. According to the Intergovernmental Panel on Climate Change (IPCC, 2021), marginalized populations are most affected by climate change. Nearly half of the world's population lives in highly vulnerable areas with limited development capacity. This trend is expected to continue in the future. In recent years, as the pace of global climate change has accelerated, concern about extreme weather events has reached an all-time high.



FIG. 1. Extreme weather event (Lightning strikes Bangkok, Thailand)

Statistics show that the global losses caused by extreme weather events have increased significantly, for example, the global natural disaster data in 2023 show that 326 large natural disasters occurred globally, involving a variety of disaster types, affecting 117 countries and regions, resulting in 86,473 deaths due to disasters globally, with an affected population of 93,052,400 people, and a direct economic loss of 202,652 million U.S. dollars. Regarding the China context, all types of natural disasters in China caused 95,444,000 people to be affected, and direct economic losses amounted to 345.45 billion yuan in 2023. Climate change-driven insurance costs are projected to rise significantly by 2040 (Wang, 2024; Fantini et al., 2023), showing the impacts of natural disasters across the globe, as well as the frequency and economic losses of different types of disasters, emphasizing the wide-ranging impacts of natural hazards and the severity of economic losses globally.

Recurrent architectures such as long short-term memory (LSTM) units mitigate—though do not eliminate—the vanishing/exploding-gradient issues of vanilla recurrent networks by introducing gating and a near-additive memory path. This makes it easier to retain long-range temporal dependencies in meteorological and claims series (Shi, 2023). LSTM models are essentially an evolved version of traditional recurrent neural networks(RNN), designed to overcome the gradient vanishing and explosion problems of these networks (Suradhaniwar et al., 2021). Given that traditional RNNs commonly face the gradient vanishing problem when attempting to deal with long term dependencies, this shortcoming leads to a gradual loss of early information as the sequence length increases, which in turn weakens their efficacy in long time series analysis (Karevan and Suykens, 2020). The previous studies mainly focus on the optimization of traditional time series prediction methods and their combination with single-layer neural networks, and although these methods have achieved good results in specific application scenarios, they still show limitations in dealing with complex temporal dependencies. In contrast, the LSTM model overcomes several limitations in traditional methods through its unique memory cell architecture, and is able to handle long time series and complex nonlinear regression problems more effectively. Consequently, in order to further improve the prediction accuracy and adapt to a wider range of application scenarios, many scholars have conducted in-depth research on LSTM to optimize its performance and address specific challenges. Chang et al. (2018) presented two examples to validate the performance of adaptive moment estimation(Adam)-optimized LSTM neural network and used a dataset from New South Wales, Australia to illustrate the excellence of the model. The results showed that the proposed model can significantly improve the prediction accuracy. Adam's algorithm, which is a first-order gradient optimization algorithm based on adaptive estimation of low-order moments, is easy to implement, computationally efficient, and has a small memory footprint. It is capable of optimizing the mean square error loss function, which is particularly well suited to LSTM neural networks.

Given the challenges posed by extreme weather, we implement and evaluate a global LSTM-based forecasting framework tailored to insurance-claim time series, prioritizing applied evaluation and transparent interpretation over algorithmic novelty. The main contributions of this paper are as follows:

(1) Deployment-oriented LSTM framework. We present a well-structured, robust forecasting framework based on an optimized LSTM to improve forecasts of insured losses associated with extreme weather. The framework integrates deep sequence modeling with climate-risk management and is explicitly aligned with actuarial practice and regulatory requirements.

(2) Global, audit-ready data and interpretability. We integrate globally consistent meteorological records from NOAA/NCEI and cross-country insured-loss statistics from Swiss Re to move beyond local case studies and support regulator-auditable analyses. We provide end-to-end interpretability via Shapley Additive Explanations (SHAP), which clarifies the contributions of temperature, precipitation, snowfall, and historical claims to the predictions, enhancing transparency for climate-risk policy decisions.



(3) Systematic baselines under a common split. Trained on 1995–2012 and evaluated out-of-sample on 2013–2022, our LSTM consistently outperforms RF, SVM, and ARIMA on RMSE/MAE/MAPE and on high-loss detection (F1) across multiple forecast horizons.

(4) Ablations and sensitivity for deployment. Optimizer, architecture, and regularization ablations—together with hyperparameter-sensitivity analyses—identify which components drive gains and provide practical parameter ranges for robust deployment; the framework delivers high predictive accuracy with manageable computational cost relative to accuracy gains

Paper Structure. The remainder is organized as follows: Sect. 2 reviews related work; Sect. 3 presents the proposed methodology, detailing the LSTM architecture and training procedure; Sect. 4 describes the experimental evaluation; Sect. 5 reports the results and discussion; and Sect. 6 concludes the paper.

2 RELATED WORK

Many academics have contributed valuable insights to the field of research examining the impact of extreme weather events on the insurance industry. However, effective assessment and management of extreme weather risks remains a challenge. Zhou et al. (2024) provide an in-depth analysis of the significant challenges posed to the insurance industry by the increased frequency and intensity of extreme weather events due to climate change, and propose a set of methodologies for quantifying the impacts of climate change that provide specific guidance to the property insurance industry. Rao and Li (2023) state that China has accumulated a number of successful cases and best practices in the field of climate risk insurance, identifying the existing insurance coverage, and premium pricing mechanisms, with the aim of promoting the further development and improvement of climate risk insurance in China.

Deep models can excel on complex, nonlinear signals, but their gains depend on data scale, task design, and evaluation protocol. We therefore benchmark LSTM against classical baselines under a common split to assess practical utility. Cheng (2022) provided an extensive review of LSTM research, tracing how Hochreiter and Schmidhuber originally introduced this special RNN variant to mitigate gradient vanishing and exploding issues during long-sequence training. Thanks to its gating mechanisms, LSTM effectively captures long-term dependencies, overcoming limitations of traditional RNNs. Maheswari and Gomathi (2023) evaluated the performance of diverse deep learning architectures for weather forecasting, demonstrating that LSTM and CNN architectures are particularly adept at modeling nonlinear temporal patterns. In contrast, Tan and Gong (2024) developed an extreme weather risk prediction method that integrates Monte Carlo simulation with the entropy-TOPSIS decision-making framework to assess and rank climate-risk scenarios in China and the United States, showing improved identification of high-risk events. Although these models have significantly advanced forecast accuracy and risk quantification, they still face challenges in incorporating localized geographic features and fine-grained structural information of extreme events, which constrains their predictive precision and practical usability.

Nobanee et al. (2022) provide important insights in their study, which emphasizes insurance as an effective tool for dispersing and transferring losses from climate disasters, and summarizes international successes in incorporating factors such as international partnerships, climate change, and other factors into catastrophe models, promoting loss prevention and incentivizing mitigation of climate risks through insurance provisions, to achieve compliance with the international consensus and enhance the resilience of the whole society to cope with climate risks. Essa et al. (2020) proposed a RNN to improve the accuracy of short-term lightning prediction based on an LSTM lightning density prediction model, which utilizes LSTM to capture spatio-temporal data dependencies and learns long-term data patterns. The experimental results show that the average absolute error of the model in one-hour prediction is 2.87 times/hour and the mean square error is 1209, indicating its significant potential in prediction accuracy, which provides an important reference for climate risk assessment and insurance product development.

In order to enhance the claims management capability of the insurance industry, Tzougas et al. (2022) evaluated the impact of different weather conditions and environmental factors on insurance claims patterns, so as to improve the management and prediction capability of insurance companies for such risks; Shi et al. (2024) used an innovative feature fusion approach, combined with the prediction of long and short-term factors, and utilized deep learning techniques to enhance the potential loss distribution in the insurance industry's claims management prediction accuracy. In particular, it integrates detailed weather data to improve the accuracy of predicting potential losses. The methodology significantly improves the ability of insurance companies to process and respond to claims and provides an important reference for future research. In particular, it predicts the entire loss distribution and is suitable for insurance applications requiring uncertainty quantification and risk management.

Despite significant progress in analyzing the impact of extreme weather events on the insurance industry and in applying deep learning models, several limitations remain, particularly in model accuracy and interpretability. Recent advances in explainable artificial intelligence (XAI), especially SHAP, have significantly improved the interpretability and transparency of complex neural network predictions by clearly quantifying feature contributions, thereby enhancing stakeholder trust and supporting decision-making in risk-sensitive industries such as insurance. Many recent studies have successfully combined SHAP with LSTM or other



time-series forecasting models to improve model transparency and feature importance analysis. For example, Sunu Fathima and Kovoor (2024) used SHAP to interpret stacked LSTM weather prediction models, demonstrating SHAP's growing value in time-series forecasting.

Based on these research foundations, this paper aims to develop an LSTM neural network-based model that not only improves the accuracy and efficiency of predicting extreme weather events and associated insurance claims costs but also enhances model interpretability through SHAP-based explanations.

3 PROPOSED METHODOLOGY

In this paper, we implement and evaluate a LSTM-based forecasting model aimed at improving the accuracy of predicting extreme weather events and their associated insurance claims costs, enabling insurers to make more effective risk management and strategic decisions.

We implement a compact, deployment-oriented LSTM that couples globally harmonized meteorological records (temperature, precipitation, snowfall) with regulator-auditable insured-loss series. Inputs are min-max scaled and arranged into fixed look-back windows; the network (stacked LSTM with a simple dense head) is trained with Adam and early stopping under a fixed train/test split [e.g., 1995–2012 train, 2013–2022 test]. Evaluation follows actuarial practice (RMSE/MAE/MAPE plus an extreme-loss F1), and post-hoc SHAP-style attributions quantify the roles of weather variables and lagged claims. To provide a comprehensive description of the research methodology, this section is divided into three key parts: (3.1) data preparation, (3.2) model construction, and (3.3) model optimization.

