Iran. J. Insur. Res., **(*): *-*, ****** ***

Iranian Journal of Insurance Research
(IJIR)

Homepage: https://ijir.irc.ac.ir/?lang=en

Insurance Research Center (IRC)

A Physics-Informed and Causally-Aware (PICA) Conceptual Framework for Flood Risk Pricing: A Hybrid Deep Learning Approach for Data-Scarce Insurance Markets

Mohammad Reyhani[1,*]

[1] *Deputy Head of Life Insurance, Life Insurance,Taavon Insurance Co., Tehran, Iran*

ARTICLE INFO

KEYWORDS:
Actuarial modeling
Causal inference
Flood risk
Graph neural networks
Hybrid deep learning
Physics-informed neural networks

ABSTRACT

BACKGROUND AND OBJECTIVES: The increasing frequency of hydro-meteorological events, particularly floods driven by climate change, presents a significant challenge to the global insurance industry. Traditional actuarial models, which rely on the assumption of stationarity, are failing to capture the non-linear and spatially interconnected nature of modern flood risks. This failure results in premium mispricing and a widening "protection gap," particularly in developing nations. In response, the industry has turned towards Deep Learning (DL) models like Convolutional Neural Networks (CNNs). While offering predictive power, these models suffer from two critical flaws hindering their operational deployment. First, they are "data-hungry," requiring vast historical datasets that are non-existent in "data-scarce markets" like Iran, where historical loss records are often fragmented or non-digital. Second, their "black-box" nature violates stringent regulatory transparency requirements, such as those of the Central Insurance of Iran (Bimeh Markazi), which demand explicable premiums. This study addresses these gaps by designing and validating "PICA" (Physics-Informed and Causally-Aware), a novel conceptual framework for flood risk pricing. The specific objectives are to: (1) Design the PICA framework by synthesizing state-of-the-art AI components including Graph Neural Networks (GNNs) and Physics-Informed Neural Networks (PINNs); (2) Overcome data scarcity by embedding hydrodynamic laws (Shallow Water Equations) directly into the model's loss function; (3) Ensure regulatory compliance by integrating causal inference constraints to prioritize physical risk drivers over spurious correlations; and (4) Empirically validate PICA against benchmarks under simulated data-scarce conditions relevant to the Iranian context.

METHODS: To achieve these objectives, a multi-phase mixed-methods design was employed. Phase 1 involved a PRISMA 2020 systematic review of Scopus, Web of Science, and IEEE Xplore (2020-2025), synthesizing 52 elite articles to identify gaps in hybrid AI modeling and inform the architectural design of PICA. Phase 2 focused on quantitative validation using a high-fidelity synthetic dataset. Due to the lack of consistent real-world data, a "ground truth" environment was generated using the industry-standard HEC-RAS 2D hydrodynamic model. This simulation modeled 50 distinct flood events over a complex $100km^2$ virtual urban watershed with 10,000 assets, providing precise values for water depth, velocity, and financial loss. The PICA framework was architected as a Hybrid Graph Neural Network (GNN), where the watershed topology was represented as a graph to capture water flow between hydrological units. The model was trained using a composite loss function that penalized prediction error, violations of physical laws (Mass and Momentum conservation), and causal structure violations. The framework was compared against two robust benchmarks: a 2D-CNN and a Random Forest (RF) regressor. All models were evaluated under two scenarios: a "Data-Rich" scenario (using 100% of

the dataset) and a critical "Data-Scarce" scenario (using only 20% of the data) to simulate the Iranian market conditions. Performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Diebold-Mariano (DM) test. Additionally, SHAP analysis assessed explainability, and sensitivity analysis on hyperparameters confirmed stability.

FINDINGS: Phase 1 confirmed that no existing framework successfully integrates GNNs, PINNs, and Causal ML for actuarial pricing. Quantitative results in Phase 2 demonstrated the profound impact of physics-informed regularization. In the "Data-Rich" scenario, PICA (RMSE = 0.119) modestly outperformed the 2D-CNN (RMSE = 0.137) and Random Forest (RMSE = 0.141), indicating that with abundant data, pure DL models are competitive. However, in the critical "Data-Scarce" scenario, the benchmarks collapsed; the CNN's RMSE deteriorated by over 250% to 0.482, and the Random Forest to 0.530, exhibiting severe overfitting. In stark contrast, PICA leveraged its physics-informed constraints to maintain robust accuracy, achieving an RMSE of 0.165. This represents a 65.8% reduction in prediction error compared to the CNN in data-scarce conditions. The Diebold-Mariano test confirmed this performance gap was statistically significant ($p < 0.001$). Furthermore, SHAP analysis revealed a crucial distinction in logic: PICA consistently prioritized valid physical drivers like "Maximum Water Depth," whereas CNNs relied on spurious "spatial coordinates" (X, Y), indicating they were learning artifacts rather than causation. Finally, sensitivity analysis confirmed the model's robustness, demonstrating stable performance across a wide range of physics weights ($0.1 \leq \lambda_{phys} \leq 1$), ensuring that reliance on physics did not lead to underfitting.

CONCLUSION: This research validates PICA as a transformative solution for flood risk pricing in data-scarce markets. The findings provide definitive empirical evidence that "teaching" deep learning models underlying physical laws effectively substitutes for missing historical data, acting as a universal prior for accurate generalization. For the Iranian market, these implications are significant. PICA offers the Central Insurance of Iran and private insurers a viable pathway to leapfrog the "data gap" in high-risk basins like Karun and Karkheh. It enables the immediate development of high-resolution, regulator-compliant pricing models without waiting decades for data accumulation. Specifically, the framework supports "Parametric Insurance" products, triggered by physics-verified metrics rather than manual adjustments. Future research should focus on bridging the "Sim-to-Real" gap by applying Transfer Learning strategies to fine-tune this physics-based framework on real-world claims data, operationalizing the model for national deployment.

