Borrowing Information Across Space and Time: Pricing Flood Risk with Physics-based Hierarchical Deep Learning Models

Yanbin Xu Ken Seng Tan Wenjun Zhu

Nanyang Technological University
yanbin001@e.ntu.edu.sg

GAIP Technical Workshop
November 17, 2023

Presentation Overview

Introduction

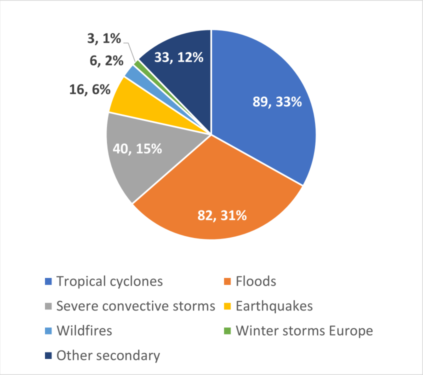


Figure: Global economic losses by peril in 2021, in USD billion and % share

Facts about flood risk:

1. Most frequent natural disasters.
2. Caused 1/3 of the natural disaster related fatalities since 2011.
3. Around 29% of the world population are exposed to flood risk.
4. Only 5% (34%) of the flood losses are insured in the emerging (developed) economies.

No-longer-“Secondary” Secondary Perils

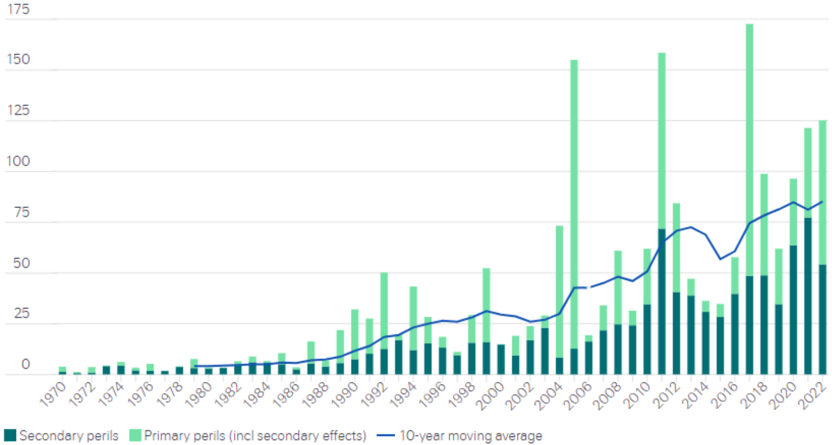


Figure: Insured natural catastrophe losses in USD billion

2022 Top Economic Loss Flood/Drought Events

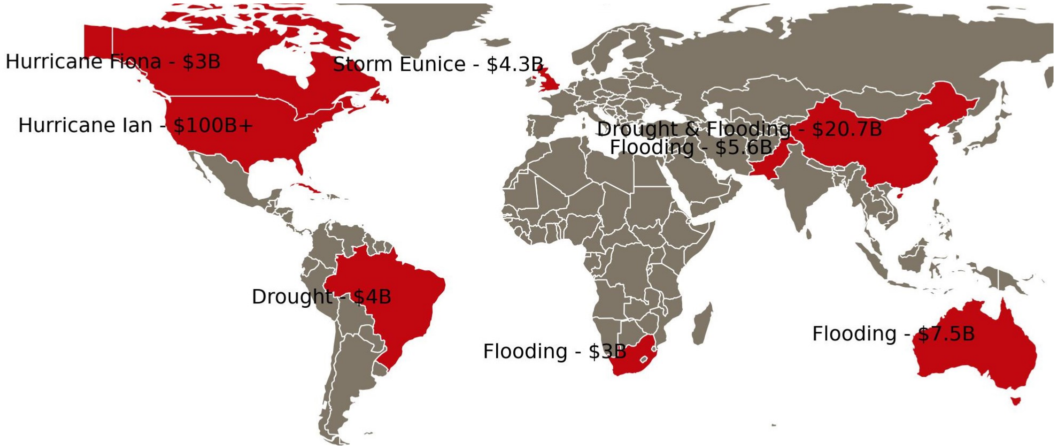


Figure: Economic Loss by Country.

2022 Top Fatality Flood/Drought Events

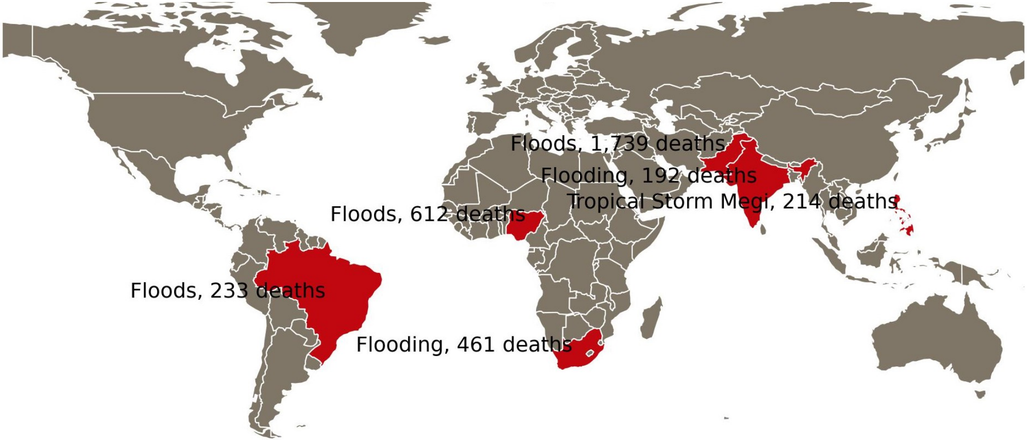


Figure: Fatality by Country.

