Journal of Management and Economic Sciences Prospects | Volume: 10 / N°: 01 (2026), p 302-317

A Hybrid Deep Learning Model for Systemic Insurance Risk Prediction: Integrating Climate Catastrophe and Telematics Data

Hamel Dalila [*1], Djafri Asma [2]

[1] University of Ferhat Abbas Sétif 1 ,Algeria, dalila.hamel@univ-setif.dz

[2] University of Ferhat Abbas Sétif 1 ,Algeria, Country, asma.djafri@univ-setif.dz

Received:15/02 /2026 *Accepted: 07/07/2026* *Published:30/07/2026*

Abstract:

In response to the growing challenges posed by climate-related catastrophes and dynamic policyholder behaviors, this study develops a hybrid deep learning framework combining Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Explainable AI (XAI) techniques.

Using multi-modal data streams from satellite climate indices, telematics driving patterns, and insurance claims, the model outperforms traditional actuarial methods with a 15–22% improvement in predictive accuracy.

The integration of SHAP-based interpretability ensures compliance with regulatory transparency requirements.

Results demonstrate the framework's potential for adaptive underwriting, catastrophe pooling, and proactive systemic risk management in the insurance sector.

Keywords: Systemic Risk; Deep Learning; Insurance Analytics; Telematics; Climate Catastrophes; Explainable AI.

Jel Classification Codes: G22, C55, Q54.

[*] ***Corresponding author***

1. INTRODUCTION

The insurance industry is increasingly exposed to **systemic risks** driven by two converging forces:

1. The intensifying frequency and severity of climate-related catastrophes.

2. The growing impact of behavioral risk factors linked to policyholder activities.

These developments challenge the adequacy of **traditional actuarial models** such as Generalized Linear Models (GLMs) and credibility theory which assume linearity, stationarity, and relatively stable risk structures. In reality, systemic insurance risks are **dynamic, nonlinear**, and often the result of complex interactions between macro-environmental factors (e.g., climate anomalies) and micro-level behavioral patterns (e.g., high-risk driving behaviors).

Systemic risk in insurance refers to the possibility that a significant, correlated shock such as a climate disaster could trigger simultaneous, widespread losses across a large portfolio, threatening solvency and market stability (Cummins & Weiss, 2014). Recent global events, including severe floods in Western Europe (2021) and wildfires in Australia (2020), have demonstrated how climate volatility can simultaneously affect thousands of policyholders, overwhelming traditional underwriting frameworks.

2. Literature Review

2.1 Systemic Risk in Insurance

Systemic risk in the insurance sector refers to the potential for a correlated event such as a natural catastrophe or macroeconomic shock to trigger simultaneous, large-scale losses across multiple insured entities, threatening the solvency of insurers and the stability of the broader financial system (Cummins & Weiss, 2014). Studies have highlighted that climate change is a key driver of systemic insurance risks, increasing both the frequency and severity of catastrophic events (IPCC, 2022). Empirical evidence from major disasters, such as the 2021 European floods and 2020

Australian wildfires, has shown how extreme weather can cause geographically correlated claims, overwhelming traditional risk models.

2.2 Telematics and Behavioral Risk Modeling

Behavioral telematics has emerged as a powerful tool for capturing granular driving data such as speed, acceleration, braking, and time-of-day driving which can be used to refine risk assessment. Baker et al. (2020) demonstrated that night driving and abrupt braking are strongly correlated with accident severity. Wüthrich (2018) emphasized that incorporating behavioral data into pricing models allows for more individualized and fair premiums. However, most telematics-based studies have focused on improving individual risk pricing rather than modeling systemic risks that combine behavioral and environmental drivers.

2.3 Climate Data and Catastrophe Modeling

The integration of climate data into actuarial modeling has grown rapidly with advances in remote sensing and satellite imagery. MODIS and ERA5 datasets provide high-resolution climate indicators, such as temperature anomalies, precipitation patterns, and wind speeds, which have been used for catastrophe risk assessment (Frees et al., 2021). While these datasets have been applied in property insurance and reinsurance, their integration with real-time behavioral data streams remains underexplored, especially for dynamic, real-time risk prediction.

2.4 Artificial Intelligence in Insurance Analytics

Artificial Intelligence (AI) has been widely applied in the insurance domain for:

- **Fraud detection** (Baesens et al., 2015).
- **Customer segmentation** (Wüthrich, 2018).
- **Dynamic pricing optimization** (Frees et al., 2021).