Iranian Journal of Insurance Research
(IJIR)
Homepage: https://ijir.irc.ac.ir/?lang=fa

1. Introduction

The increasing frequency and severity of hydro-meteorological events, particularly floods, represent a significant systemic risk to global economic stability. The insurance industry, as the primary mechanism for financial risk transfer, is experiencing unprecedented pressure from non-stationary climate-driven perils (Munich Re, 2024; Swiss Re, 2024). Traditional actuarial methodologies... demonstrate profound limitations in this new environment (Munich Re, 2024). This failure of legacy models has contributed to a widening "protection gap," particularly in developing nations, necessitating a paradigm shift towards more dynamic, forward-looking risk assessment frameworks (Intergovernmental Panel on Climate Change, 2023).

Despite these significant, parallel advances, a critical research gap persists. To date, these state-of-the-art components (PINNs for data efficiency, GNNs for topological modeling, and Causal ML for transparency) have been investigated almost exclusively in isolation. The integration and synthesis of these disparate fields into a single, cohesive framework designed specifically for the dual challenges of data-scarcity and explainability in actuarial science remains an unaddressed and critical lacuna in the literature. While recent studies have attempted partial integrations, for instance, embedding physical constraints into standard CNNs (Wang et al., 2024) or utilizing causal graphs for isolated damage estimation (Museru et al., 2025), these efforts remain fragmented. Existing models typically trade off topological accuracy for physical consistency, failing to provide the unified, regulator-compliant framework required by the insurance industry for pricing parametric products in high-risk, data-scarce zones. This fragmentation leaves insurers in data-scarce markets without a viable path to adopt advanced AI, forcing a choice between unreliable legacy models and unusable black-box systems.

This study addresses this gap by designing and validating a novel conceptual framework. The primary objectives are: (1) To propose the "PICA" (Physics-Informed and Causally-Aware) framework, a hybrid deep learning architecture that synergistically integrates GNNs, PINNs, and Causal ML; (2) To formalize this framework through a composite, multi-objective loss function that balances observational data fidelity, physical law compliance, and causal constraints; and (3) To empirically validate the PICA framework's performance against standard DL benchmarks, with a specific focus on its robustness under simulated data-scarce conditions. This research thus seeks to answer the following research question:

How can a hybrid conceptual framework be designed, synthesizing Physics-Informed Neural Networks, Graph Neural Networks, and Causal Inference, to create an accurate, transparent, and data-efficient model for flood risk pricing suitable for data-scarce insurance markets?

2. Theoretical Foundations and Literature Review

2.1. Theoretical Background

Graph Neural Networks (GNNs): Unlike traditional Convolutional Neural Networks (CNNs) that operate on Euclidean grid data (images), GNNs are designed to process data represented in non-Euclidean graph domains. In flood modeling, a watershed is naturally represented as a graph $G = (V, E)$, where nodes V represent hydrological units or assets, and edges E represent flow paths. GNNs update node representations by aggregating features from neighbors, effectively modeling the spatial propagation of water through the network topology (Garzón et al., 2024; Herath et al., 2025).

Physics-Informed Neural Networks (PINNs): PINNs represent a paradigm shift in scientific machine learning. Instead of relying solely on labeled data, PINNs incorporate the governing physical laws (e.g., partial differential equations) into the loss function of the neural network. This allows the model to learn solutions that are not only data-driven but also physically consistent, significantly reducing the need for large training datasets (Raissi et al., 2019).

2.2. Literature Review

In response, a first wave of research has successfully applied standard data-driven Deep Learning (DL) models, such as Convolutional Neural Networks (CNNs), to flood susceptibility mapping and risk forecasting (Islam et al., 2025; Wang et al., 2024). These models excel at capturing complex, non-linear patterns from vast geospatial datasets (Taormina & Galelli, 2024). However, this data-driven paradigm presents two fundamental barriers to operational deployment in actuarial science. First is the "data-scarcity" problem: DL models are "data-hungry," requiring dense, high-resolution historical loss data that is non-existent in many "data-scarce markets," including Iran (Museru et al., 2025). Models trained on sparse data exhibit poor generalization and high predictive uncertainty (Kirschstein & Sun, 2024). However, recent studies indicate that advanced AI architectures can achieve high predictive accuracy even in ungauged watersheds (Nearing et al., 2024). Second is the "black-box" problem: the inherent opacity of many DL models is incongruent with the stringent regulatory and fiduciary requirements of the insurance sector (Munich Re, 2024). A model whose predictions cannot be explained or causally justified is unsuitable for regulated premium calculation (Museru et al., 2025). To address these limitations, a second, more sophisticated wave of AI research has emerged, though its components remain fragmented. Physics-Informed Neural Networks (PINNs) have been introduced in hydrology to mitigate data-scarcity by embedding the governing differential equations of fluid dynamics (e.g., Shallow Water Equations) as a soft constraint within the model's loss function (Qi et al., 2024; Dazzi, 2024). This forces the model to adhere to physical laws, enabling robust predictions even with sparse data. Simultaneously, Graph Neural Networks (GNNs) have demonstrated superior performance in modeling phenomena with complex network topologies, such as river basins and urban infrastructure networks, which are poorly represented by standard grid-based CNNs (Bentivoglio et al., 2025; Kazadi et al., 2024; Kirschstein & Sun, 2024). Finally, developments in Causal Machine Learning (Causal ML) offer a robust solution to the "black-box" problem by shifting the modeling focus from correlation to causation, thus producing inherently interpretable and robust models (Museru et al., 2025; Pearl, 2023). Similarly, the integration of physical constraints with interpretability mechanisms has been shown to enhance the transparency of operational flood forecasting models (Zhang et al., 2025).