Global Catastrophe Reinsurance Market

The Global property catastrophe reinsurance index has risen for six consecutive years, totaling an approximate 65% cumulative increase. This marked the largest positive shift in the Index since 2006, nearly returning it to levels seen in that year.

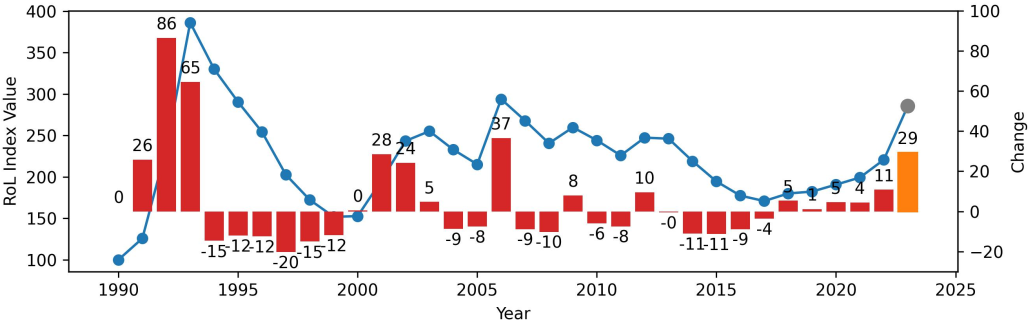


Figure: Global Property Catastrophe Rate-on-Line Index

US vs Rest APAC Flood Insurance Market

U.S. property catastrophe reinsurance rates-on-line increasing by 30% at January 1st, 2023, reaching an all-time high and marking a 97% cumulative increase since 2017, the last soft market's low point, including a 25% rise from January to July 2023.

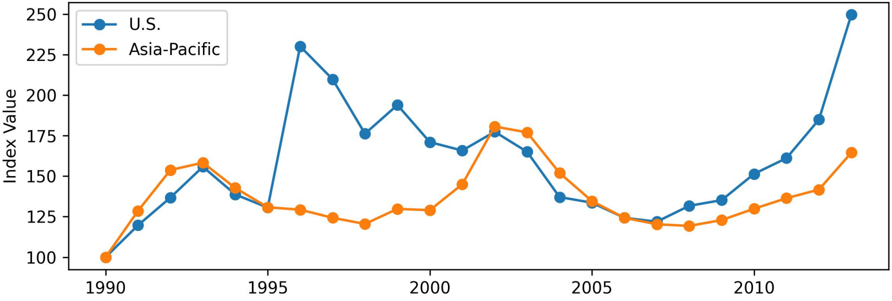
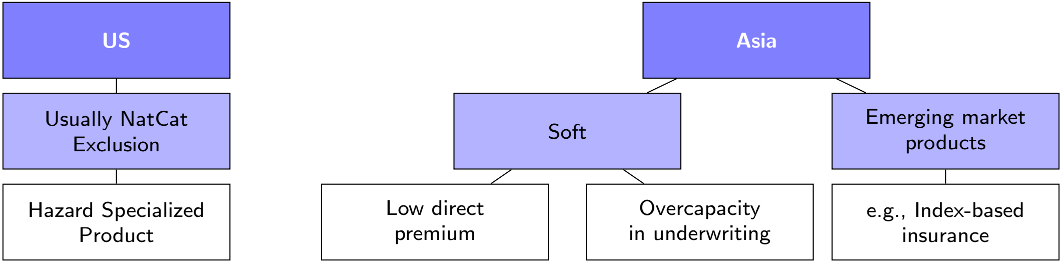


Figure: Regional Property Catastrophe Rate-on-Line Index

US vs. Asia: Property Insurance Market Differences

Aspect	US	Emerging markets
Market Maturity	Highly mature	Varies, emerging
Regulatory Environment	State-level regulation	Country-specific, varied
Product Offerings	Wide range, innovative	Traditional, diversifying
Risk Exposure	Hurricanes, wildfires	Typhoons, monsoons
Technology Adoption	Rapidly adopting	Mixed, rapid in some countries
Customer Base	High awareness	Growing awareness, price-sensitive
Distribution Channels	Mixed, incl. online	Agent/broker-dominated, digital growing

US vs. Asia: Property Insurance Market Differences



NFIP - US National Flood Insurance Program

The US approach to flood insurance has been the development of federally backed flood insurance through the NFIP.

- Managed by the Federal Emergency Management Agency (FEMA).
- Delivered through a network of over 50 insurance companies and NFIP Direct.
- Covering buildings and contents.

NATIONAL FLOOD
INSURANCE PROGRAM

NFIP - Coverage, Communities, and Requirements

- Offers flood insurance to property owners, renters, and businesses.
- Aids faster recovery post-flood events.
- Collaborates with communities for floodplain management and mitigation.
- Available in nearly 23,000 participating communities.
- Mandatory for high-risk areas with mortgages from government-backed lenders.

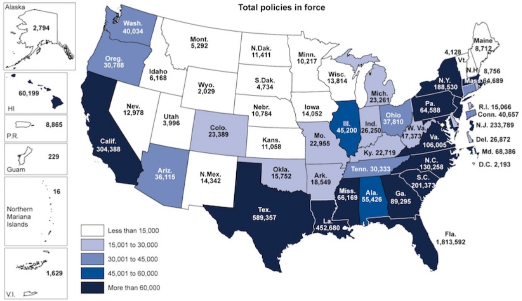


Figure: Total Policies in force 2023

NFIP Challenges and Global Flood Insurance Insights

Debt from Large-Scale Disasters: NFIP borrowed over $19 billion due to losses from 2005 and 2008 hurricanes and floods.