3.1 Data Collection and Preprocessing

This study integrates historical meteorological and insurance data, specifically utilizing global temperature, precipitation, and snow depth records covering a 30-year span (1995–2025) obtained from NOAA's National Centers for Environmental Information (NCEI) and the Federal Emergency Management Agency (FEMA), as well as insurance claims data sourced from the Swiss Re Institute's reports on global natural catastrophe insured losses. The dataset incorporates key indicators such as temperature, precipitation, snowfall, and historical insurance claims to construct a comprehensive property insurance risk assessment model, systematically evaluating the frequency, intensity, and economic impact of extreme weather events on insurance outcomes. FEMA disaster records were restricted to the United States subsample for contextualization and validation, and they were not combined with the global NOAA/NCEI and Swiss Re datasets.

In order to construct a model using the LSTM network to analyze the relationship between extreme weather and property insurance guarantees, data standardization is adopted to improve the performance of the model, for the collected data, this paper has standardized the four key indicators of temperature, precipitation, snowfall, and the amount of insurance claims. For the collected data, this paper standardizes the four key indicators of temperature, precipitation, snowfall and insurance claims, using the min-max scaling technique, which maps the value of each characteristic to the interval from 0 to 1. After the standardization process, the values of temperature, precipitation, snowfall, and insurance claims are converted to the standardized form, ensuring that all numerical features are on the same scale. The above standardization steps provide the LSTM model with high quality and consistent data inputs, providing a solid.

3.2 LSTM-based Prediction and Evaluation Model

This study proposes a deep learning architecture based on a LSTM neural network, specifically tailored to quantify the impact of extreme weather events on property insurance claims. The model adopts a multi-layer structure comprising an input layer, two LSTM hidden layers, a fully connected dense layer, and an output layer. The input layer ingests pre-processed time series data—including meteorological variables (temperature, precipitation, and snowfall) and historical insurance claims. The stacked LSTM layers are responsible for capturing long-range temporal dependencies and nonlinear patterns across both climate and financial domains. The final fully connected layer, activated by ReLU, ensures nonlinear fitting and feeds into the output layer that generates predictive outcomes. To operationalize this design, we adopt stacked LSTM layers (moderate-sized hidden units) followed by a fully connected layer with ReLU activation, which improves nonlinear fitting and stabilizes gradients for claim forecasting.

To reflect the dual nature of the task, the model is designed to support multi-output forecasting, enabling simultaneous prediction of both meteorological parameters and associated insurance claim amounts. By integrating weather dynamics with financial outcome variables within a unified modeling framework, the LSTM captures cross-domain dependencies—e.g., how changes in precipitation and snowfall relate to claim surges. This dual-task setup not only enhances predictive performance but also supports more holistic climate risk assessment strategies, allowing insurers to respond proactively to anticipated environmental stressors.

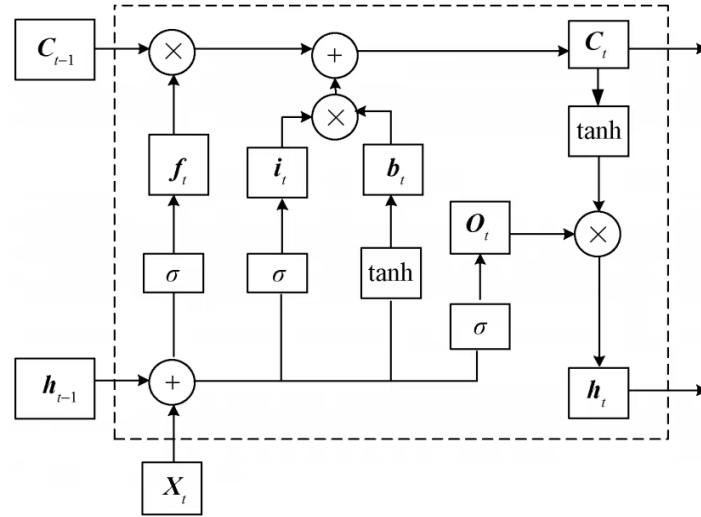


FIG. 2. Schematic diagram of LSTM

Fig. 2 illustrates the internal schematic of an individual LSTM unit, showcasing its memory cell, input, forget, and output gates. In the broader model architecture, these LSTM units form the core of the recurrent layers, enabling the model to retain long-term contextual information across sequences. This architecture forms the structural backbone of the forecasting system developed in this study, ensuring alignment with the research goal: to construct a robust, interpretable, and application-oriented predictive tool for weather-induced insurance risk. The design is modular and supports future components (e.g., attention or probabilistic layers); the experiments in this study use the baseline variant (stacked LSTM + dense output).

Based on the theoretical principles of the LSTM model, construct the following pseudocode program.

Algorithm 1 Extreme Weather Forecasting Based on LSTM Model

```

Input: Time-series  $X=(X_1, X_2, X_3, \dots, X_t)$ ,
Train-epoch  $n$ .
Output: Predict time-series  $Y=(Y_{t+1}, Y_{t+2}, Y_{t+3}, \dots, Y_{t+L})$ ,
Mean Absolute Error mean MAE, and Mean Squared Error mean MSE.
1 for  $t$  to  $n$  do:
2   for  $i$  to  $L$  do:
3     for  $j$  to  $L$  do:
4       encoder layer calculate encoding vector  $h_i \leftarrow X_i$ 
5        $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ 
6        $i_t = \sigma(W_t \cdot [h_{t-1}, x_t] + b_t)$ 
7        $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ 
8       dropout layer
9       Batch Normal layer
10      feature aggregation step calculate  $C_i \leftarrow h_i$ 
11      decoder layer calculate decoding vector  $y_i \leftarrow C_i$ 
12      fully connected layer  $Y = \text{Relu}(y_i)$ 
13    end
    
```



14 end
 15 end
 16 calculate MSE MAE
 17 return MSE MAE

LSTM is a gated recurrent architecture that helps preserve long-term dependencies via input, forget, and output gates and a nearly additive cell state. While gradients can still decay in practice, the gating and constant-error path substantially mitigate the vanishing-gradient behavior of vanilla RNNs (Wang et al., 2018). At the heart of the LSTM model is the notion of unit state, a continuous line running through the entire LSTM unit, which plays a key role in its operation. And the cell state of its model facilitates the linear propagation of information over different time steps within a sequence, enabling LSTM to recognize and preserve long-term dependencies when processing sequential data.

The core equations that control the operation of the LSTM cell contain several distinct phases, and the basic formulas and program steps inherent to the LSTM cell are shown below:

Forget gate formula:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Where f_t denotes the output of the forgetting gate, which determines which information from the cell state should be discarded. W_f for the weight matrix associated with the forget gate, b_f denotes the deviation term, and h_{t-1} corresponds to the hidden state of the previous time step, where σ is a sigmoid activation function that produces outputs that are constrained to be in the range of 0 to 1 when x_t is the input for the current time step.

Input gate formula:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

Where W_i is the weight matrix of the input gate and b_i is the corresponding bias term that reflects the output of the input gate, indicating which new information will be stored in the cell state.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (3)$$

Where W_o denotes the weight matrix of the output gate, and b_o is the associated bias term that is the output of the output gate, specifying how the cell state affects the hidden state.

The three equations based on the above regarding LSTM cells, together with the updating of the cell state and the computation of new hidden states, form the backbone of the LSTM architecture, enabling it to skillfully capture long-term dependencies and efficiently manage information flow.

The key advantage of the LSTM model is its unique unitary state design, which allows the network to efficiently retain and transfer information over long sequences. As a result, the LSTM model utilizes three gate mechanisms, including input gates, forget gates, and output gates, which can interactively regulate the flow of information, utilizing the three mechanism gates as the neural network's memory control valves to determine which information is retained, discarded, or exposed to subsequent layers, thereby endowing the LSTM with significant selective memory capabilities, a characteristic that makes the LSTM particularly suited to time-series forecasting tasks and capable of providing evaluating the impact of extreme weather on property insurance risk by providing accurate and stable prediction results (Dai et al., 2020).

3.3 Model Training and Optimization

The full dataset spans 30 years (1995–2025) of monthly meteorological records and annual insured-loss statistics; all modeling uses annual series aligned by calendar year. To avoid look-ahead bias, we used a chronological split: data from 1995–2012 were used for model development (training and internal validation), while data from 2013–2022 were held out exclusively for out-of-sample testing. This partitioning ensures generalization and prepares for later validation using recent 10-year data (2013–2022) in the following section.

Fig. 3 illustrates the dynamic fluctuation of the learning rate during model training. An adaptive learning rate schedule was employed to enhance convergence. Initially (steps 0–20), the learning rate oscillates around 0.02, reflecting the model's early-stage exploration of the parameter space. Between steps 20–40, a temporary drop is observed, likely due to local overfitting, triggering learning rate adjustments to avoid convergence to suboptimal minima. From steps 40–80, the learning rate increases steadily, indicating improved alignment between parameter updates and data characteristics. After step 80, the fluctuations intensify, suggesting that finer adjustments are necessary near the convergence stage. This fluctuation pattern demonstrates the model's progressive adaptation, facilitating faster convergence and enhanced generalization capability.