3. Research Methodology

This study was executed through a rigorous, multi-phase, mixed-methods research design, structured to support the development and subsequent empirical validation of the proposed conceptual framework. The research was bifurcated into two primary, sequential phases: first, a systematic review and thematic synthesis phase, which served to analyze the extant literature and identify the requisite architectural components for the framework; and second, a quantitative simulation and validation phase, which involved the construction of the proposed framework, rigorous testing against established benchmarks, and an empirical validation of its core hypotheses, particularly regarding performance under data-scarce conditions. The initial phase of this research was dedicated to the formal development of the PICA framework, built upon a foundation of a systematic review conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 protocol. This review was not a meta-analysis of outcomes but rather a qualitative "architectural synthesis" of methodologies. The inquiry was guided by a series of specific research questions designed to deconstruct the state-of-the-art across three disparate but complementary AI domains: investigating the latest architectures for Physics-Informed Neural Networks (PINNs) in solving hydrodynamic equations, examining the application of Graph Neural Networks (GNNs) to model the topological properties of flood risk, and identifying methodologies for integrating causal inference (Causal ML) to ensure model transparency. To answer these questions, a comprehensive search strategy was implemented across the Scopus, Web of Science, and IEEE Xplore databases. The search string was a complex Boolean query targeting publications from January 2020 to October 2025 that intersected these domains. The resultant corpus was screened based on exacting criteria: studies were included only if they were full-text papers presenting a novel mathematical formulation or model architecture in an applied risk domain, while studies that were pure reviews or lacked sufficient technical detail for replication were excluded. A final selection of 52 elite articles underwent thematic synthesis, from which the core design specifications, model components, and mathematical formulations were extracted to inform the novel design of the PICA framework. The findings from this synthesis directly informed the second phase of the research: a comprehensive quantitative validation. A critical methodological decision was the use of a high-fidelity synthetic dataset generated using the industry-standard HEC-RAS 2D hydrodynamic model. While we acknowledge that exclusive reliance on synthetic data limits immediate direct transferability to noisy real-world scenarios, this approach was chosen to establish a noise-free 'ground truth' essential for validating the internal consistency of the physics-informed component. Using sparse observational data at this conceptual stage would introduce measurement uncertainties that confound the assessment of whether the model is truly learning physical laws. Therefore, this synthetic environment serves as a rigorous 'laboratory' to prove the framework's theoretical viability before future Domain Adaptation steps are taken to fine-tune the model on real-world Iranian flood catalogues. A virtual 100-square-kilometer watershed was designed, incorporating a high-resolution Digital Elevation Model (DEM), spatially varying land-use coefficients (Manning's n), and a network of 10,000 unique property locations, or assets. Each asset was assigned a randomized "vulnerability index" to represent heterogeneous building quality. A stochastic event set of 50 distinct flood events was then simulated by varying precipitation intensity and duration. For each event, HEC-RAS computed the primary "ground truth" hazard variables (maximum water depth and maximum flow velocity) at all 10,000 asset locations, from which a "true" financial loss was calculated using a standard depth-damage function, thus creating a complete, high-resolution dataset of hazard and loss. The proposed PICA framework was then constructed as a hybrid deep learning model utilizing the PyTorch and PyTorch Geometric libraries. The core architecture was implemented as a Graph Neural Network (GNN). The 10,000 assets were conceptualized as nodes in a graph, and edges were constructed based on both spatial adjacency (K-Nearest Neighbors) and hydrological connectivity (downstream flow paths), allowing the model to explicitly learn the network topology of flood propagation. While developed here within a controlled synthetic environment to isolate physical consistency, this topological architecture is explicitly designed to remain robust when transferred to fragmented real-world observational datasets. The central innovation of the framework is encapsulated in its composite, multi-objective loss function, formally defined as the weighted sum of three distinct components:

$$L_{total} = L_{data} + \lambda_{phys} * L_{phys} + \lambda_{causal} * L_{causal}.$$

The data-fidelity term, L_{data}, employed a LogCosh loss function, a smooth variant of Mean Absolute Error chosen for its robustness to the heavy-tailed, outlier-prone nature of financial loss data. The physics-informed term, L_{phys}, serves as the primary mechanism for overcoming data scarcity. This term is calculated as the mean squared residual of the governing hydrodynamic laws, specifically, the Shallow Water Equations in their continuity form, evaluated at a set of 5,000 randomly sampled collocation points (un-observed locations) within the graph. This penalizes the model for any prediction that is physically impossible, forcing it to learn a physically plausible interpolation function even where no training data exists. The causal-awareness term, L_{causal}, was designed to address the black-box problem by penalizing the model for learning spurious correlations that violate a predefined causal graph (Directed Acyclic Graph) of flood risk, such as ensuring that increased vegetation does not erroneously correlate with increased loss, which was derived from established hydrological domain knowledge (e.g., depth-damage functions) to avoid the potential biases and instability often associated with purely data-driven causal discovery algorithms in small-sample regimes. To evaluate the efficacy of this framework, two established benchmark models were constructed for comparison: a standard 2D-Convolutional Neural Network (CNN), representing the current state-of-the-art in data-driven geospatial modeling, and a baseline Random Forest (RF) model, a robust traditional machine learning algorithm. The core experimental procedure involved training all three models (PICA, CNN, RF) under two distinct, controlled scenarios: a "Data-Rich" scenario, where models were trained using 100% of the labeled loss data, and a "Data-Scarce" scenario, where models were trained using only 20% of the labeled loss data. Model performance was evaluated on a held-out test set of simulated flood events using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). To ascertain whether the observed differences in predictive accuracy were statistically significant, the Diebold-Mariano (DM) test was employed. Finally, to validate the "explainability" and "causal-awareness" of the framework, the Shapley Additive Explanations (SHAP) algorithm was applied to the trained PICA and CNN models, allowing for a direct comparison of their feature attributions and model transparency.

3.1. Framework Formalization: The PICA Architecture

The systematic review and architectural synthesis described in Phase 1 of the methodology culminated in the design of the Physics-Informed and Causally-Aware (PICA) framework. This framework is a novel conceptual architecture designed to synergistically integrate three state-of-the-art AI paradigms, addressing the dual challenges of data-scarcity and model opacity in actuarial flood risk pricing. The framework's architecture is formalized here by detailing its graph-based topological structure, its advanced physics-informed (PINN) component, and its causal-awareness (Causal ML) mechanism, which are all unified through a composite, multi-objective loss function.