In predictive modeling, **Convolutional Neural Networks (CNNs)** have shown strong performance in extracting spatial features from satellite imagery (Krizhevsky et al., 2012), while **Long Short-Term Memory (LSTM)** networks are effective for modeling sequential dependencies in

time-series data (Hochreiter & Schmidhuber, 1997). Despite these successes, their combined application for systemic risk prediction in insurance is rare in the literature.

2.5 Explainable AI (XAI) in Regulated Insurance Markets

Transparency is a critical requirement in the insurance sector, where regulatory frameworks such as those from the European Insurance and Occupational Pensions Authority (EIOPA, 2022) demand explainable models. SHapley Additive exPlanations (SHAP) provide feature-level interpretability, while Gradient-weighted Class Activation Mapping (Grad-CAM) highlights relevant spatial regions influencing predictions. Although these methods have been tested in other financial applications, their deployment in **hybrid multi-modal insurance risk models** remains limited.

2.6 Research Gap

Based on the reviewed literature, two primary gaps emerge:

1. The lack of **integrated models** that fuse environmental (climate) and behavioral (telematics) data for **real-time systemic risk prediction**.
2. Limited adoption of **XAI frameworks** to ensure regulatory-compliant interpretability in AI-driven actuarial models.

Addressing these gaps is essential for building predictive systems that are both **accurate** and **trustworthy**, enabling insurers to anticipate systemic risks proactively.

Objective

This study proposes a **hybrid deep learning framework** that integrates:

- CNNs for spatial climate feature extraction.
- LSTMs for behavioral pattern analysis.
- XAI tools for interpretability.

By leveraging multi-modal data streams—satellite climate indices, high-frequency telematics data, and historical claims records—this research

addresses both the **predictive performance gap** and the **interpretability challenge** in systemic risk modeling.

The main contributions are:

1. Developing a novel **multi-branch hybrid deep learning model** tailored to systemic insurance risk prediction.
2. Demonstrating the effectiveness of **multi-modal data fusion** in capturing complex spatio-temporal dependencies.
3. Ensuring **regulatory-compliant interpretability** through SHAP and Grad-CAM visualizations.

This approach aims to bridge the gap between **advanced AI methods** and their **practical deployment** in insurance risk management, enabling **dynamic underwriting**, **catastrophe pooling**, and **proactive claims management** in the face of climate-driven uncertainty.

Methodology

1. Research Design

This study adopts an **applied quantitative research design** with a **case study** approach, focusing on a leading European insurer. The goal is to evaluate the predictive capacity of a **hybrid deep learning framework** for systemic insurance risks, specifically those arising from climate-linked catastrophes and behavioral patterns captured through telematics.

2. Data Collection

The dataset is multi-modal, comprising:

- **Climate Data:** ERA5 reanalysis and MODIS satellite imagery (temperature anomalies, precipitation, wind speed).
- **Telematics Data:** Trip-level GPS, speed, acceleration, braking, and night driving ratio.
- **Claims Data:** Historical records linked to telematics-enabled policyholders.

3. Analytical Methods

A dual-branch hybrid deep learning architecture was implemented:

- **CNN Branch:** For climate-related geospatial features.
- **LSTM Branch:** For telematics time-series.

Outputs are fused in a dense layer with dropout, followed by a sigmoid activation for binary classification.

Baseline models include GLM, Random Forest, and XGBoost.

Evaluation metrics: AUC, PR-AUC, Brier Score, F1-score, and calibration error.

4. Data Fusion Process

Insurance claim records were temporally and spatially matched with telematics and climate data:

- **Temporal Alignment:** Hourly synchronization.
- **Spatial Alignment:** Mapping GPS coordinates to climate grid cells using nearest-neighbor interpolation.
- **Feature Engineering:** Creation of composite indices, e.g., storm exposure score.

5. Software and Computational Environment

The study was conducted using **Python 3.10** with the following core libraries:

- **TensorFlow/Keras** – for deep learning model development.
- **Scikit-learn** – for baseline models and evaluation metrics.
- **XGBoost** – for gradient boosting model implementation.
- **Pandas** and **NumPy** – for data handling and computation.
- **SHAP** – for model interpretability and feature attribution.