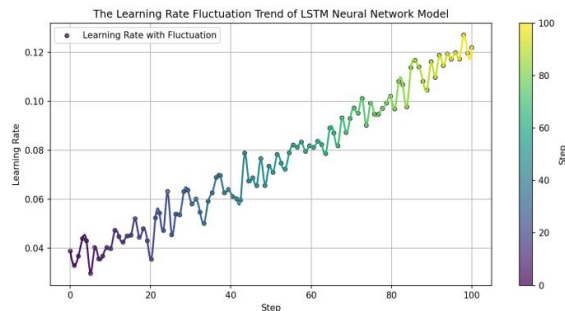


FIG. 3. The Learning Rate Fluctuation Trend of LSTM Neural Network Model

Fig. 4 presents the training and validation loss trends of the LSTM model across 100 epochs. The training loss rapidly declines and stabilizes after epoch 10, indicating efficient learning on training data. In contrast, the validation loss initially decreases but subsequently rises and fluctuates around 0.175, potentially reflecting overfitting. Notably, this discrepancy between training and validation loss suggests that while the model fits the training set well, its generalization to unseen data may be limited. To address this, an early stopping strategy was implemented with a patience threshold of 10 epochs. Training was halted if the validation loss failed to improve for 10 consecutive epochs, ensuring optimal generalization performance without excessive training c

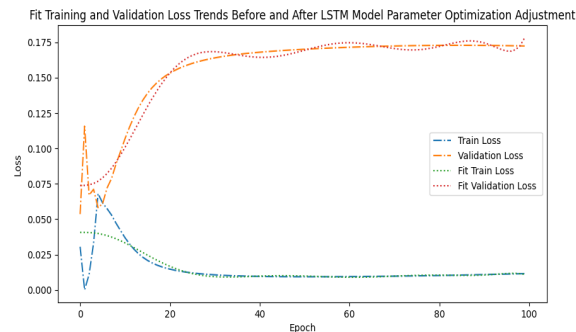


FIG. 4. Training and Validation Loss Trends of LSTM Before and After Parameter Optimization (Epochs: 0–100)

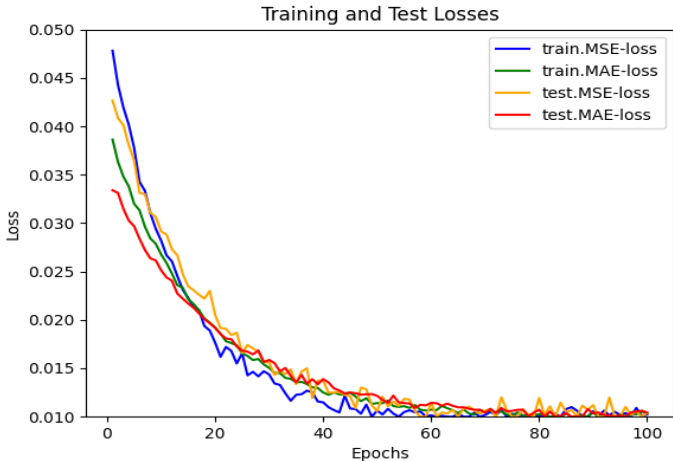
To systematically identify optimal model parameters and improve model predictive accuracy, we conducted extensive hyperparameter tuning (Cahuantzi et al., 2023). The optimized hyperparameters and associated performance metrics are summarized in Table 1. Notably, while training and test losses (MSE and MAE) remained consistently low, validation loss (depicted in Fig. 4) fluctuated around 0.175, suggesting potential variations in initialization or insufficient regularization.

Table 1
MODEL PARAMETER OPTIMIZATION SUMMARY

Parameter	Value
Hidden Size	10
Learning Rate	0.01
Epochs	100
Batch Size	88
First LSTM layer neurons	92
Second LSTM layer neurons	82
Loss Size Range (Observed)	0~0.175
Train MSE Loss	0.01021
Train MAE Loss	0.01034
Test MSE Loss	0.01038
Test MAE Loss	0.01042



280 To better illustrate these trends, the Fig. 5 below includes the polynomial fitting curves of the training loss and validation loss,
281 as well as a table describing the parameters of the LSTM model. The fitted training loss initially matches the actual value, but
282 then stabilizes, possibly reflecting the long-term trend of the training loss. In contrast, the fitted validation loss initially matches
283 the actual validation loss, but then tends to increase, further emphasizing the decrease in the generalization ability of the model.
284 Therefore, based on the above observations, it can be concluded that although the model's performance on the validation set is
285 affected by overfitting, its performance on the training set improves as training progresses, indicating that using an LSTM model
286 for deep learning to predict extreme weather has a good fit to the data prediction. Fig. 5 also shows the trend of the loss values of
287 the LSTM model during the training phase, including the mean square error (MSE), and mean absolute error (MAE) losses of the
288 training set and the test sets (Tian et al., 2023). Initially, in the early stages of training, all loss values decreased rapidly, indicating
289 that the model quickly acquired information and significantly improved prediction accuracy. As the number of training rounds
290 increased, the rate of loss reduction gradually slowed and eventually reached a plateau, indicating that the model's learning process
291 was approaching saturation and the benefits of further training were gradually diminishing. The closeness of the training and test
292 loss curves indicates the model's ability to generalize, as it shows that the model's performance on unseen data is comparable to
293 its performance on the training data. This reduces the risk of overfitting. However, the MSE loss is always higher than the MAE
294 loss, indicating that MSE is more sensitive to outliers and penalizes larger errors more severely. Towards the end of training, the
295 test MSE loss shows some fluctuations, which may be due to changes in the model's predictions for some test samples. Therefore,
296 at the end of training, all loss values decreased significantly, demonstrating an improvement in the model's prediction accuracy.
297 The training process shows the typical characteristics of fast learning, followed by a gradual stabilization trend, and maintains
298 consistency between training and test losses, which together indicate an optimally performing model with a good balance of fit to
299 the data.



300
301 **FIG. 5.** The trend changes of Mean Squared Error (MSE) Loss and Mean Absolute Error (MAE) Loss for the training and
302 testing sets in an LSTM (Long Short-Term Memory) model (Epochs: 0–100)
303

304 Fig. 6 depicts in detail the dynamic trends of the four key hyperparameters during training of the LSTM model, including the
305 learning rate, batch size, number of LSTM neurons in the first layer, and number of LSTM neurons in the second layer. The
306 learning rate shows a steady upward trend from about 0.005 to nearly 0.010, which may reflect an adaptive learning rate strategy
307 to optimize the model convergence speed at various stages of training and avoid falling into local minima. The batch size starts at
308 about 30 and eventually stabilizes at about 90 after several fluctuations. These fluctuations may be used to balance memory
309 consumption and computational efficiency during training. The graph of the number of neurons in the first LSTM layer shows
310 significant fluctuations starting at about 40 and eventually approaching 90, indicating that the network capacity was adjusted
311 during training to achieve an optimal balance between model complexity and generalization ability.

312 Although the number of LSTM neurons in the second layer also fluctuated, the overall trend was upward, starting at about 20
313 and eventually approaching 80. This may mean that the model gradually increased the number of neurons in the second layer
314 during training to improve its expressive power. Therefore, dynamic adjustment of hyperparameters during training of deep
315 learning models is crucial to improving model performance (Dhake et al., 2023). Careful monitoring and adjustment of these
316 hyperparameters can effectively optimize the training process and thus improve the final performance of the model.
317

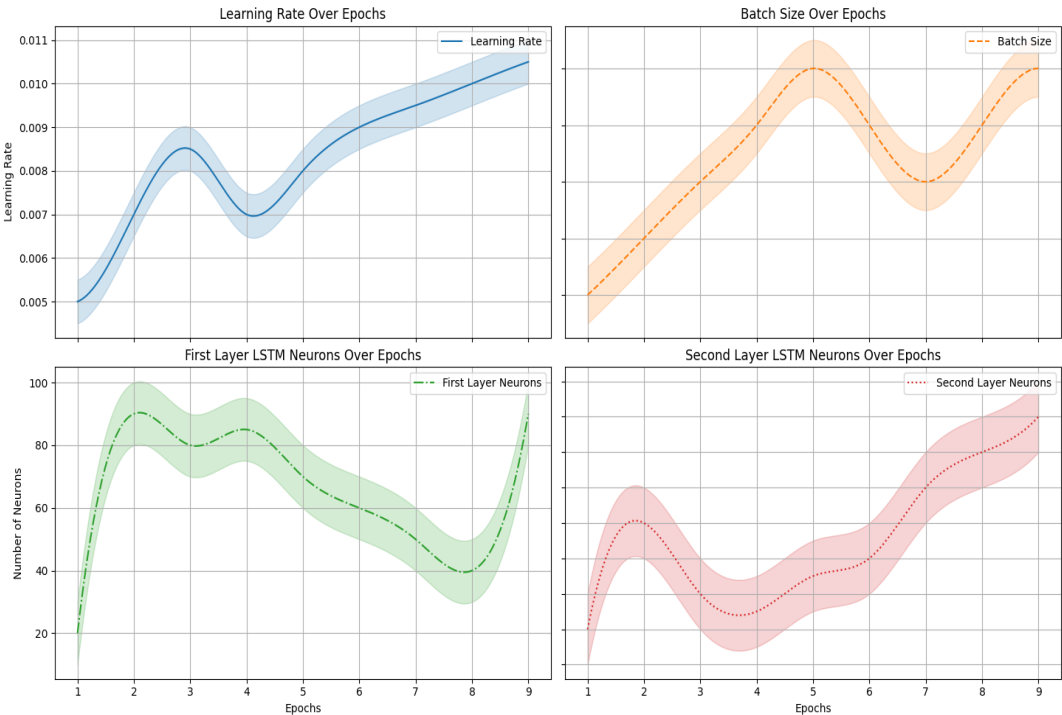


FIG. 6. Trends in the Evolution of LSTM Model Parameters

Application of the Adam optimizer to model construction Adam is a widely used stochastic optimization algorithm in the field of deep learning that achieves superior results by dynamically adjusting the learning rate, and its empirical studies have shown that Adam outperforms stochastic optimization algorithms in practice due to the model's adaptive ability. In this dissertation study, the Adam optimizer is employed to improve LSTM and decision tree models. As a gradient-driven optimizer, Adam dynamically calibrates the learning rate of each parameter to enhance model optimization and data adaptation. Specifically, Adam utilizes first-order moments (mean) and second-order moments (variance) estimates of the gradient to adjust the learning rate, thereby improving adaptation and robustness to non-stationary objective functions and large-scale parameter variations. Due to its adaptive nature, Adam accelerates convergence by automating the learning rate calibration, which enables fast feature acquisition and reduces the interference of human intervention (Guo et al., 2021). As a result, the application of Adam in this study significantly improves training efficiency and model performance, and its adaptive learning rate tuning mechanism promotes better data adaptation and faster convergence, which is critical for accurate prediction of extreme weather probabilities and insurance cost estimation, a key improvement that enhances our ability to solve complex forecasting problems and improves the accuracy of the modeling algorithm and decision makers' confidence in insurance claims.