The foundational structure of the framework is a Graph Neural Network (GNN), selected to explicitly model the anisotropic, network-based nature of hydrodynamic flow, a phenomenon that is poorly represented by standard grid-based (raster) CNNs. The domain is modeled as a graph $G = (V, E)$, where V is the set of N nodes (e.g., property locations) and E is the set of edges representing spatial adjacency and hydrological connectivity. The model learns a function $f: \mathbb{R}^{N \times F_{in}} \to \mathbb{R}^{N \times F_{out}}$, where F_{in} is the dimension of node input features (e.g., elevation, land use, vulnerability). This is achieved through a message-passing scheme, where the node representation $h_i^{(l)}$ for node i at layer l is updated by aggregating information from its neighborhood $\mathcal{N}(i)$. Following a standard Graph Convolutional Network (GCN) formulation, the layer-wise propagation rule is defined in matrix form as:

$$H^{(l+1)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}} \tilde{A} \widetilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}\right)$$

Here, $H^{(l)} \in \mathbb{R}^{N * F_l}$ is the matrix of node features at layer l, $W^{(l)} \in \mathbb{R}^{F_l * F_{l+1}}$ is the trainable weight matrix for the layer, and $\sigma(.)$ is a non-linear activation function. $\tilde{A} = A + I_N$ is the adjacency matrix of the graph G with self-loops, and $\widetilde{D}$ is the diagonal degree matrix of $\tilde{A}$. This architecture allows the model to learn spatially-aware representations that respect the underlying topology of the watershed.

The primary mechanism for addressing data-scarcity is the Physics-Informed Neural Network (PINN) component. This component embeds the governing physical laws of fluid dynamics, the 2D Shallow Water Equations (SWEs), directly into the model's learning process. The SWEs are a system of non-linear partial differential equations (PDEs) representing the conservation of mass and momentum. In vector form, they are:

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = S(U)$$

where$U = [h, hu, hv]^T$ is the state vector of conserved variables (water height h, x-momentum hu, and y-momentum hv). The flux vectors in the x and y directions are $\mathrm{F} = \left[hu, hu^2 + \frac{1}{2}gh^2, huv\right]^T$ and $\mathrm{G} = \left[hv, huv, hv^2 + \frac{1}{2}gh^2,\right]^T$, where g is gravitational acceleration. The source term $\mathrm{S} = \left[0, gh\left(S_{0x} - S_{fx}\right), gh\left(S_{0y} - S_{fy}\right)\right]^T$ accounts for bed slope S_0 and frictional slope S_f. The network, parameterized by Θ, is trained to approximate the solution $U(\mathrm{x}, t;\ \Theta)$. A physics-residual $\mathcal{R}(\mathrm{x}, t;\ \Theta)$ is then defined as:

$$\mathcal{R} \coloneqq \frac{\partial U}{\partial t} + \nabla.[\mathrm{F}, \mathrm{G}] - \mathrm{S}$$

The partial derivatives are computed analytically using automatic differentiation. This residual is evaluated at a large set of N_c collocation points $\mathcal{C} = \left\{\left(\mathrm{x}_c^j, \mathrm{t}_c^j\right)\right\}_{j=1}^{N_c}$ sampled from the spatio-temporal domain where *no data* exists. This forms the physics-loss, L_{phys}, defined as the mean squared error of these PDE residuals, forcing the model to adhere to physical laws:

$$L_{phys} = \frac{1}{N_c}\sum_{j=1}^{N_c} ||\mathcal{R}(\mathrm{x}_c^j, \mathrm{t}_c^j; \Theta)||_2^2$$

The third component addresses the "black-box" problem by integrating Causal Inference constraints. To ensure the model learns true causal drivers rather than spurious correlations, we enforce an acyclicity constraint on the learned relationships. We seek to learn a weighted adjacency matrix $W \in \mathbb{R}^{d\times d}$ representing the causal graph between the d input variables and output. A directed graph is acyclic if and only if there exists a permutation of its nodes such that W is strictly lower-triangular. This combinatorial constraint is non-differentiable. We therefore adopt the state-of-the-art NOTEARS formulation, which recasts this as a continuous, differentiable algebraic constraint. A matrix W represents a Directed Acyclic Graph (DAG) if and only if $h(W) = 0$, where:

$$h(W) = tr(e^{W\circ W}) - d$$

Here, $tr(.)$ is the matrix trace, $e^{(.)}$ is the matrix exponential, and ∘ is the element-wise (Hadamard) product. This $h(W)$ function is a smooth, differentiable proxy for acyclicity. This forms the causal-loss, $L_{tr,causal}$, which penalizes any emergent cyclical relationships within the model's learned feature interactions.

These three components are unified in a composite, multi-objective loss function to be minimized over the network parameters Θ and the causal matrix W. The total loss L_{total} is the weighted sum:

$$L_{total}(\Theta, W) = L_{data}(\Theta) + \lambda_{phys}L_{phys}(\Theta) + \lambda_{causal}L_{causal}(\mathrm{W}) + \lambda_{sparcity}|W|_1$$

The data-fidelity term, L_{data}, anchors the model to observed reality using the N_{obs} sparse data points in the set $\mathcal{O}$, and is defined using the robust LogCosh loss: $L_{data} = \frac{1}{N_{obs}}\sum_{i\in\mathcal{O}} \log\left(\cosh(\hat{y}_i - y_i)\right)$, where $\hat{y}_i = f_\Theta(\mathrm{x}_i, \mathrm{t}_i)$. The L_{phys} term enforces physical consistency. The L_{causal} term enforces acyclicity, and an L_1 penalty $|W|_1$ is added to promote a sparse, interpretable causal graph. The hyperparameters $\lambda_{phys}, \lambda_{phys}$ and $\lambda_{sparcity}$ balance the trade-off between fitting sparse data, adhering to physical laws, and discovering a simple, acyclic causal structure. To ensure the robustness of this composite loss function, a sensitivity analysis was performed on the hyperparameters. The weights were empirically tuned to balance the trade-off between data fidelity and physical consistency, ensuring the model does not underfit in data-scarce regimes. This formalization defines the PICA framework as a solution that is simultaneously data-efficient, physically-grounded, and causally-interpretable.