6. Assumptions and Limitations

Assumptions: Climate indicators are representative of local conditions; telematics data accurately reflects driver behavior.

Limitations: Single-insurer dataset limits generalizability; privacy constraints restrict socio-demographic variables; class imbalance for high-severity claims.

Results:

1. Model Performance Comparison

Table 1 summarizes the comparative performance of the proposed **Hybrid Deep Learning (HDL)** model against three baseline models: **Generalized Linear Model (GLM)**, **Random Forest (RF)**, and **XGBoost (XGB)**.

Table01: Comparative Performance of Predictive Models

Calibration Error	F1-Score (%)	Brier Score	PR-AUC (%)	AUC (%)	Model
0.082	58.7	0.182	60.1	65.4	GLM
0.065	63.5	0.163	67.8	72.6	Random Forest
0.051	68.1	0.142	71.3	76.9	XGBoost
0.021	**84.2**	**0.096**	**89.4**	**93.8**	**Hybrid DL**

Source: Authors' computation based on model evaluation results from the validation dataset (2023)

- The HDL model achieved the highest AUC (93.8%), representing a **22.9 percentage point improvement** over the GLM baseline.

- Precision-Recall AUC increased by 18.1 points compared to XGBoost, indicating superior detection of rare positive cases.
- Brier Score and Calibration Error values confirm the HDL model's **better probabilistic calibration**.

2.Statistical Significance Testing

A **paired t-test** was conducted to assess the statistical significance of the performance differences between HDL and the strongest baseline (XGBoost) across the five validation folds.

- **AUC Difference:** Mean = +16.9%, $p < 0.001$
- **PR-AUC Difference:** Mean = +18.1%, $p < 0.001$
- **F1-Score Difference:** Mean = +16.1%, $p < 0.001$

These results confirm that HDL's improvements are **statistically significant at the 1% level**.

3. Sensitivity to High-Severity Claims

Figure 1 shows the recall performance of each model for **high-severity claims** (> USD 100,000).

Results:

- HDL correctly identified **76.3%** of high-severity claims, compared to **54.7%** for XGBoost and **41.2%** for Random Forest.
- This sensitivity gain is critical for **catastrophe risk management** and **capital allocation**.

4. Explainability Outputs

SHAP value analysis identified the top five most influential predictors in the HDL model:

1. Night driving ratio.
2. Abrupt deceleration events.
3. Storm exposure score (wind × precipitation severity).

4. Average trip speed in urban areas.
5. Temperature anomaly index.

Grad-CAM visualizations of the CNN branch highlighted **geospatial clusters of elevated risk** during extreme weather conditions, particularly in flood-prone coastal regions.

5. Robustness Checks

To test model robustness:

- Training was repeated with **10% random label noise**, and HDL's AUC decreased marginally by 1.9%.
- A reduced model excluding telematics variables showed a **10.4% drop in PR-AUC**, confirming the value of behavioral data.

5. Discussion:

The results of this study demonstrate that the proposed **Hybrid Deep Learning (HDL)** model significantly outperforms traditional actuarial and machine learning baselines in predicting systemic insurance risks. The marked improvement in AUC (+22.9% over GLM) and PR-AUC (+18.1% over XGBoost) confirms the **superior discriminative power** of integrating multi-modal data streams—specifically climate indicators and telematics-based behavioral metrics—within a unified architecture.

From a theoretical standpoint, these findings reinforce the **growing consensus in the literature** that risk in modern insurance markets is inherently **non-linear** and shaped by complex interactions between environmental and behavioral variables (Cummins & Weiss, 2014; Frees et al., 2021). Traditional models such as GLMs, while interpretable and computationally efficient, lack the capacity to capture these multi-scale dependencies. The HDL model's dual-branch architecture, which separately processes spatial climate data through CNNs and temporal driving behavior through LSTMs, addresses this limitation by learning specialized representations before fusing them for joint prediction.

The explainability outputs further validate the model's alignment with domain knowledge. For instance, SHAP analysis identified **night driving ratio** and **abrupt deceleration events** as top behavioral predictors, which is consistent with prior studies linking these factors to higher accident severity (Baker et al., 2020). Similarly, the **storm exposure score** emerged as a strong environmental predictor, reflecting well-established correlations between extreme weather events and elevated claim frequencies (IPCC, 2022). The Grad-CAM visualizations added a spatial dimension to interpretability by pinpointing high-risk climate zones, thus enhancing the model's practical utility for **geospatial risk mapping**.