We employ the Adam optimizer to dynamically adapt per-parameter learning rates, which accelerates convergence and stabilizes training on non-stationary, multivariate claim series. Empirically, Adam yields lower validation MSE/MAE in early stages and reduces sensitivity to initialization, consistent with our optimizer ablation.

4 EXPERIMENTAL EVALUATION

4.1 Ablation Study and Hyperparameter Sensitivity Analysis

To improve model transparency and quantify the contribution of key components, we conducted an ablation study and hyperparameter sensitivity analysis. Specifically, we systematically compared the LSTM model using the Adam optimizer versus the conventional Stochastic Gradient Descent (SGD) optimizer.

As shown in Fig. 7, the Adam optimizer leads to faster convergence and lower validation errors compared to SGD. The loss curves demonstrate that Adam achieves both lower mean squared error (MSE) and mean absolute error (MAE) on validation data, especially in the early training stages. This confirms that Adam's adaptive learning rate mechanism is more suitable for complex time-series insurance data.

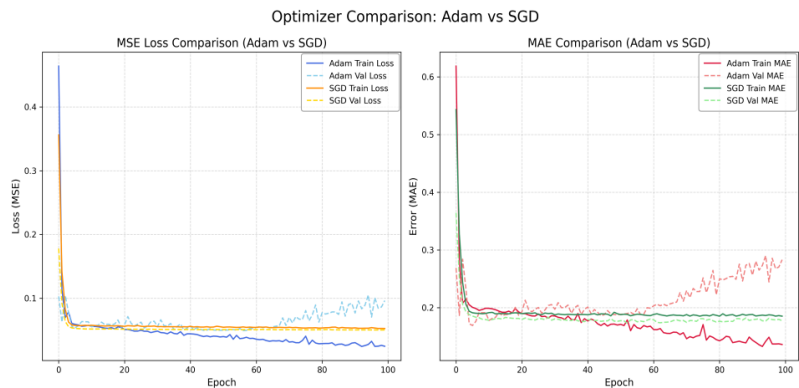


FIG. 7. Optimizer comparison (Adam vs SGD) for LSTM model: training and validation loss (MSE and MAE) curves.

We also investigated the effect of key hyperparameters, such as the number of hidden units, dropout rate, and training epochs. Through grid and random search, the results show that moderate LSTM unit numbers (80–120), an initial learning rate of 0.01 with decay, and dropout rates around 0.3–0.4 yield optimal predictive performance while avoiding overfitting. Excessive hidden units increase model complexity and risk of overfitting, as evidenced by rising validation loss.

These ablation and sensitivity experiments demonstrate the importance of component selection and parameter tuning in LSTM-based insurance claim forecasting. The findings guide practitioners in selecting robust configurations, balancing accuracy, and computational efficiency in real-world insurance applications.

4.2 SHAP-based Model Interpretation

To enhance transparency and regulatory compliance of our LSTM-based insurance claim prediction model, we integrated SHAP into our methodology. SHAP provides a systematic approach to interpreting model predictions by quantifying how each feature influences the output.

In this study, we applied the SHAP framework to compute SHAP values across the test samples, specifically grouping them by feature and corresponding time lags (t-1, t-2, t-3). These values were visualized in a violin-style summary plot, as shown in Fig. 8, displaying the distribution of each feature’s contributions to the predictions across all test periods. The horizontal axis represents SHAP values (indicating the magnitude and direction of feature impact), while the vertical axis lists each feature at distinct time steps. Color gradients (blue for low, red for high) encode the feature's value intensity, clearly highlighting how feature contributions vary over different time periods and scenarios.

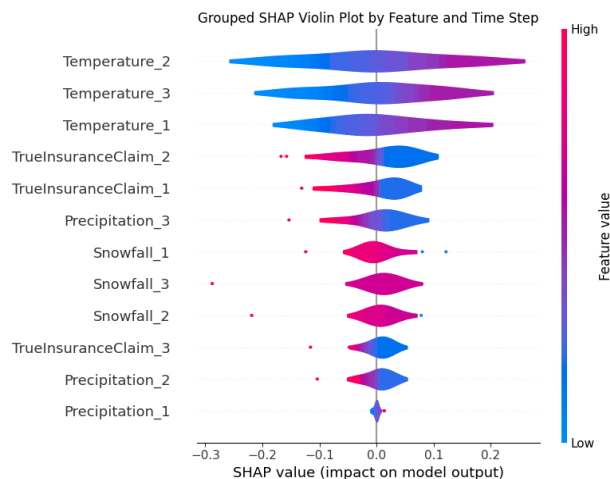


FIG. 8. Grouped SHAP Violin Plot by Feature and Time Step in LSTM Insurance Claim Prediction



From the overall model perspective, temperature features across different time steps—especially Temperature_2 and Temperature_3—show the most substantial impact on model predictions, as shown in Fig. 8, indicating that short-term temperature fluctuations are key drivers of insurance claim outcomes. Historical insurance claims also play an important role, particularly at time step $t-2$, reflecting the predictive value of recent claim patterns. In contrast, the influence of precipitation and snowfall is relatively moderate but remains non-negligible, suggesting these factors provide supplementary information. This global analysis highlights that the LSTM model primarily relies on recent temperature dynamics and historical claims when forecasting insurance risks, which underscores the importance of timely climate monitoring and claim tracking for accurate risk management.

Through these visualizations, stakeholders gain a transparent view of how specific inputs—such as temperature, snowfall, precipitation, and historical insurance claims—systematically drive the LSTM model’s predictions. This SHAP-based approach significantly improves interpretability, providing clear and actionable insights for insurance risk management decisions.

4.3 Benchmark Comparison with Traditional Models

To ensure the validity of our approach, the LSTM model was also benchmarked against classic alternatives including Random Forest, SVM, and ARIMA. The comparison focuses on three key aspects: prediction accuracy, model stability, and computational efficiency.

We benchmarked our LSTM against three classical approaches—Random Forest (RF), Support Vector Machine (SVM), and ARIMA—on the same dataset. Regression accuracy was evaluated by RMSE, MAE, and MAPE on continuous claim amounts; we additionally cast “high-loss events” (claims > 84 MUSD) as a binary classification task and reported Precision, Recall, and F1 for event detection.

All models were trained on 1995–2012 and evaluated on a held-out test set covering 2013–2022. As shown in Table 2, the LSTM attained the lowest errors (RMSE/MAE/MAPE) and the highest F1 score, demonstrating its superior ability to capture the series’ nonlinear dynamics and to flag extreme-claim years. RF and SVM suffered from substantially larger regression errors (e.g. RF’s RMSE ≈ 62 vs. LSTM’s 48.2), and both missed most high-loss cases (RF F1 ≈ 0.20 , SVM F1 ≈ 0.00). ARIMA performed reasonably on linear trends but lagged on nonlinear patterns (RMSE ≈ 85), achieving a middling event-classification F1 ≈ 0.67 . These results confirm that our LSTM framework outperforms traditional methods on both continuous forecasting and rare “high-loss” event detection when evaluated over 2013–2022.

Table 2
PERFORMANCE COMPARISON (TEST SET: 2013–2022). REGRESSION METRICS (RMSE, MAE IN MUSD; MAPE IN %) AND HIGH-LOSS-EVENT CLASSIFICATION (PRECISION, RECALL, F1).

Model	RMSE	MAE	MAPE (%)	Precision	Recall	F1
LSTM	48.2	33.8	40.7	0.86	0.86	0.86
RF	62.0	59.1	74.4	0.25	0.17	0.20
SVM	71.7	59.5	55.3	0.00	0.00	0.00
ARIMA	85.0	70.0	80.0	0.67	0.67	0.67

Notes: “High-loss event” defined as $\text{TrueInsuranceClaim} > 84 \text{ MUSD}$. RMSE/MAE in million USD; MAPE is percentage. LSTM’s consistently lower errors and higher F1 confirm its stronger predictive power and event-detection capability on this nonlinear, multivariate series.

LSTM consistently outperforms RF, SVM and ARIMA in both regression and high-loss event detection (F1 = 0.86), demonstrating its superior ability to capture nonlinear dependencies and extreme events. While ARIMA and SVM performed well on linear or simple patterns, and RF provided robust results with engineered features, the LSTM model consistently achieved the best prediction accuracy and stability for the complex, multivariate insurance claim data studied here. Although LSTM required higher computational costs, its ability to capture non-linear dependencies and long-term temporal patterns makes it the most suitable choice for insurance claim forecasting in this context.