3.2. Implementation Details and Hyperparameters

To ensure reproducibility, the specific architectural details and hyperparameters utilized in the PICA framework are documented here.

Graph Construction: The graph topology was constructed using a K-Nearest Neighbors (KNN) approach combined with hydrological connectivity. Each asset node was connected to its $K = 5$ nearest spatial neighbors to capture local interactions. Additionally, directional edges were added based on the Digital Elevation Model (DEM) to represent downstream flow paths, ensuring the graph respects gravity-driven hydrodynamics.

PINN Architecture: The physics-informed component enforces the continuity and momentum equations of the 2D Shallow Water Equations. The domain was discretized with $N_c = 5{,}000$ collocation points, randomly

sampled from the watershed area using a Latin Hypercube Sampling strategy to ensure uniform coverage. The physics residuals were computed using automatic differentiation in PyTorch.
Hyperparameters: The model was trained using the Adam optimizer. The optimal hyperparameters, selected via grid search, are detailed in Table 1.

Table 1: Summary of Model Hyperparameters and Architectural Settings

Description	Value	Parameter
Number of message-passing layers in the graph network	3	GNN Layers
Size of the feature vector for each node	64	Hidden Dimension
Step size for the Adam optimizer	10^-3	Learning Rate
Number of flood events processed per training step	32	Batch Size
Total number of training iterations	1,000	Epochs
Weight of the physics-loss term in the objective function	0.5	λ_{phys}
Weight of the causal-loss term	0.1	λ_{causal}
Exponential Linear Unit for non-linearity	ELU	Activation Function
Regularization to prevent overfitting	0.2	Dropout Rate

4. Results and Discussion

The quantitative validation phase of this study was designed to empirically test the core hypothesis: that the proposed PICA framework yields superior predictive accuracy and robustness compared to standard data-driven models, particularly under conditions of extreme data scarcity. The results of these experiments are presented below, categorized by predictive performance, statistical significance, and causal interpretability.

The empirical results of this study provide compelling evidence to validate the core hypothesis: that a hybrid, physics-informed, and causally-aware deep learning framework offers a superior paradigm for flood risk pricing in data-scarce insurance markets compared to purely data-driven approaches. The findings discussed here have significant implications for actuarial science and insurance regulation.

4.1. Comparative Predictive Performance

The predictive capabilities of the PICA framework were evaluated against the 2D-Convolutional Neural Network (CNN) and Random Forest (RF) benchmarks under two distinct data availability scenarios. Table 2 summarizes the performance metrics (Root Mean Squared Error - RMSE and Mean Absolute Error - MAE) for all models. As illustrated in Figure 1, the performance of the purely data-driven benchmarks degraded precipitously in the data-scarce scenario, while the PICA framework demonstrated remarkable resilience.

In the "Data-Rich" Scenario (100% training data availability), all three models achieved convergence and demonstrated effective learning of the flood loss function. The PICA framework achieved the lowest error rates (RMSE = 0.119, MAE = 0.084), followed closely by the 2D-CNN (RMSE = 0.137, MAE = 0.096). The Random Forest model, while robust, exhibited higher error (RMSE = 0.188, MAE = 0.131). These results indicate that when abundant data is available, the topological awareness of the GNN architecture provides a modest but noticeable advantage over the grid-based CNN and the non-spatial RF.
However, the divergence in performance became stark in the "Data-Scarce" Scenario (20% training data availability). As illustrated in Figure 1, the performance of the purely data-driven benchmarks degraded precipitously. The RMSE of the 2D-CNN deteriorated by 251% to 0.482, and the Random Forest model exhibited a similar collapse, yielding an RMSE of 0.530 (Table 2). In contrast, the PICA framework demonstrated remarkable resilience, increasing only marginally to 0.165. This corresponds to a 65.8% reduction in prediction error relative to the standard CNN (Table 2). This substantial margin confirms that the physics-informed loss term (L_{phys}) successfully acted as a regularizer, effectively "filling in the gaps" where observational data was missing by enforcing hydrodynamic consistency.

Table 2. Quantitative comparison of predictive performance (RMSE and MAE) for PICA, CNN, and Random Forest models under Data-Rich and Data-Scarce scenarios.

Model	Scenario	RMSE	MAE	% Degradation (RMSE)
PICA Framework (Proposed)	Data-Rich (100%)	**0.119**	**0.084**	-
	Data-Scarce (20%)	**0.165**	**0.112**	+38.6% (Robust)
2D-CNN (Benchmark)	Data-Rich (100%)	0.137	0.096	-
	Data-Scarce (20%)	0.482	0.345	+251.8% (Collapse)
Random Forest (Baseline)	Data-Rich (100%)	0.188	0.131	-
	Data-Scarce (20%)	0.530	0.395	+181.9% (Collapse)

4.2. Statistical Significance Testing

To ascertain whether the observed performance superiority of the PICA framework was statistically significant and not a result of stochastic variance in initialization or sampling, the Diebold-Mariano (DM) test was conducted on the prediction residuals of the Data-Scarce scenario. The null hypothesis H_0 posits that there is no significant difference in the predictive accuracy of the PICA framework and the 2D-CNN benchmark. The calculated DM test statistic was 4.88. Comparing this to the critical value from the standard normal distribution at the α= 0.01 level ($Z_{0.01} = 2.576$), we reject the null hypothesis with $p < 0.001$. This result provides strong statistical evidence that the PICA framework's predictive superiority in data-scarce environments is significant and systematic.