From an applied perspective, the HDL framework opens new opportunities for **dynamic underwriting**. Insurers could adjust premiums in near real-time based on evolving weather patterns or changes in driver behavior, thereby aligning pricing with actual risk exposure. The model's strong recall for high-severity claims (> USD 100,000) suggests that it could serve as an **early warning system** for catastrophe pooling and capital reserve planning.

Furthermore, the robustness checks indicate that the model maintains high predictive performance even in the presence of data imperfections, which is critical for deployment in real-world insurance operations where **missing or noisy data** is inevitable. The sensitivity analysis also confirmed the **indispensable role of telematics data**, as removing it resulted in a 10.4% drop in PR-AUC, underscoring the need for continued investment in connected vehicle technologies.

However, the results also reveal potential **ethical and operational challenges**. The reliance on telematics raises privacy concerns, especially under stringent regulations such as GDPR. Moreover, there is a risk of **algorithmic bias** if certain driver groups (e.g., night-shift workers, residents of high-risk areas) are systematically penalized. Addressing these concerns will require integrating **fairness-aware learning methods** and implementing **ethical AI governance frameworks** (EIOPA, 2022).

In summary, this study not only advances predictive modeling for systemic insurance risk but also contributes to the broader discourse on how

AI can be responsibly integrated into actuarial science. By combining high predictive accuracy with regulatory-compliant interpretability, the HDL framework offers a **scalable, transparent, and ethically robust** tool for the insurance sector in an era of climate uncertainty and behavioral volatility.

6. Conclusion

This study developed and validated a **Hybrid Deep Learning (HDL)** framework that integrates Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Explainable AI (XAI) techniques to predict systemic insurance risks. By fusing **multi-modal data streams** including satellite-based climate indicators, telematics-derived behavioral metrics, and historical claims the model achieved **superior predictive performance**, with an AUC of 93.8% and PR-AUC of 89.4%, significantly outperforming traditional actuarial and machine learning baselines.

The results underscore three key contributions:

1. **Methodological Innovation:** Demonstrating the effectiveness of multi-branch deep learning architectures in capturing complex spatio-temporal dependencies in systemic risk modeling.

2. **Practical Utility:** Providing insurers with a robust, real-time tool for **dynamic underwriting**, **catastrophe pooling**, and **capital reserve optimization**.

3. **Regulatory Alignment:** Ensuring transparency and interpretability through SHAP and Grad-CAM outputs, in line with regulatory expectations.

From a policy and industry perspective, the proposed framework enables **proactive risk management** by allowing insurers to adjust pricing and coverage based on emerging environmental and behavioral patterns. However, successful adoption requires addressing **data privacy concerns**, mitigating **algorithmic bias**, and ensuring compliance with local and international regulations.

Future research

should focus on:

- Expanding the dataset through **federated learning** to improve cross-jurisdictional generalizability.
- Incorporating **macro-financial indicators** to capture compound risks.
- Exploring **transformer-based architectures** for enhanced long-term dependency modeling.
- Simulating rare catastrophic events through **synthetic data generation** (e.g., GANs).

In conclusion, the HDL framework represents a **scalable, interpretable, and ethically robust** step forward in integrating AI into actuarial science, equipping the insurance sector to better navigate the intertwined challenges of **climate change** and **behavioral volatility**

Appendix A – Descriptive Statistics of the Dataset

TableA1: Summary Statistics of Key Variables

Missing (%)	Max	Min	Std. Dev.	Mean	Variable
0.0	45.6	0.0	7.2	18.0	Night driving ratio (%)
1.2	15	0	2.1	4.3	Abrupt deceleration events
0.8	92.6	18.2	12.5	54.7	Average trip speed (km/h)
0.0	2.1	-0.8	0.4	0.7	Temperature anomaly (°C)

0.0	0.92	0.05	0.15	0.36	Storm exposure score
0.0	325,000	150	11,200	4,750	Claim amount (USD)

Source: Authors' computation based on climate, telematics, and claims data from the case study insurer (2018–2023).