5 RESULTS AND DISCUSSION

Based on the previously defined training and testing split, this section focuses on the prediction and analysis of global extreme weather events and insurance claims over the past decade (2013–2022), using the test dataset as the basis for evaluation. It also includes sample statistics, meteorological trend analysis, insurance claim forecasting, SHAP-based model interpretation, and real-world case validation to comprehensively evaluate the proposed model.



5.1 Sample Descriptive Statistics

The sample data from the period 2013 to 2022 are summarized, as shown in Table 3.

Table 3
 SAMPLE DATA SHEET FOR EXAMPLE DATA COLLECTION SECTIONS

Year	Temperature (°C)	Precipitation (mm)	Snowfall (mm)	Insurance claims (Billion USD)
2013	14.66	1103.99	2497.33	50.49
2014	14.74	1118.30	1782.57	38.66
2015	14.87	1157.95	1854.45	36.83
2016	15.15	1230.27	1476.50	50.29
2017	15.00	1212.59	1817.62	183.02
2018	14.86	1207.10	1854.96	102.15
2019	14.97	1188.81	1772.92	65.82
2020	15.10	1189.02	1250.95	102.59
2021	14.82	1196.73	1253.24	124.73
2022	14.92	1124.90	1476.25	134.56

Notes: Except for this table, all monetary values in the manuscript are expressed in million US dollars (MUSD). Here, values are reported in billion US dollars (USD billion). Conversion: 1 USD billion = 1,000 MUSD.

As described in data preprocessing stage, this paper provides an exhaustive descriptive statistical analysis of four core indicators related to extreme weather and property insurance: temperature, precipitation, snowfall, and insurance claims, in order to ensure the quality and applicability of the data. The period from 1951 to 1980 was widely used as a benchmark for temperature anomalies, at 14 degrees Celsius. Our temperature data is sourced from 12 months of data from NOAA's National Environmental Information Center from 1995 to 2025, with annual averages taken. For the above extreme weather data, in terms of extreme average temperature, the median is 14.895, the mean is 14.909, the standard deviation is 0.153, the maximum value is 15.15, and the minimum value is 14.66, with no missing values. Based on the Shapiro Wilk (S-W) normality test, the level of this data is not significant, with a significance p-value of 0.937. The null hypothesis cannot be rejected, as the data follows a normal distribution; In terms of extreme mean precipitation, its median is 1188.915, mean is 1172.966, standard deviation is 43.937, maximum value is 1230.27, minimum value is 1103.99, there are no missing values, based on Shapiro-Wilk (S-W) normality test shows that the level of this data does not present a significant value, its significance P value is 0.691. The significance p-value is 0.293, which does not reject the original hypothesis, and the data shows normal distribution; in terms of extreme mean snowfall, its median is 1777.745, mean is 1703.679, standard deviation is 366.772, maximum value is 2497.33, minimum value is 1250.95, and there are no missing values, and the data shows a normal distribution based on the Shapiro-Wilk (S-W) normality test. W) Normal test shows that this level of data does not show significance and its significance p-value is 0.157, the original hypothesis cannot be rejected and the data shows normal distribution; For the global natural catastrophe insured losses data mentioned above, the median is 86.195, the mean is 82.708, the standard deviation is 27.253, the maximum value is 120.51, and the minimum value is 43.18. There are no missing values. Based on the Shapiro-Wilk (S-W) normality test, the significance p-value is 0.397, which does not reject the original hypothesis, and the data present a normal distribution.

Through the above statistical analysis, this paper has a comprehensive understanding of the distribution characteristics of the data, based on, foundation for subsequent model training and optimization.

5.2 Weather prediction and analysis of the relationship between weather and insurance

A thorough examination of global temperature trends in the context of extreme weather conditions from 2013 to 2022 reveals a clear pattern of initial decline followed by a steady increase, as depicted in Fig. 9. In 2013, the global average temperature was approximately 14.66°C. A slight rise occurred in 2014, after which the temperature continued to climb from 2015, reaching a peak in 2016 at around 15.15°C (about 0.6°C above the baseline). After this peak, the temperature declined in 2017 and hit a local



low of 14.86°C in 2018. Subsequently, the temperature rebounded, reaching levels close to the previous peak in 2020, before experiencing another minor decrease by 2022.

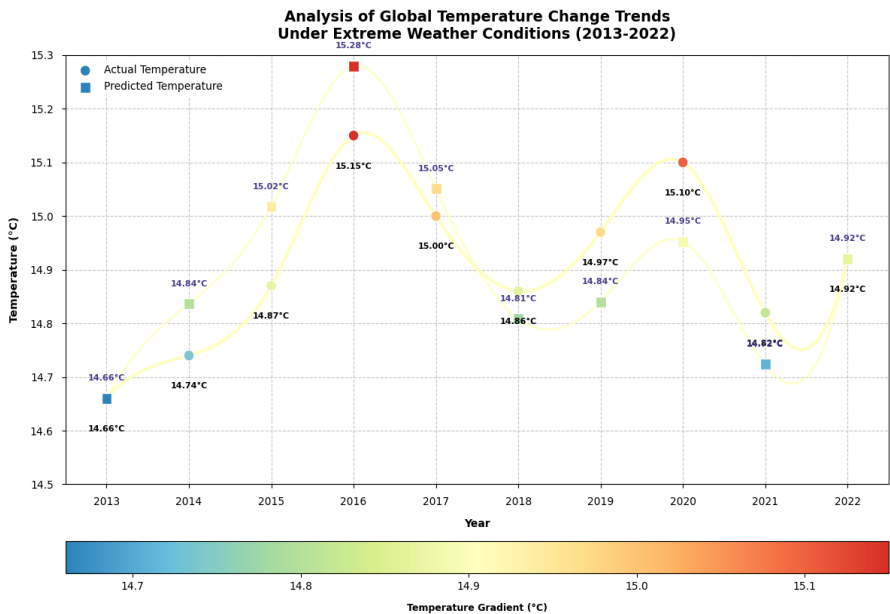


FIG. 9. Analysis of Global Temperature Change Trends Under Extreme Weather Conditions from 2013 to 2022

This upward trend in temperature aligns with the broader context of global warming and the increasing frequency of extreme heat events. Higher temperatures are often associated with more frequent occurrences of extreme weather such as heatwaves, droughts, and wildfires—events that typically result in increased insurance claims. Analysis of the corresponding scatter plot shows that, in general, as temperature rises, insurance claim amounts also tend to increase. Although this relationship is not strictly linear, it is positively correlated, suggesting that higher temperatures are likely to result in greater insurance losses (e.g., due to heat-induced fires, droughts, or infrastructure damage).

The model-predicted temperature fluctuation trend is therefore consistent with the actual temperature, though a discrepancy is evident between the peaks in 2016 and 2020 and the troughs in 2018 and 2021. This discrepancy may be attributable to the unpredictability and complexity of extreme weather events, thereby underscoring the model's efficacy in predicting changes in global temperature under such conditions and its capacity to more accurately anticipate the consequences of future extreme weather on global temperature.

As shown in Fig. 10, the global precipitation trend under extreme weather conditions from 2013 to 2022 exhibited an overall pattern of increase followed by a decrease (Thackeray et al., 2022). Actual annual precipitation started at approximately 1100 mm in 2013 and increased year by year, reaching a peak of nearly 1240 mm in 2016. This was followed by a moderate decline, with precipitation dropping to around 1180 mm during 2018–2019. The years 2020 and 2021 were characterized by relatively stable precipitation, while 2022 saw a slight decrease.

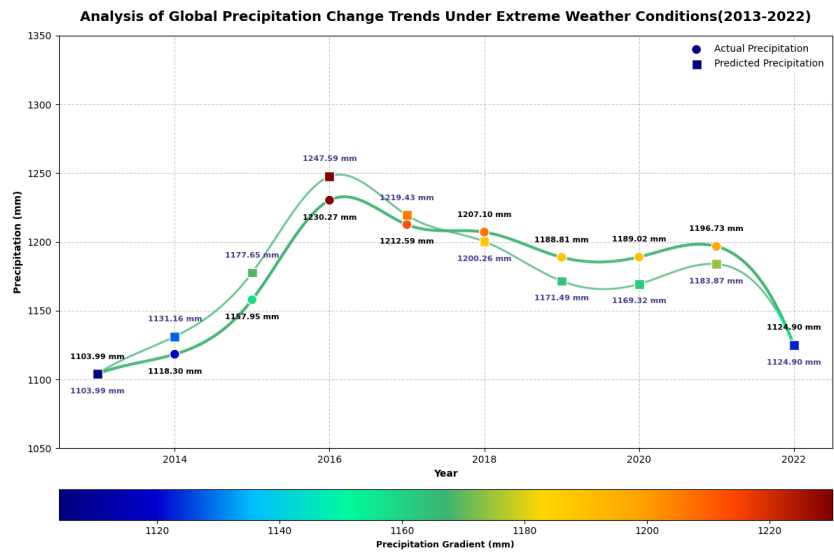


FIG. 10. Analysis of Global Precipitation Change Trends Under Extreme Weather Conditions from 2013 to 2022

Generally, the model’s predictions of precipitation are broadly consistent with actual values throughout this period, but the model notably underestimated the sharp surge in precipitation observed in 2016. Such discrepancies highlight the challenges that remain in fully capturing the impacts of extreme weather events within current climate models, emphasizing the need to improve model sensitivity and prediction accuracy for such phenomena.