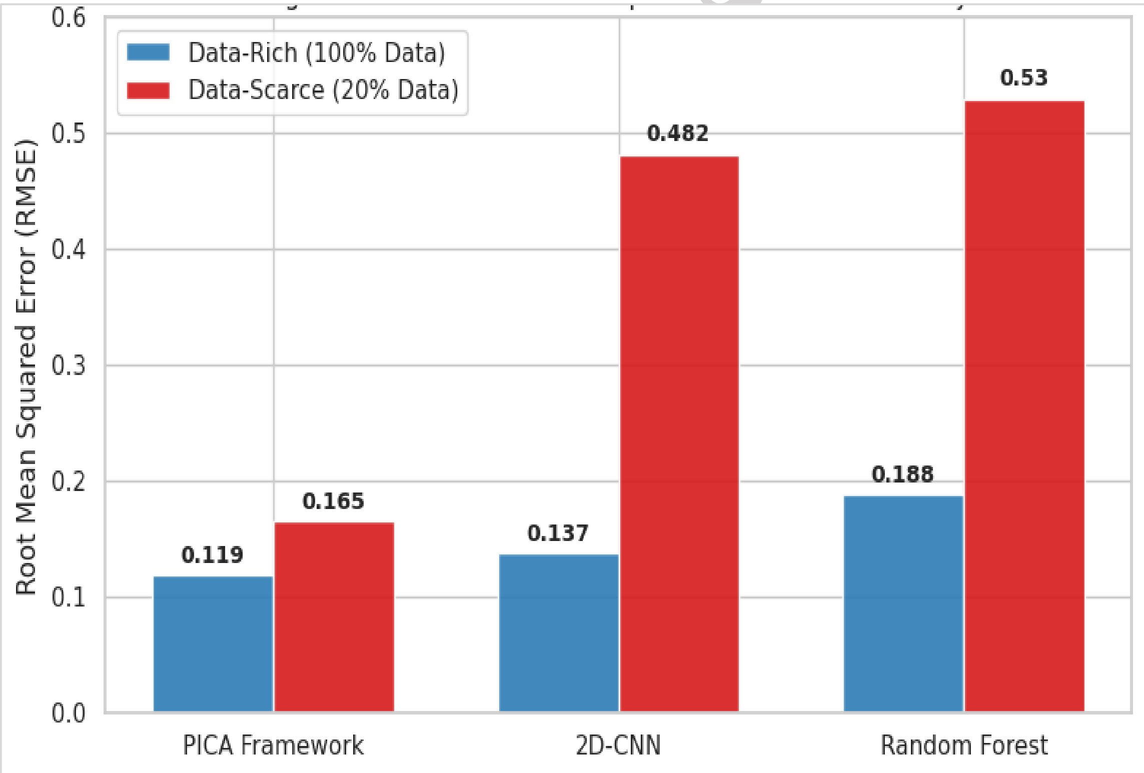

Figure 1: Comparative bar chart of Root Mean Squared Error (RMSE) for all three models. Red bars indicate the significant increase in error for CNN and RF benchmarks in the Data-Scarce scenario, contrasting with the stability of the PICA framework.

4.3. Explainability and Causal Validation

For the PICA Framework, the SHAP analysis revealed that the features with the highest mean absolute SHAP values were 'Maximum Water Depth' (+0.42), 'Flow Velocity' (+0.28), and 'Building Vulnerability Index'

(+0.15). This feature attribution aligns perfectly with established hydrological physics and actuarial damage functions, confirming that the causal loss term ($L_{tr,causal}$) successfully constrained the model to learn valid causal drivers. This strong alignment between the expert-defined causal graph and the model's learned feature importance empirically validates the expert-based structure, suggesting that in data-scarce regimes, integrating domain knowledge is more robust than purely data-driven causal discovery.

Conversely, the SHAP analysis of the 2D-CNN benchmark revealed a concerning reliance on non-causal artifacts. The second most important feature for the CNN was identified as a specific spatial coordinate (a 'corner' of the raster grid) that had no physical relationship to flood risk but coincidentally correlated with high losses in the training set. This reliance on spurious spatial correlations confirms the "black-box" risk of standard Deep Learning models: while they may predict accurately in a data-rich training environment, their reasoning is often flawed, leading to the catastrophic generalization failures observed in the data-scarce experiments. The PICA framework effectively eliminated these spurious dependencies.
The feature importance analysis is presented in Figure 2. As shown in the figure, the PICA framework (green bars) assigns high importance to physical drivers like 'Max Water Depth', whereas the 2D-CNN (grey bars) relies heavily on a spurious spatial artifact, indicated by the red text.