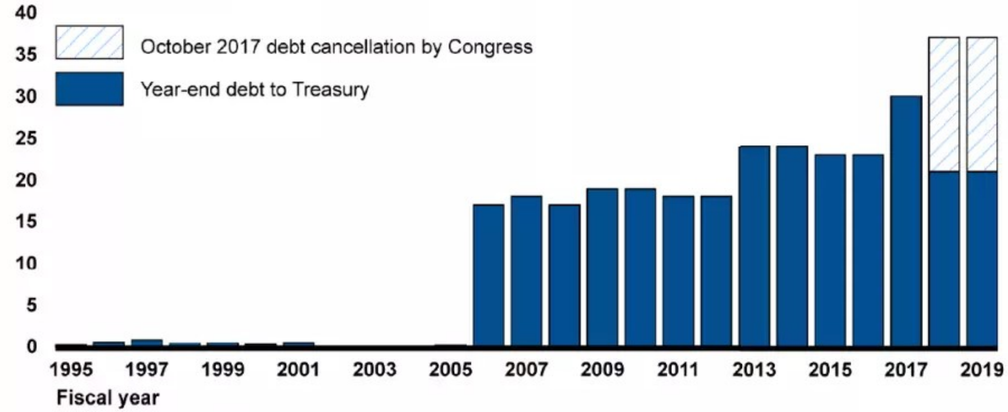


Figure: NFIP Annual Year-end Debt to Treasury in USD billion

NFIP Challenges and Global Flood Insurance Insights

- **Subsidized Insurance Issues:** Older buildings (1/4 of the policies) near or below base flood elevation before the completion of flood-risk map are undercharged to maintain the property value.
- **In-accurate risk capturing:** Many new constructions are charged premiums based only on average of historical loss, due to the low frequency of flood-risk map updating.
- **Low Insurance Penetration:** Only around 50% of homeowners in flood-prone areas have flood insurance. This lack of coverage increases the need for disaster relief.
- **Policy Lapse Rates:** On average only 74% are in force after 1 year, dropping to 36% after 5 years. The lapse rate does not vary much across flood zoon.

Literature Review and Research Gap

Difficulties in Flood Risk Modelling:

1. Empirical loss ratio models: low granularity
   - E.g., US National flood map is calculated at the county level.
2. Physical models (modelling flood mechanism): costly, hard to transfer to other market and region, hard to recalibrate.
3. Flood losses are heavily affected by anthropogenic effects.
   - In Thailand Flood 2011, water was directed to rural area by leaking several levees and dams to protect the high exposure Bangkok municipal region.

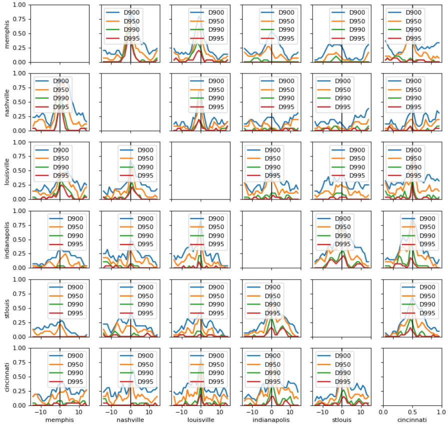
Literature Review and Research Gap

Additional difficulties in Flood Risk Modelling, for emerging markets:

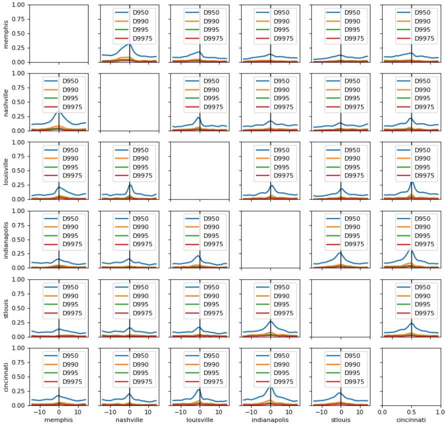
1. Data availability and resource constraints.

2. Change of land-use and urbanization.

Spatio-temporal flood risk correlation across cities



(a) Extreme flood risk scenarios



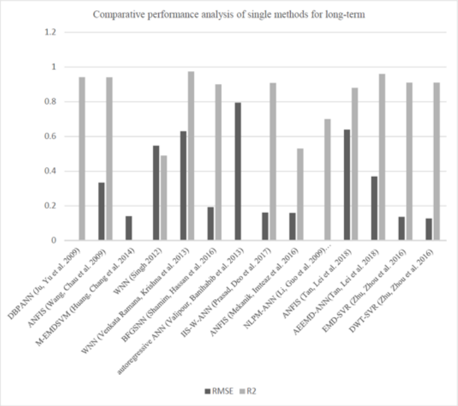
(b) Moderate flood risk scenarios

Literature Review and Research Gap

Flood insurance and its risk modellings are a rather niche topic in the actuarial science and insurance research.

- Boudreault et al. (2019) build a flood risk pricing method based on a chain of physics based modules.
- Browne et al. (2019) find that county-wide housing development in Florida is negatively associated with the mandatory National Flood Insurance Program.
- Kousky and Michel-Kerjan (2017) undertake a large-scale analysis of flood insurance claims in the United States.
- Hu (2022) finds that the peer effects affect flood insurance decisions. One's demand of flood risk insurance increase by 1-5% when her distant connection suffered a flood shock.