Appendix B– Hyperparameters of Baseline Models

TableA2: Final Hyperparameters for Each Model

Key Parameters	Model
Regularization = 0.01, Link = Logit	GLM
Trees = 300, Max Depth = 15, Features = sqrt(n)	Random Forest
Trees = 400, Max Depth = 10, Learning Rate = 0.05	XGBoost
LSTM Layers = 2, Units = 128, Dropout = 0.3, Optimizer = Adam, Learning Rate = 0.001	Hybrid DL

Source: Authors' model tuning and hyperparameter optimization results (2023)

Appendix C– Explainability Visualizations (Summary Table)

TableA3: Summary of Explainability Outputs

Key Insights	Purpose	Tool
Night driving ratio, abrupt deceleration events, storm exposure score, urban average speed, temperature anomaly are top predictors.	Feature attribution (global & instance level)	SHAP Values
Highlighted high-risk geospatial zones during extreme weather, particularly in flood-prone coastal regions.	Spatial attention mapping (CNN branch)	Grad-CAM

Source: Authors' computation based on model evaluation results from the validation dataset (2023)

Appendix D– Robustness Tests Results

TableA4: Impact of Stress Tests on Model Performance

Interpretation	Change in HDL Performance	Metric Evaluated	Scenario Tested
Model remains stable despite label noise.	-1.9%	AUC	10% Random Label Noise
Confirms strong predictive value of behavioral data.	-10.4%	PR-AUC	Removal of Telematics Features
Interpretation	**Change in HDL Performance**	**Metric Evaluated**	**Scenario Tested**

Source: Authors' robustness checks and scenario testing results (2023)

AppendixE – Summary of Dataset Used in the Study (2018–2023)

Year	Data Type	Variable	Unit / Description	Data Source
2018–2023	Climate Data	Temperature Anomaly	°C deviation from baseline average	ERA5 Reanalysis & MODIS Satellite Imagery
2018–2023	Climate Data	Precipitation Index	mm/day aggregated to monthly averages	ERA5 Reanalysis
2018–2023	Climate Data	Wind Speed	m/s at surface level	ERA5 Reanalysis
2018–2023	Climate Data	Storm Exposure Score	Composite index combining wind, rainfall, and temperature anomaly	Authors' computation based on ERA5 & MODIS
2018–2023	Telematics Data	Night Driving Ratio	% of total trips occurring between 10 PM and 5 AM	Insurer Telematics Logs
2018–2023	Telematics Data	Abrupt Braking Events	Average count per 100 km driven	Insurer Telematics Logs
2018–2023	Telematics Data	Urban Average Speed	km/h for trips within urban zones	Insurer Telematics Logs
2018–2023	Telematics Data	Total Distance Driven	km per policyholder per year	Insurer Telematics Logs
2018–2023	Claims Data	Claim Frequency	Number of claims per policyholder-year	Insurer Claims Records
2018–2023	Claims Data	Claim Severity	Average claim amount (€)	Insurer Claims Records

Source: Authors' compilation based on ERA5, MODIS, insurer telematics logs, and claims records (2018–2023).

7. Referrals and references:

01- Baker, T., McElroy, R., & Coughlan, P. (2020). The role of driving behavior in motor insurance claims: Evidence from telematics data. Journal of Risk and Insurance, 87(4), 1057. https://doi.org/10.1111/jori.12299

02- Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). Fraud analytics using descriptive, predictive, and social network techniques: A guide to data science for fraud detection. John Wiley & Sons.

03- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

04- Cummins, J. D., & Weiss, M. A. (2014). Systemic risk and the U.S. insurance sector. Journal of Risk and Insurance, 81(3), 489. https://doi.org/10.1111/jori.12039

05- European Insurance and Occupational Pensions Authority. (2022). AI governance in insurance: Regulatory perspectives. https://www.eiopa.europa.eu

06- Frees, E. W., Meyers, G., & Cummings, A. (2021). Predictive modeling of insurance claims with telematics data. North American Actuarial Journal, 25(4), 623. https://doi.org/10.1080/10920277.2020.1850843

07- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

08- Intergovernmental Panel on Climate Change. (2022). Climate change 2022: Impacts, adaptation, and vulnerability. Cambridge University Press. https://doi.org/10.1017/9781009325844

09- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097.

10- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765.

11- Ng, A. (2021). Machine learning yearning: Technical strategy for AI engineers, in the era of deep learning. deeplearning.ai.

12- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

13- Wüthrich, M. V. (2018). Machine learning in individual claims reserving. Scandinavian Actuarial Journal, 2018(6), 465. https://doi.org/10.1080/03461238.2018.1460486