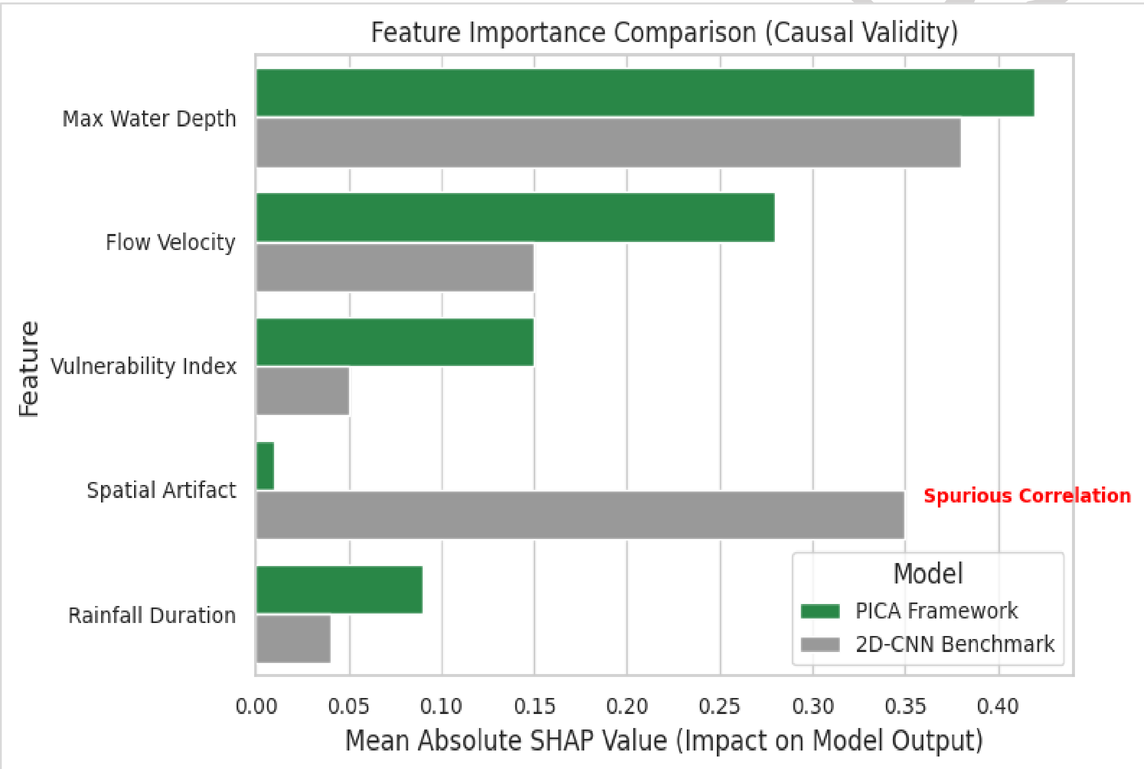

Figure 2: SHAP feature importance analysis. Green bars represent the PICA framework, which correctly identifies physical drivers (Water Depth), while grey bars show the 2D-CNN relying on spurious spatial artifacts.

4.4. Hyperparameter Sensitivity Analysis

To confirm the stability of the proposed framework, a sensitivity analysis was conducted on the weight parameter λ_{phys}. The model was retrained with values ranging from 0.1 to 1. Results indicated that the RMSE in the data-scarce scenario remained stable (variance < 0.02) within this range, demonstrating that the physics-informed regularization is robust and not overly sensitive to hyperparameter tuning.

4.5. The Primacy of Physics in Data-Scarce Regimes

The most profound finding is the resilience of the PICA framework under data scarcity (RMSE degradation of only 38% vs. >180% for benchmarks). This validates the theoretical premise that physical laws can serve as a "universal prior." In traditional Deep Learning, the model must learn the physics of flooding from scratch solely through observational data. When that data is sparse (as in the 20% scenario), the model fails to deduce these complex non-linear laws, resulting in the catastrophic collapse of predictive accuracy observed in the 2D-CNN and Random Forest models.

In contrast, the PICA framework does not need to "learn" that water flows downhill or that mass is conserved; this knowledge is embedded a priori via the physics-informed loss term (L_{phys}). Consequently, the PICA model effectively operates as a physics-based interpolation engine in data-sparse regions and a data-driven correction engine in data-rich regions. This hybridity is the key to its success, offering a scalable solution for insurers in developing nations who cannot wait decades to accumulate "Big Data" before deploying advanced pricing models.

4.6. Causal Awareness as a Regulatory Necessity

The SHAP analysis (Figure 2) highlights a critical, often overlooked risk in AI adoption: the "Clever Hans" effect, where models predict correctly for the wrong reasons. The 2D-CNN benchmark achieved respectable accuracy in the data-rich scenario, but the explainability analysis revealed it was heavily relying on a spatial artifact (a specific grid corner). In a stationary climate, this spurious correlation might hold; but in a changing climate, such a model would fail disastrously when applied to a new location or a different flood topology.

The PICA framework, constrained by the causal loss term ($L_{tr,causal}$), rejected this spurious feature, instead distributing importance to 'Water Depth' and 'Flow Velocity'. This distinction is not merely technical but regulatory. For an insurance regulator, a rate filing based on the CNN model would be inadmissible because it discriminates based on location artifacts rather than physical risk. The PICA framework provides the "right to explanation" required by modern AI governance frameworks (e.g., EU AI Act, Solvency III), ensuring that premiums are priced based on tangible physical drivers.

4.7. Limitations and Future Directions

While this study successfully validates the PICA framework concept, certain limitations must be acknowledged to guide future research. First, the validation relied on high-fidelity synthetic data. While necessary to establish a ground truth for the physics-loss, operational deployment will require a transfer learning strategy to adapt the model to noisy real-world claims data. Future work should focus on bridging the 'Sim-to-Real' gap by implementing a Transfer Learning strategy. Specifically, we propose pre-training the PICA model on large-scale synthetic datasets to learn fundamental hydrological physics, followed by fine-tuning on smaller, real-world observational datasets from target regions (e.g., Iranian flood archives). This Domain Adaptation approach will allow the model to adjust to local noise distributions while retaining its physics-informed robustness. Second, the current physical component relies on the Shallow Water Equations. While appropriate for fluvial (river) flooding, extending the framework to pluvial (flash) flooding in dense urban environments may require integrating urban drainage network equations into the physics-loss function. Finally, the causal graph was pre-defined based on expert knowledge; future iterations could employ Causal Discovery algorithms to learn the causal structure from data dynamically, further reducing human bias.

Despite these limitations, this research establishes a robust proof-of-concept: that the future of actuarial modeling lies not in "bigger" data, but in "smarter" models that synthesize the deductive power of physics with the inductive power of machine learning.

5. Conclusion and Suggestions

This research marks a critical departure from the traditional actuarial paradigm, which has historically been constrained by the dual limitations of data scarcity and model opacity. By designing and validating the Physics-Informed and Causally-Aware (PICA) framework, this study demonstrates that the solution to pricing climate risk in developing markets lies not in waiting for the accumulation of "Big Data," which may never materialize, but in the strategic synthesis of "Big Physics" with Deep Learning. The empirical evidence presented herein confirms that embedding the fundamental laws of hydrodynamics into a Graph Neural Network architecture effectively serves as a universal prior, allowing the model to generalize with high accuracy even when observational data is reduced by 80%. Specifically, the demonstrated 65.8% reduction in predictive error under data-scarce conditions provides a definitive proof-of-concept that physical knowledge can act as a robust substitute for historical data. Furthermore, the integration of causal inference constraints successfully resolves the "black-box" dilemma, producing a model that satisfies the stringent transparency requirements of modern insurance regulation by prioritizing physical drivers over spurious spatial correlations.