ML Application on Flood Modelling in Engineering Journals



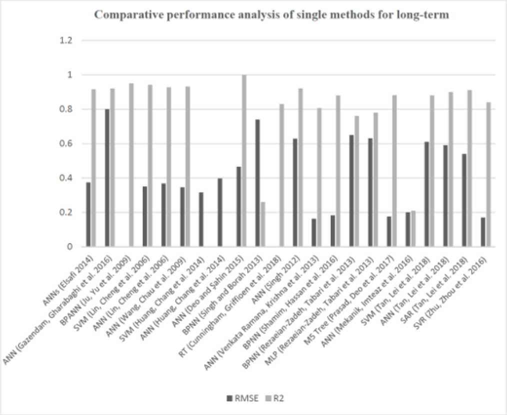


Figure: Prediction results summarized by Mosavi et al. (2018)

Research Objective

In this project, we target to provide a new deep learning structure that can produce a justifiable and transferable physics-based model for flood risk pricing.

Problem Set Up

A key to food insurance pricing is to forecast the flood risk measure.

We use water discharge as the flood risk measure. Let Q_{it} denotes the peak water discharge at gauging site $i = \{1, ..., I\}$. We seek to determine the best-estimated Q_{it} at time $t-1$, that is

$$\widehat{Q}_{it} = \mathbb{E}\left[Q_{it}|\mathbf{Q}_{t-1}, \mathbf{W}_{t-1}\right] = f\left(\mathbf{Q}_{t-1}, \mathbf{W}_{t-1}\right), \tag{1}$$

where $\mathbf{Q}_{t-1}$ and $\mathbf{W}_{t-1}$ represent the water discharge information set and the weather information set for all gauges prior to time $t-1$ until $t-l$. We use $f(\cdot)$ to denote the mean estimation of the learned deep learning model.

Model Architecture

For each city, the information integration process of the hierarchical deep learning structure is as below. The physical hierarchy of cities (or locations) is determined by their connectivity along the river channel.

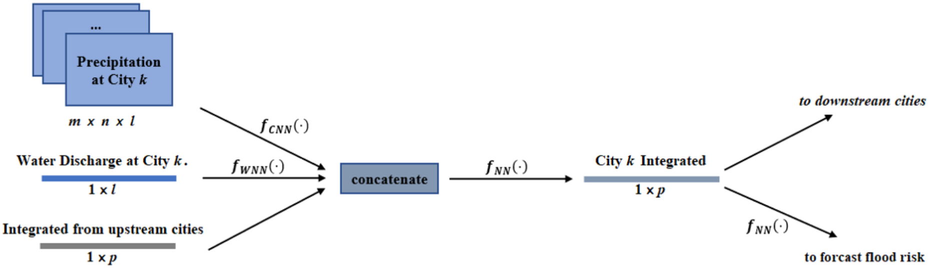


Figure: Information Integration Process of Hierarchical Deep Learning Structure.

Convolutional Neural Networks (CNNs)

- **Architecture:** CNNs are a class of deep neural networks, known for their excellence in processing data with a grid-like topology, such as images.
- **Key Components:** They utilize convolutional layers, pooling layers, and fully connected layers to capture spatial hierarchies and patterns in the data.
- **Applications:** Widely used in image and video recognition, image classification, medical image analysis, and other tasks involving visual data.

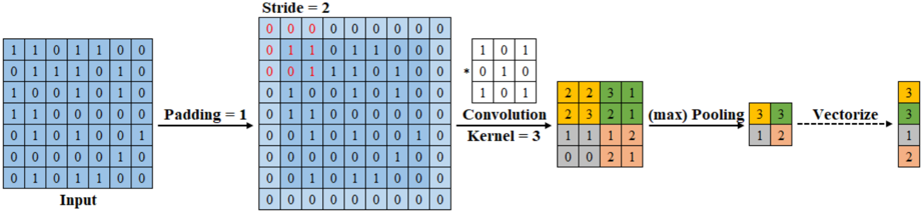


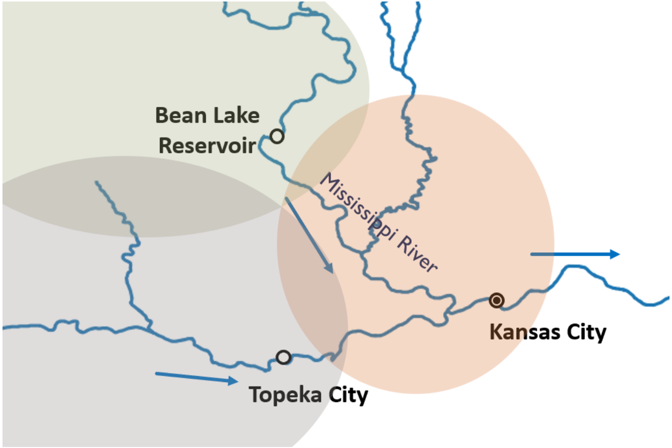
Figure: Procedure of a 2-dimensional CNN system.

A Hypothetical Example

The Proposed Structure

Climate data are split to train submodels.

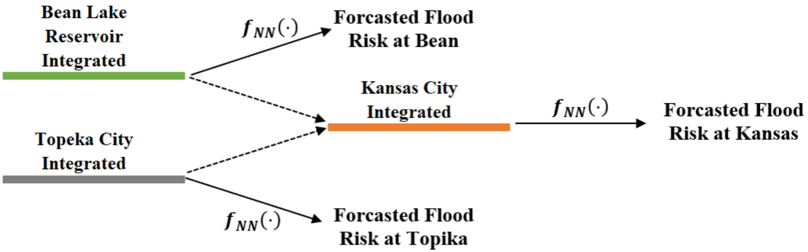
Flood risk measures at all locations are utilized to optimize the full model globally.