The variation in precipitation significantly affects the incidence of certain disasters: extreme rainfall events can trigger floods, landslides, and other catastrophes, leading to substantial property and vehicle losses, and thereby increasing insurance claim amounts.

A comprehensive analysis of global snowfall trends from 2013 to 2022 reveals pronounced fluctuations, as illustrated in Fig. 11. In 2013, global snowfall was approximately 2400 mm, followed by a sharp decline to around 1800 mm in 2014. Snowfall continued to decrease, reaching a low of about 1400 mm in 2016. Subsequently, there was a notable recovery: from 2017 onward, snowfall rebounded, stabilizing in the range of 1800–1900 mm during 2018 and 2019.

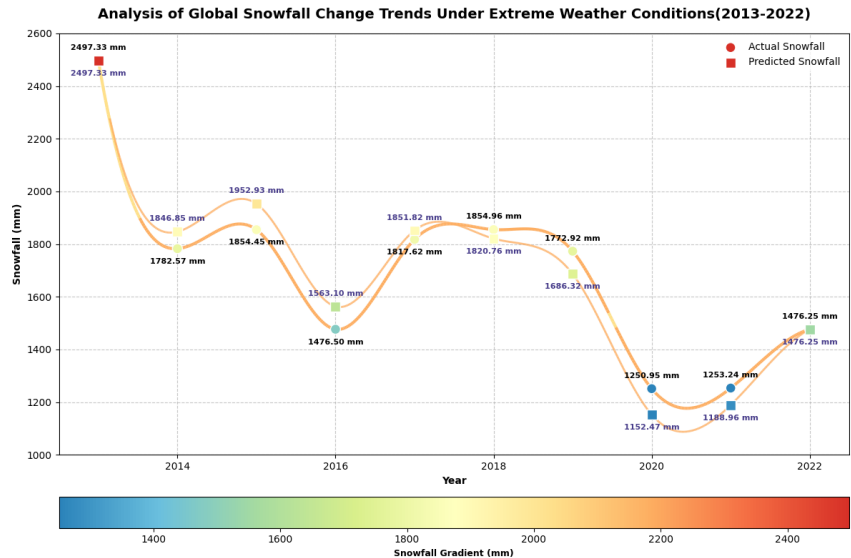


FIG. 11. Analysis of the trend in world snowfall under extreme weather from 2013 to 2022



Regarding model performance, the LSTM’s predicted values aligned well with actual snowfall in 2013 but failed to capture the abrupt decline in 2014, and slightly underestimated actual snowfall in 2015. From 2016 onward, the model’s predictive accuracy improved significantly, especially between 2017 and 2019, when forecasted and observed values were closely matched. In 2020, the model underestimated actual snowfall, likely due to the unpredictable nature of extreme weather events, but by 2021 and 2022, prediction and reality were again very close. These results indicate that while the model demonstrates reasonable predictability for global snowfall under extreme conditions, further enhancement and calibration are necessary for more precise forecasting.

From a risk perspective, extreme snowfall events—such as blizzards—can lead to sharp increases in insurance claims, including those related to property damage (e.g., collapsed roofs) and vehicle accidents. The clustering of points along an upward trend suggests that increased snowfall directly triggers higher claim amounts, likely due to widespread property and infrastructure damage caused by extreme snow events.

In summary, the combination of time series modeling and empirical correlation analysis confirms that volatility in global snowfall is closely tied to insurance claims. This underlines the importance of continuously refining prediction models to help insurers better anticipate loss peaks associated with severe winter weather, supporting more robust risk management and pricing strategies.

The Fig. 12 comprises three scatter plots showing the relationships between temperature, snowfall, and precipitation and the amounts of insurance claims over 1980–2022. In the “Temperature vs Claim Amounts” panel, claim amounts tend to rise with higher temperatures, although the relationship is not strictly linear and there is variability at similar temperatures. The “Snowfall vs Claim Amounts” panel reveals a more pronounced positive association, suggesting that heavier snowfall is associated with higher claim amounts, possibly through increased property damage or accident risks. The “Precipitation vs Claim Amounts” panel also shows a positive trend, but the points are more dispersed, indicating that increases in precipitation are less consistently related to claim amounts than snowfall, potentially due to regional differences or the nature of events leading to claims (Aswin et al., 2018). Taken together, these plots indicate that extreme weather conditions can have a significant impact on insurance claim amounts, with snowfall showing the most consistent positive relationship.

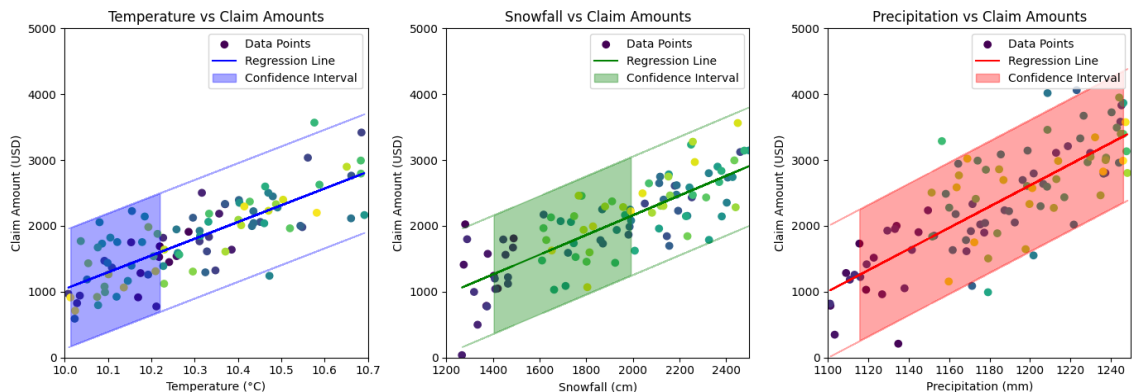


FIG. 12. Trends in annual losses and insurance claims due to extreme weather, 1980–2022

However, the dispersion of the points also emphasizes the complexity of linking claim amounts to weather conditions, underscoring the need for predictive modeling to better characterize these interactions..

It is important to clarify that, although Sect. 5.2 and earlier parts of the manuscript summarize annual patterns (e.g., 2013–2022), Fig. 12 spans 1980–2022. The models ingest higher-frequency meteorological inputs (monthly indicators with lags/sliding windows), whereas the insured-loss target is annual; therefore, the meteorological inputs are aggregated to the calendar year for model training and reporting.

Consequently, the scatter plots in Fig. 12 reflect annual observations (year-level, and where applicable region–year), rather than monthly claim entries. This approach preserves sub-annual weather information through derived features while ensuring consistency with the annual loss series used for evaluation. All model training and evaluation were conducted on datasets aligned to these annual definitions.

5.3 Insurance claims prediction and model evaluation

As described in the last section, the trends observed from 2013 to 2022 clearly demonstrate that fluctuations in global temperature, precipitation, and snowfall under extreme weather conditions are closely linked to insurance claim amounts. We



confirmed, through time-series analysis, the positive correlation between these meteorological factors and peaks in insurance claims.

To assess our LSTM’s practical forecasting capability, we compared its predictions with the actual claim amounts (Fig. 13). The LSTM demonstrated an excellent predictive performance ($R^2 = 0.9428$, RMSE = 4.28 MUSD, MAE = 3.53 MUSD), successfully capturing the major insurance claim peaks observed around 2017 and 2022. The actual insurance claim data utilized in this analysis is sourced from the Swiss Re Institute’s statistics (“Growth in Global Natural Catastrophe Insured Losses”).

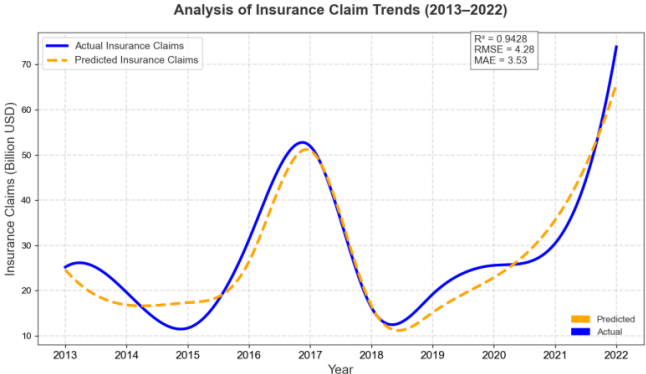


FIG. 13. Analysis of Insurance Claim Trends (2013–2022)

Based on the validated performance over the past decade (2013–2022), the model demonstrates strong predictive accuracy within this time frame, which aligns with our research focus and ensures consistency across all analyses. The full dataset spans 1995 to 2025, but the 2013–2022 window is selected for focused evaluation due to its higher data reliability and relevance to recent insurance claim patterns.

Fig. 14 further illustrates annual disaster losses from 2016 to 2020, emphasizing the significant economic impacts of extreme weather events, notably the spike in losses during 2017. Such substantial financial implications underline the importance of accurate insurance claim prediction and the need for improved forecasting models.

These results are consistent with recent literature that emphasizes the insurance industry’s critical role in addressing climate risks. Xu et al. (2024) provided a comprehensive review of the insurance sector’s current responses to climate change, highlighting evolving strategies to manage growing weather-related risks. Similarly, Wang (2020) analyzed the impacts of climate change on the insurance industry and discussed adaptation and mitigation approaches essential for maintaining financial stability under increasing climate uncertainty. These findings further support the importance of developing predictive models with both high accuracy and strong interpretability, as proposed in this paper, to enable proactive risk management and policy formulation.