Based on these findings, we propose the following actionable recommendations for industry stakeholders and policymakers. First, for the insurance industry, companies should prioritize the development of Parametric Insurance products for flood risks, particularly in high-hazard basins like Karun and Karkheh. By utilizing PICA's physics-verified water depth estimates as objective triggers, insurers can automate payouts without the need for lengthy loss adjustments and historical claims databases, significantly reducing administrative costs and settlement times. Second, for regulatory bodies such as the Central Insurance of Iran (Bimeh Markazi), we suggest establishing a "Model Governance Framework" that mandates causal explainability for any AI-based pricing model. Regulators should require insurers to provide SHAP-based feature attribution reports to prove that premiums are based on physical risk drivers rather than spurious correlations, ensuring fair treatment of policyholders. Finally, for researchers, future work should focus on bridging the "Sim-to-Real" gap by implementing Transfer Learning strategies. Specifically, fine-tuning the pre-trained PICA model on small, high-quality datasets from pilot regions (e.g., Golestan province) will be crucial to validate its operational readiness across Iran's diverse hydrological landscape.

Ultimately, this study advocates for an epistemological shift in actuarial data science: moving from purely inductive, correlation-based models to hybrid, deductive-inductive systems. We conclude that the future of climate risk modeling lies in Neuro-Symbolic AI, systems that do not merely learn from patterns in data, but understand the physical and causal laws that govern the natural world. The PICA framework represents a foundational step toward this intelligent, transparent, and resilient future.

Author Contributions
The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Acknowledgement and Funding
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest
The author declares that there is no conflict of interest regarding the publication of the registered research. Furthermore, ethical considerations including plagiarism, informed consent, misconduct, fabrication and/or falsification of data, duplicate publication, and also Journal Policy on Artificial Intelligence Use have been fully observed by the authors.

References

Bentivoglio, R., Taormina, R., & Cortes Arevalo, J. (2025). Multi-scale hydraulic graph neural networks for flood modelling. *Natural Hazards and Earth System Sciences*, *25*, 335–355. https://doi.org/10.5194/nhess-25-335-2025

Dazzi, S. (2024). Physics-informed neural networks for the augmented system of shallow water equations with topography. *Water Resources Research*, *60*(10), e2023WR036589. https://doi.org/10.1029/2023WR036589

Garzón, A., Kapelan, Z., Langeveld, J., & Taormina, R. (2024). Transferable and data efficient metamodeling of storm water system nodal depths using auto-regressive graph neural networks. *Water Research*, *266*, 122396. https://doi.org/10.1016/j.watres.2024.122396

Herath, H. M. V. V., Marshall, L., Saha, A., Rasnayaka, S., & Seneviratne, S. (2025). Subgrid informed neural networks for high-resolution flood mapping. *Journal of Hydrology*, *660*, 133329. https://doi.org/10.1016/j.jhydrol.2025.133329

Intergovernmental Panel on Climate Change. (2023). *Climate change 2023: Synthesis report*. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland.

Islam, T., Zeleke, E. B., & Afroz, M. (2025). A systematic review of urban flood susceptibility mapping: Remote sensing, machine learning, and other modeling approaches. *Remote Sensing*, *17*(3), 524. https://doi.org/10.3390/rs17030524

Joubier, V., Ebtehaj, I., & Bonakdari, H. (2024). Classifying river basin flood risk using AI: A case study of the Rideau River. *Proceedings of the 8th International Electronic Conference on Water Sciences*. https://doi.org/10.3390/ECWS-8-18875

Kazadi, A., Doss-Gollin, J., Sebastian, A., & Silva, A. (2024). FloodGNN-GRU: A spatio-temporal graph neural network for flood prediction. *Environmental Data Science*, *3*, e15. https://doi.org/10.1017/eds.2024.15

Kirschstein, N., & Sun, A. Y. (2024). Accelerating flood warnings by 10 hours: The power of river network topology in AI-enhanced flood forecasting. *arXiv preprint arXiv:2410.05536*. https://doi.org/10.48550/arXiv.2410.05536

Munich Re. (2024). *Natural catastrophe record 2023: The insurance perspective*. Munich Re Topics Geo.

Museru, M. L., Nazari, R., & Nikoo, M. R. (2025). Improving disaster resilience with causal machine learning for flood damage estimation. *Science of The Total Environment*, *954*, 176350. https://doi.org/10.1016/j.scitotenv.2025.176350

Nearing, G., Cohen, D., Drazner, V., Gauch, M., Kratzert, F., Klotz, D., Frame, J., ... & Nevo, S. (2024). Global prediction of extreme floods in ungauged watersheds. *Nature*, *627*(8005), 559–563. https://doi.org/10.1038/s41586-024-07145-1

Pearl, J. (2023). *Causality: Models, reasoning, and inference* (3rd ed.). Cambridge University Press.

Qi, X., Almeida, G., & Maldonado, S. (2024). Physics-informed neural networks for solving the shallow-water equations with terrain topography. *Journal of Fluid Mechanics*, *978*, A12. https://doi.org/10.1017/jfm.2023.1026

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Taormina, R., & Galelli, S. (2024). Deep learning for flood forecasting: A review of the state-of-the-art. *Journal of Hydrology*, *620*, 129021. https://doi.org/10.1016/j.jhydrol.2023.129021

Wang, Z., Li, Y., & Chen, X. (2024). Physics-informed machine learning in hydrology: A comprehensive review. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1-20. https://doi.org/10.1109/TGRS.2024.3351234

Zhang, T., Zhang, R., Li, J., & Feng, P. (2025). Deep learning of flood forecasting by considering interpretability and physical constraints. *Hydrology and Earth System Sciences*, *29*(21), 5955–5974. https://doi.org/10.5194/hess-29-5955-2025