A Hypothetical Example

While preserves all the information, the other advantages include

- the connectivity improves the interpretability,
- the reduced searching space leads to faster conversion time,
- the physical constraints reduce the risk of overfitting due to complication.



Flood Insurance Pricing

Conceptually, for the kth insured building, there exists a damage function $h_k : Q_{it} \to U_{kt}$, which maps the water discharge observation $\{Q_{it} = q_{it}\}$ to the corresponding loss event $\{U_{kt} = u_{kt}\}$ in a monotonically increasing manner. The pure premium can be obtained:

$$\begin{aligned} p_{kt} = \mathbb{E}[Y^L_{kt}] = \int_{Q_i^-}^{Q_i^+} u_{kt}\theta_k \cdot f_Q(q_{it})\, dq_{it} = \int_{Q_i^-}^{Q_i^+} h_k(q_{it})\,\theta_k \cdot f_Q(q_{it})\, dq_{it} \\ = \theta_k\, \mathbb{E}[h_k(Q_{it})] = \theta_k h_k(\mathbb{E}[Q_{it}]) . \end{aligned} \tag{2}$$

Flood Insurance Pricing

In this paper, we utilize a set of pooling generalized linear models (GLM) to establish the relationship between water discharge, building characteristics, and the loss ratio:

$$\mathbb{E}[Y_{kt}^{L}] = \theta_k \, \mathbb{E}[N_{kt}] \, \mathbb{E}[\overline{U}_{kt}], \tag{3}$$

where N_{kt} is the random variable denotes the number of claims incurred or frequency, and $\overline{U}_{kt}$ is the loss ratio per claim or severity.

Flood Insurance Pricing

Let $X_k = (x_{k1}, ..., x_{kp})'$ denote a vector containing individual-level building characteristics. We use separate GLMs to combine X_k with $\widehat{Q}_{it}$:

$$\nu_{kt} = \mathbb{E}\left[N_{kt} | \widehat{Q}_{it}, X_k\right] = g_n^{-1}\left(\lambda_n \widehat{Q}_{it} + X_k' \beta_n\right), \tag{4}$$

$$\mu_{kt} = \mathbb{E}\left[\overline{U}_{kt} | \widehat{Q}_{it}, X_k\right] = g_u^{-1}\left(\lambda_u \widehat{Q}_{it} + X_k' \beta_u\right). \tag{5}$$

Here, (λ_n, β_n) and (λ_u, β_u) are estimated from historical observations. It is worth noting that ν_{kt} and μ_{kt} measure the relative risk of a building without being influenced by its exposure θ_k. Finally, the net premium can be calculated as follow:

$$p_{kt} = \theta_k \cdot \nu_{kt} \cdot \mu_{kt}. \tag{6}$$

Empirical Analysis

- **Laboratory:** Mississippi River Basin.
- **City clusters** Cincinnati, Nashville, Louisville, Indianapolis, St. Louis, and Memphis.
- **Hydraulic Data:** Streamflow (USGS).
- **Meteorological data:** Daily precipitation (PRISM) Raster data.
- **Policy and Claim Data:** The National Flood Insurance Program.

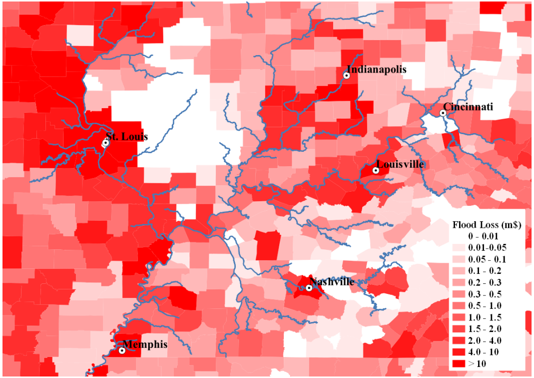
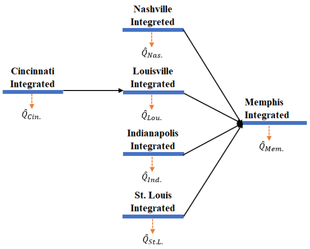


Figure: Area of Interest.

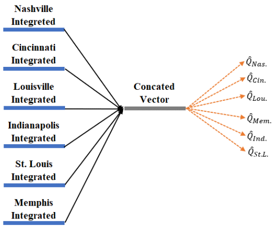
Deep Learning Structure for Area of Interest

We present two benchmark structures, as illustrated in panels (b) and (c). To maintain consistency across all DL architectures, we ensure that

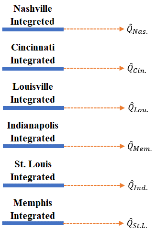
- each model has the same level of complexity in terms of parameters.
- all models use the identical pluvial and fluvial segments to incorporate both hydraulic and meteorological information.



(a) Hierarchical DL



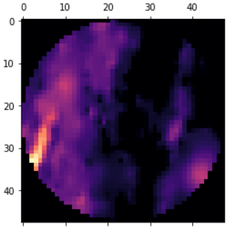
(b) No Connection DL



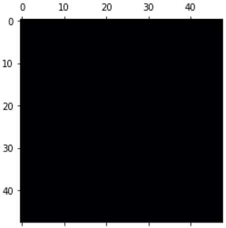
(c) No Spacial Information DL

Precipitation Data Example

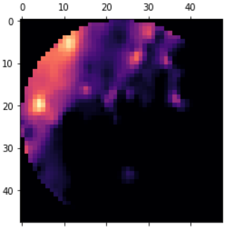
The precipitation raster of Indianapolis from December 28 to December 31 2020.