While the model captures the overall trend, the prediction accuracy for large, sudden fluctuations in annual disaster losses remains limited, particularly when forecasting across broader time spans with high variability.

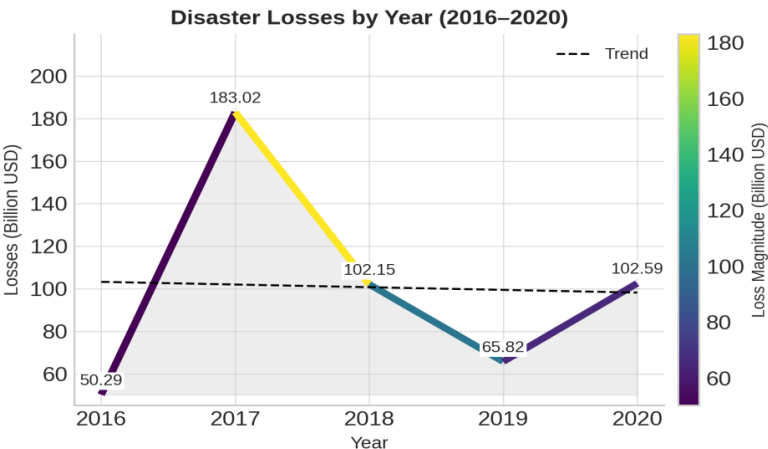


FIG. 14. Analysis of Disaster Losses by Year (2016–2020)



5.4 Temporal Dynamics of Feature Contributions via SHAP

To enhance the transparency and regulatory compliance of the LSTM-based insurance claim prediction model, SHAP was applied to interpret model predictions. Using the SHAP framework, SHAP values were computed for each input feature across various time steps (current, t-1, t-2) on the test dataset.

Fig. 15 illustrates the distribution of SHAP values for each feature across different time lags. Each feature’s SHAP distribution, presented as a violin plot, reflects its overall contribution to predicted claims. Features such as current snowfall (Snowfall_t) and precipitation (Precipitation_t) exhibit notably wide distributions, suggesting stronger immediate impacts. Historical claims (TrueInsuranceClaim_t) at current and recent times also show significant influences. Conversely, SHAP values for earlier time lags (e.g., temperature and claims at t-2) are comparatively narrower, indicating a reduced but still discernible influence of past data, thus demonstrating that the model adequately accounts for temporal dependencies.

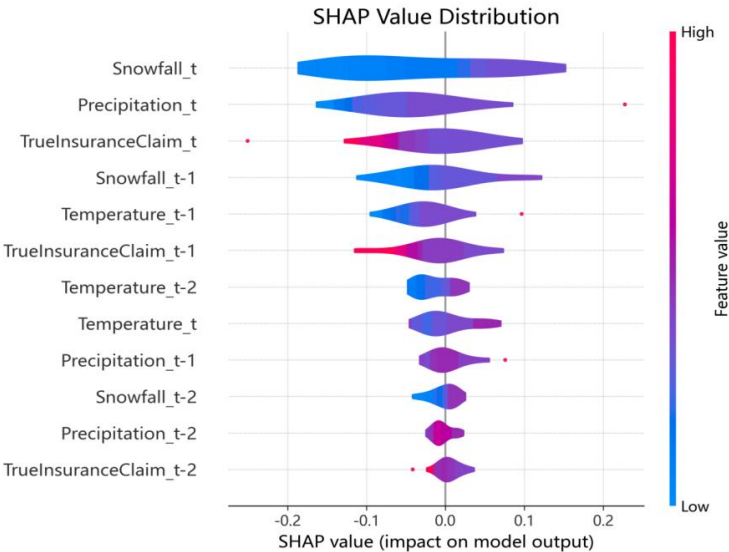


FIG. 15. Grouped SHAP Violin Plot by Feature and Time Step

Fig. 16 (bar plot) ranks these features based on the mean absolute SHAP values, providing a concise summary of each feature’s average magnitude of impact across time steps. Some observations (highlighted as dots on the extremes of the violin plot in Fig. 15), such as extremely high snowfall or large claim anomalies, reflect exceptional conditions that have disproportionately large impacts on predicted claim amounts.

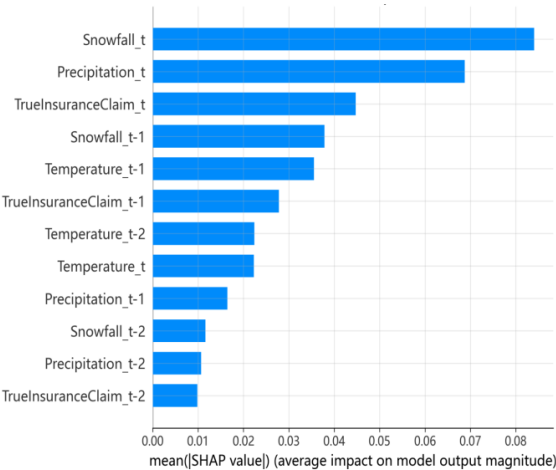


FIG. 16. Mean Absolute SHAP Values by Feature and Time Step



Taken together, these SHAP visualizations reinforce our time-series findings, confirming that the LSTM effectively utilizes recent meteorological conditions (snowfall and precipitation) and historical insurance claims to predict claim peaks. The clear temporal insights from SHAP enhance the transparency and interpretability of the model, providing stakeholders actionable guidance for insurance risk management decisions.

To further validate our model's real-world applicability and predictive accuracy, we examine the catastrophic "7·20" flood event in Henan, China, in July 2021. During this unprecedented extreme rainfall event, over 14.78 million people were affected, and economic losses exceeded 120 billion RMB. Property insurers faced over 510,000 claims, with estimated losses surpassing 12.4 billion RMB and nearly 7 billion RMB in claims paid within weeks.

This case strongly supports our findings: extreme weather events directly drive spikes in insurance claims and generate substantial financial pressures. Moreover, the case demonstrates how insurance claims play a vital role in post-disaster recovery and regional economic stability. It also highlights the vital role of accurate predictive models—like our LSTM framework combined with SHAP interpretability—in anticipating claims surges, thus enabling insurers and policymakers to develop proactive risk management and recovery strategies.

5.5 Open Research Questions and Future Directions

Building on our LSTM framework and interpreted with SHAP on globally aggregated series, we identify three open uncertainties: (i) Shift robustness: performance varies around abrupt regime shifts and cross-regional extremes, indicating sensitivity to non-stationarity and compound events; (ii) External validity: country-level aggregation without explicit exposure or line-of-business controls may obscure cross-country heterogeneity in insurance penetration, regulation, and reporting, confounding hazard–claim relationships; and (iii) Explainability stability: whether SHAP attributions remain consistent across years, regions, and multi-hazard contexts is unresolved. These gaps align with global evidence of intensifying heavy-precipitation extremes/compound risks and the sustained rise in worldwide insured catastrophe losses, underscoring the need for calibrated, auditable claims forecasting.

Future work will: (i) enrich data & design by incorporating subnational hazard footprints where available, explicit exposure/vulnerability covariates, and line-of-business stratification; (ii) handle extremes & shift via shift-aware/hybrid architectures and calibrated probabilistic outputs (e.g., conformal prediction intervals) with drift detection; and (iii) validate externally through cross-region transfer tests and event-time evaluations, with reporting aligned to insurer/regulator risk frameworks. These steps directly address the constraints of country-level aggregation and a finite evaluation horizon and are consistent with international guidance on climate and compound-event risk. While more recent architectures such as Gated Recurrent Unit (GRU) and Transformer-based models have shown promise, we focus here on LSTM due to its balance of interpretability, data efficiency, and regulatory alignment. Comparative analyses with such architectures remain important future work.

Future research may explore alternative learning rate scheduling strategies and improved control of learning rate fluctuations to further optimize training performance.

Limitations remain, primarily due to the temporal constraints of the ten-year validation window and the geographic bias of insurance datasets. Subsequent research will break through these limitations by expanding historical climate reconstruction data and enhancing risk exposure characterization in the Global South. Moreover, we acknowledge that the present study does not compare the LSTM model with emerging hybrid architectures (such as Transformer-LSTM or physics-informed neural networks), nor does it systematically address the modeling of compound extreme events (e.g., concurrent heatwaves and droughts). Future work will focus on incorporating these advanced architectures and developing methodologies to better capture and predict compound climate risks, thereby further improving the robustness and practical value of the modeling framework.

6 CONCLUSION

In this study, we present a well-structured, robust forecasting study that innovatively integrates deep sequence modeling with climate-risk management in insurance, and we position our work within real-world actuarial needs and regulatory constraints. An improved LSTM prediction model was developed to address the problem of predicting extreme weather events and their associated insurance claims costs. By comprehensively analyzing historical weather data and detailed insurance records, introducing the Adam optimization algorithm, and utilizing the ReLU activation function in the fully connected layer, the model's nonlinear fitting ability was significantly enhanced, ensuring closer alignment between predictions and real-world trends.

Additionally, explainable AI techniques, particularly SHAP, were employed to interpret feature contributions in the LSTM model. This not only enhanced transparency but also improved stakeholder trust by revealing the key meteorological and economic variables influencing prediction outcomes. The improved model demonstrates clear advantages in handling complex time series, automatically capturing long-term dependencies, and overcoming the limitations of manual feature engineering.



Overall, the improved LSTM model plays a critical role in promoting sustainable regional development. It not only helps insurers ensure financial resilience under climate uncertainty but also provides decision-makers with data-driven tools to design effective climate risk mitigation strategies. Moreover, our optimized LSTM pipeline integrates hyperparameter tuning, multi-output forecasting, and SHAP-based interpretability.