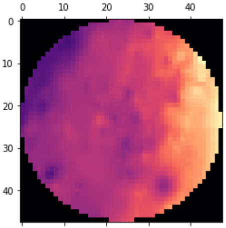
(a) Dec 28, 20



(b) Dec 29, 20



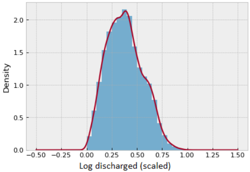
(c) Dec 30, 20



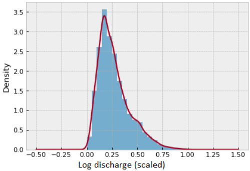
(d) Dec 31, 20

Hydraulic Data Example

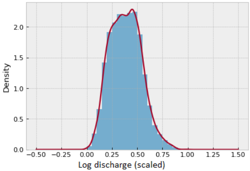
The logged water discharge distribution of Indianapolis from December 28 to December 31 2020.



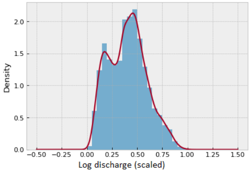
(a) Cincinnati



(b) Memphis



(c) Nashville



(d) St. Louis

Performance Measure

RMSE

To measure the severity prediction accuracy, we calculate training and testing RMSE for different exceedance probabilities of our models.

Denote $1_{Q_{it}} = 1\{Q_{it} \geq F_{Q_i}^{-1}(1-\alpha\%)\}$ as the indicator of a flood event at the most extreme $\alpha\%$ level, we have

$$RMSE(\alpha) = \sqrt{\frac{1}{\alpha \cdot IT} \sum_{i,t} \left(\frac{Q_{it} - \widehat{Q}_{it}}{Q_{it}} \right)^2 \cdot 1_{Q_{it}}}. \tag{7}$$

Performance Measure

Recall

Recall is a statistical measure commonly used in machine learning and information retrieval to evaluate the accuracy of a model. It represents the fraction of correctly predicted positive instances.

$$\begin{aligned} \text{Recall} &= \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}, \\ \text{Precision} &= \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}. \end{aligned} \tag{8}$$

Descendant and leaf cities

We define

- **Descendant cities:** cities with upper steam cities (able to use information from upper steam).
- **Leaf cities:** cities without upper steam cities.

Hypothesis

The improvement of hierarchical deep learning structure is more significant on descendant cities than leaf cities.

Training and testing sample

To facilitate the analysis, we temporally split the sample into training (70%) and testing sample (30%).

- **Training sample:** provide to the model for pattern detection.
- **Testing sample:** evaluate the performance of training sample.

Model selection

Select the model that performs the best in the testing sample.

Pluvial Segments	Exceedance Probability	Fluvial Segments: 1 hidden layer 64, Training RMSE	Training Recall	Testing RMSE	Testing Recall	Fluvial Segments: 2 hidden layer 128-64, Training RMSE	Training Recall	Testing RMSE	Testing Recall
2 convolutional hidden layers 16(3,1,1)-32(3,1,1)	10.00%	18.65%	77.87%	23.65%	75.18%	14.14%	81.08%	16.66%	78.55%
	5.00%	16.81%	72.89%	21.81%	69.78%	9.94%	86.72%	12.66%	83.26%
	1.00%	17.51%	70.02%	22.51%	66.23%	7.05%	90.45%	12.53%	82.76%
	0.50%	17.14%	68.35%	22.14%	64.99%	5.42%	92.28%	13.26%	79.54%
	0.25%	22.53%	67.27%	27.53%	60.84%	6.05%	94.07%	11.03%	86.94%
3 convolutional hidden layers 16(3,1,1)-32(3,1,1)-64(3,1,1)	10.00%	21.67%	78.47%	26.67%	75.54%	26.09%	85.27%	31.09%	82.03%
	5.00%	19.50%	72.62%	24.50%	68.27%	24.43%	82.33%	29.43%	78.50%
	1.00%	18.46%	65.99%	23.46%	61.22%	21.15%	84.51%	26.15%	78.84%
	0.50%	18.89%	67.28%	23.89%	62.13%	19.97%	83.72%	24.97%	79.22%
	0.25%	23.75%	66.96%	28.75%	59.09%	24.16%	86.25%	29.16%	79.17%
4 convolutional hidden layers 16(3,1,1)-32(3,1,1)-64(3,1,1)-128(3,1,1)	10.00%	27.97%	76.30%	32.97%	72.60%	26.47%	85.37%	31.47%	81.63%
	5.00%	23.70%	70.21%	28.70%	65.46%	21.51%	81.88%	26.51%	77.41%
	1.00%	19.10%	66.21%	24.10%	61.20%	16.81%	80.32%	21.81%	74.33%
	0.50%	16.19%	64.58%	21.19%	60.07%	15.34%	78.55%	20.34%	72.51%
	0.25%	21.35%	65.46%	26.35%	58.11%	21.26%	80.64%	26.26%	73.32%

Table: Model selection.