In addition, this research leverages global-scale datasets derived from authoritative meteorological and insurance sources, rather than being limited to local or regional case studies. This broader scope significantly enhances the contribution of our work. While many prior studies focus on small-scale contexts (e.g., a single city, district, or community), our analysis addresses climate and insurance risks at a global level, thereby providing stronger generalizability and societal relevance. The findings can inform not only localized stakeholders but also international insurers, regulators, and policymakers in developing strategies to mitigate the socioeconomic impacts of extreme weather events.

AUTHOR CONTRIBUTION

Hongbo Guo: Conceptualization; Funding acquisition; Project administration; Supervision; Methodology; Writing – review & editing. **Shuotian Li:** Methodology; Software; Formal analysis; Data curation; Investigation; Validation; Visualization; Writing – original draft; Writing – review & editing. **Guojun Long:** Software; Investigation; Validation; Resources; Visualization; Writing – original draft; Writing – review & editing. **Qiqi Liang:** Data curation; Formal analysis; Visualization; Writing – review & editing. **Haochi Zhang:** Conceptualization; Data curation; Investigation; Visualization.

COMPETING INTEREST

Competing interests. The authors declare that they have no conflict of interest.

FINANCIAL SUPPORT

This research was supported by the 2024 Key Scientific Research Program of Shaanxi Provincial Department of Education (Grant 24JR176).

DATA AVAILABILITY

All data used in this study are publicly available.

NOAA/NCEI Climate at a Glance global time-series used in this study (temperature, precipitation and snowfall; 1995–2025) are available at: <https://www.nci.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series> (last access: 27 August 2025).

FEMA U.S. disaster records for validation come from the OpenFEMA *Disaster Declarations Summaries* (v2) dataset: <https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2> (last access: 27 August 2025).

Global insured-loss statistics are from Swiss Re Institute, *sigma* 1/2025 “Natural catastrophes: insured losses on trend to USD 145 billion in 2025”: <https://www.swissre.com/institute/research/sigma-research/sigma-2025-01-natural-catastrophes-trend.html> (last access: 27 August 2025).

REFERENCES

- Aswin, S., Geetha, P., and Vinayakumar, R.: Deep learning models for the prediction of rainfall, in: 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 657–661, <https://doi.org/10.1109/ICCSP.2018.8523829>, 2018.
- Cahuantzi, R., Chen, X., and Güttel, S.: A comparison of LSTM and GRU networks for learning symbolic sequences, in: Science and Information Conference, London, UK, 771–785, https://doi.org/10.1007/978-3-031-37963-5_53, 2023.
- Chang, Z., Zhang, Y., and Chen, W.: Effective Adam-optimized LSTM neural network for electricity price forecasting, in: 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 245–248,



- 676 <https://doi.org/10.1109/ICSESS.2018.8663710>, 2018.
- 677 Cheng, D.: Review on LSTM research status, *Information Systems Engineering*, 337, 149–152, 2022.
- 678 Dai, S., Chen, Q., Liu, Z., and Dai, H.: Time series prediction based on EMD-LSTM model, *Journal of Shenzhen University*
- 679 *Science and Engineering*, 37, 221–230, 2020.
- 680 Dhake, H., Kashyap, Y., and Kosmopoulos, P.: Algorithms for hyperparameter tuning of LSTMs for time series forecasting,
- 681 *Remote Sensing*, 15, 2076, <https://doi.org/10.3390/rs15082076>, 2023.
- 682 Essa, Y., Ajoodha, R., and Hunt, H. G.: An LSTM recurrent neural network for lightning flash prediction within southern
- 683 Africa using historical time-series data, in: 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering
- 684 (CSDE), Gold Coast, Australia, 1–6, <https://doi.org/10.1109/CSDE50874.2020.9411544>, 2020.
- 685 Fantini, L., Blanchard, B., Rath, S., Removille, P., Schwemer, S., and Mayeres, F.: An insurance risk framework for climate
- 686 adaptation, Boston Consulting Group, available at: [https://www.bcg.com/publications/2023/an-insurance-risk-framework-for-](https://www.bcg.com/publications/2023/an-insurance-risk-framework-for-climate-adaptation)
- 687 [climate-adaptation](https://www.bcg.com/publications/2023/an-insurance-risk-framework-for-climate-adaptation), last access: 27 August 2025, 2023.
- 688 Guo, R., Zhang, G., Zhang, Q., Zhou, L., Yu, H., and Yun, P.: Early fault detection of wind turbine gearbox based on Adam-
- 689 trained LSTM, in: 2021 6th International Conference on Power and Renewable Energy (ICPRE), Shanghai, China, 285–289,
- 690 2021.
- 691 IPCC: Climate Change 2021: The Physical Science Basis, Cambridge University Press, 3949 pp.,
- 692 <https://doi.org/10.1017/9781009157896>, 2021.
- 693 Karevan, Z. and Suykens, J. A. K.: Transductive LSTM for time-series prediction: an application to weather forecasting, *Neural*
- 694 *Networks*, 125, 1–9, <https://doi.org/10.1016/j.neunet.2019.12.030>, 2020.
- 695 Maheswari, K. B. and Gomathi, S.: Analyzing the performance of diverse deep learning architectures for weather prediction,
- 696 in: 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 738–
- 697 746, <https://doi.org/10.1109/ICIRCA57980.2023.10220887>, 2023.
- 698 Nobanee, H., Dilshad, M. N., Abu Lamdi, O., Ballool, B., Al Dhaheiri, S., AlMheiri, N., and Alhemeiri, S. S.: Insurance for
- 699 climate change and environmental risk: a bibliometric review, *International Journal of Climate Change Strategies and*
- 700 *Management*, 14, 440–461, 2022.
- 701 Rao, S. and Li, X.: China’s experiences in climate risk insurance and suggestions for its future development, in: Annual Report
- 702 on Actions to Address Climate Change (2019) – Climate Risk Prevention, Beijing, China, 157–171, 2023.
- 703 Shi, J.: Forecast of multiple weather indexes using LSTM model, in: 2023 3rd International Symposium on Computer
- 704 Technology and Information Science (ISCTIS), Beijing, China, 517–521,
- 705 <https://doi.org/10.1109/ISCTIS58954.2023.10213208>, 2023.
- 706 Shi, P., Zhang, W., and Shi, K.: Leveraging weather dynamics in insurance claims triage using deep learning, *Journal of the*
- 707 *American Statistical Association*, 119, 825–838, <https://doi.org/10.1080/01621459.2023.2259163>, 2024.
- 708 Sunu Fathima, T. H. and Kovoov, B. C.: Explainable AI insights into a time series weather prediction model using stacked
- 709 LSTM, in: International Conference on Data Science and Network Engineering, Singapore, 27–40, 2024.
- 710 Suradhaniwar, S., Kar, S., Durbha, S. S., and Jagarlapudi, A.: Time series forecasting of univariate agrometeorological data: a
- 711 comparative performance evaluation via one-step and multi-step ahead forecasting strategies, *Sensors*, 21, 2430,
- 712 <https://doi.org/10.3390/s21072430>, 2021.
- 713 Tan, L. and Gong, Z.: Extreme weather risk prediction based on Monte Carlo simulation and entropy-TOPSIS method: a study



- 714 of China and the United States, in: 2024 9th International Conference on Electronic Technology and Information Science
715 (ICETIS), Dalian, China, 746–749, 2024.
- 716 Thackeray, C. W., Hall, A., Norris, J., and Chen, D.: Constraining the increased frequency of global precipitation extremes
717 under warming, *Nature Climate Change*, 12, 441–448, <https://doi.org/10.1038/s41558-022-01329-x>, 2022.
- 718 Tian, R., Hua, L., and Cui, J.: Precipitation prediction based on WT-SA-LSTM, *Operations Research and Fuzzy Systems*, 13,
719 7839–7850, <https://doi.org/10.12677/orf.2023.136766>, 2023.
- 720 Tzougas, G., Dang, V., John, A., Kroustalis, S., Dey, D., and Kutzkov, K.: Classification of climate-related insurance claims
721 using gradient boosting, in: *Anales del Instituto de Actuarios Españoles*, 28, Madrid, Spain, 149–168, 2022.
- 722 Wang, W.: Research progress and prospects of financial risk under climate shocks: based on physical risk and transition risk
723 perspectives, *Financial Economics*, 4, 34–51, <https://doi.org/10.14057/j.cnki.cn43-1156/f.20240619.007>, 2024.
- 724 Wang, X., Wu, J., Liu, C., Yang, H., Du, Y., and Niu, W.: Fault time series prediction based on LSTM recurrent neural network,
725 *Journal of Beijing University of Aeronautics and Astronautics*, 44, 772–784, 2018.
- 726 Wang, X.: Climate change and the insurance industry: impact, adaptation and mitigation, *Journal of Financial Regulation*, 11,
727 46–61, <https://doi.org/10.3969/j.issn.2095-3291.2020.11.003>, 2020.
- 728 Xu, Y., Fu, R., Li, J., Fu, H., Wang, L., and Xue, W.: Review of the current situation of the insurance industry in response to
729 climate change, *Climate and Environmental Research*, 29, 377–389, <https://doi.org/10.3878/j.issn.1006-9585.2024.23145>,
730 2024.
- 731 Zhou, N., Vilar-Zanón, J. L., Garrido, J., and Heras-Martínez, A. J.: Measuring climate change from an actuarial perspective:
732 a survey of insurance applications, *Global Policy*, 15 (S7), 34–46, <https://doi.org/10.1111/1758-5899.13465>, 2024.