Baseline Performance in Flood Risk Forecasting

Panel A: Cities with Upstream Nodes (Descendant Cities)									
Measure	Exceedance	WNN Water Discharge ML Model		No Spacial Information DL Model		No Connection DL Model		Hierarchical DL Model	
	Probability	Training	Testing	Training	Testing	Training	Testing	Training	Testing
RMSE	10.00%	37.26%	38.12%	22.99%	25.50%	6.57%	17.37%	10.04%	14.17%
	5.00%	31.51%	30.46%	19.50%	21.45%	7.41%	19.07%	8.42%	10.52%
	1.00%	22.93%	28.23%	16.45%	22.10%	11.96%	22.40%	7.03%	11.66%
	0.50%	13.69%	29.52%	9.69%	25.39%	10.09%	29.91%	1.46%	12.27%
	0.25%	19.41%	25.17%	18.51%	21.25%	14.17%	25.22%	4.83%	8.74%
Recall	10.00%	62.74%	61.88%	77.01%	74.50%	93.43%	82.63%	89.96%	85.83%
	5.00%	68.49%	69.54%	80.50%	78.55%	92.59%	80.93%	91.58%	89.48%
	1.00%	77.07%	71.77%	83.55%	77.90%	88.04%	77.60%	92.97%	88.34%
	0.50%	86.31%	70.48%	90.31%	74.61%	89.91%	70.09%	98.54%	87.73%
	0.25%	80.59%	74.83%	81.49%	78.75%	85.83%	74.78%	95.17%	91.26%
Panel B: Cities without Upstream Nodes (Leaf Cities)									
Measure	Exceedance	WNN Water Discharge ML Model		No Spacial Information DL Model		No Connection DL Model		Hierarchical DL Model	
	Probability	Training	Testing	Training	Testing	Training	Testing	Training	Testing
RMSE	10.00%	20.92%	22.52%	15.85%	16.57%	12.69%	27.89%	16.77%	18.26%
	5.00%	14.85%	14.92%	9.90%	12.30%	13.12%	27.70%	10.91%	14.04%
	1.00%	8.96%	12.65%	7.38%	10.30%	11.98%	34.95%	7.07%	13.09%
	0.50%	10.83%	13.33%	6.89%	12.03%	15.96%	35.01%	7.96%	13.89%
	0.25%	9.97%	18.95%	6.71%	8.12%	14.78%	36.69%	6.83%	12.50%
Recall	10.00%	79.08%	77.48%	84.15%	83.43%	87.31%	72.11%	83.23%	81.74%
	5.00%	85.15%	85.08%	90.10%	87.70%	86.88%	72.30%	89.09%	85.96%
	1.00%	91.04%	87.35%	92.62%	89.70%	88.02%	65.05%	92.93%	86.91%
	0.50%	89.17%	86.67%	93.11%	87.97%	84.04%	64.99%	92.04%	86.11%
	0.25%	90.03%	81.05%	93.29%	91.88%	85.22%	63.31%	93.17%	87.50%

Table: Risk factor accuracy measured by RMSE and recall.

Baseline Performance in Pricing

The hierarchical deep learning structure better forecasts the loss than the benchmarks.

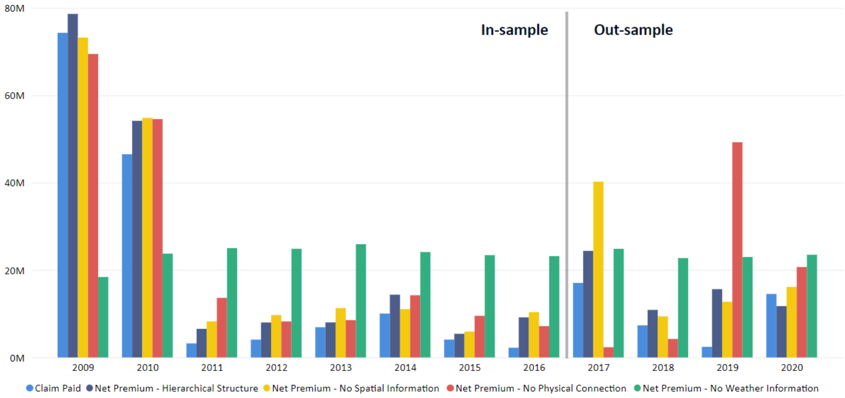


Table: Pricing performance measured by net premium in training and testing sample.

Baseline Performance in Pricing

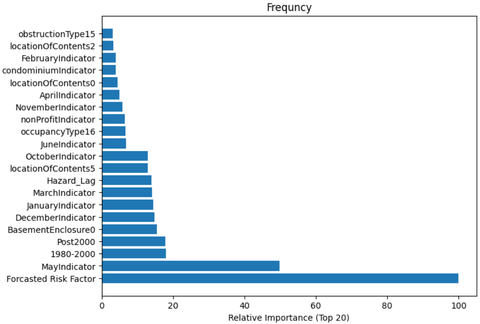
To mitigate the influence of occasional outliers, we employed bootstrap resampling on testing dataset to evaluate the pricing performance.

	No Flood Risk Factor	No Spatial Information	No Connection	Hierarchical
Panel A: Net Premium Results				
Net Premium (m\$, an.)	23.52	19.61	19.12	15.65
Net Premium Reduced (m\$, an.)	-	3.91	4.40	7.87
Net Premium Improvement (%)	-	16.62%	18.71%	33.45%
Panel B: Solvency Capital Requirement Results				
Value-at-Risk 0.5% (m\$, an.)	0.05	-0.26	-0.04	-0.23
Expected Shortfall 0.5% (m\$, an.)	-0.41	-0.58	-0.42	-0.63
BSCR (m\$, an.)	23.47	19.87	19.17	15.88
BSCR Reduced (m\$, an.)	-	3.60	4.30	7.59
BSCR Improvement (m\$, an.)	-	15.34%	18.34%	32.33%

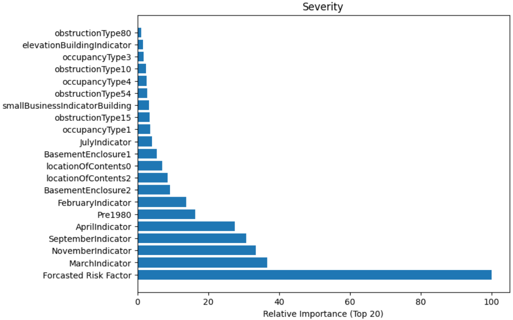
Table: Pricing performance measured by net premium and solvency capital.

Robustness Test: Importance of Generated Risk Factor

The relative importance test shows that the hierarchical DL generate risk factor is the most important factor in both the frequency and severity GLM.



(a) Frequency

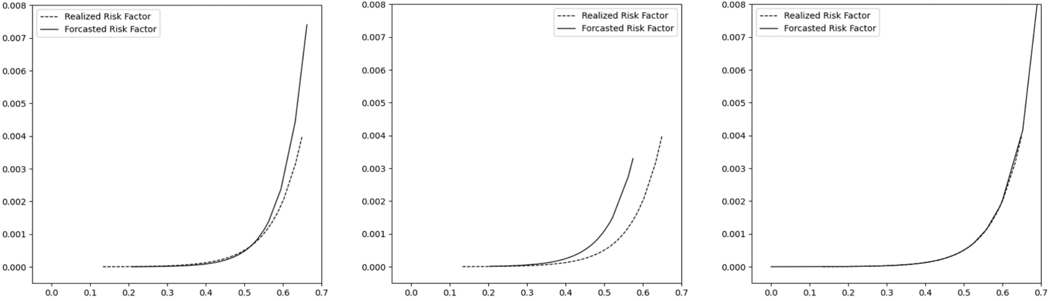


(b) Severity

Figure: Factor Relative Importance

Robustness Test: Partial Dependence

We fited a GLM model that takes the realized water discharge and plotted its partial dependence in the three Panels as dashed lines. Notably, the proposed hierarchical structure's forecasted risk factors exhibit the best approximation to the anchor.



(a) No Spatial DL **(b)** No Connection DL **(c)** Hierarchical DL

Figure: Forecasted Risk Factor Partial Dependence

Conclusion

1. We proposed a physics-based hierarchical structure that ensembles flood mechanisms into ML models.
2. The proposed structure utilize flood risk observed at secondary exposure regions to conduct global optimization.
3. In addition, the physics-based hierarchical structure improves the interpretability of the ML models and partially resolves the difficulty of flood risk prediction due to anthropogenic effects.
4. Using the Mississippi river as an example, we demonstrate that the proposed physics-based hierarchical structure has superior performance compared to conventional structure.

Next Step: Emerging Market

1. The proposed structure is intentionally tailored to fit the data and resource constraints of developing countries.

2. Our research team is ought to test and make further custermizations to the proposed framework.

Thank you!

References

Mathieu Boudreault, Patrick Grenier, Mathieu Pigeon, Jean-Mathieu Potvin, and Richard Turcotte (2022)
Pricing Flood Insurance with a Hierarchical Physics-Based Model.
North American Actuarial Journal 24(2), 251–274.

Mark J. Browne, Carolyn A. Dehring, David L. Eckles, William D. Lastrapes (2019)
Does National Flood Insurance Program Participation Induce Housing Development?
Journal of Risk and Insurance Vol.86, No.4, 835 – 859.

Carolyn Kousky, Erwann Michel-Kerjan (2017)
Examining Flood Insurance Claims in The United States: Six Key Findings.
Journal of Risk and Insurance Vol.84, No.3, 819 – 850.

Zhongchen Hu (2017)
Social Interactions and Households' Flood Insurance Decisions.
Journal of Risk and Insurance Vol.144, 414 – 432.

Amir Mosavi, , Pinar Ozturk, and Kwok-wing Chau (2018)
Flood Prediction Using Machine Learning Models: Literature Review.
Journal of Risk and Insurance Vol.86, No.4, 835 – 859.

E. Danso-Amoako.; M. Scholz; N. Kalimeris; Q. Yang; J. Shao (2012)
Predicting dam failure risk for sustainable flood retention basins: A generic case study for the wider greater Manchester area.
Computers, Environment and Urban Systems 36, 423 – 433.

Wen-Chuan Wang, Kwok-Wing Chau, Chun-Tian Cheng, Lin Qiu (2014)
A comparison of performance of several artificial intelligence. methods for forecasting monthly discharge time series.
Journal of Hydrology Vol 374, 294–306.

Shengzhi Huang, Jianxia Chang, Qiang Huang, Yutong Chen (2014)
Monthly streamflow prediction using modified EMD-based support vector machine.
Journal of Hydrology Vol 511, 764–775.

E Lechowska (2018)
What determines flood risk perception? A review of factors of flood risk perception and relations between its basic elements.
Natural Hazards Vol 94, 1341–1366.

ZW Kundzewicz, Su Budaad, Wang Yanjuna, Xia June, Huang Jinlonga, Jiang Tong (2019)
Flood risk and its reduction in China.
Advances in Water Resources Vol 130, 37–45.

Hai-Min Lyu, Shui-Long Shen, Annan Zhou, Jun Yang (2019)
Perspectives for flood risk assessment and management for mega-city metro system.
Advances in Water Resources

Brenden Jongman (2019)
Effective adaptation to rising flood risk.
Nature Communications Vol 9.

- Binh Thai Pham, Chinh Luu, Tran Van Phong, Huu Duy Nguyen, Hiep Van Le, Thai Quoc Tran, Huong ThuTa, Indra Prakash (2021)
  Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam.
  Journal of Hydrology Vol 592. 125815

- Irfan Ahmad Rana, Ali Jamsh, Zahid Irsha, Younasc Saad Saleem Bhatti(2021)
  Characterizing flood risk perception in urban communities of Pakistan.
  International Journal of Disaster Risk Reduction Vol 46.101624